



HAL
open science

Méthode de régularisation entropique et application au calcul de la fonction de distribution des ondes

Olivier Prot

► **To cite this version:**

Olivier Prot. Méthode de régularisation entropique et application au calcul de la fonction de distribution des ondes. Mathématiques [math]. Université d'Orléans, 2005. Français. NNT : . tel-00009999

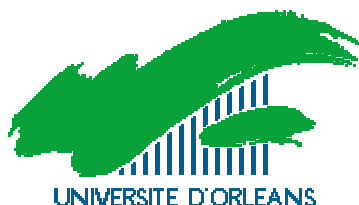
HAL Id: tel-00009999

<https://theses.hal.science/tel-00009999v1>

Submitted on 30 Aug 2005

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.



THÈSE
PRESENTÉE
À L'UNIVERSITÉ D'ORLÉANS
POUR OBTENIR LE GRADE DE
DOCTEUR DE L'UNIVERSITÉ D'ORLÉANS

Discipline : Mathématiques Appliquées

PAR

PROT OLIVIER

**Méthode de régularisation
entropique et application au calcul
de la fonction de distribution des
ondes**

Soutenue le 1 Juillet 2005

MEMBRES DU JURY :

-Jean-Gabriel Trotignon	Directeur de Thèse / Chargé de recherche, <i>LPCE/CNRS</i>
-Maïtine Bergounioux	Co-directrice de Thèse / Professeur, <i>Université d'Orléans</i>
-Dominikus Noll	Rapporteur / Professeur, <i>Université de Toulouse</i>
-Ondřej Santolík	Rapporteur / Associated professor, <i>Université de Prague</i>
-François Lefeuvre	Président / Directeur de recherche, <i>LPCE/CNRS</i>
-Aline Bonami	Professeur, <i>Université d'Orléans</i>
-Romain Abraham	Professeur, <i>Université d'Orléans (Invité)</i>

REMERCIEMENTS

J'ai commencé à travailler sur les méthodes d'analyses des ondes électromagnétiques avec Jean-Gabriel Trotignon lors de mon stage de Maîtrise au LPCE, durant l'été 2001. Je lui suis très reconnaissant de m'avoir proposé un sujet aussi intéressant, ce qui m'a conduit ensuite à m'engager dans ce travail de thèse. Jean-Gabriel a éveillé mon intérêt pour la physique des plasmas comme moyen d'investigation du milieu interplanétaire. Il m'a beaucoup aidé à prendre du recul par rapport à mon approche parfois "trop théorique" et éloignée du problème physique posé ou du dispositif expérimental. Je remercie très chaleureusement Maïtine Bergounioux qui m'a dirigé sur la partie mathématique de ce travail. Sa disponibilité, sa bonne humeur et son savoir-faire m'ont permis de travailler dans des conditions idéales.

Je remercie vivement Ondřej Santolík et le professeur Dominikus Noll d'avoir accepté d'être mes rapporteurs. Ondřej a suivi mes travaux depuis mon stage de DEA, où il m'a aidé à comprendre le modèle de fonction de distribution des ondes et les difficultés de l'utilisation d'un tel modèle pour l'interprétation d'expériences spatiales. C'est également Ondřej qui m'a proposé de tester mes méthodes sur les données de l'expérience FREJA. J'ai rencontré Dominikus Noll lors de la conférence Franco-Allemande-Espagnole d'optimisation à Avignon (2004), ses conseils et ses remarques sur mon travail m'ont beaucoup apporté pour l'achèvement de ce mémoire.

Un grand merci à Sandrine Grellier et Aline Bonami pour m'avoir "dirigé" après ma maîtrise vers la recherche. Je remercie Aline d'avoir accepté de faire partie de mon jury et de l'intérêt qu'elle porte à mes travaux.

