



**HAL**  
open science

# Modélisation mathématique des propriétés de mélanges : B-splines et optimisation avec conditions de forme

Nabih Odeh

► **To cite this version:**

Nabih Odeh. Modélisation mathématique des propriétés de mélanges : B-splines et optimisation avec conditions de forme. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1990. Français. NNT: . tel-00337111

**HAL Id: tel-00337111**

**<https://theses.hal.science/tel-00337111>**

Submitted on 6 Nov 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.

70 280

# THESE

*présentée par*

**Nabih ODEH**

pour obtenir le titre de

**DOCTEUR**

**DE L'UNIVERSITE JOSEPH FOURIER - GRENOBLE I**

(arrêté ministériel du 5 juillet 1984)

**Spécialité : Mathématiques Appliquées**

**Modélisation mathématique des propriétés de  
mélanges : B-splines et optimisation avec  
conditions de forme**

Date de soutenance : 19 Mars 1990

Composition du jury :

F.	ROBERT	Président
A.	LE MEHAUTE	
P.	VALENTIN	Examineurs
B.	LACOLLE	
P.J.	LAURENT	

**Thèse préparée au sein du laboratoire TIM3**



Ce travail a été réalisé au sein du laboratoire TIM3 (Technique de l'Informatique, de la Microélectronique et de la Microscopie quantitative) en collaboration avec le C.R.E.S (Centre de Recherche de Elf à Solaize).

J'exprime mes remerciements les plus sincères à Monsieur P.J. LAURENT, qui a dirigé mes travaux, pour le soutien et la confiance ; ses conseils et ses encouragements constants ont contribué énormément à l'accomplissement de ce travail.

Monsieur A. LE MEHAUTE, professeur à l'université de Lille, a eu l'amabilité de se déplacer pour juger cette thèse, qu'il en soit vivement remercié.

Je tiens à remercier aussi, Monsieur P. VALENTIN et Madame M.T. FEILLENS du C.R.E.S de m'avoir accueilli dans ce centre. Leur collaboration a été fructueuse.

Je remercie Monsieur F. ROBERT, professeur à l'INPG, d'avoir accepté de présider le jury, et Monsieur B. LACOLLE pour bien voulu en faire partie.

Je remercie également tous les membres de l'équipe de Géométrie Algorithmique pour leur aide et leur sympathie, et en particulier mes amis et collègues du bureau 56 de la tour IRMA.

Enfin, mes remerciements vont à tous les membres du service de Reprographie de l'IMAG pour le tirage de cette thèse.



# **Table des Matières**



## Introduction Générale

### Première Partie : B-splines et analyse de mélanges additifs

#### Chapitre I : B-splines

1.0 Introduction .....	3
1.1 Définitions et notations .....	3
1.2 Bases de fonctions puissances tronquées .....	4
1.3 B-splines .....	5
• Récurrence .....	7
• Support compact et positivité .....	7
• Propriété de l'enveloppe convexe .....	8
• Dérivées des B-splines .....	8
1.4 Splines avec conditions de forme .....	9
• Positivité .....	10
• Monotonie .....	10
• Convexité .....	12
1.5 B-splines cubiques .....	14
1.6 B-splines cubiques uniformes .....	17
1.7 Calcul des valeurs des splines cubiques uniformes .....	18
1.8 Quelques opérateurs sur l'espace $S(k,m)$ .....	21
• Opérateur du lissage .....	21
• Opérateur des moindres carrés pondérés .....	24

#### Chapitre II : Modélisation de coupes pétrolières

2.0 Introduction .....	31
2.1 Définitions et notations .....	31
2.2 Modélisation .....	33



2.3	Modèle I .....	35
2.4	Modèle II .....	37
2.5	Modèle III .....	43
2.6	Choix des poids $\{\omega_k\}_{k=1}^n$ .....	47
2.7	Modélisation du rapport de deux propriétés conservatives	48
2.8	Conclusion .....	53

Chapitre III : Algorithmes de résolution : optimisation  
convexe avec contraintes linéaires

3.0	Introduction .....	57
3.1	Formulation du problème : définitions et notations .....	58
3.2	Description de l'algorithme .....	59
3.3	Exemple simple (n=1) .....	62

Chapitre IV : Application : calcul des propriétés  
conservatives

4.0	Introduction .....	67
4.1	Analyse du pétrole brut .....	67
4.2	Logiciel C-ELF .....	69
	• Environnement .....	69
	• Modélisation et calcul numérique .....	70
	• Visualisation et représentation graphique .....	70
4.3	Analyse du brut A334 .....	72
	• Distillat .....	72
	• Densité .....	77
	• Soufre .....	77
4.4	Conclusion .....	85

## Deuxième Partie : Analyse de mélanges non additifs

### Chapitre V : Propriétés non linéaires de coupes complexes

5.0	Introduction .....	89
5.1	Première approche : équations d'indices.....	90
	• Modèle d'approximation de la fonction d'indice ..	92
	• Algorithme de résolution du problème (P'.5.1) ..	98
	• Problème additif associé .....	103
	• Evaluation de la propriété non additive .....	105
5.2	Deuxième approche : reconstitution des quantités d'indice	107
5.3	Troisième approche : reconstitution des mesures .....	110
5.4	Application au calcul de point de trouble des gazoles .	112
5.5	Conclusion .....	122

## Troisième Partie : Géométrie des ensembles de mélange

### Chapitre VI : Convexe de mélanges et opérations géométriques

6.0	Introduction .....	125
6.1	Notion de convexe de mélanges .....	126
6.2	Propriétés relatives au convexe de mélanges .....	128
6.3	Opérations géométriques sur le convexe de mélanges ..	131
6.4	Intersection d'un zonotope avec un hyperplan dans $\mathbb{R}^3$ .	132
6.5	Quelques exemples d'intersections dans $\mathbb{R}^3$ .....	149
	• Premier exemple .....	149
	• Deuxième exemple .....	156
6.6	Conclusion .....	159

<b>Conclusion</b> .....	161
<b>Bibliographie</b> .....	165

# **Introduction Générale**



Beaucoup de problèmes physiques se ramènent à des problèmes de séparation. Or, la séparation est un concept très large et peut se produire à des échelles différentes, allant de la cellule (micron) au réservoir du pétrole brut (kilomètre), avec des applications très variées dans plusieurs domaines notamment la chimie et la pétrochimie, ce qui explique l'importance industrielle du problème. Cependant, beaucoup de problèmes numériques, et en particulier dans le domaine pétrolier, demeurent non résolus jusqu'à présent. Le travail, que nous présentons ici, est une contribution à la résolution numérique de certains problèmes rencontrés dans l'étude de la séparation du pétrole.

Pour un système physico-chimique (contenu matériel et énergétique d'un volume borné dans notre espace euclidien) la séparation consiste à arranger ou déplacer des corps ou des classes d'espèces dans des régions différentes, de façon à avoir plus de facilités à évaluer leur propriétés ou bien à produire d'autres mélanges.

La première opération, dans la transformation du pétrole brut, est celle du raffinage fondé sur la distillation fractionnée qui permet de séparer un pétrole brut ou une partie de celui-ci en «coupes» d'hydrocarbures de propriétés suffisamment voisines pour qu'on puisse leur donner en bloc la même utilisation. La distillation fractionnée est un procédé de séparation qui utilise la différence des points d'ébullition entre des produits qui se trouvent mélangés et même dissous les uns dans les autres, dans la matière première.

Les problèmes étudiés, dans ce travail, concernent deux types de mélanges différents :

- **Mélanges complexes :**

Il s'agit d'une analyse d'un mélange de très grand nombre d'espèces et d'une étude de ses propriétés physiques à partir d'un nombre fini de mesures. Cependant, nous pouvons classer les propriétés physiques, concernant ce type de mélanges, en deux catégories :

- (a) propriétés additives (ou conservatives).
- (b) propriétés non additives.

L'analyse des mélanges additifs consiste à représenter un ensemble fini de mesures expérimentales, d'une propriété physique additive, par une fonction lisse continûment dérivable et permettant ainsi l'évaluation de cette propriété sur n'importe quelle coupe désirée. Ce type d'analyse a été largement abordé dans la première partie de ce travail.

Par contre, l'analyse des mélanges non additifs est beaucoup plus compliquée au niveau du traitement. En général, on admet qu'une propriété non additive  $F$  dépend d'un certain nombre de grandeurs additives et dans le cas où la propriété  $F$  ne dépend que d'une seule grandeur additive (c'est une hypothèse) ceci revient à supposer l'existence d'une fonction inversible  $I$  telle que  $I(F)$  soit additive ; c'est le cas de la viscosité, du point de trouble au refroidissement et de la propriété antidétonante (mesure d'octane). La détermination d'une telle fonction, pour une propriété non additive, permet l'application des procédés de gestion des mélanges additifs et en conséquence l'évaluation de cette propriété pour des coupes complexes. L'ensemble des procédés permettant d'obtenir une telle fonction est largement étudiés au cours de la deuxième partie.

Les outils mathématiques fondamentaux dans le traitement des mélanges complexes sont les B-splines, dont les propriétés utiles sont rappelées au premier chapitre.

L'étude effectuée, sur les différents problèmes de la gestion des mélanges complexes, a conduit à mettre au point le logiciel C-ELF ainsi que plusieurs programmes d'études et de recherches concernant l'analyse des mélanges non additifs. Ceci va permettre une meilleure exploitation numérique et graphique.

- **Mélanges simples :**

En supposant que le contenu d'un système physico-chimique est constitué d'une famille finie de produits ou d'espèces, soit  $n$  leur nombre. Pour une partition du système en plusieurs régions l'analyse de chaque espèce permet d'évaluer sa quantité dans chaque région, et donc d'associer à chacune de ces régions un scalaire non négatif. L'ensemble de ces analyses permet d'attribuer à chaque région  $n$  scalaires non négatifs, ce qui constitue un vecteur à  $n$  composantes que l'on appelle vecteur quantité. L'ensemble des mélanges faisables à partir de cette partition est appelé convexe de mélanges (ou zonotope). Cet ensemble traduit la capacité de fabrication des mélanges à partir d'une partition du système et fournit une bonne caractérisation de la répartition (séparation) des espèces dans ce système. P.VALENTIN a mis au point, ces dernières années, une représentation géométrique de la séparation par la classe des zonotopes. De cette représentation originale, il a dégagé plusieurs problèmes mathématiques concernant des opérations géométriques sur le convexe de mélanges. Ces opérations sont, essentiellement, des opérations d'inclusions, d'intersections et de projections. Nous nous sommes intéressés, dans la troisième partie de ce travail, à la géométrie des ensembles de mélanges et en particulier au problème d'intersection d'un zonotope avec un hyperplan dans  $\mathbb{R}^3$ . Ceci nous a permis de mettre au point un algorithme rapide et efficace pour calculer l'intersection.





*Première Partie : B-Splines et  
Analyse de Mélanges Additifs*

CHAPITRE I

**B-Splines**



## 1.0 Introduction

Bien que les polynômes se prêtent aisément au calcul, il n'est généralement pas possible de définir une courbe satisfaisante simplement à l'aide de polynômes. Au lieu de cela, il est usuel de diviser la courbe en un certain nombre de segments, dont chacun est défini par des polynômes particuliers, et de connecter de tels segments pour former une courbe composée. Ainsi, afin de définir une courbe, un paramètre  $u$  varie entre une valeur initiale et une valeur finale, et certaines valeurs particulières de  $u$  dites "valeurs nodales", correspondent aux points de jonction entre segments polynômiaux. La suite des valeurs nodales doit être croissante.

### 1.1 Définitions et notations

Etant donné  $\{u_i\}_0^{m+k}$  une suite de valeurs nodales vérifiant

$$u_0 < u_1 < \dots < u_{m+k}$$

et  $k \geq 1$ ,  $k-1 < m+1$ , on note  $S(k,m)$  l'espace des fonctions (splines) définies sur  $[u_{k-1}, u_{m+1}[$  de classe  $C^{k-2}$  et qui sont formées de polynômes de degré  $k-1$  sur chaque intervalle  $[u_i, u_{i+1}[$ ,  $i=k-1, \dots, m$ . On voit facilement que  $S(k,m)$  est de dimension  $m+1$ . Il est clair que toute fonction polynômiale de degré  $k-1$  appartient à  $S(k,m)$ .

## 1.2 Bases de fonctions puissances tronquées

### (1.1) Définition

On appelle fonction puissance tronquée d'ordre  $r$  une fonction de la forme:

$$(u - t)_+^r = [\max(u - t, 0)]^r = \begin{cases} 0 & u < t \\ (u - t)^r & u \geq t \end{cases}$$

dans laquelle  $r=0, 1, 2, \dots$  et  $(u - t)_+^0$  est la fonction échelon (Figure 1.1).

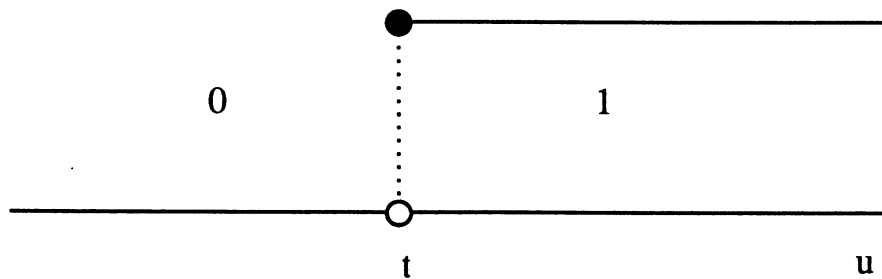


Figure 1.1 : fonction échelon  $(u - t)_+^0$

Il est facile à voir que les fonctions puissances tronquées  $(u - u_i)_+^{k-1}$   $i=0, \dots, m$  sont linéairement indépendantes et que chacune d'elles est un élément de  $S(k, m)$ . Donc, les fonctions  $\left\{ (u - u_i)_+^{k-1} \right\}_{i=0}^m$  forment une base pour  $S(k, m)$ .

Les fonctions puissances tronquées se prêtent mal au calcul. Leur utilité dans notre cas n'est pas dans la construction de fonctions splines ; elles présentent deux inconvénients pour ce genre d'opération : d'abord, une instabilité numérique et ensuite une difficulté d'effectuer des corrections locales. L'utilité de ces fonctions est qu'elles sont faciles à comprendre et à utiliser. Les fonctions puissances tronquées, d'autre part, peuvent prendre des valeurs importantes lorsque  $u$  augmente. Par conséquent, si une base de ces fonctions est utilisée pour représenter des courbes qui ne varient pas d'une façon brutale, les coefficients nécessaires peuvent varier

alternativement entre nombres positifs et négatifs de grande valeur absolue afin d'annuler numériquement certaines valeurs des fonctions de bases lorsque  $u$  augmente. Or, c'est le type d'opérations de calcul qu'il vaut mieux éviter. Nous allons, maintenant, définir les bases que nous voulons utiliser, c'est-à-dire les B-splines.

### 1.3 B-splines

Rappelons la définition des différences divisées d'une fonction  $f$  d'une variable réelle  $t$  par rapport à un ensemble d'abscisses  $u_0, u_1, u_2, \dots$

ordre 0 :

$$\delta^0 [u_i : t] f(t) = f(u_i)$$

ordre k :

$$\delta^k [u_i, u_{i+1}, \dots, u_{i+k} : t] f(t) = \frac{\delta^{k-1} [u_{i+1}, \dots, u_{i+k} : t] f(t) - \delta^{k-1} [u_i, \dots, u_{i+k-1} : t] f(t)}{u_{i+k} - u_i}$$

On montre que l'on a :

$$\delta^k [u_i, u_{i+1}, \dots, u_{i+k} : t] f(t) = \sum_{l=i}^{i+k} \frac{f(u_l)}{\prod_{\substack{j=i \\ j \neq l}}^{i+k} (u_l - u_j)}$$

(1.2) Définition

Etant donnée des valeurs nodales  $u_0, u_1, \dots, u_m, \dots, u_{m+k}$ , et en supposant que  $i \leq m$ , la B-spline d'ordre  $k$  associée aux valeurs nodales  $u_i, \dots, u_{i+k}$  est définie par :

$$B_{i,k}(u) = (-1)^k (u_{i+k} - u_i) \delta^k [u_i, u_{i+1}, \dots, u_{i+k} : t] (u - t)_+^{k-1}$$

Dans la figure 1.2 nous avons tracé les B-splines d'ordre 1, 2, 3 et 4.

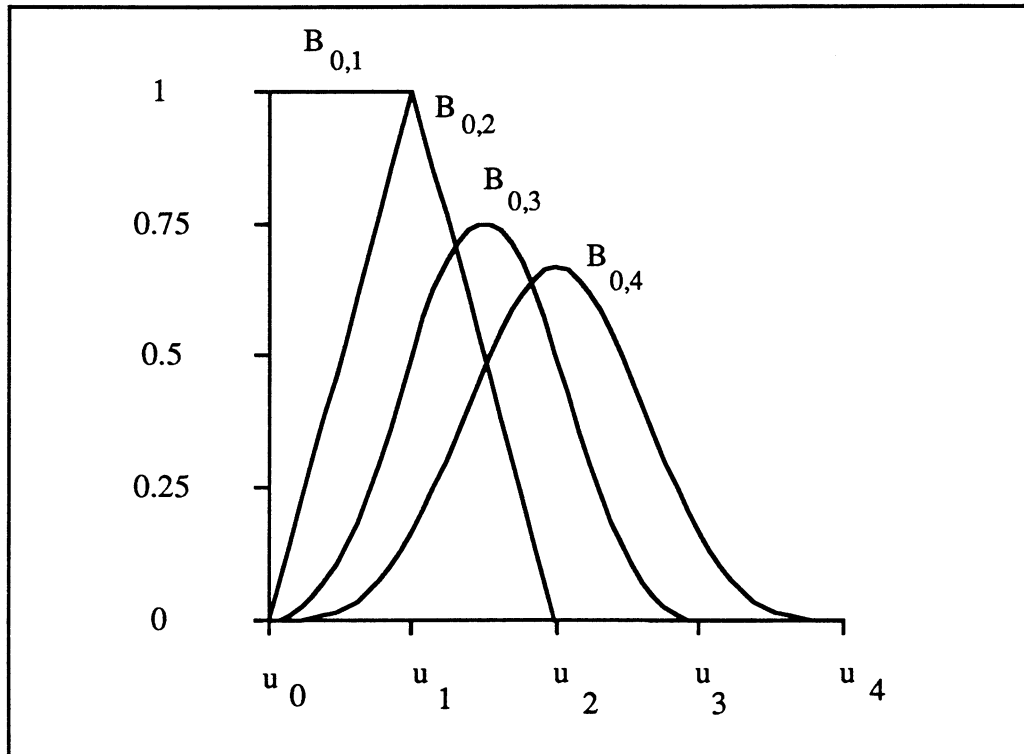


Figure 1.2 : B-splines d'ordre 1, 2, 3 et 4.

(1.3) Théorème :

Les fonctions  $B_{0,k}, \dots, B_{m,k}$  forment une base pour  $S(k,m)$ .

Cette propriété a été établie par Curry et Schoenberg, et une démonstration se trouve dans [de Boor 78].

Les B-splines d'ordre  $k$  possèdent les propriétés suivantes :

- **Récurrance :**

Les B-splines satisfont à une relation de récurrence. Ceci a été démontré indépendamment par Carl de BOOR [de Boor 72] et Maurice COX [Cox 72].

(1.4) Théorème (récurrence)

Pour tout  $i \in \{0, 1, \dots, m\}$  on a :

$$B_{i,1}(u) = \begin{cases} 1 & \text{si } u_i \leq u < u_{i+1} \\ 0 & \text{sinon} \end{cases}$$

et

$$B_{i,r}(u) = \frac{u - u_i}{u_{i+r-1} - u_i} B_{i,r-1}(u) + \frac{u_{i+r} - u}{u_{i+r} - u_{i+1}} B_{i+1,r-1}(u)$$

pour  $r=2, 3, \dots, k$  dans laquelle nous interprêtons les termes

$$\frac{u - u_i}{u_{i+r-1} - u_i} B_{i,r-1}(u) \quad \text{et} \quad \frac{u_{i+r} - u}{u_{i+r} - u_{i+1}} B_{i+1,r-1}(u)$$

comme étant nuls, respectivement, chaque fois que  $u_{i+r-1} - u_i = 0$  et  $u_{i+r} - u_{i+1} = 0$ .

- **Support compact et positivité :**

(1.5) Théorème :

$$B_{i,k}(u) > 0 \quad \text{si } u_i < u < u_{i+k}$$

et

$$B_{i,k}(u) = 0 \quad \text{si } u \leq u_i \text{ et } u \geq u_{i+k}$$

Une démonstration se trouve dans [Bartels/Beatty/Barsky 87].



Ce résultat (avec le résultat suivant) établit la validité de l'emploi des B-splines comme fonctions de pondération pour définir des courbes.

- **Propriété de l'enveloppe convexe :**

Un résultat plus important, c'est que la somme des B-splines est égale à 1. Les deux propriétés de positivité et d'égalité de la somme à 1 entraînent la propriété de l'enveloppe convexe. Pour des B-splines d'ordre  $k$  cela signifie que, si  $C$  est une courbe établie à partir de points de définition  $P_i \in \mathbb{R}^q$  ( $q \geq 1$ ), c'est-à-dire :

$$C(u) = \sum_i P_i B_{i,k}(u)$$

Alors, chaque point de la courbe  $C$  est situé à l'intérieur de l'enveloppe convexe constituée par au plus  $k$  points de définition.

(1.6) Théorème :

Pour toute valeur de  $u \in [u_{k-1}, u_{m+1}]$ , on a :

$$\sum_{i=0}^m B_{i,k}(u) = 1$$

Une démonstration par récurrence se trouve dans [Bartels/Beatty/Barsky 87].

- **Dérivées des B-splines :**

Du fait que  $\frac{\partial}{\partial u} (u-t)_+^{k-1} = (k-1) (u-t)_+^{k-2}$  pour  $k \geq 2$  et que

$$B_{i,k}(u) = (-1)^k (\delta^{k-1}[u_{i+1}, \dots, u_{i+k} : t] - \delta^{k-1}[u_i, \dots, u_{i+k-1} : t]) (u-t)_+^{k-1}$$

on peut écrire, pour  $k \geq 2$

$$\frac{d}{du} B_{i,k}(u) = (-1)^k (\delta^{k-1}[u_{i+1}, \dots, u_{i+k} : t] - \delta^{k-1}[u_i, \dots, u_{i+k-1} : t]) (u-t)_+^{k-1}$$

En utilisant la définition (1.2), on obtient la formule de dérivation suivante:

$$(1.7) \quad \frac{d}{du} B_{i,k}(u) = (k-1) \left( \frac{B_{i,k-1}(u)}{u_{i+k-1} - u_i} - \frac{B_{i+1,k-1}(u)}{u_{i+k} - u_{i+1}} \right)$$

Pour des dérivées d'un ordre plus élevé, ce résultat pourrait être prolongé récursivement, mais les expressions qui en résultent deviendraient rapidement très compliquées. Un raisonnement plus efficace, qui est exposé dans [de Boor 78] et [Schumaker 81], nécessite l'examen des dérivées des combinaisons linéaires de B-splines, c'est-à-dire :

$$s = \sum_{i=0}^m \alpha_i B_{i,k}$$

Pour une telle combinaison, la dérivée première s'écrit

$$\frac{ds}{du} = (k-1) \sum_{i=0}^m \alpha_i \left( \frac{B_{i,k-1}}{u_{i+k-1} - u_i} - \frac{B_{i+1,k-1}}{u_{i+k} - u_{i+1}} \right)$$

ou, encore

$$\frac{ds}{du} = (k-1) \left( \alpha_0 \frac{B_{0,k-1}}{u_{k-1} - u_0} + \sum_{i=1}^m \frac{\alpha_i - \alpha_{i-1}}{u_{i+k-1} - u_i} B_{i,k-1} - \alpha_m \frac{B_{m+1,k-1}}{u_{m+k} - u_{m+1}} \right)$$

or

$$B_{0,k-1}(u) = B_{m+1,k-1}(u) = 0 \quad \text{pour tout } u \in [u_{k-1}, u_{m+1}[$$

d'où la formule

$$(1.8) \quad \frac{ds}{du}(u) = (k-1) \sum_{i=1}^m \frac{\alpha_i - \alpha_{i-1}}{u_{i+k-1} - u_i} B_{i,k-1}(u) \quad u \in [u_{k-1}, u_{m+1}[, (k \geq 2)$$

#### 1.4 Splines avec conditions de forme

En général, les fonctions splines peuvent présenter des inflexions ou des oscillations indésirables et faire perdre à la courbe ses caractéristiques de positivité, de monotonie ou de convexité. Nous allons examiner ces conditions de forme dans l'espace des splines  $S(k,m)$ . Pour cela, on va

considérer un élément  $s \in S(k,m)$  tel que :

$$s(u) = \sum_{i=0}^m \alpha_i B_{i,k}(u) \quad u \in [u_{k-1}, u_{m+1}[$$

avec

$$\alpha_i \in \mathbb{R}, i=0, \dots, m$$

- **Positivité**

On considère l'ensemble

$$\text{pos}(k,m) = \left\{ \alpha \in \mathbb{R}^{m+1} : \sum_{i=0}^m \alpha_i B_{i,k}(u) \geq 0, \forall u \in [u_{k-1}, u_{m+1}[ \right\}$$

C'est un ensemble convexe qui contient l'hyperoctant positif  $\mathbb{R}_+^{m+1}$ . Pour  $k=1, 2$  on a :

$$\text{pos}(1,m) = \text{pos}(2,m) = \mathbb{R}_+^{m+1}$$

mais, pour  $k \geq 3$ , cet ensemble est strictement plus grand (au sens de l'inclusion) que  $\mathbb{R}_+^{m+1}$ . Donc, une condition suffisante pour la positivité de  $s$  dans  $[u_{k-1}, u_{m+1}[$  est que l'on a :

$$\alpha_i \geq 0 \quad (i=0, \dots, m).$$

- **Monotonie**

(1.9) Proposition (condition suffisante de monotonie)

Une condition suffisante de monotonie de  $s = \sum_{i=0}^m \alpha_i B_{i,k}$  dans l'intervalle  $[u_{k-1}, u_{m+1}[$  est que les coefficients  $\{\alpha_i\}_0^m$  soient monotones. En plus, pour  $k \in \{2,3\}$  cette condition est nécessaire.

Preuve

D'après la formule de dérivation (1.8), on a :

$$\frac{ds}{du}(u) = (k-1) \sum_{i=1}^m \frac{\alpha_i - \alpha_{i-1}}{u_{i+k-1} - u_i} B_{i,k-1}(u) \quad u \in [u_{k-1}, u_{m+1}[ , (k \geq 2)$$

Considérons le cas de la monotonie croissante. Si  $\alpha_i - \alpha_{i-1} \geq 0$ , pour  $i=1, \dots, m$ , alors on a :

$$\boxed{\frac{ds}{du}(u) \geq 0 \quad \forall u \in [u_{k-1}, u_{m+1}[}$$

En plus, si  $k=2, 3$  alors  $\frac{ds}{du}$  est formé de morceaux de polynômes de degré 0 et 1 respectivement dans chaque intervalle  $[u_i, u_{i+1}[$   $i=1, \dots, m$ . Alors, pour  $k=3$  :

$$\frac{ds}{du}(u) \geq 0 \text{ dans } [u_2, u_{m+1}[ \Leftrightarrow \frac{ds}{du}(u_j) \geq 0 \quad (j=2, \dots, m+1)$$

On a ici :

$$B_{i,2}(u_j) = \begin{cases} 1 & \text{si } i=j-1 \\ 0 & \text{sinon} \end{cases}$$

donc :

$$\frac{ds}{du}(u) \geq 0 \text{ dans } [u_2, u_{m+1}[ \Leftrightarrow \alpha_i - \alpha_{i-1} \geq 0, i=1, \dots, m$$

Pour  $k=2$  :

$$\frac{ds}{du}(u) \geq 0 \text{ dans } [u_1, u_{m+1}[ \Leftrightarrow \frac{ds}{du}(u_j) \geq 0 \quad (j=1, \dots, m)$$

or

$$B_{i,1}(u_j) = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{sinon} \end{cases}$$

Par conséquent :

$$\frac{ds}{du}(u) \geq 0 \text{ dans } [u_1, u_{m+1}[ \Leftrightarrow \alpha_i - \alpha_{i-1} \geq 0, i=1, \dots, m$$

Enfin, pour  $k=1$  le résultat est évident.

- **Convexité**

Pour une fonction spline  $s \in S(k,m)$  telle que :

$$s = \sum_{i=0}^m \alpha_i B_{i,k} \quad \text{et} \quad k \geq 2$$

on considère les quantités :

$$(1.10) \quad \beta_i = \frac{\alpha_i - \alpha_{i-1}}{u_{i+k-1} - u_i} \quad i=1, \dots, m$$

Alors, on a la proposition suivante :

(1.11) Proposition (condition suffisante de convexité)

Pour que  $s \in S(k,m)$  avec  $k \geq 2$  soit convexe dans l'intervalle  $[u_{k-1}, u_{m+1}[$ , il suffit que les quantités  $\{\beta_i\}_1^m$  définies par (1.10) soient monotones croissantes. En plus, pour  $k \in \{2, 3, 4\}$  cette condition est nécessaire.

Preuve

En se servant de la formule (1.8) avec  $k \geq 3$  on peut écrire :

$$(i) \quad \frac{d^2s}{du^2}(u) = (k-1) \sum_{i=1}^m \frac{\alpha_i - \alpha_{i-1}}{u_{i+k-1} - u_i} \frac{d}{du} B_{i,k-1}(u)$$

et en appliquant (1.7) à  $B_{i,k-1}$  on obtient :

$$(ii) \quad \frac{d}{du} B_{i,k-1}(u) = (k-2) \left( \frac{B_{i,k-2}(u)}{u_{i+k-2} - u_i} - \frac{B_{i+1,k-2}(u)}{u_{i+k-1} - u_{i+1}} \right)$$

de (i) et (ii) on déduit :

$$\frac{d^2 s}{du^2}(u) = (k-1)(k-2) \sum_{i=1}^m \left( \frac{\beta_i}{u_{i+k-2} - u_i} B_{i,k-2}(u) - \frac{\beta_i}{u_{i+k-1} - u_{i+1}} B_{i+1,k-2}(u) \right)$$

ou encore :

$$\frac{d^2 s}{du^2}(u) = (k-1)(k-2) \sum_{i=2}^m \frac{\beta_i - \beta_{i-1}}{u_{i+k-2} - u_i} B_{i,k-2}(u) \quad (k \geq 2)$$

car

$$B_{1,k-2}(u) = B_{m+1,k-2}(u) = 0 \quad \text{pour tout } u \in [u_{k-1}, u_{m+1}[$$

Par conséquent, si  $\beta_i - \beta_{i-1} \geq 0$ ,  $i=2, \dots, m$  alors

$$\frac{d^2 s}{du^2}(u) \geq 0 \quad \text{pour tout } u \in [u_{k-1}, u_{m+1}[$$

pour  $k=2$  on a

$$\beta_i = \frac{\alpha_i - \alpha_{i-1}}{u_{i+1} - u_i} = \frac{s(u_{i+1}) - s(u_i)}{u_{i+1} - u_i} \quad (i=1, \dots, m)$$

et comme dans ce cas,  $s$  est de degré 1 dans chaque intervalle  $[u_i, u_{i+1}[$ , ( $i=1, \dots, m$ ) alors on a  $\beta_i - \beta_{i-1} \geq 0$ , ( $i=2, \dots, m$ ) si et seulement si  $s$  est convexe dans  $[u_1, u_{m+1}[$ .

Pour  $k=3, 4$  on remarque que  $\frac{d^2 s}{du^2}$  est formé de polynômes de degré 0 et 1 respectivement. Alors, pour  $k=3$  on a

$$\frac{d^2 s}{du^2} \geq 0 \Leftrightarrow \frac{d^2 s}{du^2}(u_j) \geq 0 \quad (j=2, \dots, m)$$

or

$$B_{i,1}(u_j) = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{sinon} \end{cases}$$

d'où

$$\frac{d^2s}{du^2} \geq 0 \Leftrightarrow \beta_i - \beta_{i-1} \geq 0 \quad i=2, \dots, m$$

De même pour  $k=4$  on a :

$$\frac{d^2s}{du^2} \geq 0 \Leftrightarrow \frac{d^2s}{du^2}(u_j) \geq 0 \quad j=3, \dots, m+1$$

or

$$B_{i,2}(u_j) = \begin{cases} 1 & \text{si } i=j-1 \\ 0 & \text{sinon} \end{cases}$$

d'où

$$\frac{d^2s}{du^2} \geq 0 \Leftrightarrow \beta_i - \beta_{i-1} \geq 0 \quad i=2, \dots, m$$

## 1.5 B-splines cubiques

Maintenant, nous allons examiner les fonctions splines de l'espace  $S(4,m)$ , qui sont formées de segments polynômiaux cubiques, et qui sont continues ainsi que leur dérivées première et seconde (continuité  $C^2$ ), aux points de raccordement entre les segments successifs. On a ici

$$u_0 < u_1 < u_2 < u_3 = u_{\min}$$

$$u_{\min} < u_4 < \dots < u_{m+1} = u_{\max}$$

et  $u_{\max} < u_{m+2} < u_{m+3} < u_{m+4}$

Les B-splines cubiques peuvent être définies, grâce au théorème (1.4), par la récurrence :

$$B_{i,4}(u) = \frac{u - u_i}{u_{i+3} - u_i} B_{i,3}(u) + \frac{u_{i+4} - u}{u_{i+4} - u_{i+1}} B_{i+1,3}(u)$$

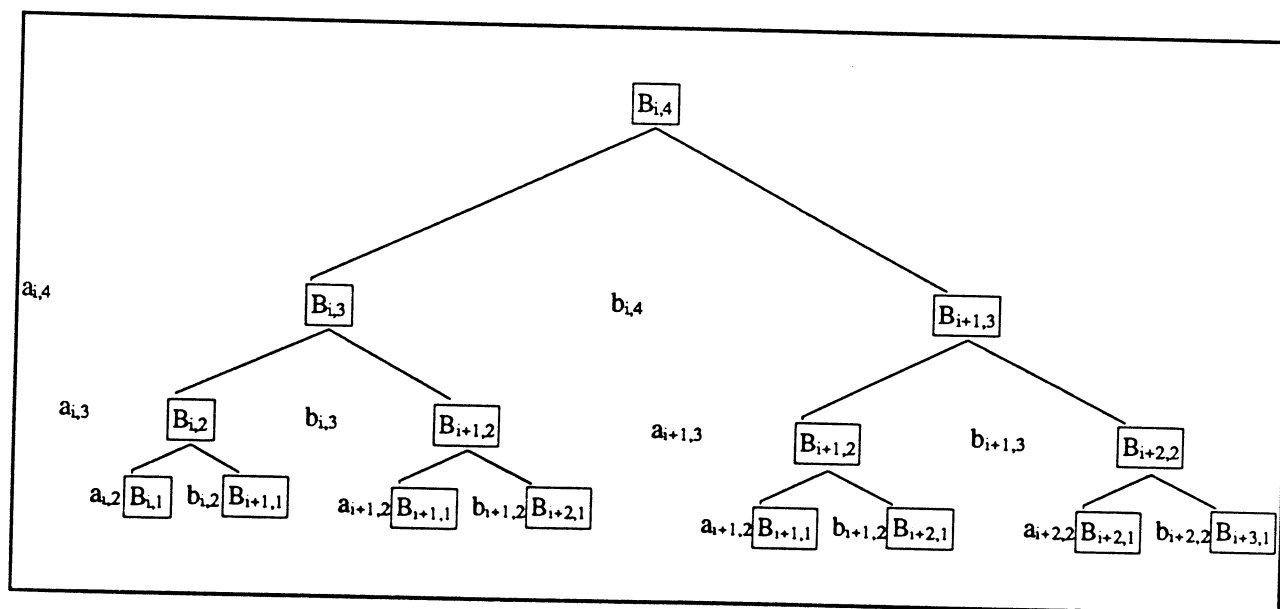
avec

$$B_{i,1}(u) = \begin{cases} 1 & \text{si } u_i \leq u < u_{i+1} \\ 0 & \text{sinon} \end{cases}$$

La B-spline  $B_{i,4}$  a pour support l'intervalle  $[u_i, u_{i+4}]$ . L'évaluation de  $B_{i,4}$  en un point  $u \in [u_i, u_{i+4}]$  peut se faire directement à partir de la relation de récurrence ci-dessus, mais cela coûte  $4(2^{k-1} - 1) = 28$  multiplications et  $5(2^{k-1} - 1) = 35$  additions. C'est la raison pour laquelle nous allons définir une nouvelle formulation des B-splines cubiques dans leurs intervalles de définitions, qui permet, à la fois, d'avoir une expression explicite et de faire un calcul économique. Pour cela on note :

$$(1.12) \quad a_{i,r} = \frac{u - u_i}{u_{i+r-1} - u_i} \quad \text{et} \quad b_{i,r} = \frac{u_{i+r} - u}{u_{i+r} - u_{i+1}} \quad (r \geq 2)$$

On remarque que  $b_{i,r} = 1 - a_{i,r}$ . Alors, on a le schéma d'évaluation suivant :



Par conséquent,  $B_{i,4}$  peut être exprimée par :



$$B_{i,4}(u) = \begin{cases} a_{i,4} a_{i,3} a_{i,2} & u \in [u_i, u_{i+1}[ \\ a_{i,4} (a_{i,3} b_{i,2} + b_{i,3} a_{i+1,2}) + b_{i,4} a_{i+1,3} a_{i+1,2} & u \in [u_{i+1}, u_{i+2}[ \\ a_{i,4} b_{i,3} b_{i+1,2} + b_{i,4} (a_{i+1,3} b_{i+1,2} + b_{i+1,3} a_{i+2,2}) & u \in [u_{i+2}, u_{i+3}[ \\ b_{i,4} b_{i+1,3} b_{i+2,2} & u \in [u_{i+3}, u_{i+4}[ \end{cases}$$

On définit sur chaque intervalle  $[u_i, u_{i+1}]$  les polynômes cubiques suivants :

$$(1.13) \quad \begin{cases} P_{0,i}(u) = a_{i,4} a_{i,3} a_{i,2} \\ P_{1,i}(u) = a_{i-1,4} (a_{i-1,3} b_{i-1,2} + b_{i-1,3} a_{i,2}) + b_{i-1,4} a_{i,3} a_{i,2} \\ P_{2,i}(u) = a_{i-2,4} b_{i-2,3} b_{i-1,2} + b_{i-2,4} (a_{i-1,3} b_{i-1,2} + b_{i-1,3} a_{i,2}) \\ P_{3,i}(u) = b_{i-3,4} b_{i-2,3} b_{i-1,2} \end{cases}$$

Alors, la B-spline  $B_{i,4}$  peut être exprimée par :

$$(1.14) \quad B_{i,4} = \begin{cases} P_{0,i} & \text{dans } [u_i, u_{i+1}[ \\ P_{1,i} & \text{dans } [u_{i+1}, u_{i+2}[ \\ P_{2,i} & \text{dans } [u_{i+2}, u_{i+3}[ \\ P_{3,i} & \text{dans } [u_{i+3}, u_{i+4}[ \\ 0 & \text{ailleurs} \end{cases}$$

Remarquons que le coût de calcul des polynômes  $P_{0,i}$  et  $P_{3,i}$  est de 5 multiplications et 6 additions, et celui des polynômes  $P_{1,i}$  et  $P_{2,i}$  est de 13 multiplications et 18 additions. Donc, en moyenne, l'évaluation de  $B_{i,4}$  coûte 9 opérations de multiplication et 12 opérations d'addition. D'où un gain de l'ordre de 65 % sur le temps de calcul par rapport à l'évaluation par récurrence.

