



HAL
open science

Application des techniques mathématiques à la gestion des mélanges : histosplines et optimisation

Khalid H. Slaoui

► **To cite this version:**

Khalid H. Slaoui. Application des techniques mathématiques à la gestion des mélanges : histosplines et optimisation. Modélisation et simulation. Institut National Polytechnique de Grenoble - INPG, 1986. Français. NNT: . tel-00320471

HAL Id: tel-00320471

<https://theses.hal.science/tel-00320471v1>

Submitted on 11 Sep 2008

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

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

THESE

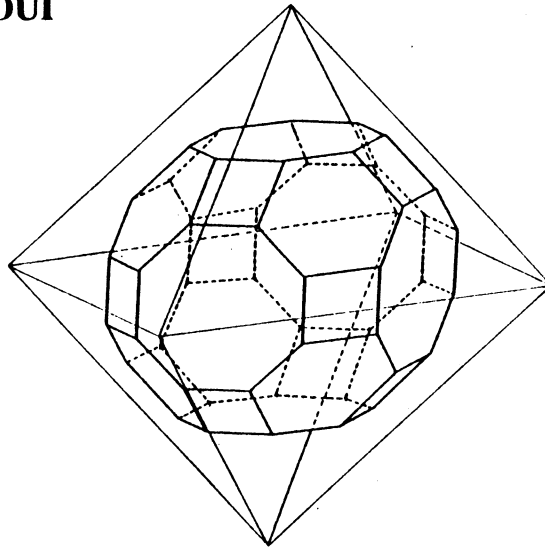
présentée à

l'Institut National Polytechnique de Grenoble

pour obtenir le titre de
Docteur Ingénieur
Mathématiques Appliquées

par

Khalid.H. SLAOUI



Application des techniques mathématique à la gestion des mélanges: histosplines et optimisation

Date de Soutenance : 27 Juin 1986

Composition du Jury :

F. ROBERT

Président

G.GUIOCHON

B.LACOLLE

P.J.LAURENT

P.VALENTIN

Examineurs



INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

Année universitaire 1982-1983

Président de l'Université : D. BLOCH

**Vice-Président : René CARRE
Hervé CHERADAME
Marcel IVANES**

PROFESSEURS DES UNIVERSITES :

ANCEAU François	E.N.S.I.M.A.G.
BARRAUD Alain	E.N.S.I.E.G.
BAUDELET Bernard	E.N.S.I.E.G.
BESSON Jean	E.N.S.E.E.G.
BLIMAN Samuel	E.N.S.E.R.G.
BLOCH Daniel	E.N.S.I.E.G.
BOIS Philippe	E.N.S.H.G.
BONNETAIN Lucien	E.N.S.E.E.G.
BONNIER Etienne	E.N.S.E.E.G.
BOUVARD Maurice	E.N.S.H.G.
BRISSONNEAU Pierre	E.N.S.I.E.G.
BUYLE BODIN Maurice	E.N.S.E.R.G.
CAVAIGNAC Jean-François	E.N.S.I.E.G.
CHARTIER Germain	E.N.S.I.E.G.
CHENEVIER Pierre	E.N.S.E.R.G.
CHERADAME Hervé	U.E.R.M.C.P.P.
CHERUY Arlette	E.N.S.I.E.G.
CHIAVERINA Jean	U.E.R.M.C.P.P.
COHEN Joseph	E.N.S.E.R.G.
COUMES André	E.N.S.E.R.G.
DURAND Francis	E.N.S.E.E.G.
DURAND Jean-Louis	E.N.S.I.E.G.
FELICI Noël	E.N.S.I.E.G.
FOULARD Claude	E.N.S.I.E.G.
GENTIL Pierre	E.N.S.E.R.G.
GUERIN Bernard	E.N.S.E.R.G.
GUYOT Pierre	E.N.S.E.E.G.
IVANES Marcel	E.N.S.I.E.G.
JAUSSAUD Pierre	E.N.S.I.E.G.
JOUBERT Jean-Claude	E.N.S.I.E.G.
JOURDAIN Geneviève	E.N.S.I.E.G.
LACOUME Jean-Louis	E.N.S.I.E.G.
LATOMBE Jean-Claude	E.N.S.I.M.A.G.

.../...

LESSIEUR Marcel	E.N.S.H.G.
LESPINARD Georges	E.N.S.H.G.
LONGEQUEUE Jean-Pierre	E.N.S.I.E.G.
MAZARE Guy	E.N.S.I.M.A.G.
MOREAU René	E.N.S.H.G.
MORET Roger	E.N.S.I.E.G.
MOSSIERE Jacques	E.N.S.I.M.A.G.
PARIAUD Jean-Charles	E.N.S.E.E.G.
PAUTHENET René	E.N.S.I.E.G.
PERRET René	E.N.S.I.E.G.
PERRET Robert	E.N.S.I.E.G.
PIAU Jean-Michel	E.N.S.H.G.
POLOUJADOFF Michel	E.N.S.I.E.G.
POUPOT Christian	E.N.S.E.R.G.
RAMEAU Jean-Jacques	E.N.S.E.E.G.
RENAUD Maurice	U.E.R.M.C.P.P.
ROBERT André	U.E.R.M.C.P.P.
ROBERT François	E.N.S.I.M.A.G.
SABONNADIÈRE Jean-Claude	E.N.S.I.E.G.
SAUCIER Gabrielle	E.N.S.I.M.A.G.
SCHLENKER Claire	E.N.S.I.E.G.
SCHLENKER Michel	E.N.S.I.E.G.
SERMET Pierre	E.N.S.E.R.G.
SILVY Jacques	U.E.R.M.C.P.P.
SOHM Jean-Claude	E.N.S.E.E.G.
SOUQUET Jean-Louis	E.N.S.E.E.G.
VEILLON Gérard	E.N.S.I.M.A.G.
ZADWORNY François	E.N.S.E.R.G.

PROFESSEURS ASSOCIES

BASTIN Georges	E.N.S.H.G.
BERRIL John	E.N.S.H.G.
CARREAU Pierre	E.N.S.H.G.
GANDINI Alessandro	U.E.R.M.C.P.P.
HAYASHI Hirashi	E.N.S.I.E.G.

PROFESSEURS UNIVERSITE DES SCIENCES SOCIALES (Grenoble II)

BOLLIET Louis
Chatelin Françoise

PROFESSEURS E.N.S. Mines de Saint-Etienne

RIEU Jean
SOUSTELLE Michel

CHERCHEURS DU C.N.R.S.

FRUCHART Robert
VACHAUD Georges

Directeur de Recherche
Directeur de Recherche

.../...

ALLIBERT Michel	Maître de Recherche
ANSARA Ibrahim	Maître de Recherche
ARMAND Michel	Maître de Recherche
BINDER Gilbert	
CARRE René	Maître de Recherche
DAVID René	Maître de Recherche
DEPORTES Jacques	
DRIOLE Jean	Maître de Recherche
GIGNOUX Damien	
GIVORD Dominique	
GÜELIN Pierre	
HOPFINGER Emil	Maître de Recherche
JOUD Jean-Charles	Maître de Recherche
KAMARINOS Georges	Maître de Recherche
KLEITZ Michel	Maître de Recherche
LANDAU Ioan-Dore	Maître de Recherche
LASJAUNIAS J.C.	
MERMET Jean	Maître de Recherche
MUNIER Jacques	Maître de Recherche
PIAU Monique	
PORTESEIL Jean-Louis	
THOLENCE Jean-Louis	
VERDILLON André	

CHERCHEURS du MINISTERE de la RECHERCHE et de la TECHNOLOGIE (Directeurs et Maîtres de Recherches, ENS Mines de St. Etienne)

LESBATS Pierre	Directeur de Recherche
BISCONDI Michel	Maître de Recherche
KOBYLANSKI André	Maître de Recherche
LE COZE Jean	Maître de Recherche
LALAUZE René	Maître de Recherche
LANCELOT Francis	Maître de Recherche
THEVENOT François	Maître de Recherche
TRAN MINH Canh	Maître de Recherche

PERSONNALITES HABILITEES à DIRIGER des TRAVAUX de RECHERCHE (Décision du Conseil Scientifique)

ALLIBERT Colette	E.N.S.E.E.G.
BERNARD Claude	E.N.S.E.E.G.
BONNET Rolland	E.N.S.E.E.G.
CAILLET Marcel	E.N.S.E.E.G.
CHATILLON Catherine	E.N.S.E.E.G.
CHATILLON Christian	E.N.S.E.E.G.
COULON Michel	E.N.S.E.E.G.
DIARD Jean-Paul	E.N.S.E.E.G.
EUSTAPOPOULOS Nicolas	E.N.S.E.E.G.
FOSTER Panayotis	E.N.S.E.E.G.

.../...

GALERIE Alain	E.N.S.E.E.G.
HAMMOU Abdelkader	E.N.S.E.E.G.
MALMEJAC Yves	E.N.S.E.E.G. (CENG)
MARTIN GARIN Régina	E.N.S.E.E.G.
NGUYEN TRUONG Bernadette	E.N.S.E.E.G.
RAVAINE Denis	E.N.S.E.E.G.
SAINFORT	E.N.S.E.E.G. (CENG)
SARRAZIN Pierre	E.N.S.E.E.G.
SIMON Jean-Paul	E.N.S.E.E.G.
TOUZAIN Philippe	E.N.S.E.E.G.
URBAIN Georges	E.N.S.E.E.G. (Laboratoire des ultra-réfractaires ODEILLON)
GUILHOT Bernard	E.N.S. Mines Saint Etienne
THOMAS Gérard	E.N.S. Mines Saint Etienne
DRIVER Julien	E.N.S. Mines Saint Etienne
BARIBAUD Michel	E.N.S.E.R.G.
BOREL Joseph	E.N.S.E.R.G.
CHOVET Alain	E.N.S.E.R.G.
CHEHIKIAN Alain	E.N.S.E.R.G.
DOLMAZON Jean-Marc	E.N.S.E.R.G.
HERAULT Jeanny	E.N.S.E.R.G.
MONLLOR Christian	E.N.S.E.R.G.
BORNARD Guy	E.N.S.I.E.G.
DESCHIZEAU Pierre	E.N.S.I.E.G.
GLANGEAUD François	E.N.S.I.E.G.
KOFMAN Walter	E.N.S.I.E.G.
LEJEUNE Gérard	E.N.S.I.E.G.
MAZUER Jean	E.N.S.I.E.G.
PERARD Jacques	E.N.S.I.E.G.
REINISCH Raymond	E.N.S.I.E.G.
ALEMANY Antoine	E.N.S.H.G.
BOIS Daniel	E.N.S.H.G.
DARVE Félix	E.N.S.H.G.
MICHEL Jean-Marie	E.N.S.H.G.
OBLED Charles	E.N.S.H.G.
ROWE Alain	E.N.S.H.G.
VAUCLIN Michel	E.N.S.H.G.
WACK Bernard	E.N.S.H.G.
BERT Didier	E.N.S.I.M.A.G.
CALMET Jacques	E.N.S.I.M.A.G.
COURTIN Jacques	E.N.S.I.M.A.G.
COURTOIS Bernard	E.N.S.I.M.A.G.
DELLA DORA Jean	E.N.S.I.M.A.G.
FONLUPT Jean	E.N.S.I.M.A.G.
SIFAKIS Joseph	E.N.S.I.M.A.G.
CHARUEL Robert	U.E.R.M.C.P.P.
CADET Jean	C.E.N.G.
COEURE Philippe	C.E.N.G. (LETI)

.../...

DELHAYE Jean-Marc	C.E.N.G. (STT)
DUPUY Michel	C.E.N.G. (LETI)
JOUBE Hubert	C.E.N.G. (LETI)
NICOLAU Yvan	C.E.N.G. (LETI)
NIFENECKER Hervé	C.E.N.G.
PERROUD Paul	C.E.N.G.
PEUZIN Jean-Claude	C.E.N.G. (LETI)
TAIEB Maurice	C.E.N.G.
VINCENDON Marc	C.E.N.G.

LABORATOIRES EXTERIEURS

DEMOULIN Eric	C.N.E.T.
DEVINE	C.N.E.T. (R.A.B.)
GERBER Roland	C.N.E.T.
MERCKEL Gérard	C.N.E.T.
PAULEAU Yves	C.N.E.T.
GAUBERT C.	I.N.S.A. Lyon



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Ce travail a été réalisé au **C.R.E.S**: Centre de Recherche de Elf à Solaize en collaboration avec le laboratoire **tim3**: Technique de l'Informatique, des mathématiques, de la Microélectronique et de la Microscopie quantitative.

Qu'il me soit permis d'exprimer mes remerciements les plus sincères à Monsieur **P. J. Laurent**, pour le soutien, la confiance et la sympathie qu'il m'a témoigné.

Je tiens à remercier aussi **P.Valentin** et Madame **M.T.Feillens** du **C.R.E.S** de m'avoir accueilli dans ce centre. Ils ont suivi ce travail avec beaucoup d'attention et un souci permanent.

Mes vifs remerciements vont également à Monsieur **B. Lacolle** qui a dirigé mes recherches au D.E.A et pour l'intérêt qu'il a manifesté à la deuxième partie de ce travail.

Je remercie Monsieur **F.Robert**, professeur à l'E.N.S.I.M.A.G d'avoir accepté de présider le jury, et Monsieur **G.Gulochon** pour bien voulu en faire partie.

Je souhaite remercier également l'ensemble du personnel de la section analytique du **C.R.E.S** pour l'accueil et la sympathie qu'ils m'ont réservé.

Je ne manquerais pas de remercier tous mes amis et frères à grenoble et particulièrement **M.Mouzali** pour tous les bons moments passés ensemble.

Enfin, mes remerciements vont à tous les membres du service de Reprographie de l'I.M.A.G pour l'excellente qualité de leur travail.



TABLE DES MATIERES

Introduction générale

Partie I

Spline quadratique et quartique, et analyse des mélanges complexes

Introduction

Chapitre 1 **Spline quadratique d'interpolation**

1- Introduction.....	9
2- Définitions et notations.....	9
3- Problème de minimisation.....	11
4- Construction numérique et algorithme de résolution..	13
4.1- Histogramme à intervalles juxtaposés.....	13
4.2- Histogrammé à intervalles chevauchés.....	18

Chapitre 2 **Spline quadratique d'ajustement**

1- Introduction.....	29
2- Construction numérique et algorithme de résolution..	30
2.1- Histogramme à intervalles quelconques.....	30
2.2- Histogramme à intervalles adjacents.....	32
3- Conclusion.....	35

Chapitre 3 **Spline quartique d'interpolation et d'ajustement**

1- Introduction.....	39
2- Spline quartique d'interpolation.....	39
2.1- Notations et définitions.....	39
2.2- Construction numérique.....	39
3- Spline quartique d'ajustement.....	45

Chapitre 4 **Modélisation des histogrammes à intervalles non bornés**

1- Introduction.....	51
2- Approche par un changement de variable.....	52
3- Paramétrisation des mesures à intervalles non bornés.....	53
4- Introduction d'une fonction-poids ou de forme.....	54
4.1- Interpolation.....	55
4.2- Ajustement.....	59

Chapitre 5 **Applications: Calcul des propriétés des coupes pétrolières et régularisation des courbes des fractions de ségrégation**

Calcul des propriétés des coupes pétrolières

1- Introduction.....	63
2- Analyse du pétrole brut.....	64
2.1- Introduction.....	64
2.2- Rappel de la distillation " T.B.P ".....	64
2.3- Analyse des quantités distillées.....	65
3- Modélisation des coupes pétrolières.....	65
3.1- Description des mesures.....	65
3.2- Représentation graphique.....	66
4- Progiciel C-elf	67
4.1- Présentation générale du progiciel.....	67
4.2- Description des modules.....	68
5- Analyse du brut A 334.....	70
5.1- Données d'un brut.....	70
5.2- Construction des fichiers de données.....	70
5.3- Interprétation des courbes.....	70
5.3.1- Distillat.....	71
5.3.2- Densité.....	72
5.3.3- Soufre.....	73
5.3.4- N-paraffines.....	74
5.4- Méthode de superposition des cartes.....	74
6- Conclusion.....	81

Régularisation des courbes des fractions de ségrégation

1- Introduction.....	82
2- Définition et rappel.....	84
2.1- Chromatogramme.....	84
2.2- Fraction de ségrégation.....	85
3- Présentation du progiciel S-elf	88
3.1- Principe général.....	88
3.2- Organigramme.....	89
3.3- Recalage global des N-paraffines.....	90
3.3.1- Repérage automatique.....	90
3.3.2- Repérage conversationnel.....	90
3.4- Fonction de recalage.....	90
3.5- Chromatogramme recalé.....	91
3.6- Traitement d'un chromatogramme.....	92
4- Application: Distillat sous-vide.....	93
5- Conclusion et recommandation.....	98
<u>Annexe</u>	99
<u>Bibliographie</u>	107



Partie II

Gestion et géométrie des ensembles de mélange

Introduction

Chapitre 1 **Formulation générale**

1- Définition préliminaires.....	115
2- Propriétés de l'ensemble des mélanges.....	116
3- Traduction algébrique des opérations physico-chimique dans un système.....	120
3.1- Opérations d'inclusion et d'intersection.....	121
3.2- Opérations de projection.....	122

Chapitre 2 **Faisabilité d'un mélange objectif**

1- Introduction.....	125
2- Caractérisation d'un mélange objectif.....	126
3- Fabrication d'un mélange objectif.....	129
3.1- Méthode de projection.....	129
1- Méthode de Barycentre.....	130
2- Méthode de projections successives.....	132
3- Test d'arrêt.....	132
3.2- Formulation de la fabrication sous la forme d'un problème d'optimisation.....	133
3.3- Méthode de décomposition.....	134
1- Norme euclidienne et forme quadratique.....	134
2- Sous-problème d'optimisation.....	134
3- Méthode en série.....	135
4- Méthode en parallèle.....	136
5- Interprétation géométrique des itérations.....	137

3.4- Formulation linéaire.....	138
1- Norme L_1 et méthode primale.....	138
2- Simplexe révisé à variables bornées.....	139
3- Interprétation géométrique des itérations.....	141
4- Méthode duale.....	143
4- Application numérique et synthèse.....	145

Chapitre 3 **Gestion des mélanges binaires**

1- Introduction.....	151
2- Bord du convexe des mélanges binaires.....	152
3- Droite d'appui du convexe des mélanges binaires.....	153
4- Faisabilité d'un mélange objectif.....	154
4.1- Possibilité de fabrication.....	154
4.2- Fabrication optimale d'un mélange objectif.....	154
4.3- Fabrication des mélanges ayant une contrainte linéaire	
5- Applications.....	159
5.1- Applications numériques.....	159
5.2- Problème de recyclage en chromatographie.....	162
<u>Bibliographie</u>	167

Introduction générale

La séparation est un phénomène quotidien, que l'on rencontre dans plusieurs domaines: chimie, pétrochimie etc... Les techniques de séparation sont bien étudiées et largement développées, ce qui explique son importance industrielle (par exemple 50 à 70% des investissements d'une usine chimique ou pétrolière portent sur la séparation). Cependant, plusieurs problèmes numériques, notamment dans le domaine pétrolier, demeurent non résolus, jusqu'à maintenant.

Nous apportons dans le présent travail une contribution à la résolution numérique de certains de ces problèmes, qui sont apparus dans l'étude de la séparation du pétrole.

La séparation consiste à arranger ou bien à déplacer des corps ou des classes d'espèces d'un système (surface fermée dans l'espace usuel) dans des régions différentes, de façon à avoir plus de facilités à évaluer leur propriétés ou à produire d'autres mélanges.

Notons n le nombre d'espèce et p le nombre de régions du système, on distingue dans ce travail deux types de mélange différents:

1- Mélanges complexes

Il consiste à analyser un très grand nombre d'espèces ($n=\infty$) et d'étudier leurs propriétés physiques dans le pétrole brut ($p=1$) à partir d'un nombre fini de mesures.

Nous avons étudié successivement deux problèmes d'une très grande importance dans l'industrie pétrolière:

- l'analyse et l'évaluation des propriétés physiques et conservatives des coupes pétrolières.
- l'ajustement et le recalage des chromatogrammes afin de rendre régulière les courbes des fractions de ségrégation.

L'outil mathématique fondamental dans ces traitements est les fonctions histosplines d'interpolation et de lissage, pour lesquelles les mesures sont les valeurs moyennes de la fonction sur des intervalles donnés. Nous avons proposé plusieurs méthodes de résolution selon la structure des intervalles (intervalles adjacents ou chevauchés et bornés ou non bornés).

Ceci nous a conduit à mettre au point deux progiciels **C-elf** et **S-elf** satisfaisant aux besoins de plusieurs secteurs de recherche pétrolière. Ils permettent une meilleure exploitation numérique et graphique des mélanges complexes.

2-Mélanges simples

En supposant le nombre d'espèce fini, 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, non résolus jusqu'à présent.

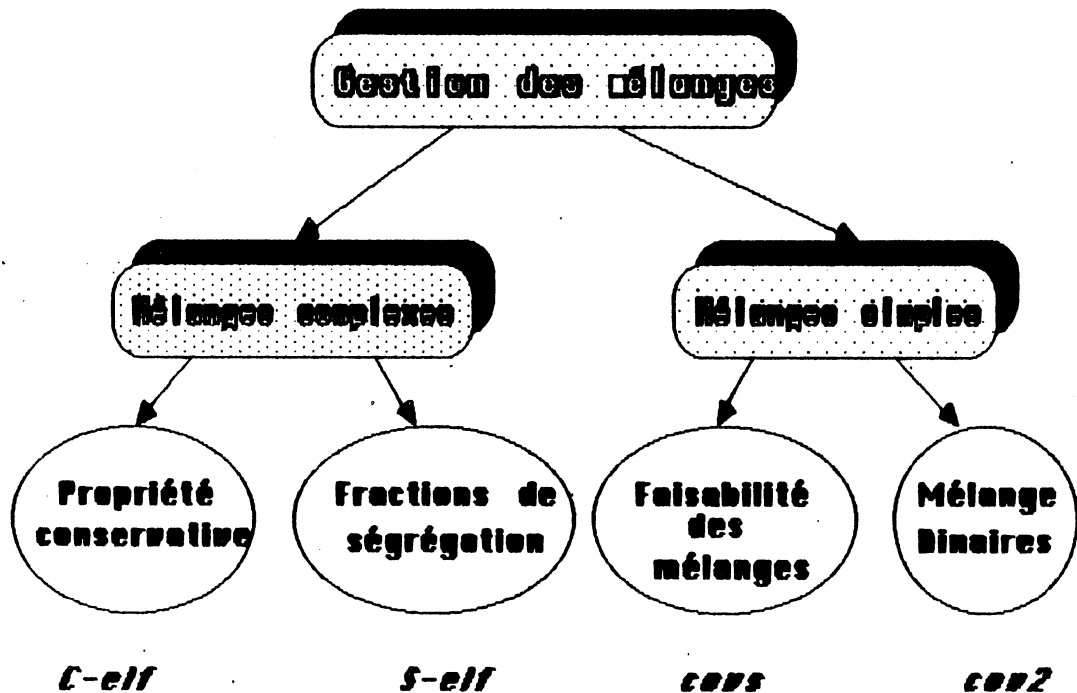
Nous nous sommes intéressés dans cette partie uniquement à deux problèmes:

-le premier concerne la fabrication d'un mélange objectif à partir d'une famille de mélange de base donnée.

Notre étude consistait à appliquer plusieurs méthodes d'optimisation convexe et à les comparer (méthode de projection, de décomposition et du simplexe).

-le deuxième problème concerne les mélanges simples ($n=2$). Plusieurs problèmes élémentaires sont résolus. Toutes les méthodes utilisées sont basées sur les techniques de la géométrie algorithmique. Ceci nous a permis, d'une part de développer un progiciel numérique et graphique regroupant toutes ces opérations, et d'autre part, de résoudre d'une façon rapide deux problèmes concrets en chromatographie: production et points de coupe optimale.

Nous résumons le plan de ce travail dans le tableau suivant:



Partie I

Spline quadratique et quartique, et analyse des mélanges complexes



Introduction

Les splines cubiques à une variable sont très utilisées. Leur théorie est largement développée dans [2],[5],[6] et [7]. Des procédures de calcul efficaces et rapides sont proposées dans [4] et [6].

Elles permettent une bonne approximation des signaux pour lesquels les mesures (observations) sont de nature ponctuelle.

Mais dans certains domaines comme la chimie, la physique, la statistique, la sociologie, etc.. on étale l'observation du phénomène ou signal sur des domaines plus larges, ce qui revient à dire que l'on mesure la valeur moyenne sur un intervalle. On représente cela graphiquement par une pile dont l'amplitude (ou hauteur) est la valeur moyenne et dont la base est l'intervalle d'observation.

L'ensemble de ces piles constitue un histogramme. On peut déjà distinguer deux classes d'histogrammes suivant la structure des intervalles de mesure:

- Histogramme à intervalles juxtaposés ou contigus.
- Histogramme à intervalles chevauchés.

Le problème consiste à reconstruire le signal à partir de son histogramme, autrement dit, de pouvoir calculer la valeur du signal en tout point de son domaine de définition.

Boneva, Kendall et Stefanov dans [1] (1970) ont largement étudié les histogrammes contigus et leur modélisation par les splines quadratiques (histosplines à une variable). L'histospline correspondant à un histogramme h donné est déterminée par la somme des histosplines unités ou "deltasplines" des intervalles de h . Cette technique locale n'est pas applicable directement à des histogrammes chevauchés.

On rencontre, des études de convergence et de lissage des histosplines par **P.J.Laurent** dans [5].

Une généralisation des histosplines à plusieurs variables et particulièrement à deux dimensions est développée dans les articles de **Nira Dyn** et **Grace Whaba** [13] et [14]. Plusieurs critères de lissage et les algorithmes correspondants y sont proposés. Il s'avère que les histosplines sont convenables pour estimer les fonctions densités.

Toutes les études de construction citées plus haut considèrent deux hypothèses essentielles: le domaine des mesures est borné et les régions des mesures (intervalles à une variable) sont linéairement indépendants.

L'objet essentiel de cette étude est la modélisation de deux problèmes pratiques dans le domaine de l'analyse des pétroles bruts:

-l'analyse et l'évaluation des propriétés physiques et conservatives des coupes pétrolières.

-l'ajustement et le recalage des chromatogrammes afin de rendre régulière les courbes des fractions de ségrégation.

les mesures des propriétés des coupes ne satisfont pas les hypothèses ci-dessus. Nous proposons dans cette partie trois techniques différentes afin de modéliser ces types de données.

Le travail présenté dans cette partie se compose de cinq chapitres:

chapitre 1,2 Développement des algorithmes efficaces et construction numérique des splines quadratiques d'interpolation et d'ajustement pour des histogrammes quelconques.

chapitre 3: Construction numérique des splines quartiques d'interpolation et d'ajustement.

chapitre 4: Les mesures des coupes pétrolières s'étalent sur des intervalles non bornés. Cette propriété rend le modèle des chapitres 1 et 2 non applicable directement. Nous avons proposé dans ce chapitre trois techniques différentes qui tiennent compte de cette particularité.

1- une approche par un changement de variable.

2- introduction de deux paramètres supplémentaires représentant cette spécificité.

3- introduction d'une fonction-poids ou de "forme" conférant au modèle cette spécificité.

chapitre 5: Dans ce chapitre nous avons décrit les deux problèmes physiques cités plus haut, et l'application du modèle des splines quadratiques. Cela a donné naissance à deux progiciels interactifs et visuels: *c-elf* et *s-elf*

Chapitre 1

Spline quadratique d'interpolation

- 1- Introduction
- 2- Définitions et notations
- 3- Problème de minimisation
- 4- Construction numérique et algorithme de résolution
 - 4.1-Histogramme à intervalles juxtaposés
 - 4.2-Histogramme à intervalles chevauchés



1) Introduction

Le travail présenté dans ce chapitre porte essentiellement sur la construction numérique des splines quadratiques correspondant aux histogrammes contigus et chevauchés.

Nous rappelons, dans la deuxième partie de ce chapitre, certaines notations et définitions concernant les histogrammes et les histosplines. Dans la troisième partie nous évoquerons le problème de minimisation conduisant à la spline quadratique en discutant son existence, son unicité et sa caractérisation.

Dans la quatrième partie, nous développerons deux algorithmes de construction numérique, efficaces et rapides, pour des intervalles juxtaposés et à pas constant. Ces deux algorithmes ne sont pas applicables pour les autres mesures. En conséquence, nous avons introduit la notion de graphe pour représenter les histogrammes, ce qui nous amène à distinguer deux sous-classes d'histogrammes chevauchés (connexe et non connexe). La première classe se ramène aux histogrammes contigus et permet de réappliquer les algorithmes précédents. Par contre, la seconde s'avère plus délicate; nous suggérons d'appliquer la méthode de projection.

2) Définitions et notations

Soit $[a,b]$ un intervalle de \mathbb{R} , et une subdivision de n points $T_l, l=1, \dots, n$ dans $[a,b]$ telle que $a=T_1 < \dots < T_{i-1} < \dots < T_i < \dots < T_n = b$.

On note:

- $I_l = [T_l, T_{l+1}]$, $h_l = T_{l+1} - T_l$, $l=1, \dots, n-1$.
- $I^k = [T_{i^k}, T_{j^k}]$, $(T_{i^k} < T_{j^k})$, $m \geq 1$, $i^k, j^k \in \{1, 2, \dots, n\}$ et $k=1, \dots, m$.
- $z = (z_k)$, $k=1, \dots, m$ un vecteur de \mathbb{R}^m qui représente les mesures des valeurs moyennes d'un signal s sur les intervalles I^k .

$$\int_{I^k} s(t) dt = z_k , \quad k=1, \dots, m,$$

m étant le nombre de mesures.

- on définit un histogramme H par l'ensemble des couples $(I^k, z_k), k=1, \dots, m$.
- \mathcal{H} l'ensemble des histogrammes définis plus haut sur $[a,b]$.
- $\mathcal{C}^1[a,b]$ l'espace des fonctions réelles absolument continues

définies sur l'intervalle $[a,b]$ et telles que leur dérivée première soit de carré sommable sur cet intervalle.

$$\int_a^b f'^2(t) dt < +\infty$$

Cet ensemble est un espace de Hilbert, avec le produit scalaire:

$$\langle f, g \rangle = \sum_{i=0}^{i-1} \int_a^b f^{(i)}(t)g^{(i)}(t)dt$$

et la norme associée:

$$\|f\|^2 = \langle f, f \rangle$$

- π_2 est le sous espace de $\mathcal{C}^1[a,b] \cap \mathcal{C}^1[a,b]$ constitué d'éléments dont la restriction à chaque $I_i, i=1, \dots, n-1$ est un polynôme de degré deux. π_2 est de dimension $n+1$.
- \mathcal{P} est l'ensemble des éléments de π_2 dont la dérivée en a et en b est nulle.

$$\mathcal{P} = \{s \in \pi_2 \mid s'(a) = s'(b) = 0\}$$

La dimension de \mathcal{P} est $n-1$.

On nomme les éléments de \mathcal{P} splines quadratiques et les contraintes $s'(a)=0$ et $s'(b)=0$ conditions de bout naturelles.

- \mathcal{P}_k est l'ensemble des polynômes de degré k .
- Soit K_i les fonctionnelles linéaires suivantes définies sur $\mathcal{C}^1[a,b]$:

$$\langle K_i, x \rangle = \int_a^b |x(t)| dt \quad i=1, \dots, m$$

- $I_z = \{x \in \mathcal{C}^1[a,b] \mid \langle K_i, x \rangle = z_i \quad i=1, \dots, m\}$
- $(e)_+ = e$ si $e > 0$, 0 sinon.
- $\chi_{[a,b]}(t) = 1$ si $t \in [a,b]$, 0 sinon.

3) Problème de minimisation

On rappelle que notre problème, consiste à reconstruire à partir d'un histogramme donné de \mathcal{H} une fonction aussi lisse que possible dans I_z .

Soit f un élément de $\mathcal{X}^1[a,b]$ et $E(f)$ le critère de lissage :

$$E(f) = \int_a^b f'^2(t) dt$$

Considérons le problème d'interpolation suivant:

$$(P.I.1) \quad \begin{aligned} & \text{minimiser } (E(f)) \\ & f \in \mathcal{X}^1[a,b] , \quad \langle K_i, f \rangle = z_i , \quad i=1, \dots, m \end{aligned}$$

Ce problème est étudié par P.J.Laurent [5] dans un contexte plus général. En se servant des théorèmes d'existence et d'unicité, on peut affirmer facilement qu'il existe une et une seule solution de (P.I.1) et que celle-ci se trouve dans \mathcal{P} .

Existence

Il suffit de voir que l'opérateur d défini sur $\mathcal{X}^1[a,b]$ et qui fait correspondre à un élément f de $\mathcal{X}^1[a,b]$ sa dérivée f' , a un noyau de dimension fini. Cette condition est suffisante pour affirmer l'existence d'une solution de (P.I.1).

Unicité

Le noyau de d est engendré par les fonctions constantes c . Or

$$\langle K_i, c \rangle = 0 \quad , \quad i=1, \dots, m \text{ et } m \geq 1$$

entraîne que c est nulle sur $[a,b]$.

Ceci est résumé dans le théorème suivant:

Théorème:

Le problème (P.I.1) admet une et une seule solution.

Caractérisation

Dans toute la suite de ce chapitre, on suppose que pour un z donné, I_z est non vide et que $m \geq 1$.

Nous avons aussi d'après [5] que σ appartenant à I_Z est solution de (P.I.1) si et seulement s'il existe des coefficients (λ_i) , $i=1, \dots, m$ tels que

$$(I.1) \quad \sum_{k=1}^m \lambda_k \langle k_k, c \rangle = 0, \text{ pour tout } c \in \mathbb{R}$$

$$(I.2) \quad \sigma'(t) = \sum_{k=1}^m \lambda_k ((T_{jk} - t)_+ - (T_{jk} - t)_+)$$

que l'on peut écrire encore

$$\sigma'(t) = \sum_{k=1}^m \lambda_k ((t - T_{jk})_+ - (t - T_{jk})_+)$$

En vertu de (I.2), on peut caractériser la spline σ par :

- 1) σ, σ' sont continues sur $[a, b]$.
- 2) La restriction de σ sur chaque intervalle simple $I_i = [T_i, T_{i+1}]$, $i=1, \dots, n-1$ est un polynôme de degré inférieur ou égal à 2.
- 3) Il existe $\lambda \in \mathbb{R}^m$ tel que $\sigma''(t) = -\sum_{k=1}^m \lambda_k \chi_{I^k}(t)$
- 4) $\sigma'(a^+) = \sigma'(b^-) = 0$

Nous signalons que pour certains cas particuliers le degré de σ se réduit à 1 dans certains intervalles. Cette restriction est due à la structure de l'histogramme des mesures. Nous allons étudier dans la suite deux classes différentes d'histogrammes et essayer d'élaborer des algorithmes de calcul de σ .

4) Construction numérique et algorithme de résolution

4.1) Histogramme à intervalles juxtaposés

a- Conditions de bout naturelles

Considérons le problème d'interpolation suivant:

$$(P.1.2) \quad \begin{aligned} & \text{minimiser} \left(\int_a^b f^2(t) dt \right) \\ & f \in \mathcal{H}^1[a,b], \quad \int_{I_i} f(t) dt = w_i, \quad i=1, \dots, n-1 \end{aligned}$$

La solution (P.1.2) est l'unique élément σ , qui est dans \mathcal{P} et qui vérifie les $n-1$ contraintes d'interpolation.

Calcul de σ :

On note $\sigma(T_i) = y_i, i=1, \dots, n$. Sur l'intervalle $[T_i, T_{i+1}]$, σ est un polynôme de degré deux. En exprimant la primitive de σ sous la forme de Newton avec des noeuds confondus $(T_i, T_i, T_{i+1}, T_{i+1})$, on obtient $\sigma(t)$ par dérivation:

$$(1.3) \quad \sigma(t) = y_i + \lambda_i (t-T_i) + u_i (t-T_i)^2 / 2 \quad t \in I_i, i=1, \dots, n-1$$

$$\begin{aligned} \text{où} \quad \lambda_i &= 2(3w_i/h_i - 2y_i - y_{i+1}) / h_i \\ u_i &= 6(y_i + y_{i+1} - 2w_i/h_i) / h_i^2 \end{aligned}$$

Il s'avère, que la connaissance du vecteur y (en plus des w_i) est suffisante pour la détermination de $\sigma(t)$ dans chaque intervalle.

D'après la continuité de σ' en chaque noeud T_i , nous avons les $n-2$ relations:

$$\sigma'(T_i^+) = \sigma'(T_i^-) \quad i = 2, \dots, n-1$$

Le développement de ces relations, plus les deux conditions de bouts naturelles, donnent le système suivant:

$$2y_1 + y_2 = 3w_1/h_1$$

$$(1.4) \quad (h_i/h_{i-1})y_{i-1} + 2(1 + h_i/h_{i-1})y_i + y_{i+1} = 3(w_i/h_i + w_{i-1}h_i/h_{i-1}^2) \quad i=2, \dots, n-1$$

$$y_{n-1} + 2y_n = 3w_{n-1}/h_{n-1}$$

La matrice correspondante à ce système est tridiagonale et à diagonale dominante.

La résolution de (1.4) est réalisée par l'élimination de Gauss, en triangularisant (1.4).

Nous allons dans la suite développer les calculs de résolution de (1.4) dans le cas particulier à pas constants.

Abscisses équidistantes ($h_i = h, i=1, \dots, n-1$)

Le système (1.4) s'écrit dans ce cas:

$$(1.5) \quad \begin{aligned} 2y_1 + y_2 &= 3w_1/h & = b_1 \\ y_{i-1} + 4y_i + y_{i+1} &= 3(w_i + w_{i-1})/h = b_i & i=2, \dots, n-1 \\ y_{n-1} + 2y_n &= 3w_{n-1}/h & = b_n \end{aligned}$$

Nous remarquons que la matrice du système linéaire (1.5) est la même que celle des splines cubiques d'interpolation.

En effet soit s une primitive de σ alors

$$s'(T_i) = \sigma(T_i) = y_i$$

$$s(T_{i+1}) - s(T_i) = \int_{T_i}^{T_{i+1}} \sigma(t) dt = w_i \quad \text{pour } i=1, \dots, n-1.$$

Par ce changement de variable le système (1.5) sera exactement identique à celui qui permet de calculer la spline cubique d'interpolation avec les mesures ponctuelles $s(T_i)$, que l'on exprimera à une constante près par:

$$s(T_i) = \sum_{k=1}^{i-1} \int_{T_k}^{T_{k+1}} \sigma(t) dt + s(T_1)$$

La triangularisation de (1.5), nous donne des relations de la forme:

$$(1.6) \quad \alpha_i y_i + \beta_i y_{i+1} = \gamma_i$$

$$\alpha_1 = 1, \quad \beta_1 = 2, \quad \gamma_1 = b_1$$

$$\alpha_{i+1} = -4\alpha_i + \beta_i$$

$$\beta_{i+1} = -\alpha_i$$

$$\gamma_{i+1} = \gamma_i - \alpha_i b_{i+1}$$

La suite (α_i, β_i) peut s'écrire sous la forme matricielle:

$$(1.7) \quad {}^t(\alpha_{i+1}, \beta_{i+1}) = A {}^t(\alpha_i, \beta_i)$$

$$A = \begin{bmatrix} -4 & 1 \\ 1 & 0 \end{bmatrix}$$

Le rapport (β_i/α_i) " converge " vers le rapport (x_2/x_1) des composantes du vecteur propre (x_1, x_2) associé à la valeur propre du plus grand module de la matrice A.

Les valeurs propres de cette transformation sont la solution de l'équation:

$$\gamma^2 + 4\gamma + 1 = 0$$

$$\gamma = -2 + \sqrt{3} \quad \text{et} \quad \gamma = -2 - \sqrt{3}$$

$$\text{On a} \quad x_2/x_1 = -\gamma = 2 - \sqrt{3} = -1/\gamma$$

Les composantes de (1.6) croissent très vite, ce qui rendra le calcul imprécis. On propose donc de diviser par β_i à chaque itération. Cela revient à fixer β_i égal à 1; nous aurons:

$$(1.8) \quad \begin{aligned} \alpha_i &= 2, & \gamma_i &= 3w_i/h \\ \alpha_{i+1} &= 4 - 1/\alpha_i, & i &= 1, \dots, n-2 \\ \gamma_{i+1} &= b_{i+1} - \gamma_i / \alpha_i \end{aligned}$$

et (α_i) converge vers $-\gamma = 2 + \sqrt{3}$, et se stabilise (à précision machine) en 13 itérations, ce qui permet d'économiser le calcul des (α_{i+13}) .

$$\alpha_{i+13} = 3.73205087$$

Algorithme :

1- calcul de (α_i) , $i=1, \dots, 13$ et (γ_i) , $i=1, \dots, n-1$ par (1.8).

2- Résolution de

$$(2+\sqrt{3})y_{n-1} + y_n = \gamma_{n-1}$$

$$y_n + 2y_{n-1} = 3w_{n-1} / h$$

$$\text{solution } y_n = w_{n-1} / h + \gamma_{n-1} / 3.$$

3- calcul de (y_i) , $i=n-1, \dots, 1$

$$y_i = (\gamma_i - y_{i+1}) / \alpha_i$$

b- Conditions de bout "cardinales"

Le système linéaire (1.5) est analogue à celui obtenu pour les splines cubiques naturelles d'interpolation en prenant comme variables les dérivées aux noeuds $(T_i), i=1, \dots, n$.

Cela nous conduit à appliquer les mêmes techniques de résolution en modifiant les relations de bouts de la spline quadratique.

On remarque, que les équations (1.5) sont vérifiées par les polynômes de degré 3 dans \mathcal{P}_3 .

$\forall p \in \mathcal{P}_3$ on a :

$$(1.9) \quad p(T_{i+1}) + 4p(T_i) + p(T_{i-1}) = (3/h) \int_{T_{i-1}}^{T_{i+1}} p(t) dt$$

Il est convenable alors de choisir les conditions de bouts sous la même forme :

$$\alpha p(T_1) + \beta p(T_2) = b_1$$

$$\beta p(T_{n-1}) + \alpha p(T_n) = b_n$$

où b_1 et b_n sont des combinaisons linéaires des (w_i)

$$(1.10) \quad \begin{aligned} b_1 &= (1/h) \sum_{i=1}^{i=k} e_i w_i & ; \\ b_2 &= (1/h) \sum_{i=1}^{i=k} e_i w_{n-i} \end{aligned}$$

et telles qu'elles soient exactes pour \mathcal{P}_3 .

Un choix convenable du rapport (β/α) rendra la suite (β_i/α_i) stationnaire.