Je remercie les membres des laboratoires LPCE et MAPMO pour leurs accueil chaleureux. En particulier, je remercie Jean-Louis Pinçon et François Lefeuvre qui ont été des interlocuteurs privilégiés tout au long de ces trois années, Pierre-Louis Blelly le directeur du LPCE, et Jean-Philippe Anker pour son écoute. Merci aussi à Philippe Jaming pour nous avoir appris à concevoir notre page web, et pour ses conseils. Je remercie également Eric Decreux pour ses discussions vraiment très instructives mais souvent bien trop conséquentes. Merci également à Romain Abraham de s'être intéressé à mon travail : j'espère qu'il ne m'en veut pas trop d'avoir "effacé" le nez de son petit bonhomme radiographié. Merci aux secrétaires du LPCE et du MAPMO : Jacqueline Nicoulaud, Isabelle Langer, Aurore Lecoustre, Laurence Chambolle, Laurent Royer et Corinne Revil, Anne Liger, Christelle Morillon, Virginie Foucault et Anne-Sophie Jaïs.

Je remercie Hervé Deféraudy pour m'avoir permis de travailler sur les données du satellite FREJA et m'avoir fait ses remarques sur mon analyse de ces données.

Je tiens à remercier les thésards du LPCE et du MAPMO avec qui ces trois années de thèse ont été très agréables. En particulier je tiens à remercier Hermine Biermé et Bruno Demange avec qui j'ai organisé le groupe de travail des thésards pendant ces deux dernières années. Avec Bruno (et sous les directives d'Aurore) nous avons tenté de transformer le groupe de travail des thésards en club de sport, mais cela n'a pas eu beaucoup de succès ! Avec Hermine nous avons régulièrement échangé nos impressions et nos points de vues sur nos travaux et sur le déroulement de la thèse, ce qui m'a bien aidé. Merci à Dominique Vieugué qui avait toujours un problème ou une anecdote mathématique intéressante ou amusante à nous soumettre (surtout lorsque nous étions tous débordés !).

Enfin, un grand merci aux membres de ma famille et a mes proches qui m'ont encouragé pendant ces trois années. En particulier je tiens à remercier Aurore pour sa compréhension et son soutien constant.

Table des matières

1	Introduction	1
I	Modèle physique et Approche théorique	7
2	Modèle	9
2.1	Description de l'environnement Terrestre	9
2.2	La FDO	10
2.2.1	Définitions et notations	10
2.2.2	L'équation de propagation des ondes	11
2.2.3	la FDO dans L^2	12
2.3	Les noyaux d'intégration du vide	14
2.3.1	Une nouvelle base	15
2.3.2	Expression du champ électrique E	16
2.3.3	Expression du champ magnétique H	17
2.3.4	Le vecteur champ électrique généralisé ε	17
3	Méthodes de Régularisation	19
3.1	Problème inverse mal posé	19
3.1.1	Exemples	20
3.2	Méthodes de Régularisation	21
3.2.1	Méthode de Tikhonov	23
3.2.2	Exemples numériques pour le calcul d'une dérivée	25
3.3	Régularisation entropique	26
3.3.1	L'entropie	27
3.3.2	Principe de la méthode	28
3.4	Formulation du problème	29
3.4.1	Algorithme 1	31
3.4.2	Second modèle	32
3.4.3	Algorithme 2	33
3.5	Dualité	34
3.5.1	Introduction et rappels	34
3.5.2	Théorème de Kuhn-Tucker	35

3.5.3	Dual du maximum d'entropie	36
3.5.4	Calcul du dual du problème (\mathcal{P}_μ)	38
3.5.5	Dual du problème de maximum d'entropie relaxé	39
4	Généralisation de la méthode	41
4.1	Introduction	41
4.2	Régularisation du problème inverse	41
4.2.1	Stabilité de la solution	44
4.3	Cas où ϕ est strictement convexe	45
4.3.1	Régularité de l'inverse γ	48
4.4	Exemples dans le cas strictement convexe	50
4.4.1	Cas fortement convexe	50
4.4.2	Cas particuliers où ϕ est non local	51
4.5	Problèmes avec contraintes	52
4.6	Résolution	55
4.6.1	Lien avec la dualité	56
4.6.2	Utilisation du théorème de Banach	57
4.6.3	Premières applications	58
4.6.4	Utilisation du théorème de Schauder	60
4.6.5	La méthode de Tikhonov dans L^p : Conclusion	62
4.6.6	Approximation de L^∞ par des normes L^p	64
II	Applications et simulation Numérique	65
5	Calcul de la FDO	67
5.1	Introduction	67
5.2	Cas du vide : données synthétiques	68
5.2.1	Mise en oeuvre numérique	68
5.2.2	Test avec trois masses de Dirac - Algorithme 1	71
5.2.3	Effet de L'entropie	72
5.2.4	Test avec une FDO continue - Algorithme 2	72
5.2.5	Test avec deux modes de polarisation	75
5.2.6	A propos de la dualité	76
5.3	Cas des plasmas	78
5.3.1	Préparation des données	79
5.3.2	Représentation des solutions obtenues	80
5.3.3	Résolution avec le champ magnétique	80
5.3.4	Calcul de la FDO avec toutes les composantes	80
5.3.5	Tests de stabilité I	82
5.3.6	Tests de stabilité II	84
5.3.7	Détection d'ondes planes	86
5.4	Discussion	87