## 1.6 B-splines cubiques uniformes

Lorsque les valeurs nodales  $\{u_i\}_0^{m+4}$  sont également espacées, les B-splines  $\{B_{i,4}\}_0^m$  sont appelées B-splines cubiques uniformes. Dans ce cas, si on note

$$h = u_i - u_{i-1} \quad \text{pour } i=1, \dots, m+4$$

alors, les termes de (1.12) s'écrivent :

$$(1.15) \quad a_{i,r} = \frac{u - u_i}{(r-1)h}, \quad b_{i,r} = \frac{u_{i+r} - u}{(r-1)h}$$

Par conséquent, les polynômes de (1.13) se simplifient :

$$(1.16) \quad \begin{cases} P_{0,i}(u) = \frac{1}{6} \left( \frac{u - u_i}{h} \right)^3 \\ P_{1,i}(u) = \frac{1}{6} \left[ -3 \left( \frac{u - u_{i+1}}{h} \right)^3 - 6 \left( \frac{u - u_{i+1}}{h} \right)^2 + 4 \right] \\ P_{2,i}(u) = \frac{1}{6} \left[ 3 \left( \frac{u - u_{i+1}}{h} \right)^3 + 3 \left( \frac{u - u_{i+1}}{h} \right)^2 - 3 \left( \frac{u - u_{i+1}}{h} \right) + 1 \right] \\ P_{3,i}(u) = - \frac{1}{6} \left( \frac{u - u_{i+1}}{h} \right)^3 \end{cases}$$

### Remarque 1.1

Si on pose  $\omega = \frac{u - u_i}{h}$  et  $\omega' = \frac{u - u_{i+1}}{h}$  alors, en utilisant le schéma de Horner, les polynômes de (1.16) peuvent être mis sous la forme:

$$(1.17) \quad \begin{cases} P_{0,i}(u) = \omega \times \omega \times \omega / 6 \\ P_{1,i}(u) = ((-0.5 \times \omega' - 1) \times \omega' + 2/3) \times \omega \\ P_{2,i}(u) = ((0.5 \times \omega' + 0.5) \times \omega' - 0.5) \times \omega + 1/6 \\ P_{3,i}(u) = - \omega' \times \omega' \times \omega' / 6 \end{cases}$$

Le coût de calcul de chacun de ces polynômes est donné par la table suivante:

	$P_{0,i}$	$P_{1,i}$	$P_{2,i}$	$P_{3,i}$
multiplications	4	4	4	4
additions	1	3	4	1

Table 1.1 : coût de calcul des polynômes de base

## 1.7 Calcul des valeurs des splines cubiques uniformes

Considérons la combinaison linéaire

$$(1.18) \quad s(u) = \sum_{i=0}^m \alpha_i B_{i,4}(u)$$

Considérons une valeur de  $u$ , telle que

$$u_i \leq u < u_{i+1} \quad \text{et} \quad 3 \leq i \leq m.$$

Dans un tel intervalle, la spline  $s$  devient simplement un polynôme de degré 3 ; par conséquent, elle peut être exprimée par :

$$(1.19) \quad s(u) = \sum_{j=0}^3 c_j [i] \left( \frac{u - u_i}{h} \right)^j$$

Les coefficients  $c_j [i]$  sont essentiellement les dérivées de  $s$  pour la valeur  $u_i$ , soit

$$c_j [i] = \frac{h^j}{j!} \frac{d^j s}{du^j} (u_i)$$

D'autre part, on a pour tout  $u \in [u_i, u_{i+1}[$  :

$$s(u) = \alpha_{i-3} B_{i-3,4}(u) + \alpha_{i-2} B_{i-2,4}(u) + \alpha_{i-1} B_{i-1,4}(u) + \alpha_i B_{i,4}(u) .$$

En utilisant (1.14) on obtient l'expression

$$(1.20) \quad s(u) = \alpha_{i-3} P_{3,i}(u) + \alpha_{i-2} P_{2,i}(u) + \alpha_{i-1} P_{1,i}(u) + \alpha_i P_{0,i}(u)$$

Or, par un calcul simple des dérivées de (1.16), on obtient :

$$(1.21) \quad \left\{ \begin{array}{ll} P_{0,i}(u_i) = 0 & P_{1,i}(u_i) = \frac{1}{6} \\ \frac{d}{du} P_{0,i}(u_i) = 0 & \frac{d}{du} P_{1,i}(u_i) = \frac{1}{2h} \\ \frac{d^2}{du^2} P_{0,i}(u_i) = 0 & \frac{d^2}{du^2} P_{1,i}(u_i) = \frac{1}{2h^2} \\ \frac{d^3}{du^3} P_{0,i}(u_i) = \frac{1}{6h^3} & \frac{d^3}{du^3} P_{1,i}(u_i) = \frac{-1}{2h^3} \\ \\ P_{2,i}(u_i) = \frac{2}{3} & P_{3,i}(u_i) = \frac{1}{6} \\ \frac{d}{du} P_{2,i}(u_i) = 0 & \frac{d}{du} P_{3,i}(u_i) = \frac{-1}{2h} \\ \frac{d^2}{du^2} P_{2,i}(u_i) = \frac{-1}{h^2} & \frac{d^2}{du^2} P_{3,i}(u_i) = \frac{1}{2h^2} \\ \frac{d^3}{du^3} P_{2,i}(u_i) = \frac{1}{2h^3} & \frac{d^3}{du^3} P_{3,i}(u_i) = \frac{-1}{6h^3} \end{array} \right.$$

Alors, les coefficients  $c_j[i]$  peuvent être calculés par :

$$(1.22) \quad \begin{pmatrix} c_0[i] \\ c_1[i] \\ c_2[i] \\ c_3[i] \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & 4 & 1 & 0 \\ -3 & 0 & 3 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{pmatrix} \begin{pmatrix} \alpha_{i-3} \\ \alpha_{i-2} \\ \alpha_{i-1} \\ \alpha_i \end{pmatrix}$$

avec un coût de calcul de 12 multiplications et 8 additions. La raison pour effectuer cette conversion provient de l'observation que la représentation de (1.19) peut être calculée plus efficacement que celle de (1.18).

Néanmoins, la conversion est coûteuse, si bien qu'elle sera seulement intéressante lorsque nous aurons à faire le calcul plusieurs fois dans l'intervalle  $[u_i, u_{i+1}[$ .

### Remarque 1.2

Pour toute valeur de  $u \in [u_i, u_{i+1}[$  on a d'après (1.6)

$$\sum_{j=0}^3 P_{j,i}(u) = 1$$

d'où

$$P_{2,i}(u) = 1 - P_{0,i}(u) - P_{1,i}(u) - P_{3,i}(u)$$

En portant ceci dans (1.20) on obtient

$$(1.23) \quad s(u) = (\alpha_{i-3} - \alpha_{i-2}) P_{3,i}(u) + (\alpha_{i-1} - \alpha_{i-2}) P_{1,i}(u) + (\alpha_i - \alpha_{i-1}) P_{0,i}(u) + \alpha_{i-2}$$

en conséquence, si on calcule  $s(u)$  par (1.23) en utilisant (1.17) pour l'évaluation des polynômes de base  $P_{0,i}$ ,  $P_{1,i}$  et  $P_{3,i}$  alors, le coût de l'opération est de 13 multiplications et 9 additions. D'autre part, une fois calculés les coefficients  $\{c_j^{[i]}\}_{j=0}^3$ ,  $s(u)$  peut être évalué à partir de l'expression (1.19) avec un coût de 4 multiplications et 4 additions.

## 1.8 Quelques opérateurs sur l'espace $S(k,m)$

Maintenant, nous allons étudier quelques opérateurs définis sur l'espace des splines  $S(k,m)$ . cela sera utile pour le développement des chapitres suivants.

- **Opérateur du lissage**

Pour toute fonction spline  $s \in S(k,m)$ , on fait correspondre l'intégrale :

$$(1.24) \quad L_d(s) = \int_{u_{k-1}}^{u_{m+1}} (s^{(d)}(u))^2 du \quad \text{avec } d \in \{1, 2, \dots, k-1\}$$

D'autre part,  $s$  est une combinaison linéaire de B-splines de la forme

$$s = \sum_{i=0}^m \alpha_i B_i$$

d'où, on peut écrire

$$L_d(s) = \int_{u_{k-1}}^{u_{m+1}} (s^{(d)}(u))^2 du = \sum_{i=0}^m \sum_{j=0}^m \alpha_i \alpha_j \int_{u_{k-1}}^{u_{m+1}} B_i^{(d)} B_j^{(d)} du$$

Si on pose pour tout  $i, j = 0, \dots, m$

$$(1.25) \quad \Omega_{i,j} = \int_{u_{k-1}}^{u_{m+1}} B_i^{(d)} B_j^{(d)} du$$

alors, on a

$$(1.26) \quad L_d(s) = \langle \alpha, \Omega \alpha \rangle_{\mathbb{R}^{m+1}}$$

avec  $\alpha = (\alpha_i)_{i=0, \dots, m}$  et  $\Omega = (\Omega_{i,j})_{\substack{i=0, \dots, m \\ j=0, \dots, m}}$

et  $\langle \cdot, \cdot \rangle_{\mathbb{R}^{m+1}}$  dénote le produit scalaire euclidien dans  $\mathbb{R}^{m+1}$ .

On remarque que  $\Omega$  est une matrice carrée symétrique semi-définie positive. De plus, on a

$$\Omega_{i,j} = 0 \quad \text{si} \quad |i - j| > k-1$$

Il sera donc suffisant de calculer les termes  $\Omega_{i,i}, \Omega_{i,i+1}, \dots, \Omega_{i,i+k-1}$ .

Dans le cas de B-splines cubiques uniformes et pour  $d=2$  on peut expliciter les termes de la matrice  $\Omega$  comme suit :

On utilise (1.16) pour déterminer les dérivées secondes des polynômes de base dans l'intervalle  $[u_i, u_{i+1}]$  :

$$(1.27) \quad \begin{cases} \tilde{P}_{0,i}(u) = \frac{1}{h^3}(u - u_i) \\ \tilde{P}_{1,i}(u) = \frac{-1}{h^3}[3(u - u_{i+1}) + 2h] \\ \tilde{P}_{2,i}(u) = \frac{1}{h^3}[3(u - u_{i+1}) + h] \\ \tilde{P}_{3,i}(u) = \frac{-1}{h^3}(u - u_{i+1}) \end{cases}$$

Si on pose pour  $0 \leq j_1, j_2 \leq 3$  :

$$\pi (P''_{j_1,i}, P''_{j_2,i}) = \begin{cases} \int_{u_i}^{u_{i+1}} P''_{j_1,i}(u) P''_{j_2,i}(u) du & \text{si } 3 \leq i < m+1 \\ 0 & \text{sinon} \end{cases}$$

alors, on peut écrire :

$$(1.28) \quad \begin{cases} \Omega_{i,i} = \pi(P''_{0,i}, P''_{0,i}) + \pi(P''_{1,i}, P''_{1,i}) + \pi(P''_{2,i}, P''_{2,i}) + \pi(P''_{3,i}, P''_{3,i}) \\ \Omega_{i,i+1} = \pi(P''_{1,i+1}, P''_{0,i+1}) + \pi(P''_{2,i+2}, P''_{1,i+2}) + \pi(P''_{3,i+3}, P''_{2,i+3}) \\ \Omega_{i,i+2} = \pi(P''_{2,i+2}, P''_{0,i+2}) + \pi(P''_{3,i+3}, P''_{1,i+3}) \\ \Omega_{i,i+3} = \pi(P''_{3,i+3}, P''_{0,i+3}) \end{cases}$$

On peut calculer facilement les quantités  $\pi (P''_{j_1,i}, P''_{j_2,i})$  pour  $0 \leq j_1, j_2 \leq 3$  et  $3 \leq i < m+1$  ; on obtient les valeurs suivantes :

$$(1.29) \quad \begin{cases} \pi(P''_{0,i}, P''_{0,i}) = \frac{1}{3h^3} & \pi(P''_{2,i}, P''_{1,i}) = -\frac{1}{2h^3} \\ \pi(P''_{1,i}, P''_{1,i}) = \frac{1}{h^3} & \pi(P''_{3,i}, P''_{2,i}) = -\frac{1}{2h^3} \\ \pi(P''_{2,i}, P''_{2,i}) = \frac{1}{h^3} & \pi(P''_{2,i}, P''_{0,i}) = 0 \\ \pi(P''_{3,i}, P''_{3,i}) = \frac{1}{3h^3} & \pi(P''_{3,i}, P''_{1,i}) = 0 \\ \pi(P''_{1,i}, P''_{0,i}) = -\frac{1}{2h^3} & \pi(P''_{3,i}, P''_{0,i}) = \frac{1}{6h^3} \end{cases}$$

d'où la forme explicite de  $\Omega$  :





i) cas de mesures ponctuelles :

Sur l'espace de splines  $S(k,m)$ , on définit l'opérateur des moindres carrés pondérés pour des mesures ponctuelles  $(t_j, Z_j)_{j=1, \dots, n}$  par

$$(1.31) \quad M_p^2(s) = \sum_{j=1}^n \omega_j (s(t_j) - Z_j)^2$$

Soit  $s = \sum_{i=0}^m \alpha_i B_i$ . Alors, si on pose :

$$M = (M_{ij})_{\substack{i=0, \dots, m \\ j=1, \dots, n}} \quad \text{avec} \quad M_{ij} = B_i(t_j)$$

alors, on peut écrire

$$M_p^2(s) = \langle W(M^t \alpha - Z), W(M^t \alpha - Z) \rangle_{\mathbb{R}^n}$$

ou, encore

$$(1.32) \quad M_p^2(s) = \langle \alpha, MW^2 M^t \alpha \rangle_{\mathbb{R}^{m+1}} - 2 \langle MW^2 Z, \alpha \rangle_{\mathbb{R}^{m+1}} + \langle Z, W^2 Z \rangle_{\mathbb{R}^n}$$

ii) cas de mesures par intervalles :

Les mesures  $\{Z_j\}_{j=1, \dots, n}$  représentent les intégrales d'une certaine fonction inconnue sur les intervalles  $\{I_j\}_{j=1, \dots, n}$ . Alors, pour tout  $s \in S(k,m)$  l'opérateur des moindres carrés pondérés est défini par :

$$(1.33) \quad M_1^2(s) = \sum_{j=1}^n \omega_j \left( \int_{I_j} s(t) dt - Z_j \right)^2$$

Si on pose :

$$M_{i,j} = \int_{I_j} B_i(t) dt$$

alors

$$M_I^2(s) = \langle W(M^t\alpha - Z), W(M^t\alpha - Z) \rangle_{\mathbb{R}^n}$$

ou encore

$$(1.34) \quad M_I^2(s) = \langle \alpha, MW^2M^t\alpha \rangle_{\mathbb{R}^{m+1}} - 2 \langle MW^2Z, \alpha \rangle_{\mathbb{R}^{m+1}} + \langle Z, W^2Z \rangle_{\mathbb{R}^n}$$

Dans le cas de B-splines cubiques uniformes, on peut expliciter la matrice M en utilisant les polynômes de base (1.16) :

$$\begin{aligned} M_{i,j} &= \int_{I_j} B_i(t) dt \\ &= \int_{I_j \cap [u_i, u_{i+1}]} P_{0,i}(t) dt + \int_{I_j \cap [u_{i+1}, u_{i+2}]} P_{1,i+1}(t) dt + \\ &\quad \int_{I_j \cap [u_{i+2}, u_{i+3}]} P_{2,i+2}(t) dt + \int_{I_j \cap [u_{i+3}, u_{i+4}]} P_{3,i+3}(t) dt \end{aligned}$$

Il sera donc suffisant de savoir calculer les intégrales de la forme :

$$\int_{t_1}^{t_2} P_{r,i}(t) dt \quad 0 \leq r \leq 3$$

avec  $[t_1, t_2] \subseteq [u_i, u_{i+1}]$ .

En utilisant (1.16), on trouve facilement les formules d'intégration suivantes :

$$(1.35) \quad \left\{ \begin{array}{l} \int_{t_1}^{t_2} P_{0,i}(t) dt = \frac{h}{24} \left[ \left( \frac{t_2 - u_i}{h} \right)^4 - \left( \frac{t_1 - u_i}{h} \right)^4 \right] \\ \int_{t_1}^{t_2} P_{1,i}(t) dt = \frac{h}{24} \left( -3 \left[ \left( \frac{t_2 - u_{i+1}}{h} \right)^4 - \left( \frac{t_1 - u_{i+1}}{h} \right)^4 \right] - 8 \left[ \left( \frac{t_2 - u_{i+1}}{h} \right)^3 - \left( \frac{t_1 - u_{i+1}}{h} \right)^3 \right] + 16 \left( \frac{t_2 - t_1}{h} \right) \right) \\ \int_{t_1}^{t_2} P_{2,i}(t) dt = \frac{h}{24} \left( 3 \left[ \left( \frac{t_2 - u_{i+1}}{h} \right)^4 - \left( \frac{t_1 - u_{i+1}}{h} \right)^4 \right] + 4 \left[ \left( \frac{t_2 - u_{i+1}}{h} \right)^3 - \left( \frac{t_1 - u_{i+1}}{h} \right)^3 \right] - 6 \left[ \left( \frac{t_2 - u_{i+1}}{h} \right)^2 - \left( \frac{t_1 - u_{i+1}}{h} \right)^2 \right] + 4 \left( \frac{t_2 - t_1}{h} \right) \right) \\ \int_{t_1}^{t_2} P_{3,i}(t) dt = \frac{-h}{24} \left[ \left( \frac{t_2 - u_{i+1}}{h} \right)^4 - \left( \frac{t_1 - u_{i+1}}{h} \right)^4 \right] \end{array} \right.$$



## CHAPITRE II

### **Modélisation de coupes pétrolières**



## 2.0 Introduction

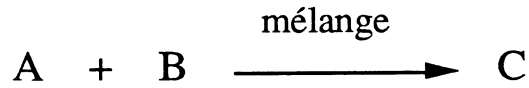
Avant de traiter un brut donné dans une raffinerie, on le soumet à une analyse de laboratoire dont le but est de guider les opérations industrielles en déterminant en quelles fractions il est le plus avantageux de le "découper". L'analyse la plus directe est la distillation fractionnée ; c'est un procédé de séparation qui utilise la différence des points d'ébullition entre des produits qui se trouvent mélangés et même dissous les uns dans les autres, dans la matière première. On chauffe une certaine quantité du brut ; le liquide se vaporise et par condensation on recueille au cours du temps des fractions de volume égale ; pour chacune d'elles, on note les températures correspondant au début et à la fin du passage de la fraction ; la moyenne de ces deux valeurs est appelée température T.B.P. (True Boiling Point). Aussi, on fait des mesures pour voir si la densité, l'indice d'octane, etc., sont convenables. En opérant de cette manière, sur chaque fraction, on obtient des courbes qui reflètent en un certain sens la composition et la répartition des hydrocarbures dans l'échantillon analysé.

### 2.1 Définitions et notations

#### (2.1) Définition

Si on note  $F$  une propriété physique d'un mélange, et  $m$  sa masse, alors on dit que  $F$  est additif (ou conservatif) si l'on a :





$$m_A F_A + m_B F_B = (m_A + m_B) F_C$$

où A et B sont des mélanges quelconques.

Certaines propriétés ne sont pas additives ; c'est le cas de la viscosité et du point de trouble. Nous rappelons que tout au long de ce chapitre nous supposons que les propriétés sont additives. Nous désignons par analyse additive la mesure quantitative d'une espèce dans chaque fraction du brut. Pour une analyse additive, on dispose d'un ensemble fini de couples  $(I_k, Z_k)$ ,  $k=1, \dots, n$ , avec  $I_k$  l'intervalle entre la température initiale et la température finale de la coupe  $k$ , et  $Z_k$  est la mesure expérimentale de la quantité de propriété additive dans  $I_k$ . Alors, il existe une fonction densité  $f$  de la propriété  $F$  telle que :

$$Z_k = \int_{I_k} f(T) \cdot dm(T) \quad k=1, \dots, n$$

### Remarque 2.1

En général, on peut avoir deux types de collections de coupes : d'une part, des coupes adjacentes qui se succèdent, et d'autre part, des coupes chevauchées réparties d'une manière quelconque sur l'ensemble des températures. L'ensemble des couples  $\{(I_k, Z_k)\}_{k=1}^n$  peuvent être représentés graphiquement par un histogramme (figure 2.1).

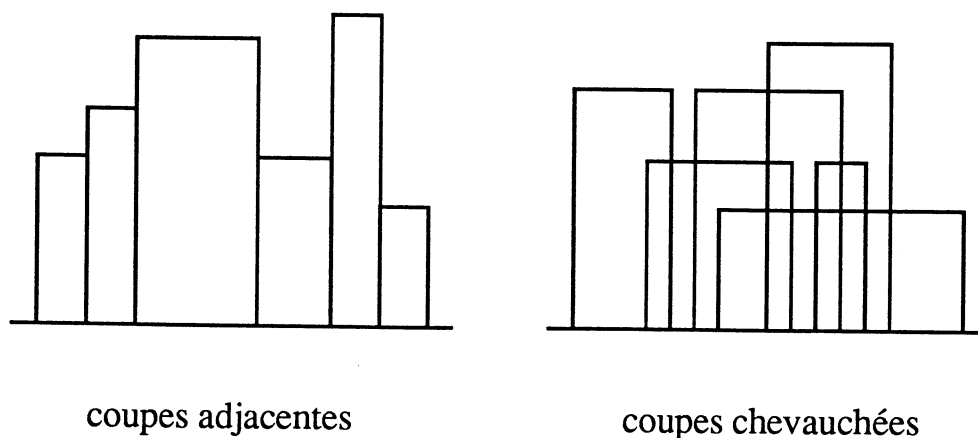


Figure 2.1 : Histogrammes des mesures

## (2.2) Problème

Pour l'ensemble de coupes  $I_k = [T_{i_k}, T_{f_k}]$ ,  $k=1, \dots, n$ , on pose :  
 $T_{\min} = \min \{T_{i_k}\}_{k=1}^n$  et  $T_{\max} = \max \{T_{f_k}\}_{k=1}^n$ . Alors, notre problème consiste à déterminer une fonction approximant la densité des mesures  $\frac{dm}{dT} \cdot f$  dans l'intervalle  $[T_{\min}, T_{\max}]$ .

### Remarque 2.2

Par sa nature physique, la densité des mesures  $\frac{dm}{dT} \cdot f$  est positive ; par conséquent, toute solution du problème (2.2) ne sera pas valable si elle ne satisfait pas cette condition de positivité.

## 2.2 Modélisation

Notre but est, d'une part de représenter les propriétés conservatives par des fonctions lisses continûment dérivables et permettant ainsi à l'utilisateur de les calculer sur n'importe quelle coupe désirée, et d'autre part de lui permettre une exploitation synthétique par une visualisation

graphique des différentes courbes, surfaces et courbes de niveaux. Dans [Slaoui 86], on propose une modélisation par des histosplines ; néanmoins, ceci n'a pas permis d'assurer les caractéristiques de forme (positivité et monotonie) propres aux propriétés physiques.

Dans [Boneva/Kendall/Stefanov 71] on a largement étudié les histogrammes adjacents et leur modélisation par des splines quadratiques (histosplines à une variable).

Pour atteindre ce but, nous proposons trois modèles dont le premier est un modèle de base (modèle sans conditions de forme). Ces modèles sont basés sur l'approximation de données au sens des moindres carrés, la fonction approximante étant une combinaison linéaire de B-splines cubiques. On introduit un terme de régularisation basé sur l'intégrale du carré de la dérivée seconde de la spline. Ce terme permet de mieux doser le lissage et il permet de plus de reconstituer correctement la fonction sous-jacente aux données même quand celles-ci manquent dans certains sous-intervalles de l'intervalle  $[T_{\min}, T_{\max}]$  du problème (2.2).

Dans toute la suite, on considère l'ensemble de valeurs nodales :

$$u_0 < u_1 < u_2 < u_3 < \dots < u_{m+1} < u_{m+2} < u_{m+3} < u_{m+4}$$

telles que :

$$u_3 = T_{\min} \quad ; \quad u_{m+1} = T_{\max}$$

et

$$u_i - u_{i-1} = h \quad i=1, \dots, m+4.$$

on utilisera l'espace  $S(4,m)$  des splines cubiques uniformes définies sur  $[T_{\min}, T_{\max}]$  avec  $m \geq 3$ . Alors, toute fonction  $s \in S(4,m)$  peut être mise sous la forme :

$$s = \sum_{i=0}^m \alpha_i B_i$$

en plus, on suppose que les intervalles  $\{I_k\}_{k=1}^n$  ne sont pas tous centrés sur un même point de l'intervalle  $[T_{\min}, T_{\max}]$ , (ceci intervient dans la suite, pour assurer l'unicité de la solution).

### 2.3 Modèle I

Dans ce modèle, on considère le problème de minimisation suivant :

$$(P.2.1) \quad \left\{ \min_{s \in S(4,m)} \left\{ \tau \int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt + \sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 \right\} \right\}$$

où  $\tau$  est le paramètre du lissage et  $\omega_k, k=1, \dots, n$  sont des poids dépendant de l'erreur expérimentale sur les mesures  $Z_k, k=1, \dots, n$  (cf. § 2.7).

Ce problème est étudié par P.J. LAURENT dans [Laurent 72] dans un contexte plus général. En se servant des théorèmes d'existence et d'unicité, on peut affirmer facilement qu'il existe une solution unique du (P.2.1).

En utilisant (1.26), le terme de régularisation s'écrit :

$$(2.3) \quad \int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt = \langle \alpha, \Omega \alpha \rangle_{\mathbb{R}^{m+1}}$$

$\Omega$  est la matrice carré symétrique semi-définie positive donnée par (1.30).

De même, la relation (1.34) permet d'écrire :

$$(2.4) \quad \sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 = \langle W(M^t \cdot \alpha - Z), W(M^t \cdot \alpha - Z) \rangle_{\mathbb{R}^n}$$

avec

$$M = (M_{ij})_{\substack{i=0, \dots, m \\ j=1, \dots, n}} \quad ; \quad M_{ij} = \int_{I_j} B_i(t) dt$$

et

$$W^2 = \text{diag} (\omega_1, \dots, \omega_n).$$

les éléments  $\{M_{ij}\}_{\substack{i=0, \dots, m \\ j=1, \dots, n}}$  se calculent facilement, grâce aux formules d'intégration (1.35). Alors, si on pose :

$$(2.5) \quad \begin{cases} \psi = \tau \Omega + MW^2M^t \\ b = MW^2Z \\ c = Z^tW^2Z \end{cases}$$

Le problème (P.2.1) est équivalent au problème :

$$(P'.2.1) \quad \begin{cases} \min \{ \langle \alpha, \psi.\alpha \rangle - 2 \langle b, \alpha \rangle + c \} \\ \alpha \in \mathbb{R}^{m+1} \end{cases}$$

La solution du problème (P'.2.1) est caractérisée par la condition d'optimalité variationnelle :

$$\text{grad}_{\alpha} \{ \langle \alpha, \psi.\alpha \rangle - 2 \langle b, \alpha \rangle + c \} = 0$$

c'est-à-dire par l'équation :

$$(2.6) \quad \psi.\alpha - b = 0$$

La matrice  $\psi$  étant définie positive, on peut écrire :

$$(2.7) \quad \alpha = \psi^{-1}.b$$

### Remarque 2.3

Ce modèle ne peut pas garantir la positivité de la fonction spline obtenue. D'autre part, ce modèle ne tient pas compte d'une éventuelle information sur la forme (monotonie) du signal à reconstituer.

## 2.4 Modèle II

Nous avons vu au § 1.4 que l'ensemble  $\text{pos}(4,m)$  est convexe et contient, d'une façon stricte, l'hyperoctant positif  $\mathbb{R}_+^{m+1}$ . Une manière simple pour imposer la positivité d'une fonction spline  $s \in S(4,m)$  telle que  $s = \sum_{i=0}^m \alpha_i B_i$  consiste à choisir des coefficients  $\{\alpha_i\}_{i=0}^m$  positifs. Nous rappelons que la positivité des  $\{\alpha_i\}_{i=0}^m$  est une condition suffisante, pour assurer la positivité de  $s$ , mais pas nécessaire. D'autre part, la proposition (1.9) fournit une condition suffisante de monotonie, par le choix des coefficients monotones.

Dans ce modèle, nous allons tenir compte de la positivité de la solution du problème (2.2) ainsi que d'une éventuelle caractéristique de forme concernant la monotonie de cette solution. Alors, trois cas peuvent se présenter :

(i) Positivité sans monotonie : on a les contraintes simples

$$(2.8) \quad \alpha_i \geq 0, i=0, \dots, m$$

(ii) Positivité avec monotonie croissante : On a les contraintes suivantes :

$$(2.9) \quad \begin{cases} \alpha_i \geq 0 & i=0, \dots, m \\ \alpha_i - \alpha_{i-1} \geq 0 & i=1, \dots, m \end{cases}$$

(iii) Positivité avec monotonie décroissante : l'ensemble des contraintes à imposer est :

$$(2.10) \quad \begin{cases} \alpha_i \geq 0 & i=0, \dots, m \\ \alpha_i - \alpha_{i-1} \leq 0 & i=1, \dots, m \end{cases}$$

On remarque que les contraintes du cas (ii), peuvent se mettre sous la forme:

$$0 \leq \alpha_0 \leq \alpha_1 \leq \dots \leq \alpha_m$$

ou encore

$$(2.11) \quad \begin{cases} \alpha_0 \geq 0 \\ \alpha_i - \alpha_{i-1} \geq 0, i=1, \dots, m \end{cases}$$

De même, pour le cas (iii) on a :

$$\alpha_0 \geq \alpha_1 \geq \dots \geq \alpha_m \geq 0$$

ou encore

$$(2.12) \quad \begin{cases} \alpha_{i-1} - \alpha_i \geq 0, i=1, \dots, m \\ \alpha_m \geq 0 \end{cases}$$

On considère le problème d'optimisation suivant :

$$(P.2.2) \quad \begin{cases} \min \left\{ \tau \int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt + \sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 + Q(G^{-1} \cdot \alpha) \right\} \\ s \in S(4,m) \\ s = \sum_{i=0}^m \alpha_i B_i \\ G^{-1} \cdot \alpha \geq 0 \end{cases}$$

où  $G$  est une matrice  $(m+1) \times (m+1)$  permettant d'exprimer les conditions de monotonie et de positivité des coefficients  $\{\alpha_i\}_{i=0}^m$  sous la forme :

$$G^{-1} \cdot \alpha \geq 0$$

et ceci dans les trois cas suivants :

- positivité seulement :

$$G = I_{m+1} = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & 1 \end{pmatrix}$$

- positivité et monotonie croissante :

$$G = \begin{pmatrix} 1 & & & & \\ 1 & 1 & & & \\ \cdot & \cdot & \cdot & & \\ \cdot & & & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & 1 \end{pmatrix} \quad \text{et} \quad G^{-1} = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & & -1 & 1 \end{pmatrix}$$



- positivité et monotonie décroissante :

$$G = \begin{pmatrix} 1 & 1 & \cdot & \cdot & \cdot & 1 \\ & 1 & 1 & \cdot & \cdot & 1 \\ & & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot \\ & & & & & 1 \end{pmatrix} \quad \text{et} \quad G^{-1} = \begin{pmatrix} 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & -1 \\ & & & & & 1 \end{pmatrix}$$

et Q est une application de  $(\mathbb{R}_+^*)^{m+1}$  dans  $\mathbb{R}_+$  définie par :

$$(2.13) \quad Q(X) = \sum_{i=0}^m P(X_i)$$

avec P une fonction réelle définie sur  $]0, +\infty[$  (on supposera que  $P(t) = +\infty$  si  $t \notin ]0, +\infty[$ ) ayant les propriétés suivantes :

$$(2.14) \quad \begin{cases} (a) & P \text{ est convexe et deux fois continûment dérivable sur } ]0, \infty[ \\ (b) & P(t) \geq 0, \forall t \in ]0, \infty[ \\ (c) & \lim_{t \rightarrow 0^+} P(t) = +\infty \\ (d) & \lim_{t \rightarrow +\infty} P(t) = 0 \end{cases}$$

L'allure de cette fonction est représentée par la figure 2.2. Le terme  $Q(G^{-1}.\alpha)$  est un terme de pénalisation intérieure qui "repousse" la fonction de façon infinie lorsque une contrainte est en train de devenir active. Donc, à l'optimum, les contraintes seront strictes.

Comme dans le modèle I, le terme  $\int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt$  peut s'exprimer en

fonction de  $\alpha$  par :

$$\int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt = \langle \alpha, \Omega.\alpha \rangle_{\mathbb{R}^{m+1}}$$

avec  $\Omega$  est la matrice de (1.30). De même, on a :

$$\sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 = \langle W(M^t \cdot \alpha - Z), W(M^t \cdot \alpha - Z) \rangle_{\mathbb{R}^n}$$

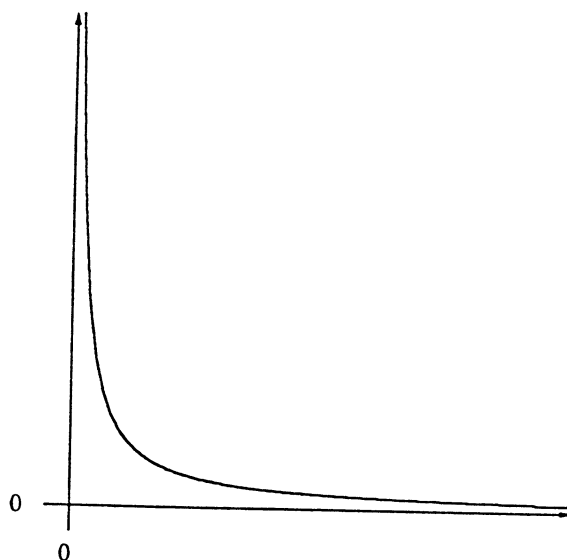


Figure 2.2 : allure de la fonction P.

En faisant le changement de variable suivant :

$$(2.15) \quad X = G^{-1} \cdot \alpha \quad \Leftrightarrow \quad \alpha = G \cdot X$$

on peut écrire :

- $\int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt = \langle X, G^t \Omega G \cdot \alpha \rangle_{\mathbb{R}^{m+1}}$
- $\sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 = \langle W(M^t G \cdot X - Z), W(M^t G \cdot X - Z) \rangle_{\mathbb{R}^n}$

- $Q(G^{-1}.\alpha) = Q(X)$

Donc, le problème (P.2.2) se transforme en :

$$(P'.2.2) \quad \begin{cases} \min \{ \langle X, \Psi.X \rangle - 2 \langle b, X \rangle + c + Q(X) \} \\ X \in \mathbb{R}^{m+1} \\ X \geq 0 \end{cases}$$

avec

$$(2.16) \quad \begin{cases} \Psi = G^t [\tau \Omega + M W^2 M^t] G \\ b = G^t M W^2 Z \\ c = Z^t W^2 Z \end{cases}$$

La matrice  $\Psi$  est une matrice symétrique définie positive. Ainsi, l'application :

$$X \rightarrow \langle X, \Psi.X \rangle - 2 \langle b, X \rangle + c$$

est strictement convexe sur  $\mathbb{R}_+^{m+1}$ . L'application  $Q$  est aussi convexe (convexité de la fonction  $P$  sur  $\mathbb{R}_+$ ) et prend la valeur  $+\infty$  sur le bord de l'hyperoctant positif  $\mathbb{R}_+^{m+1}$ . Par conséquent, le problème (P'.2.2) a une solution unique appartenant à l'intérieur de  $\mathbb{R}_+^{m+1}$ . Cette solution peut être obtenue en résolvant (P'.2.2) par une méthode itérative (par exemple une méthode de descente [voir chapitre suivant]).

#### Remarque 2.4

Cette modélisation permet d'obtenir une fonction spline d'approximation qui assure la positivité (et éventuellement la monotonie) dans l'intervalle  $[T_{\min}, T_{\max}]$ . D'autre part, comme on a vu au §1.4, l'ensemble de splines de l'espace  $S(4,m)$  à coefficients positifs (resp.

monotones) est plus petit, au sens de l'inclusion, que l'ensemble de splines de  $S(4,m)$  positives (resp. monotones) dans l'intervalle  $[T_{\min}, T_{\max}]$ , ce qui limite la possibilité de reconstituer, d'une façon optimale, la fonction sous-jacente aux données.

## 2.5 Modèle III

Malgré la simplicité de la manière dont on a imposé les conditions de forme dans le modèle précédent, elle est restrictive dans le sens où elle exclut une partie de l'espace  $S(4,m)$  susceptible de contenir la meilleure solution répondant aux caractéristiques de forme souhaitées. Une idée simple consiste à considérer ponctuellement des conditions de forme nécessaires sur un ensemble de points régulièrement distribués dans l'intervalle du problème. On considère donc une subdivision régulière de  $[T_{\min}, T_{\max}]$  que l'on note  $\{v_j\}_{j=1}^v$  telle que :

$$\{u_i : i=3, \dots, m+1\} \subseteq \{v_j : j=1, \dots, v\}$$

De plus, on considère

$$\varepsilon_1 = \begin{cases} 1 & \text{s'il y a monotonie} \\ 0 & \text{sinon} \end{cases}$$

et

$$\varepsilon_2 = \begin{cases} 1 & \text{s'il s'agit d'une monotonie croissante} \\ -1 & \text{s'il s'agit d'une monotonie décroissante} \end{cases}$$

Maintenant, soit  $s \in S(4,m)$  telle que  $s = \sum_{i=0}^m \alpha_i B_i$  et  $s(t) \geq 0$   $\forall t \in [T_{\min}, T_{\max}]$ , alors, on a nécessairement :

$$(2.17) \quad s(v_j) \geq 0, j=1, \dots, v$$

ou

$$\sum_{i=0}^m \alpha_i B_i(v_j) \geq 0, j=1, \dots, v$$

On remarque que cette condition est cette fois nécessaire, mais qu'elle n'est pas suffisante. Néanmoins, pour une subdivision  $\{v_j\}_{j=1}^v$  suffisamment fine on peut se rapprocher de la condition exacte de positivité.

On note

$$C_{ji}^{[0]} = B_i(v_j) \quad \text{pour } i=0, \dots, m \quad \text{et} \quad j=1, \dots, v.$$

On pose

$$C^{[0]} = \left( C_{ji}^{[0]} \right)_{\substack{i=0, \dots, m \\ j=1, \dots, v}}$$

Alors, les contraintes (2.17) se mettent sous la forme :

$$(2.18) \quad \boxed{C^{[0]} \cdot \alpha \geq 0}$$

De même, si  $s$  est monotone dans l'intervalle  $[T_{\min}, T_{\max}]$ , on a alors :

$$\varepsilon_2 \cdot \frac{ds}{dt}(t) \geq 0 \quad \forall t \in [T_{\min}, T_{\max}]$$

d'où

$$(2.19) \quad \varepsilon_2 \cdot \frac{ds}{dt}(v_j) \geq 0 \quad j=1, \dots, v$$

ou encore

$$\varepsilon_2 \cdot \sum_{i=0}^m \alpha_i B_i(v_j) \geq 0 \quad j=1, \dots, v$$

On pose

$$C^{[1]} = \left( C_{ji}^{[1]} \right)_{\substack{i=0, \dots, m \\ j=1, \dots, v}}$$

avec

$$C_{ji}^{[1]} = \frac{dB_i}{dt}(v_j) \text{ pour } i=0, \dots, m \quad \text{et} \quad j=1, \dots, v$$

Les conditions (2.19) peuvent donc s'écrire sous la forme :

$$(2.20) \quad \boxed{\varepsilon_2 \cdot C^{[1]} \cdot \alpha \geq 0}$$

On considère le problème de minimisation suivant :

$$(P.2.3) \quad \left\{ \begin{array}{l} \min \left\{ \tau \int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt + \sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 + Q(\alpha) \right\} \\ s \in S(4, m) \\ s = \sum_{i=0}^m \alpha_i B_i \\ C^{[0]} \cdot \alpha \geq 0 \\ \varepsilon_1 \cdot \varepsilon_2 \cdot C^{[1]} \cdot \alpha \geq 0 \end{array} \right.$$

où

$$(2.21) \quad Q(\alpha) = \sum_{j=1}^v \left[ P \left( \sum_{i=0}^m \alpha_i B_i(v_j) \right) + \varepsilon_1 P \left( \varepsilon_2 \sum_{i=0}^m \alpha_i \frac{dB_i}{dt}(v_j) \right) \right]$$

et P est une fonction réelle vérifiant (2.14). Comme dans les deux modèles précédemment décrits, on peut exprimer les différents termes du (P.2.3) en fonction des coefficients  $\{\alpha_i\}_{i=0, \dots, m}$  de la spline s sur les B-splines cubiques :

$$\bullet \quad \int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt = \langle \alpha, \Omega \alpha \rangle_{\mathbb{R}^{m+1}}$$

- $\sum_{k=1}^n \omega_k \left( \int_{I_k} s(t) dt - Z_k \right)^2 = \langle W(M^t \cdot \alpha - Z), W(M^t \cdot \alpha - Z) \rangle_{\mathbb{R}^n}$
- $Q(\alpha) = \sum_{j=1}^v \left[ P(C_j^{[0]} \cdot \alpha) + \varepsilon_1 \cdot P(\varepsilon_2 \cdot C_j^{[1]} \cdot \alpha) \right]$

alors, le problème (P.2.3) est équivalent au problème suivant :

$$(P'.2.3) \quad \begin{cases} \min \{ \langle \alpha, \Psi \cdot \alpha \rangle - 2 \langle b, \alpha \rangle + c + Q(\alpha) \} \\ \alpha \in \mathbb{R}^{m+1} \\ C^{[0]} \cdot \alpha \geq 0 \\ \varepsilon_1 \cdot \varepsilon_2 \cdot C^{[1]} \cdot \alpha \geq 0 \end{cases}$$

avec

$$(2.22) \quad \begin{cases} \Psi = \tau \cdot \Omega + M W^2 M^t \\ b = M W^2 Z \\ c = Z^t W^2 Z \end{cases}$$

L'application  $Q$  est positive et strictement convexe. D'autre part, la matrice  $\Psi$  est symétrique définie positive. Par conséquent, le problème (P'.2.3) admet une solution unique  $\alpha^*$  qui appartient en fait à l'ensemble :

$$\left\{ \alpha \in \mathbb{R}^{m+1} : C^{[0]} \cdot \alpha > 0 \text{ et } \varepsilon_1 \cdot \varepsilon_2 \cdot C^{[1]} \cdot \alpha > 0 \right\}$$

Cette solution peut être obtenue par un algorithme itératif d'optimisation convexe (voir chapitre suivant).