En fixant ce rapport à $-\gamma$ le système (I.6) s'écrit:

$$(I.11) \quad \begin{aligned} (-1/\gamma) y_1 + y_2 &= b_1 \\ (-1/\gamma) y_i + y_{i+1} &= \gamma_i \end{aligned}$$

où $\gamma_i = \gamma \gamma_{i-1} + b_i \quad i = 2, \dots, n-1$

Le plus petit entier k , que l'on puisse prendre est 4, et les coefficients (e_i) sont alors égaux à :

$$(I.12) \quad \begin{aligned} e_1 &= (53 + 25\sqrt{3}) / 12 \\ e_2 &= -(33 + 23\sqrt{3}) / 12 \\ e_3 &= (21 + 13\sqrt{3}) / 12 \\ e_4 &= -(5 + 3\sqrt{3}) / 12 \end{aligned}$$

Algorithme 2

1- calcul de b_1 et b_n

2- calcul des γ_i par (I.11)

3- résolution du système

$$(-1/\gamma) y_{n-1} + y_n = \gamma_{n-1}$$

$$y_{n-1} + (-1/\gamma) y_n = b_n$$

solution $y_n = \sqrt{3} (b_n - (2-\sqrt{3})\gamma_{n-1}) / 6.$

4- calcul des (y_i) , $i = n-1, \dots, 1$

$$y_i = \gamma(y_{i+1} - \gamma_i).$$

4.2) Histogramme à intervalles chevauchés

Nous avons traité, jusqu'à présent, des intervalles adjacents de mesures, ce qui confère au système associé la propriété d'être tridiagonal. Dans cette partie, nous allons aborder le même problème, mais, dans le cas d'intervalles qui se chevauchent de façon quelconque.

4.2.1) Représentation des histogrammes par des graphes

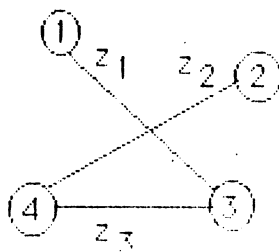
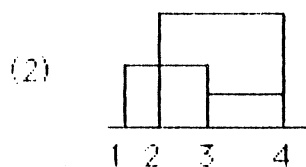
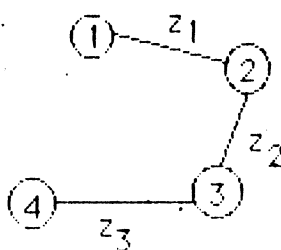
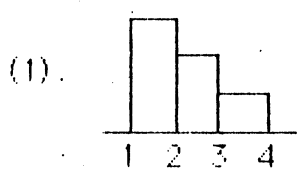
Avant d'aborder le problème du calcul de σ , nous allons essayer de déceler certaines propriétés des histogrammes qu'on ne peut pas voir directement à partir de leur représentation précédente.

Nous suggérons une autre façon de représenter les mesures à valeur moyenne, équivalente, mais plus commode, pour élucider certaines propriétés des histogrammes.

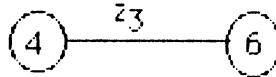
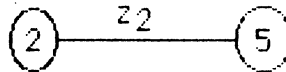
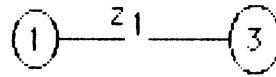
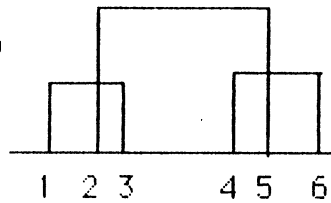
Notations:

Soit $G(n,T,m,A,Z)$ un graphe défini par un ensemble T de n noeuds $(T_i), i=1, \dots, n$. Ces noeuds sont reliés par un ensemble A d'arcs et chaque arc est affecté d'un poids ou valeur de Z . m est le nombre d'arcs de A .
On voit que chaque élément de \mathcal{H} peut être représenté par un graphe.

Exemples:



(3)



L'objet de cette représentation est l'introduction de certains termes et outils de la théorie des graphes comme chaîne, cycle, connexe, etc.. pour interpréter certaines opérations algébriques (Base, dépendance, redondance, ...).

Les deux représentations sont équivalentes, et le passage de l'un à l'autre est évident.

Rappel:

- chaîne: Soit T_i et T_j deux noeuds de T . Une chaîne joignant T_i et T_j dans G est une séquence d'arcs adjacents. Le premier arc de la séquence est adjacent à T_i et le dernier à T_j .

- Cycle: C'est une chaîne dont les extrémités coïncident ($T_i = T_j$).

- Connexe: Un graphe est connexe si pour tous couples (T_i, T_j) de T il existe une chaîne qui les joint.

Propriétés:

Si G est un graphe connexe alors:

1- m est supérieur ou égal à $n-1$.

2- Il existe au moins un arbre de G (sous-graphe connexe constitué de $(n-1)$ arcs). Cet arbre est obtenu en éliminant les cycles de G .

La démonstration de ces propriétés est immédiate.

Revenons maintenant au calcul de σ . Nous rappelons que σ est formée, de morceaux de polynômes de degré deux, sur chaque intervalle I_i , $i=1, \dots, n-1$

et pour la calculer il suffit de connaître ses valeurs moyennes (W_i), $i=1, \dots, n-1$ sur les intervalles I_i , $i=1, \dots, n-1$ et ses valeurs (y_i) en chaque noeuds (T_i), $i=1, \dots, n$.

L'idée, est de définir certaines hypothèses sur l'histogramme de mesure ou son graphe correspondant pour, d'une part assurer un calcul direct des (w_i) et d'autre part utiliser le procédé de calcul des (y_i) décrit dans 4.1.

4.2.2) Graphe connexe

Soit H un histogramme de f et supposons que les deux hypothèses suivantes soient vérifiées:

$$1- I_Z \neq \emptyset$$

2- Son graphe G est connexe.

Les mesures (Z_k) s'expriment en fonction des w_i $i=1, \dots, n$ sous forme matricielle par:

$$(1.13) \quad EW = Z$$

E une matrice $m \times n-1$ et $E_{k,l} = 1$ si $T_l \in [T_{j^k}, T_{j^k-1}]$, 0 sinon.

Chaque ligne de E correspond à une mesure et chaque colonne de E correspond à un intervalle de base $I_i = [T_i, T_{i+1}]$.

Les matrices E correspondant aux trois graphes précédents sont:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{et} \quad E_3 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

En effet, chaque mesure z_k s'écrit:

$$\int_{I^k} f(t) dt = \sum_{e=I^k} \int_{I^e} f(t) dt = z_k, \quad k=1, \dots, m$$

Les propriétés relatives à un graphe connexe se traduisent algébriquement par:

- 1- Le nombre de lignes de E est supérieur au nombre de colonnes.
- 2- Les valeurs moyennes non mesurées, peuvent être obtenues par des combinaisons linéaires simples des autres mesures.
- 3- le rang de la matrice E est n-1.

La première hypothèse signifie qu'il existe au moins une solution de (I.13). Nous allons montrer que la seconde hypothèse signifie qu'il existe un et un seul vecteur W de \mathbb{R}^{n-1} solution de (I.13).

Soit φ la fonction définie de \mathbb{R}^{n-1} dans \mathbb{R}^m telle que:

$$(I.14) \quad \varphi(w) = Ew$$

Théorème

Soit H un histogramme, G son graphe, E la matrice correspondante et φ la transformation linéaire qui lui est associée. Alors G est connexe si et seulement si φ est injective.

corollaire1: G est connexe $\Leftrightarrow {}^tEE$ est bijective.

démonstration:

- G est connexe $\Rightarrow {}^tEE$ est inversible

${}^tEEw = 0 \Rightarrow {}^t(Ew)Ew = 0 \Rightarrow Ew = 0$ et d'après le théorème $W = 0$ d'où tEE est bijective.

- tEE est inversible \Rightarrow G est connexe

$Ew = 0 \Rightarrow {}^tEEw = 0$ et comme tEE est inversible alors w est nul ce qui implique que E est injective et d'après le théorème G est connexe.

corollaire2: Si les deux hypothèses (1) et (2) sont vérifiées alors il existe un et un seul vecteur w dans \mathbb{R}^{n-1} tel que $Ew = Z$.

($I_Z \neq \emptyset \Rightarrow$ l'existence et E injective \Rightarrow l'unicité)

tEE est une matrice carré symétrique dont les coefficients sont des entiers positifs.

$({}^tEE)_{e1} =$ le nombre d'intervalles $I^k = [T_{1k}, T_{jk}]$, tels que $i^k \leq e$ et $1 < j^k$.

Finalement on résume sous les hypothèses (1) et (2) que le problème de minimisation (P.I.1) se ramène (pour $n > 1$) à:

(P.1.3) *minimiser* (E(f))

$$f \in \mathcal{X}^1[a,b], \quad (({}^tEE)^{-1} {}^tEZ)_i = \int_{I_i} f(t) dt, \quad i=1, \dots, n-1$$

On pose $w = (({}^tEE)^{-1} {}^tEZ)$

On voit facilement que l'algorithme de résolution se décompose alors en deux étapes.

Algorithme:

- 1- Résolution du système linéaire $Ew = Z$.
- 2- Application de l'Algorithme 1 en prenant comme données W_i $i=1, \dots, n-1$ pour calculer les y_i $i=1, \dots, n$.

On voit deux méthodes différentes de résolution de (I.13) :

La première se base sur la transformation du graphe G par les opérations définies plus haut. L'idée est de constituer à partir de G la chaîne joignant les noeuds 1 et n-1 et passant une et une seule fois d'une façon croissante par tous les noeuds compris entre 1 et n-1.

L'avantage de ce procédé est qu'il fait intervenir uniquement les opérations d'additions; par contre il ne permet pas d'affirmer que le vecteur W obtenu est la bonne solution de (I.13). Autrement dit, si les mesures z sont bruitées et que $I_z = \emptyset$, ce procédé donnera toujours une solution.

La deuxième méthode est la résolution directe du système (${}^tEEw = {}^tEZ$). (voir la méthode de Gauss, Jordan ou Cholesky).

Remarques:

1- Supposons que z soit perturbé ou bruité - ce cas sera traité dans le chapitre des splines quadratiques d'ajustement- ce qui détruit la validité de l'hypothèse 1 ($I_z \neq \emptyset$). Alors, la solution correspondante au système (${}^tEEw = {}^tEZ$) semble convenable.

En effet, il est préférable dans ce cas de chercher le vecteur z^* le plus proche de z dans $\text{Im}(\varphi)$, au sens des moindres carrés et de le prendre comme vecteur de mesure.

$$(P.1.4) \quad \begin{aligned} & \textit{minimiser} \ \|x-z\| \\ & x \in \text{Im}(\varphi) \end{aligned}$$

La solution de (P.I.4) est la solution du système suivant:

$${}^tEE w = {}^tEz \quad \text{et} \quad z^* = Ew$$

Ce qui, d'une part évite de vérifier l'hypothèse (1) et d'autre part, de calculer directement le vecteur W.

2- Résolution matricielle:

$$Ew = z \quad \text{et} \quad Ay = Bw \quad \Rightarrow \quad y = A^{-1}B({}^tEE)^{-1} {}^tEz$$

A (respectivement B) est la matrice du premier membre (respectivement du seconde membre) du système linéaire (I.15).

4.2.3 Graphe non connexe ou histogramme quelconque

Les méthodes numériques décrites plus haut découlent et dépendent essentiellement de la structure des intervalles de mesure qui confèrent certaines simplifications aux algorithmes de calcul. En fait, on peut toujours, indépendamment de cette structure, décrire un algorithme général de construction. (Voir dans [5] la méthode de projection ou bien les méthodes utilisant une base des splines). En pratique, et pour ce type d'histogramme, (non connexe), cela conduit à des erreurs inacceptables. La méthode de projection consiste à reporter les conditions d'interpolation sur la dérivée de la spline.

Supposons que les fonctionnelles $(K_i, i=1, \dots, m)$ soient linéairement indépendantes (on peut toujours se ramener à ce cas, en éliminant les mesures redondantes représentées par des cycles dans le graphe G). Il en résulte que $m < n-1$.

soit f une fonction de $\mathcal{C}^1[a,b]$:

$$f(x) = f(a) + \int_{[a,x]} f'(t) dt = f(a) + \int_{[a,b]} \chi_{[a,x]}(t) f'(t) dt$$

Soit l'opérateur L_k défini par

$$L_k = c_k K_k + d_k K_{k+1} \quad \text{et} \quad \langle L_k, c \rangle = 0 \quad \forall c \in \mathbb{R}$$

Pour que cette dernière condition soit vérifiée il suffit de prendre $c_k = H^{k+1}$ et $d_k = -H^k$ ($H^k = T_{j,k} - T_{i,k}$).

En notant la fonction $\mathfrak{X}_t(x) = \chi_{[a,x]}(t)$ et en appliquant L_k à f on obtient:

$$(1.15) \quad \langle L_k, f \rangle = \int_{[a,b]} \langle L_k, \mathfrak{X}_t \rangle f'(t) dt = v_k$$

où $v_k = H^{k+1}z_k - H^k z_{k+1}$, $k=1, \dots, m-1$

On pose $\varphi_k(t) = \langle L_k, \mathfrak{X}_t \rangle$. Le problème se ramène alors à chercher la dérivée σ' de norme $E(\sigma)$ minimum et vérifiant les conditions:

$$(1.16) \quad \int_{[a,b]} \varphi_k(t) \sigma'(t) dt = v_k, \quad k=1, \dots, m-1$$

Notons F l'espace (de dimension $m-1$) engendré par les φ_k . Soit F^\perp le complémentaire orthogonal (dans $\mathfrak{H}^*[a,b]$) de F . C'est l'ensemble des g de $\mathfrak{H}^*[a,b]$ tels que:

$$(1.17) \quad \int_{[a,b]} \varphi_k(t) g(t) dt = 0, \quad k=1, \dots, m-1$$

L'ensemble Δ des g de $\mathfrak{H}^*[a,b]$ vérifiant les conditions (1.17) est un translaté de F^\perp . L'élément de Δ qui réalise le minimum est la projection de l'origine sur Δ . On doit donc avoir $\sigma' \in \Delta$ et $\sigma' \in F$.

Il en résulte que:

$$\sigma'(t) = \sum_{k=1}^{m-1} \mu_k \varphi_k(t)$$

Pour obtenir les μ_k on exprime que σ' vérifie les conditions (1.17).

d'où

$$\sum_{i=1}^{m-1} \mu_i \int_{[a,b]} \varphi_i(t) \varphi_k(t) dt = v_k, \quad k=1, \dots, m-1$$

C'est un système linéaire de $m-1$ équations à $m-1$ inconnues. On ne peut dire grand chose sur ses propriétés, à partir de la collection des intervalles $(I^k, k=1, \dots, m)$. Par la résolution de ce système, nous obtenons l'expression de $\sigma'(t)$, en particulier, les dérivées de σ en chaque point (T_l) .

Pour déterminer complètement la spline σ , outre le vecteur des dérivées, il suffit de calculer le vecteur des valeurs de σ en chaque point $(T_l, l=1, \dots, n)$.

Par intégration et utilisation d'une seule condition d'interpolation on peut avoir complètement σ .

Lors du calcul des $\sigma'(T_i)$ et $\sigma(T_i)$, σ s'écrit pour $t \in [T_i, T_{i+1}]$:

$$\sigma(t) = \sigma(T_i) + \sigma'(T_i)(t - T_i) + (\sigma'(T_{i+1}) - \sigma'(T_i))(t - T_i)^2 / 2h$$

Remarque:

La méthode que nous venons de décrire donne en pratique des résultats satisfaisants; nous voulons dans cette remarque, présenter les grandes lignes de la méthode utilisant la base canonique des splines quadratiques. Soit $(\sigma_i), i=1, \dots, n-1$ la base canonique des splines quadratiques (σ_i est un élément de \mathcal{P} tel que son intégrale est égal à 1 sur I_i et 0 sur $I_j, j \neq i$; dans [3] on l'appelle fonction deltaspline). Pour tout élément σ de \mathcal{P} (ensemble des splines quadratiques) on a:

$$\sigma = \sum_{i=1}^{n-1} w_i \sigma_i$$

où w_i est la valeur moyenne de σ sur I_i . Il en résulte:

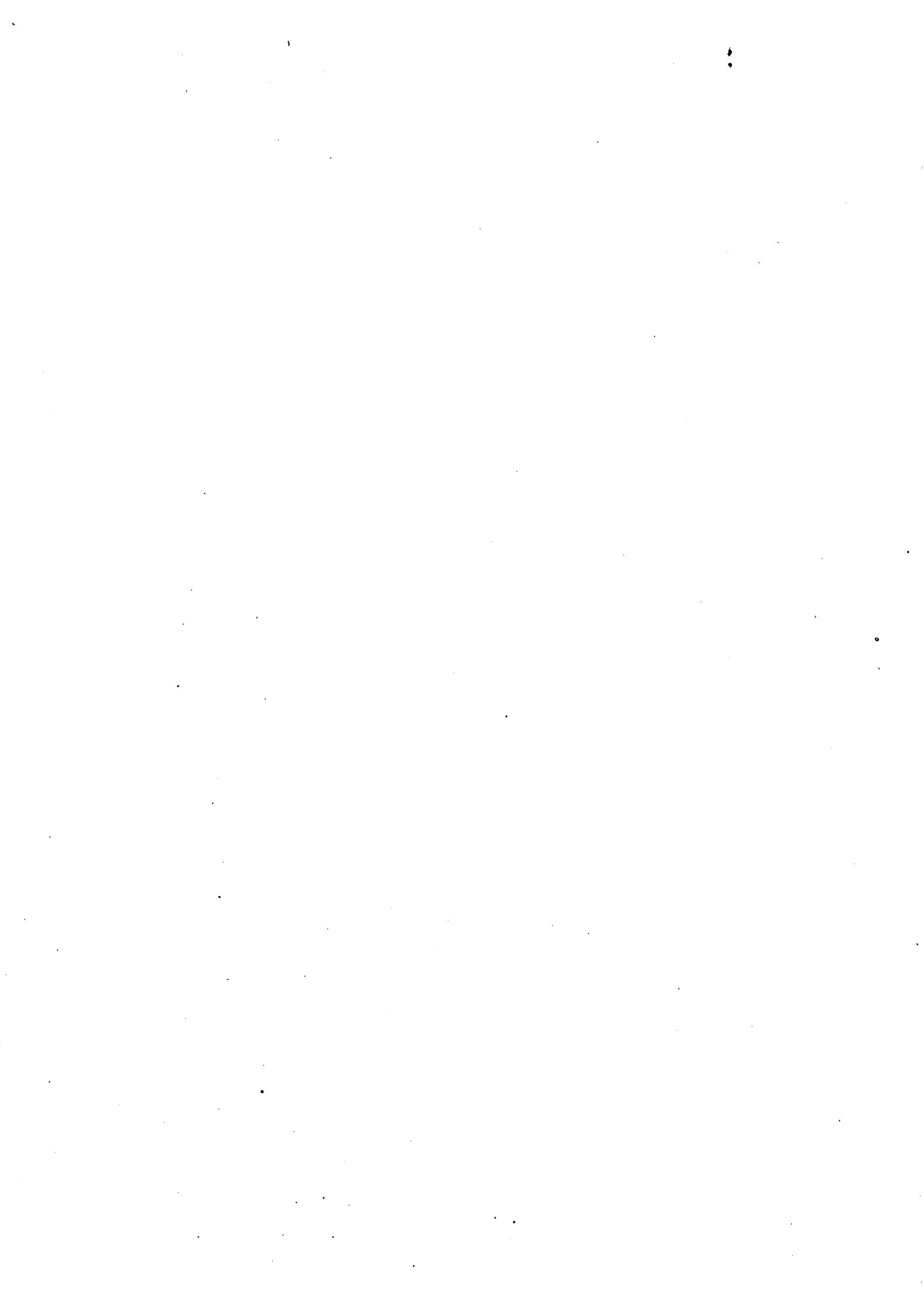
$$E(\sigma) = \langle w, \Omega w \rangle \quad \text{et} \quad Ew = z$$

Ω est une matrice carré $(n-1) \times (n-1)$ (voir le chapitre 2).

D'où le problème de minimisation:

$$(P.1.5) \quad \begin{aligned} & \text{minimiser } \langle w, \Omega w \rangle \\ & w \in \mathbb{R}^{n-1}, \quad Ew = z \end{aligned}$$

C'est un problème d'optimisation quadratique dans un sous espace de \mathbb{R}^{n-1} . L'avantage de (P.1.5) est qu'il permet de calculer directement w . Nous pensons à la méthode de gradient conjuguée avec des contraintes linéaires. C'est un procédé efficace; il a l'avantage de converger en un nombre fini d'itérations. En dépit de cela, cette méthode reste coûteuse par rapport à la précédente.



Chapitre 2

Spline quadratique d'ajustement

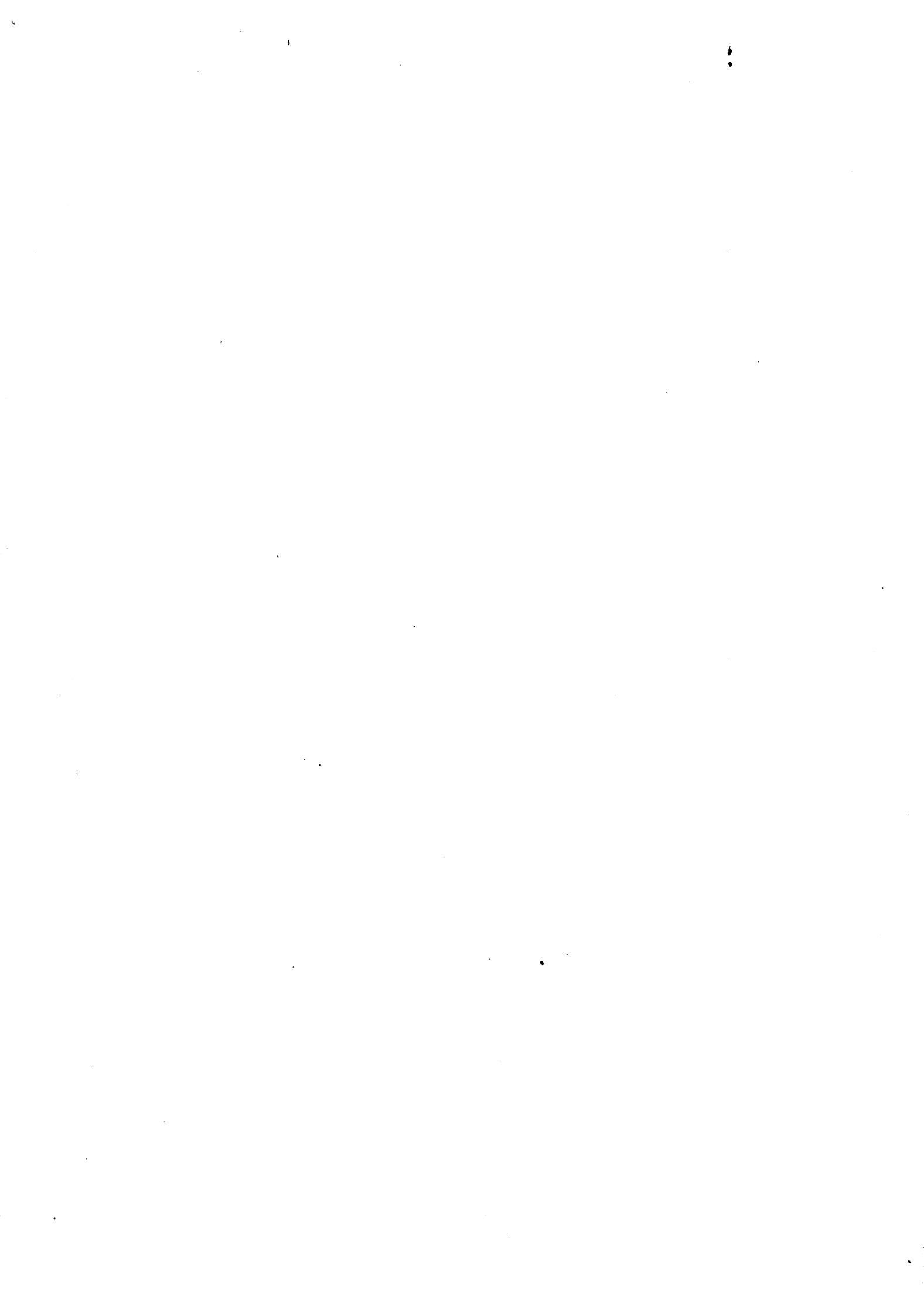
1-Introduction

2-Construction numérique et algorithme de résolution

2.1-Histogramme à intervalles quelconques

2.2-Histogramme à intervalles adjacents

3-Conclusion



1) Introduction

En pratique, les mesures expérimentales sont entachées d'erreurs d'où l'introduction d'un terme de contrôle d'ajustement, représentant l'écart global entre les mesures expérimentales et ajustées. Ce terme est affecté d'un poids ρ permettant de doser le lissage et l'ajustement des données.

On conserve les mêmes notations définies dans (1.2). On note (z_i) , $i=1, \dots, m$ les valeurs mesurées et on considère le problème de minimisation suivant:

$$E(f) = \int_a^b f'^2(t) dt, \quad J(f) = \sum_{k=1}^m \left(\int_{t_k} f(t) dt - z_k \right)^2$$

$$(P.II) \quad \text{minimiser } (E(f) + \rho J(f))$$

$$f \in \mathcal{X}^1[a,b]$$

Les conditions d'existence et d'unicité étant assurées, l'espace des splines quadratiques d'ajustement est le même que celui des splines quadratiques d'interpolation. On retrouvera la même caractérisation de (1.3), seules les conditions d'interpolation seront remplacées par les conditions d'ajustement définies dans [5].

Dans ce chapitre on distingue aussi les deux classes d'histogrammes: adjacents et chevauchés. Des algorithmes de calcul de σ sont étudiés.

Dans la première partie nous évoquerons le problème dans sa généralité pour un histogramme quelconque et dans la deuxième partie, on se restreint au cas simple des histogrammes adjacents vu la simplification des calculs qu'ils introduisent.

Il en résulte:

$$(11.4) \quad (M^t E E - L / \omega) w = M^t E z$$

où $\omega = \rho h^3 / 6$ et $\rho > 0$

La matrice correspondant au système (11.4) est symétrique. Nous allons montrer que l'on peut retrouver ces relations à partir d'une formulation matricielle préalable du problème de minimisation (P.II).

Formulation matricielle

Soit deux fonctionnelles Ψ et Φ définies sur \mathbb{R}^{n-1} :

- $\Phi(w) = \min (E(f)) / f \in \mathcal{X}^1[a,b] , \int_{11} f(t) dt = w_i, i=1, \dots, n-1$
- $\Psi(w) = \langle Ew - z, Ew - z \rangle$

On voit facilement que (P.II) se ramène à chercher un w de \mathbb{R}^{n-1} solution du problème suivant:

$$\underset{x \in \mathbb{R}^{n-1}}{\text{minimiser}} (\Phi(x) + \rho \Psi(x))$$

En notant $(\sigma_i), i=1, \dots, n-1$ la base canonique des splines quadratiques.

$$\int_{11} \sigma_j(t) dt = \delta_{ij}, \quad i, j=1, \dots, n-1$$

La spline σ se décompose en fonction des σ_i :

$$\sigma = \sum_{i=1}^{n-1} w_i \sigma_i$$

et

$$\Phi(w) = \langle w, \Omega w \rangle$$

où Ω est une matrice symétrique semi-définie positive:

$$\Omega_{i,j} = \int_a^b \sigma_i'(t) \sigma_j'(t) dt, \quad i, j=1, \dots, n-1$$

La fonctionnelle $\Phi + \rho \Psi$ est une fonction quadratique dérivable. La solution du (P.II) est caractérisée par:

$$\text{grad} (\Phi(w) + \rho \Psi(w)) = 0$$

qui s'écrit encore

$$\Omega w + \rho^t E(Ew - z) = 0$$

La matrice Ω dépend uniquement du pas de mesure. Le calcul direct de ses coefficients $\Omega_{i,j}$ est délicat. Par contre il suffit de voir que le vecteur Ωw n'est autre que le vecteur des dérivées secondes affecté du signe moins.

$$(11.5) \quad -\Omega w = U$$

Ce qui permet de retrouver les relations de (11.2), puis celles de (11.4). Nous résumons le calcul de σ dans l'algorithme suivant qui se décompose en la résolution de deux systèmes linéaires.

Algorithme 4

- 1- Résolution de (11.4) et calcul de w .
- 2- Application de l'Algorithme en prenant w comme vecteur de données.

2.2) Histogramme à intervalles adjacents

Les mesures ($z_i, i=1, \dots, n-1$) se font sur des intervalles contigus, la matrice se réduit à la matrice identité (carré $(n-1) \times (n-1)$). Cela rend les relations (11.1) et (11.4) plus simples et nous permet de déterminer directement un système en fonction des y_i (valeurs de σ en $T_i, i=1, \dots, n$). Les relations (11.2) s'écrivent:

$$(11.6) \quad \sigma''(t) = \rho \left(\int_{I_i} \sigma(t) dt - z_i \right) \quad t \in [T_i, T_{i+1}] \quad \text{et } i=1, \dots, n-1$$

En développant la dérivée seconde à partir de (1.1):

$$(11.7) \quad y_{i+1} + y_i - 2w_i / h_i = \rho h_i^3 (w_i / h_i - z_i / h_i) / 6 \quad i=1, \dots, n-1$$

ce qui permet d'exprimer w_i / h_i :

$$(11.8) \quad w_i / h_i = (y_{i+1} + y_i + \omega_i z_i / h_i) / (2 + \omega_i)$$

$$\omega_i = \rho h_i^3 / 6 \quad i=1, \dots, n-1$$

En remplaçant $w_i/h_i + w_{i-1}h_i/h_{i-1}^2$ dans (I.2) nous obtenons le système suivant dans le cas d'abscisses équidistantes.

$$(II.9) \quad \begin{aligned} (2+1/\omega)y_1 + (1-1/\omega)y_2 &= 3z_1/h \\ (1-1/\omega)y_{i-1} + 2(2+1/\omega)y_i + (1-1/\omega)y_{i+1} &= 3(z_i + z_{i-1})/h \\ (1-1/\omega)y_{n-1} + (2+1/\omega)y_n &= 3z_{n-1}/h \end{aligned}$$

où $\omega = \rho h^3/6$

Nous remarquons que :

- le système (II.9) est tridiagonal, symétrique et analogue à (I.2).
- Quand ω tend vers l'infini la solution de (II.9) tend vers celle de (I.2). En effet la spline quadratique d'ajustement converge vers la spline d'interpolation.

L'élimination de Gauss se traduit par les relations suivantes:

$$(II.10) \quad \alpha_i y_i + \beta_i y_{i+1} = \gamma_i \quad i = 1, \dots, n-1$$

$$\alpha_1 = 2 + 1/\omega, \quad \beta_1 = 1 - 1/\omega$$

$$\gamma_1 = 3z_1/h, \quad \theta = (4\omega + 2) / (\omega - 1)$$

$$\alpha_{i+1} = -\theta\alpha_i + \beta_i$$

$$\beta_{i+1} = -\alpha_i$$

$$\gamma_{i+1} = \gamma_i - \alpha_i \beta_i / \beta_i$$

La suite (α_i, β_i) s'écrit sous la forme matricielle

$${}^t(\alpha_{i+1}, \beta_{i+1}) = A(\omega) {}^t(\alpha_i, \beta_i)$$

$$A(\omega) = \begin{bmatrix} -\theta & 1 \\ 1 & 0 \end{bmatrix}$$

Les valeurs propres de $A(\omega)$ sont les racines de l'équation:

$$\gamma^2 + \theta\gamma + 1 = 0$$

$$\gamma = (-\theta + \sqrt{(\theta^2 - 4)}) / 2 \quad \text{et} \quad \gamma' = (-\theta - \sqrt{(\theta^2 - 4)}) / 2$$

le rapport (β_i / α_i) converge vers (x_2/x_1) et (x_1, x_2) est le vecteur propre associé à la valeur propre du plus grand module de la matrice $A(\omega)$.

On a:

$$x_2 / x_1 = -\gamma$$

et si on fixe β_i à $\beta = (1 - 1/\omega)$, α_i converge vers $-\gamma\beta$.

Les équations (II.9) deviennent:

$$\begin{aligned} \alpha_1 y_1 + \beta y_2 &= 3z_1/h \\ \text{(II.11)} \quad \alpha_i y_i + \beta y_{i+1} &= \gamma_i \quad i=2, \dots, n-1 \\ \beta y_{n-1} + \alpha_1 y_n &= 3z_{n-1}/h \end{aligned}$$

où

$$\begin{aligned} \alpha_i &= (4\omega + 2)/\omega - \beta^2 / \alpha_{i-1} \\ \text{(II.12)} \quad \gamma_{i+1} &= b_{i+1} - \gamma_i \beta / \alpha_i \end{aligned}$$

STABILITE DE LA SUITE (α_i)

- Pour $(\omega > 1)$ la suite se stabilise après une dizaine d'itérations (pour une précision machine de 8 chiffres).

- Pour $\omega = 1$ la suite est stationnaire, et la solution du (II.11) est directe:

$$\begin{aligned} \text{(II.13)} \quad y_1 &= z_1 / h & w_1 &= (5z_1 + z_2)/6 \\ y_i &= (z_i + z_{i-1}) / 2h \quad \text{et} & w_i &= (z_{i+1} + 4z_i + z_{i-1}) / 6 \\ y_n &= z_{n-1} / h & w_{n-1} &= (w_{n-2} + 5w_{n-1}) / 6 \end{aligned}$$

W correspond à la convolution de z et d'un filtre linéaire invariant en position.

- Mais pour ω très faible, la suite (α_i) est instable et nécessite, un nombre très grand d'itérations pour atteindre son point de convergence.

Algorithme 5

1- résolution de (II.11) même structure que l'Algorithme 1

2- calcul des w_i $i = 1, \dots, n-1$ par (13):

$$w_i = h(y_{i+1} + y_i + \omega z_i/h) / (2 + \omega)$$

3) Conclusion

Le système qu'on vient de résoudre a les mêmes propriétés que celui obtenu par les splines d'interpolation. En conséquence, nous avons pu appliquer le même procédé de résolution. Cette simplification résulte du type de mesures qu'on rencontrera aussi dans la construction des splines quartiques d'ajustement (voir chapitre 3).



Chapitre 3

Spline quartique d'interpolation et d'ajustement

- 1- Introduction
- 2- Spline quartique d'interpolation
 - 2.1-Notations et définitions
 - 2.2-Construction numérique
- 3- Spline quartique d'ajustement



1) Introduction

Plusieurs exemples de mesure à valeur moyenne ont été traités, il s'avère que l'approximation par les splines quadratiques est satisfaisante. Cela nous a conduit à développer deux progiciels (chapitre 5) se basant sur ce modèle. Néanmoins, nous allons étudier dans ce chapitre un modèle dont le degré de continuité est supérieur: les splines quartiques.

Dans la première section, nous présenterons l'algorithme de construction d'une spline quartique d'interpolation, en précisant certains calculs préliminaires concernant les conditions de bout cardinales. La construction des splines quartiques d'ajustements fera l'objet de la deuxième section.

2) Spline quartique d'interpolation

2.1) Notations et définitions

On conserve toutes les notations et définitions du chapitre 1. On suppose dans toute la suite que $m \geq 1$ et que $n \geq 2$.

On note:

- $\mathcal{C}^2[a,b]$ l'ensemble des fonctions réelles définies sur $[a,b]$ dont la dérivée première est absolument continue et la dérivée seconde est de carré sommable.
- π_4 désigne le sous espace de $\mathcal{C}^2[a,b] \cap \mathcal{C}^3[a,b]$, formé d'éléments, dont la restriction à chaque intervalle I_i est un polynôme de degré 4. La dimension de π_4 est $2n+1$.
- $\mathcal{Q} = \{ s \in \pi_4 / s^{(l)}(T_1^+) = s^{(l)}(T_n^-) = 0, \quad l=2,3 \}$

La dimension de \mathcal{Q} est de $2n-3$.

2.2) Construction numérique

Problème de minimisation

Soit un histogramme donné H défini par les couples (I^k, w_k) , $k=1, \dots, m$.

Soit f un élément de $\mathcal{C}^2[a,b]$ et $E(f)$ le critère de lissage:

$$E(f) = \int_{[a,b]} f''^2(t) dt$$

Considérons le problème d'interpolation suivant:

$$(P.III.1) \quad \begin{aligned} & \text{minimiser } E(f) \\ & f \in \mathcal{H}^2[a,b], \quad \langle K_i, f \rangle = w_i, \quad i=1, \dots, m \end{aligned}$$

Ce problème a un sens physique, pour la représentation de l'histogramme H. En effet la quantité à minimiser est une approximation de l'énergie de flexion d'une barre mince assujettie à produire exactement l'histogramme de donnée. La position d'équilibre de cette barre correspondra donc, à une approximation près, à la solution du problème (P.III.1).

Pour que les théorèmes généraux sur la fonction spline s'appliquent, il faut que les intervalles I^k ne soient pas tous centrés sur un même point, de façon à avoir pour tous polynômes p de degré 1:

$$(\langle K_i, p \rangle = 0, \quad i=1, \dots, m) \Rightarrow p=0$$

Ainsi, l'existence et l'unicité d'une solution de (P.III.1) sont assurées.

caractérisation:

La même technique décrite dans (I.3) pour caractériser les splines quadratiques est appliquée pour caractériser σ .

σ est une solution de (P.III.1), si et seulement s'il existe des coefficients $(\mu_i), i=1, \dots, m$ tel que:

$$\sum_{k=1}^m \mu_k \langle k_k, p \rangle = 0, \quad p \in \mathcal{P}_1$$

et

$$\sigma''(t) = -\sum_{k=1}^m \mu_k \left((t - t_{jk})_+^2 - (t - T_{jk})_+^2 \right)$$

Il en résulte :

- 1- la restriction de σ à chaque intervalle $I_k, k=1, \dots, m$ est un polynôme de degré 4.
- 2- $\sigma, \sigma', \sigma'', \sigma'''$ sont continues sur $[a,b]$
- 3- Il existe $\mu \in \mathbb{R}^m$ tel que

$$\sigma^{(4)}(t) = \sum_{k=1}^m \mu_k \chi_{I^k}(t)$$

- 4- $\sigma^{(j)}(T_1^+) = \sigma^{(j)}(T_n^-) = 0, \quad j=2,3$

Nous supposons que les intervalles I^k soient juxtaposés, et que $m=n-1$. L'algorithme de construction se réduit à la résolution d'un système linéaire de $2(n-1)$ équations à $2(n-1)$ inconnues.

Notons $\sigma(T_i) = y_i$ et $\sigma'(T_i) = \lambda_i$, $i=1, \dots, n$

La connaissance de $(y_i, y_{i+1}, \lambda_i, \lambda_{i+1}, w_i)$ sur I_i est suffisante pour la détermination de $\sigma(t)$. En effet, il suffit d'exprimer la primitive de σ sous la forme de Newton, avec des noeuds confondus en (T_i, T_{i+1}) , puis de procéder par dérivation.

Ainsi, pour des intervalles à pas h constant, nous obtenons les équations suivantes en développant les relations de raccordement en T_i de σ'' et σ''' .

$$(III.1) \quad \begin{aligned} 8(y_{i+1} - y_i) - h(\lambda_{i+1} - 6\lambda_i + \lambda_{i-1}) &= 20(w_i - w_{i-1}) / h \\ 7y_{i+1} + 16y_i + 7y_{i-1} - h(\lambda_{i+1} - \lambda_{i-1}) &= 15(w_i + w_{i-1}) / h, \quad i=2, \dots, n-1 \end{aligned}$$

Conditions de bout:

Nous rappelons qu'il y a quatre relations supplémentaires inhérentes à la minimisation de (P.III.1).

En T_1 : $\sigma''(T_1) = \sigma'''(T_1) = 0$, ce qui donne:

$$(III.2) \quad \begin{aligned} 4(2y_2 + 3y_1) - h(\lambda_2 - 3\lambda_1) &= 20w_1 / h \\ 7y_2 + 8y_1 + h(1.5\lambda_1 - \lambda_2) &= 15w_1 / h \end{aligned}$$

En T_n : $\sigma''(T_n) = \sigma'''(T_n) = 0$, ce qui donne:

$$(III.3) \quad \begin{aligned} -4(3y_n + y_{n-1}) + h(3\lambda_n - \lambda_{n-1}) &= -20w_{n-1} / h \\ 8y_n + 7y_{n-1} - (1.5\lambda_n - \lambda_{n-1}) &= 15w_{n-1} / h \end{aligned}$$

Résolution du système (III.1), (III.2) et (III.3)

On écrit le système sous la forme:

$$(III.4) \quad \begin{aligned} B \Lambda_1 + C \Lambda_2 &= D_1 \\ A \Lambda_{i-1} + B \Lambda_i + C \Lambda_{i+1} &= D_i \quad i=2, \dots, n-1 \\ A \Lambda_{n-1} + B \Lambda_n &= D_n \end{aligned}$$

où $\Delta_i = {}^t(h\lambda_i, \gamma_i)$,

$$A = \begin{bmatrix} -1/8 & -1 \\ 1 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 3/4 & 0 \\ 0 & 16 \end{bmatrix}, \quad C = {}^tA$$

$$B' = \begin{bmatrix} 3/8 & 3/2 \\ 3/2 & 8 \end{bmatrix}, \quad B'' = \begin{bmatrix} -3/8 & -3/2 \\ -3/2 & 8 \end{bmatrix}$$

$$D_1 = (w_1/h) {}^t(2.5, 15)$$

$$D_i = (1/h) {}^t(2.5(w_i - w_{i-1}), 15(w_i + w_{i-1}))$$

$$D_n = (w_{n-1}/h) {}^t(-2.5, 15)$$

On procède par élimination de Gauss par bloc, le système (III.4) s'écrit sous la forme:

$$(III.5) \quad U_i \Delta_i + \Delta_{i+1} = K_i, \quad i=1, \dots, n-1$$

$$U_1 = \begin{bmatrix} 9 & 20 \\ 1.5 & 4 \end{bmatrix}$$

$$U_{i+1} = \begin{bmatrix} 42 & -128 \\ 6 & -16 \end{bmatrix} + \begin{bmatrix} 15 & 112 \\ 2 & 15 \end{bmatrix} U_i^{-1}$$

$$K_{i+1} = C^{-1} D_{i+1} + (U_{i+1} - C B)^{-1} K_i$$

Algorithme

- 1- calcul de U_i et K_i par (III.6)
- 2- calcul de Λ_n en résolvant le système suivant:

$$U_{n-1} \Lambda_{n-1} + \Lambda_n = K_{n-1}$$

$$A \Lambda_{n-1} + B'' \Lambda_n = D_n$$

- 3- calcul des Λ_i , $i=n-1, \dots, 1$
 $\Lambda_i = U_i^{-1} (K_i - \Lambda_{i+1})$

Remarque 1:

Nous remarquons que la matrice correspondante au système (III.4) est identique à celle obtenue pour les splines quintiques. En effet les deux systèmes sont analogues, seul le second membre est différent. Ceci nous amène à introduire les conditions de bout cardinales. Cette idée est largement développée dans la thèse de R.P Kulkarni, et elle aboutit à des résultats très satisfaisants. (au niveau du temps, d'encombrement de la mémoire et de l'ordre de convergence aux points (T_i)). Pour plus de détail, nous y renvoyons le lecteur. On se contentera ici de résumer brièvement le principe et de faire les calculs préalables à cet algorithme.