6	Autres opérateurs	89
6.1	Inversion dans $H_0^2(E)$	89
6.2	Inversion dans $H^2(E)$ avec des conditions périodiques	93
6.3	Inversion dans $L^p(E)$	94
7	Conclusion et perspectives	99
A	Entropy Regularization Method	103
A.1	Introduction	103
A.2	Mathematical model	105
A.3	About entropy	106
A.4	A penalized problem	107
A.4.1	Relaxation of the feasible set	109
A.4.2	An infinite dimensional algorithm	110
A.4.3	A finite dimensional algorithm	112
A.5	Computing the probability density	116
A.6	Numerical tests	120
A.6.1	Wave distribution function in vacuum	120
A.6.2	Test of A1	122
A.6.3	Test of A2 on a continuous density	126
A.7	Conclusion	127
B	PROCEEDING WDS 2004	129
B.1	Introduction	129
B.2	Method	131
B.2.1	First minimization problem	132
B.2.2	Second optimization problem	132
B.3	Results	133
B.3.1	Analysis of an ELF hiss event recorded by Freja	133
B.3.2	Study of the ability to detect a plane wave	135
B.4	Discussion	136
C	ENTROPY REGULARIZATION METHOD APPLICATION	139
C.1	Introduction	140
C.2	Methods	141
C.2.1	Description of two other WDF methods	141
C.2.2	Entropy Regularization	142
C.2.3	Choice of the regularization parameter	144
C.2.4	Stability Analysis	145
C.3	ERA Implementation on the FREJA Data	146
C.4	FREJA data analysis results	147
C.4.1	Computed WDF	148
C.4.2	Stability Analysis	149
C.5	Discussion and Conclusions	149

D Programmes	155
D.1 Programme ERA I et ERA II	155
D.1.1 Description des fichiers	155
D.1.2 Listings des programmes	157
D.2 Conversion des données de FREJA	165
D.2.1 Listing des programmes	166

Table des figures

2.1	Système de coordonnées	15
3.1	Problème mal-posé : la dérivation	21
3.2	Transformée de Radon	22
3.3	Calcul numérique d'une dérivée	26
3.4	Non convexité de la fonction coût	33
5.1	Principe de décalage	71
5.2	FDO avec 3 pics	73
5.3	Effet de l'entropie	73
5.4	Variation de la solution en fonction de μ	74
5.5	FDO continue	77
5.6	Deux modes de polarisation	77
5.7	Spectrogramme	78
5.8	Représentation des solutions	81
5.9	Résultats avec le champ magnétique	81
5.10	Resultats avec toutes les composantes	83
5.11	Test de stabilité II	85
5.12	Détection d'une onde plane	86
6.1	Solution obtenue en utilisant la norme du gradient comme opérateur de régularisation, avec une contrainte de nullité au bord.	96
6.2	Solution obtenue en utilisant la norme H^1 comme opérateur de régularisation, sous une condition de périodicité.	96
6.3	Cette fonction a été obtenue en utilisant comme opérateur de régularisation, la norme H^1 translatée par la valeur moyenne de y	97
6.4	Solution obtenue à l'aide de la méthode de Tikhonov dans L^p	97
B.1	Coordinate system and representation of the WDF	134
B.2	Results with ERA, GP and ME methods	135
B.3	Plane wave direction determination	136
C.1	Coordinate system	147
C.2	Representation of the WDF	148

C.3	Results obtained with ERA, GP and ME	153
C.4	Stability analysis	154
D.1	Contenus du dossier ERA_I, avec la disposition des différents fichiers du programme.	156
D.2	Diagramme du programme ERA	158
D.3	Conversion des données FREJA	167