### Remarque 2.5

Ce modèle fournit une solution qui respecte les conditions de positivité

et de monotonie sur une grille de points uniformément distribués dans l'intervalle  $[T_{\min}, T_{\max}]$ . Cependant, lorsque le nombre de points  $v$  est faible (proche de  $m-1$ ) il y a un risque assez grand, d'avoir une solution qui ne vérifie pas la positivité et la monotonie entre ces points, et lorsque  $v$  est grand, le volume de calcul nécessaire pour résoudre le problème (P'.2.3) est considérable: ceci résulte du fait que le nombre de contraintes dans (P'.2.3) est proportionnel à  $v$ . En pratique, le choix du  $v$  doit être ni trop petit ni trop grand afin de réaliser un compromis entre volume de calcul et caractéristiques de forme.

## 2.6 Choix des poids $\{\omega_k\}_{k=1}^n$

Etant donnée que l'erreur relative sur les mesures expérimentales est supposée constante, on la note  $e$ . Il est, donc, naturel de choisir des poids  $\omega_k$  inversement proportionnels au carré de l'erreur absolue sur  $Z_k$ , et ceci lorsque l'erreur absolue sur  $Z_k$  est supérieure à un certain seuil  $\Delta Z$  directement proportionnel à la valeur moyenne des  $Z_k$  :

$$(2.23) \quad \Delta Z = \delta \cdot Z_{\text{moy}}$$

avec

$$Z_{\text{moy}} = \frac{1}{n} \sum_{k=1}^n Z_k$$

En pratique, on utilisera un  $\delta$  de l'ordre de  $10^{-2}$ . Alors, les poids  $\omega_k$  seront calculés par la relation suivante :

$$(2.24) \quad \omega_k = \frac{\omega_0}{[\max(\Delta Z, e \cdot Z_k)]^2} \quad k=1, \dots, n$$

où  $\omega_0$  est une constante de normalisation telle que :



$$\sum_{k=1}^n \omega_k / n = 1$$

d'où

$$(2.25) \quad \omega_0 = \frac{n}{\sum_{k=1}^n \left( \frac{1}{\max(\Delta Z, e \cdot Z_k)} \right)^2}$$

## 2.7 Modélisation du rapport de deux propriétés conservatives

Dans ce paragraphe, nous allons étudier les propriétés physiques s'exprimant par le quotient des quantités conservatives ou additives (par exemple la densité et la teneur en soufre total). Alors, si on considère les propriétés conservatives  $F_1$  et  $F_2$  alors, il s'agit d'une modélisation de la propriété :

$$F = \frac{F_1}{F_2}$$

On dispose de deux ensembles de coupes  $(I_k^{(j)}, Z_k^{(j)})_{k=1, \dots, n_j}$  avec :

$$I_k^{(j)} = [T_{i_k}^{(j)}, T_{f_k}^{(j)}], \quad k=1, \dots, n_j; j=1, 2.$$

on pose :

$$T_{\min}^{(j)} = \min_{k=1, \dots, n_j} \{T_{i_k}^{(j)}\} \quad \text{et} \quad T_{\max}^{(j)} = \max_{k=1, \dots, n_j} \{T_{f_k}^{(j)}\}$$

Notre objectif est alors de déterminer deux fonctions  $f_1$  et  $f_2$ , densités des propriétés  $F_1$  et  $F_2$  respectivement, telles que :

$$(2.26) \quad f = \frac{f_1}{f_2}$$

soit la densité de la propriété non additive F. On considère ici :

$$T_{\min} = \max \{ T_{\min}^{(1)}, T_{\min}^{(2)} \}$$

et

$$T_{\max} = \min \{ T_{\max}^{(1)}, T_{\max}^{(2)} \}$$

Alors, (2.26) est définie dans l'intervalle  $[T_{\min}, T_{\max}]$ , et on a pour tout  $[T_i, T_f] \subseteq [T_{\min}, T_{\max}]$  :

$$(2.27) \quad F(T_i, T_f) = \frac{\int_{T_i}^{T_f} f_1(t) \, dm(T)}{\int_{T_i}^{T_f} f_2(t) \, dm(T)}$$

Tout d'abord, on traite les propriétés conservatives  $F_1$  et  $F_2$  par l'un des modèles I, II ou III précédemment décrits, ce qui nous permet d'obtenir les splines d'approximation :

$$\hat{f}_1 = \sum_{i=0}^m \hat{\alpha}_i^{(1)} \cdot B_i^{(1)} \quad \text{sur} \quad [T_{\min}^{(1)}, T_{\max}^{(1)}]$$

et

$$\hat{f}_2 = \sum_{i=0}^m \hat{\alpha}_i^{(2)} \cdot B_i^{(2)} \quad \text{sur} \quad [T_{\min}^{(2)}, T_{\max}^{(2)}]$$

pour les densités  $f_1$  et  $f_2$  respectivement. De même, on considère :

$$(2.28) \quad \hat{f} = \frac{\hat{f}_1}{\hat{f}_2}$$

La densité  $\hat{f}$  est généralement inacceptable (oscillations possibles, pas de monotonie). D'autre part, cette densité n'est pas valable, en général, pour évaluer localement la propriété F (par exemple les coupes étroites). C'est la raison pour laquelle on s'intéresse à la modification de la forme de  $\hat{f}$  définie par (2.28) ; pour cela, nous considérons un ensemble de points  $\{T_r\}_{r=1, \dots, \mu}$  tels que :

$$T_r = T_{\min} + (r-1) h \quad r=1, \dots, \mu$$

avec

$$h = \frac{T_{\max} - T_{\min}}{\mu - 1}$$

$\mu$  est un entier positif suffisamment grand. On pose :

$$\hat{f}_r = \hat{f}(T_r), \quad r=1, \dots, \mu$$

et on cherche une fonction  $\tilde{f} = \sum_{i=0}^m \tilde{\alpha}_i B_i$  solution du problème :

$$(P.2.4) \quad \left\{ \begin{array}{l} \min \left\{ \tau \int_{T_{\min}}^{T_{\max}} [s''(t)]^2 dt + \sum_{r=1}^{\mu} (s(T_r) - \hat{f}_r)^2 + Q(G^{-1} \cdot \alpha) \right\} \\ s \in S(4, m) \\ s = \sum_{i=0}^m \alpha_i B_i \\ G^{-1} \cdot \alpha \geq 0 \end{array} \right.$$

où la matrice G et l'application Q ont les mêmes significations que dans §2.5. Le paramètre  $\tau$  permet de doser le lissage de la spline  $\tilde{f}$  qui aura la forme souhaitée pour la densité f de (2.26). On pose :

$$\tilde{f}_r = \tilde{f}(T_r), \quad r=1, \dots, \mu$$

Ensuite, nous faisons le traitement d'une manière simultanée des deux propriétés conservatives  $F_1$  et  $F_2$  en tenant compte de la forme modifiée de  $\tilde{f}$ , obtenue par la résolution du (P.2.4), à l'aide de l'ensemble des points de contrôle :

$$\left\{ \tilde{f}_r \right\}_{r=1, \dots, \mu}$$

Nous avons à résoudre le problème suivant :

$$(P.2.5) \quad \left\{ \begin{array}{l} \min \left( \begin{array}{l} \tau_1 \int_{T_{\min}^{(1)}}^{T_{\max}^{(1)}} [s_1^{\sim}(t)]^2 dt + \tau_2 \int_{T_{\min}^{(2)}}^{T_{\max}^{(2)}} [s_2^{\sim}(t)]^2 dt + \sum_{k=1}^{n_1} \omega_k^{(1)} \left( \int_{I_k^{(1)}} s_1(t) dt - Z_k^{(1)} \right)^2 + \\ \sum_{k=1}^{n_2} \omega_k^{(2)} \left( \int_{I_k^{(2)}} s_2(t) dt - Z_k^{(2)} \right)^2 + Q(\alpha^{(1)}) + Q(\alpha^{(2)}) + \rho \sum_{r=1}^{\mu} (s_1(T_r) - \tilde{f}_r s_2(T_r))^2 \end{array} \right) \\ s_1, s_2 \in S(4, m) \\ s_1 = \sum_{i=0}^m \alpha_i^{(1)} B_i^{(1)} \\ s_2 = \sum_{i=0}^m \alpha_i^{(2)} B_i^{(2)} \\ \left. \begin{array}{l} \alpha_i^{(1)} \geq 0 \\ \alpha_i^{(2)} \geq 0 \end{array} \right\} i=0, \dots, m \end{array} \right.$$

avec

$$Q(\alpha^{(j)}) = \sum_{i=0}^m P(\alpha_i^{(j)}) \quad j=1, 2$$

et  $P$  une fonction réelle vérifiant (2.14). Comme dans (2.3) on peut écrire :

$$\bullet \int_{T_{\min}^{(j)}}^{T_{\max}^{(j)}} [s_j^{\sim}]^2 dt = \langle \alpha^{(j)}, \Omega_j \cdot \alpha^{(j)} \rangle_{\mathbb{R}^{m+1}} \quad j=1, 2$$

avec  $\Omega_1, \Omega_2$  des matrices symétriques semi-définie positives définies par (1.30).

On peut aussi exprimer les autres termes comme suit :

$$\bullet \sum_{k=1}^{n_j} \omega_k^{(j)} \left( \int_{I_k^{(j)}} s_j(t) dt - Z_k^{(j)} \right)^2 = \left\langle W_j(M_j^t \cdot \alpha^{(j)} - Z^{(j)}), W_j(M_j^t \cdot \alpha^{(j)} - Z^{(j)}) \right\rangle_{\mathbb{R}^{n_j}}$$

pour  $j=1, 2$ .

$$\bullet \sum_{r=1}^{\mu} \left( s_1(T_r) - \tilde{f}_r \cdot s_2(T_r) \right)^2 = \sum_{r=1}^{\mu} \left[ \sum_{i=0}^m \left( \alpha_i^{(1)} \cdot B_i^{(1)}(T_r) - \tilde{f}_r \cdot \alpha_i^{(2)} \cdot B_i^{(2)}(T_r) \right) \right]^2$$

$$= \sum_{r=1}^{\mu} \left( V_r^t \cdot \gamma \right)^2$$

$$= \left\langle \gamma, V \cdot \gamma \right\rangle_{\mathbb{R}^{2(m+1)}}$$

avec :

$$\gamma = \begin{pmatrix} \alpha^{(1)} \\ \alpha^{(2)} \end{pmatrix} \in \mathbb{R}^{2(m+1)}$$

$$V_r^t = \left( B_0^{(1)}(T_r), \dots, B_m^{(1)}(T_r), -\tilde{f}_r \cdot B_0^{(2)}(T_r), \dots, -\tilde{f}_r \cdot B_m^{(2)}(T_r) \right)$$

et

$$V = \sum_{r=1}^{\mu} V_r \cdot V_r^t$$

une matrice carrée symétrique semi-définie positive d'ordre  $2(m+1)$ .

### Remarque 2.6

La résolution du problème (P.2.5) avec des paramètres  $\tau_1$ ,  $\tau_2$  et  $\rho$  fournit deux fonctions splines  $\bar{s}_1$  et  $\bar{s}_2$  telles que :

$$\bar{s}_j = \sum_{i=0}^m \bar{\alpha}_i^{(j)} \cdot B_i^{(j)} \quad j=1, 2$$

soient des approximations des densités  $f_1$  et  $f_2$ . Pour un bon choix de ces paramètres, le rapport  $\frac{\bar{s}_1}{\bar{s}_2}$  représente une bonne approximation de la densité  $f$  de la propriété F.

## 2.8 Conclusion

Nous avons proposé dans ce chapitre trois modèles pour traiter les propriétés conservatives. Ces modèles sont basés sur l'approximation de données au sens des moindres carrés régularisés. Les données sont fournies sous formes d'histogrammes contigus ou chevauchés. Des conditions de forme (positivité et monotonie) peuvent être imposées dans les modèles II et III; c'est la raison pour laquelle nous avons un coût de calcul croissant suivant le modèle utilisé I, II ou III. Enfin, nous avons présenté au §2.8 un procédé de modélisation des propriétés physiques s'exprimant par le quotient des quantités conservatives. Ce procédé permet d'obtenir, selon le cas, une courbe densité lisse, positive et monotone dans l'intervalle de définition.



## CHAPITRE III

### **Algorithmes de résolution : Optimisation convexe avec contraintes linéaires**





### 3.0 Introduction

Nous avons vu au chapitre II que la plupart des problèmes d'optimisation rencontrés, lors de la modélisation de coupes pétrolières, se ramènent à un problème de minimisation de la forme :

$$(P.3.1) \quad \begin{cases} \text{minimiser} & \Phi(x) + Q(x) \\ x \in D = \{x \in \mathbb{R}^n : A.x \geq 0\} \end{cases}$$

$\Phi$  étant une fonction quadratique de la forme :

$$\Phi(x) = \frac{1}{2} \langle x, \psi.x \rangle + \langle b, x \rangle + c$$

(avec :  $\psi$  symétrique définie positive), et  $Q$  est une application strictement convexe définie sur  $\overset{\circ}{D} = \{x \in \mathbb{R}^n : A.x > 0\}$ , et telle que  $Q(x) = +\infty$  si  $x \in \partial D = D \setminus \overset{\circ}{D}$ .  $A$  est une matrice de type  $(p,n)$  où  $p$  est le nombre de contraintes linéaires dans (P.3.1). Nous allons présenter, dans ce chapitre, un algorithme dit de directions réalisables, pour la résolution numérique de (P.3.1). L'idée consiste à choisir un point de départ satisfaisant les contraintes et à chercher une direction de déplacement réalisable telle que:

- i) un petit déplacement dans cette direction ne fait pas sortir de l'ensemble des solutions  $\overset{\circ}{D}$ .
- ii) la fonctionnelle  $\Phi + Q$  décroît d'une façon stricte.

On se déplace alors, avec un certain pas, dans cette direction pour obtenir un nouveau point meilleur que le précédent.

### 3.1 Formulation du problème : définitions et notations

La fonctionnelle  $f = \Phi + Q$  est supposée, deux fois continûment différentiable. Son gradient et son hessien, en un point courant  $x^k$  sont notés  $\nabla f(x^k)$  et  $\nabla^2 f(x^k)$  respectivement. On considère les formes linéaires:

$$(3.1) \quad g_i : x \rightarrow \langle A_i \cdot, x \rangle_{\mathbb{R}^n} \quad (i=1, \dots, p)$$

Par sa nature particulière, le problème (P.3.1) admet une solution unique  $x^* \in \overset{\circ}{D}$  caractérisée par la condition nécessaire et suffisante d'optimalité globale :

$$(3.2) \quad \nabla f(x^*) = 0 \quad (\text{stationnarité})$$

Le hessien  $\nabla^2 f(x)$  est une matrice définie positive pour tout  $x \in \overset{\circ}{D}$ . Ce problème est équivalent à la résolution du système d'équations non linéaires:

$$(3.3) \quad \frac{\partial f}{\partial x_j}(x) = 0 \quad (j=1, \dots, n)$$

dans le domaine  $D$  définie par les contraintes linéaires :

$$(3.4) \quad g_i(x) \geq 0 \quad (i=1, \dots, p)$$

On peut chercher à résoudre directement ce système, ce qui conduit à la méthode de Newton. Cependant, la méthode de Newton n'est pas applicable directement, à cause des contraintes (3.4) ; mais, on peut agir sur le pas de déplacement pour que la suite des itérés  $x^k : k=1, \dots$  reste à l'intérieur du domaine  $D$ .

### 3.2 Description de l'algorithme

Une idée assez naturelle, pour adapter la méthode de Newton au problème (P.3.1), est de modifier le déplacement en faisant intervenir une suite de scalaires  $0 < \lambda^k \leq 1$  permettant, dans un premier temps, la localisation de l'optimum à l'intérieur du domaine  $D$  et de poursuivre, dans un second temps, les itérations de Newton classique dans le voisinage de la solution (les  $\lambda^k$  deviennent identiquement égaux à 1 à partir d'une certaine valeur de  $k$ ). En modifiant ainsi, le moins possible, la méthode d'origine, on peut espérer conserver son efficacité.

Alors, les différentes étapes de l'algorithme se présentent comme suit:

(a) *point de départ :*

On suppose que l'on connaît une solution réalisable  $x^0 \in \overset{\circ}{D}$  (un tel point peut toujours être obtenu par programmation linéaire).

(b) *direction de déplacement ( $d^k$ )*

Solution de l'équation :

$$\nabla f(x^k) + \nabla^2 f(x^k).d^k = 0$$

Soit

$$d^k = -[\nabla^2 f(x^k)]^{-1}.\nabla f(x^k)$$

$d^k$  est une direction de descente car :

$$\langle d^k, \nabla f(x^k) \rangle = -\langle [\nabla^2 f(x^k)]^{-1}.\nabla f(x^k), \nabla f(x^k) \rangle < 0$$

( $[\nabla^2 f(x^k)]^{-1}$  étant définie positive ).

**(c) pas de déplacement ( $\lambda^k$ )**

On pose :

$$I^k = \{i \in \{1, 2, \dots, P\} : g_i(d^k) < 0\}$$

et

$$\lambda_{\max}^k = \begin{cases} \min \left\{ -\frac{g_i(x^k)}{g_i(d^k)} \right\} & \text{si } I^k \neq \emptyset \\ i \in I^k \\ +\infty & \text{sinon} \end{cases}$$

où  $\lambda_{\max}^k$  détermine le déplacement maximal, dans la direction  $d^k$ , autorisé par les contraintes. Alors, le pas de déplacement est obtenu par :

$$\lambda^k = \min(1, \beta \cdot \lambda_{\max}^k)$$

avec  $\beta \in ]0, 1[$  fixe.

**(d) itération**

$$x^{k+1} = x^k + \lambda^k \cdot d^k \quad (k \geq 0)$$

**(e) test d'arrêt**

Comme la convergence n'est pas finie, on doit définir un test d'arrêt. Le critère retenu pour cet algorithme est le suivant :

$$\|\nabla f\|^2 = \sum_{j=1}^n \left( \frac{\partial f}{\partial x_j} \right)^2 < \varepsilon_1 \quad \text{et} \quad \|x^{k+1} - x^k\|^2 < \varepsilon_2$$

avec  $\varepsilon_1, \varepsilon_2 > 0$  donnés.

### Remarque 3.1

A l'étape (c) et lorsque  $\lambda_{\max}^k$  est fini la fonction à une variable  $\varphi(\lambda) = f(x^k + \lambda.d^k)$  prend l'allure suivante :

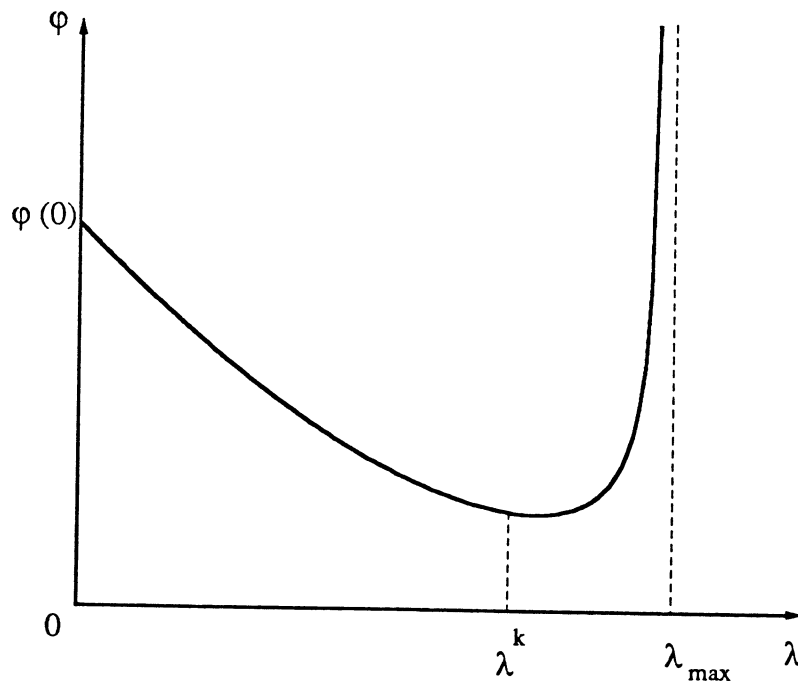


Figure 3.1 :  $\varphi(\lambda) = f(x^k + \lambda.d^k)$

le choix optimal pour  $\lambda^k$  consiste à prendre

$$(3.5) \quad \varphi(\lambda^k) = \min_{\lambda \in [0, \lambda_{\max}]} \varphi(\lambda)$$

Néanmoins, ceci coûte cher ; mais, en utilisant l'une des méthodes de recherche "économiques" proposées dans [Armijo 66], [Wolfe 69], [Goldstein/Price 67] et [Powell 76] on peut trouver une bonne approximation du minimum de  $\varphi$  dans l'intervalle  $[0, \lambda_{\max}]$ . En pratique, on utilise un  $\beta = 0.75$ .

### 3.3 Exemple simple (n=1)

On considère les fonctions réelles à une variable :

$$\Phi(x) = x^2 + 3x + 3$$

$$Q(x) = \frac{\eta}{x} \quad (\eta > 0)$$

et on pose le problème de minimisation suivant :

$$(P.3.2) \quad \min_{x \geq 0} \left\{ x^2 + 3x + 3 + \frac{\eta}{x} \right\}$$

on a ici :

$$(3.6) \quad \begin{cases} f(x) = x^2 + 3x + 3 + \frac{\eta}{x} \\ f'(x) = 2x + 3 - \frac{\eta}{x^2} \\ f''(x) = 2 \left( 1 + \frac{\eta}{x^3} \right) \end{cases}$$

d'où la formule itérative :

$$(3.7) \quad x^{k+1} = x^k - \lambda^k \frac{2x^k + 3 - \frac{\eta}{(x^k)^2}}{2 \left( 1 + \frac{\eta}{x^3} \right)} \quad (k \geq 0)$$

pour  $x^0 = 1$ ,  $\eta = 10^{-3}$  et  $\varepsilon_1 = \varepsilon_2 = 10^{-10}$  on a la séquence d'itérations suivante :

$k$	$x^k$	$f(x^k)$	$\ \nabla f(x^k)\ ^2$	$\lambda^k$	$\ x^{k+1} - x^k\ ^2$
0	1	7.001	24.99	0.3	0.563
1	0.25	3.8165	12.1	0.115	3.52 E-2
2	6.25 E-2	3.2074	8.23	0.167	2.20 E-3
3	1.5625 E-2	3.1111	1.13	1.0	4.09 E-6
4	1.764813 E-2	3.1099	3.08 E-2	1.0	2.299 E-7
5	1.812762 E-2	3.10987591	4.707 E-5	1.0	4.13 E-10
6	1.814793 E-2	3.10987584	1.31 E-10	1.0	1.16 E-15
7	1.814797 E-2	3.10987584	1.32 E-21		

Table 3.1 : séquence d'itérations de l'algorithme.





## CHAPITRE IV

### **Application : Calcul des propriétés conservatives**



## 4.0 Introduction

L'objet de ce travail est, d'une part de représenter un ensemble fini de mesures expérimentales, d'une propriété physique additive, par une fonction lisse continûment dérivable, et d'autre part de permettre une exploitation synthétique par une visualisation graphique des courbes (T.B.P.), surfaces et cartes iso-niveau inhérentes à cette fonction. Nous avons présenté au chapitre II trois modèles pour traiter les propriétés conservatives, ainsi qu'un procédé de traitement des propriétés s'exprimant par le quotient des quantités conservatives, tout en assurant une éventuelle caractéristique de forme (positivité et monotonie) ; ce travail a permis la mise au point du logiciel C-ELF qui permet de traiter, d'une manière interactive, une ou plusieurs analyses conservatives. Ce logiciel est mis au point dans le laboratoire TIM3 de l'institut IMAG (Informatique et Mathématiques Appliquées de Grenoble) et implanté au CRES sur un micro VAX (une version PC développée sous TURBO PASCAL est actuellement disponible).

### 4.1 Analyse du pétrole brut

La première opération, dans la transformation du pétrole brut, est celle du raffinage, fondé sur la distillation fractionnée qui permet de séparer un pétrole brut ou une partie de celui-ci en «coupes» d'hydrocarbures de propriétés suffisamment voisines pour qu'on puisse leur donner en bloc la même utilisation. On distingue, dans un brut, trois types de produits essentiels :

- (i) produits volatils ou gazeux : gaz et essence légère.
- (ii) produits légers : essence lourde, kérosène, gazoil.
- (iii) produits lourds : résidus.

La distillation industrielle est réalisée dans un four à tubes «TOPPING» que le pétrole brut traverse en circulation forcée et à grande vitesse. Au sortir du four dans laquelle la totalité du liquide à distiller a été portée à plus de 300 °C, celui-ci est brusquement détendu dans une capacité de quelques mètres cubes située vers la partie inférieure d'une vaste colonne de fractionnement. Les parties du pétrole brut vaporisable à la température réalisée s'échappent vers le haut de la colonne en se fractionnant par échange de température avec le liquide contenu dans 20 ou 30 plateaux à barbotage qu'elles traversent successivement. La colonne est alimentée simultanément en tête par un reflux constitué généralement par une partie de la condensation des vapeurs sortant en haut de la colonne. Ce reflux s'enrichissant en fractions lourdes au fur et à mesure qu'il descend d'un plateau à l'autre rejoint les parties du pétrole brut qui n'ont pas été vaporisées à leur arrivée dans la colonne et le tout repris comme résidu à la partie inférieure de la colonne. On extrait du sommet de la colonne, par prélèvement latéral continu dans les plateaux, des coupes extrêmement serrées et correspondant aux sélections désirées.

Un autre procédé de transformation du pétrole brut s'appelle «Craquage». Ce procédé permet de faire éclater ou «craquer» les grosses molécules des hydrocarbures lourds pour en faire des molécules plus légères.

Le pétrole brut est analysé, au préalable, au laboratoire afin de permettre au raffineur de le caractériser, d'évaluer sa composition des produits et mesurer sa densité, sa viscosité et sa teneur de chaque espèce (soufre, mercaptan, n-parafines).

## 4.2 Logiciel C-ELF

La structure du C-ELF est principalement constituée de trois unités étroitement liées par des échanges d'informations (Figure 4.1) :

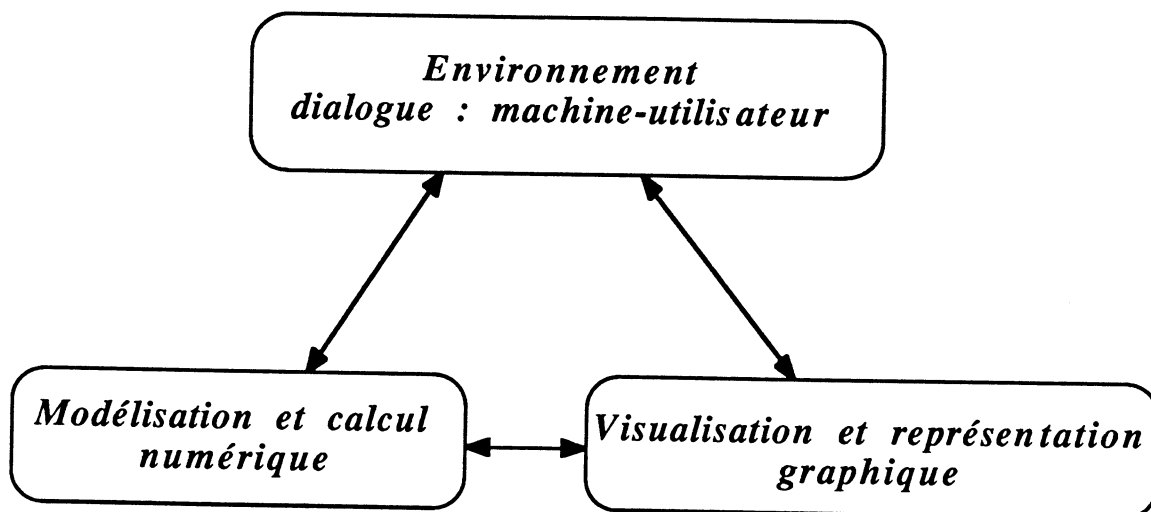


Figure 4.1 : Vue globale du C-ELF

- **Environnement**

L'environnement se présente sous forme d'un menu principal et d'un ensemble de sous-menus. Chacun des menus propose un certain nombre d'options, numérotées (1, 2, 3, ...etc.), permettant d'effectuer le choix, et de parcourir l'ensemble de structure du logiciel. L'utilisateur est guidé par une séquence de dialogue afin d'exécuter des tâches précises selon les options choisies. Les options du menu principal sont les suivantes :

- (1) Une seule analyse
- (2) Rapport de deux analyses
- (3) Liste d'analyses traitées
- (4) Suppression d'une analyse

- **Modélisation et calcul numérique**

Dans cette partie du logiciel se trouvent concrétisés les différents modèles de traitement des propriétés conservatives exposés au chapitre II, ainsi que le traitement du rapport de deux propriétés conservatives.

La construction numérique des fonctions splines est basée sur la résolution d'un problème d'optimisation du type (P.3.1). Un paramètre de lissage, dont la valeur est choisie par l'utilisateur, permet de régler le degré de lissage et de reconstituer correctement la fonction sous-jacente aux données. Une fois que le traitement numérique est terminé, C-ELF permet d'exploiter numériquement la solution obtenue, soit par une évaluation interactive de coupes ( $T_{\text{initial}}$  ,  $T_{\text{final}}$ ) arbitrairement choisies par l'utilisateur, soit par une création d'un fichier d'information contenant l'ensemble de paramètres utilisés, ainsi que les données du problème et les différentes coupes recalculées à partir de la fonction trouvée, et d'un fichier numérique contenant l'ensemble des coefficients de la spline solution ainsi que les températures limites du problème traité, et ceci pour une utilisation extérieure par d'autres applications.

- **Visualisation et représentation graphique**

Les mesures d'une propriété physique doivent être exploitées sous forme de courbes continues permettant une utilisation plus rationnelle. C-ELF propose trois types de représentation graphique :

- Une représentation classique par une courbe continue (température , propriété).
- Une représentation à deux dimensions par une surface (point initial , point final , propriété).

- Une représentation projective de la surface précédente par des courbes de niveaux (point initial , point final). Chaque courbe correspond à un iso-niveau.

La visualisation de la courbe T.B.P différentielle de la propriété, en cours de traitement, permet à l'utilisateur de bien choisir, d'une façon interactive, d'une part le modèle numérique approprié et d'autre part la bonne valeur de paramètre du lissage généralement situé dans l'intervalle  $]0,10^5]$  et qui est fortement liée à l'estimation de l'erreur relative en les mesures expérimentales de la propriété conservative considérée.

Dans le cadre de l'exploitation graphique de la fonction obtenue par modélisation numérique, C-ELF propose dans un menu graphique un éventail de tracés que l'on peut citer ici :

- Courbe : T.B.P. différentielle
- Courbe : T.B.P.
- Courbes : coupes à point initial fixe
- Courbes : coupes à largeur constante
- Surface : quantité en fonction de  $(T_i, T_f)$
- Courbes : iso-quantité en projection

Ces tracés sont aussi bien accessible dans le cas d'une seule analyse que dans le cas du rapport de deux analyses. Dans le cas du rapport de deux analyses additives, nous citons deux exemples courants dans l'analyse d'un brut :

- (i) La densité du brut correspond au rapport poids sur volume du distillat.
- (ii) La teneur en soufre total correspondant au rapport de deux quantités relatives à une coupe contenue dans une unité du brut : la première est la quantité du soufre contenue dans cette coupe ; la deuxième est sa masse.



### 4.3 Analyse du brut A334

Les résultats d'analyse d'un brut et ses différentes caractéristiques se trouve regroupés dans un cahier (cahier des bruts) sous forme de tableaux. Ces tableaux sont formés de l'ensemble de coupes réalisées (température initiale , température finale) et les quantités d'analyse correspondantes, ainsi que certaines caractéristiques de ces coupes. Nous présentons, dans ce qui suit, un traitement complet de certaines propriétés concernant le brut A334.

- **Distillat**

Nous rappelons que la distillation initiale permet de fractionner le brut en plusieurs coupes, et de mesurer les fractions (%poids et %volume) correspondantes.

(a) T.B.P. différentielle :

L'additivité nous permet d'obtenir (par modélisation) à partir des coupes initiales, toutes les coupes infiniment étroites (coupes différentielles), ce que nous appelons la T.B.P. différentielle (Figure 4.2).

(b) T.B.P. :

La sommation de toutes les fractions étroites (T.B.P. différentielle) d'une coupe ( $T_i$  ,  $T_f$ ) donne la valeur de la fraction de cette coupe. La sommation de toutes les fractions différentielles à partir de début de l'analyse jusqu'à un point T donne la valeur de T.B.P. du distillat en ce point, ce qui constitue la courbe T.B.P. (Figure 4.3). On remarque que la fraction du distillat entre  $T_i$  et  $T_f$  est représentée :

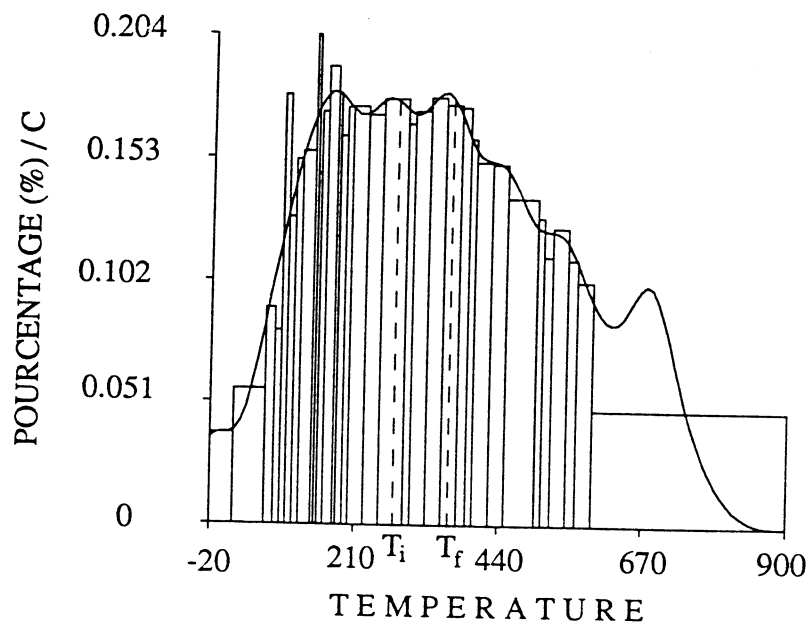


Figure 4.2 : Brut A334 - distillat  
T.B.P. différentielle

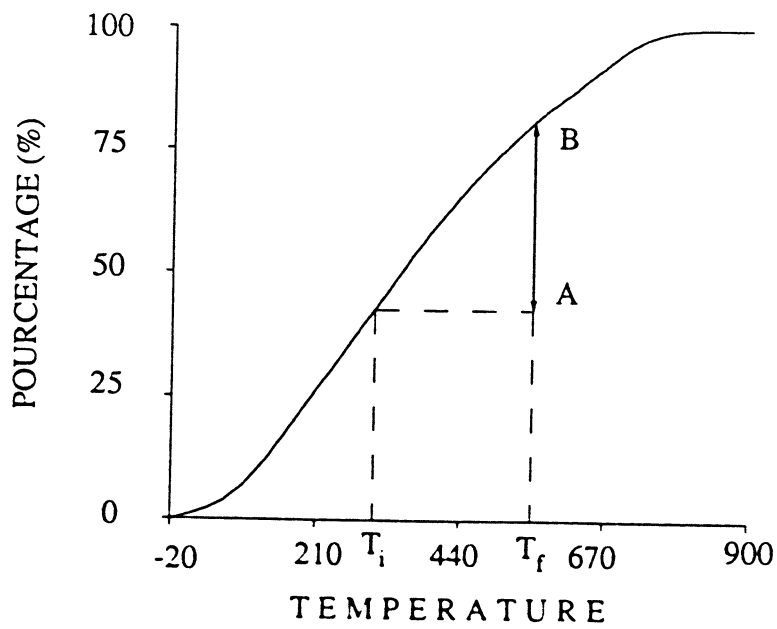


Figure 4.3: Brut A334 - distillat  
T.B.P. totale

- sur la figure (4.2) par la surface délimitée par les deux droites

verticales, passant par  $T_i$  et  $T_f$ , et la courbe.

- sur la figure (4.3) par le segment [A,B] (dénivellation entre A et B).

(c) Coupes à point initial fixe :

C'est un ensemble de courbes T.B.P. à point initial fixe. Il permet la localisation de toutes les coupes dont la quantité ou le point final est défini à priori. Dans la figure (4.4), et sur la droite D tous les points d'intersections correspondent aux coupes ( $T_i$ ,  $T_f$ ) avec  $T_i$  le point initial de la courbe et  $T_f$  l'abscisse de D. (exemple ( $440^\circ$ ,  $540^\circ$ , 13.26%)).