L'idée consiste à chercher deux matrices B' et B'' (2×2) afin que la résolution du système soit plus simple et directe. Ensuite, on choisit les vecteurs D_1 et D_n en fonction des mesures W_i pour que les deux relations soient exactes \mathcal{P}_4 .

B' et B'' sont déterminées de la même façon que pour les splines quintiques.

On note

$$B' = \begin{bmatrix} s & r \\ r & u \end{bmatrix}$$

Alors

$$D_1 = \left(\sum_{i=1}^5 d_i w_i, \sum_{i=1}^5 d'_i w_i \right) / h \quad \text{et} \quad D_n = \left(-\sum_{i=1}^5 d_i w_{n-i}, \sum_{i=1}^5 d'_i w_{n-i} \right) / h$$

où

$$\begin{aligned} d_1 &= (-1800s + 1096r + 146) / 480 \\ d_2 &= (4360s + -1304r + 591) / 480 \\ d_3 &= (-4200s + 1096r - 389) / 480 \\ d_4 &= (2040s - 504r + 161) / 480 \\ d_5 &= (-400s + 96r - 29) / 480 \end{aligned}$$

$$d^5 = (-50r + 12u - 26) / 60$$

$$d^4 = (255r - 63u + 144) / 60$$

$$d^3 = (-525r + 137u - 246) / 60$$

$$d^2 = (545r - 163u + 61) / 60$$

$$d^1 = (-225r + 137u + 134) / 60$$

remarque 2

Les données à valeurs moyennes sur des intervalles juxtaposés conduisent à un système relativement simple et dont la matrice est symétrique et pentadiagonale. Par contre, pour des histogrammes chevauchés et connexes, le procédé n'est plus applicable directement, mais on peut toujours se ramener à cette situation, après la résolution du système ($Ew = z$) qui a une solution unique.

3) Spline quartique d'ajustement

On suppose que les mesures z_i sont bruitées ($z_i = w_i + \epsilon_i$, $i=1, \dots, n-1$), et que les intervalles de mesure sont juxtaposés. Rappelons que pour un élément f de $\mathcal{H}^2[a,b]$, $J(f)$ contrôle l'ajustement des données.

$$J(f) = \sum_{i=1}^{n-1} \left(\int_{I_i} f(t) dt - z_i \right)^2$$

Le problème d'ajustement s'écrit:

$$(P.III.2) \quad \begin{array}{l} \text{minimiser } (E(f) + \rho J(f)) \\ f \in \mathcal{H}^2[a,b] \end{array}$$

Construction numérique de la spline quadratique d'ajustement

Les équations (III.1) sont encore vérifiées, seules les conditions d'interpolation sont remplacées par les relations d'ajustement:

$$(III.6) \quad \sigma'''(t) = \rho(z_i - \int_{I_i} \sigma(t) dt) \quad t \in I_i \text{ et } i=1, \dots, n-1$$

En développant $\sigma'''(t)$ et en vertu de (III.6):

$$3(y_{i+1} - 2w_i/h + y_i) - h(\lambda_{i+1} - \lambda_i)/2 = \omega(w_i - z_i)/h \quad i=1, \dots, n-1$$

d'où

$$(III.7) \quad w_i(\omega+6)/h = 3(y_{i+1} + y_i) + h(\lambda_i - \lambda_{i+1})/2 + \omega z_i/h \quad i=1, \dots, n-1$$

$$\omega = \rho h^5 / 120$$

Nous remarquons, comme pour les splines quadratiques, qu'on peut réduire le système précédent, en éliminant les w_i dans (III.4) par (III.7) en un système analogue.

Nous avons:

$$(w_i + w_{i-1})(\omega+6)/h = (\lambda_{i-1} - \lambda_{i+1})h/2 + 3(y_{i+1} + 2y_i + y_{i-1}) + \omega(z_i + z_{i-1})/h$$

$$(w_i - w_{i-1})(\omega+6)/h = (-\lambda_{i+1} + 2\lambda_i - \lambda_{i-1})h/2 + 3(y_{i+1} - y_{i-1}) + \omega(z_i - z_{i-1})/h$$

$i=2, \dots, n-1$

Nous remplaçons dans (III.1), (III.2) et (III.3):

$$(8-12/\omega)(y_{i+1}-y_{i-1})-h((1-4/\omega)\lambda_{i+1}-(6+16/\omega)\lambda_i+(1-4/\omega)\lambda_{i-1})=20(z_i-z_{i-1})/h$$

(III.8)

$$(7-3/\omega)y_{i+1}+2(8+3/\omega)y_i+(7-3/\omega)y_{i-1}-h(1-3/2\omega)(\lambda_{i+1}-\lambda_{i-1})=15(z_i+z_{i-1})/h$$

Les conditions de bout $\sigma''(T_1)=\sigma'''(T_1)=0$ s'écrivent:

$$(8-12/\omega)y_2 + 12(1+1/\omega)y_1 - h((1-4/\omega)\lambda_2 + (3+8/\omega)\lambda_1) = 20z_1/h$$

(III.9)

$$(7-3/\omega)y_2 + (8+3/\omega)y_1 + h((1.5+3/\omega)(\lambda_1 - (1-1.5/\omega)\lambda_2)) = 15z_1/h$$

et de $\sigma''(T_n) = \sigma'''(T_n) = 0$:

$$-12(1+1/\omega)y_n - (8-12/\omega)y_{n-1} + h((3+8/\omega)\lambda_n - (1-4/\omega)\lambda_{n-1}) = -20z_1/h$$

(III.10)

$$(7-3/\omega)y_{n-1} + (8+3/\omega)y_n - h(1.5(1+1/\omega)\lambda_n - (1-1.5/\omega)\lambda_{n-1}) = 15z_1/h$$

Résolution du système (III.8), (III.9) et (III.10)

De même que (III.4) le système s'écrit:

$$(III.18) \quad \begin{aligned} B' \Lambda_1 + C \Lambda_2 &= D_1 \\ A \Lambda_{i-1} + B \Lambda_i + C \Lambda_{i+1} &= D_i, \quad i=2, \dots, n-1 \\ A \Lambda_{n-1} + B'' \Lambda_n &= D_n \end{aligned}$$

où $\Lambda_i = {}^t(h\lambda_i, y_i)$,

$$A = \begin{pmatrix} (-1+4/\omega)/8 & -1+3/2\omega \\ 1-3/2\omega & 7-3/\omega \end{pmatrix}, \quad B = \begin{pmatrix} 6+16/\omega & 0 \\ 0 & 16+6/\omega \end{pmatrix}$$

$$B' = \begin{pmatrix} 3+8/\omega & 12(1+1/\omega) \\ 1.5-6/\omega & 8+3/\omega \end{pmatrix}, \quad B'' = \begin{pmatrix} 3+8/\omega & -12(1+1/\omega) \\ -1.5(1+1/\omega) & 8+3/\omega \end{pmatrix}$$

$$D_1 = (z_1/h) {}^t(2.5, 15)$$

$$D_i = (1/h) {}^t(2.5(z_i - z_{i-1}), 15(z_i + z_{i-1}))$$

$$D_n = (z_{n-1}/h) {}^t(-2.5, 15)$$

Par élimination de Gauss par bloc:

$$(III.19) \quad U_i \Delta_i + \Delta_{i+1} = K_i, \quad i=1, \dots, n-1$$

$$U_1 = C^{-1}B, \quad K_1 = C^{-1}D_1$$

$$U_{i+1} = C^{-1}(B - AU_i^{-1})$$

$$K_{i+1} = C^{-1}(D_{i+1} - AU_i^{-1}K_i)$$

Remarque:

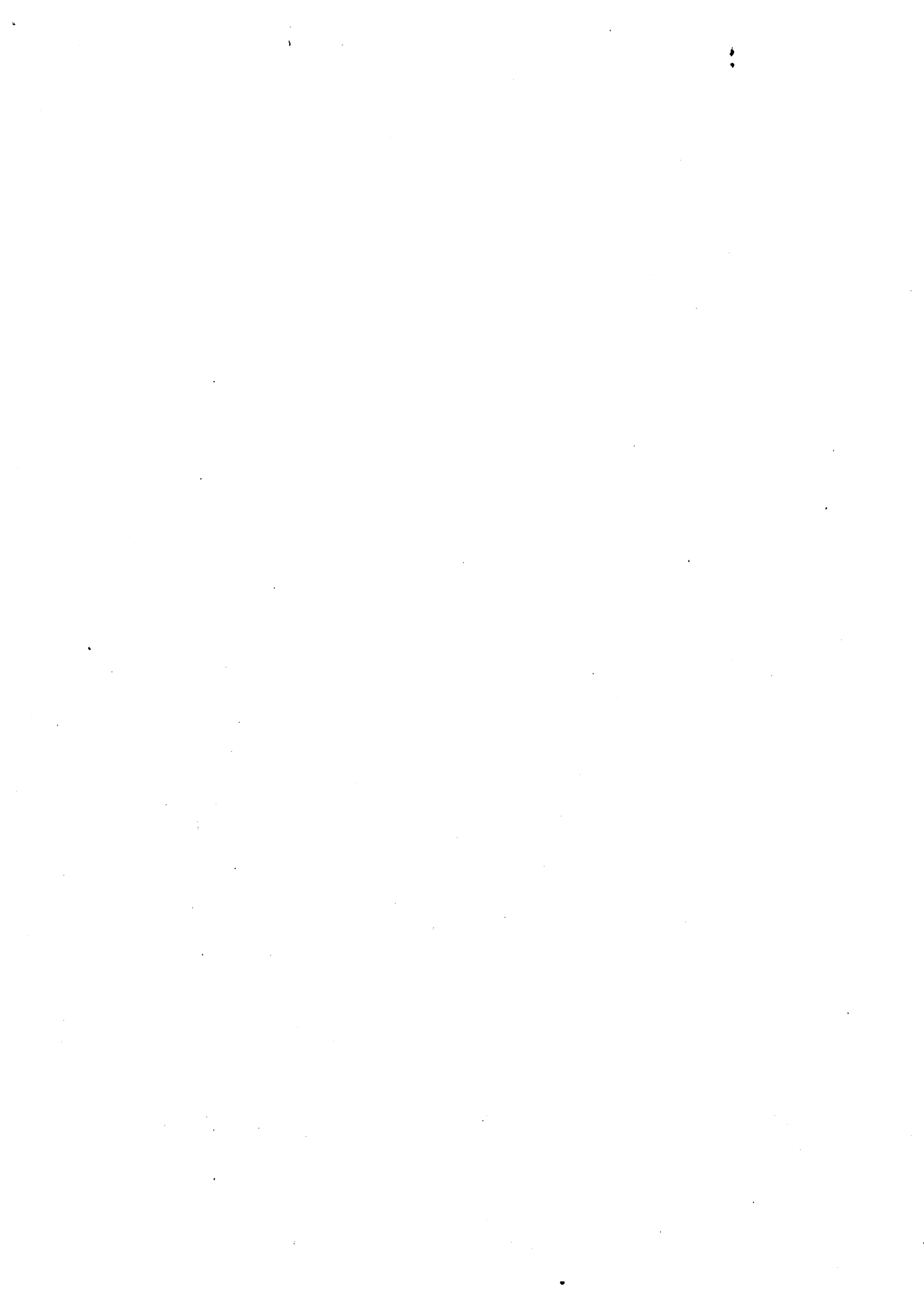
On voit bien que les systèmes de résolution pour les splines quartiques d'interpolation et d'ajustement sont analogues. Ce qui permet d'utiliser le même procédé de résolution, donc d'avoir le même ordre de grandeur du nombre d'opérations élémentaires et du temps de calcul.



Chapitre 4

Modélisation des histogrammes à intervalles non bornés

- 1- Introduction
- 2- Approche par un changement de variable
- 3- Paramétrisation des mesures à intervalles non bornés
- 4- Introduction d'une fonction-poids ou de forme.
 - 4.1- Interpolation
 - 4.2- Ajustement



1) Introduction

Nous rappelons que ce travail a été motivé, essentiellement, par la modélisation des propriétés conservatives des coupes pétrolières qui fera l'objet du chapitre suivant. L'histogramme des mesures correspondant à cette analyse est défini sur \mathbb{R} . Il en résulte que certains intervalles de mesure sont non bornés. Le modèle des histosplines, proposé dans les chapitres 1,2 et 3, est adapté spécialement aux mesures dont les intervalles sont bornés.

Pour ce faire, nous suggérons trois procédés différents qui feront l'objet de ce chapitre.

1- La première idée qui vient à l'esprit est de se ramener aux intervalles bornés, par un changement de variables, puis de procéder par la méthode des histosplines.

2- Un deuxième procédé consiste à introduire deux paramètres supplémentaires représentant les données à l'infini. Autrement dit le problème se décompose en deux sous-problèmes de minimisation: la recherche de l'histospline σ sur $[T_1, T_n]$ (T_1 (resp T_n) est le plus petit (resp le plus grand) point de coupe fini des mesures), et la détermination de ces deux paramètres. Le modèle final d'approximation est déterminé par une extrapolation convenable de σ sur \mathbb{R} .

3- Une troisième idée consiste à remplacer les histosplines par un modèle dont les fonctions sont intégrables sur \mathbb{R} . Nous proposons les fonctions de la forme:

$$g(t) = v(t)f(t)$$

avec v une fonction poids donnée et qui permet à g d'être intégrable sur \mathbb{R} .

2) Approche par un changement de variables

Le modèle d'approximation, développé dans les chapitres 1,2 et 3, suppose que les mesures soient définies sur des intervalles bornés. Dans ce chapitre, on aborde le problème avec des intervalles non bornés. On suppose que le domaine de l'histogramme des données soit \mathbb{R} .

Dans cette section, nous procédons par une approche conservative en se ramenant à des intervalles non bornés, donc au modèle déjà développé. Bien que cette transformation apparaisse simple, nous ne savons pas trop l'interpréter physiquement lors de la modélisation des coupes pétrolières.

Notons φ une fonction continue transformant l'intervalle $[0,1]$ en \mathbb{R} . (On peut évidemment choisir un intervalle quelconque fini $[a,b]$ au lieu de $[0,1]$).

$$(IV.1) \quad \begin{array}{ccc} \varphi [0,1] & \text{-----} & \mathbb{R} \\ t & \text{-----} & \varphi(t) = T \end{array}$$

et vérifiant les conditions suivantes:

$$\varphi(0) = -\infty \quad \text{et} \quad \varphi(1) = +\infty$$

On peut citer l'exemple suivant: $\varphi(t) = \text{tg}(\pi t/2)$.

Notons s la fonction à modéliser et fournie par les données:

$$(IV.2) \quad \int_{I^k} s(T) dT = z_k \quad , \quad k = 1, \dots, m.$$

Notons aussi que la structure des intervalles I^k peut être quelconque

En remplaçant T dans (IV.2) par $\varphi(t)$, nous aurons:

$$(IV.3) \quad \int_{J^k} s\varphi(t)\varphi'(t) dt = z_k \quad , \quad k = 1, \dots, m$$

où I^k est l'image de J^k par la fonction φ .

En posons $r(t) = s\varphi(t)\varphi'(t)$, il résulte de (IV.3) les relations suivantes:

$$(IV.4) \quad \int_{J^k} r(t) dt = z_k \quad , \quad k = 1, \dots, m$$

Le domaine de l'histogramme défini par les couples (J^k, z_k) est borné et satisfait les hypothèses des histosplines. Ainsi on peut appliquer le modèle des histosplines sur les données de (IV.4).

Enfin, on obtient la fonction s par le calcul de $r(t)$ par interpolation ou ajustement:

$$s(T) = (r / \varphi') \circ \varphi^{-1}(T)$$

Pour l'exemple cité plus haut: $s(T) = \pi r(2\arctg(T)/\pi) (1+T^2)/2$.

3) Paramétrisation des mesures à intervalles non bornés

Rappelons certaines notations définies dans le chapitre 1; soit s la fonction à modéliser, et $w'_i = \int_{I_i} s(t)dt$, $i=0, \dots, n$.

où $I_0 = [-\infty, T_1]$, $I_i = [T_i, T_{i+1}]$ et $I_n = [T_n, +\infty]$. $w' = (w'_0, w'_1, \dots, w'_n)$ est un vecteur de \mathbb{R}^{n+1} et $w = (w'_1, w'_2, \dots, w'_{n-1})$ est un vecteur de \mathbb{R}^{n-1} .

Notons E' la matrice associée à l'histogramme des mesures. $E' = [E^0, E, E^n]$ où E^0 (resp E^n) est un vecteur colonne de \mathbb{R}^m désignant les intervalles non bornés inférieurement (resp supérieurement).

Les mesures z peuvent encore s'exprimer en fonction du vecteur w' par la relation $E'w' = w'_0 E^0 + Ew + w'_n E^n = z$.

Dans les deux problèmes de minimisation (interpolation et lissage), la première étape consiste à calculer le vecteur w' et la seconde étape consiste à déterminer les valeurs de la solution aux noeuds (T_i) en prenant w comme vecteur de mesure.

Nous allons développer dans la suite uniquement la première étape.

Interpolation

Dans ce cas il suffit de résoudre le système $E'w' = z$

Rappelons les différents cas traités dans le chapitre 1:

-Histogramme connexe:

* adjacent: $w' = z$.

* chevauché: $w' = ({}^t E' E')^{-1} {}^t E' z$.

-Histogramme non connexe: méthode de projection.

Ajustement

Posons directement le problème de minimisation sous forme matricielle:

$$\begin{aligned} & \text{minimiser } \langle w, \Omega w \rangle + \rho \|E'w' - z\|^2 \\ & w' \in \mathbb{R}^{n+1} \end{aligned}$$

où la matrice Ω est définie dans le chapitre 2.

Par la dérivation partielle, et outre les relations (II.4) nous aurons:

$$\begin{aligned} \langle E^0, E'w' - z \rangle &= 0 \\ \Omega w + \rho^t E(E'w' - z) &= 0 \\ \langle E^n, E'w' - z \rangle &= 0 \end{aligned}$$

Le même procédé, que celui du chapitre 2, est adopté pour résoudre le système qui découle de ce problème.

Le système linéaire de résolution en w' est:

$$M^t E E' w' - L w / \omega = M^t E z \quad \text{où } \omega = \rho h^3 / 6$$

Modèle d'approximation

Après la détermination du vecteur w , nous calculons les valeurs de l'histosplines σ sur $[T_1, T_n]$, ayant $w_l, l=1, \dots, n-1$ comme valeurs moyennes, et nous extrapolons σ sur \mathbb{R} par deux fonctions φ_1 et φ_n dont les propriétés sont les suivantes:

$$\begin{aligned} \sigma^{(1)}(T_1) &= \varphi_1^{(1)}(T_1) \\ \sigma^{(1)}(T_n) &= \varphi_n^{(1)}(T_n), \quad i=0,1. \\ \int_{I_0} \varphi_1(t) dt &= w_0, \quad \int_{I_n} \varphi_n(t) dt = w_n. \end{aligned}$$

Le modèle se traduit en raccordant σ et φ au niveau de T_1 et T_n :

$$s(t) = \begin{cases} \varphi_1(t) & \text{si } t \leq T_1 \\ \sigma(t) & \text{si } T_1 < t < T_n \\ \varphi_n(t) & \text{si } T_n \leq t \end{cases}$$

4) Introduction d'une fonction poids ou de forme

Jusqu'ici, nous avons développé les algorithmes de construction numérique des histosplines: splines quadratique et quartique afin d'estimer des fonctions à une variable en utilisant les mesures de leur intégrales sur un nombre fini d'intervalles bornés. En pratique, on rencontre des fonctions à estimer dont les mesures sont déterminées à partir des intégrales d'autres fonctions. Citons un type d'exemple courant étudié par Boneva, Kendall et Stefanov dans [1]. On cherche à estimer le rapport de deux fonctions mesurées par des données d'intégration.

Notons f le rapport de g et de v : $f(t) = g(t)/v(t)$.

Dans [1] la fonction $v(t)$ représente le nombre de femmes ayant l'âge t dans un pays donné et $g(t)$ le nombre de maternité dans ce groupe. L'objectif est d'estimer le taux de maternité en utilisant les données suivantes:

$$\int_{I_i} v(t)f(t)dt = z_i \quad \text{et} \quad \int_{I_i} v(t)dt = v_i, \quad i=1, \dots, m$$

On s'intéresse particulièrement au calcul de la fonction à deux variables:

$$F(x_1, x_2) = \int_{x_1}^{x_2} v(t)f(t)dt$$

On retrouve aussi dans l'article [13] et [14] de Nira Dyn et Grace Wahba l'estimation du même type de fonction, mais à deux variables. Nous allons voir dans le chapitre 5 d'autres exemples intéressants, tels que la masse volumique, la teneur locale d'un produit dans un mélange et les fractions de ségrégation.

La méthode d'estimation proposée dans ces deux articles consiste à procéder en deux étapes séparées. On estime $v(t)$ en utilisant les données v_i , $g(t) = f(t)v(t)$ en utilisant les données z_i et f sera le quotient de ces deux estimations.

Cette étude séparée a un inconvénient dans la mesure où on ne peut pas contrôler la forme de f ainsi que sa régularité. Ce qui a été constaté en pratique et surtout dans le traitement des fractions de ségrégation (chapitre 5 application 2).

En conséquence, dans cette section nous reposons le problème sous une autre forme, en essayant d'une part, de doser la régularité de f et de l'estimer en utilisant simultanément la fonction $v(t)$ et les données z_i et, d'autre part, de modéliser les histogrammes des intervalles non bornés. Nous citons dans la suite un exemple simple de la fonction poids v permettant cette modélisation.

4.1) Interpolation

On pose le problème de minimisation sous la forme suivante en supposant d'abord que les mesures soient exactes et que la fonction $v(t)$ soit donnée ou calculée a priori.

$$(P) \quad \text{minimiser} \left(\int_{\mathbb{R}} f'(t)^2 dt \right)$$
$$\langle K_i, f \rangle = \int_{I_i} v(t)f(t)dt = z_i, \quad i=0, \dots, n$$

Nous supposons dans cette section que les intervalles soient adjacents, ce qui ne réduit en rien la méthode décrite ci-dessous. Le nombre de mesure est égale à $n+1$, et les intervalles de mesure sont $I_0 =]-\infty, T_1]$, $I_i = [T_i, T_{i+1}]$ et $I_n = [T_n, +\infty[$. Nous allons étudier la caractérisation de σ solution de (P) pour une fonction v non nulle. Puis nous présenterons dans la suite un cas simple, mais important, de v afin de modéliser les histogrammes à intervalles non bornés.

L'existence et l'unicité de la solution de (P) sont assurées, nous renvoyons le lecteur à la référence [5] de P.J.Laurent. Nous savons aussi que σ est la solution de (P) si et seulement s'il existe des scalaires α_i , $i=0, \dots, n$ de \mathbb{R} tels

$$\text{que (IV.5) } \sigma'(t) = \sum_{i=0}^n \alpha_i \langle K_i, \mathfrak{R}_t \rangle \quad \text{et} \quad \sum_{i=0}^n \alpha_i \langle K_i, c \rangle = 0 \quad \forall c \in \mathbb{R}$$

Il est convenable de ramener les mesures z_i sur la dérivée de σ . On montre qu'il existe n scalaires μ_i de \mathbb{R} tels que

$$\text{(IV.6) } \sigma'(t) = \sum_{i=1}^n \mu_i \varphi_i(t)$$

où $\varphi_i(t) = \langle L_i, \mathfrak{R}_t \rangle$, $L_i = -V_i K_{i-1} + V_i K_i$, $V_i = V(T_{i+1}) - V(T_i)$ et V est une primitive de v . Cette formulation est analogue à celle exposée dans le chapitre 1 paragraphe 4.2.3.

La fonctionnelle φ_i est définie sur le domaine $I_{i-1} \cup I_i$ et s'exprime en fonction de V par:

$$\begin{aligned} \varphi_1(t) = & \begin{array}{ll} -V_1 (V(-\infty) - V(t)) & \text{si } t < T_1 \\ V_0 ((V(T_2) - V(t))) & \text{si } T_1 \leq t \leq T_2 \\ 0 & \text{si } T_2 < t \end{array} \\ \\ \varphi_i(t) = & \begin{array}{ll} 0 & \text{si } t < T_{i-1} \\ -V_i (V(T_{i-1}) - V(t)) & \text{si } T_{i-1} \leq t \leq T_i \\ V_{i-1} ((V(T_{i+1}) - V(t))) & \text{si } T_i \leq t \leq T_{i+1} \\ 0 & \text{si } T_{i+1} < t \end{array} \end{aligned}$$

$$\varphi_n(t) = \begin{cases} 0 & \text{si } t < T_{n-1} \\ -V_n (V(T_{n-1}) - V(t)) & \text{si } T_{n-1} \leq t \leq T_n \\ V_{n-1} ((V(+\infty) - V(t))) & \text{si } T_n < t \end{cases}$$

A partir des conditions d'interpolation, nous obtenons les relations suivantes:

$$(IV.7) \quad \langle L_1, \sigma \rangle = \int_{\mathbb{R}} \varphi_i(t) \sigma'(t) dt = v_i = -V_i z_{i-1} + V_{i-1} z_i, \quad i=1, \dots, n$$

Ce qui permet de formuler le système de résolution de μ

$$\Omega \mu = v$$

où Ω est une matrice carré tridiagonale et symétrique.

$$\Omega_{i,i} = \int_{\mathbb{R}} \varphi_i(t)^2 dt, \quad \Omega_{i,i+1} = \int_{\mathbb{R}} \varphi_i(t) \varphi_{i+1}(t) dt$$

En notant $\sigma'(T_i) = \lambda_i, i=1, \dots, n$, ces derniers s'écrivent en fonction de μ

$$(IV.8) \quad \lambda_i = \mu_i V_i V_{i-1}, \quad i=1, \dots, n$$

et on peut exprimer la restriction de $\sigma'(t)$ sur l'intervalle I_i de la façon suivante, en supposant que les V_i soient non nuls:

$$(IV.9) \quad \sigma'(t) = \begin{cases} \lambda_1 / V_0 (V(t) - V(-\infty)) & \text{si } t \leq T_1 \\ (\lambda_{i+1} - \lambda_i) / V_i (V(t) - V(T_i)) + \lambda_i & \text{si } T_i \leq t \leq T_{i+1}, \quad i=1, \dots, n-1 \\ \lambda_n / V_n (-V(t) + V(+\infty)) & \text{si } T_n \leq t \end{cases}$$

La fonction σ sera déterminée par intégration de σ' et utilisation d'une seule condition d'interpolation.

Exemple simple de $v(t)$

Définissons v de la façon suivante:

$$(IV.10) \quad v(t) = \begin{cases} (-\alpha(t-T_1) + 1) e^{\alpha(t-T_1)} & \text{si } t \leq T_1 \\ 1 & \text{si } T_1 \leq t \leq T_n \\ (\alpha(t-T_n) + 1) e^{-\alpha(t-T_n)} & \text{si } T_n \leq t \end{cases}$$

où α est un paramètre positif de \mathbb{R} .

Le choix de cet exemple est important et permet de modéliser les histogrammes dont le domaine est non borné, il constitue en quelque sorte une généralisation de la modélisation présentée dans la section précédente de ce chapitre. Nous pouvons dire aussi que v est une fonction poids pénalisant les données mesurées sur des intervalles non bornés.

Il résulte de ce choix de v que la solution σ a encore une forme quadratique dans le domaine $[T_1, T_n]$ et une forme exponentielle sur $]-\infty, T_1] \cup [T_n, +\infty[$.

Après un calcul intermédiaire analogue au chapitre 1 et 2, on retrouve les mêmes relations liant les valeurs y_i de σ aux noeuds (T_i) , sauf évidemment la première et la dernière correspondant à T_1 et T_n .

Si on suppose que l'histogramme de mesure soit à pas constant et égal à h le système en y_i s'écrit:

$$(IV.11) \quad \begin{aligned} a(\alpha)y_1 + y_2 &= (b(\alpha)z_0 + 3z_1) / h \\ y_{i-1} + 4y_i + y_{i+1} &= 3(z_{i-1} + z_i) / h \\ y_{n-1} + a(\alpha)y_n &= (3z_{n-1} + b(\alpha)z_n) / h \end{aligned}$$

où $a(\alpha) = 2(1 + 4h\alpha/13)$ et $b(\alpha) = (4h\alpha^2/13)$

On remarque, que pour $\alpha = 0$ on retrouve la spline quadratique associée uniquement aux mesures définies sur $[T_1, T_n]$. Par contre pour α très grand, nous avons constaté des oscillations très fortes aux noeuds T_1 et T_n , d'où la nécessité d'un choix convenable.

La fonction σ s'écrit en fonction des valeurs y_i :

$$\sigma(t) = \begin{cases} c_1(3 - \alpha(t - T_1)) e^{\alpha(t - T_1)} + c_2 & \text{si } t \leq T_1 \\ y_i + \lambda_i(t - T_i) + u_i(t - T_i)^2 / 2 & \text{si } T_i \leq t \leq T_{i+1}, i = 1, \dots, n-1 \\ c_3(3 + \alpha(t - T_n)) e^{-\alpha(t - T_n)} + c_4 & \text{si } T_n \leq t \end{cases}$$

$$\text{où } \begin{aligned} c_1 &= (8y_1 - 4\alpha z_0) / 13 & , & \quad c_2 = (-11y_1 + 12\alpha z_0) / 13 \\ c_3 &= (8y_n - 4\alpha z_n) / 13 & , & \quad c_4 = (-11y_n + 12\alpha z_n) / 13 \end{aligned}$$

4.2) Ajustement

On traite dans ce paragraphe, le cas où les mesures ne sont pas exactes et soumises à des erreurs aléatoires. On conserve les mêmes notations que précédemment. Le problème d'ajustement s'écrit:

$$(P) \underset{f \in \mathcal{X}^1(\mathbb{R})}{\text{minimiser}} \left(\int_{\mathbb{R}} f'(t)^2 dt + \rho \sum_{i=0}^n \left(\int_{I_i} v(t)f(t)dt - z_i \right)^2 \right)$$

On abordera directement dans la suite la construction numérique de la solution σ de (P) en calculant le vecteur des dérivées de σ et les valeurs moyennes notées $w_i = \int_{I_i} v(t)\sigma(t)dt$, $i=0, \dots, n$.

Les relations (IV.5), (IV.6), (IV.8) et (IV.9) sont encore vérifiées par σ , par contre les relations d'interpolation (IV.7) sont remplacées par les relations d'ajustement suivantes:

$$(IV.12) \quad \alpha_i = \sigma''(T_i^+) / V_i = \rho (z_i - w_i), \quad i=1, \dots, n$$

et (IV.13) $\langle L_i, \sigma \rangle = -V_i w_{i-1} + V_{i-1} w_i, \quad i=1, \dots, n$

Les relations (IV.12) permettent d'exprimer le vecteur w des valeurs moyennes de $v(t)\sigma(t)$ en fonction des dérivées premières de σ aux noeuds T_i .

$$(IV.14) \quad \begin{aligned} w_0 &= z_0 + \lambda_1 / \rho V_0 \\ w_i &= z_i + (\lambda_{i+1} - \lambda_i) / \rho V_i \\ w_n &= z_n - \lambda_n / \rho V_n \end{aligned}$$

que l'on peut exprimer encore en fonction du vecteur μ par les relations (IV.8). Notons L la matrice carrée telle que $w = z + (1/\rho)L\mu$.

On développe les relations (IV.13) et on utilise l'expression du vecteur w , ce qui conduit au système linéaire en μ suivant:

$$(\Omega - (1/\rho)L)\mu = z$$

La résolution de ce système nous permet d'obtenir directement les dérivées λ_i (IV.8) de σ et les valeurs moyennes w_i (IV.14), ce qui suffit pour déterminer entièrement la spline σ .

Remarque

Dans l'exemple de v traité plus haut, on peut se ramener directement au calcul du vecteur y . Pour cela, il suffit d'exprimer les valeurs w_i en fonction des y_i . Nous avons :

$$w_0 = (c(\alpha) y_1 + z_0) / d(\alpha)$$

$$w_i = h(y_{i+1} + y_i + \omega z_i/h) / (2 + \omega) \quad , \quad i=1, \dots, n-1$$

$$w_n = (c(\alpha) y_n + z_n) / d(\alpha)$$

$$\text{ou } d(\alpha) = 1 + 4\alpha^3/13\rho, \quad c(\alpha) = 8\alpha^2/13\rho \quad \text{et} \quad \omega = \rho h^3/6$$

En remplaçant les z_i du système IV.11 par les w_i , on obtiendra un système linéaire tridiagonal en y , analogue à celui de (IV.11).

Chapitre 5

Applications: Calcul des propriétés des coupes pétrolières et régularisation des courbes des fractions de ségrégation

Application 1: Calcul des propriétés des coupes pétrolières

- 1- Introduction
- 2- Analyse du pétrole brut
 - 2.1- Introduction
 - 2.2- Rappel de la distillation "T.B.P"
 - 2.3- Analyse des quantités distillées
- 3- Modélisation des coupes pétrolières
 - 3.1- Description des mesures
 - 3.2- Représentation graphique
- 4- Progiciel **C-elf**
 - 4.1- Présentation générale du progiciel
 - 4.2- Description des modules
- 5- Analyse du brut A 334
 - 5.1- Données d'un brut
 - 5.2- Construction des fichiers de données
 - 5.3- Interprétation des courbes
 - 5.3.1-Distillat
 - 5.3.2-Densité
 - 5.3.3- Soufre
 - 5.3.4-N-paraffines
 - 5.4- Méthode de superposition des cartes
- 6- Conclusion

Application 2: Régularisation des courbes des fractions de ségrégation

- 1- Introduction
- 2- Définition et rappel
 - 2.1- Chromatogramme
 - 2.2- Fraction de ségrégation
- 3- Présentation du progiciel **S-elf**
 - 3.1- Principe général
 - 3.2- Organigramme
 - 3.3- Recalage global des N-paraffines
 - 3.3.1-Repérage automatique
 - 3.3.2-Repérage conversationnel
 - 3.4- Fonction de recalage
 - 3.5- Chromatogramme recalé
 - 3.6- Traitement d'un chromatogramme
- 4- Application: Distillat sous-vide
- 5- Conclusion et recommandation

Application 1: Calcul des propriétés des coupes pétrolières

1) INTRODUCTION

La distillation fractionnée des mélanges liquides d'hydrocarbures demeure la base fondamentale du raffinage. C'est le plus ancien procédé de séparation du pétrole brut en coupes élémentaires ou d'isolation d'un produit particulier à traiter dans le brut.

Les données brutes d'une propriété physique de ces coupes sont désignées par un ensemble de couples finis (I_i, Z_i) où Z_i représente la mesure expérimentale de cette propriété et I_i l'intervalle de température de la coupe analysée. On s'intéresse dans ce travail aux propriétés physiques s'exprimant par le quotient des quantités **conservatives** ou **additives**, par exemple la densité et la teneur en soufre total.

L'objet de ce travail est, d'une part de représenter ces propriétés par une fonction lisse continûment dérivable 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 cartes inhérentes à cette fonction.

La réalisation de ce travail a été menée selon trois axes différents:

- 1) Une étude mathématique: la formulation et le développement de certains algorithmes de calcul des fonctions splines quadratiques d'interpolation et d'ajustement, ainsi que leur adaptation aux propriétés conservatives, sont présentés dans les chapitres 1, 2 et 4.
- 2) Une étude graphique: le développement de procédures de visualisation de courbes, de surfaces et de cartes d'iso-niveau; les méthodes et algorithmes utilisés sont expliqués dans le manuel d'utilisation.
- 3) Un développement d'un progiciel souple et modulaire, combinant les résultats des deux études, est présenté dans la troisième partie de ce chapitre.

Ce progiciel "**C-ELF**" est mis au point dans le laboratoire **Tim3** de l'institut **I.M.A.G.** "Informatique et Mathématiques Appliquées de Grenoble" et implanté au **C.R.E.S** sur le calculateur **HP1000** de la chromatographie en section analytique.

L'interprétation des résultats numériques et graphiques d'une analyse du brut fait l'objet de la quatrième partie.

2) ANALYSE DU PETROLE BRUT

2.1) INTRODUCTION

Le pétrole brut est un mélange très complexe composé d'un grand nombre de produits que l'on sépare dans des classes, suivant leur volatilité. On distingue trois types de produits essentiels dans le brut :

- Produits volatils ou gazeux : gaz et essence légère.
- Produits légers : essence lourde, kérosène, gasoil..
- Produits lourds : résidus.

Dans l'industrie du raffinage, on dispose d'un certain nombre de techniques de séparation (distillation..) et de transformation (craquage, synthèse..) permettant d'obtenir ces produits et de fournir, après transformation, des produits commerciaux, issus du pétrole brut.

2.2) RAPPEL DE LA DISTILLATION "T.B.P"

Une analyse préalable de chaque pétrole dans les laboratoires d'analyse permet au raffineur de caractériser le brut, d'évaluer la composition des produits et de mesurer la densité, la viscosité et la teneur de chaque espèce (soufre, mercaptan, n-paraffines..) dans le brut.

En vue de déterminer les rendements et les propriétés des différentes coupes susceptibles d'être obtenues à partir d'un brut, le laboratoire d'analyse procède à la "distillation **T.B.P**" (true boiling point) du pétrole brut. Auparavant les gaz auront été séparés et analysés par chromatographie. La distillation **T.B.P** est une distillation en discontinu réalisée dans une colonne à haute sélectivité qui permet de séparer les divers hydrocarbures constituant le brut en fonction de leur température d'ébullition.

On utilise en discontinu une colonne à garnissage pouvant traiter de 1 à 10 litres de pétrole, et dont le facteur de séparation doit être équivalent à environ 10 plateaux théoriques et fonctionner avec un taux de reflux de l'ordre de 5. L'appareil de distillation comporte un bouilleur équipé d'un chauffage pour compenser les pertes thermiques, un condenseur et un système de reflux en tête avec prise de température. Le distillat est recueilli à l'état liquide ou gazeux.

En tête de la colonne, on recueillera les produits ou les fractions de distillation successives; on notera les températures initiales et finales de passage de chacune des fractions, T_i et T_f . Ces premiers renseignements permettront de tracer la courbe de distillation **T.B.P** du brut.

Cette courbe reflète en un certain sens la composition du produit et la répartition des hydrocarbures dans l'échantillon analysé.

2.3) ANALYSE DES QUANTITES DISTILLEES

La distillation initiale permet de fractionner le brut en des coupes classées en fonction des températures d'ébullition des hydrocarbures. Ces analyses permettent de déterminer les rendements des différents bruts. Les fractions (1 litre) obtenues sont analysées pour mesurer la densité, le pourcentage de soufre, la viscosité, le point trouble et le point d'écoulement, etc...Elles sont ensuite mélangées pour obtenir des fractions de 2,4 et 6 litres qui à leur tour seront analysées. Ceci permet de connaître les propriétés susceptibles d'être obtenues et cela pour différents points de coupe.

3) MODELISATION DES COUPES PETROLIERES

3.1) DESCRIPTION DES MESURES

Nous rappelons que, tout au long de cette partie nous supposons que les analyses sont conservatives au sens large (i.e si on note P la propriété physique d'un mélange et m sa masse alors: $m_1 P_1 + m_2 P_2 = (m_1 + m_2) P$) ce qui n'est pas le cas pour certaines analyses comme la viscosité et le point de trouble.

Un brut contient un très grand nombre d'espèce chimique (des milliards) que l'on regroupe par points d'ébullition.

Nous désignons par analyse additive la mesure quantitative d'une espèce dans chaque fraction du brut. On normalise ces quantités en divisant par la charge totale du brut et en multipliant par 100. Nous obtenons des valeurs strictement positives notées Z_i qui sont les

Intégrales de la fonction à estimer sur des intervalles de températures notés $I_i=(T_{i1},T_{i2})$ où T_{i1} et T_{i2} sont respectivement le point initial et le point final de la coupe.

Nous ajoutons aussi que nous avons deux types de collections de coupes sur l'ensemble des températures (voir figure 1) :

d'une part, toutes les coupes adjacentes et qui se succèdent (i.e le point final de la i ème coupe coïncide avec le point initial de la $i+1$ ème coupe), d'autre part, des coupes entrelacées ou chevauchées réparties d'une manière quelconque sur l'ensemble des températures et dont le graphe correspondant est connexe. La réunion de ces intervalles est non bornée et peut être égale à l'ensemble \mathbb{R} .

Ainsi notre problème consiste à déterminer une fonction approximant la propriété ci-dessus en connaissant ses mesures sur des coupes (contigues ou chevauchées). Cette propriété s'exprime en fonction des mesures additives (Intégrales), ce qui permet d'utiliser le modèle des splines quadratiques développé dans les chapitres précédents.

Le calcul du paramètre ρ optimal du lissage est déterminé par la méthode de Reinch (cette méthode est décrite en annexe), qui se ramène à contrôler l'écart global entre les mesures et leur valeurs ajustées (somme des carrés des écarts entre mesures et valeurs ajustées) en supposant avoir une bonne estimation de la variance de la méthode de mesure.

3.2) REPRESENTATION GRAPHIQUE

Les mesures additives sont fournies dans un tableau sous forme d'intervalles et de pourcentages que l'on peut traduire graphiquement par un histogramme (figure 1).

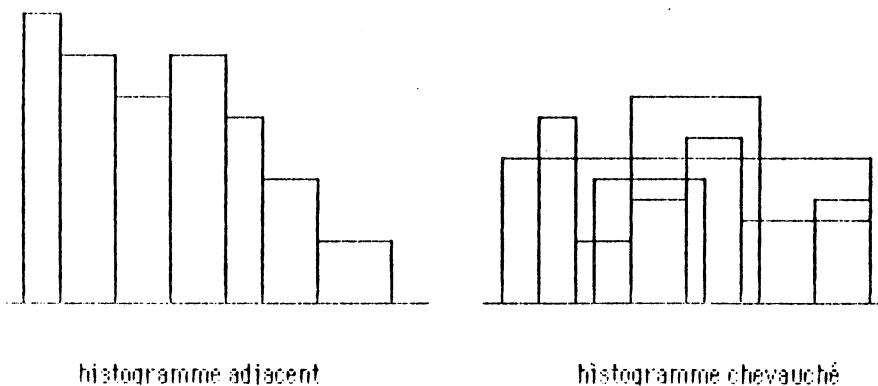


figure : 1

Les mesures de la propriété physique doivent être exploitées sous forme de courbes continues qui permettent une utilisation plus rationnelle.