Liste des tableaux

5.1	Résultats pour différentes valeur de μ	75
5.2	Résultats avec les données FREJA	84
5.3	Deux ondes planes	87
A.1	WDF with 3 peaks	124
C.1	Results for each points	152

Chapitre 1

Introduction

Le but de cette thèse est la résolution d'un problème inverse hautement indéterminé, intervenant en physique spatiale lors de l'interprétation de données expérimentales. Le problème est celui de la détermination de la (ou des) direction(s) de propagation d'une onde électromagnétique à partir des mesures locales du champ. La résolution d'un tel problème inverse ne peut se faire qu'en rajoutant des hypothèses de manière à déterminer une solution unique et stable. Les méthodes de régularisation sont des méthodes très souvent utilisées pour résoudre les problèmes mal posés. Le principe de ces méthodes est de relaxer les contraintes du problème afin de déterminer une solution stable. Dans ce mémoire, nous proposons une méthode de régularisation entropique pour résoudre le problème considéré, ainsi que diverses généralisations. Mais, avant d'entrer dans les détails, voyons comment ce problème est historiquement apparu.

Depuis le début de l'ère spatiale, l'analyse des signaux électromagnétiques enregistrés par les satellites ou les sondes spatiales, est un des principaux moyens d'investigation de l'environnement terrestre. Les observations effectuées dans la magnétosphère terrestre, ont permis de découvrir différents types d'émissions naturelles. Les mécanismes générateurs et les sources d'émissions de ces ondes étant encore mal connus, il est primordial de pouvoir déterminer localement, à partir des mesures du champ, les directions de propagation de ces ondes [61].

Pour déterminer la direction de propagation d'une onde électromagnétique à partir de mesures simultanées du champ électromagnétique, la plupart des auteurs utilisent l'approximation "onde plane". Dans cette approximation l'onde se déplace dans une direction unique caractérisée par le vecteur d'onde k . Ce vecteur se calcule aisément, moyennant quelques hypothèses sur les caractéristiques de l'onde, en analysant les valeurs et vecteurs propres de la matrice spectrale. Il existe de nombreuses méthodes pour effectuer cette détermination [32, 2]. Toutefois, l'approximation "onde plane" apparaît trop restrictive car, dans la réalité il est peu probable que l'onde n'ait qu'une seule direction de propagation. Il y a, en effet, toujours des réflexions et des réfractions dans le milieu qui empêchent l'onde de se propager dans une direction unique [61]. Pour vérifier l'hypothèse de la présence d'une seule onde plane, le concept de degré de polarisation a été introduit [49, 40]. C'est un indicateur qui nous permet de dire si l'onde peut effectivement être considérée comme plane. Dans le cas

contraire, il faut utiliser d'autres méthodes d'investigation.

Pour une analyse plus réaliste, nous avons utilisé le concept de Fonction de Distribution des Ondes (FDO). Le concept de FDO a été introduit par *Storey et Lefeuvre* [61, 62, 63], il y a environ trente ans, pour étudier les ondes dans un plasma homogène. A une fréquence fixe ω , la FDO est une fonction $F : S \rightarrow \mathbb{R}^+$, où S est la sphère unité de \mathbb{R}^3 , qui représente la distribution d'énergie dans toutes les directions. Autrement dit, si l'on représente les points de la sphère en coordonnées sphériques grâce à l'angle polaire θ et l'angle azimutal ϕ , $F(\theta, \phi)$ est la quantité d'énergie de l'onde se propageant dans la direction (θ, ϕ) . La contrainte de positivité sur la fonction F provient clairement du fait qu'elle représente une énergie. *Storey et Lefeuvre* [61] ont montré qu'une équation intégrale liait la FDO F à la matrice spectrale des mesures du champ électromagnétique ; cette équation est

$$V = \int_0^{2\pi} \int_0^\pi q(\theta, \phi) F(\theta, \phi) \sin \theta \, d\phi \, d\theta \quad (1.1)$$

où $q : S \rightarrow \mathbb{C}^n$ est le noyau d'intégration, i.e. une fonction connue qui dépend de la nature du milieu de propagation. La donnée $V \in \mathbb{C}^n$ est la matrice spectrale que nous avons identifiée à un vecteur à n composantes complexes. L'équation (1.1) est très importante car c'est elle qui va définir le problème inverse que nous voulons étudier. Notons qu'il existe différentes définitions de la fonction de distribution des ondes, par exemple *Rönmark* [48] la définit à partir de la transformée de Wiener.