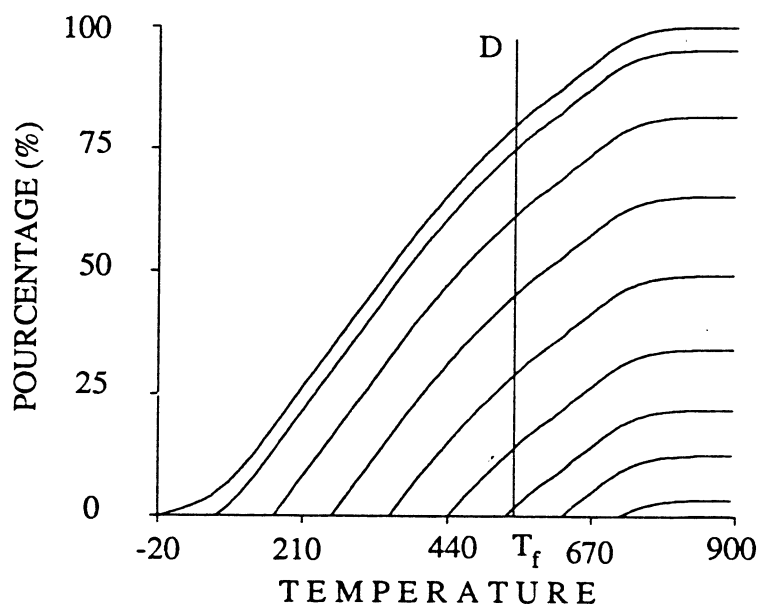


Figure 4.4: Brut A334 - distillat  
coupes à point initial fixe

(d) Coupes à largeur constante :

La figure (4.5) visualise un ensemble de courbes représentant des coupes à largeur constante ( $\Delta T = T_f - T_i$ ). Sur la droite D on peut aussi lire

les différentes coupes ( $T_i$ ,  $T_f$ ) dont la fraction du distillat est 60% avec  $T_i$  abscisse du point d'intersection et  $T_f = T_i + \Delta T$  (exemple ( $130^\circ, 480^\circ, 59\%$ ) avec  $\Delta T = 350^\circ$ )

(e) Surface de fractions ( $T_i$ ,  $T_f$ ) :

A partir de la T.B.P. nous pouvons avoir une représentation bidimensionnelle de la T.B.P. en fonction des points initiaux et finaux. La figure (4.6) visualise cette surface.

(f) Courbe iso-niveau en projection :

La figure (4.7) visualise la carte d'état major de la T.B.P. surface. Chaque courbe correspond à la projection d'un niveau (obtenu en intersectant la T.B.P. surface par le plan horizontal de ce niveau) sur le plan des températures.

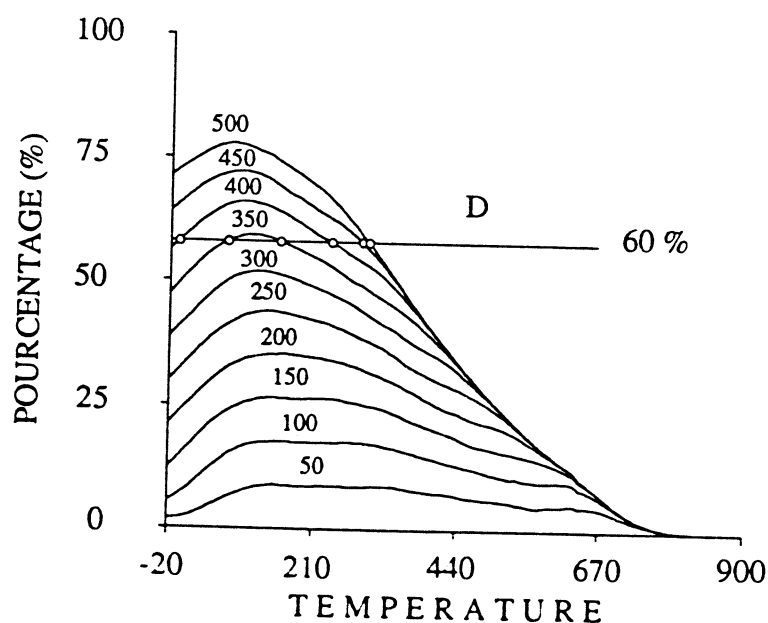


Figure 4.5: Brut A334 - distillat coupes à largeur constante

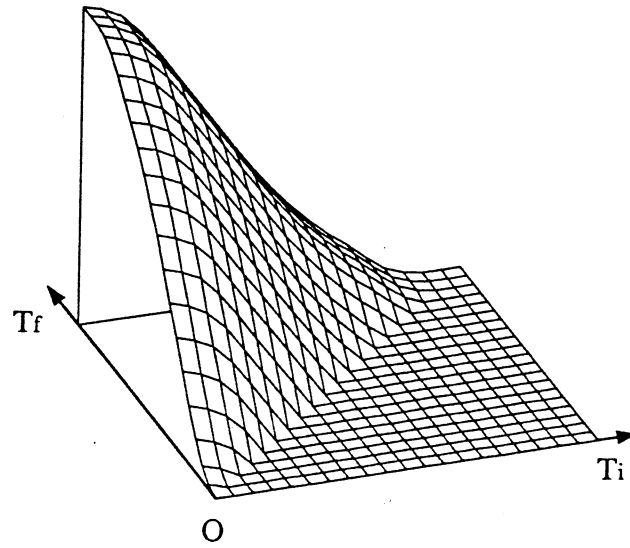


Figure 4.6 : Brut A334 - distillat  
surface T.B.P en fonction de ( $T_i$ ,  $T_f$ )

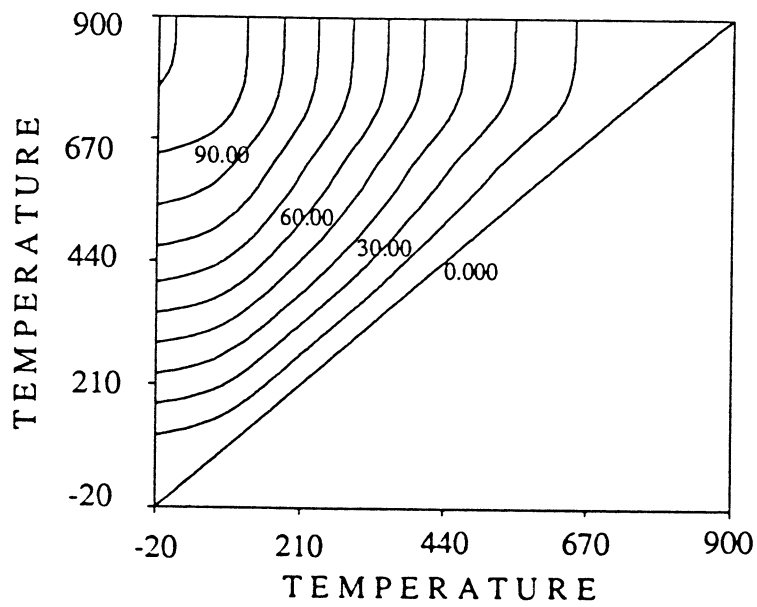


Figure 4.7 : Brut A334 - distillat  
courbes iso-niveau en projection

- **Densité**

La densité ou la masse volumique d'un brut relative à une coupe ( $T_i$ ,  $T_f$ ) se traduit par le rapport de deux quantités :

$$d(T_i, T_f) = \frac{\Delta P(T_i, T_f)}{\Delta V(T_i, T_f)}$$

avec :

$$\begin{aligned}\Delta P(T_i, T_f) &= \%P(T_i, T_f) \cdot P \\ \Delta V(T_i, T_f) &= \%V(T_i, T_f) \cdot V\end{aligned}$$

$P$ ,  $V$  : poids et volume total du brut analysé.

$\%P$ ,  $\%V$  : pourcentage en poids et en volume du distillat dans la coupe ( $T_i$ ,  $T_f$ ).

On remarque que la densité est un rapport, à un coefficient près, de deux fonctions T.B.P ( $\%P$  et  $\%V$ ).

C-ELF propose un procédé de traitement, présenté au §2.8, permettant d'avoir une courbe densité lisse, positive et monotone. Cette fonction (densité du brut) peut être aussi exploitée numériquement et graphiquement d'une façon interactive, comme dans le cas du distillat.

- **Soufre**

(i) Teneur en Soufre :

On rappelle que la distillation secondaire de chaque fraction fournit le pourcentage de différent produits par rapport à ces coupes, en particulier le soufre. Les quantités du soufre sont désignées dans le cahier des bruts par la teneur en soufre total ( $\%TST$ ). On précise ici que ces quantités (pour tout brut) ne satisfont pas la propriété d'additivité (il est facile de vérifier que la

somme totale des teneurs en soufre ne donne pas la teneur en soufre total dans la coupe somme). Il est nécessaire alors de se ramener au brut en les pondérant par les coefficients de rendement (%distillat). Ces nouvelles valeurs représentent la quantité de soufre par rapport au brut et sont additives. La figure 4.8 représente la T.B.P différentielle du soufre, tandis que la figure 4.9 montre sa courbe T.B.P totale.

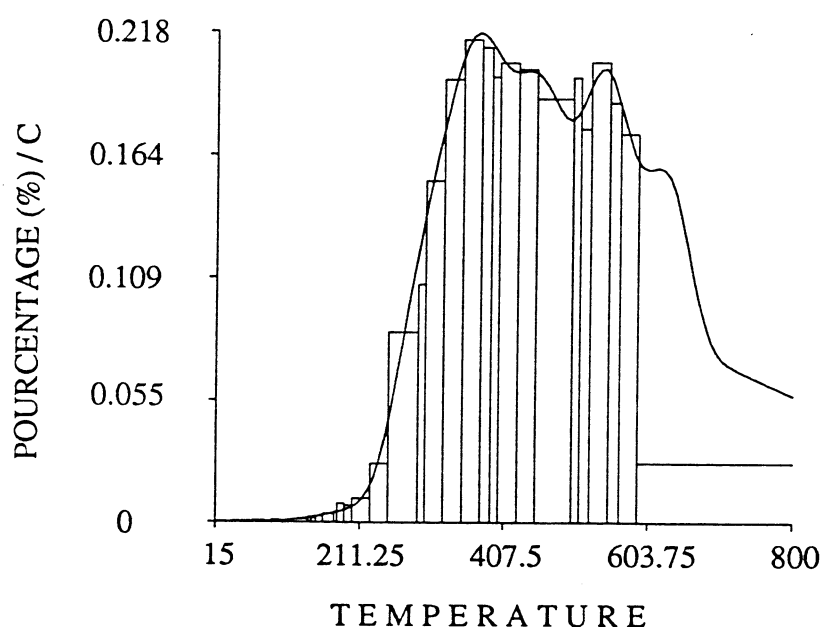


Figure 4.8 : Brut A334 - soufre  
T.B.P. différentielle

C-ELF propose les mêmes visualisations graphiques décrites au §4.2:

- quantité de soufre pour les coupes à point initial fixe (figure 4.10).
- quantité de soufre pour les coupes à largeur constante (figure 4.11).
- quantité de soufre en fonction de  $T_i$  et  $T_f$  (figure 4.12).
- carte d'état major des quantités de soufre (figure 4.13).

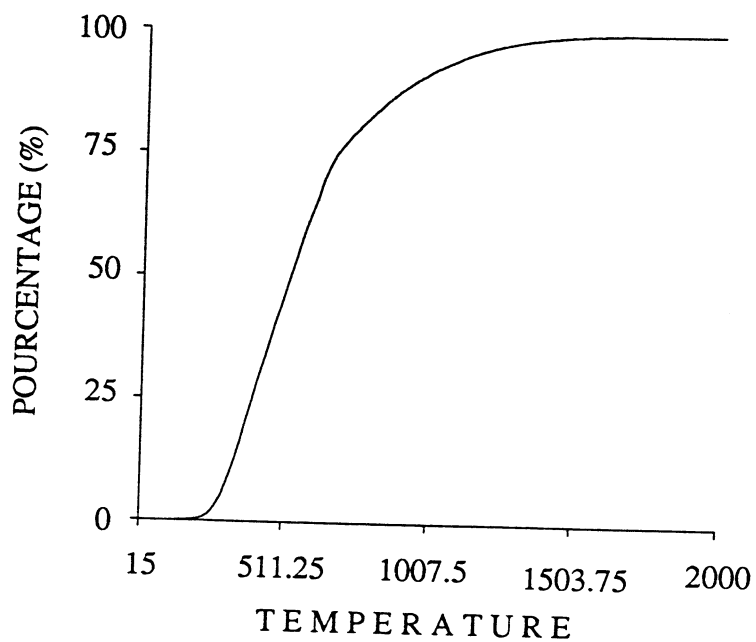


Figure 4.9 : Brut A334 - soufre  
T.B.P. totale

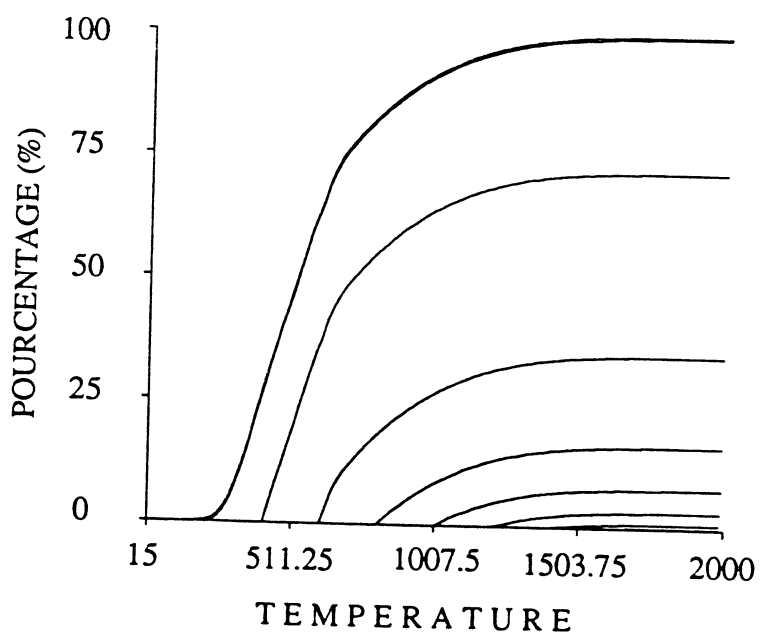


Figure 4.10: Brut A334 - soufre  
coupes à point initial fixe



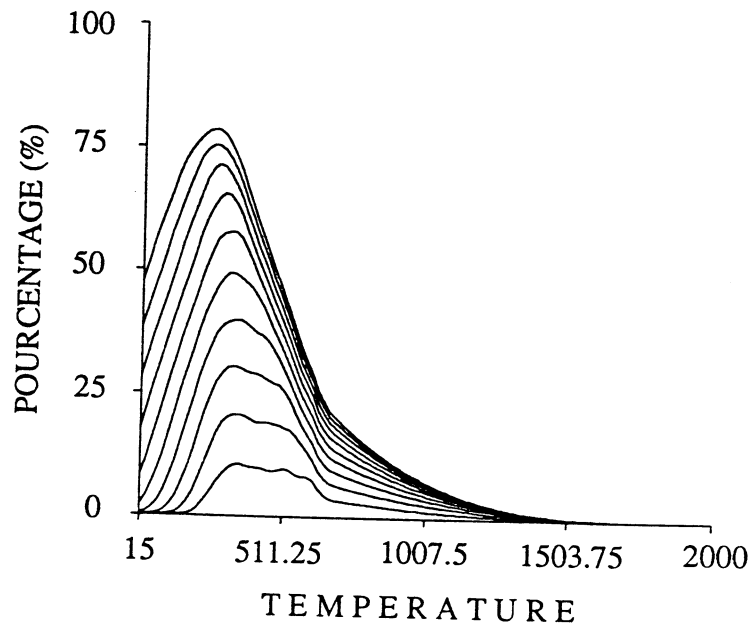


Figure 4.11 : Brut A334 - soufre  
coupes à largeur constante

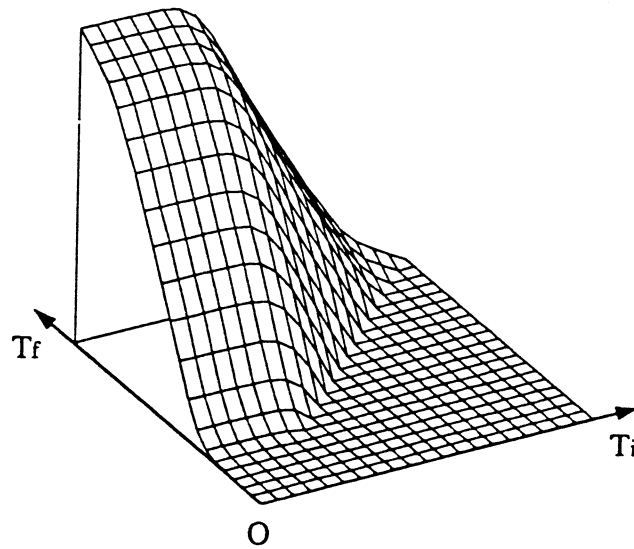


Figure 4.12 : Brut A334 - soufre  
surface T.B.P en fonction de  $(T_i, T_f)$

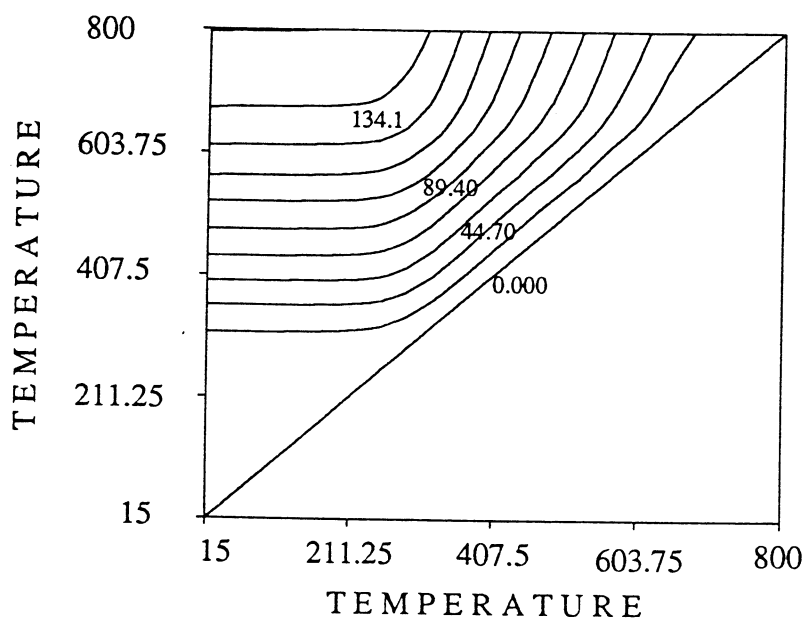


Figure 4.13 : Brut A334 - soufre  
courbes iso-niveau en projection

(ii) Teneur en Soufre Total :

En disposant de la T.B.P soufre et la T.B.P distillat, nous pouvons nous ramener au calcul de la teneur en soufre total (T.S.T) par le rapport des deux T.B.P. Cette opération est analogue à la détermination de la densité ci-dessus. Dans la figure 4.14 nous avons tracé la courbe T.B.P différentielle de la T.S.T, et dans la figure 4.15 sa T.B.P totale. Les figures 4.16, 4.17, 4.18 et 4.19 montrent les différents tracés de coupes à point initial fixe, de coupes à largeur constante, surface et carte iso-niveau de la T.S.T.

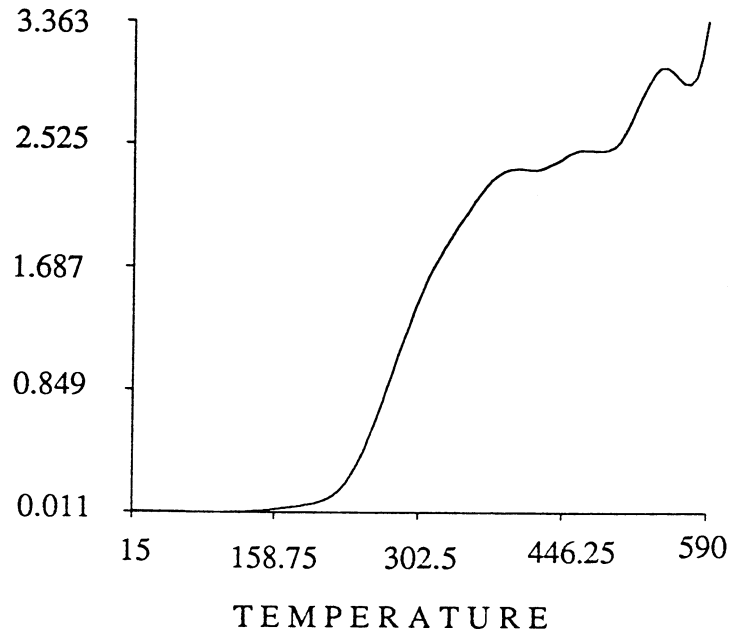


Figure 4.14 : Brut A334 - T.S.T  
T.B.P. différentielle

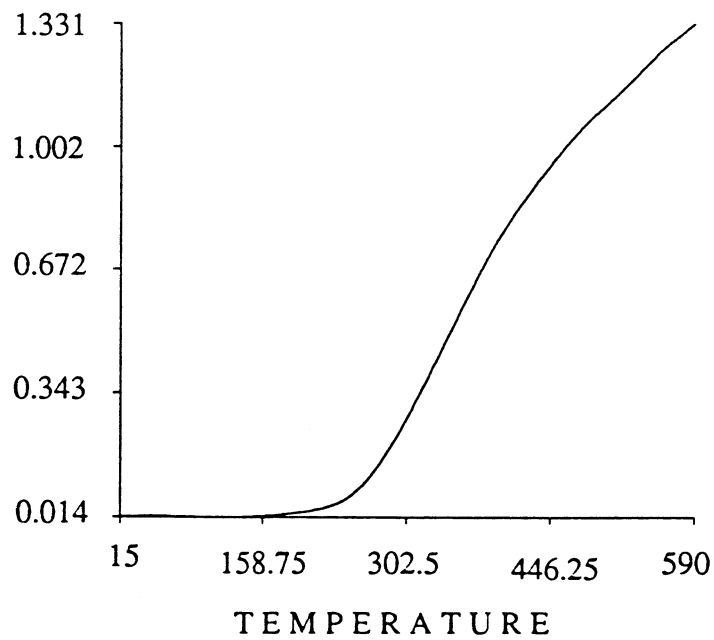


Figure 4.15 : Brut A334 - T.S.T  
T.B.P. totale

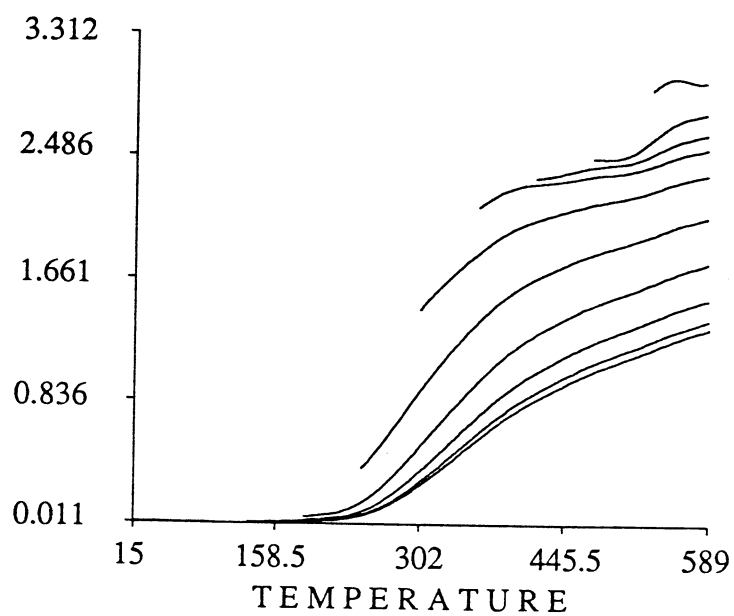


Figure 4.16 : Brut A334 - T.S.T  
coupes à point initial fixe

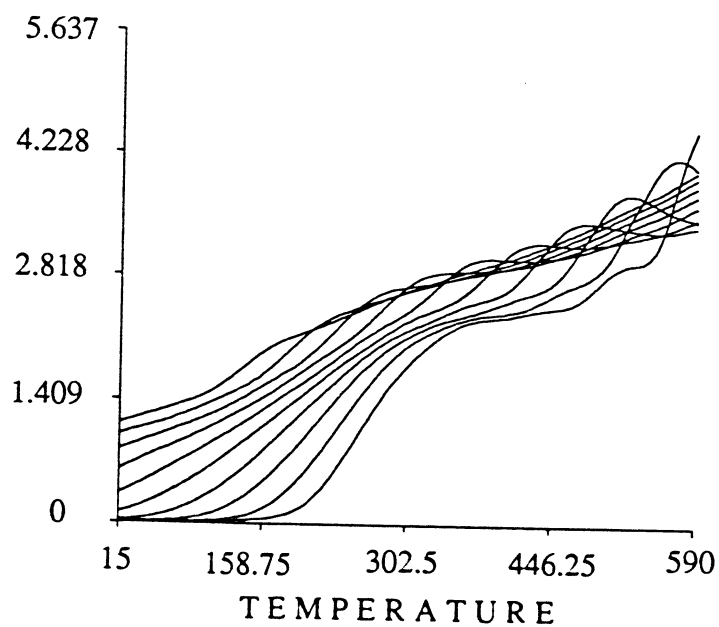


Figure 4.17 : Brut A334 - T.S.T  
coupes à largeur constante

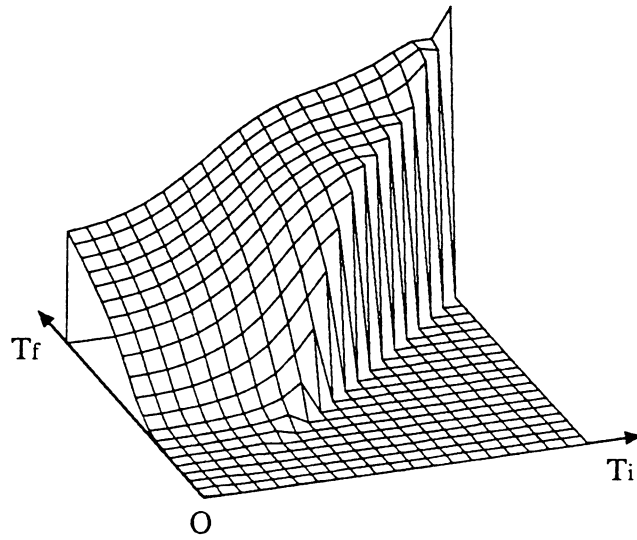


Figure 4.18 : Brut A334 - T.S.T  
surface T.B.P en fonction de  $(T_i, T_f)$

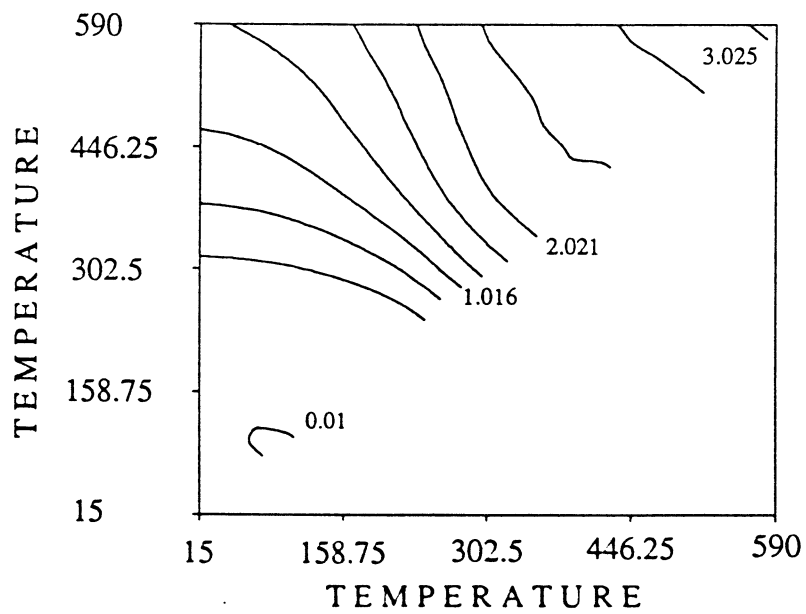


Figure 4.19 : Brut A334 - T.S.T  
courbes iso-niveau en projection

#### 4.4 Conclusion

L'analyse du pétrole brut fournit des résultats discrets et imprécis représentés par un ensemble de coupes pour lesquelles on a mesuré, au laboratoire, plusieurs propriétés qui caractérisent le brut. Nous disposons d'un logiciel qui permet de rendre continues et dérivables ces analyses, en utilisant les B-splines comme éléments de bases d'approximation. Ce logiciel permet aussi de reconstituer toutes les coupes possibles et calculer leurs quantités et propose plusieurs représentations visuelles à une et deux dimensions par exemple la T.B.P en fonction du point initial et sa carte d'état-major. Cet ensemble de représentations permet à l'utilisateur une évaluation commode des propriétés conservatives d'une coupe ou d'un brut.



*Deuxième Partie : Analyse de  
Mélanges Non Additifs*

CHAPITRE V

**Propriétés non linéaires de  
coupes complexes**





## 5.0 Introduction

Les propriétés conservatives (linéaires) représentent une part importante dans la gestion des mélanges complexes. Leur modélisation a été suffisamment abordée dans la première partie. D'autres propriétés physiques, concernant le pétrole et les coupes pétrolières, ne sont pas additives au mélange ; c'est le cas de la viscosité, du point de trouble au refroidissement et de la propriété antidétonante (mesure d'octane). En général, on admet qu'une propriété non additive  $F$  dépend d'un certain nombre de grandeurs additives et dans le cas où la propriété  $F$  ne dépend que d'une seule grandeur additive (c'est une hypothèse) ceci revient à supposer l'existence d'une fonction inversible  $I$  telle que  $I(F)$  soit additive ; cette fonction est appelée loi d'indice ou indice de la propriété. Il existe alors, pour un brut donné, une fonction  $f(T)$  densité de la propriété  $F$ . De même, le rendement poids sur brut est donné par la mesure  $m(T)$ , de densité  $dm/dT$ .

On veut calculer la fonction de deux variables  $F(T_i, T_f)$  telle que :

$$(5.1) \quad F(T_i, T_f) = I^{-1} \left( \frac{\int_{T_i}^{T_f} I(f(T)) \cdot dm(T)}{\int_{T_i}^{T_f} dm(T)} \right)$$

$F$  donnera, respectivement, la viscosité, le trouble et l'octane de la coupe

distillant entre les températures  $T_i$  et  $T_f$ . La fonction  $f(T)$  n'est pas accessible directement, mais doit être calculée à partir de mesures de  $F$  sur des coupes du même brut et définies par des intervalles de températures  $[T_i, T_f]$  généralement chevauchés. Les mesures, ainsi que la loi d'indice  $I$  sont sujettes à des erreurs.

Pour un brut donné, on dispose d'un ensemble de mesures  $\{F_{ij}\}$  et  $\{m_{ij}\}$  de la propriété non additive  $F$  correspondant à un ensemble de coupes:

$$(5.2) \quad \{(T_i, T_j) : 1 \leq i \leq n_1 \text{ et } n_1+1 \leq j \leq n_1+n_2\}$$

avec  $T_1 < T_2 < \dots < T_{n_1} < T_{n_1+1} < T_{n_1+2} < \dots < T_{n_1+n_2}$

ce qui constitue un ensemble de  $(n_1 \times n_2)$  mesures. On pose :

$$(5.3) \quad Z_{ij} = m_{ij} \cdot I(F_{ij})$$

C'est la quantité d'indice associée à la coupe  $(T_i, T_j)$ .

La recherche de l'indice de mélange peut se faire selon le critère que l'on choisit pour sa détermination effective. Dans ce qui suit, nous présentons trois approches: la première est basée sur les équations d'indices (additivité), la deuxième consiste à reconstituer les quantités d'indice, tandis que la troisième cherche à reconstituer les mesures de la propriété  $F$ .

### 5.1 Première approche : équations d'indices

Les  $\{Z_{ij}\}$  sont additives et vérifient, donc, les relations du mélange additif concernant les coupes (5.2). Pour tout  $(i_1, i_2) \in \{1, \dots, n_1\}^2$  et  $(j_1, j_2) \in \{n_1+1, \dots, n_1+n_2\}^2$  tels que :

$$i_1 < i_2 \quad \text{et} \quad j_1 < j_2$$

il existe une relation linéaire (équation d'indices) du type :

$$(5.4) \quad Z_{i_1, j_2} - Z_{i_1, j_1} = Z_{i_2, j_2} - Z_{i_2, j_1}$$

Cette relation exprime le fait qu'on puisse obtenir, par différence entre  $Z_{ij}$ , la quantité d'indices associée à la coupe étroite  $(T_{i_1}, T_{i_2})$  ou  $(T_{j_1}, T_{j_2})$ , (voir figure 5.1).

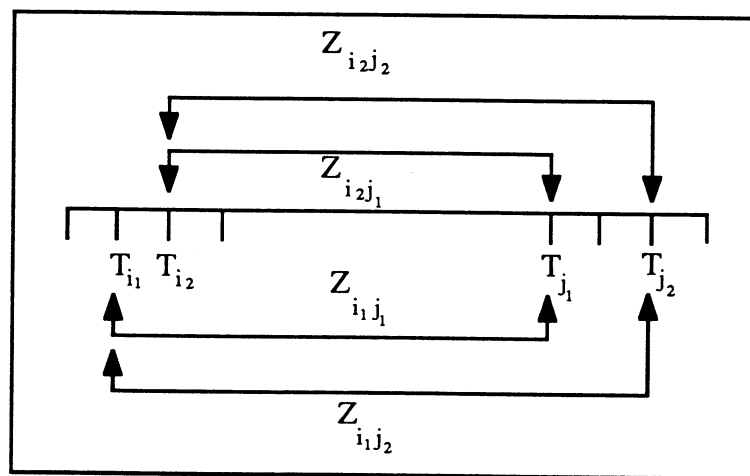


Figure 5.1 : équation d'indices et coupes étroites

L'ensemble de tous les couples  $(i_1, i_2) \in \{1, \dots, n_1\}^2$  tels que  $i_1 < i_2$  est égal à l'ensemble suivant :

$$(5.5) \quad \left\{ \begin{array}{l} (1,2), (1,3), \dots, (1, n_1) \\ (2,3), \dots, (2, n_1) \\ \vdots \\ (n_1-1, n_1) \end{array} \right\}$$

De même, l'ensemble des couples  $(j_1, j_2) \in \{n_1+1, \dots, n_1+n_2\}^2$  tels que  $j_1 < j_2$  est donné par :

$$(5.6) \quad \left\{ \begin{array}{l} (n_1+1, n_1+2), (n_1+1, n_1+3), \dots, (n_1+1, n_1+n_2) \\ (n_1+2, 3), \dots, (n_1+2, n_1+n_2) \\ \vdots \\ (n_1+n_2-1, n_1+n_2) \end{array} \right\}$$

Par conséquent, il existe au total  $n_1(n_1-1)n_2(n_2-1)/4$  équations du type (5.4). Ces équations ne sont pas indépendantes et sont générées par  $(n_1-1)(n_2-1)$  équations de base ; par conséquent  $(n_1-1)(n_2-1) \left[ \frac{n_1 n_2}{4} - 1 \right]$  équations sont redondantes. Les indices  $(i_1, i_2) \times (j_1, j_2)$  correspondant aux équations de base sont :

$$\{ (i_1, i_1+1) \times (j_1, j_1+1) : 1 \leq i_1 \leq n_1-1 \quad \text{et} \quad n_1+1 \leq j_1 \leq n_1+n_2-1 \}$$

- **Modèle d'approximation de la fonction d'indice**

On note dans toute la suite :

$$F_{\min} = \min\{F_{ij}\} \quad \text{et} \quad F_{\max} = \max\{F_{ij}\}$$

et on considère l'espace  $S(4, d-1)$ , des splines cubiques uniformes définies sur  $[F_{\min}, F_{\max}]$ , dans lequel on va chercher une approximation de la fonction indice  $I$  vérifiant au mieux les relations (5.4) entre les grandeurs additives  $\{Z_{ij}\}$ . La dimension de cet espace est égale à  $d$ .

On pose :

$$(5.7) \quad \delta_{(i_1, i_2) \times (j_1, j_2)} = Z_{i_1, j_2} - Z_{i_1, j_1} - Z_{i_2, j_2} + Z_{i_2, j_1}$$

et

$$(5.8) \quad \Delta_{\text{brut}}^2 = \sum_{\substack{(i_1, i_2) \\ i_1 < i_2}} \sum_{\substack{(j_1, j_2) \\ j_1 < j_2}} \delta_{(i_1, i_2) \times (j_1, j_2)}^2$$

La fonction d'indice doit être inversible, et convexe ; on lui impose donc d'être monotone (croissante) et convexe. D'autre part, puisque les

relations (5.4) sont définies à un facteur multiplicatif et à une translation près, toute combinaison de la forme :

$$\mu I + \varepsilon \quad \mu, \varepsilon \in \mathbb{R}^* \times \mathbb{R}$$

est une fonction d'indice pour la propriété F. Une manière de faire un choix unique de I consiste à imposer les conditions de normalisation suivantes :

$$(5.9) \quad \begin{cases} I(F_{\min}) = 0 \\ I(F_{\max}) = 1 \end{cases}$$

On considère l'ensemble de valeurs nodales :

$$u_0 < u_1 < u_2 < u_3 < \dots < u_d < u_{d+1} < u_{d+2} < u_{d+3}$$

tels que :

$$u_3 = F_{\min} \quad \text{et} \quad u_d = F_{\max}$$

et

$$h = u_i - u_{i-1} = \frac{F_{\max} - F_{\min}}{d-3} \quad i=1, \dots, d+3$$

Alors, toute fonction  $s \in S(4, d-1)$  peut être mise sous la forme :

$$s = \sum_{i=0}^{d-1} \alpha_i B_i$$

On cherche I comme étant la solution du problème d'optimisation suivant:

$$(P.5.1) \quad \begin{cases} \text{minimiser} \quad \left\{ \Delta_{\text{brut}}^2 + k \left[ I^2(F_{\min}) + (I(F_{\max})-1)^2 \right] \right\} \\ I \in S(4, d-1) \\ I' \geq 0 \\ I'' \geq 0 \end{cases}$$

(k étant un réel positif assez grand).  $\Delta_{\text{brut}}^2$  s'écrit sous la forme :

$$(5.10) \quad \Delta_{\text{brut}}^2 = \langle \alpha, U\alpha \rangle_{\mathbb{R}^d}$$

$U$  est une matrice symétrique semi-définie positive.

On a en effet :

$$\delta_{(i_1, i_2) \times (j_1, j_2)} = \sum_{r=0}^{d-1} \alpha_r E_{(i_1, i_2) \times (j_1, j_2)}^r$$

avec

$$(5.11) \quad E_{(i_1, i_2) \times (j_1, j_2)}^r = \begin{cases} m_{i_1, j_2} \cdot B_r(F_{i_1, j_2}) - m_{i_1, j_1} \cdot B_r(F_{i_1, j_1}) \\ - m_{i_2, j_2} \cdot B_r(F_{i_2, j_2}) + m_{i_2, j_1} \cdot B_r(F_{i_2, j_1}) \end{cases}$$

On pose

$$E_{(i_1, i_2) \times (j_1, j_2)} = \left( E_{(i_1, i_2) \times (j_1, j_2)}^r \right)_{r=0}^{d-1} \in \mathbb{R}^d$$

d'où

$$\delta_{(i_1, i_2) \times (j_1, j_2)} = \langle E_{(i_1, i_2) \times (j_1, j_2)}, \alpha \rangle_{\mathbb{R}^d}$$

On a donc :

$$\Delta_{\text{brut}}^2 = \sum_{i_1 < i_2} \sum_{j_1 < j_2} \langle E_{(i_1, i_2) \times (j_1, j_2)}, \alpha \rangle_{\mathbb{R}^d}^2$$

ou encore :

$$\Delta_{\text{brut}}^2 = \sum_{i_1 < i_2} \alpha^t \cdot E_{(i_1, i_2)} \cdot E_{(i_1, i_2)}^t \cdot \alpha$$

avec

$$(5.12) \quad E_{(i_1, i_2)} = [E_{(i_1, i_2) \times (n_1+1, n_1+2)}, E_{(i_1, i_2) \times (n_1+1, n_1+3)}, \dots, E_{(i_1, i_2) \times (n_1+n_2-1, n_1+n_2)}]$$

qui est une matrice de type  $(d, \frac{n_2(n_2-1)}{2})$ .