On distingue trois types de représentations graphiques proposées par le progiciel **C-ELF**

- 1). Une représentation classique par une courbe continue (température, propriété).
- 2). Une représentation à deux dimensions par une surface (point initial, point final, propriété).
- 3). Une représentation projective des surfaces précédentes par les courbes de niveaux. C'est un ensemble de courbes (point initial, point final), Chaque courbe correspond à un iso-niveau.

4) PROGICIEL C-ELF

4.1) PRESENTATION GENERALE DU PROGICIEL

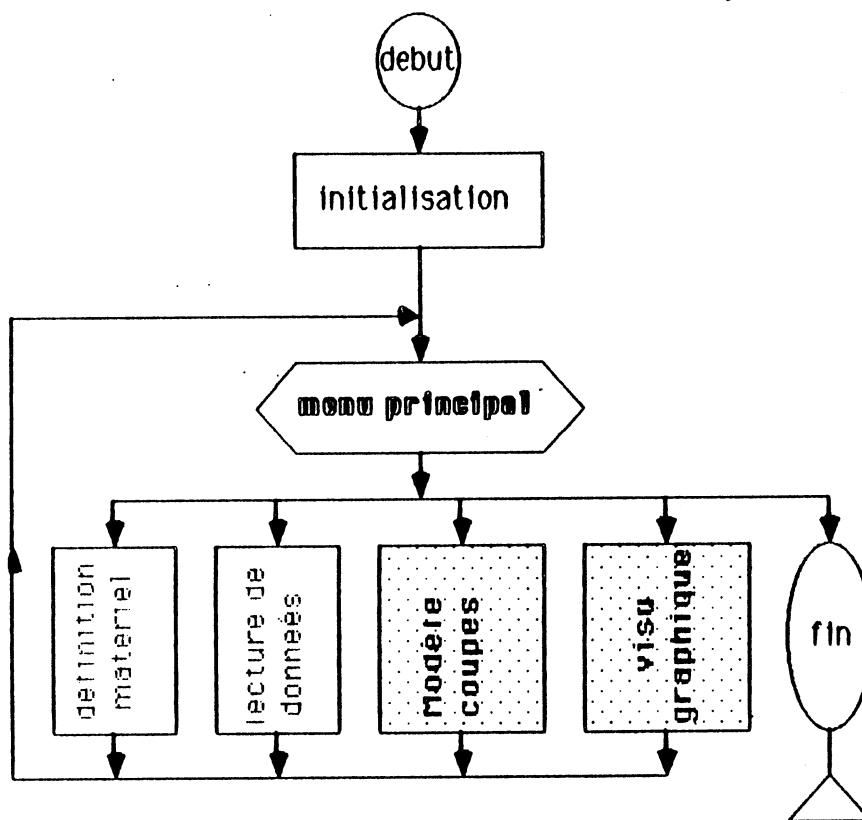
La structure du progiciel repose sur deux parties essentielles:

- Un module est consacré à la construction numérique de la ou les fonctions splines quadratiques et à la détermination du paramètre de lissage.

- Un autre module est réservé aux traitements et visualisations graphiques (courbe, surface et courbes de niveaux..).

Ces deux modules se composent de plusieurs procédures ou routines , chacune effectuant une opération particulière, utilisable ou appelable dans d'autres programmes. Ceci à pour but de montrer la souplesse de leur structure et de permettre à l'utilisateur d'ajouter ou de soustraire des procédures sans souci d'interférence ou d'influence (voir l'organigramme global de chaque module et d'appel des routines de bases).

organigramme *C-ELF*



4.2) DESCRIPTION DES MODULES

1. MODULE DEFINITION DU MATERIEL

Ce module permet à l'utilisateur, d'une part de définir le support sur lequel il préfère visualiser ses figures graphiques en choisissant entre deux terminaux graphiques (2623B et 2648B) et une table traçante (9872) disponibles au CRES, et d'autre part de définir le nombre d'ANALYSES prévu (1 ou 2). (ANALYSE correspond aux mesures additives définissant la propriété qu'on traite)

2. MODULE LECTURE DES DONNEES

Ce module permet le chargement des vecteurs températures, valeurs moyennes suivant le nombre et le type de coupes d'analyses à traiter. Les

fichiers de données ont une structure bien définie et sont constitués par l'utilisateur à l'aide d'un programme utilitaire conçu spécialement à cet effet.

3. MODULE TRAITEMENT NUMERIQUE

Ce module calcule la (ou les) fonctions spline(s) quadratique(s) et le paramètre optimal du lissage. Ce paramètre est déterminé par la méthode de Reinch et nécessite la connaissance de la variance des erreurs sur les mesures (voir annexe).

4. MODULE VISUALISATION GRAPHIQUE

Ce module est exclusivement consacré aux visualisations graphiques. Il se décompose en deux sous-modules graphiques suivant le nombre d'ANALYSES de brut (1 ou 2): **GRAPHEI** et **GRAPHEII**. Chacun d'eux est desservi par un menu proposant un éventail de tracés que l'on peut accroître selon les besoins de l'utilisateur.

GRAPHEI

Le module **GRAPHEI** assure le traitement graphique d'une seule ANALYSE ADDITIVE du brut notée f . (Exemple le %poids ou le %volume du distillat, la teneur en soufre ou d'un produit du pétrole brut). L'utilisateur doit définir la fonction à traiter, le signal f ou sa primitive F puis choisir le tracé désiré.

Exemple T.B.P %poids du distillat:

- $f(t)$ est la quantité distillée dans une coupe étroite $(t, t+dt)$.
- $F(t)$ est la quantité distillée du début de l'analyse jusqu'à t .

Tracés possibles:

- les courbes $f(t)$ et $F(t)$.
- les courbes: quantité distillée pour des coupes à point initial fixe.
- les courbes: quantité distillée pour des coupes à largeur constante.
- la surface : quantité distillée en fonction du point initial et final.
- la carte d'état major de la surface précédente.

GRAPHEII

Ce module traite une propriété physique donnée par le rapport de deux analyses additives. Nous citons deux exemples courants dans l'analyse d'un brut:

- La densité du brut correspondant au rapport poids sur volume du distillat.
- 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.

Nous aurons à traiter séparément deux analyses additives f1 et f2 et leurs primitives F1 et F2.

Tracés possibles: (densité)

- les courbes: densité pour des coupes à point final (point initial fixe).
- les courbes: densité pour des coupes à largeur constante.
- surface : densité en fonction du point initial et du point final.
- carte : les coupes à densité constante.

5) ANALYSE DU BRUT A 334

5.1) DONNEES D'UN BRUT

L'ensemble des résultats d'analyse d'un brut et ses caractéristiques sont groupés dans un cahier (cahier des bruts) sous forme de tableaux pour chaque produit du brut. Ces tableaux contiennent les points initiaux et finaux de chaque coupe et la quantité d'analyse correspondante, ainsi que certaines caractéristiques de ces coupes.

5.2) CONSTITUTION DES FICHIERS DE DONNEES

En général les plages de température sont prédéfinies pour tous les produits issus du brut, sauf éventuellement si le raffineur souhaite plus de précisions. Pour cela nous avons écrit un programme utilitaire permettant de constituer les fichiers nécessaires au progiciel **C-ELF**. Ce programme conversationnel propose un ensemble prédéfini des noeuds T_i (point de température) et demande à lire les quantités correspondant aux différentes coupes.

Une description détaillée est présentée dans le manuel d'utilisation.

5.3) INTERPRETATION DES COURBES

Nous proposons dans ce chapitre un traitement complet des différents produits du brut A 334.

5.3.1) DISTILLAT

Rappelons que la distillation initiale fractionne le brut en plusieurs coupes. Chaque coupe est représentée par un point initial, un point final et sa fraction (en %poids et en %volume).

a- La propriété d'extensivité nous permet de ramener l'analyse à des fractions étroites dont le point initial est très proche du final. Ce que nous désignons par la T.B.P différentielle (fig:2 (a)) ou la dérivée de la T.B.P

b- Le mélange physique ou bien la reconstitution mathématique de toutes les coupes à partir des coupes de mesures (qui se traduit par une sommation des quantités de distillat des coupes) nous donne la T.B.P du brut (fig:2 (b)).

Sur la figure (a) la fraction (T_i , T_f) est représentée par la surface hachurée délimitée par les deux droites verticales passant par T_i et T_f et la courbe.

Sur (b), elle est représentée par le segment [A, B], dénivellation entre A et B

Le progiciel **C-ELF** nous propose quatre types de représentations graphiques différentes.

c- 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 (2.c) visualise cette surface et marque à la demande, la projection d'un niveau, précisée par l'utilisateur, sur le plan des T_i et T_f .

d- La figure (2.d) 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.

e- La figure (2.e) visualise la projection de la T.B.P sur le plan d'axes (T_f , %distillat). C'est un ensemble de courbes T.B.P à point initial fixe. Elle permet la localisation de toutes les coupes dont la quantité est définie a priori. Sur la droite D tous les points d'intersections A correspondent aux coupes (T_i , T_f) avec T_i le point initial de la courbe et T_f l'abscisse de A.
A = (150°, 520°, 60%).

f- La figure (2.f) visualise la projection orthogonal de la T.B.P sur le plan d'axes %distillat et la bissectrice du plan des températures. Chaque courbe correspond à une largeur $\Delta T = T_f - T_i$ constante. Sur la droite D on peut aussi lire les différentes coupes (T_i, T_f) dont la fraction est 60% avec T_i abscisse de A et $T_f = T_i + dt$.
A = (210°, 610°, 60%).

5.3.2) DENSITE

La densité (ou masse volumique) d'un brut sur chaque coupe (T_i, T_f) se traduit par la relation :

$$d(T_i, T_f) = (\%p/\%v).(p/v)$$

p, v : poids et volume total du brut.

%p, %v: pourcentage en poids et en volume du distillat dans la coupe (T_i, T_f)

En fait la densité est un rapport, à un coefficient près, de deux fonctions T.B.P.

C-*ELF* nous propose les mêmes visualisations graphiques décrites en 5.3.1.

- 1- la densité en fonction de T_i et T_f (figure (3.a)).
- 2- la carte d'état major (figure (3.b)).
- 3- la densité des coupes à point initial fixe (figure (3.c)).
- 4- la densité des coupes à largeur fixe (figur (3.d)).

On voit sur les quatre figures que la densité croît en fonction de la température. En effet, on récupère le résidu à la fin de l'analyse pour des températures relativement élevées et dont le poids est supérieur au volume.

En outre ce qui intéresse le raffineur est la densité différentielle ($T_i = T_f$) telle qu'on peut la visualiser sur la figure (3.a) par la bissectrice et sur la figure (3.b) par la courbe la plus basse.

Nous remarquons aussi que cette densité n'est pas monotone et qu'il y a des fluctuations au voisinage de la bissectrice sur la surface (figure (3.a)). Ceci peut être interprété par l'imprécision des mesures des températures qui décale les produits sortants des deux T.B.P (c'est un problème analogue à celui des fractions de ségrégation, apparition des \sqrt au niveau des N-paraffines).

Par un lissage accentué, mais non souhaitable, on peut évidemment supprimer ces fluctuations.

5.3.3) SOUFRE

a- Teneur en Soufre par rapport au brut

On rappelle que la distillation secondaire de chaque fraction fournit le pourcentage des différents produits par rapport à ces coupes, en particulier le soufre. Ces quantités sont désignées dans le cahier des bruts par la teneur en soufre total (%TS). On précise aussi que ces quantités (pour tout produit) ne satisfont pas la propriété d'extensivité (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 linéairement additives.

On a les deux configurations d'intervalle de mesures de soufre du même brut. Ce qui nous permet de tester et de comparer les deux procédures de calcul des splines quadratiques.

La première figure(4.a) donne la T.B.P différentielle du soufre à partir des intervalles chevauchés :

La deuxième figure-(4.b) donne la T.B.P différentielle du soufre à partir des intervalles adjacents.

Les deux courbes sont très voisines. Des petites fluctuations apparaissent au voisinage de 200° et 500°. Elles peuvent être expliquées par le fait que le deuxième histogramme contient plus d'informations que le premier en ces endroits.

b- Teneur en Soufre total

Disposant de la T.B.P soufre et de 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 (3.2). C'est pour cette raison que nous avons conçu un progiciel indépendant des configurations d'analyse et des propriétés du brut.

5.3.4) N-PARAFFINES

Les données des n-paraffines s'expriment par des mesures de Dirac. A chaque température d'ébullition correspond une n-paraffine et sa quantité présente dans le brut.

Dans ce cas les mesures se traduisent graphiquement par des bâtons et non pas par un histogramme. Pour rester dans le domaine d'application du **C-ELF** nous suggérons, d'éclater ces mesures en histogramme contigu tel que le centre de chaque intervalle soit une température d'ébullition.

Ainsi **C-ELF** est applicable. Il permet d'obtenir une approximation grossière, dont le seul intérêt est le passage du discret au continu. Mais, il existe des procédures plus efficaces, qui permettent d'une part cette conversion et d'autre part de préserver les pics en chaque température d'ébullition. La figure (5.a) donne la quantité des n-paraffines pour toute températures t.

5.4) METHODE DE SUPERPOSITION DES CARTES.

Nous avons indiqué que **C-ELF** exploite les histosplines uniquement pour des représentations visuelles. Nous allons proposer une solution à partir de ces représentations afin de résoudre un problème de détermination de coupes.

Pour une densité et une teneur en soufre données, on souhaite localiser les différentes coupes et calculer leur rendements correspondants. Il est facile de voir que la superposition des deux niveaux respectifs de la densité et la teneur en soufre, nous permettent de localiser, par leur intersection les coupes en question.

Sur la figure (5.b) on a $d = 0.93$ et $\%S = 0.5$ l'intersection des deux niveaux donne les coupes $(81^\circ, 335^\circ)$ et $(335^\circ, 81^\circ)$.

Sur la carte d'état major des rendements il suffit de lire les niveaux passant par les points correspondant aux coupes $(81^\circ, 335^\circ)$ et $(335^\circ, 81^\circ)$: c'est 40,5%.

En fait, il suffit de superposer les trois cartes : densité, teneur en soufre et rendement. Puis de localiser les points d'intersection des trois types de niveaux.

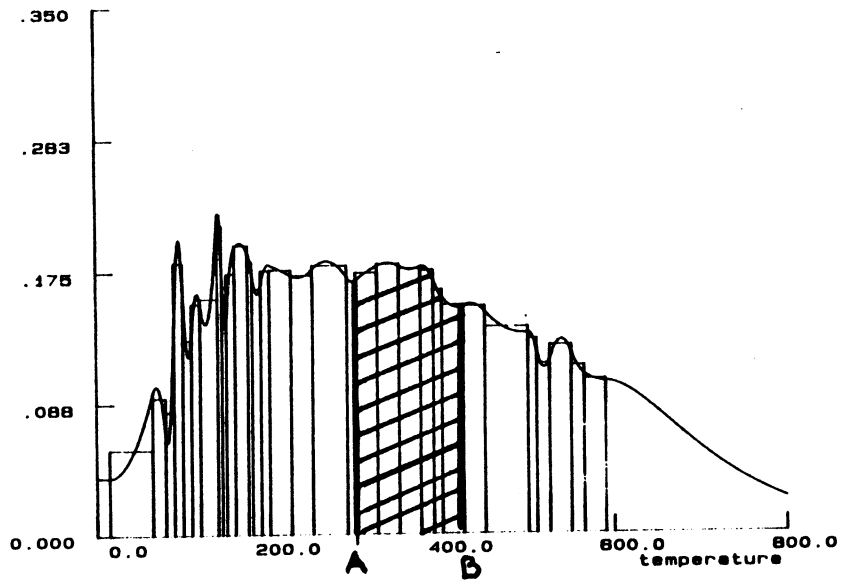


figure 2.a: Brut A 334
T.B.P distillation différentielle

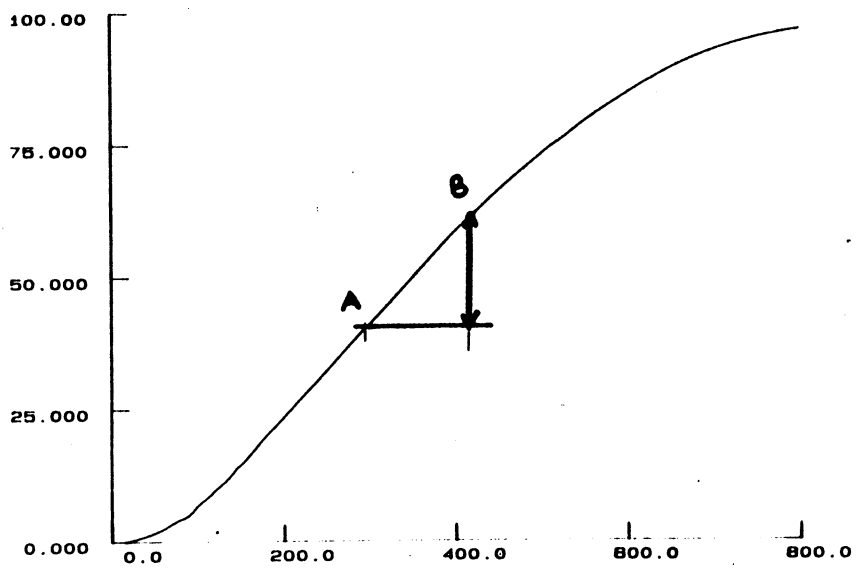
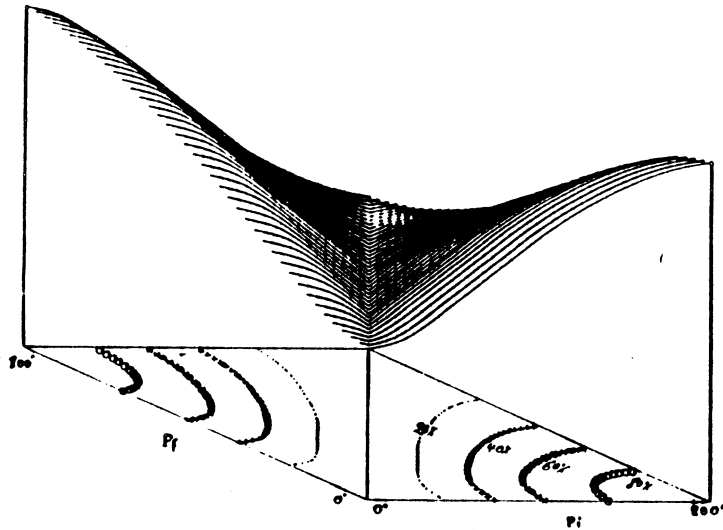


figure 2.b: Brut A334
T.B.P distillation totale



T.B.P a334

figure 2.c: Brut A 334
T.B.P (%distillat) en fonction (T₁, T₂)

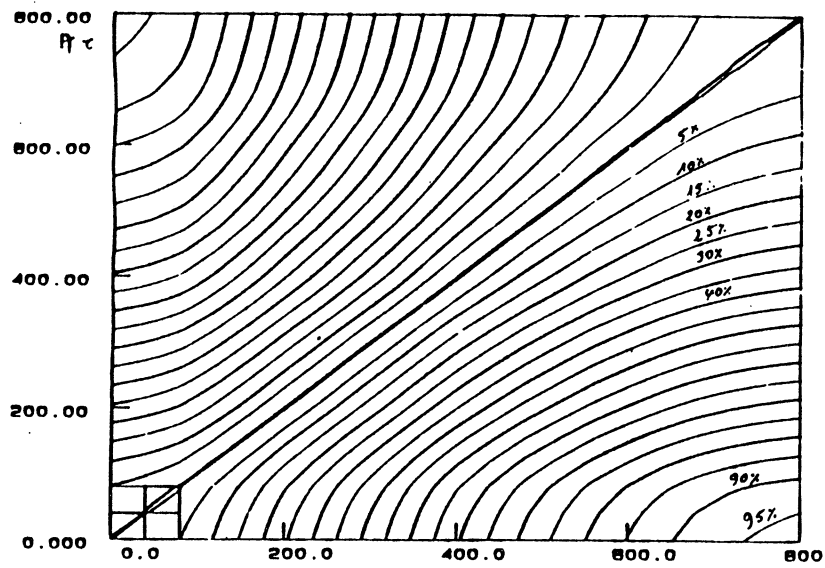


figure 2.d: Brut A 334
Carte d'Etat Major (%distillat)

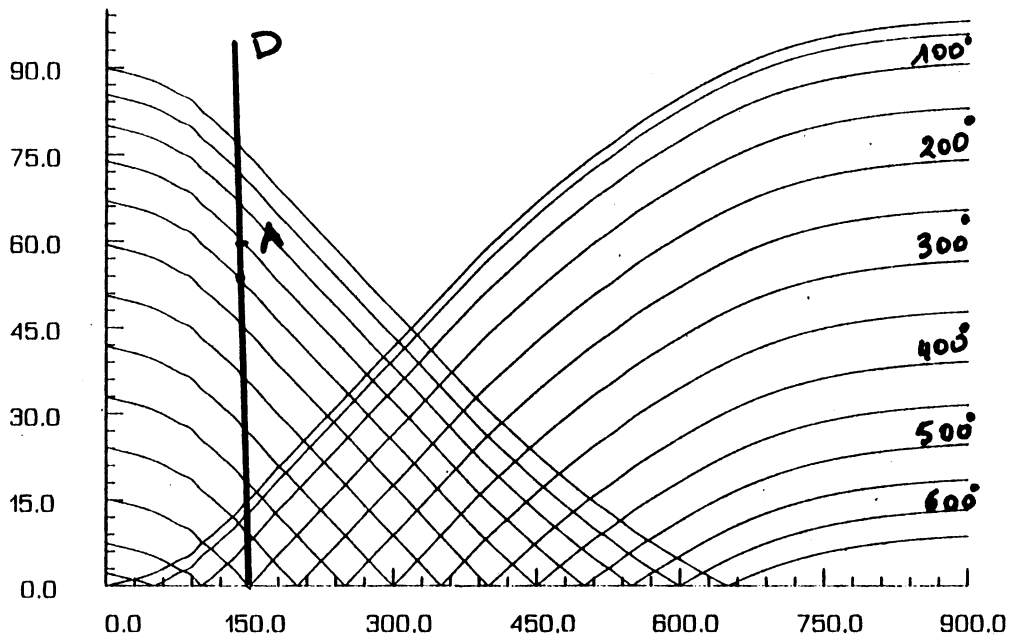


figure 2.e: Brut A 334
T.B.P pour des coupes à point initial fixe
Les T_i sont indiqués sur les courbes

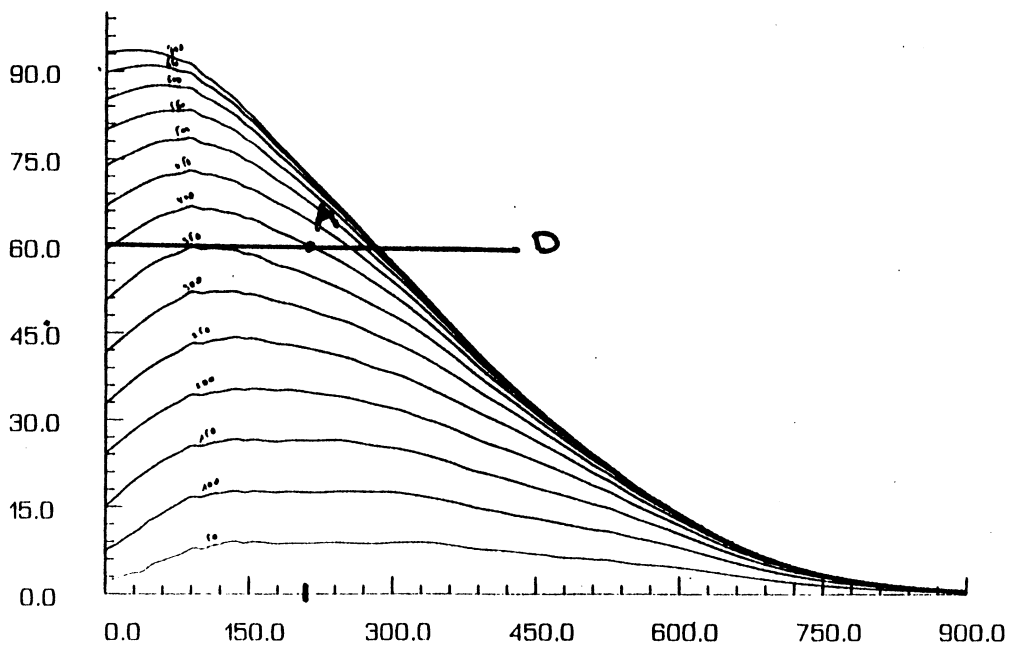


figure 2.f: Brut A 334
T.B.P pour des coupes à largeur fixe
Les largeurs sont indiquées sur les courbes

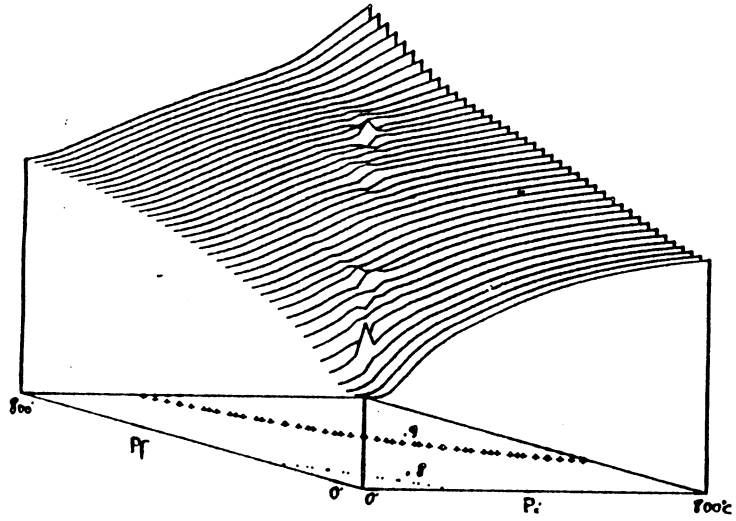


figure 3.a: Brut A 334
Surface Masse Volumique (densité) à 15°C (T_i, T_f)

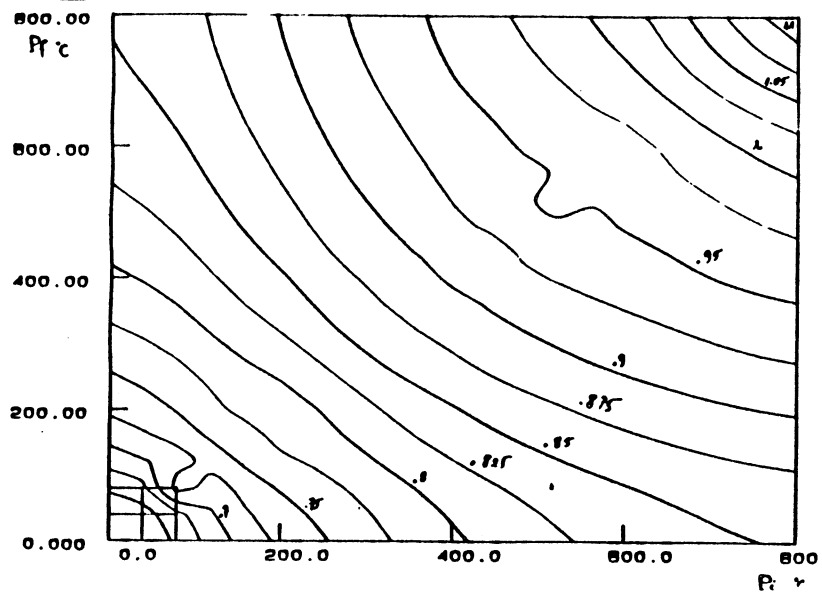


figure 3.b: Brut A 334
Carte d'Etat Major iso-MV₁₅ (T_i, T_f)

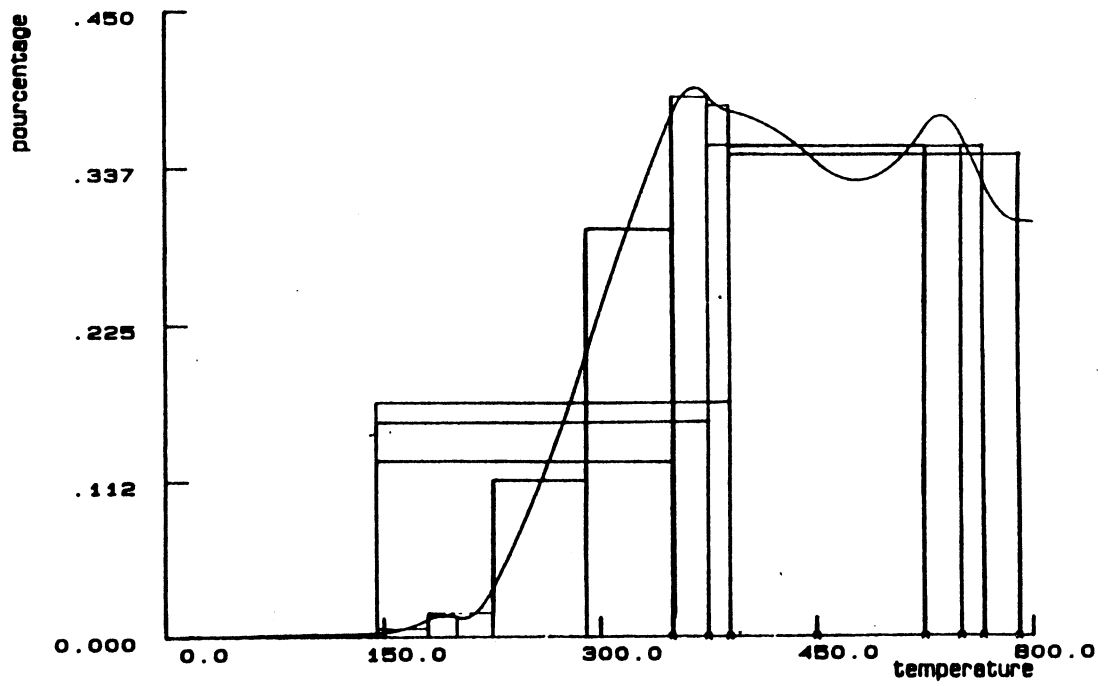


figure 4.a: Brut A 334
T.B.P soufre différentielle
(coupes chevauchées)

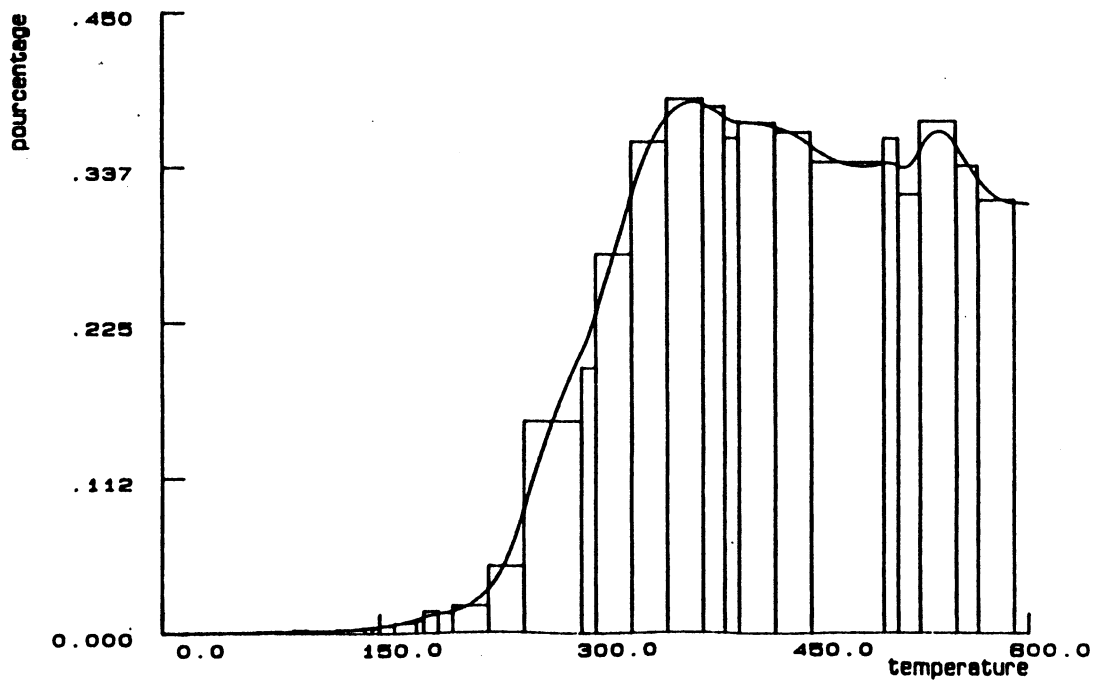


figure 4.b: Brut A 334
T.B.P soufre différentielle
(coupes contigues)

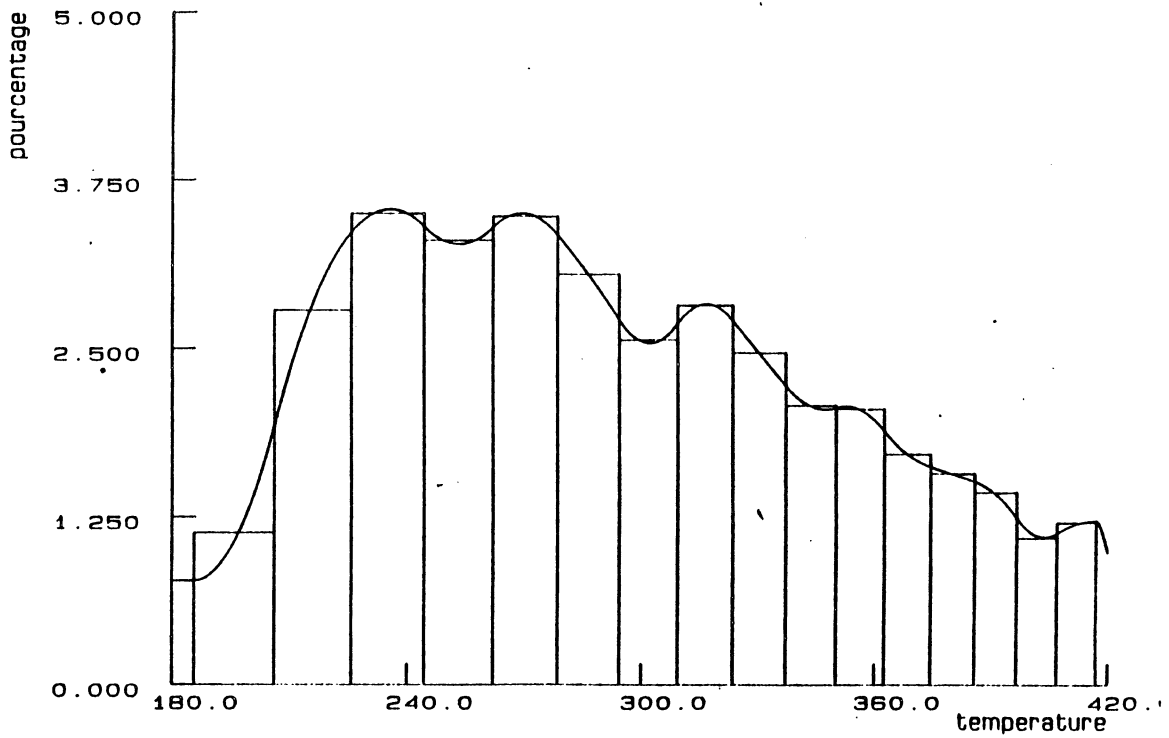
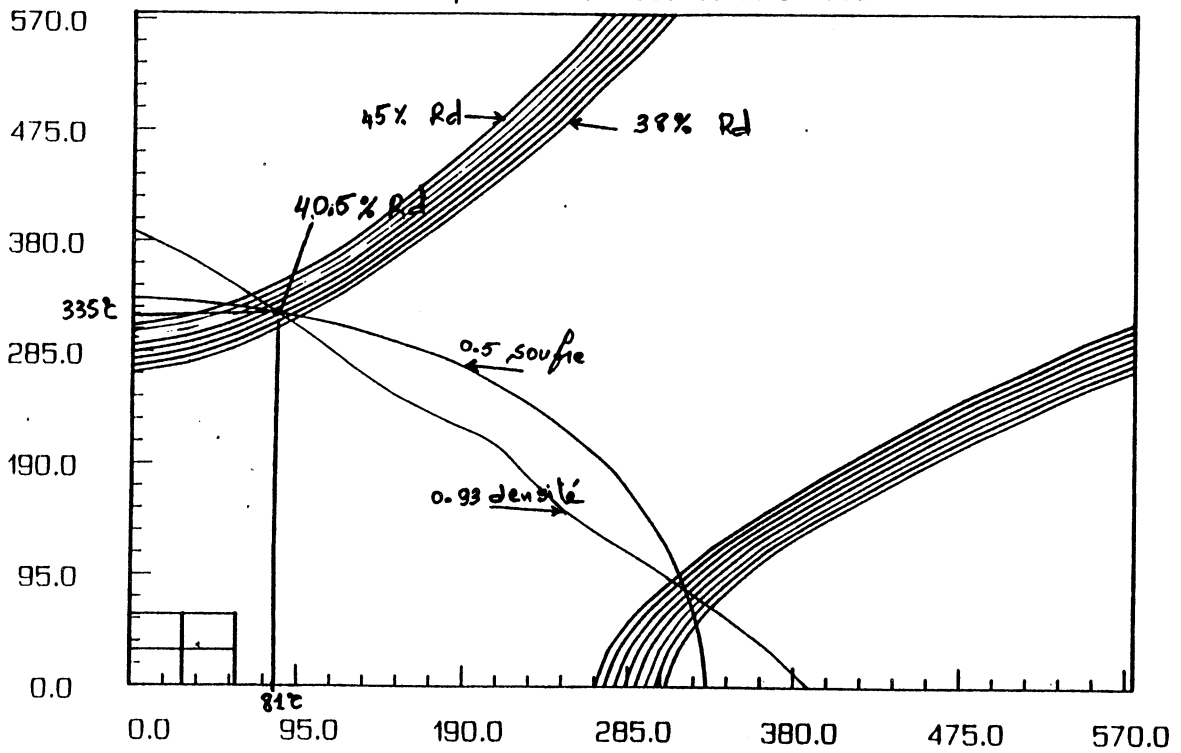


figure 5: Brut A 334
Courbe continue des N-paraffines
obtenue à partir des mesures de Diracs



Résolution graphique superposition des cartes
soufre, densité et rendement

6) CONCLUSION

L'analyse des coupes pétrolières fournit des données discrètes et imprécises sur des intervalles adjacents ou bien chevauchés. Nous disposons d'un progiciel opérationnel sur le HP1000 de la section analytique qui :

- 1) permet de rendre continues et dérivables ces analyses, en utilisant les histosplines comme éléments de bases d'approximation.
 - 2) propose une extrapolation conforme aux propriétés des coupes, pour des températures relativement grandes: supérieur à 590°.
 - 3) permet de reconstituer toutes les coupes possibles et calculer leurs quantités, par exemple le % de distillat entre 291° et 415°.
 - 4) propose certaines représentations visuelles à une et deux dimensions par exemple la T.B.P en fonction du point initial et du point final et sa carte d'état major.
- Ces représentations permettent à l'utilisateur une évaluation commode des propriétés conservatives d'une coupe ou d'un brut.

Application 2: Régularisation des courbes des fractions de ségrégation

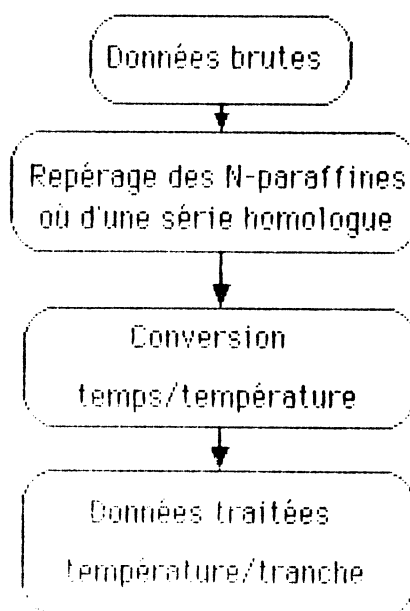
1) INTRODUCTION

La chromatographie est une méthode de séparation permettant l'analyse des mélanges très complexes comme le pétrole brut. Ce procédé de séparation fournit à la fin de l'analyse, grâce à un enregistreur, un signal ou un spectre discret correspondant aux quantités cumulées sur une suite d'intervalles de temps de rétention disjoints.

La théorie des fractions de ségrégation développée par **P.Valentin**, nous permet de déterminer à partir de ces spectres les points de coupes et de caractériser la qualité du fractionnement, et couramment appliqués par **elf-France** pour les bilans de distillation industrielles.

Malheureusement, la visualisation des courbes de ségrégation, sans traitement préalable, montre l'existence de fluctuations au niveau des températures d'ébullition de chaque N-paraffine. Ces fluctuations rendent moins précise la caractérisation citée ci-dessus.

Le traitement, préalable des données expérimentales est décrit par l'organigramme suivant:



Ces fluctuations sont provoquées par les décalages des temps de rétentions et les erreurs qui découlent de leur conversion en température. Nous avons conçu un traitement a posteriori pour les rectifier lors du bilan par un procédé de recalage des spectres par rapport à un spectre de référence défini par l'utilisateur.

L'objet de ce travail se résume en trois points essentiels:

- 1- Rendre continus les chromatogrammes et leurs dérivées, en faisant appel aux histosplines.
- 2- Essayer d'éliminer ces fluctuations par le procédé de recalage, automatique et conversationnel.
- 3- Avoir un programme permettant d'assurer ces deux dernières étapes et de tracer les différentes courbes de ségrégation.

Dans la première section de cette partie, nous rappellerons la définition des fractions de ségrégation, leur intérêt et les irrégularités apparentes. Dans la deuxième section, nous présenterons le principe général du recalage et ébaucherons les différents organigrammes du logiciel et de ses modules. Nous terminerons par une applications directes: distillation sous vide.

Remarque On ne résoud ainsi qu'une partie du problème. Les fluctuations dues à des différences de formes entre les chromatogrammes (par exemple deux analyses sur deux colonnes différentes) ne peuvent être traitées directement par ce programme. Le recalage des N-paraffines n'assure pas le recalage des parties intermédiaires (c'est par exemple le cas du bilan F.C.C).

2) DEFINITION ET RAPPEL

2.1) CHROMATOGRAPHIE

La chromatographie est une méthode de séparation des composés d'un mélange complexe.