Le problème direct est le suivant : à partir d'une FDO $F : S \rightarrow \mathbb{R}^+$ donnée, déterminer la matrice spectrale V correspondante. Pour résoudre le problème direct, il suffit donc de déterminer les noyaux d'intégration $q : S \rightarrow \mathbb{C}^n$ du milieu considéré. Ce problème a été résolu par *Storey et Lefeuvre* [62, 63], qui ont donné l'expression analytique de ces noyaux sous l'hypothèse d'un plasma homogène et froid. Dans sa thèse, *Santolík* [50] a résolu le problème direct dans le cas d'un plasma chaud avec prise en compte de l'effet Doppler [51, 52]. Dans ce cas les noyaux ne sont plus connus de façon analytique, mais ils sont déterminés à partir de la relation de dispersion du milieu considéré [50].

Le problème inverse est celui de la détermination de F à partir de V en résolvant l'équation (1.1). C'est un problème mal posé, car il existe une infinité de solutions vérifiant cette équation puisque l'opérateur intégral n'est pas bijectif. Il transforme un élément d'un espace de dimension infinie F en un vecteur complexe de dimension finie. La détermination d'une solution unique et stable à ce problème ne peut se faire qu'en rajoutant des critères de sélection sur la fonction recherchée. Pour ce faire, la plupart des méthodes utilisées sont basées sur le choix d'un modèle de FDO. Le calcul de F revient donc à un problème d'identification de paramètres, ce qui est fait en pratique via des méthodes classiques de minimisation, telles que les méthodes de descente ou de quasi-Newton. *Lefeuvre* [24], a étudié différentes méthodes de résolution pour ce problème mal-posé.

Une des méthodes les plus utilisées pour traiter des données expérimentales est la méthode de Maximum d'Entropie (ME) de *Lefeuvre et Delannoy* [25, 29]. Cette méthode a pour principe de tronquer les noyaux d'intégration afin de ne conserver que ceux qui sont réellement significatifs pour l'interprétation physique. Un modèle visant à maximiser l'entropie est alors utilisé et les paramètres (multiplicateurs de Lagrange) identifiés par un algorithme de minimisation. Dans le programme MAXENTWDF de *Lefeuvre et Delannoy* [25] cette minimisation est

faite à l'aide de la méthode de Levenberg-Marquardt. L'entropie est un critère de sélection qui permet de déterminer une solution, unique lisse et positive. De plus, on peut interpréter l'entropie d'une fonction comme l'opposée de la quantité d'information contenues dans cette fonction [57]. C'est pourquoi la maximisation de l'entropie permet de sélectionner la solution du problème inverse qui contient le moins d'information. Dans le cas d'un problème mal posé tel que (1.1), il est préférable de chercher une solution contenant un minimum d'information. Il existe en effet une infinité de solutions dont certaines apportent une information considérable, mais nous n'avons aucun moyen de dire si l'information apportée par une solution est pertinente ou pas. Toutefois, la stabilité de la solution n'est pas assurée par la maximisation de l'entropie. C'est la troncature des noyaux les moins significatifs qui va stabiliser la solution. Cette méthode a l'avantage de déterminer la solution de façon rapide, car il y a peu de paramètres et on les obtient en minimisant une fonction convexe.

Différentes méthodes ont été conçues pour déterminer la FDO en privilégiant quelques directions de propagations [12, 24, 50]. La méthode des pics gaussiens (GP), décrite dans la thèse de *Santolík* [53], consiste à choisir un modèle de FDO constitué d'une somme de gaussiennes. Les paramètres de position, énergie et largeur du faisceau sont ensuite déterminés pour chacune d'entre elles, par une minimisation au sens des moindres carrés. Notons que la détermination de ces paramètres n'est pas une chose aisée puisque la fonction à minimiser n'est pas convexe. L'identification des paramètres du modèle devient très difficile lorsque le nombre de pics gaussiens est choisi trop grand. C'est pourquoi, en pratique cette méthode est utilisée pour déterminer un, deux ou trois pics.