Si on pose :

$$(5.13) \quad U_{(i_1, i_2)} = E_{(i_1, i_2)} \cdot E_{(i_1, i_2)}^t$$

(matrice de type (d,d) symétrique semi-définie positive), on a alors :

$$\Delta_{\text{brut}}^2 = \alpha^t \cdot U \cdot \alpha$$

avec

$$(5.14) \quad \boxed{U = \sum_{i_1 < i_2} U_{(i_1, i_2)}}$$

Les conditions de normalisation (5.9) s'écrivent :

$$(5.15) \quad \left. \begin{array}{l} I(F_{\min}) = 0 \\ I(F_{\max}) = 1 \end{array} \right\} \Leftrightarrow \begin{cases} \frac{1}{6} \alpha_0 + \frac{2}{3} \alpha_1 + \frac{1}{6} \alpha_2 = 0 \\ \frac{1}{6} \alpha_{d-3} + \frac{2}{3} \alpha_{d-2} + \frac{1}{6} \alpha_{d-1} = 1 \end{cases}$$

Si on pose :

$$G = \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix}$$

et

$$V_{01} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

alors, on peut écrire :

$$I^2(F_{\min}) + (I(F_{\max}) - 1)^2 = (G \cdot \alpha - V_{01})^t \cdot (G \cdot \alpha - V_{01})$$

ou encore :

$$(5.16) \quad \boxed{I^2(F_{\min}) + (I(F_{\max}) - 1)^2 = \alpha^t \cdot G^t G \cdot \alpha - 2 V_{01}^t \cdot G \cdot \alpha + 1}$$



La convexité de  $I$  dans l'intervalle  $[F_{\min}, F_{\max}]$  peut être exprimée par la condition nécessaire et suffisante suivante :

$$\forall t \in [F_{\min}, F_{\max}] : I''(t) \geq 0$$

$I$  étant de degré 3 sur chaque sous intervalle  $[u_i, u_{i+1}]$ , ( $i=3, \dots, d-1$ ), alors, il suffit que :

$$I''(u_i) \geq 0 \quad i=3, \dots, d$$

ou encore

$$(5.17) \quad \boxed{\frac{1}{h^2} (\alpha_{i-1} - 2\alpha_i + \alpha_{i+1}) \geq 0} \quad i=1, \dots, d-2$$

Pour la monotonie de  $I$ , il suffit d'imposer la condition suivante :

$$(5.18) \quad I'(F_{\min}) \geq 0$$

Cette condition, avec les conditions de convexité (5.17), implique la monotonie croissante de  $I$  dans l'intervalle de validité  $[F_{\min}, F_{\max}]$ . Cette condition s'exprime par :

$$(5.19) \quad \alpha_0 \cdot B'_0(F_{\min}) + \alpha_1 \cdot B'_1(F_{\min}) + \alpha_2 \cdot B'_2(F_{\min}) \geq 0$$

Or on a :

$$\begin{cases} B'_0(F_{\min}) = \frac{-1}{2h} \\ B'_1(F_{\min}) = 0 \\ B'_2(F_{\min}) = \frac{1}{2h} \end{cases}$$

d'où la condition équivalente :

$$(5.20) \quad \boxed{\alpha_2 - \alpha_0 \geq 0}$$

Les conditions (5.17) et (5.20) peuvent s'écrire sous la forme :

$$(5.21) \quad \boxed{A.\alpha \geq 0}$$

avec A matrice du type (d-1, d) donnée par

$$(5.22) \quad A = \begin{pmatrix} -1 & 0 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & 1 & -2 & 1 & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \cdot & \cdot & \cdot \\ & & & & & & 1 & -2 & 1 \end{pmatrix}$$

On pose :

$$(5.23) \quad \begin{cases} \psi = U + k.G^t G \\ b = -k.G^t.V_{01} \end{cases}$$

Compte tenu de (5.10), (5.16) et de (5.21) le problème (P.5.1) se transforme en :

$$(P'.5.1) \quad \begin{cases} \text{minimiser} & \{ \frac{1}{2} \langle \alpha, \psi.\alpha \rangle + \langle b, \alpha \rangle \} \\ \alpha \in \mathbb{R}^d \\ A.\alpha \geq 0 \end{cases}$$

### Remarque 5.1

En général, la matrice U de (5.14) n'est pas définie positive: ceci dépend de la valeur de d qui doit être inférieure ou égale à  $(n_1-1)(n_2-1)$  pour que U devienne définie positive. Lorsque cette condition est remplie, la matrice  $\psi$ , elle aussi, devient définie positive ; dans ce cas, l'application:

$$\alpha \longmapsto \frac{1}{2} \langle \alpha, \psi \cdot \alpha \rangle + \langle b, \alpha \rangle$$

est strictement convexe et le problème (P'.5.1) admet une solution unique.

- **Algorithme de résolution du problème (P'.5.1)**

Le problème (P'.5.1) peut se mettre sous la forme :

$$(P0) \quad \begin{cases} \min & \frac{1}{2} \langle x, \psi \cdot x \rangle + \langle b, x \rangle \\ x \in & \mathbb{R}^d \\ A \cdot x & \leq g \end{cases}$$

La matrice  $\psi$  étant supposée symétrique définie positive, alors on peut la factoriser par la méthode de Cholesky et écrire :

$$\psi = B^t \cdot B$$

$B$  : matrice réelle triangulaire supérieure. En faisant le changement de variable suivant :

$$(5.24) \quad y = B \cdot x \quad \Leftrightarrow \quad x = B^{-1} \cdot y$$

le problème (P0) se transforme en :

$$(P1) \quad \begin{cases} \min & \frac{1}{2} \langle y, y \rangle + \langle r, y \rangle \\ y \in & \mathbb{R}^d \\ C \cdot y & \leq g \end{cases}$$

avec : 
$$\begin{cases} \mathbf{g} = (\mathbf{B}^{-1})^t \cdot \mathbf{b} \\ \mathbf{C} = \mathbf{A} \cdot \mathbf{B}^{-1} \end{cases}$$

Nous proposons de résoudre (P1) par la méthode du gradient projeté [Rosen 60] qui consiste à projeter le gradient sur la frontière du domaine défini par les contraintes. Ceci donne un cheminement le long de la frontière dans la direction de la plus forte pente "relative" (figure 5.2), c'est-à-dire autorisée par les contraintes. La méthode du gradient projeté est essentiellement intéressante dans le cas des contraintes linéaires.

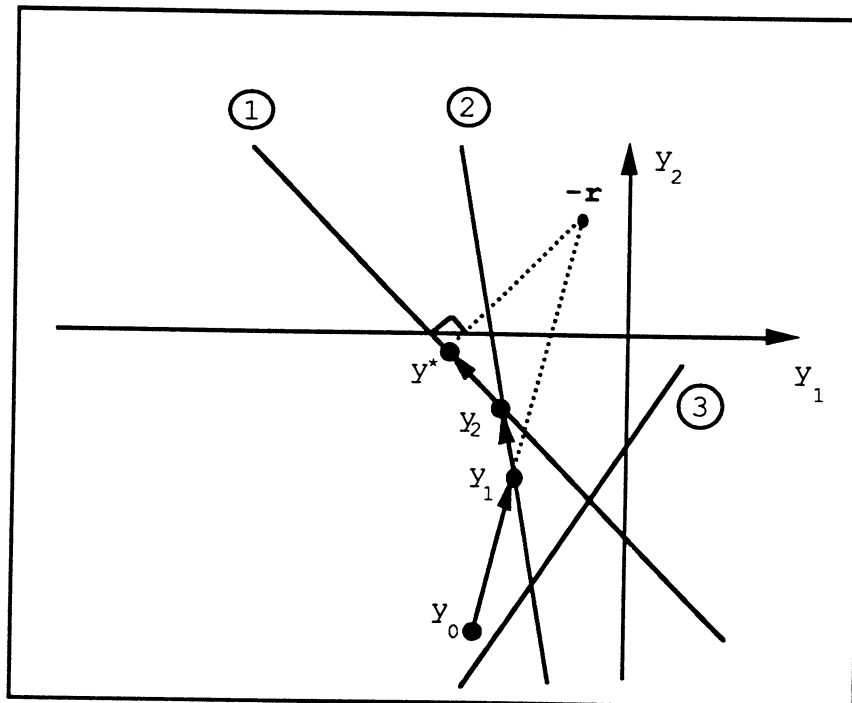


Figure 5.2 : Méthode du gradient projeté appliquée sur un problème du type (P1)

On suppose que l'on connaît une solution réalisable  $y_0$  (un tel point peut toujours être obtenu par programmation linéaire).

Soit  $I^\circ(y) = \{i : C_i \cdot y = g_i\}$  l'ensemble des indices des contraintes saturées au point  $y$ . En un point courant  $y_k$  on va chercher une direction de déplacement  $\delta_k$  ( $\|\delta_k\| = 1$ ) qui permet de diminuer le plus possible la fonction  $\Phi(y) = \frac{1}{2} \langle y, y \rangle + \langle r, y \rangle$  ( $\langle \nabla \Phi(y_k), \delta_k \rangle$  minimal), mais qui, au moins pour un petit déplacement, permet de rester dans l'ensemble des solutions réalisables. Ceci conduit à imposer à la direction  $\delta_k$  de satisfaire les relations :

$$(5.25) \quad C_i \cdot \delta_k = 0 \quad \text{pour tout } i \in I^\circ(y)$$

On note  $C^\circ$  la sous-matrice de  $C$  constituée par les lignes  $i \in I^\circ(y)$  de  $C$ . On suppose que  $C^\circ$  est de plein rang c'est-à-dire  $\text{rang}(C^\circ) = |I^\circ(y)| = q$ , ce qui revient à supposer qu'il n'y a pas de dégénérescence au point  $y_k$ . L'ensemble des vecteurs  $\delta_k$  satisfaisant (5.25) est le sous-espace vectoriel  $S^\circ$  de dimension  $d-q$  défini par  $S^\circ = \{y : C^\circ \cdot y = 0\}$ .

On montre (cf. [Minoux 83]) que la direction de déplacement  $\delta_k \in S^\circ \neq \{0\}$  tel que  $\langle \nabla \Phi(y_k), \delta_k \rangle$  soit minimal 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  $-\nabla \Phi(y_k)$  sur  $S^\circ$  donnée par :

$$(5.26) \quad \begin{cases} d_k = -P^\circ \cdot \nabla \Phi(y_k) = -P^\circ \cdot (y_k + r) \\ P^\circ = I - C^{\circ t} [C^\circ \cdot C^{\circ t}]^{-1} \cdot C^\circ \end{cases}$$

$P^\circ$  est la matrice de projection orthogonale sur  $S^\circ$ .

Après le calcul de la direction de déplacement  $d_k$  par (5.26) on détermine le pas de déplacement maximal autorisé par les contraintes

c'est-à-dire  $\alpha_{\max}$  tel que :

$$(5.27) \quad \alpha_{\max} = \max \{ \alpha > 0 : y_k + \alpha \cdot d_k \in D \}$$

où

$$D = \{ y \in \mathbb{R}^d : C \cdot y \leq g \}$$

Le pas de déplacement optimal est alors obtenu par :

$$(5.28) \quad \alpha_k^* = \min (1, \alpha_{\max})$$

et le point suivant  $y_{k+1}$  est obtenu par :

$$(5.29) \quad y_{k+1} = y_k + \alpha_k^* \cdot d_k$$

L'algorithme se poursuit de cette façon tant que  $d_k = -P^\circ \cdot (y_k + r) \neq 0$ .

Lorsque  $d_k = 0$  au point courant  $y_k$ , on a dans ce cas :

$$(5.30) \quad \nabla \Phi (y_k) + C^{\circ l} \cdot u = 0$$

avec

$$(5.31) \quad u = - \left[ C^\circ \cdot C^{\circ l} \right]^{-1} \cdot C^\circ \cdot \nabla \Phi (y_k)$$

Si  $u \geq 0$  alors il s'agit d'un optimum du problème (condition de Kuhn et Tucker). Sinon, le point courant n'est pas optimal et il faut trouver une autre direction de déplacement. On supprime dans  $I^\circ$  l'une des contraintes  $i$  pour lesquelles  $u_i < 0$  ; on obtient alors une nouvelle matrice  $C^\circ$  et une nouvelle matrice de projection  $P'^\circ$ , qui permet de trouver une nouvelle direction de déplacement :

$$(5.32) \quad d'_k = -P'^{\circ} \cdot (y_k + r)$$

$d'_k$  ainsi obtenu vérifie nécessairement :

- (i)  $d'_k \neq 0$ .
- (ii)  $d'_k$  est une direction de descente.

L'algorithme du gradient projeté appliqué au problème (P1) est donc conforme à la procédure suivante :

$y_0$  tel que  $C \cdot y_0 \leq 0$ ; {point de départ}

$k \leftarrow 0$ ;

fin  $\leftarrow$  FAUX;

**tantque** fin  $\neq$  VRAI **faire**

**début**

$L^{\circ} \leftarrow I^{\circ}(y_k)$ ; {ensemble de contraintes saturées en  $y_k$ }

$e : C^{\circ}$ ; {sous-matrice de  $C$  correspondant aux contraintes  $i \in L^{\circ}$ }

$P^{\circ} = I - C^{\circ t} \cdot [C^{\circ} \cdot C^{\circ t}]^{-1} \cdot C^{\circ}$ ; {matrice de projection orthogonale sur le  
sous espace  $S^{\circ}$ }

$d_k \leftarrow -P^{\circ} \cdot (y_k + r)$ ; {direction du déplacement}

**si**  $d_k = 0$  **alors**

**début**

$u \leftarrow -[C^{\circ} \cdot C^{\circ t}]^{-1} \cdot C^{\circ} \cdot (y_k + r)$ ;

**si**  $u_i \geq 0, (\forall i)$  **alors** fin  $\leftarrow$  VRAI

**sinon**

**début**

choisir  $i$  tel que  $u_i$  la composante la plus négative de  $u$ ;

$L^{\circ} \leftarrow L^{\circ} - \{i\}$ ;

**allera**  $e$

```

    fin_sinon
  fin_si
  sinon
  début
     $\alpha_{\max} \leftarrow \max \{ \alpha > 0 : y_k + \alpha \cdot d_k \in D \};$ 
     $\alpha_k^* \leftarrow \min (1, \alpha_{\max});$ 
     $y_{k+1} \leftarrow y_k + \alpha_k^* \cdot d_k;$ 
     $k \leftarrow k+1$ 
  fin_sinon
fin_tantque;
```

- **Problème additif associé**

La résolution du problème d'optimisation (P'.5.1) permet d'obtenir la meilleure approximation de la fonction d'indice concernant les données  $\{m_{ij}, F_{ij}\}$  ; cette approximation est représentée par l'ensemble des coefficients  $\{\alpha_i\}_{i=0}^{d-1}$  sur les B-splines cubiques uniformes de l'espace  $S(4, d-1)$ . On peut alors évaluer les grandeurs additives  $Z_{ij} = m_{ij} \cdot I(F_{ij})$  correspondant aux coupes  $\{(T_i, T_j) : 1 \leq i \leq n_1 \text{ et } n_1+1 \leq j \leq n_1+n_2\}$ . On pose :

$$T_{\min} = T_1 \quad \text{et} \quad T_{\max} = T_{n_1+n_2}$$

On s'intéresse dans ce paragraphe à rendre continue les quantités  $\{Z_{ij}\}$  en recherchant une fonction spline  $\sigma^*$  définie sur l'intervalle  $[T_{\min}, T_{\max}]$  et qui vérifie :

$$(5.33) \quad Z_{ij} \approx \int_{T_i}^{T_j} \sigma^*(t) dt$$

On remarque que les grandeurs additives  $\{Z_{ij}\}$  ne sont pas indépendantes, car elles sont générées par  $\{Z_{i,i+1}\}_{i=1}^{n_1+n_2-1}$ . Donc, on dispose de  $(n_1+n_2-1)$



degrés de liberté ; c'est la raison pour laquelle on considère l'espace  $S(4, n_1+n_2-2)$  dans lequel on va chercher :

$$\sigma^* = \sum_{i=0}^{n_1+n_2-2} \beta_i^* \cdot B_i$$

la meilleure approximation de  $I(f) \cdot dm/dT$  au sens des moindres carrés. Nous avons alors à résoudre le problème de minimisation suivant :

$$(P.5.2) \quad \left\{ \begin{array}{l} \text{minimiser} \quad \sum_{\substack{i \in \{1, \dots, n_1\} \\ j \in \{n_1+1, \dots, n_1+n_2\}}} \left( \int_{T_i}^{T_j} \sigma(t) dt - Z_{ij} \right)^2 \\ \sigma \in S(4, n_1+n_2-2) \\ \sigma = \sum_{i=0}^{n_1+n_2-2} \beta_i B_i \end{array} \right.$$

En utilisant (1.34) on peut écrire :

$$(5.34) \quad \left( \int_{T_i}^{T_j} \sigma(t) \cdot dt - Z_{ij} \right)^2 = \langle \beta, R R^t \cdot \beta \rangle_{\mathbb{R}^{n_1+n_2-1}} - 2 \langle R \cdot Z, \beta \rangle_{\mathbb{R}^{n_1+n_2-1}} + \langle Z, Z \rangle_{\mathbb{R}^{n_1 \cdot n_2}}$$

avec

$$(5.35) \quad Z = \begin{pmatrix} Z_{1, n_1+1} \\ Z_{1, n_1+2} \\ \vdots \\ Z_{1, n_1+n_2} \\ \vdots \\ \vdots \\ Z_{n_1, n_1+n_2} \end{pmatrix} \quad \text{et} \quad R = \left( R_{r k} \right)_{\substack{r=0, \dots, n_1+n_2-2 \\ k=1, \dots, n_1 \cdot n_2}}$$

$$R_{r k} = \int_{T_i}^{T_j} B_r(t) dt$$

$$\begin{array}{l} i = (k-1) \operatorname{div} n_2 + 1 \\ j = (k-1) \operatorname{mod} n_2 + 1 + n_1 \end{array}$$

Le problème (P.5.2) est alors équivalent au problème suivant :

$$(P'.5.2) \quad \begin{cases} \text{minimiser } \left\{ \frac{1}{2} \langle \beta, \psi \cdot \beta \rangle + \langle b, \beta \rangle + c \right\} \\ \beta \in \mathbb{R}^{n_1+n_2-1} \end{cases}$$

avec

$$(5.36) \quad \begin{cases} \psi = R \cdot R^t \\ b = -R \cdot Z \\ c = \frac{1}{2} Z^t \cdot Z \end{cases}$$

Si on note  $\beta^* = -\psi^{-1} \cdot b$  solution du (P'.5.2), alors  $\sigma^* = \sum_{i=0}^{n_1+n_2-2} \beta_i^* \cdot B_i$  est la solution du (P.5.2). De plus,  $\sigma^*$  vérifie :

$$(5.37) \quad \sigma^*(t) \approx [I(f) \cdot dm/dT](t) \quad t \in [T_{\min}, T_{\max}]$$

Ceci va permettre d'écrire :

$$(5.38) \quad \int_{T_i}^{T_f} I(f(t)) \cdot dm(t) \approx \int_{T_i}^{T_f} \sigma^*(t) \cdot dt$$

- **Evaluation de la propriété non additive**

Nous avons vu ci-dessus que la spline  $\sigma^*$  représente la meilleure approximation de  $I(f) \cdot dm/dT$  au sens du problème d'optimisation (P.5.2), et que la relation (5.38) permet d'évaluer les quantités additives :

$$\int_{T_i}^{T_f} I(f(t)) \cdot dm(t)$$

pour des coupes  $(T_i, T_f)$  telles que :

$$T_{\min} \leq T_i < T_f \leq T_{\max}$$

D'autre part, par additivité on peut écrire :

$$(5.39) \quad \int_{T_i}^{T_f} I(f(t)).dm(t) = \left( \int_{T_i}^{T_f} dm(t) \right) \cdot I(F(T_i, T_f))$$

En utilisant (5.38), on obtient :

$$(5.40) \quad F(T_i, T_f) \approx I^{-1} \left( \frac{\int_{T_i}^{T_f} \sigma^*(t).dt}{\int_{T_i}^{T_f} dm(t)} \right)$$

### Remarque 5.2

La fonction d'indice est définie sur l'intervalle  $[F_{\min}, F_{\max}]$ , et prend ses valeurs dans  $[0,1]$  ; par conséquent, la relation (5.40) ne peut pas être utilisée pour des coupes  $(T_i, T_f)$  pour lesquelles :

$$\frac{\int_{T_i}^{T_f} \sigma^*(t).dt}{\int_{T_i}^{T_f} dm(t)} \notin [0,1]$$

donc, ce procédé permet d'évaluer  $F$  pour des coupes  $(T_i, T_f)$  dont les valeurs  $F(T_i, T_f)$  sont situées dans l'intervalle de validité  $[F_{\min}, F_{\max}]$ . En pratique, ceci revient à considérer des coupes larges, c'est-à-dire :

$$T_i \in [T_1, T_{n1}] \quad \text{et} \quad T_f \in [T_{n1+1}, T_{n1+n2}]$$

bien que certaines coupes étroites "médianes" puissent être également calculées.

## 5.2 Deuxième approche: reconstitution des quantités d'indice

Pour une fonction d'indice donnée I, on note:

$$(5.41) \quad \hat{Z}_{ij} = \int_{T_i}^{T_j} \sigma^*(t) dt$$

la quantité d'indice reconstituée pour la coupe  $(T_i, T_j)$ .  $\sigma^*$  étant la solution du problème additif (P.5.2). Dans cette approche, on remplace le terme  $\Delta_{\text{brut}}^2$ , qui est la somme des carrés des erreurs sur les équations d'indice, par la somme des carrés des écarts entre quantités d'indice  $Z_{ij}$  et quantité d'indice reconstituée  $\hat{Z}_{ij}$ :

$$(5.42) \quad \text{Ecart}_{\text{brut}}^2 = \sum_{i,j} (\hat{Z}_{ij} - Z_{ij})^2$$

Par conséquent, nous avons à résoudre le problème de minimisation suivant:

$$(P.5.3) \quad \left\{ \begin{array}{l} \text{minimiser} \left\{ \text{Ecart}_{\text{brut}}^2 + k [ I^2(F_{\min}) + (I(F_{\max})-1)^2 ] \right\} \\ I \in S(4,d-1) \\ I' \geq 0 \\ I'' \geq 0 \end{array} \right.$$

Pour une fonction spline I appartenant à  $S(4,d-1)$  telle que:

$$I = \sum_{i=0}^{d-1} \alpha_i B_i \quad \text{définie sur } [F_{\min}, F_{\max}]$$

la solution du problème (P.5.2) correspondant est donnée par:

$$\sigma^* = \sum_{i=0}^{n_1+n_2-2} \beta_i^* \cdot B_i$$

avec:

$$(5.43) \quad \beta^* = (R \cdot R^t)^{-1} \cdot R \cdot Z$$

où R et Z sont définis par (5.35). Nous avons alors:

$$(5.44) \quad \hat{Z}_{ij} = \langle R_{\cdot k}, \beta^* \rangle_{\mathbb{R}^{n_1+n_2-1}}$$

avec

$$(5.45) \quad k = (i-1) n_2 + (j-n_1)$$

On note:

$$M = \text{diag} (m_{1,n_1+1}, m_{1,n_1+2}, \dots, m_{1,n_1+n_2}, \dots, m_{n_1,n_1+1}, \dots, m_{n_1,n_1+n_2})$$

la matrice carrée diagonale ayant  $m_{ij}$ ,  $i=1, \dots, n_1$  et  $j=n_1+1, \dots, n_1+n_2$  comme éléments diagonaux. Soit :

$$(5.46) \quad E = (E_{kq})_{\substack{k=1, \dots, n_1 \cdot n_2 \\ q=0, \dots, d-1}}$$

tel que :

$$(5.47) \quad \begin{cases} E_{kq} = B_q (F_{ij}) \\ i = (k-1) \text{ div } n_2 + 1 \\ j = (k-1) \text{ mod } n_2 + 1 + n_1 \end{cases}$$

Alors, on peut écrire :

$$(5.48) \quad Z = M.E.\alpha$$

Donc, on a:

$$(5.49) \quad \widehat{Z} = R^t.(R.R^t)^{-1}.R.M.E.\alpha$$

Par conséquent, le terme Ecart<sup>2</sup><sub>brut</sub> peut s'écrire sous la forme:

$$(5.50) \quad \text{Ecart}_{\text{brut}}^2 = \langle \alpha, L.\alpha \rangle_{\mathbb{R}^d}$$

avec:

$$(5.51) \quad L = \left( R^t.(R.R^t)^{-1}.R - \text{ID} \right).M.E \quad \text{ID: matrice identité.}$$

Le problème (P.5.3) est alors équivalent au problème suivant:

$$(P'.5.3) \quad \left\{ \begin{array}{l} \text{minimiser} \quad \left\{ \frac{1}{2} \langle \alpha, \Omega.\alpha \rangle + \langle b, \alpha \rangle \right\} \\ \alpha \in \mathbb{R}^d \\ A.\alpha \geq 0 \end{array} \right.$$

avec:

$$(5.52) \quad \left\{ \begin{array}{l} \Omega = L + k.G^tG \\ b = -k.G^t.V_{01} \end{array} \right.$$

A est la matrice (d-1) x d définie par (5.22).

Remarque 5.3

Les quantités d'indice reconstituées  $\{\widehat{Z}_{ij}\}$  définies par (5.41) vérifient les équations d'indice (5.4), c'est-à-dire:

$$\widehat{Z}_{i_1, j_2} - \widehat{Z}_{i_1, j_1} = \widehat{Z}_{i_2, j_2} - \widehat{Z}_{i_2, j_1}$$

Alors, la minimisation du terme Ecart<sup>2</sup> du (5.42), c'est aussi une minimisation (en quelque sort) du terme  $\Delta_{\text{brut}}^2$  du (5.8), car on a:

$$\delta_{(i_1, i_2) \times (j_1, j_2)} = (\widehat{Z}_{i_1, j_2} - Z_{i_1, j_2}) - (\widehat{Z}_{i_1, j_1} - Z_{i_1, j_1}) - (\widehat{Z}_{i_2, j_2} - Z_{i_2, j_2}) + (\widehat{Z}_{i_2, j_1} - Z_{i_2, j_1})$$

Donc, on peut s'attendre à une solution du problème (P.5.3) très proche (quasi identique) de celle du problème (P.5.1). Ceci est vérifié en pratique, dans le cas de l'application au calcul de point de trouble des gazoles (voir plus loin). Autre point distinctif de cette approche est qu'elle n'exige pas des mesures chevauchées de la propriété non additive F, et à titre d'exemple, cette approche est applicable dans le cas des mesures adjacentes.

**5.3 Troisième approche: reconstitution des mesures**

Etant donné une fonction indice I pour la propriété non additive F. On note :

$$(5.53) \quad \widehat{F}_{ij} = I^{-1} \left( \frac{\widehat{Z}_{ij}}{\int_{T_i}^{T_j} dm(T)} \right)$$

la valeur reconstituée pour la coupe  $(T_i, T_j)$ .  $\hat{Z}_{ij}$  est la quantité d'indice reconstituée définie par (5.41). L'idée dans cette approche consiste à chercher une fonction indice, convexe et monotone, qui minimise l'erreur relative en  $\{\hat{F}_{ij}\}$  par rapport aux données  $\{F_{ij}\}$ , c'est-à-dire :

$$(5.54) \quad \text{Err}_{\text{brut}}^2 = \frac{\sum_{i,j} (\hat{F}_{ij} - F_{ij})^2}{\sum_{i,j} F_{ij}^2}$$

Par conséquent, on a à résoudre le problème de minimisation suivant :

$$(P.5.4) \quad \left\{ \begin{array}{l} \text{minimiser} \left\{ \text{Err}_{\text{brut}}^2 + k [ I^2(F_{\min}) + (I(F_{\max})-1)^2 ] \right\} \\ I \in S(4, d-1) \\ I' \geq 0 \\ I'' \geq 0 \end{array} \right.$$

C'est un problème d'optimisation non linéaire. La valeur reconstituée  $\hat{F}_{ij}$  est l'image réciproque par  $I$  d'une grandeur dépendant linéairement des coefficients  $\{\alpha_i\}_{i=0}^{d-1}$ . Ceci découle du fait que  $\hat{Z}_{ij}$  dépend linéairement de  $\alpha = \{\alpha_i\}_{i=0}^{d-1}$  (voir expression (5.49)). Par conséquent, on peut écrire:

$$(5.55) \quad I(\hat{F}_{ij}) = \Phi_{k.} \cdot \alpha \quad k = (i-1)n_2 + (j-1)n_1$$

avec:

$$(5.56) \quad \Phi = M^{-1} \cdot R^t \cdot (R \cdot R^t)^{-1} \cdot R \cdot M \cdot E$$

une matrice du type  $(n_1 \cdot n_2) \times d$ .

Par conséquent, le terme  $\text{Err}_{\text{brut}}^2$  peut s'écrire sous la forme :



$$(5.57) \quad \text{Err}_{\text{brut}}^2 = \frac{1}{\|F_{\text{brut}}\|^2} \sum_{i,j} \left[ I^{-1}(\Phi_{k.} \cdot \alpha) - F_{ij} \right]^2$$

Dans cette expression l'indice  $k$  est déterminé, à partir de  $i$  et  $j$ , par la relation (5.45) et  $\|F_{\text{brut}}\|^2$  représente la somme des carrés des  $\{F_{ij}\}$ .

#### 5.4 Application au calcul de point de trouble des gazoles

Pour un brut donné, on dispose d'un ensemble de seize mesures expérimentales  $\{F_{ij}, m_{ij}\}$  de point de trouble (propriété non additive) correspondant à seize coupes ( $T_i, T_j$ ) tels que  $T_i \in \{T_1, \dots, T_4\}$  et  $T_j \in \{T_5, \dots, T_8\}$  (ici  $n_1 = n_2 = 4$ ). On peut donc appliquer les procédés de recherche de fonction d'indice présentés aux paragraphes §5.1, §5.2 et §5.3, ce qui permet d'obtenir une fonction d'indice propre au brut considéré. Ce type de traitement est appelé traitement local à l'encontre du traitement global qui consiste à traiter une famille de bruts simultanément avec pour but d'obtenir une fonction d'indice commune. Dans le cas du traitement global, le terme  $\Delta_{\text{brut}}^2$  (resp.  $\text{Ecart}_{\text{brut}}^2$  ou  $\text{Err}_{\text{brut}}^2$ ) sera remplacé par  $\sum_{\text{brut}} \Delta_{\text{brut}}^2$  (resp.  $\sum_{\text{brut}} \text{Ecart}_{\text{brut}}^2$  ou  $\sum_{\text{brut}} \text{Err}_{\text{brut}}^2$ ).

L'intérêt du traitement global réside dans la possibilité d'évaluer la propriété non additive pour un mélange de plusieurs coupes concernant des bruts différents en se servant de l'additivité qui s'exprime par une relation du type :

$$(5.58) \quad \sum_{\text{brut}} \left[ m_{\text{brut}} \cdot I_{\text{global}}(F_{\text{brut}}) \right] = \left( \sum_{\text{brut}} m_{\text{brut}} \right) \cdot I_{\text{global}}(F_{\text{mélange}})$$

Une expérimentation sur trente six bruts a été réalisée pour les deux

types de traitement local et global. Le traitement local ne présente aucun problème et les résultats obtenus sont encourageants. Cependant la détermination de point de trouble d'un mélange de deux bruts différents n'est pas possible (aux deux bruts différents correspond deux fonctions d'indice différentes). C'est la raison pour laquelle l'utilité de ce type de traitement est limitée à l'évaluation de la propriété non additive pour un mélange de coupes de même brut.

Afin de comparer les résultats obtenus, nous avons associé à chacune des approches étudiées, une fonction erreur :

- première approche :  $\text{Erreur}_{\text{brut}} \text{ (I)} = \sqrt{\Delta_{\text{brut}}^2}$
- deuxième approche :  $\text{Erreur}_{\text{brut}} \text{ (II)} = \sqrt{\text{Ecart}_{\text{brut}}^2}$
- troisième approche :  $\text{Erreur}_{\text{brut}} \text{ (III)} = \sqrt{\text{Err}_{\text{brut}}^2} * 100$  (pourcentage)

Ainsi, nous avons évalué les erreurs correspondantes à chaque type de traitement local et global. D'autre part, nous avons comparé ces résultats avec ce qu'on obtient en utilisant l'indice de mélange exponentiel (actuellement utilisé). Cet indice est donné par l'expression suivante :

$$(5.59) \quad I(t) = 10000 \left( \frac{t + 273.15}{333.15} \right)^{\frac{4000}{200 - t}}$$

Il est clair que cet indice peut être bien représenté par sa meilleure approximation dans  $S(4, d-1)$  sur  $[F_{\min}, F_{\max}]$  sous forme d'une combinaison linéaire de B-splines (voir figure 5.3). Les différents indices trouvés (par traitement global) ainsi l'indice exponentiel normalisé ont été tracés (figure 5.4).

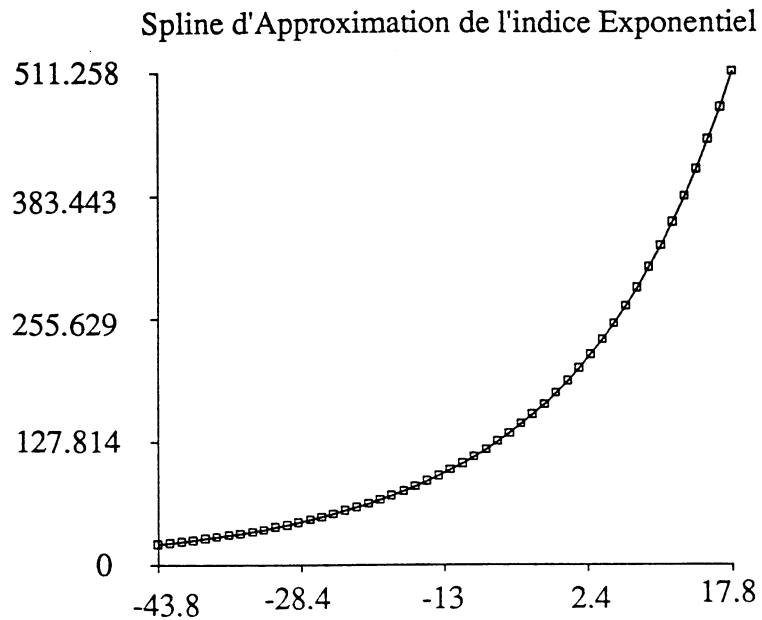


Figure 5.3-(a) : Spline d'approximation et l'ensemble de points pris sur la courbe exponentielle.

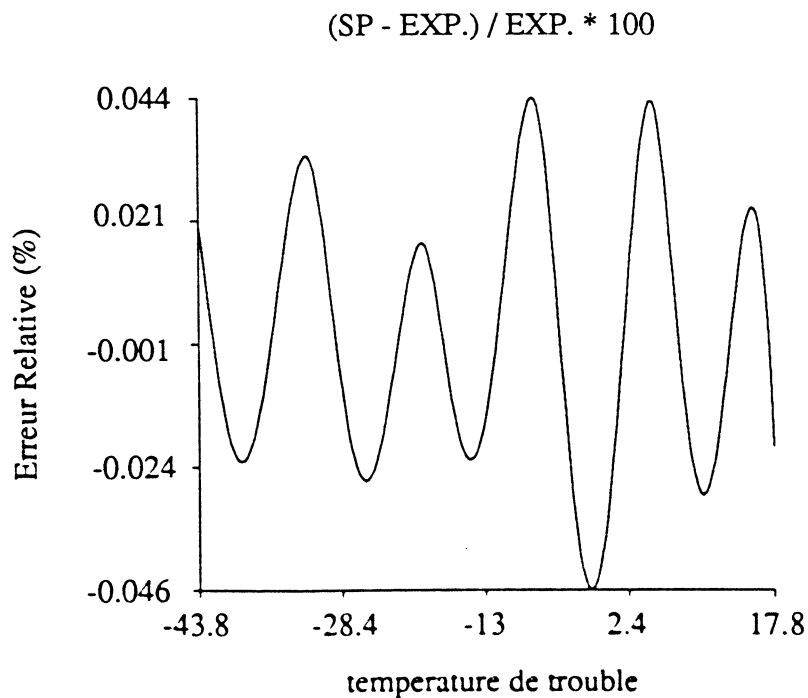


Figure 5.3-(b) : Erreur relative (%) dans l'approximation de l'indice exponentiel par une spline de  $S(4,d-1)$ .

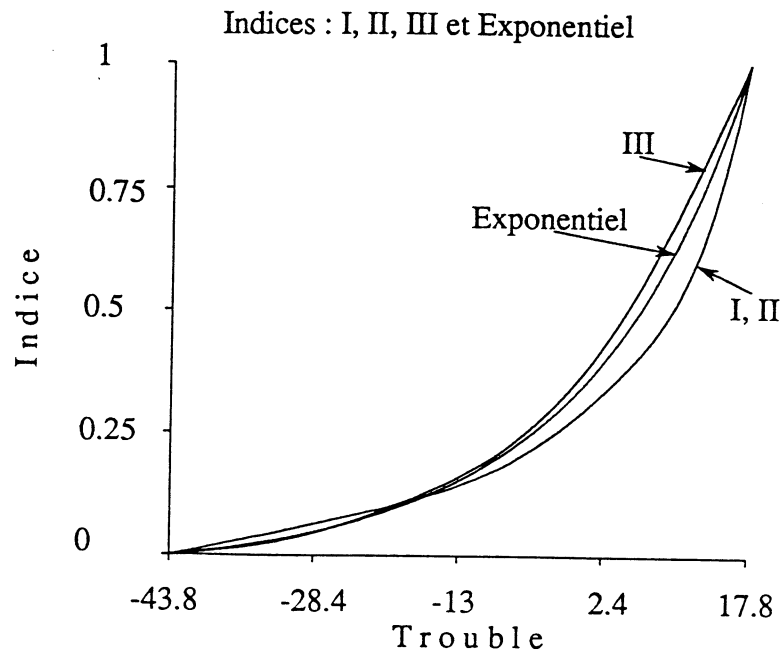


Figure 5.4 : Courbes des différents indices trouvés et la courbe de l'indice exponentiel normalisé.