Le principe et les différentes techniques ne sont pas décrits ici. Etant donné la multitude d'ouvrages traitant de la chromatographie, nous laisserons au lecteur le libre choix de les consulter.

Rappelons que toute analyse chromatographique fournit un spectre discret $Q(t)$, sous forme d'un fichier composé d'un grand nombre (plusieurs milliers) de temps de rétention et d'intégrales du signal entre ces points. La conversion temps/température transforme ce fichier en intégrales et en points de température.

L'analyse de ces fichiers d'une manière manuelle est très délicate. La première étape de notre étude consiste à rendre continus ces spectres et leurs dérivées, et à pouvoir calculer la valeur du signal à n'importe quelle température; cela permet une analyse et un traitement directs et commodes. Nous faisons appel aux histosplines, développés dans la première application [9] et utilisés par le logiciel *C-BLF*

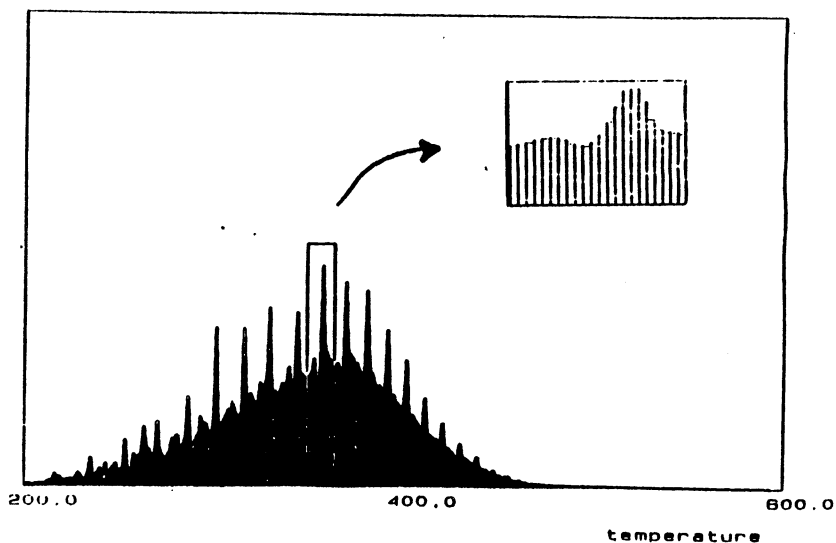


figure 7 : Spectre des données brutes graphiques

2.2) FRACTIONS DE SEGREGATION

La séparation de tout mélange est caractérisée par un vecteur entrant, représentant la charge (F_0), et un ensemble de p vecteurs représentant les différentes coupes (F_j) sortantes (dans l'espace vectoriel des débits de grandeurs conservatives). Théoriquement le bilan vectoriel s'écrit:

$$F_0 = \sum_{j=1}^p F_j$$

En normant chaque ligne de la matrice $[F_j]$ $j=1, \dots, p$ par la somme de ses éléments, on obtiendra la matrice des fractions de ségrégation. Ces vecteurs discrets ont un intérêt important pour l'évaluation de la qualité du fractionnement.

Il est possible, grâce à l'étape précédente, d'obtenir des courbes de ségrégation continues:

$$f_j(t) = F_j(t) / \sum_{i=1}^p F_i(t)$$

Cette quantité représente la fraction de ségrégation du corps de point d'ébullition t dans la coupe j . La forme de la courbe correspondante permet de :

- caractériser la qualité du fractionnement indépendamment de la nature et de la quantité de la charge du mélange.
- définir un moyen comparatif de mesure de fractionnement.

D'autres propriétés concernant la détermination des points de coupe et le nombre de plateaux sont décrites dans [11]. Cette courbe possède ainsi l'avantage d'être lisse (régulière) et elle ne dépend pas des composés majoritaires, les N-paraffines.

Remarque: les F_j sont reliées aux chromatogrammes normés X_j par la relation $F_j = p_j X_j$, où p_j = rendement en coupe j .

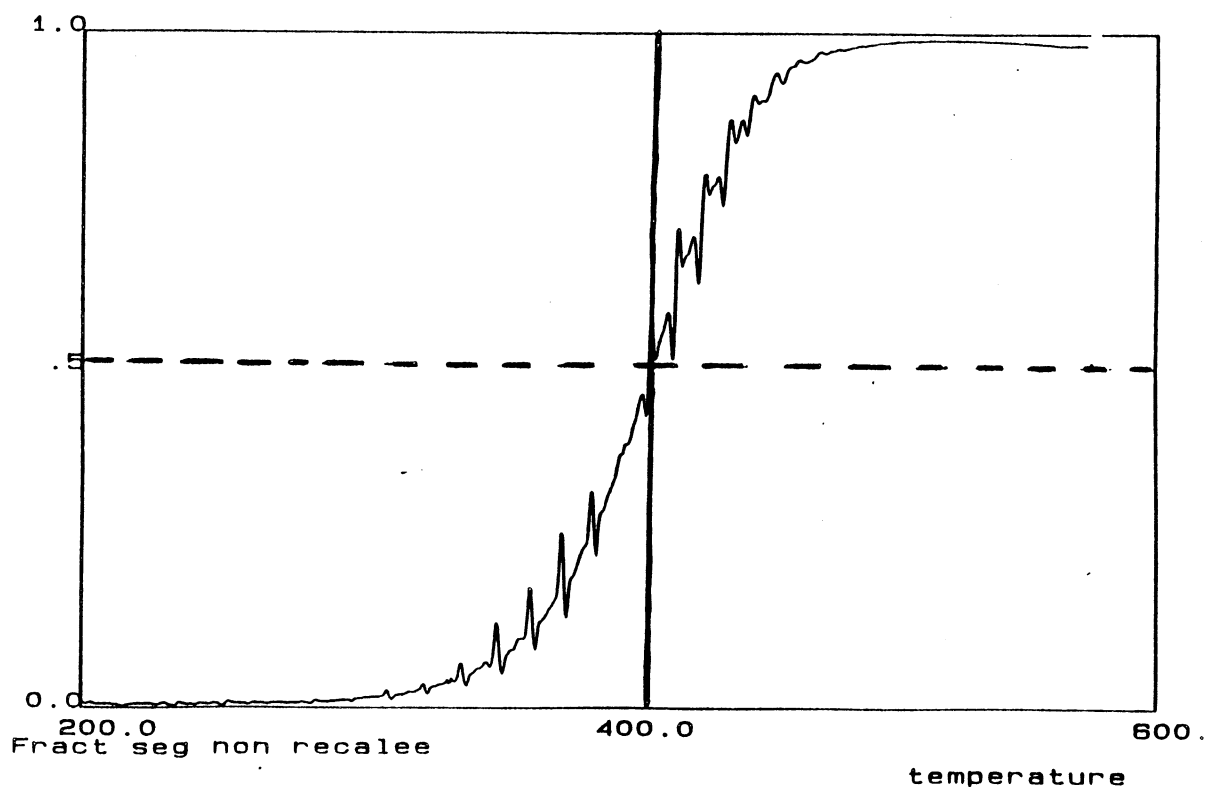


figure 8 : Courbe des fractions de ségrégation (rapport de deux spectres)
Fluctuations sous formes de λ

La pente (et plus particulièrement la forme) de la courbe indique la qualité du fractionnement.

La première idée consiste à localiser les N-paraffines par leurs températures d'ébullition dans deux spectres différents et de comparer leurs positions.

Nous remarquons alors qu'une N-paraffine dans deux coupes, n'est pas localisée à la même température d'ébullition, autrement dit, que le pic correspondant dans la première coupe est décalé par rapport à celui de la deuxième coupe (figure 9), par suite des inévitables fluctuations des temps de rétention, et des imperfections du logiciel standard de dépouillement des analyses. Dans le dépouillement initial on transforme l'échelle des temps en température. Pour cela, on utilise un "étalon interne" de température, les N-paraffines que l'on repère dans le signal, d'où la relation $T(t)$. Si on n'a pas de N-paraffines (coupes lourdes), on fait une expérience au préalable en injectant des N-paraffines et on espère que $T(t)$ sera la même pour l'analyse réelle. L'écart de t entre deux analyses de N-paraffines est de 1 à 5 sur 300 secondes.

D'où l'idée de fixer un spectre, par exemple la charge comme référence, et de recalibrer simultanément toutes les N-paraffines des autres spectres par rapport à celui-ci.

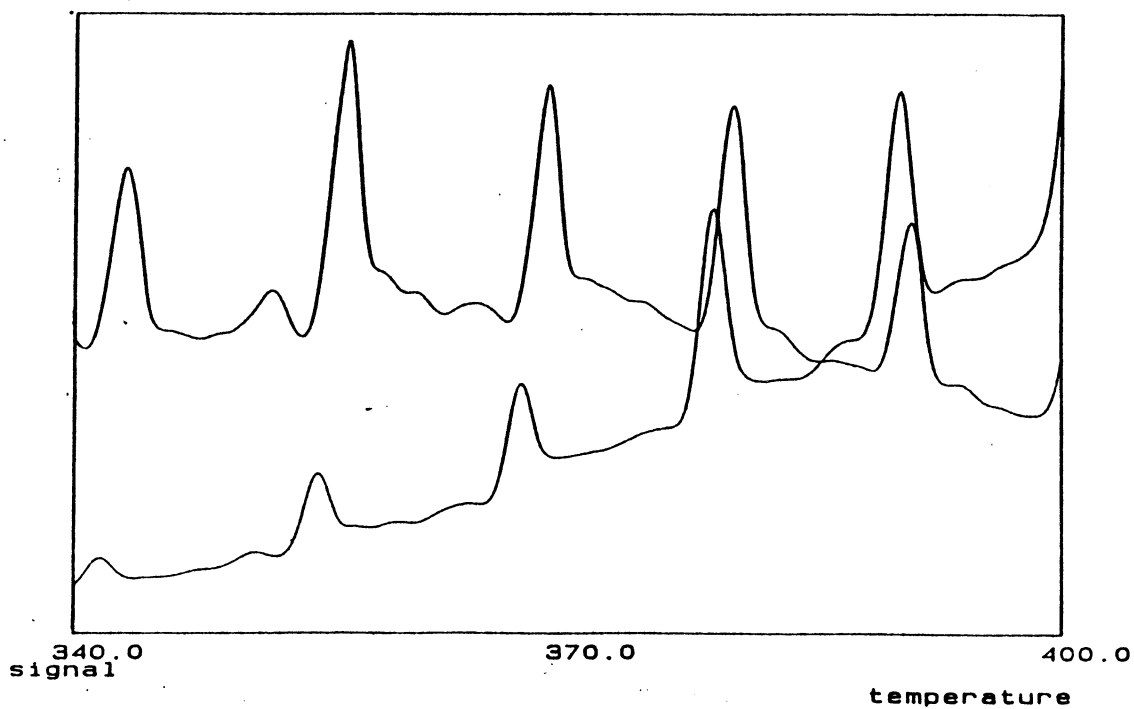


figure 9.a : Décalage des N-paraffines de deux spectres différents.

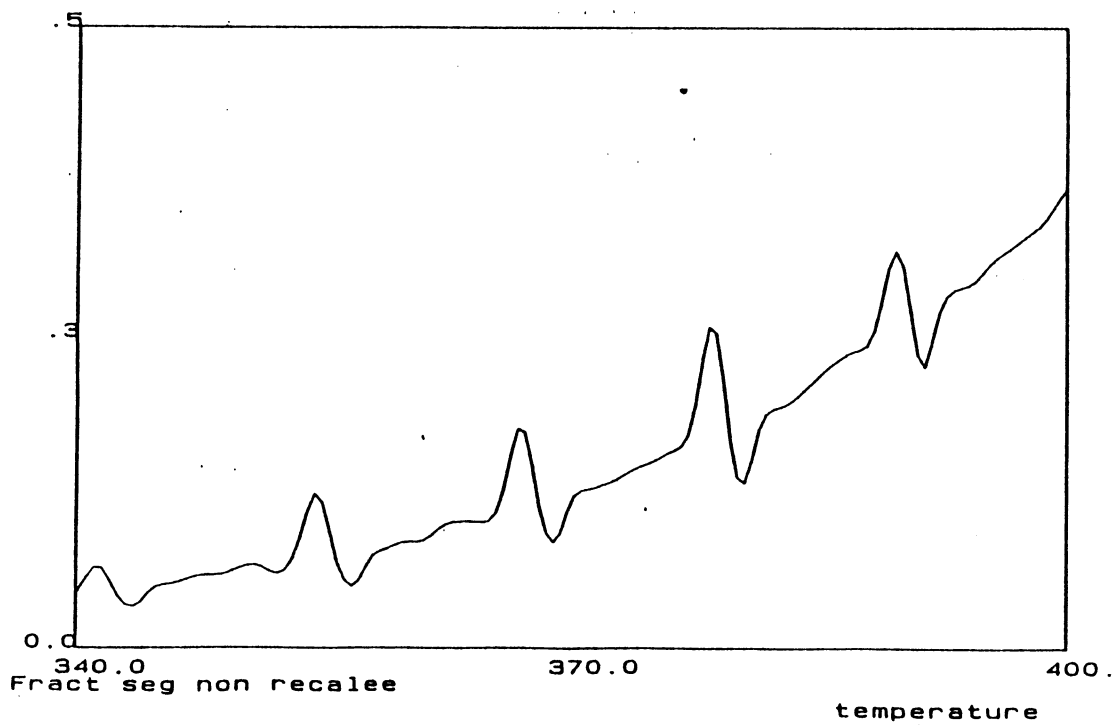


figure 9.b : Fluctuations apparaissant au niveau des N-paraffines (rapport des signaux décalés du (a)).

3) PRESENTATION DU PROGICIEL S_BLF

3.1) Principe général

Le progiciel **S-BLF** se compose de plusieurs modules automatiques et conversationnels. Il nécessite en entrée:

- Un vecteur de température d'ébullition.
- Quatre chromatogrammes : la charge et les trois effluents.

On distingue cinq étapes faisant l'objet de traitements différents:

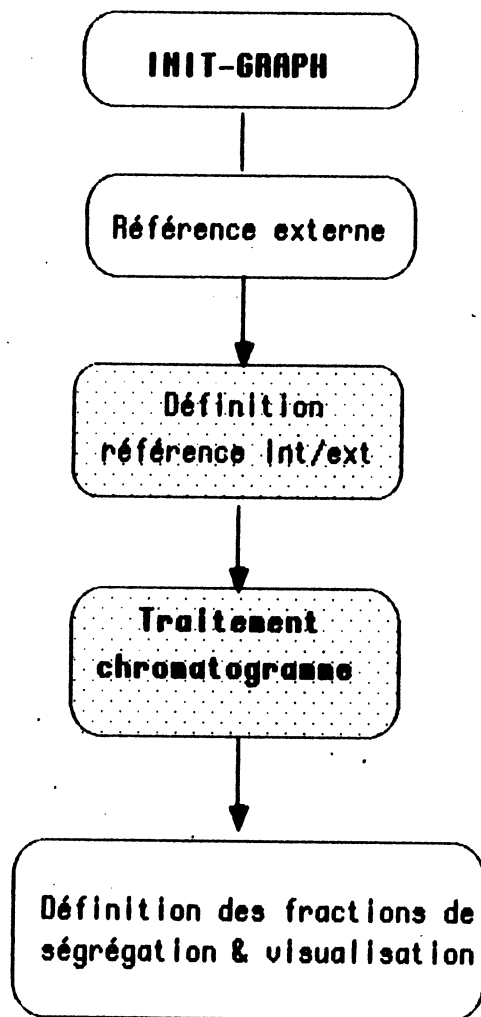
- Lecture des fichiers binaires des chromatogrammes
- Utilisation des histosplines pour rendre continus les chromatogrammes
- Repérage automatique des pics puis des N-paraffines.
- Repérage conversationnel des N-paraffines.
- Calcul de la fonction du recalage.

A la fin, on obtient les fonctions qui généralisent les fractions de ségrégation sous la forme:

$$f(t) = \frac{\sum_{i \in I} F_i(t)}{\sum_{j \in J} F_j(t)}$$

I et J \subset {1,2,3}

3.2) ORGANIGRAMME S-ELF



3.3) Recalage global des N-paraffines

3.3.1) Repérage automatique des N-P

Les composants du mélange se traduisent par des pics sur le spectre chromatographiques. Les N-P majoritaires dans ce mélange se présentent sous forme de pics émergeant nettement du signal.

La première étape consiste à localiser tous les pics du spectre par les températures correspondantes aux sommets. Puis on retient ceux qui émergent le plus en se fixant un seuil du rapport de l'amplitude maximum et minimum défini empiriquement par D.Valera [12].

$$\frac{Z_{\max}}{Z_{\min}} > 1.2$$

Nous remarquons que ce repérage automatique permet de reconnaître, dans l'intervalle [200°,550°] presque toutes les N-P présentes dans les signaux provenant de la distillation sous-vide.

3.3.2) Repérage conversationnel

Certaines N-paraffines sont éluées avec des produits voisins, ou ne sont pas reconnues par le procédé automatique. En conséquence, nous avons conçu une étape conversationnelle, permettant à l'utilisateur compétent d'ajouter des pics ou d'éliminer les candidats mal reconnus.

3.4) FONCTION DE RÉCALAGE

Après ce traitement de repérage, on récupère deux vecteurs de températures des N-P, de même dimension, correspondants aux candidats.

Le premier vecteur (TR) désigne les températures des N-P de la charge: c'est le vecteur de référence. Le deuxième vecteur (TS) correspond aux températures des N-P du chromatogramme en cours de traitement (Fr).

La fonction de recalage ou de changement de variable est la spline cubique d'interpolation, passant par les points (TS_i,TR_i).

Cette fonction présentée dans la figure 10, nous permet de conserver la continuité des signaux et de leurs dérivées. Sa courbe suit en général l'allure de la bissectrice du cadran positif (TR,TS).

3.5) CHROMATOGRAMME RECALE

Le signal recalé n'est que la composée du chromatogramme et de la fonction (Fr) de recalage.

La figure 11 montre le chromatogramme avant et après le recalage.

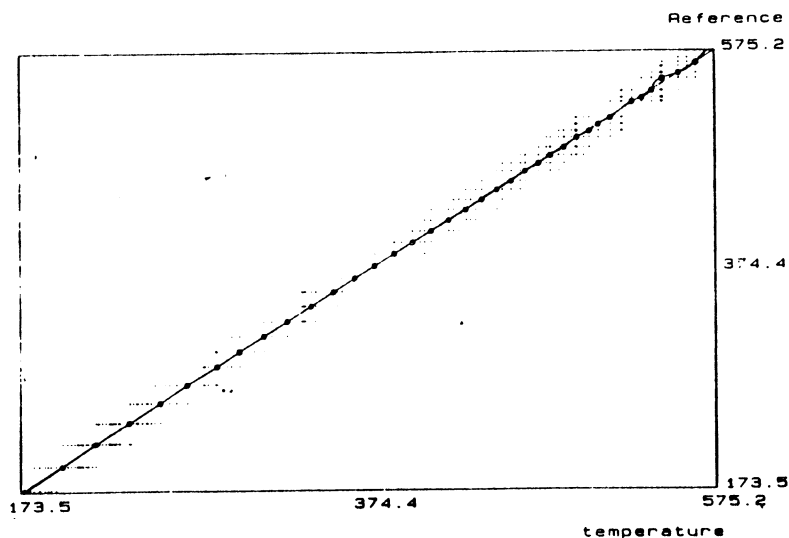


figure 10 : - fonction de recalage global.

... indique les pics de l'effluent sur l'axe température et les points de référence sur l'axe référence.

□ candidat repéré.

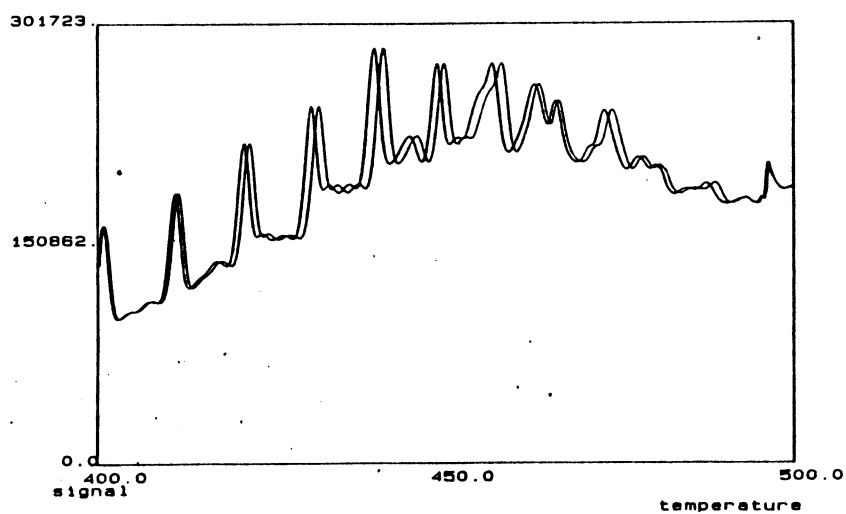
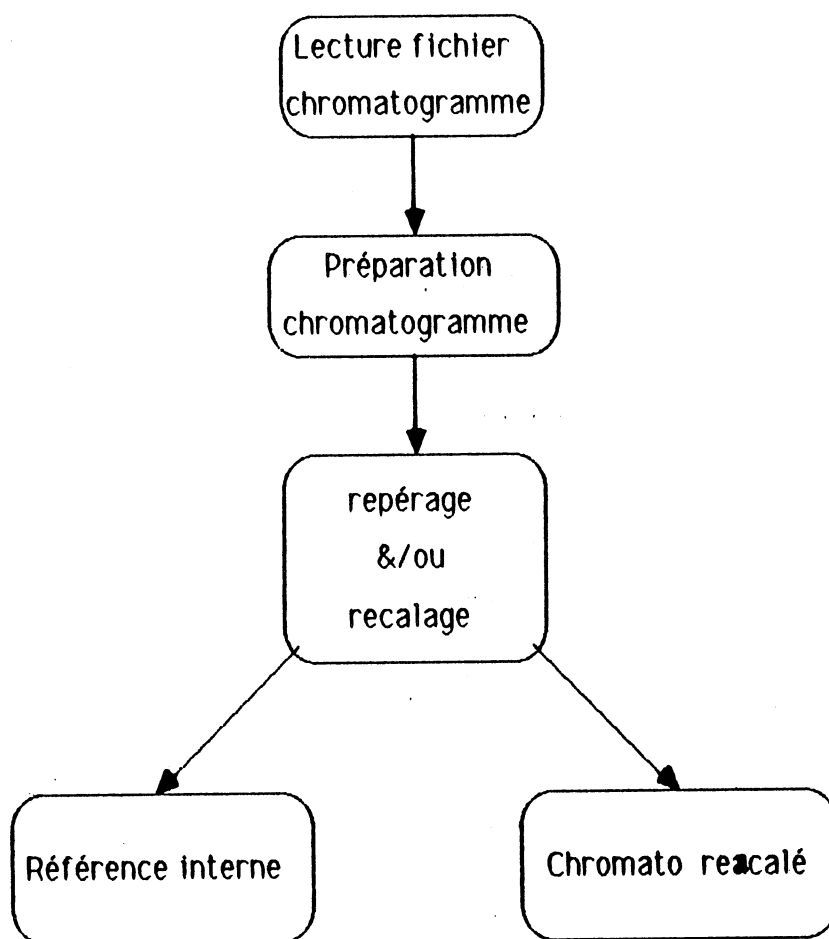


figure 11 : Chromatogramme d'un effluent avant et après recalage.

Traitement d'un chromatogramme:



Nous avons ajouté trois options supplémentaires dans le module traitement d'un chromatogramme: Normalisation, Pondération et Filtrage. Le filtrage utilisé est un filtre linéaire dans le domaine spatial, modélisé sous la forme d'une convolution entre le signal et un opérateur ou masque H défini sur un voisinage du point considéré.

$$S_f(t_i) = S(t_i) * H(t_i) \\ = .25 (S(t_{i-1}) + 2 S(t_i) + S(t_{i+1}))$$

S désigne la fonction du chromatogramme. H est linéaire et invariant en position. Cette méthode, principalement fondée sur des calculs de moyenne, aboutit au lissage du signal par une réduction du bruit.

4) APPLICATIONS

4.1) DISTILLATION SOUS VIDE

La distillation sous-vide de la coupe lourde, obtenue à partir de la distillation atmosphérique, comprend une entrée: la charge (RA), et trois sorties: le distillat DSV, le gasoil GOSV et le résidu RSV (voir la figure 12).

Les figures (13),(14) et (15) montrent le tracé de certaines courbes de fraction de ségrégation obtenues avant et après l'application du processus de recalage.

Remarques.

1- *S-ELF* donne des résultats satisfaisant quand les N-P émergent nettement des spectres qui interviennent dans l'expression de la fraction de ségrégation.

exemple: $DSV / (GSV + DSV)$ figure (13).

$(DSV + RSV) / (GSV + DSV + RSV)$ figure (14).

2- On constate une différence marquée des courbes de fraction de ségrégation quand ces dernières contiennent le spectre du résidu (RSV).

En effet les N-P sont diluées dans cette coupe, donc mal localisées par *S-ELF*. exemple : $RSV / (RSV \text{ et } GSV)$ figure (15)

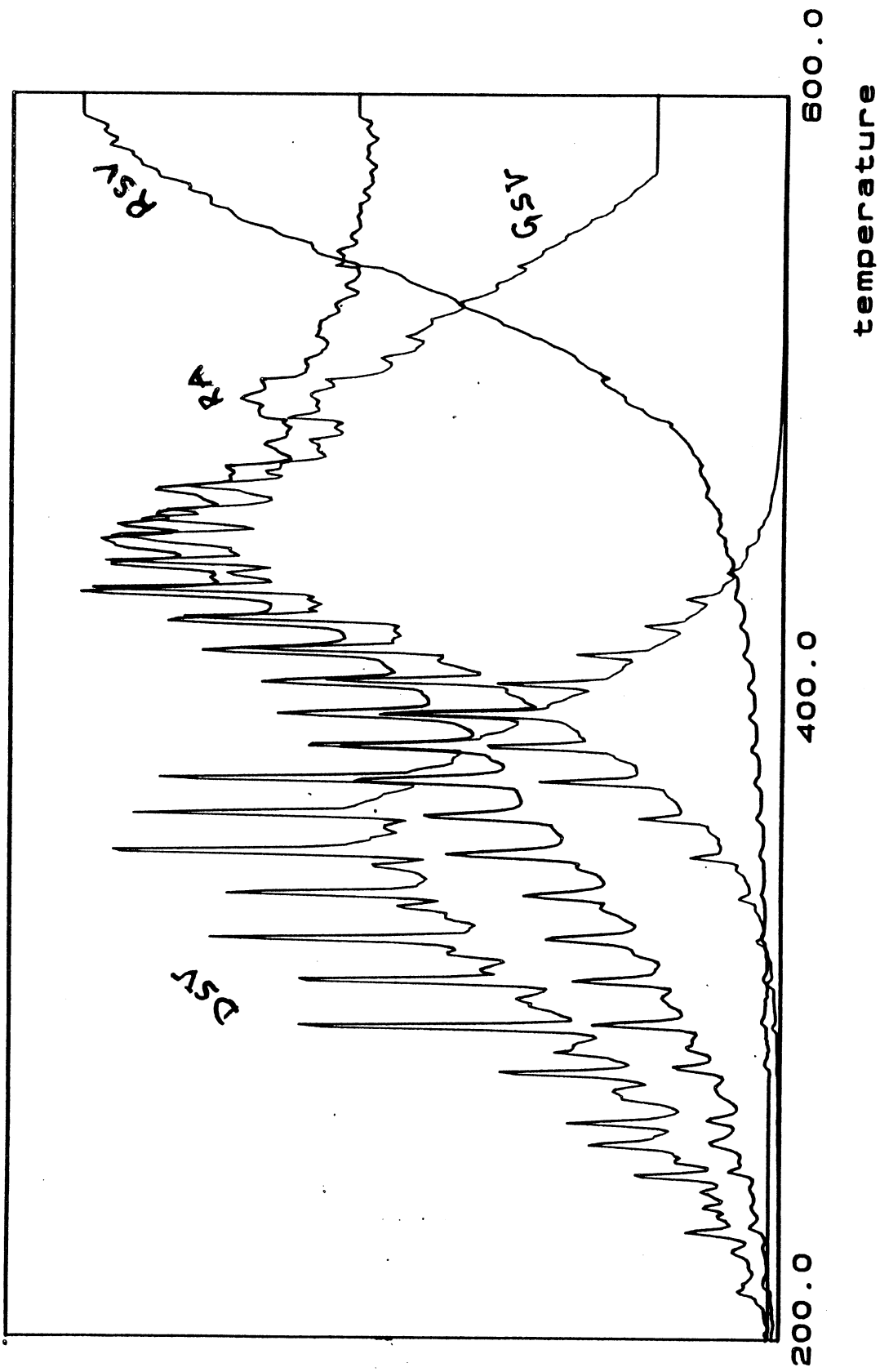


figure 12: Chromatogrammes obtenus à partir de la Distillation Atmosphérique

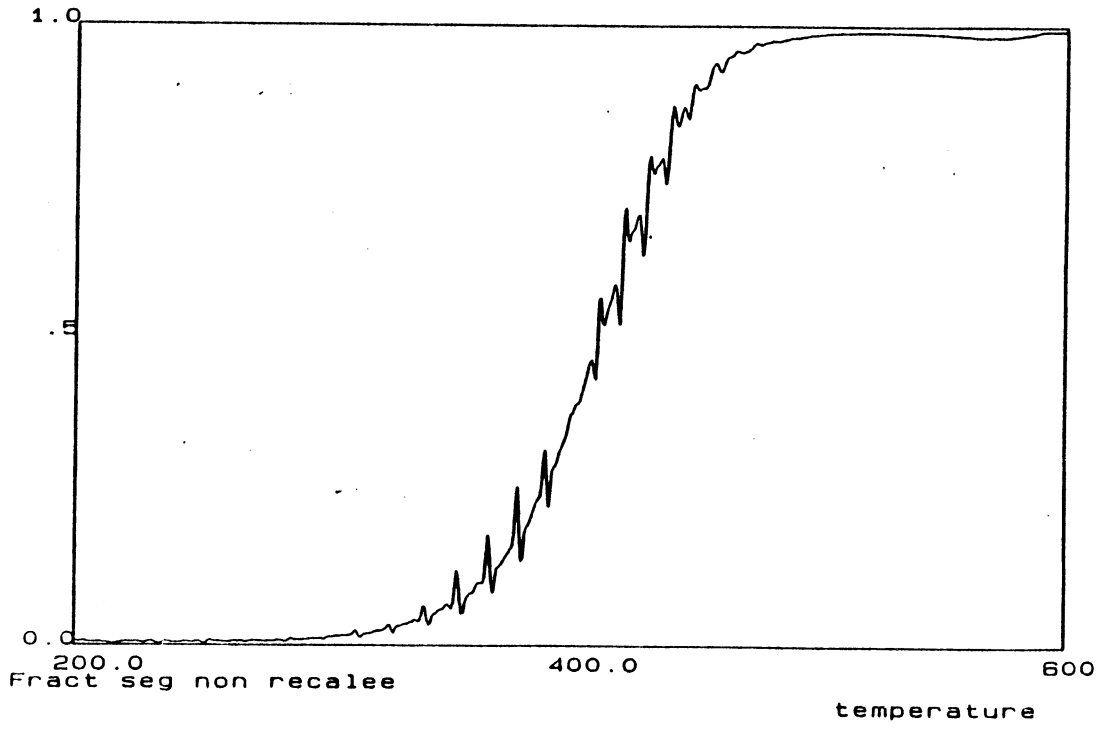


figure 13 : fraction de ségrégation avant recalage
 $DSV / (GSV + DSV)$

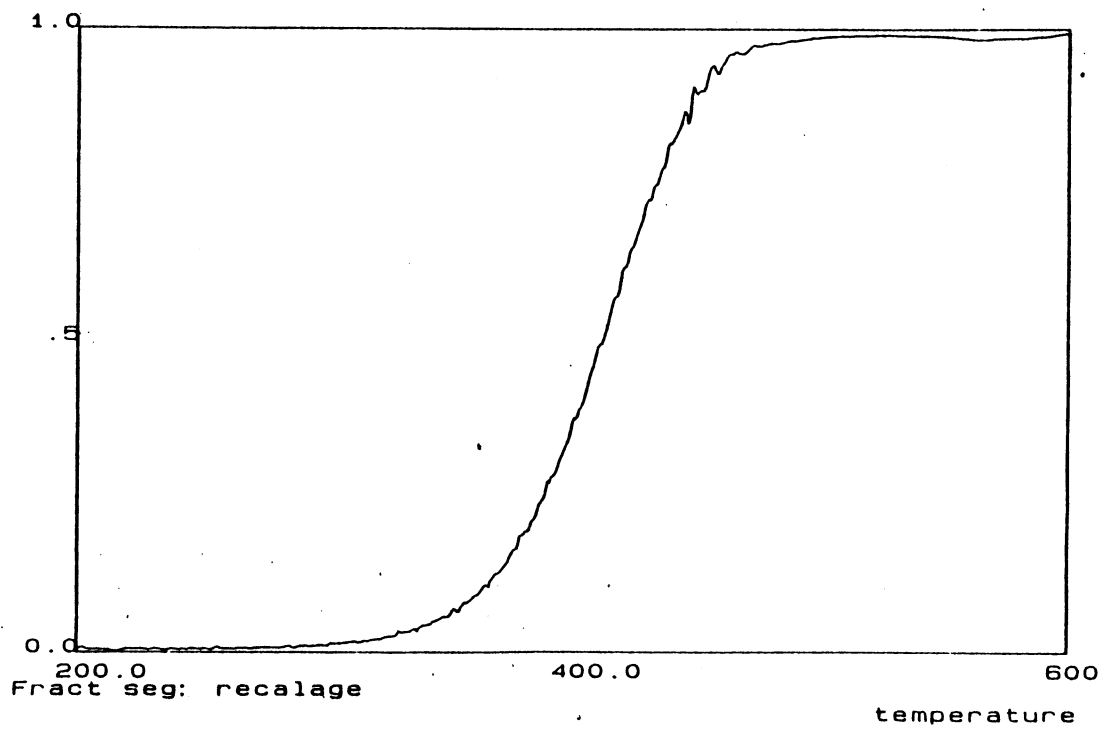


figure 13 : fraction de ségrégation après recalage
 $DSV / (GSV + DSV)$

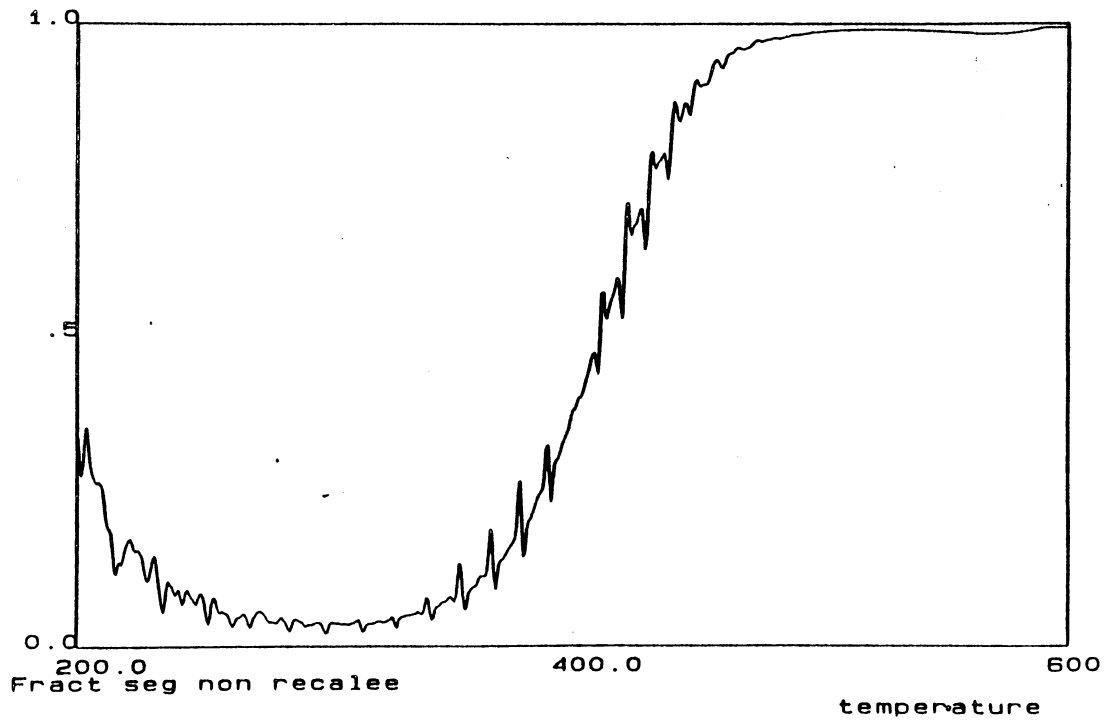


figure 14 : fraction de ségrégation avant recalage
(DSV + RSV) / (GSV + DSV + RSV)

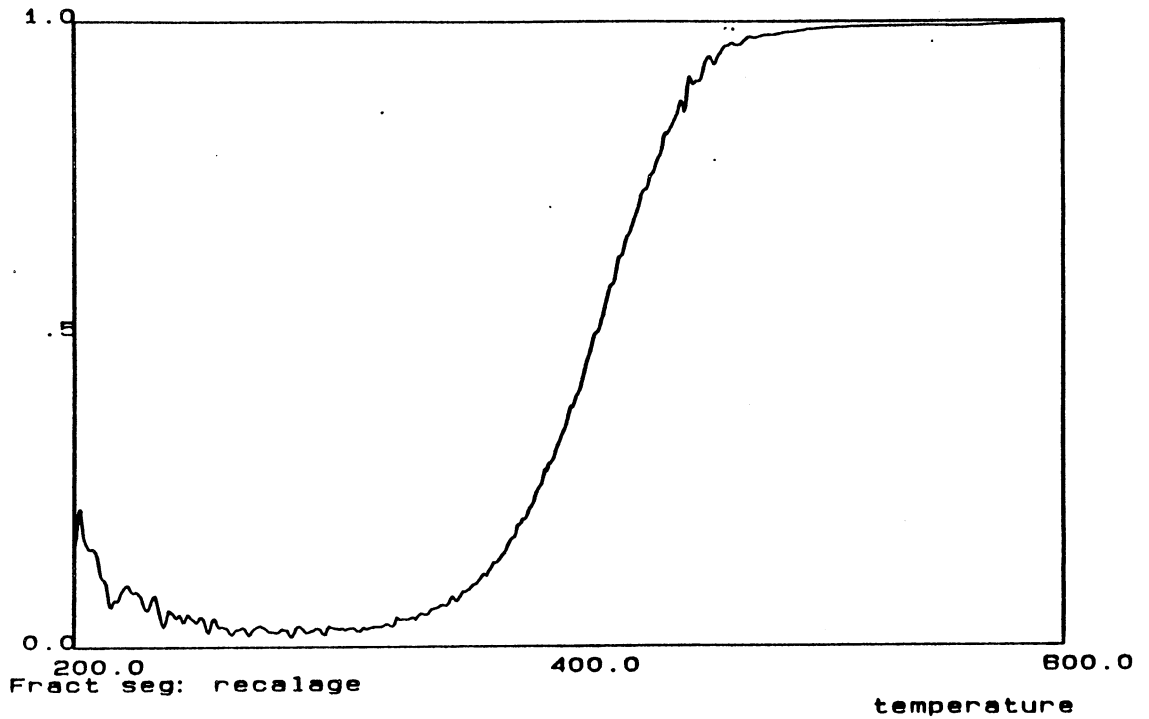


figure 14 : fraction de ségrégation après recalage
(DSV + RSV) / (GSV + DSV + RSV)

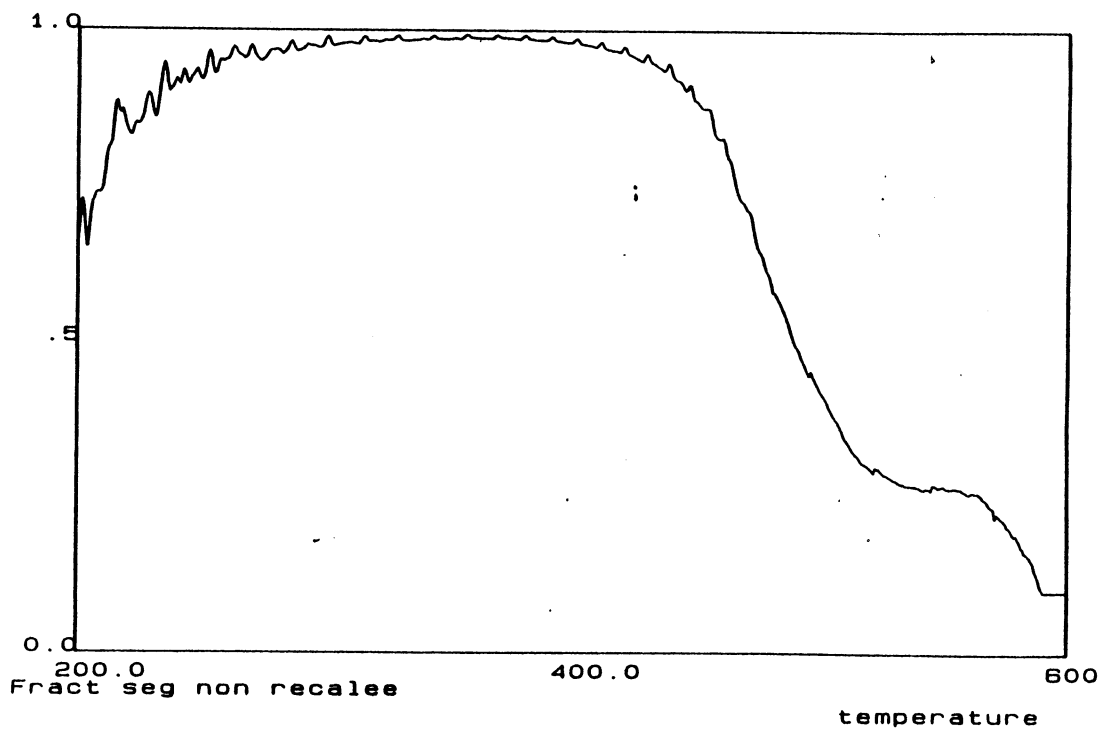


figure 15 : fraction de ségrégation avant recalage
 $RSV / (DSV + RSV)$

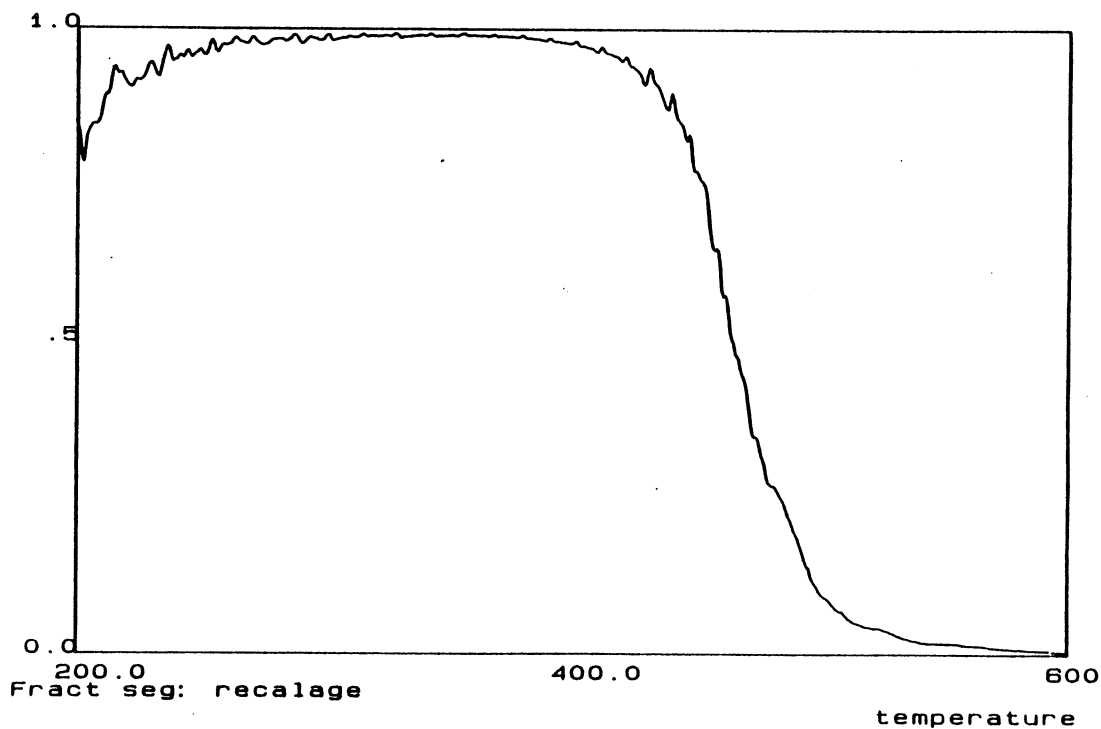


figure 15 : fraction de ségrégation après recalage
 $RSV / (DSV + RSV)$

5) CONCLUSION

Nous disposons, actuellement, d'un progiciel opérationnel sur le **HP1000** au **C.R.E.S.** Il permet d'une part une bonne modélisation des chromatogrammes par les histosplines, malgré le grand nombre de données, d'autre part un procédé de repérage automatique et conversationnel pour rendre régulières les courbes de ségrégation. Ceci n'est possible qu'au voisinage des N-paraffines émergeant convenablement du massif hydrocarboné.