Comme nous l'avons signalé, dans la remarque 5.3, les indices trouvés par les deux premières approches sont pratiquement identiques ; c'est la raison pour laquelle nous ne présentons que les résultats d'évaluation des erreurs pour les approches I et III. Ces résultats sont présentés dans les tables 5.1 et 5.2. Dans la table 5.1 se trouvent cinq colonnes représentant les erreurs, de la première approche, correspondantes aux traitements local, global, par l'indice exponentiel, par la spline d'approximation de l'indice exponentiel et par l'indice III. On remarque dans cette table que les résultats de l'indice exponentiel et ceux de son approximation sont très proches, ceci justifie l'utilisation de cette approximation à la place de l'indice exponentiel. La table 5.2 contient quatre colonnes représentant les erreurs pour les traitements local, global, par l'approximation de l'indice exponentiel et par l'indice I. En bas de chaque table se trouvent les valeurs moyennes des erreurs ainsi que leurs écarts-type. Une visualisation graphique, sous forme d'histogrammes, des différentes colonnes de la table 5.1 est présentée par

les figures 5.5-(a,b,c,d). De même, nous avons visualisé les résultats de la table 5.2 sur les figures 5.6-(a,b,c,d).

n°	Brut	LOCAL	GLOBAL	EXP.	SP (EXP)	III
1	A340	0.278	0.732	1.059	1.043	0.893
2	ARABE LEGER	1.639	6.095	7.507	7.483	7.341
3	ARABE LOURD	0.410	3.400	3.643	3.642	4.884
4	ARABE MOYEN	0.913	1.583	1.607	1.592	2.201
5	ASHTART	0.389	5.157	3.416	3.393	3.352
6	BACHAQUERO	0.365	0.709	0.892	0.897	0.771
7	DUBAI	0.610	2.191	3.294	3.296	3.064
8	EKOFISK	0.909	1.990	2.659	2.678	2.499
9	EKOFISK 78	0.949	1.525	2.287	2.280	2.085
10	FLOTTA	0.674	0.924	1.405	1.388	1.638
11	FORTIES	0.511	1.171	0.943	0.953	1.678
12	IRAN LOURD	1.086	2.414	3.006	3.026	4.021
13	IRAN LEGER	1.092	4.107	5.300	5.302	5.267
14	ISTHMUS	0.406	1.971	2.231	2.201	2.072
15	KIRKUK	0.394	0.898	0.945	0.954	1.834
16	KOLE	2.132	2.973	3.768	3.775	4.313
17	KOWEIT	0.166	0.823	1.664	1.669	1.194
18	LAGO MEDIO	1.091	2.555	3.156	3.156	2.904
19	LEONA	1.019	2.202	2.155	2.156	3.034
20	MANDJI	1.305	15.407	12.051	12.050	10.910
21	NIGER BRASS	1.815	8.183	12.810	12.854	15.060
22	NIGER FOR	1.194	2.810	2.969	3.004	4.243
23	NIGER LIGHT	1.496	2.351	4.069	4.076	4.354
24	QATAR	1.708	3.885	5.392	5.387	5.113
25	QUA IBOE	0.398	2.105	1.605	1.627	1.841
26	RIO	0.442	1.190	1.603	1.630	2.455
27	SIRRI	1.081	1.577	2.301	2.295	2.161
28	SOUEDIE	0.329	1.204	1.172	1.186	1.971
29	SOVIET	0.395	0.833	0.905	0.903	1.114
30	SUEZ	0.282	1.261	2.078	2.086	1.584
31	SUEZ 325	0.588	1.241	2.084	2.104	1.886
32	TIA LIGHT	0.642	1.424	1.875	1.906	2.743
33	TIA MEDIUM	0.576	1.815	1.268	1.270	1.561
34	UMM SHAIFF	0.321	0.929	0.792	0.772	1.264
35	ZAKHUM	0.789	2.568	3.180	3.196	4.519
36	ZUEITINA	0.787	4.441	5.759	5.772	5.975
	moyenne	0.811	2.685	3.135	3.139	3.439
	ecart type	0.488	2.687	2.722	2.725	2.819

Table 5.1 : Erreurs de la première approche.

n°	Brut	LOCAL (%)	GLOBAL (%)	EXP. (%)	SP (EXP) (%)	I (%)
1	A340	1.16	1.56		2.40	1.57
2	ARABE LEGER	5.65	11.19		12.07	13.36
3	ARABE LOURD	1.05	8.59		7.70	8.70
4	ARABE MOYEN	1.92	3.14		2.53	3.37
5	ASHTART	1.73	5.74		5.52	9.77
6	BACHAQUERO	0.89	2.02		1.63	2.87
7	DUBAI	2.78	7.31		7.99	7.36
8	EKOFISK	2.40	3.51		4.50	3.71
9	EKOFISK 78	2.09	3.24		3.82	3.12
10	FLOTTA	1.85	2.04		2.13	1.77
11	FORTIES	1.10	2.05		1.51	2.20
12	IRAN LOURD	3.32	4.57		3.82	4.16
13	IRAN LEGER	2.84	12.93		13.47	13.59
14	ISTHMUS	1.07	2.31		2.71	2.55
15	KIRKUK	0.97	2.69		1.70	2.30
16	KOLE	5.70	6.83		6.90	6.04
17	KOWEIT	0.62	2.48		2.98	2.21
18	LAGO MEDIO	1.96	3.80		4.99	4.29
19	LEONA	2.00	3.39		2.64	4.38
20	MANDJI	1.56	17.90		19.90	26.09
21	NIGER BRASS	4.56	13.65		13.79	11.20
22	NIGER FOR	2.80	3.88		3.30	4.68
23	NIGER LIGHT	3.01	4.16		4.69	2.97
24	QATAR	4.27	8.15		9.23	8.32
25	QUA IBOE	1.36	3.40		3.13	6.42
26	RIO	1.74	3.05		2.56	3.11
27	SIRRI	2.78	3.53		3.63	3.47
28	SOUEDIE	1.72	4.35		3.22	4.27
29	SOVIET	0.96	2.25		1.76	2.25
30	SUEZ	0.95	3.06		3.28	3.91
31	SUEZ 325	2.87	3.64		4.98	3.87
32	TIA LIGHT	1.75	2.48		2.13	2.18
33	TIA MEDIUM	1.85	3.77		3.32	6.66
34	UMM SHAIFF	1.19	1.88		1.83	1.78
35	ZAKHUM	1.65	5.81		4.45	5.39
36	ZUEITINA	2.36	8.76		10.27	9.17
	moyenne	2.18 %	5.09 %		5.18 %	5.64 %
	ecart type	1.24 %	3.73 %		4.12 %	4.69 %

Table 5.2 : Erreurs de la troisième approche.

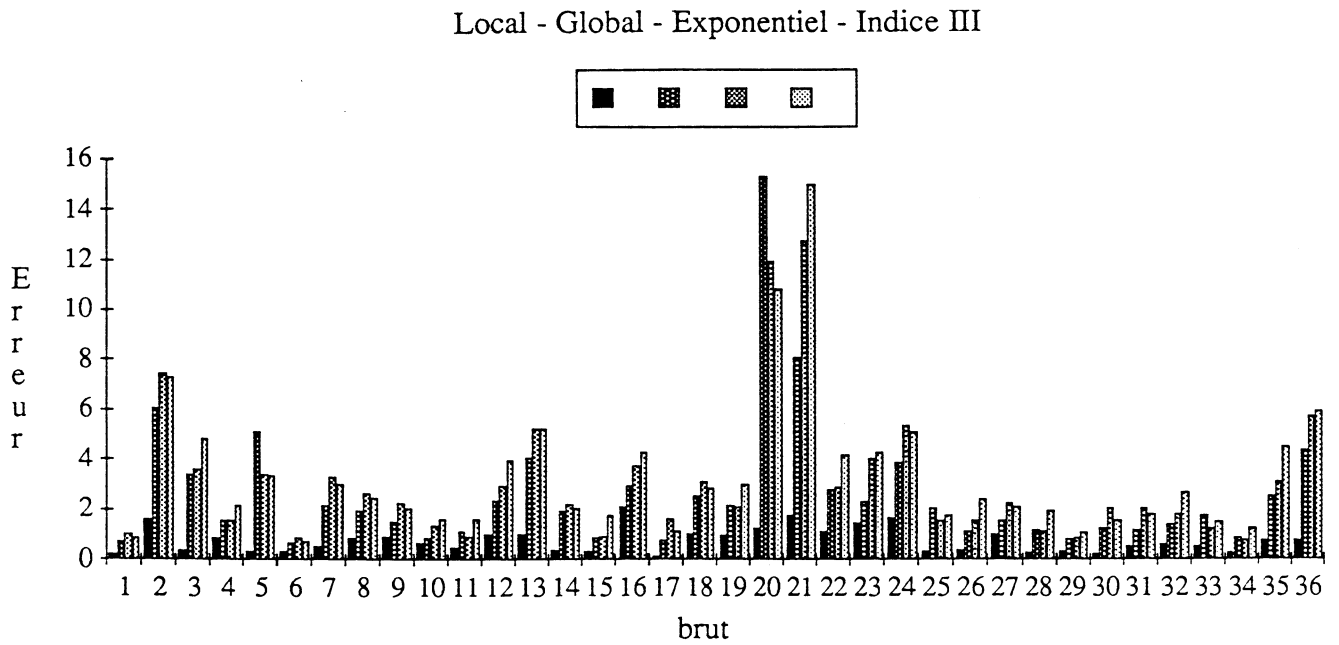


Figure 5.5-(a) : Première approche : histogramme d'erreurs pour les traitements local, global, exponentiel et par l'indice III.

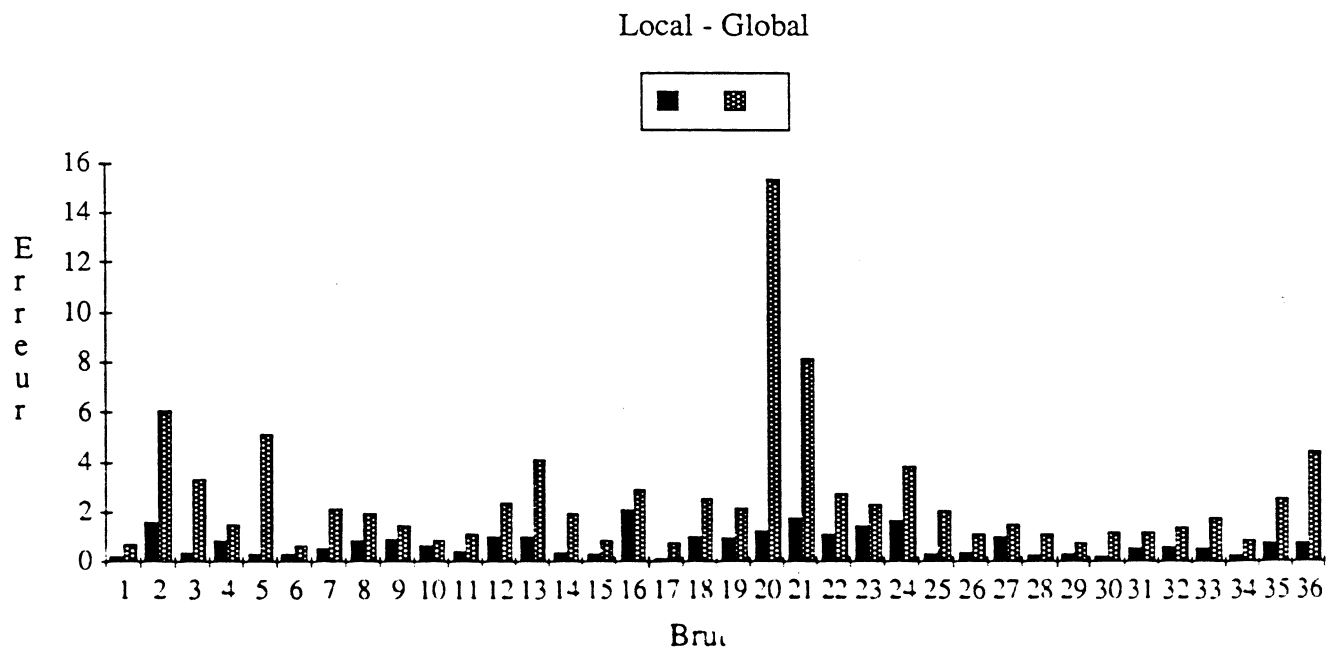


Figure 5.5-(b) : Première approche : histogramme d'erreurs pour les traitements local et global.

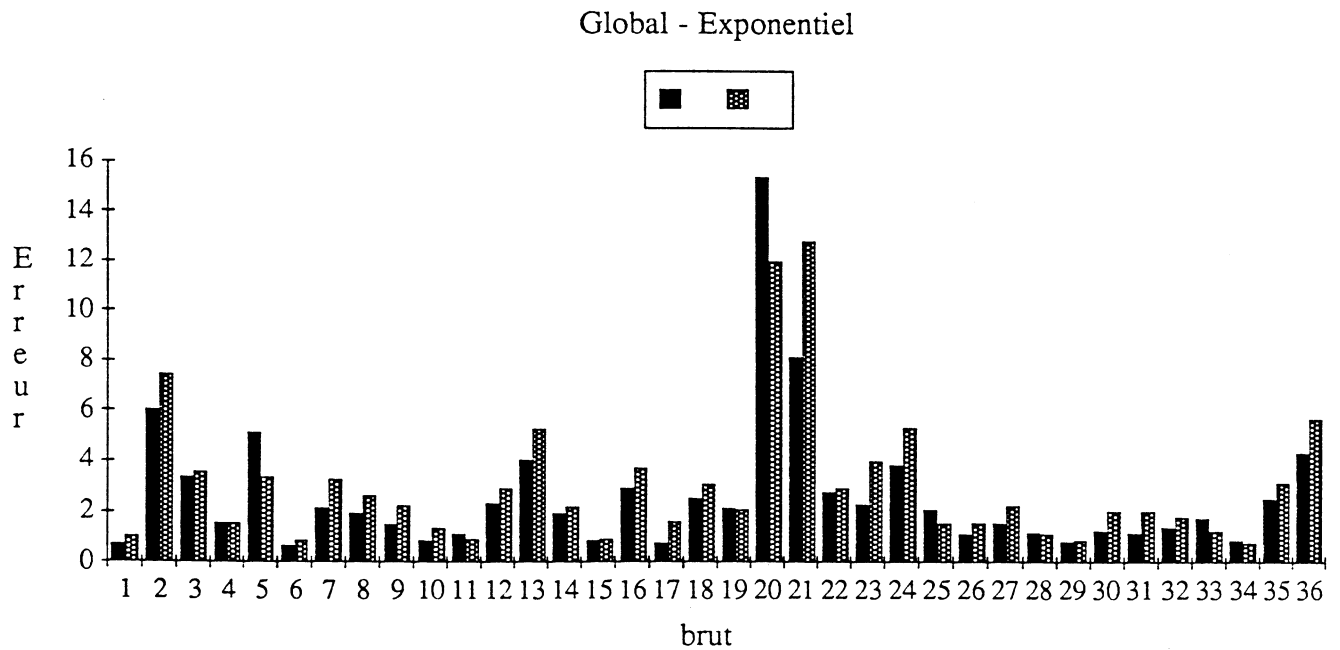


Figure 5.5-(c) : Première approche : histogramme d'erreurs pour les traitements global et exponentiel.

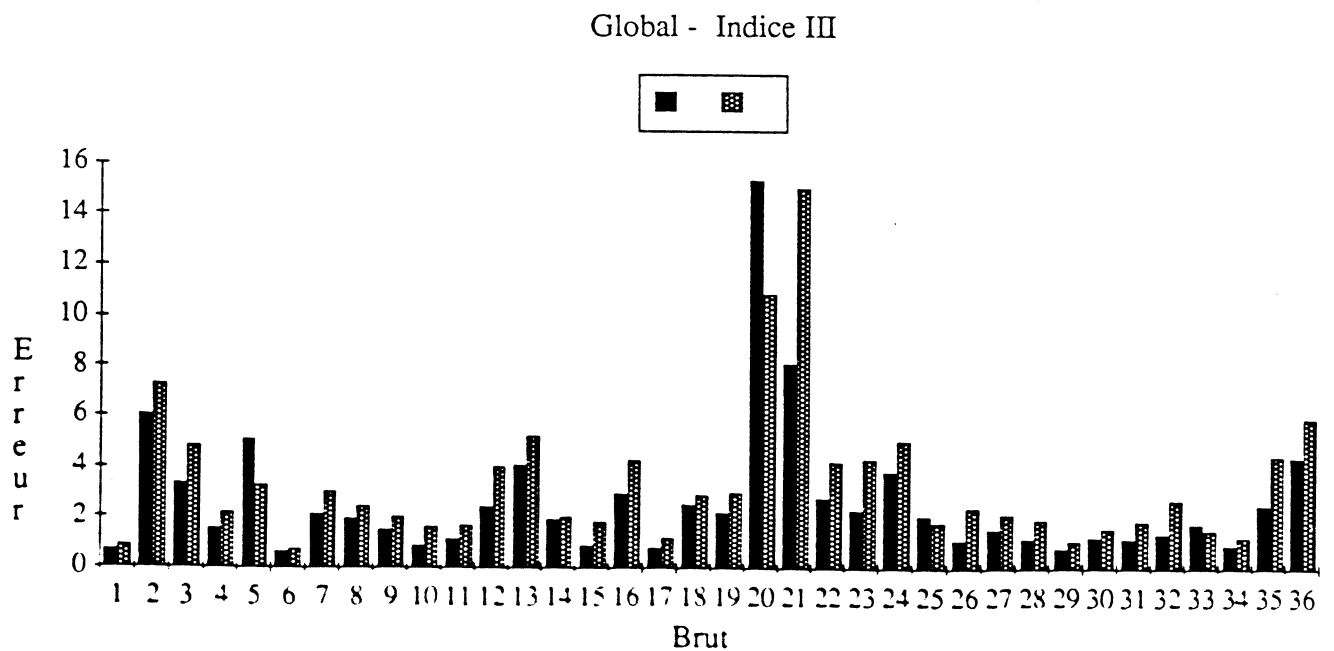


Figure 5.5-(d) : Première approche : histogramme d'erreurs pour les traitements global et par l'indice III.



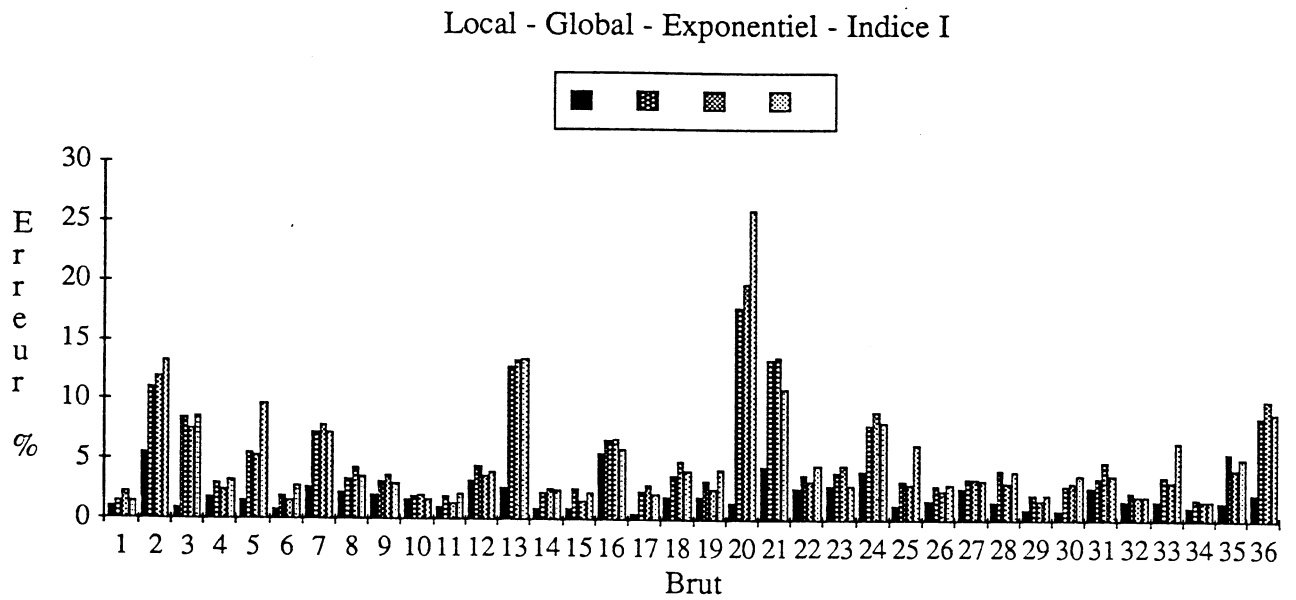


Figure 5.6-(a) : Troisième approche : histogramme d'erreurs pour les traitements local, global, exponentiel et par l'indice I.

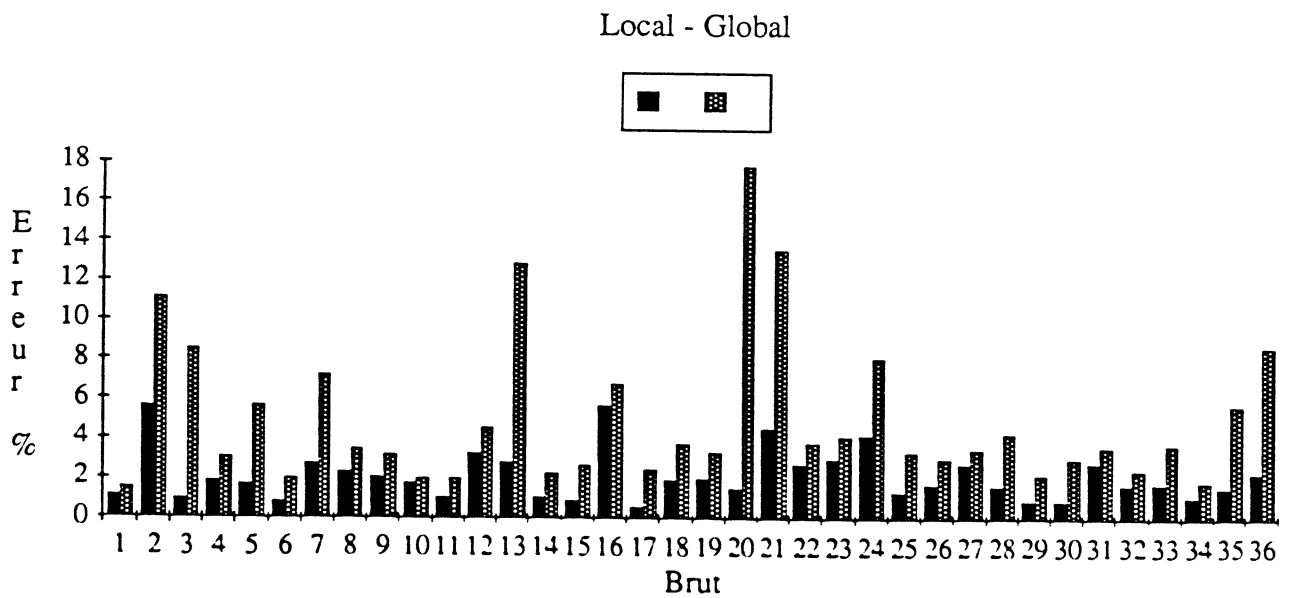


Figure 5.6-(b) : Troisième approche : histogramme d'erreurs pour les traitements local et global.

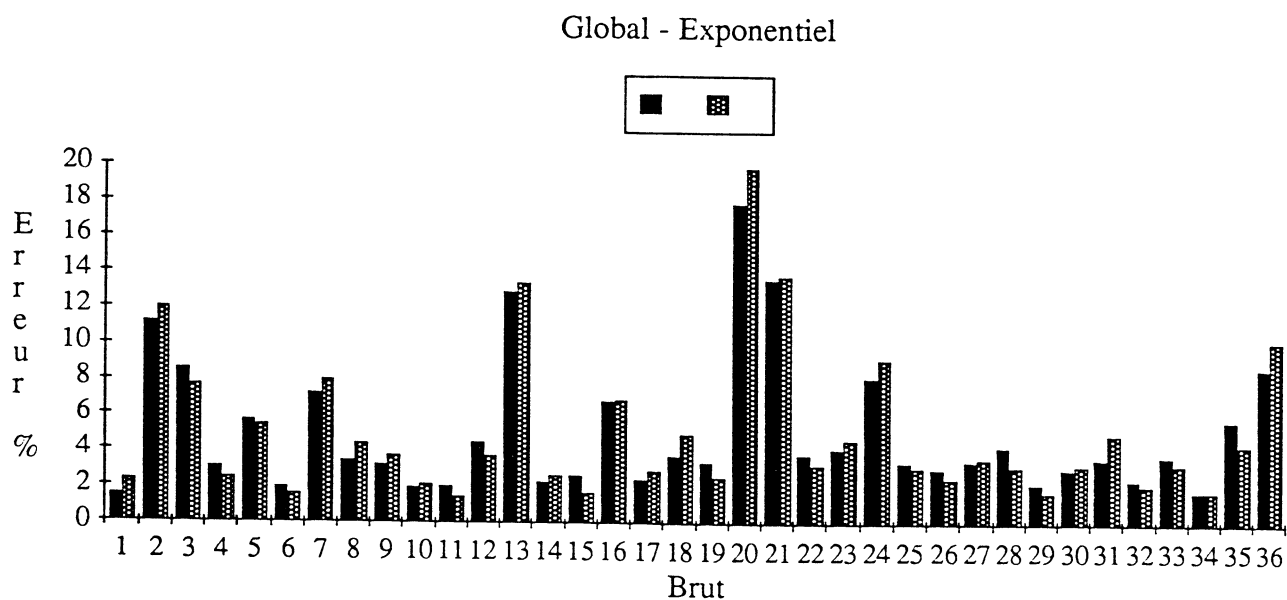


Figure 5.6-(c) : Troisième approche : histogramme d'erreurs pour les traitements global et exponentiel.

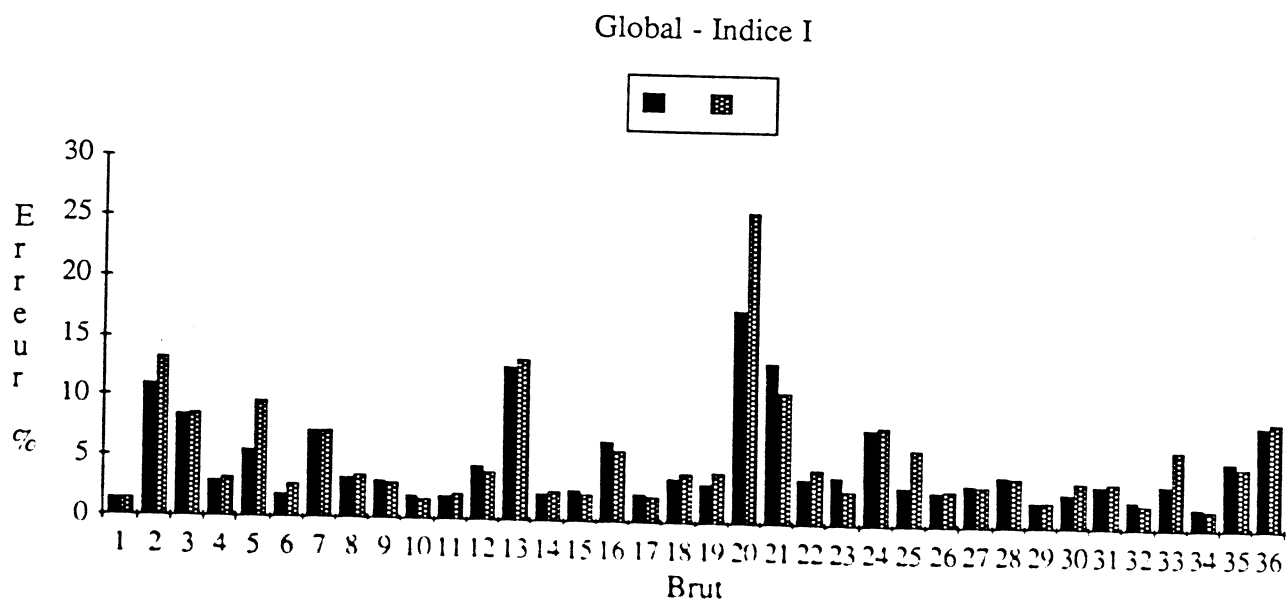


Figure 5.6-(d) : Troisième approche : histogramme d'erreurs pour les traitements global et par l'indice I.

Enfin, dans le cas du traitement global, un certain nombre de bruts se distinguent du reste par une forte erreur . Ces bruts sont représentés sur les figures 5.5 et 5.6 par les abscisses 2, 3, 5, 7, 13, 20, 21, 24 et 36.

## 5.5 Conclusion

La détermination de l'indice de mélange pour les propriétés non additives permet l'application des procédés de gestion de mélanges additifs et en conséquence, l'évaluation de ces propriétés pour des mélanges complexes.

La fonction d'indice, étant une combinaison linéaire de B-splines, est un indice de mélange universel dans le sens où elle est parfaitement définie par ces coefficients sur les B-splines et son intervalle de validité ; cependant, seules les coupes ayant des valeurs situées dans l'intervalle de validité peuvent être évaluées ; ces coupes sont en général des coupes larges.

Enfin, Les approches étudiées dans ce chapitre, pour la détermination de l'indice de mélange, sont applicables sur des données concernant un seul brut ou un ensemble de bruts.

*Troisième Partie : Géométrie des  
Ensembles de Mélange*

CHAPITRE VI

**Convexe de mélanges et  
opérations géométriques**



## 6.0 Introduction

Un système physico-chimique est le contenu matériel et énergétique d'un volume borné dans notre espace euclidien. Pour caractériser un tel système on procède par une division en sous-systèmes ; cela revient à définir une partition du volume du système en régions par une famille de volumes disjoints. Notons  $\Sigma$  un système physico-chimique de volume  $\Omega$ , alors, une partition de  $\Sigma$  en  $q$  volumes ( $\Omega_1, \Omega_2, \dots, \Omega_q$ ) ( $q \geq 2$ ) vérifie :

$$\Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_q = \Omega \quad \text{et} \quad \Omega_i \cap \Omega_j = \emptyset \quad (i \neq j)$$

Nous supposons que le contenu d'un système est constitué d'une famille finie de produits ou d'espèces ; soit  $n$  leur nombre. L'analyse de chaque espèce permet d'évaluer sa quantité dans chaque région, et donc, d'associer à chacune de ces régions un scalaire non négatif. De plus, cette analyse doit répondre aux conditions d'extensivité suivantes :

- Le résultat d'analyse de l'union de deux régions est égal à la somme des résultats d'analyse de deux régions.
- Le résultat d'analyse d'une fraction  $\lambda$  de volume d'une région est égal à  $\lambda$  fois le résultat d'analyse de cette région.

L'ensemble de ces analyses permet d'attribuer à chaque région  $n$  scalaires non négatifs, ce qui constitue un vecteur à  $n$  composantes que l'on appelle vecteur quantité ; chaque composante mesure la quantité d'une espèce dans cette région. Prenons un exemple relatif aux gazoles. Si l'on considère comme spécifications la masse, le % soufre, le point de trouble et la température de filtrabilité, alors, un vecteur quantité est composé des quatre analyses extensives suivantes :

masse
quantité de soufre
quantité d'indice de trouble
quantité d'indice de filtrabilité

L'ensemble de vecteurs quantités est inclus dans un espace vectoriel de dimension inférieure ou égale à  $n$ .

### 6.1 Notion de convexe de mélanges

Pour une partition de  $\Omega$  en  $q$  régions  $\{\Omega_j\}_{j=1}^q$  on dispose de  $q$  vecteurs quantités  $\{N_j\}_{j=1}^q$ .

#### (6.1) Définition

On appelle convexe de mélanges associé à la partition  $(\Omega_1, \Omega_2, \dots, \Omega_q)$  l'ensemble des mélanges faisables par opérations de sous-partitionnement et d'union des régions ayant un sens physique (on suppose que toute région est homogène et que la proportion de chaque vecteur quantité intervenant dans le mélange est définie par le volume de la partie de la région correspondante). L'ensemble de ces mélanges se traduit algébriquement par l'ensemble  $Z$  suivant :

$$(6.2) \quad Z = \left\{ M \in \mathbb{R}^n : M = \sum_{j=1}^q \lambda_j N_j, \lambda_j \in [0,1] \right\}$$

Cet ensemble peut aussi s'écrire comme somme (*de Minkowsky*) des segments :

$$Z_j = [O, N_j] = \{ \lambda \cdot N_j, \lambda \in [0,1] \} \quad j=1, \dots, q.$$

On a donc :

$$(6.3) \quad Z = Z_1 + Z_2 + \dots + Z_q$$

Cet ensemble traduit la capacité de fabrication des mélanges à partir d'une partition du système et fournit une bonne caractérisation de la répartition (séparation) des espèces dans ce système.

Notons la différence entre cet ensemble et les enveloppes convexe et positive de  $\{N_1, N_2, \dots, N_q\}$ . Nous avons en fait les relations d'inclusion (généralement strictes) :

$$\text{Conv}\{N_1, N_2, \dots, N_q\} \subset Z \subset \text{Pos}\{N_1, N_2, \dots, N_q\}$$

L'ensemble  $Z$  est un polytope particulier, connu en géométrie sous le nom de zonotope. Ce terme signifie que  $Z$  est constitué par des zones que l'on définit par l'ensemble des facettes jointes par une arête équipollente à un vecteur de la famille  $\{N_j\}_{j=1}^q$ .

La forme de l'ensemble  $Z$  est représentée par la figure 6.1 sur un exemple simple ( $q=3$ ) concernant des vecteurs quantités à deux composantes conservatives.

Le concept de convexe de mélanges est développé par P.Valentin (cours de génie de la séparation, Ecole Polytechnique) et dans [Valentin 86] il a proposé une modélisation géométrique de la séparation par des zonotopes. Ces derniers sont étudiés par plusieurs auteurs. Nous citons [Coxeter 62] qui développe particulièrement les zonotopes à trois dimensions (zonohedra), [McMullen 71,75] et [Shephard 74a,74b] qui introduisent la notion de zonotope associé.



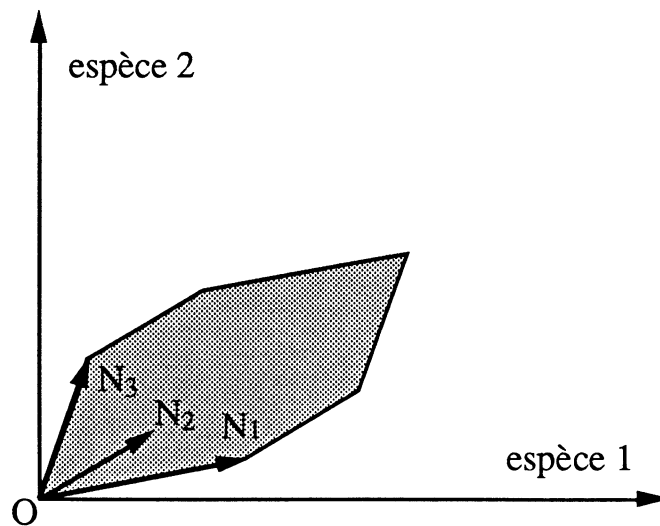


Figure 6.1 : convexe de mélanges binaires ( $n=2$ ) associé  
à la famille  $\{N_j\}_{j=1}^3$

Signalons que des concepts voisins de celui de convexe de mélanges sont utilisés dans d'autres domaines. En cristallographie, des cristaux à trois dimensions ont la même structure que l'ensemble  $Z$ . En automatique, l'ensemble des états atteignables a une représentation du même type qu'un convexe de mélanges [Hermes 69]. Plus récemment, en économie, l'ensemble de production d'un groupe d'entreprises a été modélisé par un convexe analogue [Hildenbrand 81].

## 6.2 Propriétés relatives au convexe de mélanges

Les convexes de mélanges possèdent toutes les propriétés concernant les polyèdres convexes et en particulier celles des zonotopes. Nous citons ici les plus intéressantes, parmi ces propriétés, celles qui sont utiles pour la fabrication des mélanges ou celles qui pourraient faciliter la représentation graphique de ces convexes. Pour les démonstrations de ces propriétés on peut se référer à l'article de synthèse de [Bolker 69].

Proposition 6.1

Le convexe de mélanges est un ensemble convexe, compact de  $\mathbb{R}^n$  à symétrie centrale.

La convexité de l'ensemble  $Z$  signifie que si  $M_1$  et  $M_2$  sont deux mélanges faisables, c'est-à-dire :

$$M_1 = \lambda_1.N_1 + \dots + \lambda_q.N_q \quad \lambda_j \in [0,1], j=1, \dots, q$$

$$M_2 = \mu_1.N_1 + \dots + \mu_q.N_q \quad \mu_j \in [0,1], j=1, \dots, q$$

alors, le mélange  $\theta.M_1 + (1-\theta).M_2$  avec  $\theta \in [0,1]$ , peut être fait avec  $[\theta.\lambda_1 + (1-\theta).\mu_1].N_1, \dots, [\theta.\lambda_q + (1-\theta).\mu_q].N_q$ .

La symétrie centrale est la symétrie par rapport au point  $I = \frac{1}{2} \sum_{j=1}^q N_j$ .

Si  $M \in Z$  alors le point  $M^*$  tel que  $M - I = I - M^*$  représente un mélange faisable et appartient à l'ensemble  $Z$ , ceci découle du fait que  $M^* = 2I - M = \sum_{j=1}^q N_j - M$  s'obtient en mélangeant tout ce qui n'a pas été consommé pour produire  $M$ . D'autre part, la symétrie centrale d'un polyèdre convexe, compact de  $\mathbb{R}^n$  ne suffit pas (en général) pour montrer qu'il est un convexe de mélanges (ou zonotope). Par contre, il est facile de voir qu'en dimension 2 tout polyèdre convexe, compact à symétrie centrale est un zonotope, c'est-à-dire qu'il existe un entier  $q$  et une famille de vecteurs  $\{N_j\}_{j=1}^q$  de  $\mathbb{R}^2$  telle que ce polyèdre coïncide, à une translation près, avec le zonotope  $Z$  associée à cette famille. En dimension 3, Alexandrov dans [Alexandrov 33] a caractérisé les zonotopes comme des polyèdres convexes dont les faces sont à symétrie centrale (donc des zonotopes de dimension 2).

Rappelons que l'on appelle face d'un polyèdre convexe toute intersection de ce polyèdre avec un de ses plans support ; la dimension

polyèdre convexe, les faces de dimension 0 sont les sommets, celles de dimension 1 les arêtes, etc.). En dimension quelconque, mais finie, le théorème suivant caractérise l'ensemble des zonotopes.

### Théorème 6.1

Etant donné un polyèdre  $Z$  convexe, compact de  $\mathbb{R}^n$ . Alors, les assertions suivantes sont équivalentes :

- (i)  $Z$  est un zonotope.
- (ii)  $Z$  est l'image linéaire du cube unité dans  $\mathbb{R}^q$  pour un certain  $q$ .
- (iii) Toute face de dimension  $n-1$  de  $Z$  est un zonotope ( $n \geq 3$ ).
- (iv) Toute face de dimension 2 de  $Z$  est à symétrie centrale.

Une manière simple pour décrire un polyèdre convexe consiste à décrire l'ensemble de ses sommets. Or, on voit facilement que le convexe de mélange  $Z$  est l'enveloppe convexe des  $2^q$  points suivants :

$$(6.4) \quad \sum_{j=1}^q \lambda_j \cdot N_j, \quad \lambda_j \in \{0,1\}$$

On définit un sommet  $s$ , d'un polyèdre convexe  $Z$ , comme étant un point extrême du  $Z$  ; c'est-à-dire : si  $x, y \in Z$  et  $0 < \lambda < 1$  tels que  $s = \lambda x + (1-\lambda) y$  alors, on a :  $s = x = y$ .

Un sommet est alors un point de la frontière tel que toute droite passant par ce point ne recoupe pas le polyèdre de part et d'autre de ce point. Les sommets d'un zonotope  $Z$  sont les faces de dimension 0 et sont de la forme (6.4). Cependant, les points définis par (6.4) peuvent être classés en trois groupes non disjoints : sommets, points frontières et points intérieurs (les sommets sont des points frontières).

D'autre part, il est facile de voir que l'ensemble  $Z$  est l'intersection d'un nombre fini de demi-espaces fermés de  $\mathbb{R}^n$ . Pour cela, il suffit de prendre une sous-famille de  $(n-1)$  vecteurs  $N_j$  indépendants et de considérer l'hyperplan contenant ces vecteurs et puis de déterminer l'hyperplan d'appui de  $Z$  parallèle à ce dernier. Le convexe  $Z$  est alors contenu dans un demi-espace défini par cet hyperplan. L'intersection de  $Z$  avec cet hyperplan d'appui n'est pas vide ; elle est appelée une facette du convexe  $Z$ .

### 6.3 Opérations géométriques sur le convexe de mélanges

Plusieurs opérations géométriques sur le convexe de mélanges ont leurs interprétations physiques relatives au système correspondant. Ces opérations sont, essentiellement, des opérations d'inclusions, d'intersections, et de projections.

L'appartenance d'un point  $M$  au convexe de mélanges  $Z$  signifie que le mélange représenté par ce point est faisable à partir des mélanges de base représentés par la famille  $\{N_j\}_{j=1}^q$ .

L'intersection d'un convexe de mélanges  $Z$  avec un ensemble quelconque de  $\mathbb{R}^n$  correspond à un ensemble de mélanges répondant à une ou plusieurs contraintes. Il est clair que l'intersection ne conserve pas la propriété de symétrie.

Le problème d'intersection d'une droite avec le convexe de mélanges est d'un intérêt important, en pratique ceci peut être utile pour la détermination de :

- (i) la quantité maximale d'un mélange de composition donnée, ceci revient à chercher l'intersection de  $Z$  avec la demi-droite issue de l'origine et correspondant à cette composition.

- (ii) la quantité maximale d'une espèce en sachant les quantités des (n-1) espèces restantes.

L'intersection de convexe de mélanges avec un plan fournit un moyen de visualisation du convexe  $Z$  sur un écran en coupant  $Z$  par des plans parallèles. Aussi, l'intersection de  $Z$  avec un hyperplan donne l'ensemble des mélanges ayant une quantité donnée formée à partir d'une combinaison linéaire des espèces de base.

Enfin, la projection du convexe de mélanges sur un sous-espace de  $\mathbb{R}^n$  parallèlement à un sous-espace complémentaire, est aussi un convexe de mélanges engendré par la famille des vecteurs projections des  $\{N_j\}_{j=1}^q$ .

#### 6.4 Intersection d'un zonotope avec un hyperplan dans $\mathbb{R}^3$

Dans ce paragraphe, nous développons un algorithme rapide et efficace permettant la détermination de l'intersection d'un zonotope avec un hyperplan dans notre espace euclidien usuel. La démarche suivie est tout à fait générale et par conséquent, elle est valable pour une dimension ( $n > 3$ ). Cependant, la technique nécessaire pour décrire l'ensemble d'intersection avec  $n=3$ , qui est un polygone convexe dans ce cas, est plus simple que pour celui avec  $n > 3$ , qui est en général un polyèdre convexe de  $\mathbb{R}^{n-1}$ .

Pour ce faire, nous considérons d'abord le problème de l'hyperplan d'appui à un zonotope. Soit  $C \in \mathbb{R}^3$ , on cherche un scalaire  $d \in \mathbb{R}$  tel que l'hyperplan :

$$H = \{ X \in \mathbb{R}^3 / \langle C, X \rangle = d \}$$

soit d'appui au convexe  $Z$  défini par (6.2), c'est-à-dire :

$$\begin{cases} \langle C, X \rangle \geq d & \forall X \in Z \\ H \cap Z \neq \emptyset \end{cases}$$

Il est clair que l'on doit prendre :

$$(6.5) \quad d = \min_{X \in Z} \langle C, X \rangle$$

et que  $H \cap Z$  est l'ensemble des solutions de ce problème de minimisation. On a encore

$$(6.6) \quad d = \min_{\substack{\lambda_j \in [0,1] \\ j=1, \dots, q}} \sum_{j=1}^q \lambda_j \langle C, N_j \rangle$$

Si on pose :

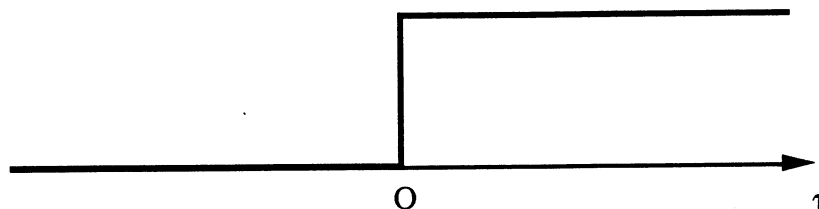
$$(6.7) \quad \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_q \end{pmatrix} \quad \text{avec} \quad \gamma_j = \langle C, N_j \rangle \quad j=1, \dots, q.$$

Alors, l'ensemble des solutions de (6.6) est le suivant :

$$(6.8) \quad H \cap Z = \left\{ X = \sum_{j=1}^q \lambda_j N_j / \begin{array}{ll} \lambda_j = 0 & \text{si } \gamma_j > 0 \\ \lambda_j = 1 & \text{si } \gamma_j < 0 \\ \lambda_j \in [0,1] & \text{si } \gamma_j = 0 \end{array} \right\}$$

On considère la multi-application  $\mathfrak{Z}$  (pour échelon) de  $\mathbb{R}$  dans  $\mathbb{R}$  définie par :

$$\mathfrak{Z}(\tau) = \begin{pmatrix} \{0\} & \text{si } \tau < 0 \\ \{0,1\} & \text{si } \tau = 0 \\ \{1\} & \text{si } \tau > 0 \end{pmatrix}$$

Figure 6.2 : graphe de la multi-application  $\xi$ 

La figure (6.2) représente le graphe de cette multi-application. Alors, on peut écrire :

$$(6.9) \quad H \cap Z = \left\{ X = \sum_{j=1}^q \lambda_j N_j / \lambda_j \in \xi(-\gamma_j), j=1, \dots, q \right\}$$

Notons :

$$J_1 = \{j / \gamma_j > 0\}$$

$$J_2 = \{j / \gamma_j < 0\}$$

$$J_3 = \{j / \gamma_j = 0\}$$

On a alors :

$$d = \sum_{j \in J_2} \gamma_j$$

Maintenant on considère un plan de  $\mathbb{R}^3$  :

$$(6.10) \quad \mathcal{P} = \{X \in \mathbb{R}^3 / \langle A, X \rangle = b\}$$

avec  $A \in \mathbb{R}^3$  et  $b \in \mathbb{R}$ . Il est facile de déterminer si  $\mathcal{P}$  coupe  $Z$  ou non. En effet, par ce qui précède, on peut calculer :

$$(6.11) \quad \begin{cases} b_{\min} = \min_{X \in Z} \langle A, X \rangle \\ b_{\max} = \max_{X \in Z} \langle A, X \rangle \end{cases}$$

Alors,  $\mathcal{P}$  coupe  $Z$  si et seulement si on a :

$$(6.12) \quad \boxed{b_{\min} \leq b \leq b_{\max}}$$

Si l'on note :

$$(6.13) \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_q \end{pmatrix} \quad \text{avec} \quad \mu_j = \langle A, N_j \rangle \quad j=1, \dots, q.$$

et

$$K_1 = \{j / \mu_j > 0\}$$

$$K_2 = \{j / \mu_j < 0\}$$

$$K_3 = \{j / \mu_j = 0\}$$

On a alors :

$$(6.14) \quad \begin{cases} b_{\min} = \sum_{j \in K_2} \mu_j \\ b_{\max} = \sum_{j \in K_1} \mu_j \end{cases}$$

En pratique on pourra :

- soit se donner un plan  $\mathcal{P}$  et vérifier que  $\mathcal{P}$  coupe  $Z$ .
- soit se donner une direction  $A$  et calculer  $b_{\min}$  et  $b_{\max}$  et choisir  $r$  niveaux intermédiaires  $b_i$  ( $i=1, \dots, r$ ). On calculera alors les  $r$  intersections  $\mathcal{P}_i \cap Z$ , où  $\mathcal{P}_i$  est le plan associé à  $A$  et  $b_i$ .



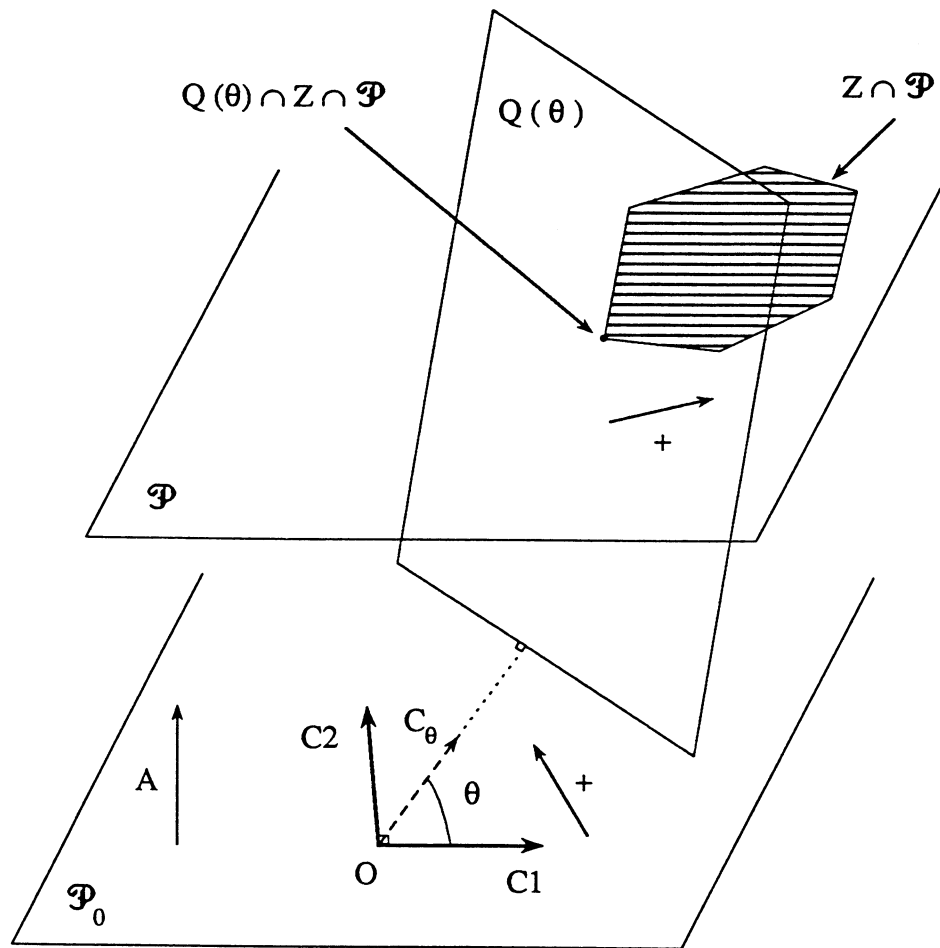


Figure 6.3 : illustration géométrique de la détermination du  $Z \cap \mathcal{P}$  en utilisant les plans d'appui  $Q(\theta)$

Pour déterminer  $Z \cap \mathcal{P}$  on va procéder à nouveau par plan d'appui. Soit  $\mathcal{P}_0 = \{X \in \mathbb{R}^3 / \langle A, X \rangle = 0\}$  sous espace de dimension 2 de  $\mathbb{R}^3$  (plan parallèle à  $\mathcal{P}$  et passant par l'origine). On note  $C_1$  et  $C_2$  deux vecteurs orthonormés de  $\mathcal{P}_0$ . Pour déterminer les plans d'appui de  $Z \cap \mathcal{P}$  on peut se contenter de prendre des vecteurs directions de la forme :

$$(6.15) \quad C_\theta = \cos \theta \cdot C_1 + \sin \theta \cdot C_2$$

C'est-à-dire que  $C_\theta \in \mathcal{P}_0$  et que  $\|C_\theta\|_2 = 1$  ; ainsi le plan  $\langle C_\theta, X \rangle = d$  est perpendiculaire à  $\mathcal{P}_0$  (voir figure 6.3). On aura alors une famille de plans d'appui à  $Z \cap \mathcal{P}$  paramétrés par le seul paramètre  $\theta \in [0, 2\pi]$  et on pourra résoudre le problème :

$$(P.6.1) \quad \begin{cases} d(\theta) = \min \langle C_\theta, X \rangle \\ X \in Z \cap \mathcal{P} \end{cases}$$

et déterminer l'ensemble  $(Z \cap \mathcal{P}) \cap Q(\theta)$  où

$$(6.16) \quad Q_\theta = \left\{ X \in \mathbb{R}^3 / \langle C_\theta, X \rangle = d(\theta) \right\}$$

On remarque que  $C_1$  et  $C_2$  et  $A$  sont orthogonaux. De plus  $C_1$  et  $C_2$  sont de norme 1 ; on n'a pas supposé que  $A$  était de norme égale à 1. Donc, un  $X \in \mathbb{R}^3$  donné peut être représenté dans ce repère :

$$X = \xi_1.C_1 + \xi_2.C_2 + \xi_3.A$$

avec

$$\xi_1 = \langle X, C_1 \rangle, \xi_2 = \langle X, C_2 \rangle, \xi_3 = \frac{\langle X, A \rangle}{\langle A, A \rangle}$$

Un tel  $X$  appartient à  $\mathcal{P}$  si et seulement si  $\xi_3 \langle A, A \rangle = b$ .

Le problème (P.6.1) peut s'écrire de façon équivalente :

$$(P'.6.1) \quad \begin{cases} d(\theta) = \min (\lambda, \gamma) \\ \lambda_j \in [0, 1], j=1, \dots, q \\ (\lambda, \mu) = b \end{cases}$$

où  $(\cdot, \cdot)$  dénote le produit scalaire dans  $\mathbb{R}^q$  et  $\gamma, \mu \in \mathbb{R}^q$  définis par (6.7) et (6.13) respectivement.

On remarque que  $\mu$  est fixe (pour un plan  $\mathcal{P}$  donné) alors que  $\gamma$  va

varier, puisque la direction  $C_\theta$  est associée à l'angle  $\theta$  qui variera de 0 à  $2\pi$ .

Le problème (P'.6.1) ne se décompose plus composante à composante (à cause de la contrainte  $(\lambda, \mu) = b$ ). Pour détourner cette difficulté on introduit le  $\mathfrak{L}$ agrangien :

$$(6.17) \quad \mathfrak{L}(\lambda, k) = (\lambda, \gamma) + k \cdot [(\lambda, \mu) - b], \quad \begin{cases} \lambda \in [0,1]^q \\ k \in \mathbb{R} \end{cases}$$

Alors, le problème précédent peut s'écrire :

$$d(\theta) = \min_{\lambda \in [0,1]^q} \max_{k \in \mathbb{R}} \mathfrak{L}(\lambda, k) \quad (\text{Primal})$$

Car, on a :

$$\max_{k \in \mathbb{R}} \mathfrak{L}(\lambda, k) = \begin{cases} (\lambda, \gamma) & \text{si } (\lambda, \mu) = b \\ +\infty & \text{sinon} \end{cases}$$

Et le problème dual associé est le suivant :

$$d'(\theta) = \max_{k \in \mathbb{R}} \min_{\lambda \in [0,1]^q} \mathfrak{L}(\lambda, k) \quad (\text{Dual})$$

Ce problème se décompose ...

Ce  $\mathfrak{L}$ agrangien correspond à la perturbation suivante du problème primal :

$$\begin{aligned}
 h(u) &= \min_{\lambda \in [0,1]^q} (\lambda, \gamma) \\
 (\lambda, \mu) &= b + u
 \end{aligned}
 \qquad \mu \in \mathbb{R}$$

Si l'on fait l'hypothèse (6.18) que le plan  $\mathcal{P}$  coupe l'intérieur du zonotope  $Z$ , alors le problème primal est stable (i.e.  $h$  est continue en 0). Cela revient à dire que l'on suppose :

$$(6.18) \quad \begin{cases} Z \text{ a un intérieur non vide (les } N_j \text{ ne sont pas coplanaires)} \\ b_{\min} < b < b_{\max} \end{cases}$$

Il est clair que c'est bien le cas auquel on s'intéresse en pratique. Cette hypothèse entraîne que le problème dual a des solutions et que l'on a  $d(\theta) = d'(\theta)$ .

### Remarque 6.1

Il est probable que par d'autres arguments (programmation linéaire) on pourrait déduire que  $d(\theta) = d'(\theta)$ , même si l'on ne suppose que  $b_{\min} < b < b_{\max}$ .

La dualité peut être utilisée de deux façons (qui ne sont pas indépendantes) :

- Première façon : Appelons  $S$  l'ensemble des solutions du problème primal et  $S(k)$  l'ensemble des solutions du sous-problème :

$$\begin{aligned}
 \min \quad & \mathcal{L}(\lambda, k) \\
 \lambda \in & [0,1]^q
 \end{aligned}$$

Comme on a :

$$\mathfrak{L}(\lambda, k) = (\lambda, \gamma + k \cdot \mu) - k \cdot b$$

Alors,  $\lambda = \lambda(k)$  est solution de ce sous-problème si et seulement si  $\lambda_j \in \mathfrak{X}(-(\gamma_j + k \cdot \mu_j))$ ,  $j=1, \dots, q$  et le montant de ce sous-problème (qui n'est autre que la fonction objective du problème dual) est donné par :

$$\Phi(k) = \sum_{j \in J_2(k)} (\gamma_j + k \cdot \mu_j) - k \cdot b$$

où

$$J_2(k) = \{ j / \gamma_j + k \cdot \mu_j < 0 \}$$

Alors, on a le résultat suivant : l'hypothèse (6.18) étant vérifiée, si  $\bar{k}$  est solution du problème Dual (i.e. maximise  $\Phi(k)$ ) alors  $S \subset S(\bar{k})$ .

• Deuxième façon : Nous avons toujours (même sans l'hypothèse (6.18)):

$$(6.19) \quad \left. \begin{array}{l} (\bar{\lambda}, \bar{k}) \text{ point de} \\ \text{selle de } \mathfrak{L} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \bar{\lambda} \text{ solution du Primal} \\ \bar{k} \text{ solution du Dual} \\ d(\theta) = d'(\theta) \end{array} \right.$$

Or, on a par définition,  $(\bar{\lambda}, \bar{k})$  est un point de selle de  $\mathfrak{L}$ agrangien  $\mathfrak{L}$  si et seulement si :

$$(6.20) \quad \begin{array}{ccc} \mathfrak{L}(\bar{\lambda}, k) \leq \mathfrak{L}(\bar{\lambda}, \bar{k}) \leq \mathfrak{L}(\lambda, \bar{k}) \\ \forall k \in \mathbb{R} \quad \uparrow \quad \quad \quad \uparrow \quad \forall \lambda \in [0,1]^q \\ \text{(I)} \quad \quad \quad \text{(II)} \end{array}$$

L'inégalité (I) est équivalente à :

$$(I') \quad (\bar{\lambda}, \mu) = b$$

et l'inégalité (II) est équivalente à :

$$(II') \quad \bar{\lambda}_j \in \mathfrak{X} (-(\gamma_j + \bar{k} \cdot \mu_j)), j=1, \dots, q.$$

En réunissant (I') et (II') on obtient la multi-équation :

$$(6.21) \quad \boxed{b \in \sum_{j=1}^q \mu_j \cdot \mathfrak{X} (-(\gamma_j + \bar{k} \cdot \mu_j))}$$

On pose

$$(6.22) \quad G(k) = \sum_{j=1}^q \mu_j \mathfrak{X} (-(\gamma_j + k \cdot \mu_j))$$

Il faut déterminer  $\bar{k}$  tel que l'on ait  $b \in G(\bar{k})$ . En général  $G(k)$  est réduit à un point, sauf lorsque pour l'un des indices  $j$  on a  $\gamma_j + k \cdot \mu_j = 0$ , auquel cas on a un intervalle.

Rappelons que les  $\mu_j$  sont fixés lorsque  $\mathfrak{P}$  est donné. Par contre  $\gamma_j$  va varier en fonction de  $\theta$ . On a déjà noté  $K_1 = \{j / \mu_j > 0\}$ ,  $K_2 = \{j / \mu_j < 0\}$ ,  $K_3 = \{j / \mu_j = 0\}$ .

Pour mieux comprendre, prenons un exemple simple ( $q=5$ ). Dans cet exemple, on suppose  $\mu_3 = 0$ ,  $\mu_2$  et  $\mu_5 > 0$ ;  $\mu_1$  et  $\mu_4 < 0$ . On trace le graphe des cinq droites  $D_j : k \rightarrow \gamma_j + k \cdot \mu_j$ ,  $j=1, \dots, 5$ , ainsi que le graphe de la multi-application  $G$  (voir figure 6.4). Lorsque  $j \in K_1 \cup K_2$ , on a un zéro pour  $D_j$  donné par  $K_j = -\frac{\gamma_j}{\mu_j}$ .

Le graphe de  $G$  est donné en trait gras, et on a pour :

$$\begin{array}{ll} k < k_2, & \text{on a} \quad G(k) = \{\mu_2 + \mu_5\} = \{b_{\max}\} \quad (\text{voir ci-dessous}) \\ k \in ]k_2, k_4[ & \Rightarrow \quad G(k) = \{\mu_5\} \\ k \in ]k_4, k_5[ & \Rightarrow \quad G(k) = \{\mu_5 + \mu_4\} \end{array}$$

$$\begin{aligned}
 k \in ]k_5, k_1[ &\Rightarrow G(k) = \{\mu_4\} \\
 k > k_1, &\text{ on a } G(k) = \{\mu_1 + \mu_4\} = \{b_{\min}\}
 \end{aligned}$$

et lorsque :

$$\begin{aligned}
 k = k_2 &: G(k) = [\mu_5, \mu_2 + \mu_5] \\
 k = k_4 &: G(k) = [\mu_5 + \mu_4, \mu_5] \\
 k = k_5 &: G(k) = [\mu_4, \mu_5 + \mu_4] \\
 k = k_1 &: G(k) = [\mu_1 + \mu_4, \mu_4]
 \end{aligned}$$

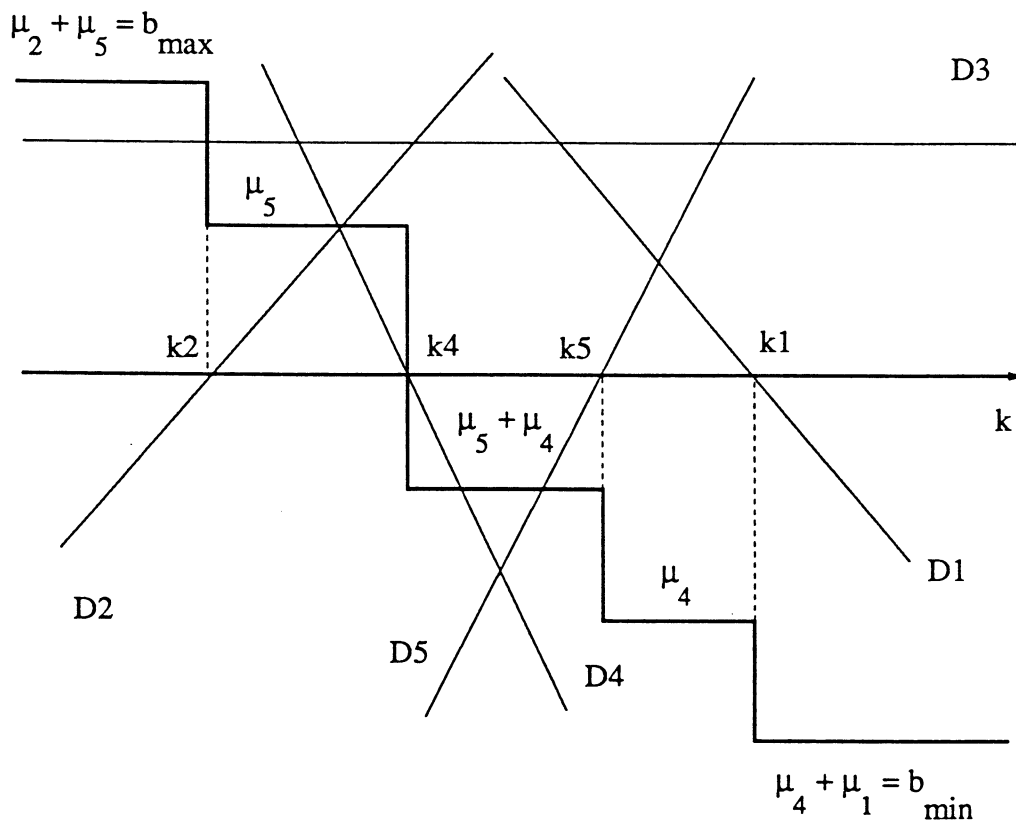


Figure 6.4 : graphe de la multi-application  $G$  et les droites  $D_j, j=1, \dots, 5$ .

On voit que la multi-application  $G$  est non croissante (décroissante au sens large). En fait, elle est constante par morceaux, sauf en un nombre fini de points où elle vaut un intervalle (entre la valeur à gauche et la valeur à

droite).

Si  $b \in [b_{\min}, b_{\max}]$ , on a toujours une solution. En général, la solution  $\bar{k}$  tel que  $b \in G(\bar{k})$  est unique, sauf pour un nombre fini de valeurs, par exemple si  $b = \mu_5$ , on aura  $[k_2, k_4]$  comme solutions en  $k$  ; mais, il faut remarquer que la solution correspondante en  $\lambda$  est unique même dans ce cas. Par exemple, dans l'exemple présent :

- si l'on choisit  $\bar{k} = k_4$ , tous les  $\bar{\lambda}_j$  sauf pour  $j = 4$  sont déterminés par  $\bar{\lambda}_j \in \mathfrak{X}(-(\gamma_j + \bar{k} \cdot \mu_j))$ ,  $\bar{\lambda}_1 = 0$ ,  $\bar{\lambda}_2 = 0$ ,  $\bar{\lambda}_3 = 0$  et  $\bar{\lambda}_5 = 1$  et l'on doit avoir  $\bar{\lambda}_4 = 0$  pour vérifier alors  $(\bar{\lambda}, \mu) = b$ .
- si l'on prend  $\bar{k} = k_2$ , tous les  $\bar{\lambda}_j$  sauf pour  $j=2$  sont déterminés par (II'), et l'on doit prendre  $\bar{\lambda}_2 = 0$  pour avoir  $(\bar{\lambda}, \mu) = b$ .
- si l'on prend  $\bar{k} \in ]k_2, k_4[$ , tous les  $\bar{\lambda}_j$  sont donnés par (II'), et on a toujours le même résultat.

Donc, on peut prendre pour  $\bar{k}$  l'une des extrémités ; par conséquent,  $\bar{k}$  est toujours l'un des  $k_j$ .

L'algorithme consiste à découvrir les sommets (et les arêtes) de  $Z \cap \mathfrak{P}$  en faisant varier  $\theta$ . Pour un  $\theta$  donné, on suppose que l'on ait déterminé  $\bar{k} = k_{\bar{j}}$ ,  $\bar{j} \in K_1 \cup K_2$  tel que  $b \in G(\bar{k})$ . Alors, deux cas se présentent:

**cas I :**

$$\forall j \neq \bar{j} \quad \text{on a } \gamma_j + \bar{k} \cdot \mu_j \neq 0$$

Dans un tel cas, l'ensemble  $Z \cap \mathfrak{P} \cap Q(\theta)$  est réduit à un point qui est l'un des sommets du polygone  $Z \cap \mathfrak{P}$  (voir figure 6.5).



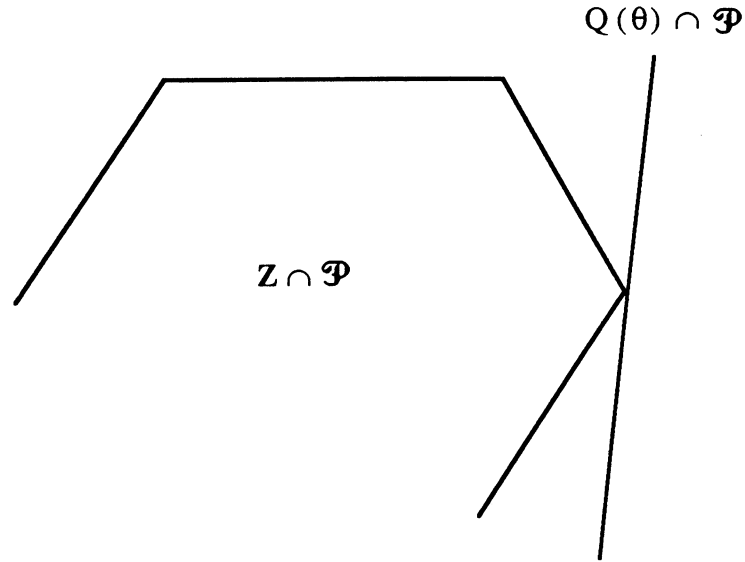


Figure 6.5 :  $Q(\theta) \cap \mathcal{P}$  est une droite d'appui à  $Z \cap \mathcal{P}$ . On remarque que  $Z \cap Q(\theta) \cap \mathcal{P}$  est réduit à un sommet de  $Z \cap \mathcal{P}$ .

Ce sommet peut être déterminé explicitement par les relations :

$$(6.23) \quad \begin{cases} \bar{\lambda}_j \in \mathfrak{X}(-(\gamma_j + \bar{k} \cdot \mu_j)) & j \neq \bar{j} \\ \bar{\lambda}_{\bar{j}} = \frac{b - \sum_{j \neq \bar{j}} \bar{\lambda}_j \mu_j}{\mu_{\bar{j}}} \end{cases}$$

**cas II :**

$$\exists j_1, \dots, j_r \neq \bar{j} \text{ tels que } \gamma_{j_i} + \bar{k} \cdot \mu_{j_i} = 0 \quad (i=1, \dots, r)$$

Tout d'abord, on remarque que les vecteurs  $N_{j_1}, N_{j_2}, \dots, N_{j_r}; N_{\bar{j}}$  sont coplanaires ; ceci découle du fait que :

$$(6.24) \quad \gamma_j + \bar{k} \cdot \mu_j = 0 \Leftrightarrow \langle C_\theta + \bar{k} \cdot A, N_j \rangle = 0$$

Alors, pour tout  $j \in \{j_1, j_2, \dots, j_r; \bar{j}\}$  on détermine  $\bar{\lambda}_j$  en utilisant la

relation :

$$(6.25) \quad \bar{\lambda}_j \in \mathfrak{X}(-(\gamma_j + \bar{k} \cdot \mu_j))$$

Et pour l'ensemble  $\{j_1, j_2, \dots, j_r; \bar{j}\}$  on a l'équation :

$$(6.26) \quad \bar{\lambda}_{j_1} \cdot \mu_{j_1} + \dots + \bar{\lambda}_{j_r} \cdot \mu_{j_r} + \bar{\lambda}_{\bar{j}} \cdot \mu_{\bar{j}} = b - \sum_{j \notin \{j_1, j_2, \dots, j_r; \bar{j}\}} \bar{\lambda}_j \cdot \mu_j$$

Alors, si  $\{j_1, j_2, \dots, j_r\} \cap K_3 \neq \emptyset$  alors, il existe  $i \in \{1, \dots, r\}$  tel que  $\mu_{j_i} = \gamma_{j_i} = 0$ , d'où :

$$\bar{\lambda}_{j_i} \in \mathfrak{X}(0) = [0,1]$$

Il s'agit donc d'une arête de  $Z \cap \mathfrak{P}$ . Par contre, si l'on a  $\mu_{j_i} \neq 0$ , ( $i=1, \dots, r$ ) alors, on peut noter :

$$(6.27) \quad b' = b - \sum_{j \notin \{j_1, j_2, \dots, j_r; \bar{j}\}} \bar{\lambda}_j \cdot \mu_j$$

Alors, l'équation (6.26) s'écrit

$$(6.28) \quad \mu_{j_1} \cdot \bar{\lambda}_{j_1} + \dots + \mu_{j_r} \cdot \bar{\lambda}_{j_r} + \mu_{\bar{j}} \cdot \bar{\lambda}_{\bar{j}} = b'$$

Cette équation représente un hyperplan dans  $\mathbb{R}^{r+1}$  que l'on note  $H^{r+1}$  ; alors l'ensemble de solutions  $\{\bar{\lambda}_j / j = j_1, j_2, \dots, j_r; \bar{j}\}$  est égal à

$$(6.29) \quad H^{r+1} \cap [0,1]^{r+1}$$

On pose :

$$(6.30) \quad \left\{ \begin{array}{l} b'_{\min} = \sum_{\substack{j \in \{j_1, j_2, \dots, j_r; \bar{j}\} \\ j \in K_2}} \mu_j \\ b'_{\max} = \sum_{\substack{j \in \{j_1, j_2, \dots, j_r; \bar{j}\} \\ j \in K_1}} \mu_j \end{array} \right.$$

Dans l'équation (6.28), si  $b' = b'_{\min}$  ou  $b' = b'_{\max}$  alors on a un sommet de  $Z \cap \mathcal{P}$  et l'ensemble (6.29) est réduit à un point qui est l'un des sommets de  $[0,1]^{r+1}$ . On remarque ici, que le point  $\bar{\lambda}$  obtenu a des composantes 0 ou 1 et par conséquent, c'est l'un des points (6.4). Lorsque  $b' \in ]b'_{\min}, b'_{\max}[$  l'ensemble (6.29) n'est pas réduit à un point et par conséquent, il s'agit d'une arête de  $Z \cap \mathcal{P}$ .

L'algorithme prend alors la forme suivante :

$$\left\{ \begin{array}{l} \theta \leftarrow 0; \\ i \leftarrow 1; \end{array} \right. \quad \{ \text{initialisation} \}$$

**tantque**  $\theta < 2\pi$  **faire**

**début**

- **si** (il s'agit d'une arête) **alors**

**début**

- ajouter à  $\theta$  une petite variation  $d\theta$  :  
 $\theta \leftarrow \theta + d\theta$ ;
- l'ensemble  $Z \cap \mathcal{P} \cap Q(\theta)$  est réduit à un point (voir figure 6.6), c'est le sommet  $S_i$ .
- $i \leftarrow i + 1$ ;

**fin\_si**

- *calcul du  $\theta$  suivant.*

**fin\_tantque.**

En pratique, on utilisera un  $d\theta$  de l'ordre  $10^{-5}$  degré.

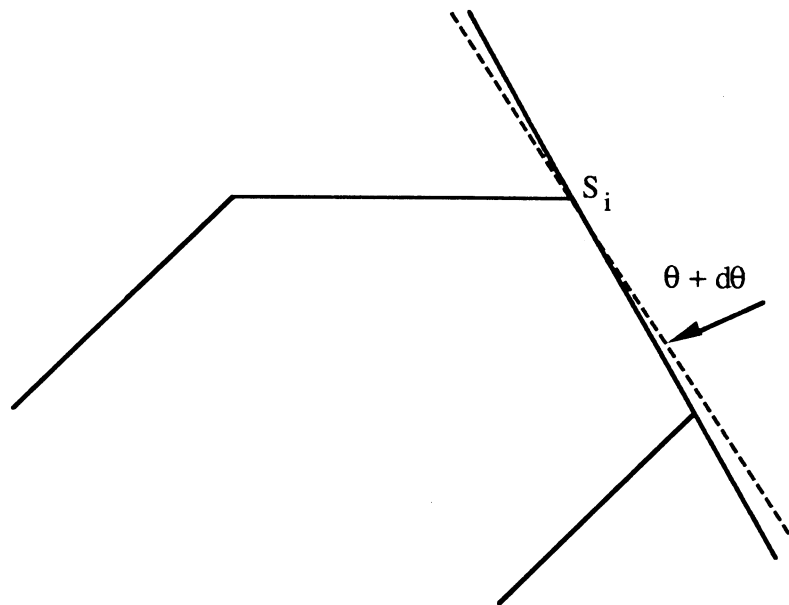


Figure 6.6 : une petite variation  $d\theta$  permet de calculer un nouveau sommet  $S_i$ .

La recherche du  $\theta$  suivant consiste à déterminer la plus petite rotation angulaire  $\Delta\theta$  pour laquelle la racine  $k_{\bar{j}} = \bar{k}$  vient coïncider avec l'une des racines  $k_j$ ,  $j \neq \bar{j}$  ( $j \in K_1 \cup K_2$ ), ou bien l'un des  $\gamma_j$  ( $j \in K_3$ ) devient nul. Le calcul de l'angle de croisement du  $k_{\bar{j}}$  avec une racine  $k_j$ ,  $j \in K_1 \cup K_2$  peut se faire de la façon suivante :

Nous avons :

$$k_j = k_{\bar{j}} = \bar{k} \quad \Leftrightarrow \quad \begin{cases} \gamma_{\bar{j}} + \bar{k} \cdot \mu_{\bar{j}} = 0 \\ \gamma_j + \bar{k} \cdot \mu_j = 0 \end{cases}$$

d'où

$$(i) \quad \frac{\gamma_{\bar{j}}}{\mu_{\bar{j}}} = \frac{\gamma_j}{\mu_j}$$

D'autre part, nous avons pour tout  $j$  :

$$(ii) \quad \gamma_j = \cos\theta \cdot \langle C_1, N_j \rangle + \sin\theta \cdot \langle C_2, N_j \rangle$$

On pose :

$$\eta_{1j} = \langle C_1, N_j \rangle \quad \text{et} \quad \eta_{2j} = \langle C_2, N_j \rangle$$

Alors, en utilisant (i) et (ii) on obtient :

$$(6.31) \quad \boxed{\operatorname{tg} \theta = - \frac{(\mu_{\bar{j}} \cdot \eta_{1j} - \mu_j \cdot \eta_{1\bar{j}})}{(\mu_{\bar{j}} \cdot \eta_{2j} - \mu_j \cdot \eta_{2\bar{j}})}}$$

Cette relation permet de calculer deux angles de croisement  $\theta_j^1$  et  $\theta_j^2 = \theta_j^1 + \pi$  ( $\theta_j^1$  est déterminé dans l'intervalle  $[0, \pi]$ ). De même, pour un  $j \in K_3$ ,  $\gamma_j$  s'annule si et seulement si :

$$\cos \theta \cdot \eta_{1j} + \sin \theta \cdot \eta_{2j} = 0$$

ou encore :

$$(6.32) \quad \boxed{\operatorname{tg} \theta = - \frac{\eta_{1j}}{\eta_{2j}}}$$

Cette relation, comme celle de (6.31), fournit deux valeurs  $\theta_j^1$  et  $\theta_j^2 = \theta_j^1 + \pi$ .

Donc, pour calculer le  $\theta$  suivant, on procède comme suit :

On pose pour tout  $j \neq \bar{j}$  :

$$(6.33) \quad \Delta\theta_j = \begin{cases} \theta_j^1 - \theta & \text{si } \theta_j^1 > \theta \\ \theta_j^2 - \theta & \text{si } \theta_j^1 \leq \theta \leq \theta_j^2 \\ (2\pi - \theta) + \theta_j^1 & \text{si } \theta \geq \theta_j^2 \end{cases}$$

Et on prend :

$$(6.34) \quad \Delta\theta = \min_{j \neq \bar{j}} \Delta\theta_j$$

Alors, le  $\theta$  suivant est égal à  $\theta + \Delta\theta$ .

## 6.5 Quelques exemples d'intersections dans $\mathbb{R}^3$

Dans ce paragraphe, nous présentons quelques exemples d'intersections de zonotopes avec des plans dans  $\mathbb{R}^3$ .

- **Premier Exemple :**

Etant donnée un entier positif non nul  $n$ . On considère l'ensemble de vecteurs :

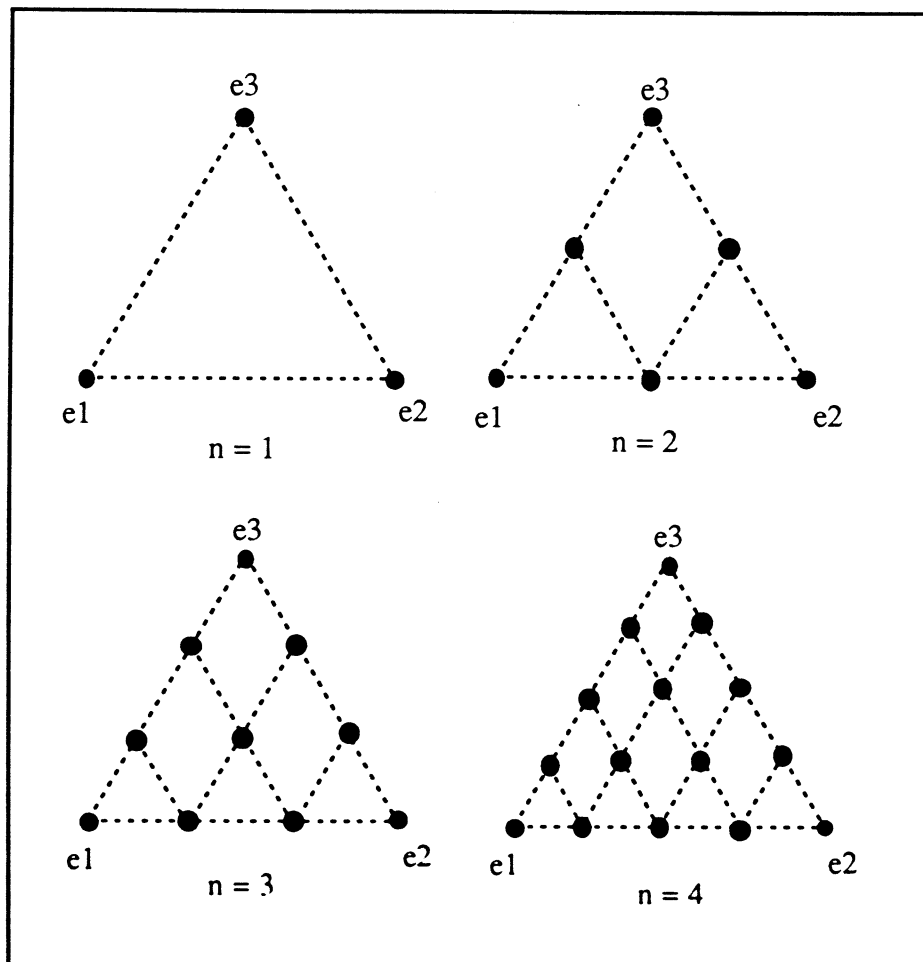
$$(6.35) \quad \left\{ N_j^q \right\}_{j=1}^q = \left\{ \begin{pmatrix} \frac{j_1}{n} \\ \frac{j_2}{n} \\ 1 - \left( \frac{j_1}{n} + \frac{j_2}{n} \right) \end{pmatrix} : j_1=0, \dots, n \text{ et } j_2=0, \dots, n-j_1 \right\}$$

On remarque ici que le nombre de vecteurs  $q$  est égal à  $(n+1)(n+2)/2$  et

que les composantes de chaque vecteur de cette famille sont positives et que leur somme vaut 1. Par conséquent, les points représentés par les vecteurs (6.35) sont situés dans le plan :

$$\left\{ X \in \mathbb{R}^3 : \sum_{i=1}^3 X_i = 1 \right\}.$$

Si on note  $e_1$ ,  $e_2$  et  $e_3$  les points correspondant aux trois vecteurs de la base canonique de  $\mathbb{R}^3$  alors les points correspondants aux vecteurs  $\left\{ N_j \right\}_{j=1}^q$  sont situés dans le triangle représenté par  $e_1$ ,  $e_2$  et  $e_3$  comme le montre les schémas suivants :



Dans la figure (6.7) nous présentons une visualisation des quatre

premiers zonotopes de cette famille. Cette visualisation est obtenue par la projection d'une centaine de coupes parallèles dans la direction  $A = (1,1,1)^t$  sur le plan  $\mathcal{P}_0$  (voir figure 6.3).