Cela montre une fois de plus, l'utilité des splines et des procédés mathématiques dans la résolution des problèmes chromatographiques par traitement en temps différé (modification du dépouillement d'une analyse sans la refaire).

RECOMMANDATION

Nous ne prétendons pas avoir résolu entièrement le problème des fluctuations des courbes de ségrégation. Nous proposons seulement un procédé de recalage global, basé sur le repérage des sommets des N-paraffines. Cette dernière localisation semble insuffisante pour certaines d'entre elles et inapplicable au début et à la fin de l'analyse chromatographique, où elles sont en concentration trop faible. On notera aussi que la méthode demande une certaine certitude dans les conditions d'obtention des chromatogrammes: colonne de même type.

Nous suggérons:

1- de conserver le processus de recalage, et d'ajouter aux sommets les points d'inflexion des pics dans le repérage des N-P pour obtenir une meilleure précision.

2- d'exploiter cette modélisation des fractions de ségrégation pour la détermination du point de coupe ainsi que la visualisation de la surface et de la carte d'état major des fractions à deux dimensions.

3- de sortir les fichiers transformés.

Annexe

- 1- Construction numérique de la spline quadratique par les dérivées premières
- 2- Construction numérique de la spline quadratique par les dérivées secondes ;
- 3- Détermination du paramètre de lissage



1) Construction numérique de la spline quadratique par les dérivées premières

a-Interpolation

On considère une subdivision de l'intervalle $[a,b]$ par des abscisses équidistantes de pas h ($a=T_1 < \dots < T_i < \dots < T_n = b$). On conserve toutes les notations définies dans les chapitres 1 et 2. On s'intéresse dans ce paragraphe aux histogrammes à intervalles juxtaposés et aux valeurs de σ' aux noeuds (T_i) $i=1, \dots, n$.

Rappelons que : $y_i = \sigma(T_i)$, $\lambda_i = \sigma'(T_i)$, $i=1, \dots, n$

$$u_i = \sigma''(t) \quad \text{pour } t \in I_i = [T_i, T_{i+1}]$$

$w_i = \int_{I_i} \sigma(t) dt$, $i=1, \dots, n-1$ sont données.

On va expliciter des relations liant y_i, λ_i, u_i et w_i puis on procédera par élimination des variables y_i et u_i .

1- $\sigma'(t)$ est un polynôme de degré 1 sur $[T_i, T_{i+1}]$ Il en résulte:

$$(A.1) \quad \lambda_i - \lambda_{i-1} = hu_{i-1}$$

$$(A.2) \quad h(\lambda_i + \lambda_{i-1})/2 = y_i - y_{i-1}$$

2- Pour $t \in [T_{i-1}, T_i]$ on a :

$$\sigma(t) = y_{i-1} + \lambda_{i-1}(t - T_{i-1}) + u_{i-1}(t - T_{i-1})^2/2$$

On rappelle les deux relations suivantes déjà vues dans le chapitre 1:

$$(A.3) \quad u_{i-1} = 6(y_i + y_{i-1} - 2w_{i-1}/h) / h^2$$

$$(A.4) \quad \lambda_{i-1} = 2(3w_{i-1} - 2y_{i-1} - y_i) / h$$

Il en résulte:

$$(A.5) \quad y_i - y_{i-1} = h(\lambda_i + \lambda_{i-1})/2$$

$$\text{et } y_i + y_{i-1} = h(\lambda_i - \lambda_{i-1})/6 + 2w_{i-1}/h$$

d'où:

$$(A.6) \quad y_i = h(2\lambda_i + \lambda_{i-1})/6 + w_{i-1}/h$$

$$\text{et } y_{i-1} = -h(\lambda_i + 2\lambda_{i-1})/6 + w_{i-1}/h$$

La relation (A.6) donne deux expressions différentes de y_i , on en déduit :

$$y_i = h(2\lambda_i + \lambda_{i-1}) / 6 + w_{i-1} / h$$

$$y_i = -h(\lambda_{i+1} + 2\lambda_i) / 6 + w_i / h$$

D'où (A.7)
$$\lambda_{i+1} + 4\lambda_i + \lambda_{i-1} = 6(w_i - w_{i-1}) / h^2$$

Conditions de bout

par $\sigma'(t_1^-) = \sigma'(t_n^+) = 0$

On a : (A.8) $\lambda_1 = \lambda_n = 0$

La spline σ sera déterminée entièrement par:

- 1- Résolution du système (A.7) et (A.8).
- 2- calcul des (y_i) par (A.6)

b-Ajustement

On suppose que les mesures z_i ne sont pas exactes. Les relations d'ajustements donnent:

(A.9)
$$u_i = \rho(w_i - z_i) \quad , \quad i=1, \dots, n-1$$

A partir de la relation (A.1) et (A.9) on peut exprimer w_i en fonction de λ_i et λ_{i+1} , on a

(A.10)
$$w_i = (\lambda_{i+1} - \lambda_i) / \rho h + z_i$$

que l'on remplace dans le système (A.7) et (A.8), il en résulte:

$$\lambda_1 = 0$$

(A.11)
$$a(\omega) \lambda_{i-1} + b(\omega) \lambda_i + a(\omega) \lambda_{i+1} = 6(z_i - z_{i-1}) / h^2$$

$$\lambda_n = 0$$

où $a(\omega) = 1 - 1/\omega$ et $b(\omega) = 4 + 2/\omega$.

La spline d'ajustement est déterminée par :

- 1- Résolution du système (A.11)
- 2- Calcul des w_i par (A.10)
- 3- Calcul des y_i par (A.6)

La résolution des systèmes (A.7) (A.8) et (A.11) est décrite en détail dans le chapitre 1.

2) Construction numérique de la spline quadratique par les dérivées secondes

a-Interpolation

Dans ce paragraphe on s'intéresse aux calculs des dérivées secondes de σ dans les intervalles I_j .

$\sigma(t)$ peut être formulée sur l'intervalle I_{j-1} par deux expressions différentes:

$$(A.12) \quad \sigma(t) = y_j + \lambda_j(t-T_j) + u_{j-1}(t-T_j)^2 / 2$$

$$\sigma(t) = y_{j-1} + \lambda_{j-1}(t-T_{j-1}) + u_{j-1}(t-T_{j-1})^2 / 2$$

La valeur moyenne de $\sigma(t)$ sur I_{j-1} s'écrit alors:

$$(A.13) \quad y_j h - \lambda_j h^2 / 2 + u_{j-1} h^3 / 6 = w_{j-1}$$

$$(A.14) \quad y_{j-1} h + \lambda_{j-1} h^2 / 2 + u_{j-1} h^3 / 6 = w_{j-1}$$

w_j s'écrit à partir de (A.14):

$$(A.15) \quad y_j h + \lambda_j h^2 / 2 + u_j h^3 / 6 = w_j$$

En additionnant (A.15) et (A.13) nous aurons

$$(A.16) \quad y_j = -(u_j + u_{j-1}) h^2 / 12 + (w_j + w_{j-1}) / 2h$$

que l'on remplacera dans (A.3); nous obtenons:

$$(A.17) \quad u_{j-1} + 4u_j + u_{j+1} = 6(w_{j-1} - 2w_j + w_{j+1}) / h^3$$

Conditions de bout:

Par la relation (A.13) et $\lambda_1 = 0$ nous avons

$$y_1 + u_1 h^2 / 6 = w_1 / h$$

et par les relations (A.1) et (A.2) nous avons $y_2 - y_1 = u_1 h^2 / 2$

d'où $y_2 = u_1 h^2 / 3 + w_1 / h$

et d'après (A.16) $y_2 = -(u_2 + u_1) h^2 / 12 + (w_1 + w_2) / 2h$

ce qui donne

$$(A.18) \quad 5u_1 + u_2 = 6(w_2 - w_1) / h^3$$

Les mêmes calculs se font pour $\sigma'(T_n) = 0$

$$(A.19) \quad u_{n-2} + 5u_{n-1} = 6(w_{n-2} - w_{n-1})/h^2$$

Le système de résolution en (u_i) se résume à:

$$(A.20) \quad \begin{aligned} 5u_1 + u_2 &= 6(w_2 - w_1)/h^3 \\ u_{i-1} + 4u_i + u_{i+1} &= 6(w_{i-1} - 2w_i + w_{i+1})/h^3 \\ u_{n-2} + 5u_{n-1} &= 6(w_{n-2} - w_{n-1})/h^3 \end{aligned}$$

b-Ajustement

Il suffit de remplacer w_i de (A.10) dans le second membre de (A.20), le système final est:

$$(A.21) \quad \begin{aligned} c(\omega)u_1 + a(\omega)u_2 &= 6(z_2 - z_1)/h^3 \\ a(\omega)u_{i-1} + b(\omega)u_i + a(\omega)u_{i+1} &= 6(z_{i-1} - 2z_i + z_{i+1})/h^3 \\ a(\omega)u_{n-2} + c(\omega)u_{n-1} &= 6(z_{n-2} - z_{n-1})/h^3 \end{aligned}$$

où $c(\omega) = 5 + 1/\omega$ et $\omega = \rho h^3 / 6$

On pourrait évidemment prendre (w_i) comme inconnues. On obtiendrait alors le système:

$$(A.22) \quad \begin{aligned} c(\omega)w_1 + a(\omega)w_2 &= 5z_1 + z_2 \\ a(\omega)w_{i-1} + (4+2/\omega)w_i + a(\omega)w_{i+1} &= z_{i-1} + 4z_i + z_{i+1} \\ a(\omega)w_{n-2} + c(\omega)w_n &= z_{n-2} + 5z_{n-1} \end{aligned}$$

3) Détermination du paramètre de lissage

Les problèmes de minimisation des chapitres 2 et 3 sont résolus pour un ρ fixé au préalable. Ce paramètre contrôle aussi bien le lissage que l'infidélité des mesures (z_i) .

Il ne peut être évalué auparavant par l'utilisateur, compte tenu du nombre et de la diversité des données traitées. On remarque facilement, que pour un ρ très faible la solution σ tend vers une constante ($\sigma'(t) = 0$) et pour un ρ très grand la solution converge vers une spline quadratique au moindre carré minimisant la quantité $J(f)$ ($f \in \mathcal{C}^1[a,b]$). Le paramètre convenable se situe alors entre les deux bornes 0 et l'infini.

Méthode de REINSCH

La méthode de Reinsch et Schoenberg est basée sur la minimisation de $J(f)$, en imposant la contrainte suivante:

$$(A.23) \quad \sum_{k=1}^m ((\int_1^k f(t)dt - z_k) / v_k)^2 \leq S$$

La même question se pose encore: comment choisir le paramètre S ? Reinsch (référence [7]) propose de choisir S dans l'intervalle $[m-\sqrt{2m}, m+\sqrt{2m}]$ dans le cas où v_i est une bonne estimation de l'écart de mesure z_i .

Ce choix lui a permis d'obtenir des résultats satisfaisants (référence[7],[8]).

Dans ce qui suit, nous supposons que v_i est constante et nous notons $F(\rho)$ le premier membre de la relation (A.23) multiplié par v . On montre facilement que d'une part la fonction $F(\rho)$ est continue, convexe et décroissante et d'autre part que l'introduction de la contrainte (A.23) doit être saturée à l'optimum ($F(\rho) = v S^2$ (voir [8])).

Algorithme de calcul de ρ

Sachant que la fonction $F(\rho)$ est décroissante et convexe, on peut procéder de deux façons pour converger vers le bon paramètre.

Méthode dichotomique

Cette méthode consiste à localiser itérativement la solution ρ dans un sous intervalle jusqu'à ce que ce dernier se réduise en un point machine. C'est une méthode précise et facile à programmer, mais elle est lente.

Méthode itérative de Newton

On propose d'utiliser le procédé de Newton. La fonction utilisée est $1/F(\rho)$. Soit la $k^{\text{ème}}$ itération de Newton:

$$(A.24) \quad \rho_{k+1} = \rho_k + F(\rho_k) (v^2 S - F(\rho_k)) / (v^2 S F'(\rho_k))$$

Cette relation nécessite l'évaluation de la dérivée $F'(\rho_k)$.

Calcul de la dérivée de F

Rappelons les notations matricielles:

$$F(\rho) = \langle Ew - z, Ew - z \rangle$$

La solution w correspondante à ρ vérifie:

$$(A.25) \quad \Omega w + \rho {}^t E(Ew - z) = 0$$

En vertu de cette dernière relation et de sa dérivée par rapport à ρ nous aurons:

$$\rho(\Omega + \rho {}^t E E) \partial w / \partial \rho = \Omega w$$

$$\text{d'où} \quad F'(\rho) = - {}^t z' E (\Omega + \rho {}^t E E)^{-1} {}^t E z' \quad \text{et} \quad z' = Ew - z$$

Notons w' tel que $(1/\rho)(\Omega + \rho {}^t E E)w' = {}^t E z'$. w' n'est que le vecteur ajusté de z' que l'on peut calculer en remplaçant z par z' .

$$F'(\rho) = - (1/\rho) {}^t z' E w'$$

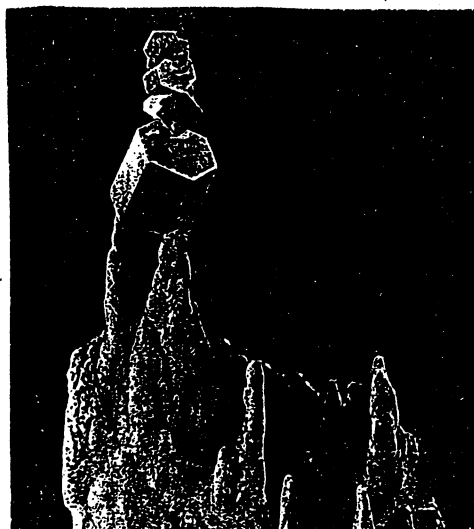
BIBLIOGRAPHIE

- [1] **Boneva, L.D.Kendall, and Stefanov**
splines transformations: Three New Dianostic aids for the
statistical Data Analyst.
J.Roy.Statist. soc. B 33: 1-70 (1971)
- [2] **C.de Boor**
A pratical guide splines
Springer-Verlag. New York 1978.
- [3] **R.Kulkarni**
truncated cardinal quintic spline
Reseach report, Univ of Grenoble 1985
- [4] **P.J.laurent**
Fonctions spline cardinales tronquées.
Reseach report, Univ of Grenoble 1983.
- [5] **P.J.laurent**
Approximation et optimisation
Hermann (Paris) 1972.
- [6] **M.Paihua**
Quelques méthodes numériques pour le calcul des fonctions
splines à une et plusieurs variables
Thèse de 3^{eme} cycle université de grenoble 1978.
- [7] **C.H Reinsch**
smoothing by spline fonction I
Numer.Math 10, 177-183 1967.
- [8] **C.H Reinsch**
smoothing by spline function II
Numer.Math 16, 177-183 1971.
- [9] **K.H Slaoui**
progiciel *c-elf* de traitement interactif d'analyse complexes
avec visualisations. Application à l'analyse des pétroles bruts.
Rapport technique elf-solaize 1985.

- [10] **P.Valentin**
Caractérisation de l'état d'un système physico-chimique, mesure de la séparation.
Partie I, en parution.
- [11] **P.Valentin**
Bilan de fractionnement DSV-test 3/03/83
RCH/CRES/ACP - Note n° 3/616 - 12/08/83
- [12] **D.Valera Garcia**
Application de l'inf-convolution spline au traitement des chromatogrammes de gasoils
thèse 3^{ème} cycle université de grenoble 1984
- [13] **G.Wahba**
Histosplines with knots which are order statistics
Ann.stat 2: 140-151 1976
- [14] **G.Wahba**
On the estimation of functions of several variables from aggregated data
University of Wisconsin-Madison MRC.TSR. 1974.
- [15] **P.Whithier**
Le pétrole Raffinage et Génie Chimique
technip Paris 1966

Partie II

Gestion et géométrie des ensembles de mélange



Que de signes miraculeux dans les cieux et sur la terre que les hommes côtoient en s'en détournant

le saint coran.



Introduction

Dans certains domaines et particulièrement en chimie, pétrochimie, raffinage etc... les techniques de séparation des corps dans un système ou dans un mélange sont bien étudiées et largement développées. Cependant plusieurs problèmes mathématiques ne sont pas résolus jusqu'à présent. Dernièrement, **P. Valentin** a proposé une modélisation géométrique de la séparation par la classe des **zonotopes**. Ces derniers, assez rares, sont étudiés par plusieurs auteurs. Nous citons **Coxeter** (1948) dans [5] qui développe essentiellement ces ensembles à trois dimensions: "Zonohedra" ; **McMullen** (1971) dans [13] et **Shephard** (1972) dans [23] introduisent la notion de zonotope associé

Dans le chapitre 1, nous commençons par une présentation générale des zonotopes, on s'intéresse particulièrement à leur structure et propriétés géométriques.

Un zonotope correspond à un état de la séparation dans un système donné, et plus précisément, à l'ensemble des mélanges que l'on peut fabriquer à partir de cet état. Dans le chapitre 2, nous nous intéressons essentiellement au problème de la faisabilité d'un mélange objectif à partir d'un zonotope, ou bien à partir de l'état de séparation associé. Plusieurs méthodes d'optimisation sont appliquées et comparées.

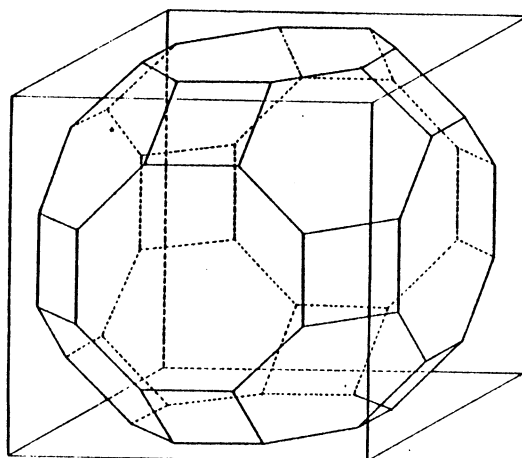
Nous terminerons par le chapitre 3 dans lequel nous exposons une étude détaillée des résolutions numériques des opérations de base concernant la gestion des mélanges binaires. Toutes les techniques utilisées sont basées sur la méthode de la recherche dichotomique.



Chapitre 1

FORMULATION GENERALE

- 1) Définitions préliminaires
- 2) Propriétés de l'ensemble des mélanges
- 3) Traduction géométrique des opérations physico-chimiques dans un système physico-chimique
 - Opérations d'inclusion et d'intersection
 - Opérations de projection



grand rhombicuboctaèdre



1) Définitions préliminaires

Toutes les définitions citées dans ce chapitre sont développées dans les articles de P.Valentin [20] et [21].

Système physico-chimique

Un système est le contenu matériel et énergétique d'un volume déterminé et fini dans notre espace usuel. Ce volume peut être de forme compliquée et enclos dans une surface fermée. Nous considérons plus spécialement les systèmes formés de mélanges de fluides et de solides. Cette définition est globale et ne permet pas une caractérisation précise du contenu du système. Pour y parvenir on procède par une division du système en sous-systèmes dont les interactions doivent être spécifiées. Cela revient à définir une partition du volume du système en régions par une famille de volumes disjoints.

Notons Σ un système et V son volume. Ω une partition de Σ en p volumes $\Omega = (\Omega_1, \Omega_2, \dots, \Omega_p)$ avec $\Omega_i \cap \Omega_j = \emptyset$ et $\Omega_1 \cup \Omega_2 \dots \cup \Omega_p = V$

Nous nous intéressons dans ce travail uniquement aux partitions des volumes. La partition doit être la plus simple possible, on choisira par exemple les régions dont les propriétés sont homogènes dans le système.

Analyse

Notre objectif est de caractériser le système, donc d'évaluer le contenu de chaque région de la partition. Sans détailler la façon d'examiner les régions, nous supposons que le contenu d'un système est constitué d'une famille finie de produits ou d'espèces. On note n leur nombre.

L'analyse d'une espèce i se ramène à évaluer sa quantité dans chaque région. Une analyse sera alors une application qui associe à une région un scalaire non négatif et elle devra satisfaire aux conditions d'extensivité suivantes:

- L'union de deux ou plusieurs régions donne pour résultat d'analyse la somme des résultats des analyses des régions initiales.

- La fraction λ de volume d'une région donne un résultat d'analyse égal à λ fois plus petit que celui de la région initiale.

A partir de ces n analyses on voit apparaître deux espaces vectoriels de dimension finie. La détermination de n est très difficile en pratique que nous ne discuterons pas dans cette étude.

Espace des quantités

A chaque région du volume Ω_i correspond un vecteur N_i de n scalaires positifs que l'on appelle: vecteur quantité. Chaque scalaire mesure la quantité d'une espèce dans cette région. L'ensemble de ces vecteurs est inclus dans un espace vectoriel de dimension n que l'on note E . A la partition Ω , on fait correspondre p vecteurs quantités.

Espace de ségrégation

On appelle ségrégation l'opération de partition d'une espèce en p quantités à distribuer entre p régions. Cette opération engendre des vecteurs à p scalaires positifs. On associe aux espèces présentes dans le système n vecteurs de ségrégation. L'ensemble de ces vecteurs est inclus dans un espace vectoriel de dimension p que l'on note S .

Ensemble des mélanges associé à une partition

L'ensemble des mélanges est l'ensemble des vecteurs quantités que l'on peut former par des opérations de sous-partition et d'union des régions ayant un sens physique. On considère que toute région est homogène et que la quantité de chaque vecteur intervenant dans le mélange est définie par le volume de la partie de la région correspondante. L'ensemble de ces vecteurs se traduit algébriquement par l'ensemble C suivant:

$$C = \left\{ M / M = \sum_{i=1}^p \lambda_i N_i, 0 \leq \lambda_i \leq 1 \right\}$$

2) Propriétés de l'ensemble des mélanges

L'ensemble défini plus haut, traduit la capacité de fabrication des mélanges que l'on peut avoir à partir d'une partition donnée du système. Plus la partition est fine, plus le degré de liberté d'action est grand. Cet ensemble fournit aussi une bonne caractérisation de l'état ou bien de la répartition: "séparation" des espèces dans ce système. On montre que toutes les opérations physico-chimiques, pour analyser les différentes espèces, se traduisent par des opérations algébriques simple à formuler sur l'ensemble des mélanges correspondants. Nous nous intéressons particulièrement dans ce travail à la faisabilité des mélanges fixés a priori que l'on appelle mélanges objectifs.

L'ensemble **C** des mélanges possède plusieurs propriétés géométriques et topologiques intéressantes. On le rencontre d'abord dans la nature en cristallographie. Des cristaux à trois dimensions ont la même structure que l'ensemble **C**. En économie, **W.Hildenbrand** (1981) dans [11] montre son utilité dans le calcul de la fonction de production efficace. En automatique il représente l'ensemble des états atteignables en un temps donné pour toutes les commandes réalisables.

Dans le domaine d'analyse convexe cet ensemble est appelé zonotope. Plusieurs auteurs ont étudié ses différentes propriétés et transformations. Nous citons **Coxeter** (1948) dans [5] qui développe essentiellement ces ensembles à trois dimensions: "Zonohedra". **McMullen** (1971) dans [13] et **Shephard** (1972) dans [23] introduisent la notion de zonotope associé. La technique consiste à transformer un zonotope à n dimensions et engendré par p vecteurs en un zonotope associé à $p-n$ dimensions et engendré par p vecteurs. Ce dernier permet de décrire facilement toutes les propriétés combinatoires du zonotope initial. Malheureusement, dans les problèmes de la séparation, le nombre p de régions est souvent très grand.

Nous rappelons dans la suite certaines propriétés élémentaires, mais nécessaires, pour avoir une idée sur la structure géométrique de cet ensemble.

Remarque

Dans toute la suite, on identifie l'espace vectoriel des quantités à l'espace affine issu de l'origine.

C est un zonotope

On voit facilement que **C** est un convexe et qu'il est contenu dans l'octant positif de **E**. Si on note D le point distal obtenu par la sommation de tous les vecteurs N_i on montre que **C** est inclus dans le paralléloéde de **E**, dont la diagonale principale est le segment $[0,D]$. Il est clair aussi que **C** est l'intersection d'un nombre fini de demi-espaces fermés. L'ensemble **C** est alors un polytope de **E**.

Pour le voir, il suffit de prendre une sous-famille de $(n-1)$ -vecteurs indépendants de $(N_i, i=1, \dots, p)$, de considérer l'hyperplan contenant ces vecteurs et de déterminer l'hyperplan d'appui de **C** parallèle à ce dernier. Le convexe **C** est alors contenu dans un demi-espace de cet hyperplan. L'intersection de ces deux ensembles n'est pas vide et est appelée une

facette du convexe C. On verra plus loin que celle-ci est le translaté d'un convexe des mélanges.

La définition précédente des facettes permet de représenter le convexe C d'une façon extérieure par des inégalités. Chaque inégalité correspond à un hyperplan et traduit que C est contenu dans un demi-espace borné par celui-ci.

Le convexe C est appelé **zonotope**. Ce dernier terme signifie que C 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 $\{N_i\}$.

C possède une structure symétrique

Tout mélange M de C s'exprime par définition:

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

Notons A la matrice $n \times p$ dont les colonnes sont les vecteurs N_i :

$$A = [N_1, N_2, \dots, N_p]$$

Un point M de C doit alors vérifier la relation $A\lambda = M$ avec $\lambda \in [0,1]^p$. Il est clair que le mélange $A(1-\lambda) = M^*$ appartient aussi à C et que $(M + M^*)/2$ est un point fixe pour tout point M de C. Ce point est le milieu de la diagonale principale de C que l'on note I. C possède alors le point I comme centre de symétrie. En outre, toute facette de C est le translaté d'un convexe de mélange engendré par une sous-famille de $\{N_i\}$ et possède aussi un centre de symétrie.

Structure géométrique et topologique de C

Le convexe des mélanges peut être défini comme la somme de Minkowski des segments $[N_i]$.

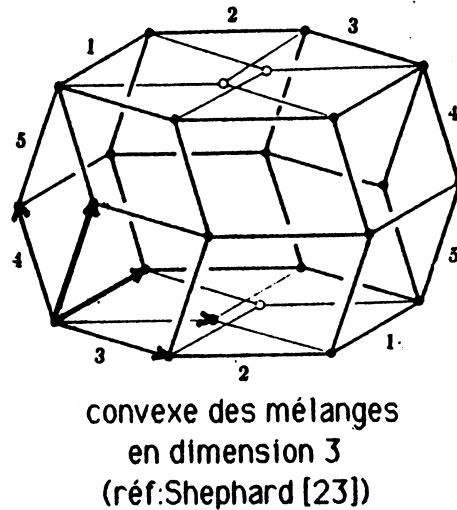
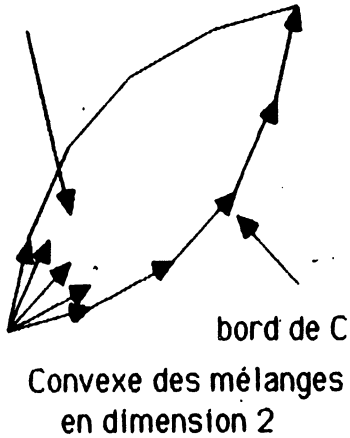
$$C = [N_1] + [N_2] + \dots + [N_p]$$

L'opération somme est appliquée sur les ensembles ou segments $[N_i]$. Cette définition de C est équivalente à la première et traduit le fait que C est l'enveloppe convexe des 2^p points suivants:

$$\sum_{i=1}^p \varepsilon_i N_i \quad , \quad \varepsilon_i = 0 \text{ ou } 1$$

Les figures suivantes montrent des convexes de mélanges à 2 et à 3 dimensions.

vecteurs quantités



Sommets de C

On définit un sommet d'un polytope C de E par un point de la frontière de C telle que toute droite passant par ce point ne recoupe pas C de part et d'autre de ce point. Nous allons voir que les sommets sont les 0-faces de C, et qu'ils sont de la forme:

$$\sum_{i=1}^p \varepsilon_i N_i, \quad \varepsilon_i = 0 \text{ ou } 1$$

Nous classons les points de cette forme en trois groupes: les sommets, les points frontières et les points intérieurs.

Faces de C

D'une manière générale, l'intersection de C et d'un hyperplan d'appui peut être de dimension inférieure à n-1. On définit une r-face de C par l'ensemble des points du bord de C, contenus dans une variété linéaire de dimension r. Tous les sous-espaces de E de dimension r+1 et qui contiennent cet ensemble (r-face) ne recoupent pas le convexe C de part et d'autre de celui-ci.

Soit F une face de C; on montre que F est de la forme

$$F = [N_{i_1}] + [N_{i_2}] + \dots + [N_{i_r}] + \varepsilon_{r+1} N_{i_{r+1}} + \dots + \varepsilon_p N_{i_p} \quad \text{où } \varepsilon_i = 0, 1$$

En effet, soit H un hyperplan d'appui $H = \{x \in E / \langle x, u \rangle = \alpha\}$. Soit H^*

l'hyperplan parallèle à H et passant par l'origine. Certains vecteurs de la famille (N_j) notés N_{1k} , $k=1, \dots, r$ sont dans H^* et les autres, sont dans les demi-espaces ouverts de E et borné par H^* .

Notons $(\epsilon_j N_{1j}, j=r+1, \dots, p)$ ceux qui appartiennent à H^* où $\epsilon_j = 1$ si $N_{1j} \in H^*$, 0, sinon. Alors $\forall x \in C \quad \langle x, u \rangle \leq \epsilon_{r+1} \langle N_{1r+1}, u \rangle + \dots + \epsilon_p \langle N_{1p}, u \rangle$. c.q.f.d.

Chaque face est le translaté d'un convexe de mélange. La propriété de symétrie centrale du convexe s'étend alors à toutes les faces associées aux sous-familles de la famille initiale.

Les 0-faces sont les sommets de C, les 1-faces ses arêtes, les 2-faces des polygones à symétrie centrale, etc...

Caractérisation de C

La symétrie centrale de C ne suffit pas pour montrer qu'il est un convexe de mélange (ou zonotope). Pour s'en convaincre, il suffit de coller par la base deux pyramides (dans \mathbb{R}^3) identiques. Leur base est un carré. Cet ensemble possède une symétrie centrale et il n'est pas un convexe de mélange.

Nous citons ici le théorème d'Aleksandrov (1933) pour $n = 3$ et généralisé par Shephard (191976) dans [15]:

théorème

Un polytope dont toutes les facettes ((n-1)-faces) sont de symétrie centrale est lui même de symétrie centrale.

On peut appliquer ce théorème sur toutes les faces de C, ce qui permet de caractériser la classe des zonotopes.

corollaire

C est un convexe de mélange si et seulement si toutes ses 2-faces possèdent une symétrie centrale.

3) Traduction algébrique des opérations physico-chimiques dans un système

On rappelle que toute opération mathématique sur le convexe des mélanges est la traduction d'une opération sur le système. On s'intéresse particulièrement aux opérations d'inclusions, d'intersections et de projections de C par rapport à un autre ensemble de points. Nous citons

dans la suite plusieurs exemples permettant de saisir l'importance pratique de la résolution de ces problèmes.

Opérations d'inclusion et d'intersection

L'intersection du convexe C avec un sous-ensemble de E ou bien avec un sous-espace vectoriel correspond à l'ensemble des mélanges répondant à une ou plusieurs contraintes. On s'intéresse dans ce travail particulièrement aux contraintes linéaires. L'intersection ne conserve pas la propriété de symétrie. Nous citons dans la suite certaines applications pratiques rencontrées en gestion des mélanges.

1- Faisabilité d'un mélange M quelconque

En pratique, on souhaite savoir si un mélange M quelconque, fixé a priori, est fabricable à partir de la famille initiale (N_i) . Algébriquement, cela revient à chercher si le point M appartient ou non au convexe C . Ce problème de recherche est largement discuté dans le chapitre 2. Nous l'avons transformé en plusieurs problèmes d'optimisation, ce qui nous a permis d'utiliser les méthodes et les algorithmes d'optimisation convexe.

2-Fabrication optimale d'un mélange

Ce problème nécessite d'abord la définition d'un critère d'optimalité de fabrication. En pratique et en l'absence de critère économique, on souhaite réaliser M de façon qu'on puisse fabriquer le plus possible de mélanges à partir du reste de la famille. La solution est simple pour $n = 2$ et revient à chercher un arc du bord de C dont la corde est le segment M . L'algorithme de résolution est développé dans le chapitre 3. Cette opération de troncature n'est applicable que pour $n=2$, par contre la situation devient plus compliquée pour $n>2$.

3- Faisabilité des mélanges ayant $n-1$ contraintes linéaires

Ce problème d'intersection est d'un intérêt important du point de vue pratique et théorique. Il est équivalent à la détermination de l'intersection d'une droite et du convexe des mélanges. Nous trouvons, en pratique, plusieurs applications importantes:

- On peut s'intéresser à la détermination de la quantité maximale d'un mélange de composition donnée. Il revient au même de chercher l'intersection de C et de la droite passant par l'origine et correspondant à cette composition.
- On peut aussi s'intéresser à la détermination de la quantité maximale d'une espèce en sachant les quantités des $n-1$ espèces restantes.

4- Faisabilité des mélanges ayant $n-2$ contraintes linéaires

Ceci revient à la détermination de l'intersection d'un plan et du convexe C . La résolution de ce problème, outre son application pratique, fournit un moyen de représentation du convexe C sur un écran en coupant C par des plans parallèles.

5- Faisabilité des mélanges ayant une contrainte linéaire

L'ensemble des points soumis à une contrainte linéaire est un hyperplan de E . L'intersection du convexe C et d'un hyperplan ne passant pas par l'origine est l'ensemble des mélanges, ayant une quantité donnée formée d'une combinaison linéaire des espèces de base.

Opérations de projection

La projection du convexe C sur un sous-espace de E , parallèlement à un sous-espace complémentaire, est aussi un convexe de mélange. Ce dernier est engendré par la famille des vecteurs projections des N_i .

Il est clair que la projection de C sur les sous-espaces de base ou canoniques, de E , parallèlement à leurs supplémentaires, fournit une caractérisation restrictive du système.

Conclusion

Il est clair qu'une maîtrise géométrique et mécanique de ces opérations élémentaires constitue un outil de base pour résoudre des problèmes plus complexes.

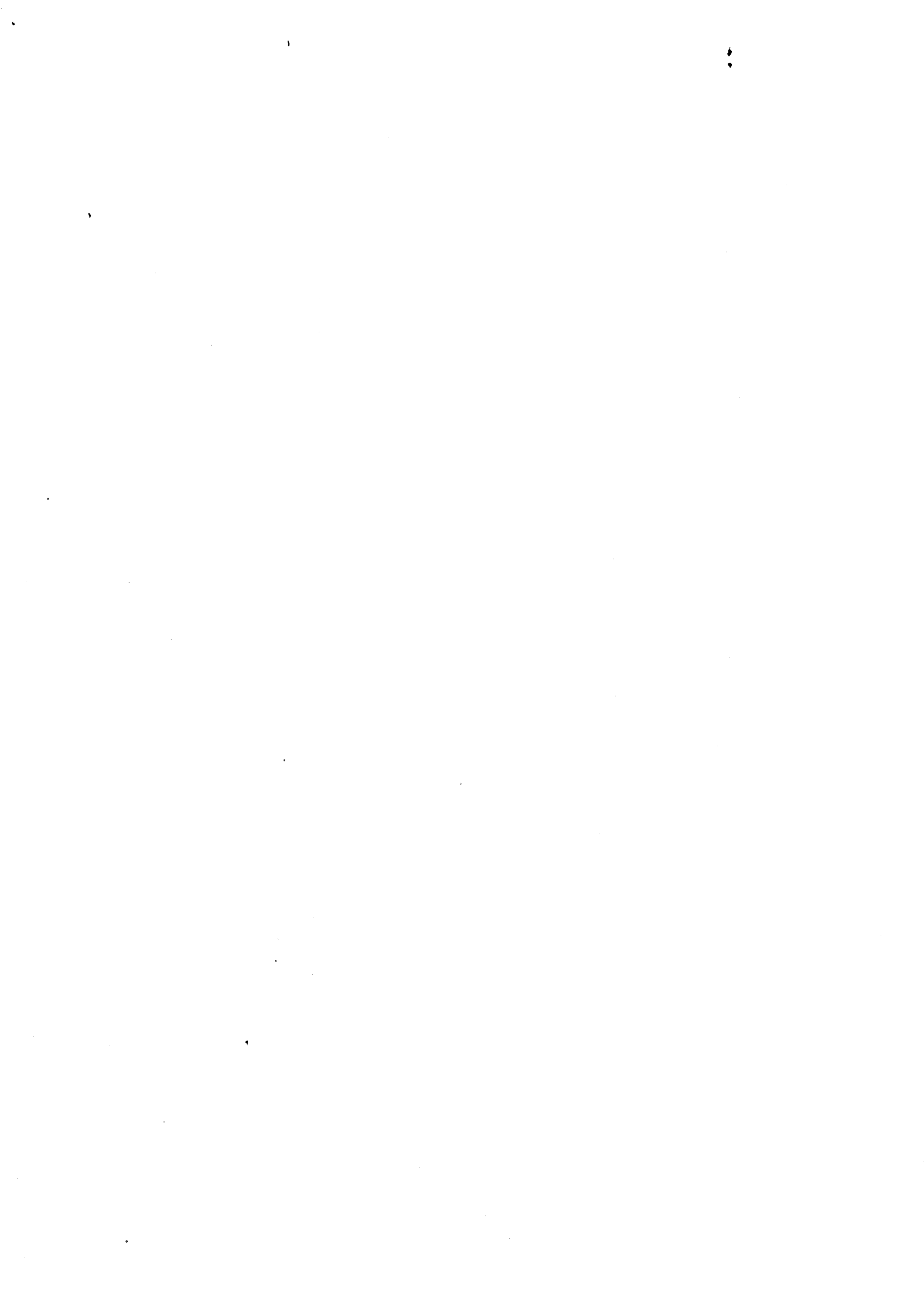
Nous nous limiterons dans le cadre de ce travail, d'une part, à la résolution de la faisabilité et d'autre part à la résolutions des problèmes plans ($n=2$).

On notera enfin que E ne possède pas de structure euclidienne, on ne sait pas définir physiquement deux mélanges "orthogonaux".

Chapitre 2

Faisabilité d'un mélange objectif

- 1- Introduction
- 2- Caractérisation d'un mélange objectif
- 3- Fabrication d'un mélange objectif
 - 3.1- Méthode de projection
 - 1- Méthode de Barycentre
 - 2- Méthode de projections successives
 - 3- Test d'arrêt
 - 3.2- Formulation de la fabrication sous la forme d'un problème d'optimisation
 - 3.3- Méthode de décomposition
 - 1- Norme euclidienne et forme quadratique
 - 2- Sous-problème d'optimisation
 - 3- Méthode en série
 - 4- Méthode en parallèle
 - 5- Interprétation géométrique des itérations
 - 3.4- Formulation linéaire
 - 1- Norme L_1 et méthode primale
 - 2- Simplexe révisé à variables bornées
 - 3- Interprétation géométrique des itérations
 - 4- Méthode duale
- 4- Application numérique et synthèse



1) Introduction

Tout procédé de séparation est caractérisé par son convexe de mélange. Ce dernier fournit un moyen efficace et commode pour évaluer la qualité de la séparation. Pour ce faire, on fixe a priori un objectif ou un mélange de qualité spécifiée, et on veut savoir si ce procédé peut générer ou produire cet objectif. Ce problème ainsi posé, nous incite à répondre à deux questions:

1) Est-il possible de fabriquer un mélange M à partir d'une famille de vecteurs quantités produites par la séparation et qui génère le convexe de mélange correspondant ? (Ceci, sans chercher la manière de le fabriquer).

2) Comment peut-on mélanger ces vecteurs quantités pour atteindre cet objectif ? Autrement dit, quelle est la technique de fabrication et la décomposition de M par rapport à ces vecteurs?

Dans la deuxième section, nous présenterons une caractérisation de l'appartenance d'un objectif à un convexe de mélange et par conséquent la possibilité de le réaliser. Nous ferons intervenir les théorèmes classiques de l'analyse convexe.