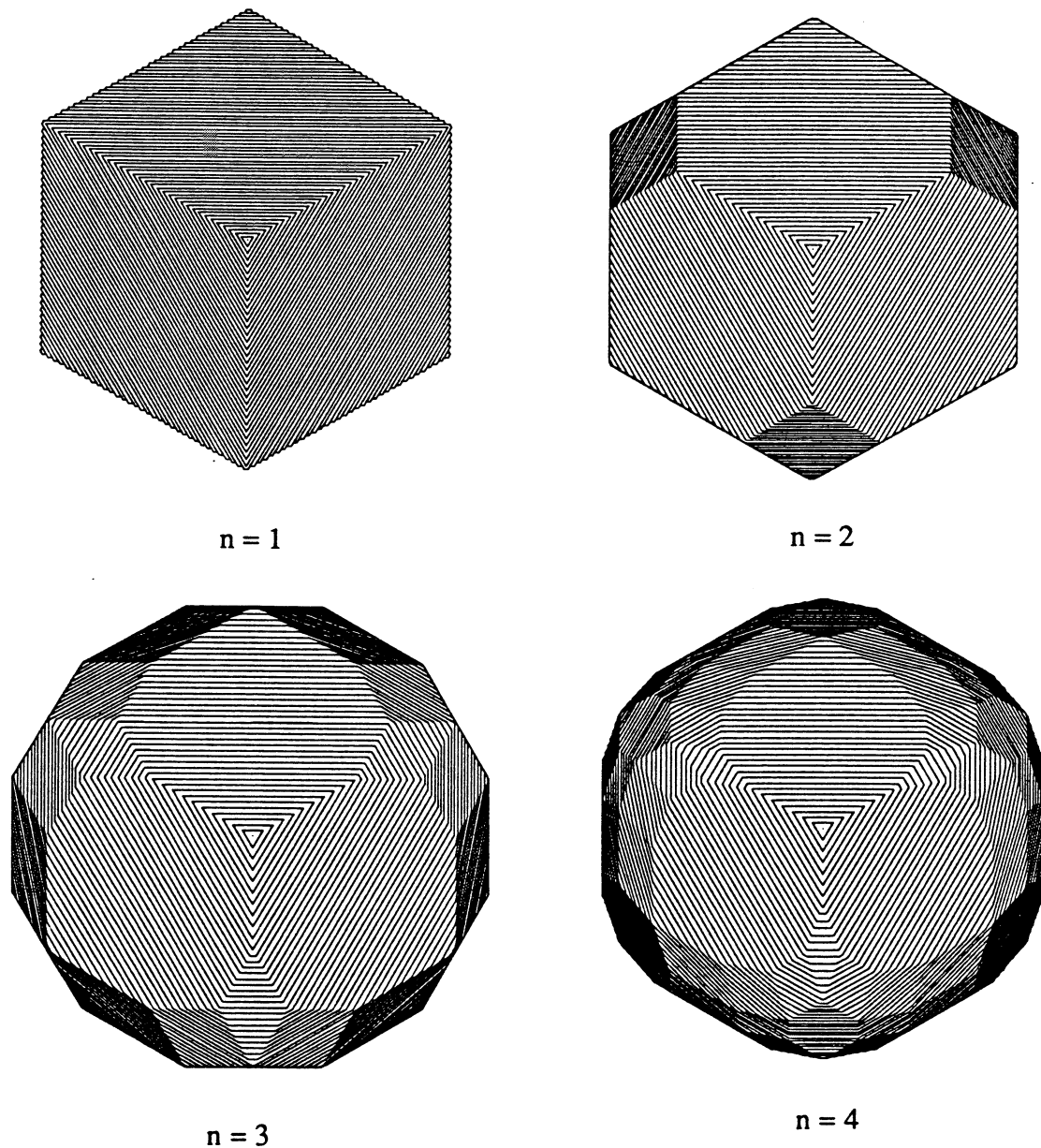


Figure 6.7 : Visualisation par coupes parallèles des quatre premiers zonotopes générés par les vecteurs (6.35).

Pour  $n=10$ , la famille (6.35) comprend 66 vecteurs et le zonotope correspondant est représenté, dans la direction  $A = (1,1,1)^t$ , par la figure



suivante :

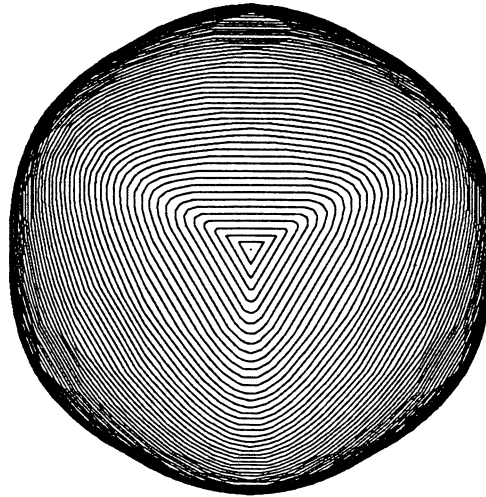


Figure 6.8 : ( $n = 10$ ) Exemple de zonotope généré par 66 vecteurs de base.

Nombre de Sommets par Coupe	Temps (Secondes) par Coupe
3	3.40
6	5.99
12	10.16
13	15.10
18	15.10
19	15.98
24	20.16
36	30.05
42	34.88
43	35.76
60	49.81
61	50.65
67	55.53
72	59.71
78	64.48
79	65.42
84	69.42
85	70.36
90	74.37
91	75.36
96	79.42
97	80.25
109	90.03
126	103.97
127	104.86

n	q	Nombre Moyen de Sommets par Coupe	Temps Moyen (Sec.) par Coupe
1	3	4.0	0.04
2	6	7.0	0.14
3	10	11.2	0.41
4	15	16.4	1.08
5	21	23.1	2.58
6	28	30.5	5.49
7	36	38.9	10.79
8	45	48.4	19.96
9	55	58.8	34.86
10	66	70.4	58.41

Table 6.1 : (a)- Nombre de sommets et temps de calcul par coupe ( $n=10$ ).  
 (b)- Nombre moyen de sommets et temps moyen de calcul par coupe pour les dix premiers zonotopes ( $n=1, \dots, 10$ ).

En ce qui concerne l'étude de la complexité de l'algorithme d'intersection, présenté au § 6.4, nous avons calculé, pour un zonotope donné ( $n$  fixe) et pour une direction donnée, le nombre de sommets et le temps de calcul nécessaire pour chaque coupe intermédiaire. Nous avons constaté que le temps de calcul dépend linéairement du nombre de sommets dans la coupe considérée. Dans la table 6.1-(a) on trouve les résultats obtenus pour  $n=10$ .

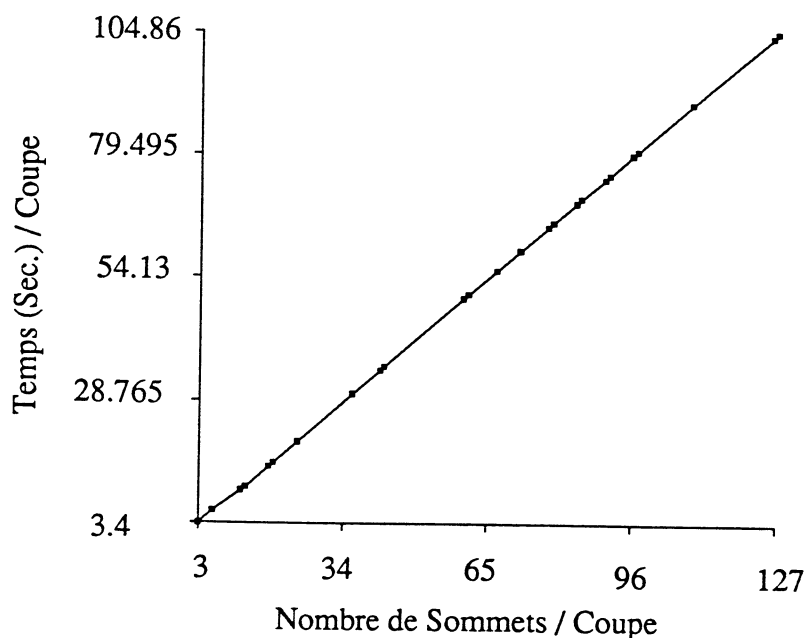


Figure 6.9 : Relation entre temps de calcul et nombre de sommets par coupe.

La figure 6.9 montre le caractère linéaire de la relation entre le temps de calcul et le nombre de sommets par coupe. De même nous avons calculé, pour les dix premiers zonotopes ( $n=1, \dots, 10$ ) et pour une direction donnée, le nombre moyen de sommets ainsi que le temps moyen de calcul par coupe ; les résultats obtenus sont présentés dans table 6.1-(b). On constate que le nombre moyen de sommets par coupe évolue d'une façon linéaire par rapport au nombre de vecteurs de base  $q$  (voir figure 6.10). Par contre, la relation entre le temps moyen par coupe et le nombre de vecteurs de base est loin d'être linéaire ; elle est polynomiale de degré 3 comme le montre la figure 6.11.

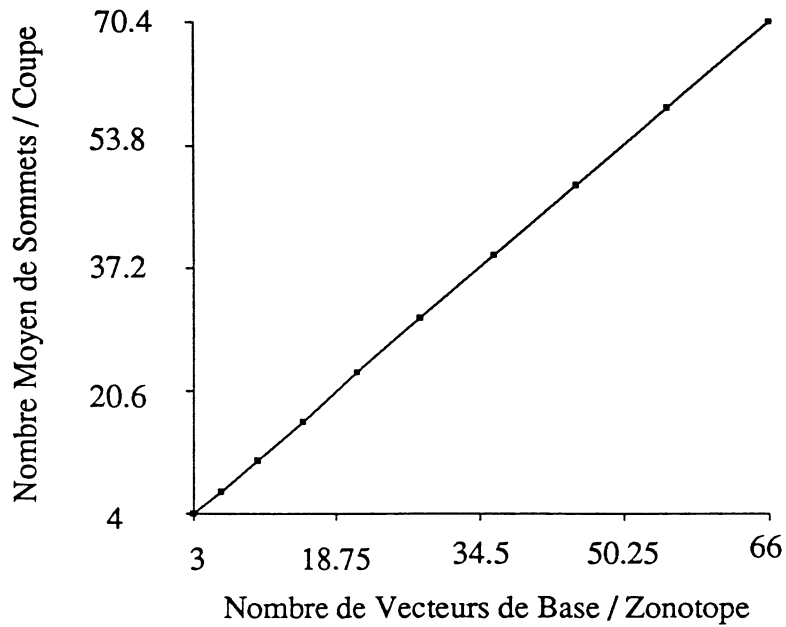


Figure 6.10 : Relation entre nombre moyen de sommets par coupe et nombre de vecteurs de base par zonotope.

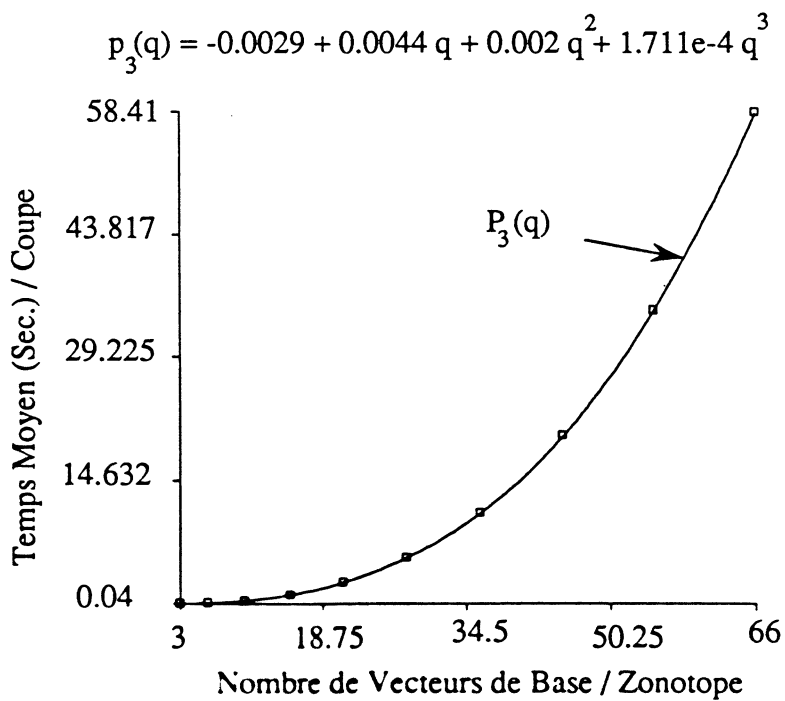


Figure 6.11 : Relation entre temps moyen de calcul par coupe et nombre de vecteurs de base par zonotope.

Enfin, la figure 6.12 présente une visualisation d'une dizaine de coupes, dans la direction  $A = (1,0,0)^t$ , des zonotopes correspondant aux valeurs  $n=1, \dots, n=10$ .

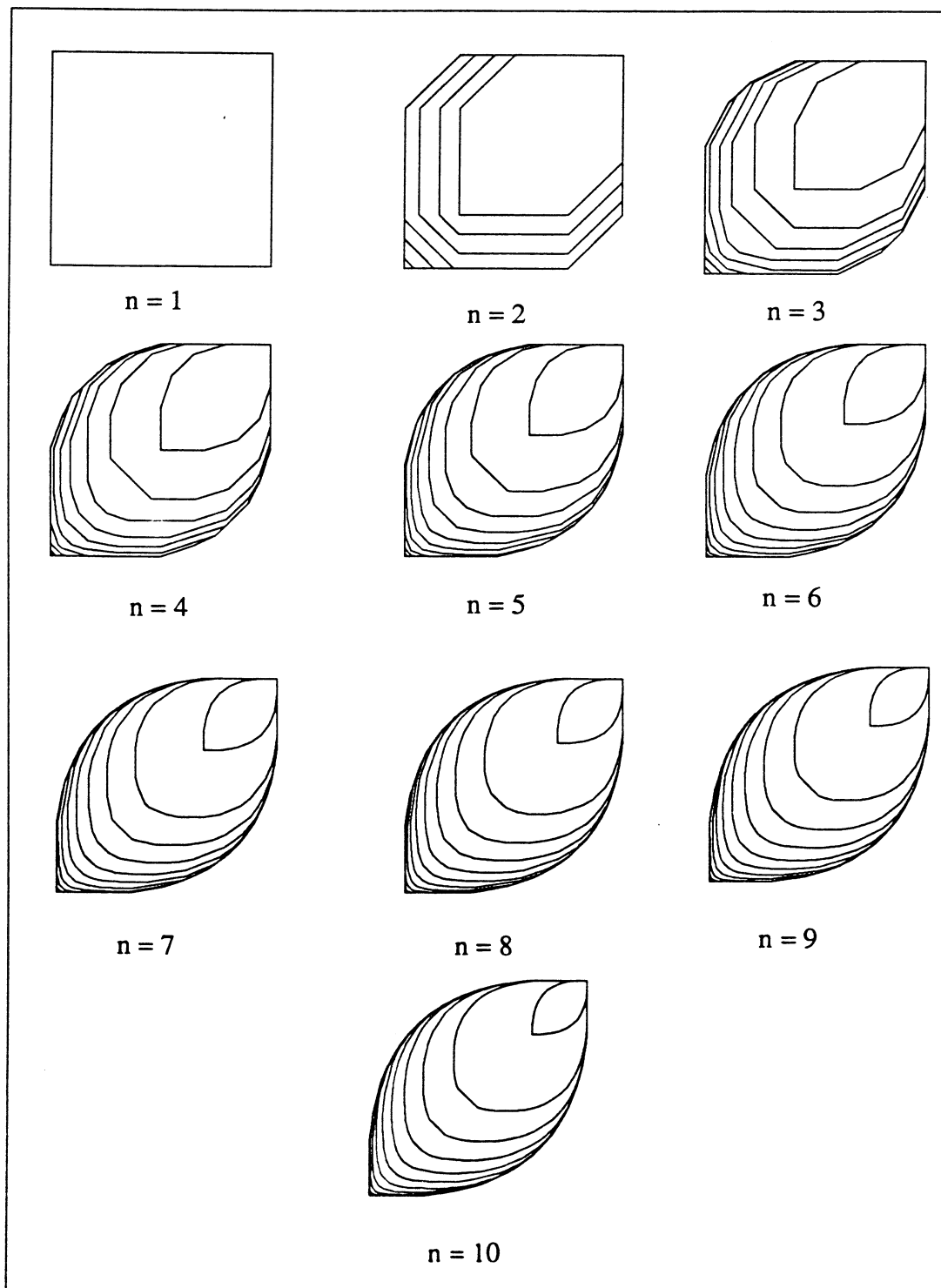


Figure 6.12 : Dix coupes dans la direction  $A = (1,0,0)^t$ .

- *Deuxième Exemple :*

Cet exemple est extrait de [Valentin 86]. On dispose de dix régions issues de la discrétisation dans le temps d'une opération de distillation. Pour chaque région on a évalué les quantités de trois espèces. Les vecteurs quantités obtenus sont présentés dans la table suivante :

région	espèce (1)	espèce (2)	espèce (3)
1	0.1891	0.0807	0.0301
2	0.1782	0.0872	0.0345
3	0.1647	0.0949	0.0403
4	0.1476	0.1041	0.0482
5	0.1257	0.1148	0.0596
6	0.0975	0.1263	0.0761
7	0.0632	0.1352	0.1015
8	0.0282	0.1315	0.1403
9	0.0055	0.0971	0.1973
10	0.0001	0.0279	0.2719

Table 6.2 : Vecteurs quantités associés aux différentes régions du système considéré.

Sur le zonotope généré par les vecteurs quantités ci-dessus, nous avons effectué plusieurs coupes dans des directions variées (voir figures 6.13-(a) et 6.13-(b)).

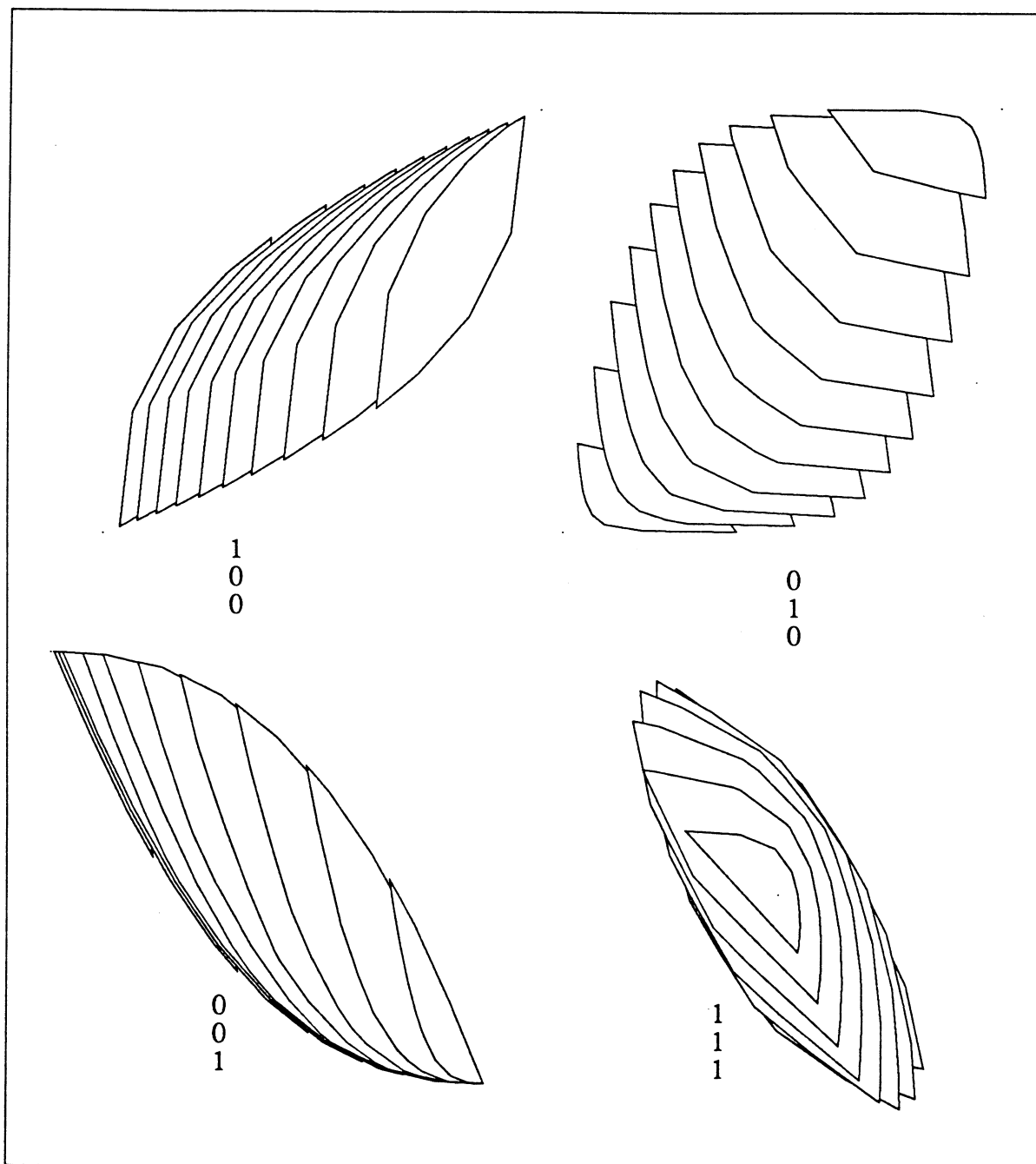


Figure 6.13 : (a)

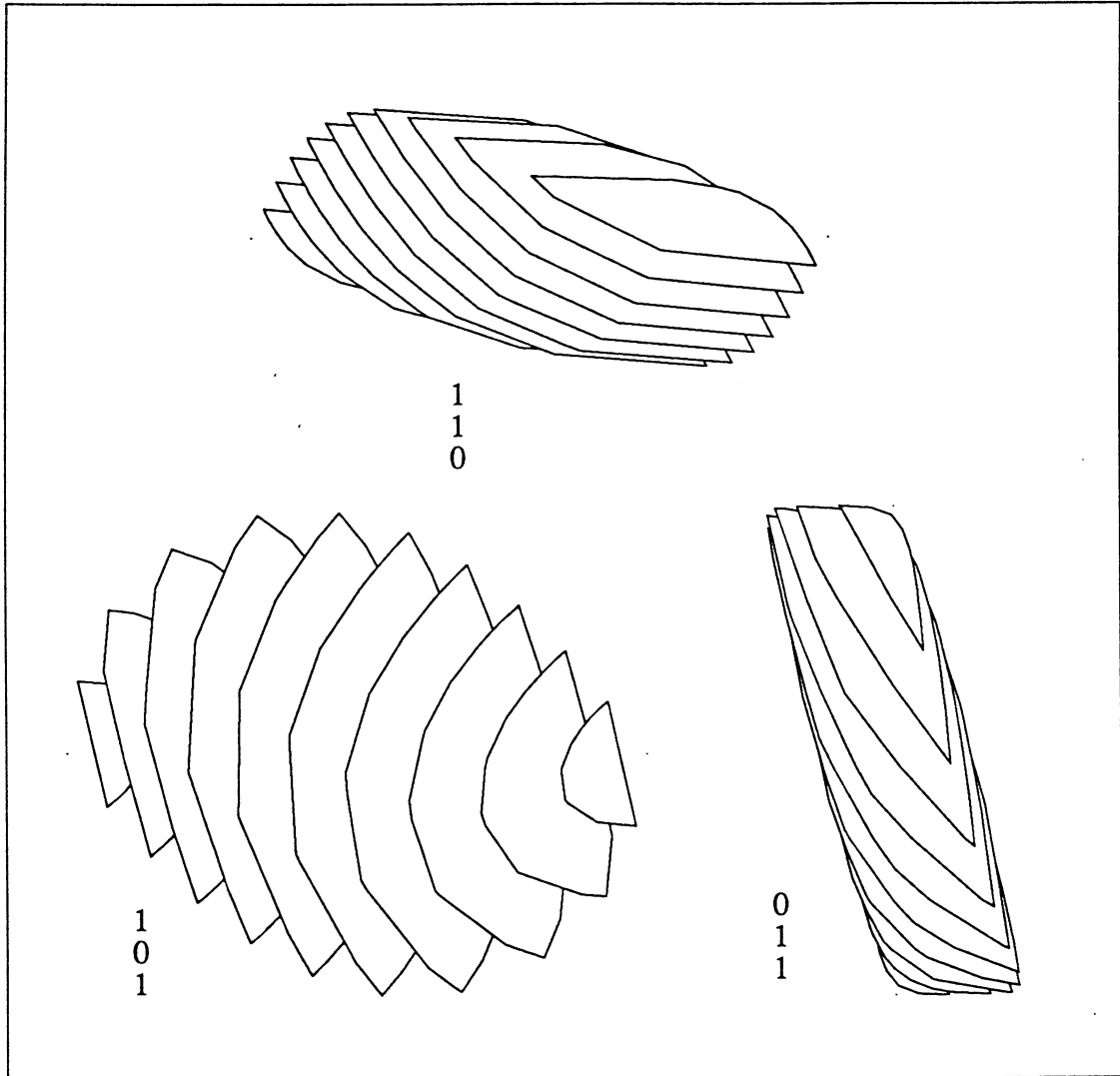


Figure 6.13 : (b)

## 6.6 Conclusion

Il est clair qu'une maîtrise géométrique des opérations géométriques sur le convexe de mélanges permet une meilleure gestion de la fabrication des mélanges et l'acquisition des outils efficaces pour guider les opérations physico-chimiques.

Le convexe de mélanges traduit la capacité de fabrication des mélanges à partir d'une partition du système physico-chimique correspondant et fournit une bonne caractérisation de la répartition (séparation) des espèces dans ce système.

Nous avons développé, essentiellement, dans ce chapitre un algorithme rapide et efficace permettant la détermination de l'intersection d'un zonotope avec un hyperplan dans notre espace euclidien usuel. Enfin, une expérimentation, sur cet algorithme, concernant sa complexité et son efficacité est présenté au § 6.5.





# **Conclusion**



L'analyse du pétrole brut fournit des résultats discrets et entachés d'erreurs représentés par un ensemble de coupes pour lesquelles on a mesuré, au laboratoire, plusieurs propriétés qui caractérisent le brut. Ces propriétés sont classées en deux catégories :

- **Propriétés conservatives (additives)** : comme la mesure du distillat (%poids et %volume) et la teneur en soufre ; nous avons proposé dans le chapitre II plusieurs modèles mathématiques permettant de représenter ces propriétés par des fonctions (splines) lisses continûment dérivables (passage du discret au continu) et de visualiser graphiquement les différentes courbes, surfaces et courbes de niveaux. La construction numérique des fonctions splines est basée sur la résolution d'un problème d'optimisation du type (P.3.1). Les études effectuées sur les propriétés conservatives ont permis de mettre au point le logiciel C-ELF qui permet de rendre continues et dérivables ces analyses, en utilisant les B-splines comme éléments de bases d'approximation. Ce logiciel permet aussi de reconstituer toutes les coupes possibles, de calculer leurs quantités, et propose plusieurs représentations visuelles à une et deux dimensions, par exemple la T.B.P. en fonction du point initial et sa carte d'état-major. Cet ensemble de représentations permet à l'utilisateur une évaluation commode des propriétés conservatives. Nous avons aussi présenté au chapitre II, §2.8 un procédé de modélisation des propriétés physiques s'exprimant par le quotient des quantités conservatives. Ce procédé permet d'obtenir, selon le cas, une courbe densité lisse, positive et monotone dans l'intervalle de définition.

- **Propriétés non conservatives** : ce sont les propriétés qui ne sont pas additives au mélange ; c'est le cas de la viscosité, du point de trouble au refroidissement et de la propriété antidétonante. Pour une propriété non

additive  $F$  dépendant d'une seule grandeur additive, le traitement consiste à chercher une fonction inversible  $I$  (appelée fonction indice) telle que  $I(F)$  soit additive. Nous avons présenté, dans la deuxième partie plusieurs approches permettant de trouver une telle fonction. Ces approches se différencient par le critère du choix utilisé.

La détermination de l'indice de mélange pour les propriétés non additives permet l'application des procédés de gestion de mélanges additifs. La fonction d'indice, étant une combinaison linéaire de B-splines, est un indice de mélange universel dans le sens où elle est parfaitement définie par ces coefficients sur les B-splines et son intervalle de validité.

La troisième partie de ce travail est consacrée à la géométrie des ensembles de mélanges. Nous avons présenté la notion de convexe de mélanges (zonotope) comme un ensemble de mélanges simples associés à une partition finie du système physico-chimique considéré. Cette représentation géométrique de l'ensemble de mélanges est très importante, car plusieurs opérations géométriques sur le convexe de mélanges ont leurs interprétations physiques relatives au système correspondant. Dans cette partie nous avons développé, essentiellement, un algorithme d'intersection du convexe de mélanges avec un hyperplan dans  $\mathbb{R}^3$ .

Enfin, il est clair qu'une maîtrise mathématique des opérations de mélanges (simples ou complexes) permet une meilleure gestion de la fabrication des mélanges et l'acquisition des outils efficaces pour guider les opérations industrielles.

# **Bibliographie**



- [Alexandrov 33] Alexandrov A.D. (1933), A theorem on convex polyhedra, Trudy Mat. Inst. Steklov. Sect. Math. 4, 87. (Russian)
- [Armijo 66] ARMIJO L. (1966), Minimization of Functions having Continuous Partial Derivatives, Pacific J.Math., vol.16, pp. 1-3.
- [Bartels/Beatty/Barsky 87] BARTELS R.H., BEATTY J.C., and BARSKY B.A. (1987), An Introduction to Splines for Use in Computer Graphics and Geometric Modeling, Morgan Kaufmann publishers, Los Altos, California.
- [Bolker 69] BOLKER E.D. (1969), A Classe of Convex Bodies, Trans. Amer. Math. Soc., vol. 145, pp. 323-345.
- [Boneva/Kendall/Stefanov 71] BONEVA L.I., KENDALL D. and STEFANOV I. (1971), Spline Transformations : Three New Diagnostic Aids for the Statistical Data-Analyst, J. Roy. Stat. Soc. B 33, PP. 1-70.
- [de Boor 78] de BOOR C. (1978), A practical Guide to Splines, Applied Mathematical Sciences Volume 27, Springer-Verlag, New York.
- [de Boor 72] de BOOR C. (1972), On Calculating with B-Splines, Journal of Approximation Theory 6(1), July, 50-62.



- [Ciarlet 82] CIARLET P.G. (1982), Introduction à l'Analyse Numérique Matricielle et à l'Optimisation, Masson, Paris.
- [Cox 72] COX M.G. (1972), The Numerical Evaluation of B-Splines, Journal of the Institute of Mathematics and its Applications 10(2), Octobre, 134-149.
- [Coxeter 62] COXETER H.S.M. (1962), The Classification of Zonohedra by means of Projective diagrams, J. Math. Pures Appl. (9) 41, pp. 137-156.
- [Gastinel 66] GASTINEL N. (1966), Analyse Numérique Linéaire, Hermann, Paris.
- [Girard 86] GIRARD D. (1986), Convexe Residu et Gestion de Mélanges sur une Plateforme, Rapport de Recherche Elf-Aquitaine n° 6/577, Juin.
- [Girard/Valentin 89] GIRARD D. and VALENTIN P. (1989), Zonotopes and Mixtures Management, International Series of Numerical Mathematics, vol.87, Birkhäuser Verlag Basel.
- [Goldstein/Price 67] GOLDSTEIN A.A. & PRICE J.F. (1967), An Effective Algorithm for Minimization, Numer. Math., vol.10, pp. 184-189.
- [Grünbaum 67] GRÜNBAUM B. (1967), Convex Polytopes, John Wiley and Sons, London-NewYork-Sydney.
- [Hermes/Lasalle 69] HERMES H. and LASALLE J.P., Functional Analysis and Time Optimal Control, Academic Press, NewYork, pp. 20-32.

- [Hildenbrand 81] HILDENBRAND W. (1981), Short-Run Productions Based on Microdata, *Econometrica*, vol. 49, n°5, pp. 1095-1125.
- [Laurent 72] LAURENT P.J. (1972), *Approximation et Optimisation*, Hermann, Paris.
- [Laurent 87] LAURENT P.J. (1987), *Courbes Ouvertes ou Fermées par B-Splines Régularisées*, Rapport de Recherche n° 652, Mars, Université de Grenoble.
- [McMullen 71] McMULLEN P. (1971), On Zonotopes, *Trans. Amer. Math. Soc.* vol. 159, pp. 91-110.
- [McMullen 73] McMULLEN P. (1973), Representations of Polytopes and Polyhedral Sets, *Geom. Ded.* 2, pp. 83-99.
- [McMullen/Schneider/Shephard 74] McMULLEN P., SCHNEIDER R. and SHEPHARD G.C. (1974), Monotypic Polytopes and their Intersection Properties, *Geom.Ded.* 3, pp. 99-129.
- [McMullen 75] McMULLEN P. (1975), Space Tiling Zonotopes, *Mathematika*, vol. 22, pp. 202-211.
- [McMullen/Shephard 76] McMULLEN P. and SHEPHARD G.C. (1976), Polytopes with Centrally Symmetric Facets, *Israël Journal of Mathematics* 23, PP. 3-4.
- [Medina 85] MEDINA J. (1985), *Fonctions Splines avec Conditions de Forme*, thèse U.S.T.M.G., Grenoble.
- [Minoux 83] MINOUX M. (1983), *Programmation Mathématique (tome 1)*, Dunod, Paris.

- [Powell 76] POWELL M.J.D. (1976), Some Global Convergence Properties of Variable Metric Algorithm for Minimization without Exact Line Searches, in Nonlinear Programming, SIAM-AMS Proceedings 9, R.W. Cottle et C.E. Lemke, Eds., Providence R.I.
- [Preparata/Shamos 85] PREPARATA F.P. and SHAMOS M.I. (1985), An Introduction to Computational Geometry, Springer-Verlag, New York.
- [Reinsch 67] REINSCH C.H. (1967), Smoothing by Spline Function I, Numer. Math. 10, 177-183.
- [Reinsch 71] REINSCH C.H. (1971), Smoothing by Spline Function II, Numer. Math. 16, 451-454.
- [Rosen 60] ROSEN J.B. (1960), The Gradient Projection Method for Non Linear Programming, part 1 : Linear Constraints, Journal S.I.A.M., 8, P.181-217.
- [Rosen 61] ROSEN J.B. (1961), The Gradient Projection Method for Non Linear Programming, part 2 : Non Linear Constraints, Journal S.I.A.M., 9, P.514-532.
- [Sajus 65] SAJUS L. (1965), LE PETROLE : Raffinage et Pétrochimie, Collection Armand Colin, n° 385, Section de chimie, Paris.
- [Schneider/Weil 82] SCHNEIDER R. and WEIL W. (1982), Zonoides and Related Topics, In Applications of Convexity, Birkhäuser Verlag, Basel.
- [Schumaker 81] SCHUMAKER L.L. (1981), Splines Functions: Basic Theory, John Wiley & Sons, New York.

- [Shephard 74a] SHEPHARD G.C. (1974), Combinatorial Properties of Associated Zonotopes, *Canad. J. Math.*, vol. 26, pp. 302-321.
- [Shephard 74b] SHEPHARD G.C. (1974), Space-Filling Zonotopes, *Mathematika*, vol. 21, pp. 261-269.
- [Slaoui 86] SLAOUI K.H. (1986), Application des Techniques Mathématiques à la Gestion des Mélanges : Histosplines et Optimisation, thèse I.N.P.G., Grenoble.
- [Valentin 86] VALENTIN P. (1986), Convex Set of Mixtures : A Geometric Approach to Separation Engineering, Elf-Solaize Research Center.
- [Vasseur 65] VASSEUR L.V. (1965), Pétrole et Gaz Naturel dans le Monde, Collection Armand Colin, n° 328, Section Géographie, Paris.
- [Vellinger 70] VELLINGER E. (1970), Industrie du Pétrole, DUNOD, Paris.
- [Wolfe 69] WOLFE P. (1969), Convergence Conditions for Ascent methodes, *SIAM Review*, 11, pp. 226-235.



AUTORISATION DE SOUTENANCE

DOCTORAT 3ème CYCLE, DOCTORAT INGENIEUR,  
DOCTORAT DE L'UNIVERSITE JOSEPH FOURIER - GRENOBLE 1

Vu les dispositions de l'Arrêté du 16 avril 1974,

Vu les dispositions de l'Arrêté du 5 juillet 1984,

Vu les rapports de Mr. .... **A. LE MÉHAUTÉ** .....

Mr. .... **P. VALENTIN** .....

Mr. .... **ODEH Nabih** ..... est autorisé(e)  
à présenter une thèse en vue de l'obtention du .. **DU 1 -** .....  
..... **Mathématiques Appliquées** .....

07 MARS 1990

Grenoble, le .....

Le Président de l'Université  
Joseph Fourier - Grenoble 1



A. NEMOZ







## **Résumé**

La séparation dans le domaine pétrolier consiste à arranger ou déplacer des corps ou des classes d'espèces dans des régions différentes de façon à avoir plus de facilités à évaluer leurs propriétés ou bien à produire d'autres mélanges. Le travail, que nous présentons ici, est une contribution à la résolution numérique de certains problèmes rencontrés dans l'étude de séparation du pétrole. Les problèmes étudiés, dans ce travail, concernent deux types de mélanges différents : mélanges complexes et mélanges simples. L'étude effectuée, sur différents problèmes de la gestion des mélanges, a conduit à mettre au point un logiciel interactif et visuel ainsi que plusieurs programmes d'expérimentation ; ceci va permettre une exploitation numérique et graphique meilleure.

## **Abstract**

Separation in the oil field consists in arranging or displacing bodies or some categories of species into different regions so that there are more facilities to evaluate their properties or to make other mixtures. This work is a contribution to the numerical resolution of some problems arising in the study of oil separation. The problems considered here are concerning two different kinds of mixtures : complex mixtures and simple mixtures. The performed study, of different problems of mixing management, have led to implement an interactive and visual software and many experimental programs ; this will permit a better numerical and graphical exploitation.

## **Mots clés**

B-splines, Mélanges, Lissage, Modélisation, T.B.P., Optimisation, Approximation, Zonotope.