Dans la troisième section nous aborderons les différentes techniques ou procédés produisant la décomposition d'un mélange objectif en fonction de la famille des vecteurs quantités.

Dans un premier temps, nous essaierons de faire la synthèse de deux classes de méthodes: la première est horizontale et se base sur la méthode des "projections successives", la deuxième est verticale et se base sur la décomposition du problème primal en sous-problèmes de mélange simple à résoudre. Dans un dernier temps, nous le ramènerons à un problème de programmation linéaire à variables bornées. Cette dernière est en quelque sorte le couplage des deux précédentes et se fait en bloc.

Enfin, on essaiera de faire la synthèse et de comparer les trois techniques à partir de certains résultats et applications numériques.

2) Caractérisation d'un mélange objectif

Soit M un mélange de E , (N_i) , $i=1, \dots, p$ une famille de vecteurs quantités de E et C le convexe des mélanges correspondant.

Dans cette partie, nous étudions la question de l'appartenance de M à C sans tenir compte de sa décomposition par rapport à la famille (N_i) , $i=1, \dots, p$.

Rappelons que C est un polytope de E et qu'il est l'intersection d'un nombre fini de demi-espaces fermés. Il suffit alors d'avoir un moyen pour générer ces demi-espaces et de vérifier que M leur appartient.

Rappelons d'abord certains résultats de l'analyse convexe sur la séparation d'un point et d'un convexe fermé.

Théorème (séparation d'un point et d'un convexe fermé)

Soit C un convexe fermé non vide dans E et $M \notin C$. Alors il existe un vecteur u et un scalaire α tel que

$$\langle u, M \rangle > \alpha \quad \text{et} \quad \langle u, x \rangle \leq \alpha, \quad \forall x \in C$$

Ce théorème, montre que si M n'appartient pas à C alors il existe un hyperplan séparant le point M et C . En outre, si C est engendré par un nombre fini d'hyperplans alors il existe un hyperplan de cette famille qui les sépare.

Corollaire

Soit C un convexe non vide et fermé de E , alors 1) et 2) sont équivalents.

1) $M \in C$

2) $\forall u \in E, \langle u, M \rangle \leq \sup\{\langle u, x \rangle, x \in C\}$

démonstration:

2) \Rightarrow 1)

Supposons que $M \notin C$; alors il existe un α_1 et u_1 tel que

$\langle u_1, M \rangle > \alpha_1$ et $\langle u_1, x \rangle \leq \alpha_1, \forall x \in C$. Par 2), on a $\alpha \leq \alpha_1$ et $\langle u_1, M \rangle \leq \alpha$ d'où $\langle u_1, M \rangle \leq \alpha_1$, ce qui est contradictoire; d'où $M \in C$.

1) \Rightarrow 2)

Il suffit d'utiliser la contraposée du théorème.

Remarques:

1- On peut remplacer la proposition 2) du corollaire par

$$(1) \quad \forall u \in E, \inf(\langle u, x \rangle, x \in C) \leq \langle u, y \rangle \leq \sup(\langle u, x \rangle, x \in C)$$

Ceci donne directement une première caractérisation de l'appartenance d'un mélange donnée M au convexe des mélanges C

2- Pour un vecteur u donné, le calcul de la quantité $\inf(\langle u, x \rangle, x \in C)$ ou de la quantité $\sup(\langle u, x \rangle, x \in C)$ est simple et nécessite np multiplications et (n-1)p additions.

En effet $\sup(\langle u, x \rangle, x \in C) = \sup(\langle u, A\lambda \rangle, \lambda \in [0,1]^p)$
 et la quantité $\langle u, A\lambda \rangle = \langle {}^tAu, \lambda \rangle$

$$\text{d'où (2) } \sup(\langle u, x \rangle, x \in C) = \sum_{i=1}^p (\langle u, N_i \rangle)_+$$

$$\text{et de même } \inf(\langle u, x \rangle, x \in C) = \sum_{i=1}^p (\langle u, N_i \rangle)_-$$

3- Signalons que le théorème de *CARATHEODORY* fournit une autre caractérisation de l'appartenance de M à C. mais elle est moins intéressante car elle ne permet pas une vérification directe.

Dans le chapitre 1, nous avons montré qu'à chaque sous-famille de (n-1) vecteurs linéairement indépendants correspond deux hyperplans d'appuis de C, parallèles. Notons q le nombre minimal de ces hyperplans. Ce nombre est borné par $2C_p^{n-1}$. Il est clair que le problème de l'appartenance de M à l'intersection d'un nombre fini de demi-espaces est un problème de recherche. Il consiste à chercher s'il existe un hyperplan de la famille précédente séparant M et C. Dans le domaine de l'algorithme géométrique, on envisage deux types de méthodes de recherche:

1- Une recherche séquentielle qui consiste à vérifier les q inégalités associées aux hyperplans générant C. Cette méthode nécessite une place mémoire très grande, par contre le temps de recherche est de l'ordre de O(npq).

2- Une recherche rapide et généralisant la recherche binaire (voir [6],[18] et [25]). Cette méthode suppose que toutes les informations nécessaires aux hyperplans soient stockées et organisées convenablement pour faciliter la recherche. Le temps de recherche est de l'ordre de

$O(f(n)\log(q))$, où f est une fonction de n [6].

Il s'avère que les deux méthodes nécessitent la détermination de toutes les facettes de C , ce qui est coûteux. En conséquence, nous préconisons, pour répondre à cette question d'utiliser les méthodes développées dans la suite.

Remarque

Dans le cas des convexes de mélange de dimension 2, nous allons proposer, dans le chapitre 3, un algorithme rapide en $O(p\log(p))$ pour répondre à la question de l'appartenance.

3) Fabrication d'un mélange objectif

La fabrication d'un mélange objectif telle qu'elle est posée à l'expérimentateur est très compliquée. En effet, ce dernier espère disposer d'une technique dynamique permettant, au cours du processus de la séparation, de décider si le produit sortant doit participer à la constitution de M ou non.

Si on se contente de certains mélanges de base (vecteurs quantités), définis a priori pour atteindre M, dans ce cas le problème devient purement algébrique, dans le sens où on cherche si M est décomposable ou non par rapport à ces vecteurs.

Dans cette section, nous allons formuler la faisabilité de M sous plusieurs formes, ce qui nous a conduit à considérer plusieurs méthodes vues sous un angle différent (méthode de projection, de décomposition et de programmation linéaire). Nous essaierons dans ces deux dernières méthodes d'exploiter la dualité et de montrer son apport par rapport au problème primal. Nous terminerons par une synthèse et une comparaison de ces trois classes en décrivant les avantages et les inconvénients de chacune d'elles quand elle est adaptée au problème.

3.1) Méthode de projection.

On rappelle qu'un mélange M est fabricable ou faisable à partir d'une famille de mélanges $\{N_i, i=1, \dots, p\}$ s'il est décomposable par rapport à cette dernière avec des coefficients appartenant à $[0, 1]$.
Le problème de la fabricabilité peut s'exprimer par:

$$(Q) \quad \exists ? \lambda \in [0, 1]^p / A\lambda = M$$

La première idée intuitive est de décomposer le problème en sous-problèmes, chaque d'eux correspondant à la fabrication d'une espèce (ou composante) de M. Puis de faire évoluer cette technique d'une manière séquentielle (en série) ou simultanée (en parallèle) jusqu'à ce que soit atteint notre objectif: fabrication totale de toutes les espèces de M. (si M est fabricable).

Nous allons voir que ce procédé peut se formuler par des projections et qu'il converge dans les deux cas (série ou parallèle) vers M.

En effet, ces idées nous font penser aux méthodes de *Cimmino* et de *Kaczmarz* qui consistent à résoudre des systèmes linéaires de grande taille en utilisant la projection des itérés sur les hyperplans correspondant aux équations du système (voir *N. Gastine* [8]).

Avant de décrire ces méthodes, précisons quelques notations. Posons:

$$C_0 = [0, 1]^p \quad \text{et} \quad C_i = \{ \lambda \in \mathbb{R}^p \mid \langle A_i, \lambda \rangle = M_i \}, \quad i=1, \dots, n$$

où A_i désigne le vecteur de ségrégation de l'espèce i . Les composantes de A_i sont les quantités de l'espèce i dans les p vecteurs quantités.

Le problème (Q) se ramène à la recherche d'un point λ appartenant à l'intersection des convexes $C_i, i=0, \dots, n$.

$$" \exists \lambda \in S \mid \lambda \in \varphi(M) = \bigcap_{i=0}^n C_i "$$

Nous signalons que pour un λ donné, la projection sur $C_i, i=0, \dots, n$ est explicite et facile à formuler:

$$P_{C_i}(\lambda) = \lambda + (M_i - \langle A_i, \lambda \rangle / \|A_i\|^2) A_i, \quad i=1, \dots, n$$

$$(P_{C_i}(\lambda))_k = \begin{cases} 0 & \text{si} \quad \lambda_k \leq 0 \\ \lambda_k & \text{si} \quad 0 \leq \lambda_k \leq 1 \\ 1 & \text{si} \quad 1 \leq \lambda_k \end{cases}$$

Ainsi la méthode de projection s'avère très facile à mettre en oeuvre. En conséquence, et vu l'abondante littérature traitant de ces méthodes, cela nous a conduit d'une part à faire une synthèse, en ébauchant uniquement les grandes lignes des algorithmes et, d'autre part, de renvoyer pour toutes propriétés et démonstrations aux articles correspondants.

1- Méthodes de Barycentre.

Le processus itératif s'inspire du procédé de *Cimmino* (*N. Gastine* [8]), généralisé pour des convexes quelconques dans *Auslender* [1] *G. Pierra* [17], *Yair. Censor* [4] et *Alvad. Pierró* [2].

Ce procédé consiste à projeter un même point λ^k sur les n convexes C_i et C_0 . Puis à corriger ces projections en prenant leur barycentre. En effet, il suffit de remarquer que le barycentre de ces points est plus près de la solution que λ^k .

On peut formuler ces étapes dans l'algorithme suivant:

Algorithme 1

- 1- Soit λ^0 quelconque.
- 2- le passage de λ^k à λ^{k+1} s'effectue de la façon suivante.

on pose $\lambda^{k,i} = P_{C_i}(\lambda^k)$, $i=0, \dots, n$

$$\lambda^{k+1} = \left(\sum_{i=0}^n \lambda^{k,i} \right) / (n+1)$$

- 3- si test d'arrêt stop, sinon aller en 2.

On a donc successivement une phase 1 de calcul des $n+1$ projections, puis une phase 2 de coordination. On désigne cet algorithme sous le nom de "méthode du barycentre".

propriété

Si $\varphi(M) \neq \emptyset$ alors la suite (λ^k) converge fortement vers un élément de $\varphi(M)$.

Démonstration (voir **G.Pierra** [17]).

Dans notre problème, on distingue deux types de convexes de nature différentes, C_0 et C_i , $i=1, \dots, n$. On relève dans la thèse de **G.Pierra** une variante de l'algorithme précédent, destinée à exploiter cette spécificité.

Algorithme 2

- 1- partant de λ^0 quelconque.
- 2- le passage de λ^k à λ^{k+1} s'effectue de la façon suivante:

$$\lambda^{k+1} = P_{C_0} \left(\left(\sum_{i=1}^n \lambda^{k,i} \right) / n \right) \text{ et } \lambda^{k,i} = P_{C_i}(\lambda^k), \quad i=1, \dots, n$$

- 3- si test d'arrêt stop, sinon aller en 2.

On a dans ce cas une phase de résolution en parallèle spécifique aux hyperplans C_i , $i=1, \dots, n$, une phase 2 de coordination, puis une phase 3 qui consiste à projeter les itérés dans C_0 .

La propriété de convergence précédente est toujours vérifiée pour la suite

(λ^k) généré par l'algorithme 2.

2- Méthode de projections successives.

Cette méthode, de *Gubin et al* généralise celle de *Kacmarz (N. Gastinel 8)*), qui consiste à résoudre un système linéaire de grande taille et la méthode de relaxation de *Agmon, Motzkin et Schoenberg* qui consiste à chercher un point d'intersection d'un ensemble fini de demi-espaces. Le processus est très simple et consiste à projeter successivement le point λ^k sur un ensemble $C_{i_k}, i_k=0, \dots, n$ et prendre un point dans le segment connectant λ^k et son projeté sur C_{i_k} . On prend en général les projections dans un ordre cyclique.

L'algorithme est décrit dans *Yair Censor* [4], il suffit de remplacer l'étape 2 dans l'algorithme 2 par:

2- le passage de λ^k à λ^{k+1} s'effectue par :

$$\lambda^{k+1} = \lambda^k + \mu_k (P_{C_{i_k}}(\lambda^k) - \lambda^k)$$

$$\varepsilon_1 < \mu_k < 2 - \varepsilon_2, \quad \varepsilon_1 > 0, \quad \varepsilon_2 > 0 \quad \text{et} \quad i_k = k \bmod(n) + 1$$

Le paramètre de relaxation μ_k est une suite de réels appartenant à $[0, 2]$. Ce paramètre ajoute un degré de liberté du choix de l'itération suivante. Il n'existe pas une stratégie optimale du choix de ce paramètre; la plupart du temps on le détermine d'une façon empirique.

Ces méthodes sont dites horizontales. Le passage d'un itéré à un autre nécessite toutes les informations d'une ligne (ou même toute la matrice A). En outre la matrice A n'est pas creuse ce qui est susceptible de perturber les déplacements des itérés et par conséquent de ralentir la convergence de ces méthodes (ce qui a été constaté en pratique en comparaison avec les autres méthodes de fabrication).

Cependant, ces méthodes sont faciles à programmer et nécessitent peu de place mémoire.

3- Test d'arrêt

Lors de la résolution d'un problème d'optimisation, les critères d'arrêt généralement utilisés pour arrêter le processus itératif sont:

$$|\lambda_i^{k+1} - \lambda_i^k| \leq \varepsilon_a \quad \text{ou} \quad |\lambda_i^{k+1} - \lambda_i^k| \leq \varepsilon_r |\lambda_i^k|, \quad i=1, \dots, p$$

ε_a et ε_r étant respectivement la précision absolue et relative sur la composante i du vecteur λ .

En pratique, nous avons constaté que ces critères sont mauvais car ils

imposent un choix arbitraire de ϵ_a et ϵ_r . S'ils sont choisis trop grands, alors le processus itératif est arrêté avant l'obtention de la solution et le résultat obtenu est mauvais. Par contre, s'ils sont choisis trop petits, le processus n'est pas arrêté assez tôt et parfois même, il boucle.

Or, ce test n'est valable que si l'on connaît a priori la précision avec laquelle sera trouvée la solution. Ce qui est le cas dans nos problèmes de faisabilité si $M \in C$. Dans ce cas nous avons utilisé le critère d'arrêt suivant:

$$|M_i^k - M_i| \leq \epsilon_a, \quad i=1, \dots, n$$

M^k étant le mélange associé au vecteur λ^k . Dans tous les programmes, nous avons combiné ces deux critères afin d'obtenir un test optimal en réduisant le nombre des itérations.

3.2) Formulation de la fabrication sous la forme d'un problème d'optimisation.

Soit λ un vecteur quelconque de $[0,1]^p$ et le mélange $A\lambda$ de C correspondant. Notons v le vecteur résidu $M - A\lambda$.

En utilisant la norme $\| \cdot \|$ définie dans E , on peut poser le problème de la recherche d'une décomposition de M sous la forme d'un problème d'optimisation. En fait, il suffit de calculer, selon la norme $\| \cdot \|$, la distance entre l'ensemble C et le point M . Cette dernière est définie par :

$$d(M, C) = \min(\| M - X \|), \quad X \in C$$

que l'on peut écrire

$$(P) \quad \begin{array}{ll} \text{minimiser}(\|v\|) & \Leftrightarrow \text{minimiser}(\|M - A\lambda\|) \\ A\lambda + v = M & 0 \leq \lambda \leq 1 \\ 0 \leq \lambda \leq 1 & \end{array}$$

Cette formulation, d'une part conduit à la proposition suivante:

$$M \in C \Leftrightarrow v^* = 0, \quad v^* \text{ est la solution de (P)}$$

et d'autre part permet d'appliquer les algorithmes d'optimisation, suivant la norme choisie, pour la recherche d'une solution de (P).

Dans la suite, nous allons étudier deux normes différentes: la norme euclidienne et la norme de la somme des modules, et les algorithmes qui en découlent.

Nous allons voir que la première norme conduit à des méthodes dites de décomposition (plusieurs variétés ont été développées). La deuxième norme conduit à une formulation linéaire du problème ce qui permet d'adapter, d'une façon optimale, la méthode du simplexe à variables bornées.

3.3) Méthode de décomposition.

Soit la fonctionnelle définie sur \mathbb{R}^p par

$$\Psi: \mathbb{R}^p \longrightarrow \mathbb{R}$$

$$\lambda \longmapsto \|A\lambda - M\|^2$$

Cette fonction est convexe et dépend du choix de la norme $\|\cdot\|$.

Le problème (P) s'écrit:

$$(P) \quad \text{Min} (\Psi(\lambda_1, \lambda_2, \dots, \lambda_p))$$

$$0 \leq \lambda_i \leq 1, \quad i = 1, \dots, p$$

Le nombre de contraintes est très grand (p est grand en général). En revanche, elles sont simples et indépendantes (chaque contrainte dépend d'une seule variable). Ainsi, nous avons pensé à exploiter cette spécificité et à décomposer le problème (P) en une succession de sous-problèmes d'optimisation, mais plus simples (i.e. que l'on sait résoudre plus facilement ou d'une manière explicite). Cette façon de résoudre de (P) nous amène à étudier les différentes méthodes de décomposition suivant le choix et le critère itératif.

1- Norme euclidienne et forme quadratique.

On prend la norme euclidienne dans E . Cette dernière mesure la distance entre deux points de E . Par contre elle n'a pas de signification physique pour un mélange. Cependant, ce choix confère à Ψ d'être dérivable et quadratique.

$$\Psi(\lambda) = \langle \lambda, {}^t A A \lambda \rangle - 2 \langle {}^t A M, \lambda \rangle + \langle M, M \rangle$$

ce qui, évidemment, simplifiera la résolution de ces sous-problèmes et permettra d'avoir une solution explicite de ces derniers.

2- Sous-problème d'optimisation de (P)

L'idée de décomposition consiste à remplacer (P) par une succession de problèmes plus simples (P_i). Autrement dit, de se ramener à des sous-espaces de dimension inférieure et de résoudre le problème localement.

Soit Ψ_i la fonctionnelle de \mathbb{R} dans \mathbb{R}^+ définie par:

$$\Psi_i(\xi) = \Psi(\lambda_1, \dots, \lambda_{i-1}, \xi, \lambda_{i+1}, \dots, \lambda_p)$$

et posons le problème (P_i):

$$(P_i) \quad \text{minimiser } (\Psi_i(\xi))$$

$$0 \leq \xi \leq 1$$

Résolution de (Pi).

C'est un problème d'optimisation à une variable sur \mathbb{R} .
La solution de (Pi) est explicite. Par dérivation on a:

$$\Psi_i'(\xi) = 0 \Leftrightarrow \xi = \langle N_i, M - \sum_{j=1}^p \lambda_j N_j \rangle / \|N_i\|^2, \quad i=1, \dots, p$$

Il en résulte que ξ^* solution de (Pi) s'écrit:

$$\xi^* = \begin{cases} 0 & \text{si } \xi \leq 0 \\ \xi & \text{si } 0 < \xi < 1 \\ 1 & \text{si } 1 \leq \xi \end{cases}$$

Après cette définition simple des (Pi), la question qui découle immédiatement c'est comment les combiner pour résoudre (P)?
On distingue deux classes différentes, que l'on reconnaît par la méthode en série et la méthode en parallèle.

3-Méthode en série.

La méthode de décomposition en série consiste à résoudre (Pi) en modifiant à chaque pas la composante correspondante. Il existe plusieurs façons pour choisir l'indice i dans chaque itération. Nous avons expérimenté deux cas, la première, dite de Gauss Seidel, est simple et consiste à choisir les indices dans un ordre cyclique; la deuxième, connue sous le nom de Southwell, détermine la composante qui nécessite le plus d'effort.

Notons g^k le gradient de Ψ en λ^k et I^k l'ensemble des indices
 $I^k = \{1, 2, \dots, p\} \setminus \{i / \lambda_i^k = 1 \text{ et } g_i^k < 0\} \setminus \{i / \lambda_i^k = 0 \text{ et } g_i^k > 0\}$

On détermine la composante i^* telle que $|g_{i^*}^k| = \max_{i \in I^k} |g_i^k|$, $i \in I^k$.

Algorithme 3: (Gauss Seidel)

- 1- initialiser λ^0 par 0.
- 2- le passage de λ^k à λ^{k+1} s'effectue par:
 - résoudre (Pi) pour $i = 1, \dots, p$
 - (Pi) $\min (\Psi(\lambda_1^{k+1}, \dots, \lambda_{i-1}^{k+1}, \xi, \lambda_{i+1}^k, \dots, \lambda_p^k))$
 - $0 \leq \xi \leq 1$

- 3- si test d'arrêt ($\lambda^k = \lambda^{k+1}$) stop, sinon aller en 2.
 L'algorithme de Soutwell consiste à remplacer l'étape 2 par
 2 * le passage de λ^k à λ^{k+1} s'effectue par:
- détermination de la composante i^*
 - calcul de g^k et I^k
 - calcul de i^* , $|g_{i^*}^k| = \max |g_i^k|, i \in I^k$
- * résolution de (P_{i^*})
 * modifier la composante i^* de λ^k

4-Méthode en parallèle.

Cette méthode consiste à résoudre tous les (P_i) pour un λ^k fixé. On a plusieurs variétés, pour passer de l'itération λ^k à l'itération λ^{k+1} suivante.

Notons $\lambda^{k,i}$ la solution du problème (P_i) . *Auslender* dans [1] cite deux possibilités. La première consiste à corriger ces solutions en prenant $\lambda^{k+1} = \sum \alpha_i \lambda^{k,i}$ où (α_i) est une suite de réels tels que: $\alpha_i > 0$ et $\sum \alpha_i = 1$.

On prend en général $\alpha_i = 1/p, i=1, \dots, p$.

La deuxième consiste à choisir la meilleure direction parmi ces p problèmes (i.e celle qui minimise le plus Ψ). Posons:

$$\Psi_{i^*}(\xi_{i^*}^*) = \min_{i=1, \dots, p} (\Psi_i(\xi_i^*))$$

et $\lambda_j^{k+1} = \lambda_j^k$ si $j \neq i^*$, $\xi_{i^*}^*$ sinon.

Algorithme 4

- 1- Initialiser λ^0 à 0.
- 2- résoudre pour $i=1, \dots, p$

$$(P_i) \min (\Psi(\lambda_1^k, \dots, \lambda_{i-1}^k, \xi, \lambda_{i+1}^k, \dots, \lambda_p^k))$$

$0 \leq \xi \leq 1$

- 3- soit ξ_i la solution de (P_i) , déterminer i^* tel que :

$$\Psi_{i^*}(\xi_{i^*}^*) = \min_{i=1, \dots, p} (\Psi_i(\xi_i))$$

et $\lambda_j^{k+1} = \lambda_j^k$ si $j \neq i^*$, $\xi_{i^*}^*$ sinon

(3'- $\lambda^{k+1} = \sum \alpha_i \lambda^{k,i}$)

- 4- si test d'arrêt stop, sinon aller en 2.

4- Interprétation géométrique des itérations.

Les itérations décrites plus haut, consistent à parcourir un chemin dans C suivant les directions (N_i) (vecteurs quantités).

Soit M^k un mélange intermédiaire. A la $k^{\text{ème}}$ itération, la résolution de (P_i) consiste à procéder par les opérations de mélange et de séparation entre les deux vecteurs quantités $M^k - \lambda_i^k N_i$ et N_i afin de fabriquer un mélange meilleur M^{k+1} (i.e résidu $\|M - M^{k+1}\| \leq \|M - M^k\|$). Ceci se traduit dans E par la projection orthogonale de M sur le segment $[M^k - \lambda_i^k N_i, M^k + (1 - \lambda_i^k) N_i]$.

Ces méthodes sont du même type que les méthodes de projection, simples à programmer et nécessitent peu de place mémoire.

3.4) Formulation linéaire

1- Norme L_1 et méthode primale

On prend comme norme dans E , la norme L_1 , et on la note $\| \cdot \|$:

$$\text{si } v \in E, \quad \|v\| = \sum_{i=1}^n |v_i|$$

Le choix de cette norme a une signification physique. En effet, soit N un mélange quelconque, $\|N\|$ représente la quantité totale des espèces du mélange N .

Dans l'espace E des quantités, la boule unité est un cross-polytope à $2n$ sommets (sur chaque axe se trouve deux sommets symétriques par rapport à l'origine), dont le bord de l'octant positif de E est le simplexe des fractions molaires.

Nous avons proposé, dans le paragraphe 3.2, de calculer la distance entre M et C afin de déterminer une décomposition $\lambda \in [0,1]^p$ tel que $A\lambda = M$. Sans réduire en rien la décomposition de M , on peut remplacer C par l'intersection de C et de l'hypercube $[0,M]^n$. Ceci nous permet d'ajouter les contraintes $0 \leq v_i, i=1, \dots, n$ et par conséquent de supprimer les valeurs absolues dans la fonction objectif de (P). Ce dernier devient:

$$(P.L) \quad \underset{i=1}{\text{minimiser}} \sum_{i=1}^n v_i \quad \Leftrightarrow \quad \underset{i=1}{\text{minimiser}} \langle M - A\lambda, 1 \rangle$$

$$\begin{array}{ll} A\lambda + v = M & A\lambda \leq M \\ 0 \leq v \leq M, \quad 0 \leq \lambda \leq 1 & 0 \leq \lambda \leq 1 \end{array}$$

La fonction objectif est linéaire et (P.L) est un programme linéaire écrit sous forme standard. L'ensemble des variables (λ, v) est dans $S \times E$. L'application directe de la méthode du simplexe est très coûteuse, vu le nombre de contraintes et de variables, (p est en général très grand) et peut conduire à des instabilités.

Or nous avons n contraintes d'égalités (n est faible) et les $n+p$ contraintes restantes sont simples et limitent le domaine des (λ, v) . Ainsi, nous pouvons résoudre (P.L) comme un programme linéaire ne comportant que n contraintes d'égalités et $p+n$ variables. Pour ce faire, il suffit de modifier très peu l'algorithme du simplexe révisé et de considérer uniquement les bornes des $n+p$ contraintes:

$$0 \leq v \leq M \quad \text{et} \quad 0 \leq \lambda \leq 1$$

L'adaptation du simplexe fera l'objet du paragraphe suivant.

2-Simplexe révisé à variables bornées

Posons le problème suivant:

$$\text{minimiser } \langle c, x \rangle$$

$$Bx = b$$

$$\alpha \leq x \leq \beta$$

où $B = [A, I]$ est une matrice $(n, n+p)$

$$x = {}^t(\lambda, v), \quad \alpha = 0, \quad \beta = {}^t(1, M), \quad b = M \text{ et } c = {}^t(0, 1).$$

Définitions

1- soit $I = (1, 2, \dots, p, \dots, p+n)$,

$$J = (i_1, \dots, i_n), \quad i_k \in I \quad \text{et} \quad J' = I \setminus J.$$

2- On dit que J est une base si la sous-matrice de B correspondant à J notée B^J est carré non-singulière.

3- Une solution du système $Bx = b$ telle que pour tout $j \in J'$ on a $x_j = \alpha_j$ ou bien β_j est dite solution de base associée à J .

Propriété: Toute solution de base est un sommet du polyèdre :

$$\{ x \in Sx_E / Bx = b \}$$

Démonstration [7].

Pour chaque base, on peut transformer le problème (P.L) de la façon suivante:

$$\text{minimiser } \langle c^J - \pi B^J, x_J \rangle - \pi b$$

$$x_J + (B^J)^{-1} B^J x_{J'} = (B^J)^{-1} b$$

$$\alpha \leq x \leq \beta$$

où π est le multiplicateur de Lagrange et $\pi B^J = c^J$. (π est vecteur ligne de E)

théorème: Une solution de base relative à J est une solution optimale de (P.L) si les conditions suivantes sont satisfaites:

$$1) \quad (c^J - \pi B^J) > 0 \Rightarrow x_j = \beta_j$$

$$2) \quad (c^J - \pi B^J) < 0 \Rightarrow x_j = \alpha_j$$

Démonstration voir [19].

Processus itératif

On part d'une base réalisable et de sa solution associée. Si les conditions d'optimalité sont vérifiées, on est à un optimum, sinon on pose $s (s \in J')$ tel que (1) ou (2) soient non vérifiées.

Supposons que $x_s = \alpha_s$ et $(c^s - \pi B^s) > 0$ (le même raisonnement s'effectuera dans le cas où $x_s = \beta_s$ et $(c^s - \pi B^s) < 0$).

Le processus consiste à faire croître x_s à partir de α_s sans changer les valeurs des variables hors base et en ajustant les variables de base de telle sorte que x reste solution de $Bx = b$ et qu'il reste entre les bornes α et β .

Deux cas peuvent se présenter:

1- x_s atteint sa borne supérieure β_s , dans ce cas J reste inchangé et la condition (2) est satisfaite pour s .

2- L'une des variables de base x_r atteint une borne. Dans ce cas r sort de la base et s est introduit dans J .

$$J = J \cup \{s\} / \{r\}.$$

On s'arrête dès que (1) et (2) sont satisfaites.

Remarque: La formulation du problème permet au simplexe de tourner sans passer par la première phase (qui consiste à calculer une base et une solution de base réalisable).

On initialise: $J = \{p+1, p+2, \dots, p+n\}$, $v = M$ et $\lambda = 0$.

Algorithme.

1- **initialisation :**

$$\alpha = 0 \quad , \quad \beta = [1, M] \quad \text{et} \quad \lambda = [0, M]$$

2- **calculer π :** $\pi A^J = c^J$.

3- **chercher s tel que:**

$$dc = c^s - \pi A^s = \text{Max} (c^j - \pi A^j / \lambda_j = \beta_j , \quad -c^j + \pi A^j / \lambda_j = \alpha_j)$$

si $dc < 0$ stop

4- **calculer T :** $A^J T = A^s$ et $B = \lambda^J + \lambda_s T$

5- **si $(\lambda_s = \beta_s)$ alors**

$$\lambda'_s = \min (B_j - \beta_j / T_{j < 0} , B_j / T_{j > 0})$$

si $\lambda'_s > \alpha_s$ alors

- $\exists r$ tel que λ_r atteint une borne
 $J = J U(s)/(r)$, $\lambda_s = \lambda'_s$ et calculer λ^j , aller en 2-
sinon ($\lambda_s = \alpha_s$), J reste fixe et calculer λ^j , aller en 3-
 6- **si** ($\lambda_s = \alpha_s$) **alors**
 $\lambda'_s = \text{Max}(B_j - \beta_j / T_j > 0, B_j / T_j < 0)$
si $\lambda'_s < \beta_s$ **alors**
 $\exists r$ tel que λ_r atteint une borne
 $J = J U(s)/(r)$, $\lambda_s = \lambda'_s$ et calculer λ^j , aller en 2-
sinon ($\lambda_s = \beta_s$), J reste fixe et calculer λ^j , aller en 3-
 7- **fin.**

3- Interprétation géométrique des itérations

a- Espace des ségrégations

L'espace des ségrégations est défini par l'ensemble des λ (de dimension p). Les variables de (P.L) sont dans $S \times E$. Le domaine des contraintes est un polyèdre borné.

$$Q = \{ (\lambda, v) \in [0, 1]^p \times [0, M]^n / A\lambda + v = M \}$$

On montre que les solutions de base définies plus haut sont des sommets de Q . Chaque étape ou itération décrite ci-dessus se traduit par un passage d'un sommet de Q , à un autre, adjacent au premier.

Si on définit Q^* par :

$$Q^* = \{ \lambda \in S / \exists v \in E \text{ et } (\lambda, v) \in Q \}$$

comme projection de Q dans S parallèlement à E .

On peut énoncer la propriété suivante:

A toute solution de base de (P.L) correspond un sommet de Q^* . (qui n'est autre que sa projection dans S).

En effet, soit (λ, v) une solution de base de Q . Montrons que λ est un sommet de Q^* .

Supposons que λ ne soit pas un extrémum de Q^* , alors il existe λ^1 et λ^2 deux éléments de Q^* tel que:

$$\lambda = \theta \lambda^1 + (1-\theta) \lambda^2, \quad 0 < \theta < 1$$

Il existe v_1 et v_2 tel que:

$$A \lambda^1 + v_1 = M \quad \text{et} \quad A \lambda^2 + v_2 = M$$

On en déduit que $v = \theta v_1 + (1-\theta) v_2$

d'où $(\lambda, v) = \theta (\lambda^1, v_1) + (1-\theta) (\lambda^2, v_2)$

C.Q.F.D

Cette propriété nous permet de conserver l'interprétation ci-dessus dans l'espace des ségrégations.

Les λ de la forme $\lambda_{1_1}, \dots, \lambda_{1_n} \in [0,1]$; $\lambda_{1_k} \in (0,1)$, $k=1, \dots, n$ sont les sommets de Q^* .

b- Espace des quantités

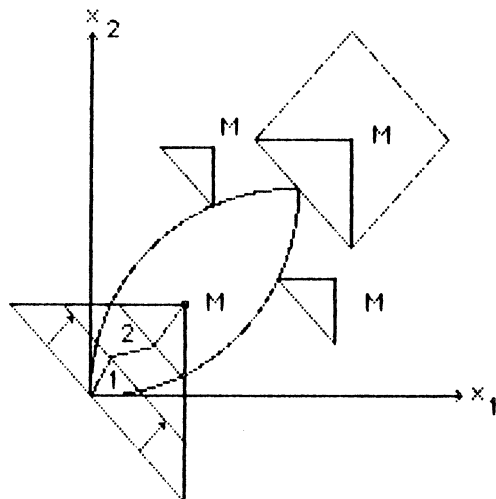
L'interprétation précédente des itérations du simplexe est horizontale et explique l'évolution des λ dans S .

Cependant nous cherchons à trouver un sens physique à cette évolution et surtout au niveau des opérations de mélange et de séparation des coupes N_i , $i=1, \dots, p$.

Cette analyse est verticale et considère les colonnes de A et leur combinaison.

Le processus évolue en minimisant la quantité molaire totale du résidu v restant à fabriquer à partir du mélange courant $A\lambda$.

Ceci se traduit dans E par une évolution (rétrécissement) de la boule de centre M et de rayon $\|v\|$ que l'on peut représenter dans \mathbb{R}^2 par:



Le mélange $A\lambda$ fabriqué au cours de chaque itération est sur le bord de cette boule. Si M est dans C cette boule se réduit à la fin des itérations au point M .

Nous distinguons trois types de coupes dans chaque itération du simplexe:

- 1- les mélanges qui sont consommés entièrement pour fabriquer M .
- 2- les mélanges qui ne sont pas consommés.
- 3- les n mélanges restants qui sont consommés partiellement ou totalement (éléments de la base J).

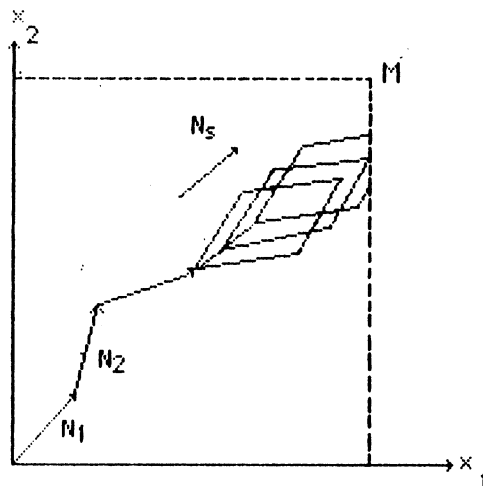
Le processus consiste à choisir un mélange parmi (1) et (2) (c'est à dire une direction N_k dans E) qui puisse améliorer le plus possible, le mélange $A\lambda$ (sa quantité molaire croît) en laissant fixe les mélanges hors base et en remélangeant les éléments de la base.

En fait, le problème se ramène dans chaque étape au remélange de $(n+1)$ vecteurs quantités à partir desquels on fabrique le mélange résidu

$$(M - A^J \lambda^J + A^S \lambda^S)$$

Cette opération ne nécessite pas d'appeler le simplexe pour $(n+1)$ mélanges mais de garder la décomposition courante et de mélanger ou de séparer l'élément N_S .

Géométriquement, cela consiste à translater le convexe de mélange correspondant aux n éléments de la base sur la direction (N_S) tout en restant dans l'hypercube $[0, M]^n$. La figure suivante montre dans \mathbb{R}^2 la translation du parallélogramme associé aux vecteurs de la base sur le vecteur N_S .



4- Méthode duale

La méthode linéaire primale engendre une suite de solutions réalisables, en assurant une décroissance monotone de la fonction à minimiser. En revanche, la méthode duale est plus facile à mettre en oeuvre et fournit de bonnes interprétations sur la propriété de l'appartenance de M à C .

Considérons le Lagrangien classique correspondant au problème (P.L). On dualise uniquement les contraintes $A\lambda - M \leq 0$.

$$L(\lambda, \mu) = \begin{cases} \langle M - A\lambda, 1 \rangle + \langle \mu, A\lambda - M \rangle & \text{si } \mu \in \mathbb{R}_+^n, \lambda \in [0, 1]^p \\ +\infty & \text{si } \lambda \notin [0, 1]^p \\ -\infty & \text{si } \mu \notin \mathbb{R}_+^n, \lambda \in [0, 1]^p \end{cases}$$

Le problème (P.L) s'écrit alors

$$(P.P) \quad a = \inf_{\lambda \in [0, 1]^p, \mu \in \mathbb{R}_+^n} (\sup (L(\lambda, \mu)))$$

Rappelons d'abord que M est positif ($M \geq 0$), sinon le domaine des contraintes de (P.P) se réduit à l'ensemble vide et M n'appartiendrait pas à C . Il en résulte que l'ensemble des solutions de (P.L) est non vide et que a est fini. Sans perte de généralité, on suppose que les composantes de M sont strictement positives. En effet si la $j^{\text{ème}}$ composante de M est nulle, il suffit de supprimer tous les vecteurs N_j de la famille initiale dont la $j^{\text{ème}}$ composante est nulle et d'en éliminer cette composante.

Sous cette hypothèse, il existe alors un λ de $[0, 1]^p$ tel que $A\lambda < M$. En conséquence et d'après [10], les conditions de stabilité de (P.P) sont vérifiées.

Notons (P.D) le problème dual de (P.P).

$$(P.D) \quad b = \sup_{\mu \in \mathbb{R}_+^n, \lambda \in [0, 1]^p} (\inf (L(\lambda, \mu)))$$

En vertu des hypothèses précédentes et d'après les conditions de stabilité de (P.P) on peut affirmer que l'ensemble des solutions de (P.D) est non vide et que $b = a$.

Notons w la fonctionnelle définie sur \mathbb{R}_+^n par:

$$w(\mu) = \inf_{\lambda \in [0, 1]^p} (L(\lambda, \mu))$$

et rappelons aussi que sous les conditions de stabilité, (λ^*, μ^*) est un point de selle si et seulement si λ^* est une solution de (P.P), μ^* est une solution de (P.D) et $a = b$.

D'où $a = \langle M - A\lambda^*, 1 \rangle = b = w(\mu^*)$

On en déduit une caractérisation de l'appartenance de M à C :

$$(1) \quad (M \in C) \Leftrightarrow (w(\mu^*) = 0)$$

$$\text{et } (2) \quad (M \notin C) \Leftrightarrow (\exists \mu \in \mathbb{R}_+^n / w(\mu) > 0)$$

La première proposition est évidente.

Démonstration de la deuxième proposition:

si il existe $\mu \in \mathbb{R}_+^n$ tel que $w(\mu) > 0$ alors $w(\mu^*) > 0$ et d'après (1) $M \notin C$.
si $M \notin C$ alors $a > 0$ et comme $a = w(\mu^*)$ on a $w(\mu) > 0$ pour $\mu = \mu^*$.

Interprétation géométrique

La proposition (2) se traduit géométriquement par l'existence d'un hyperplan séparant le convexe C et le point M , ce qui rejoint la caractérisation géométrique de la faisabilité de M évoquée au début de ce chapitre. En effet, si on pose pour un μ appartenant à \mathbb{R}_+^n

$$u = 1 - \mu, \alpha = \langle u, M \rangle - w(\mu) \text{ et } X = A\lambda$$

on aura d'après la définition de w , $\langle u, X \rangle \leq \alpha \quad \forall X \in C$.

Et si μ vérifie la proposition (2) alors $\langle u, M \rangle > \alpha$.

L'hyperplan $(X, \langle X, u \rangle \leq \alpha)$ de E sépare le point M et le convexe C .

4) Application numérique et synthèse

Dans cette section, nous allons comparer les quatre méthodes de faisabilité décrites dans ce chapitre, à partir d'un modèle de la séparation en distillation. On simule les quantités des différentes espèces séparées dans une colonne à m étages, fonctionnant à reflux total, par le modèle suivant:

$$F(t) = (A^m e^{-A^m t}) \vec{N}_0$$

N_0 est le vecteur charge injecté dans la colonne, dont chaque composante désigne la quantité totale d'espèce associée. m est le nombre d'étages de la colonne, t le paramètre temps, et A une matrice carrée de volatilités relatives aux espèces. $(F(t)dt)$ est la famille différentielle des vecteurs quantités récupérés à la sortie de la colonne.

Nous supposons que le modèle est idéal, c'est à dire que la matrice A est diagonale. Le vecteur quantité associé à l'intervalle de temps $[t_1, t_2]$ est déterminé directement par

$$(e^{-A^m t_1} - e^{-A^m t_2}) \vec{N}_0$$

Nous fixons les paramètres suivants:

$n = 3, m = 3, N_0 = (1.0, 1.0, 1.0)$ et

$$A = \begin{bmatrix} 2.49 & .0 & .0 \\ .0 & 1.0 & .0 \\ .0 & .0 & .364 \end{bmatrix}$$

Nous avons déterminé la décomposition des mélanges $M_1 = (.1, .2, .5)$ et $M_2 = (.5, .4, .7)$ par les quatre méthodes de fabrication. Le tableau suivant montre, selon le nombre p des vecteurs quantités, le temps d'exécution mis par chaque méthode sur **HP 1000.A**.

p	100	200	300	400	500
A	0:01:05	0:03:36	0:08:95	0:11:77	0:14:61
	0:02:51	0:05:44	0:14:65	0:19:98	0:24:59
B	0:00:61	0:01:00	0:01:30	0:01:68	0:02:06
	0:00:75	0:01:01	0:01:37	0:02:73	0:02:10
C	0:13:55	0:13:95	0:37:80	0:52:84	1:11:57
	0:30:60	1:26:75	1:55:92	5:13:59	5:18:00
D	1:26:34	1:47:73	1:46:29	2:30:74	5:42:88
	1:30:59	2:29:14	2:01:51	3:16:26	4:07:01

- A: méthode linéaire du simplexe.
 B: méthode de décomposition en série.
 C: méthode de décomposition en parallèle.
 D: méthode des projections successives

a:b:c signifie a minutes, b secondes et c centièmes. Il désigne le temps d'exécution. Dans chaque méthode, la première ligne désigne le temps de décomposition de M_1 , la seconde celui de M_2 .

On remarque que les méthodes "A et B" sont plus rapides que les deux autres. En général, nous avons constaté à partir des exemples numériques traités que:

- 1- la méthode du simplexe est la plus rapide.
- 2- le temps d'exécution de chaque méthode dépend essentiellement de l'erreur absolue ou relative fixée a priori et du mélange à fabriquer.

Nous remarquons que pour les mélanges M_1 et M_2 , la méthode B est plus rapide que A, ce qui n'est pas le cas en général.

On retient deux remarques importantes: La méthode du simplexe fournit une solution exacte en un nombre fini d'itération et les autres méthodes sont lentes, par contre elles fournissent une bonne approximation de la décomposition en peu d'itérations.



Chapitre 3

Gestion des mélanges binaires

- 1- Introduction
- 2- Bord du convexe des mélanges binaires
- 3- Droite d'appui du convexe des mélanges binaires
- 4- Faisabilité d'un mélange objectif
 - 4.1- Possibilité de fabrication
 - 4.2- Fabrication optimale d'un mélange objectif
 - 4.3- Fabrication des mélanges ayant une contrainte linéaire
- 5- Applications
 - 5.1- Applications numériques
 - 5.2- Problème de recyclage en chromatographie



1) Introduction

Dans ce chapitre, nous nous intéressons uniquement aux mélanges à deux espèces. Nous traitons les différentes opérations citées dans le chapitre 1, en exploitant les propriétés et les simplifications introduites par cette spécificité. Cette dernière conduit à une gestion simple des mélanges binaires et donne naissance à un progiciel interactif et visuel recouvrant toutes ces opérations élémentaires, nécessaires à la gestion des mélanges (domaine complètement vierge actuellement).

Le convexe des mélanges binaires est un cas particulier des polygones. En conséquence, toutes les techniques adoptées, dans ce chapitre, sont inspirées des procédés de **M.L.SHAMOS & F.PREPARATA** [18]. Nous avons introduit certaines améliorations dues à la spécificité de chaque problème. Tous les algorithmes proposés procèdent par la méthode de la recherche binaire, ce qui les qualifie d'être rapides.

Dans la deuxième section, nous aborderons le problème de la visualisation du bord du convexe par ses sommets. Nous présenterons dans la section 3 un algorithme permettant de déterminer à partir d'une droite donnée ses deux droites d'appui du convexe. Dans la section 4, nous traiterons le problème de la faisabilité des mélanges. On y distingue trois parties: la possibilité de fabrication d'un mélange objectif, la fabrication optimale d'un mélange et la recherche des mélanges ayant une contrainte linéaire.

Nous terminerons dans la section 5 par deux applications très importantes du point de vue pratique. La première concerne la fabrication optimale de deux produits par le recyclage; la deuxième consiste à chercher les points de coupe optimale d'un système donné.

2) Bord du convexe des mélanges binaires.

Soit p vecteurs quantités (N_i) , $i=1, \dots, p$ dans E à deux dimensions. Le convexe des mélanges binaires C se réduit à un polygone à symétrie centrale (voir chapitre 1). Notre objectif est de déterminer tous les sommets ou bien tous les segments composant le bord. Cela permet d'une part à l'utilisateur de visualiser le convexe sur un écran, donc d'évaluer la qualité de la séparation et, d'autre part, de faciliter la résolution des différents problèmes traités ci-dessous.

Ce problème sera posé lorsqu'une famille de vecteurs quantités n'est pas de sélectivité régulière, c'est à dire, lorsque la courbe formée par leur juxtaposition, sortie d'un séparateur, dans leur ordre initial, n'est pas convexe. On l'appelle "**courbe initiale**".

Le bord sera alors la "convexifiée" de la courbe initiale, ou bien la juxtaposition des vecteurs quantités dans l'ordre de sélectivité régulière, lequel sera donné par l'ordre des fractions de ségrégation d'une espèce par rapport à l'autre. En effet, on voit apparaître une liste de p points ordonnés sur le simplexe (segment d'extrémité $(1,0)$ et $(0,1)$) lorsqu'on normalise les vecteurs N_i par la somme des deux composantes. On détermine cet ordre en parcourant cette liste d'une extrémité à l'autre.

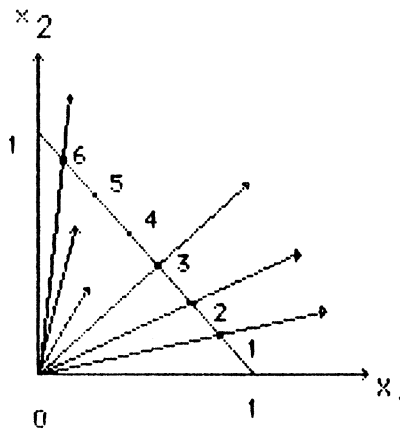


figure 3.1

théorème.

Soit p vecteurs quantités (N_i) de E . On peut déterminer le bord du convexe de mélange (la liste des sommets ou bien des segments du bord) en $O(p \log(p))$ opérations.

Démonstration:

Le problème se ramène à trier p constantes (les fractions de

ségrégation ou bien les arguments des N_i). Des algorithmes comme la recherche binaire procèdent en $O(p \log(p))$. Après avoir ordonné les vecteurs (N_i) , les sommets s'obtiennent par $S_1 = 0$ et $S_{i+1} = S_i + N_i$, $i = 1, \dots, p$

$$S_{p+i+1} = S_{p+i} - N_{p+1-i}, \quad i = 1, \dots, p-1$$

Nous avons $2p$ sommets et $2p$ segments définissant le bord du convexe.

Des exemples numériques sont présentés à la fin de ce chapitre.

Pour tout ce qui suit on supposera que les vecteurs quantités sont ordonnés, et qu'il sont deux à deux linéairement indépendants. Dans le cas général, on pourra toujours se ramener à cette situation.

3) Droite d'appui du convexe des mélanges binaires

Soit p vecteurs quantités N_i ordonnés de E , C le convexe associé et d une direction donnée de E . Il existe deux droites d'appui de C parallèles à d . On souhaite dans ce paragraphe déterminer leur équations et les deux sommets d'appui associés. Notons d^\perp un vecteur orthogonal à d et $\langle x, d^\perp \rangle = c$ l'équation d'une droite parallèle à d (c étant un scalaire de \mathbb{R}).

Les deux droites d'appui sont déterminées par :

$$\max_{X \in C} (\langle X, d^\perp \rangle) \quad \text{et} \quad \min_{X \in C} (\langle X, d^\perp \rangle)$$

La résolution de ces deux problèmes est explicite, nous obtenons

$$c_{\max} = \sum_{i=1}^p (\langle N_i, d^\perp \rangle)_+ \quad \text{et} \quad c_{\min} = \sum_{i=1}^p (\langle N_i, d^\perp \rangle)_-$$

Les points d'appuis associés que l'on note S_{\max} et S_{\min} sont:

$$S_{\max} = \sum_{i=1}^p \lambda_i N_i \quad \text{tel que} \quad \lambda_i = 1 \quad \text{si} \quad \langle N_i, d^\perp \rangle > 0, \quad 0 \quad \text{sinon}$$

S_{\min} est le symétrique de S_{\max} par rapport au centre de symétrie de C . Cette méthode est linéaire en $O(p)$ et ne dépend pas de l'ordre des vecteurs N_i .

Par contre s'ils sont ordonnés, il suffit de localiser la position du vecteur d par rapport aux N_i . Ceci se fait en procédant par la recherche binaire sur les arguments des N_i . Notons j le rang tel que l'argument de d est compris

entre ceux de N_j et N_{j+1} . On en déduit facilement les points S_{\max} et S_{\min} et par conséquent les scalaires c_{\max} et c_{\min} .

$$S_{\max} = \sum_{i=1}^J N_i, \quad c_{\max} = \langle S_{\max}, d^{\perp} \rangle$$

Ainsi la détermination des droites d'appuis se fait en $O(\log(p)) + 2J$ opérations.

4) Faisabilité d'un mélange objectif

4.1) Possibilité de fabrication

Le problème de la faisabilité d'un mélange objectif est traité d'une façon générale dans le chapitre 2. La particularité des mélanges binaires nous permet, d'une part, de fabriquer d'une manière optimale ce mélange et, d'autre part, d'appliquer un algorithme rapide M.L.Shamos[18] pour résoudre le problème d'inclusion.

Soit M un mélange objectif et (N_i) une famille de p vecteurs ordonnés.

M.L.Shamos[18] propose une méthode en $O(\log(m))$ pour résoudre le problème d'inclusion d'un point dans un polygone à m sommets.

Le procédé consiste à:

- 1-déterminer un point intérieur au polygone.
- 2-faire passer de ce point à chaque sommet une droite.
(ce qui subdivise le polygone en m triangles).
- 3-utiliser ce point comme origine et localiser par une recherche binaire, en se servant des coordonnées polaires, le triangle contenant M .

Dans notre cas, on considère le centre de symétrie comme origine. Nous avons $2p$ sommets, donc la question d'inclusion est résolue en $O(\log(p))$.

4.2) Fabrication optimale d'un mélange objectif

A partir d'une famille donnée, nous cherchons à constituer un mélange M d'une manière "optimale". Autrement dit, et sans contraintes économiques, nous souhaitons fabriquer M de façon qu'on puisse constituer, à partir du reste de la famille, le plus de mélanges possible.

propriétés:

1- A chaque décomposition de M correspond un convexe résidu.

En effet soit $\lambda \in [0,1]^p$, $\sum_{i=1}^p \lambda_i N_i = M$ alors la famille des vecteurs

$((1-\lambda_i)N_i, i=1, \dots, p)$ constitue les vecteurs quantités résidu et l'ensemble de leur mélange constitue le convexe résidu C_{R_λ} .

La fabrication optimale se traduit par la recherche de λ tel que le volume de C_{R_λ} soit maximal.

2- Soit M un mélange donné dans C et soit C_{R_λ} le convexe résidu de M correspondant à la décomposition λ on a:

$$C_{R_\lambda} \subset (C - M)$$

3- Dans le cas binaire: $\exists \lambda / C_{R_\lambda} = (C - M) \cap C$.

En effet, cela vient du fait que d'une part, $(C - M) \cap C$ est un convexe de mélange et d'autre part, il contient tous les convexes résidus associés à M. La décomposition est telle que M soit une corde de la famille initiale. Il sera formé par les vecteurs ou morceaux de vecteurs intérieurs à l'arc de la corde M.

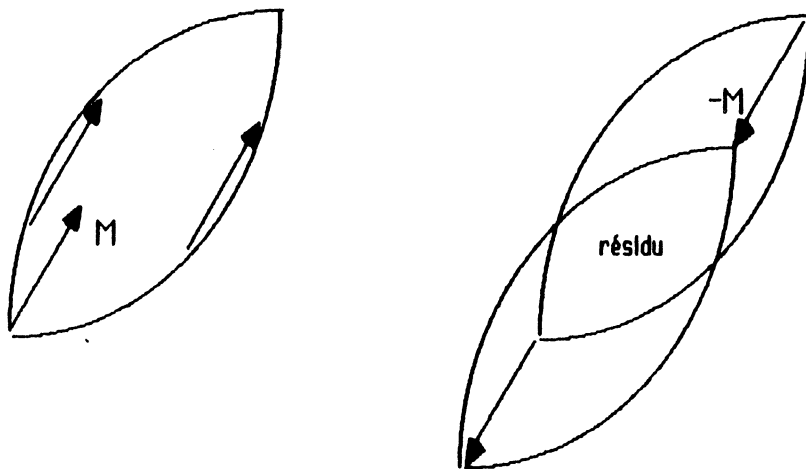


figure 3.2

Il s'avère que, dans le cas des mélanges binaires, la fabrication optimale d'un objectif est caractérisée par la troncature, ou bien encore, par la détermination du convexe résidu, lequel est donné par l'intersection de C

et $C-M$. Ce qui n'est pas le cas (sauf exception) dans un espace de dimension supérieure à 2.

On relève plusieurs algorithmes dans M.L.Shamos[18], dus à Shamos Hoey(1976) et O'Rourke-chien-olson-Naddor(1982), permettant de déterminer l'intersection de deux polygones quelconques. Ces méthodes sont linéaires et nécessitent un passage en revue de toutes les facettes de chacun des deux polygones. Par contre, dans notre cas : intersection d'un convexe et de son translaté, nous proposons une méthode se basant sur la recherche binaire en exploitant la structure particulière de C .

Notons la liste des sommets inférieurs de C par L_1 constituée par $p+1$ points et L_2 le translaté de L_1 par M . La troncature revient à déterminer l'intersection de ces deux lignes brisées.

Nous remarquons que si $M \in C$ et ne constitue pas une fraction d'un vecteur quantité quelconque, alors il existe un seul point d'intersection. (On exclut le cas où M est nul).

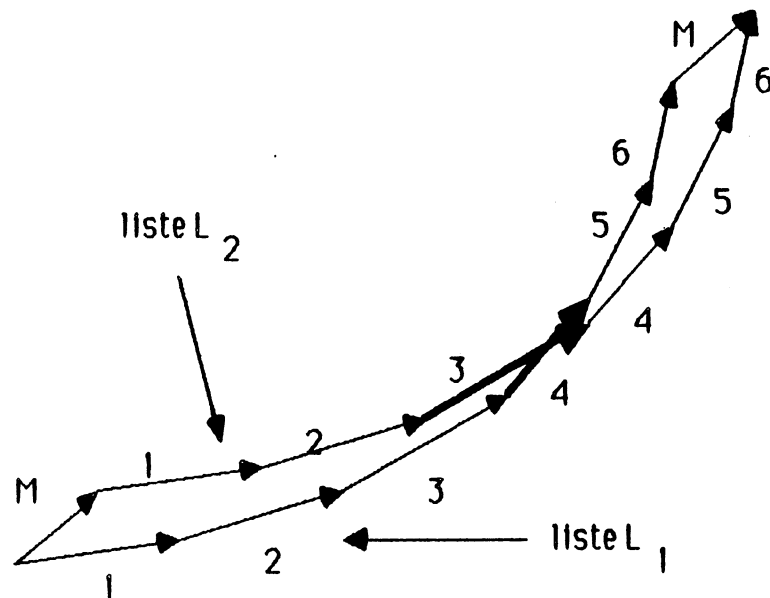


figure 3.3

Le principe de l'algorithme se base sur la remarque suivante: La liste L_2 peut être divisée en deux sous-listes ordonnées telles que les sommets de la première sous-liste soient dans C et ceux de la deuxième à l'extérieur de C . Le segment raccordant ces deux sous-listes coupe la liste L_1 .

Ainsi, par une recherche binaire, et en conservant cette propriété en chaque itération, on localise le segment de raccordement en $\log(p)$. Dans chaque itération, on a besoin de tester si un sommet extrémité est dans C ou non. Cette recherche, d'après le paragraphe 4.1, se fait en $\log(p)$. La localisation du segment de L_2 qui coupe L_1 se fait alors en $O(\log(p)^2)$. On applique le même processus en renversant le rôle de L_1 et L_2 , ce qui nous permet d'avoir le segment de L_1 qui coupe la liste L_2 .

A partir de ces deux informations (les deux segments précédents), on peut déterminer facilement le point d'intersection, la composition du mélange M et le convexe résidu. En effet, à chaque segment correspond un rang dans la liste et son vecteur quantité équipollent. Soit $(S_i, S_j) \in L_1 \times L_2$ tel que S_i coupe L_2 et S_j coupe L_1 . (i (respectivement j) indique le rang du segment S_i dans L_1 (respectivement dans L_2):

- Le point d'intersection est égal à $S_i \cap S_j$.
- Le mélange M est constitué par les mélanges:
 $\lambda_i N_i, N_{i+1}, \dots, N_{j-1}, \lambda_j N_j$. (λ_i, λ_j indiquent les fractions de N_i, N_j)
- La famille des vecteurs non consommés est :
 $\{N_1, N_2, \dots, N_{i-1}, (1 - \lambda_i)N_i, (1 - \lambda_j)N_j, N_{j+1}, \dots, N_p\}$

Application

Cette technique de fabrication fournit un moyen efficace pour montrer la coexistence de plusieurs mélanges dans C . Donc elle permet leur fabrication d'une manière indépendante. Elle permet ainsi de tester l'inclusion d'un convexe objectif dans C .

Nous signalons que l'ordre de fabrication n'intervient pas, et que, si à une étape, on n'arrive pas à fabriquer un mélange de la famille objectif on déduit directement que le convexe des mélanges objectif n'est pas inclus dans C . En effet, il suffit de fabriquer d'une manière successive et dans un ordre quelconque les mélanges objectifs. Dans chaque étape on utilise la famille résultante de la fabrication précédente.

4.3) Fabrication des mélanges ayant une contrainte linéaire

La fabrication d'un mélange ou de tous les mélanges ayant une contrainte linéaire se traduit d'une façon générale par la recherche de l'intersection de C et d'une droite. Nous citons plusieurs applications

physiques très importantes:

* On peut s'intéresser aux mélanges contenant une proportion donnée des deux espèces, ou une quantité donnée d'une seule espèce.

* On s'intéresse aussi à déterminer la quantité maximale d'un mélange de composition donnée. Il est équivalent à la recherche de l'intersection d'une droite passant par l'origine et le bord de C.

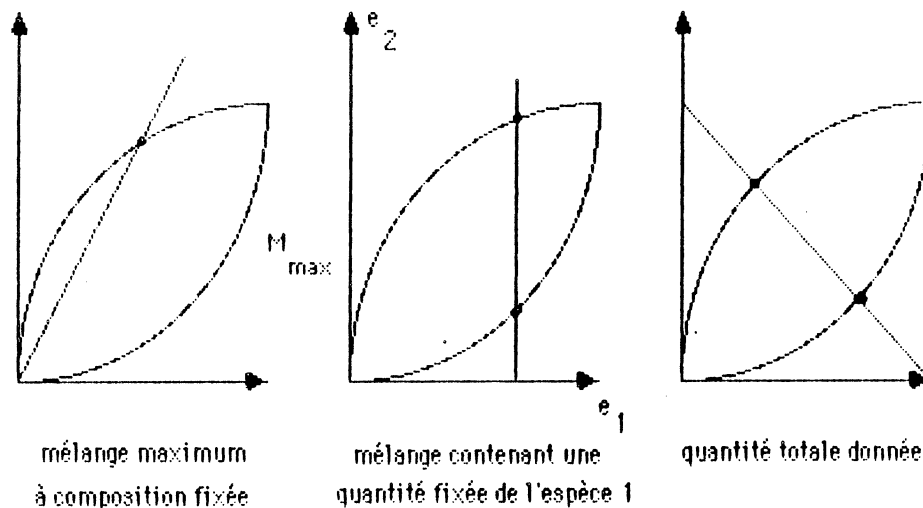


figure 3.4

Le principe de l'algorithme est très simple, il se décompose en deux étapes essentielles. Soit D une droite quelconque, et C le convexe des mélanges:

1-détermination des deux droites d'appuis de C parallèles à D .

Cette étape permet, d'une part, de tester si $D \cap C \neq \emptyset$ et, d'autre part, de diviser la liste des sommets du convexe C en deux sous-listes telle que chacune d'elle coupe la droite une seule fois (si $D \cap C \neq \emptyset$).

2-traitement de chaque liste par une recherche binaire afin de localiser le segment qui coupe D .

La recherche de l'intersection d'une droite se fait en $O(\log(p))$.

5) Applications

1- Applications numériques

Nous présentons dans ce paragraphe l'application des différentes opérations: la visualisation du bord, l'intersection et la troncature, sur le convexe C des mélanges. Les vecteurs quantités sont simulés par une primitive de la fonction $f(t) = |\cos(t + \alpha_i)|$, $i=1,2$ où α_i est un paramètre associé à la composante i . Physiquement, on peut dire que cette simulation correspond à un procédé en régime transitoire, décrivant des débits variant d'une manière périodique au cours du temps. Le vecteur quantité associé à l'intervalle $[t_1, t_2]$ sera

$$\text{déterminé par } \left(\int_{t_1}^{t_2} |\cos(t + \alpha_i)| dt \right), i=1,2.$$

où t_1 et α_1 sont donnés en radian.

-La figure 3.5 représente la courbe initiale des vecteurs quantités par la ligne brisée. Le bord du convexe des mélanges est représenté par la courbe continue. Dans (a) on a fixé les données à: $\alpha_1=0$, $\alpha_2=1.07$, $I = [0, \pi]$ et $p=50$.

Dans (b) on a pris $\alpha_2 = .7$. Dans (c) on a $I = [0, 50]$, $\alpha_1=0$, $\alpha_2 = 20$ et $p = 100$.

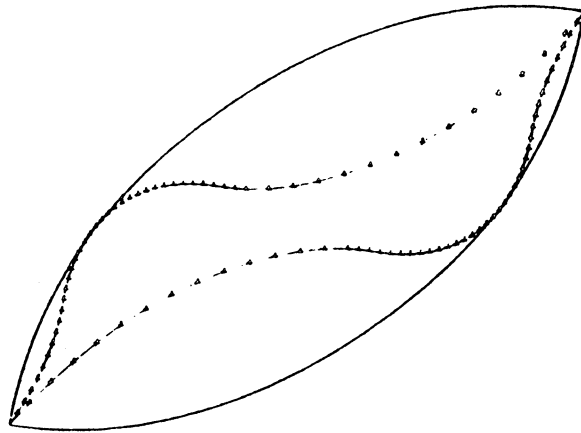
-La figure 3.6 montre l'intersection de quatre droites de directions différentes et le convexe C. Les coordonnées des points du bord sont déterminés par l'algorithme de la paragraphe 4.3 de ce chapitre:

$$A = (1.233, 6.5617), \quad B = (19.5430, 15.7710), \quad C = (15.77, 19.543)$$

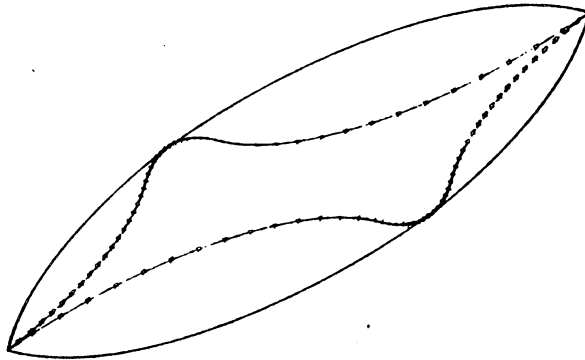
$$A' = (8.0, 1.6763), \quad B' = (8.0, 15.969)$$

$$A'' = (2.6887, 10.0), \quad B'' = (17.31, 10.0)$$

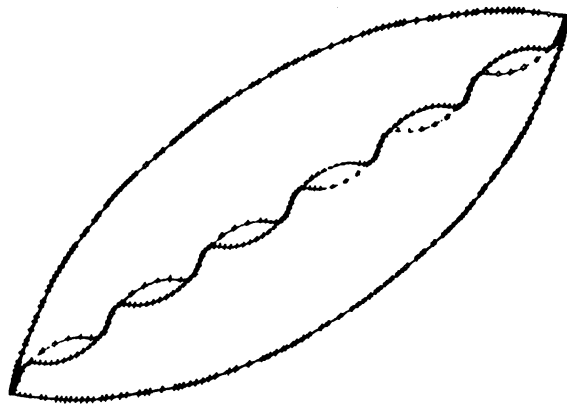
-La figure 3.7 montre la troncature du convexe C respectivement par les mélanges M_1 et M_2 , et leur convexe résidu après chaque fabrication.



(a)



(b)



(c)

figure 3.5 Visualisation du bord du convexe des mélanges après un réordonnement de la famille des vecteurs quantités. + désigne le point de juxtaposition de deux vecteurs quantités

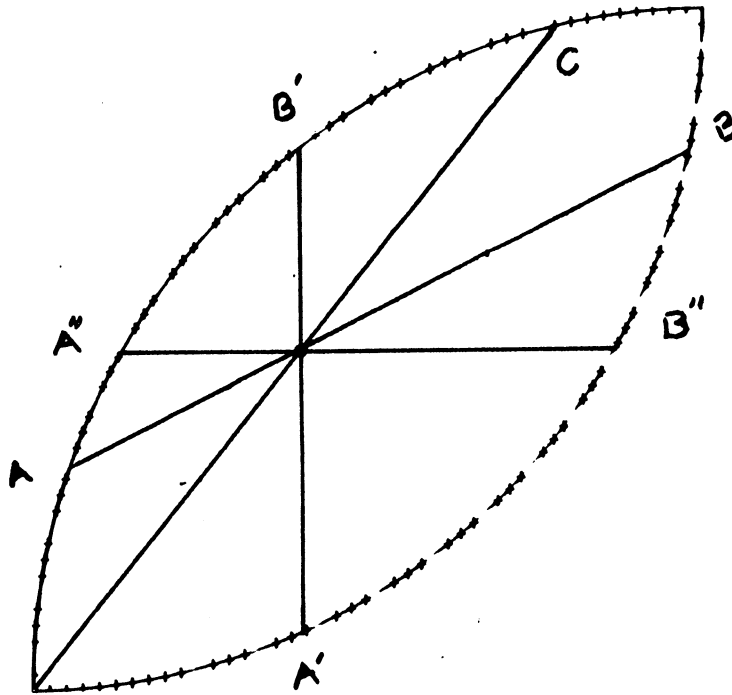


figure 3.6 Faisabilité des mélanges soumis à des contraintes linéaires

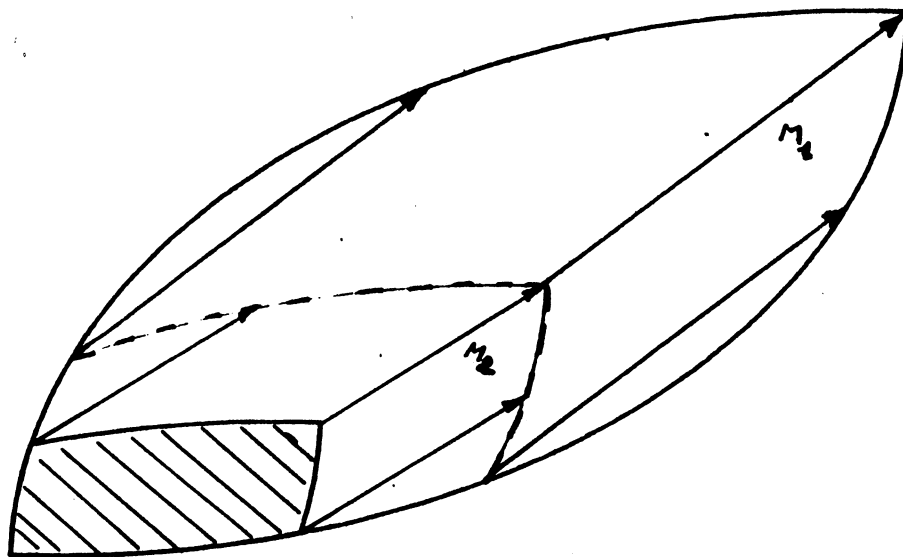


figure 3.7 Fabrication optimale des mélanges M_1 et M_2 .
La zone hachurée indique le convexe résidu.

2- Problème de recyclage en chromatographie

Dans ce paragraphe, nous allons étudier deux problèmes très importants en chromatographie préparative (par exemple en biotechnologie): la production optimale de deux produits soumis à des spécifications données et la détermination des points de coupe optimale. L'objet essentiel de ces deux exemples est de montrer l'utilité et l'efficacité des opérations, décrites dans ce chapitre, dans la résolution des problèmes plus compliqués de la gestion des mélanges binaires.

En chromatographie, le procédé de séparation est souvent cyclique et suppose en entrée un vecteur charge N : "feed" qui alimente, d'une façon périodique en fonction du temps, la colonne de séparation. A la sortie de cette dernière, on récupère les deux produits N_1 , N_2 et le résidu R qui sera recyclé c'est à dire mélangé avec la charge N . Nous allons voir qu'il y aura plusieurs critères du choix de R dans le recyclage.

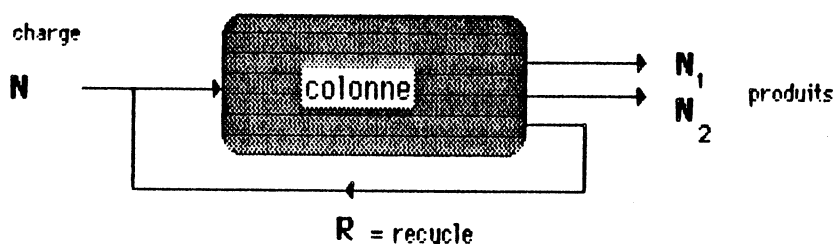


figure: processus du recyclage

A la sortie de la colonne, un détecteur enregistre les chromatogrammes associés aux espèces analysées, à partir desquels on calcule la famille des vecteurs quantités. On suppose dans toute la suite que cette famille est différentielle et que le convexe C des mélanges associé est lisse.

Production optimale des produits N_1 et N_2

On s'intéresse dans cette section à produire au maximum les deux produits N_1 et N_2 , soumis à des spécifications données.

Pour que cette opération soit possible, il est nécessaire que l'intersection de C et des droites de directions N_1 et N_2 soit non vide. En outre, on suppose que la diagonale principale de C est dans le cône engendré par ces deux droites.

On distingue deux façons différentes de recyclage: la première consiste à mélanger en tête de la colonne avec la charge N le recycle R qui est

parallèle à ce dernier, la deuxième consiste à recycler un vecteur non parallèle à la charge net. Dans ce cas on maintient la quantité totale injectée par cycle au cours du temps.

Recyclage à stabilité forcée

On suppose que la charge est fixe. Cette technique nécessite un seul cycle pour déterminer la production optimale des produits 1 et 2. Notons que le parallélogramme formé par \mathbf{N}_1 et \mathbf{N}_2 n'est pas inclus dans \mathbf{C} et qu'on peut fabriquer seulement des fractions de \mathbf{N}_1 et de \mathbf{N}_2 (sinon la production totale sera satisfaite et le recyclage ne sera pas utile).

On souhaite déterminer λ et μ tel que le volume du parallélogramme $(\lambda\mathbf{N}_1, \mu\mathbf{N}_2)$ soit maximal et que le recycle soit colinéaire à la charge.

L'algorithme de résolution se résume en trois étapes:

- 1- Intersection de \mathbf{C} et de \mathbf{N}_1 ce qui donne λ .
- 2- Intersection de \mathbf{C} et de \mathbf{N}_2 ce qui donne μ .
- 3- correction de λ ou bien de μ tel que $\lambda\mathbf{N}_1 + \mu\mathbf{N}_2$ soit colinéaire à \mathbf{N} .

La figure suivante montre les différentes constructions: $\lambda\mathbf{N}_1$, $\mu\mathbf{N}_2$ et \mathbf{R} .

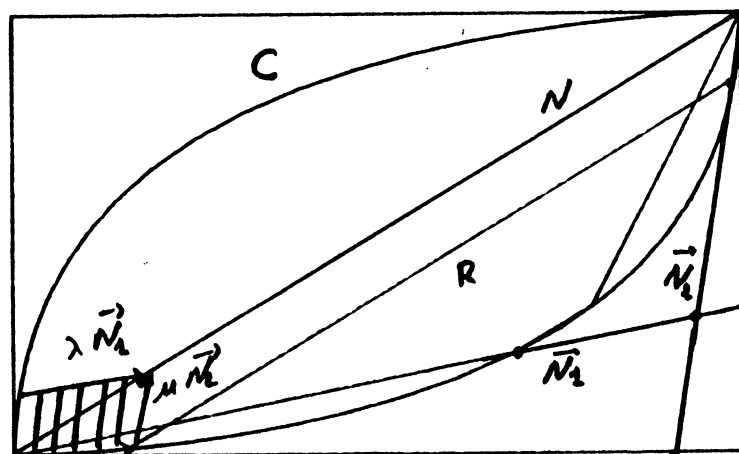


figure 3.8 Recyclage stable et fabrication maximale des produits \mathbf{N}_1 et \mathbf{N}_2 .

Recyclage évolué

Cette technique consiste à laisser évoluer la charge jusqu'à ce que le cycle et le convexe des mélanges associé soient stables.

Notons R_{k-1} le recycle du au $k-1^{\text{ème}}$ cycle, la charge N_k suivante est obtenue par la relation suivante:

$$N_k = \lambda N + R_{k-1} \quad \text{et} \quad \|N_k\|_1 = \|N\|_1$$

Ceci consiste à projeter le recycle R_{k-1} sur le simplexe:

$$\{(x,y) / x+y = \|N\|_1, x \geq 0 \text{ et } y \geq 0\}$$

On injecte de nouveau N_k et on refait les mêmes opérations. En pratique la stabilité est atteinte en 4 itérations.

La figure suivante montre les cycles intermédiaires et les parallélogrammes de production associés à N_1 et N_2 .

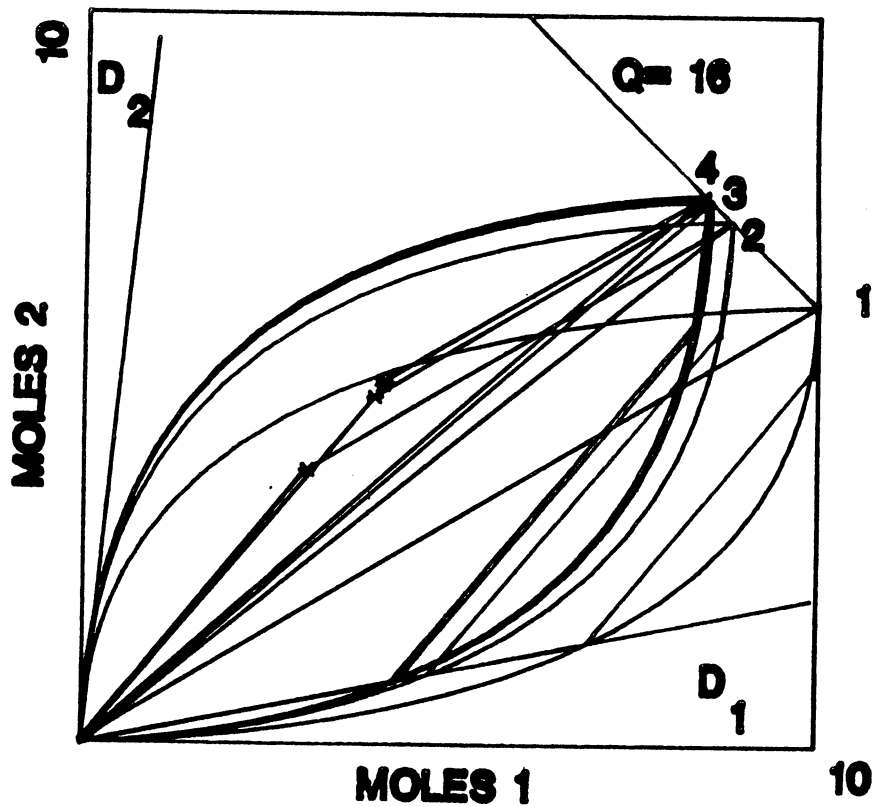


figure 3.9 Evolution des cycles

Cette technique, plus complexe, a l'avantage de mieux utiliser la séparation par la colonne.

Le tableau suivant montre les calculs numériques des volumes et des produits N_1 et N_2 dans les deux cas de recyclage.

	recyclage forcé			recyclage évolué		
	N	N ₁	N ₂	N	N ₁	N ₂
1	1.658	1.568	.090	4.330	4.110	0.180
2	0.995	0.093	.902	2.620	0.820	1.800
1+2	2.653	1.661	.992	6.950	4.930	1.980
volume	1.406			7.25		

On remarque que la deuxième technique est nettement meilleure que la première, on obtient 4 fois plus de produits.

Points de coupe optimale

Il consiste à déterminer q points de coupe au cours du temps (on note t_1, t_2, \dots, t_q les temps associés) de telle façon que la production soit optimale. Ceci se traduit par la recherche de $q-1$ points sur le bord de C tel que le volume du polygone correspondant soit maximal. Dans [20] on montre qu'il existe une et une seule séquence de $q-1$ points qui maximise le volume résultant de la séparation.

On se contente, dans la suite, de présenter le cas où $q=2$ pour avoir une idée sur la caractérisation de ces points. On montre facilement dans [20] que, dans ce cas, le point de coupe correspond au point d'appui de C dont la tangente est parallèle à la diagonale principale de celui-ci. La figure suivante représente la position de ce point.

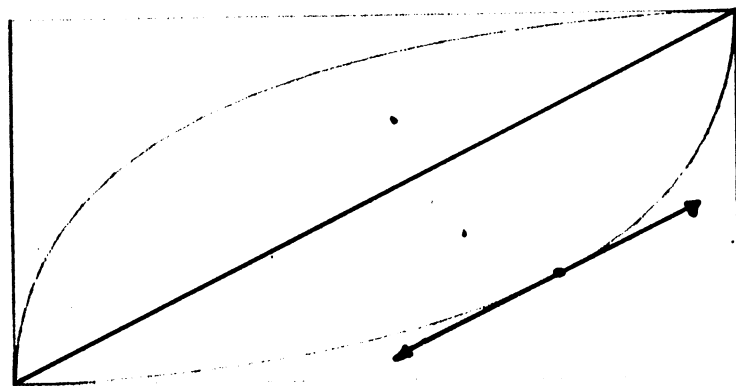


figure 3.10 2 points de coupe optimale



BIBLIOGRAPHIE

- [1] **A. Auslender**
Optimisation, Méthodes Numériques
Masson, Paris, 1976
- [2] **Alvaro R. De pierro**
A simultaneous Projections method for linear inequalities
linear Algebra and its Applications 64:243-253 1983.
- [3] **Y. Censor and Tommy Elfving**
New méthodes for linear inequalities
Linear Algebra and its Application 42: 199-211 1982.
- [4] **Y. Censor**
Row-Action Methods for Huge and sparse systems and their
application
SIAM Review vol 23 n° 4 october 1981.
- [5] **H.S.M. Coxeter**
The classification of zonohedra by means of projective diagrams
J.Math.pures.appliquées 41(1962) 137-156
- [6] **D. Dobkin and R.J. Lipton**
Multidimensional searching problems
SIAM J Comput vol 5 n°2 June 1976
- [7] **P.Chrétienne, Y.Pesqueux et J-C. Grandjean**
Algorithmes et pratique de la programmation linéaire
Technip Paris 1980.
- [8] **N. Gastinel**
Analyse numérique linéaire
Hermann Paris 1966.
- [9] **B. Grunbaum**
Convex Plytopes
Interscience 1967

- [10] **P.J. Laurent**
Approximation et optimisation
Hermann (Paris) 1972.
- [11] **W. Hildenbrand**
Short-run production functions based on microdata
Econometrica vol 49 n° 5 (1981) 1095-1125.
- [12] **T.Husain and Tweddle**
On the Extreme Points of the Sum of Two Compact Convex Sets
Math Ann 188,113-122 1970
- [13] **P.McMullen**
On zonotopes
Trans.Amer.Math.Soc 159(1971) 91-110.
- [14] **P.McMullen**
Space Tiling Zonotopes
Mathematika 22 (1975) 202-211
- [15] **P.McMullen, G.C.Shephard**
Polytopes with centrally symmetric facets
Israel Journal of Mathematics 23 (1976) 3-4.
- [16] **P.McMullen, G.C.Shephard**
Diagrams for centrally symmetric polytopes
Mathematika 15 (1968) 123-138
- [17] **G. Pierra**
Méthodes de projections parallèles extrapolées relatives à une
intersection de convexes
R.R n° 7 septembre 1975.
- [18] **F.P. Preparata M.I. Shamos**
Computational Geometry (an introduction)
Springer-verlag New york 1985.
- [19] **M. Sakarovitch**
Techniques mathématiques de la recherche opérationnelle
Poly E.N.S.I.M.A.G octobre 1979.

- [20] P. Valentin**
Zonoids, a geometric approach to Separation Engineering
with application to chromatography
Elf-solaize Research Center (à paraître)
- [21] P. Valentin**
Génie de la séparation
à paraître dans Hermann
- [22] J. Vignes**
Algorithmes numériques Analyse et mise en oeuvre
Technip Paris 1980
- [23] G.C. Shephard**
Combinatorial properties of associated zonotopes
Can. J. Math vol XXVI 2 (1974) 302-321.
- [24] G.C. Shephard**
Space-Filling Zonotopes
Mathematika 21 (1974), 261-269.
- [25] A.C. Yao and R. Rivest**
On the polyhedral decision problem
SIAM J Comput vol 9 n° 2 May 1980



A U T O R I S A T I O N de S O U T E N A N C E

VU les dispositions de l'article 3 de l'arrêté du 16 avril 1974

VU les rapports de présentation de Messieurs

- . P.J LAURENT, Professeur
- . P. VALENTIN, Ingénieur

Monsieur SLAOUI HASNAOUI Khalid

est autorisé à présenter une thèse en soutenance en vue de l'obtention du diplôme de
DOCTEUR-INGENIEUR, spécialité "Mathématiques appliquées".

Fait à Grenoble, le 13 juin 1986

D. BLOCH
Président
de l'Institut National Polytechnique
de Grenoble

P.O. le Vice-Président.





Khalid H. SLAOUI

DOCTORAT D'INGENIEUR

27 Juin 1986

I.N.P.G.

E.N.S. Informatique et Mathématiques Appliquées de Grenoble

Résumé

Ce travail se compose de deux parties complémentaires:

La première concerne la modélisation des histogrammes à intervalles adjacents ou chevauchés, dans un domaine borné, par les histosplines, ainsi que trois approches pour les histogrammes à intervalles non bornés. La motivation de cette étude, est le calcul des propriétés conservatives des coupes pétrolières, et la régularisation des courbes de ségrégation. Ceci a donné naissance à deux progiciels interactifs et visuels.

Dans la seconde partie, on présente la faisabilité d'un mélange objectif à partir d'une famille donnée de mélanges; plusieurs méthodes d'optimisation ont été testées (projections, décomposition et simplexe). On verra aussi la résolution numérique des opérations de base concernant la gestion des mélanges binaires par des techniques de géométrie algorithmique.

Mots clés

Convexe des mélanges, Chromatographie, Fonction Histospline, Fraction de ségrégation, Histogramme, N-paraffine, T.B.P, Zonotope.