Santolík [50] a également décrit une méthode d'inversion avec un modèle de FDO discret. Comme la FDO est définie sur la sphère unité, sa discrétisation est un problème délicat. *Santolík* [50] a utilisé une discrétisation de la sphère en forme "d'igloo" de manière à ce que toutes les briques aient à peu près la même taille. Les valeurs de la FDO sur le maillage sont ensuite déterminées par une minimisation au sens des moindres carrés. Cette méthode a l'avantage de n'avoir aucun paramètre à fixer avant les calculs, si ce n'est la finesse du maillage utilisé.

Dans la méthode du maximum d'entropie [25], le nombre de noyaux à tronquer doit être fixé de façon à obtenir une solution stable. Notons que l'analyse des données de la mission GEOS-1 [39, 60] a par ailleurs montré qu'un grand nombre de données ne pouvaient être analysées à l'aide la méthode du maximum d'entropie : en effet, soit l'erreur avec la donnée était trop grande, soit la solution était instable. Une autre approche consiste à garder tous les noyaux du modèle et de laisser à la méthode le soin de sélectionner automatiquement ceux qui sont les plus importants pour l'inversion. C'est cette dernière qui a été suivie dans ce mémoire. D'autres méthodes ont été proposées et utilisées avec succès, en particulier les modèles paramétriques de FDO. Ces dernières nécessitent de fixer un certain nombre de paramètres pour pouvoir effectuer les calculs. Par exemple, avec la méthode des Pics Gaussiens (GP), il faut fixer le nombre de pics à identifier avant de calculer la solution. La méthode présentée dans ce mémoire permet d'obtenir une solution plus générale, ne se limitant pas à un modèle donné, et sans avoir à fixer de paramètres à l'avance.

La méthode préconisée est une méthode de régularisation. Les méthodes de régularisation sont très employées pour résoudre les problèmes mal posés [22]. Le principe de ces méthodes est d'utiliser une famille d'opérateurs, indexée par un paramètre $\mu > 0$, pour déterminer la

solution. Ce paramètre μ est appelé paramètre de régularisation et sert à créer un équilibre entre la stabilité de la solution et l'erreur. Si μ est choisi petit, l'erreur commise est petite mais la solution est instable. A l'inverse lorsque μ est plus grand, l'erreur est plus grande mais la stabilité est améliorée. La méthode de régularisation la plus connue est la méthode de *Tikhonov* [65]. Le principe de cette méthode est de déterminer la solution du problème en choisissant l'élément de norme minimale situé dans un voisinage donné de la solution. Toutefois, la méthode de Tikhonov ne fournit pas de fonctions positives ; or la contrainte de positivité est primordiale dans notre modèle puisque la FDO représente une distribution d'énergie.

La méthode de régularisation entropique est similaire à celle de Tikhonov, sauf que ce n'est pas la norme qui est minimisée mais l'entropie que l'on maximise. L'utilisation de l'entropie a plusieurs avantages : premièrement l'entropie agit comme une fonction barrière ce qui oblige la solution à être positive, deuxièmement on montre que la solution obtenue a la même régularité que les noyaux d'intégration du modèle. Comme dans la méthode de maximum d'entropie, l'entropie permet de déterminer une solution qui contient peu d'information. Il est connu que l'entropie permet de définir une stratégie de régularisation dans L^1 [57, 16]. Nous avons plutôt étudié le problème dans L^2 et cherché à déterminer sous quelles hypothèses l'inversion était possible. Nous avons également cherché à savoir comment déterminer, numériquement et de façon efficace la solution à partir de la condition d'optimalité. Résoudre le problème de cette façon n'est pas totalement satisfaisant car la solution ne dépend pas de façon linéaire de la donnée. Ce problème a été résolu en modifiant légèrement le problème de minimisation utilisé. Puis une méthode numérique a été développée pour résoudre ce nouveau problème.

Le cadre de l'étude précédente étant particulièrement lié au le calcul de la FDO, nous avons cherché à généraliser cette méthode dans un cadre plus abstrait. Un espace de Banach réflexif X est considéré à la place de l'espace L^2 , et nous avons remplacé l'entropie par un opérateur de régularisation ϕ s.c.i., Gâteaux-différentiable et convexe. Nous avons alors cherché à déterminer sous quelles hypothèses nous pouvions résoudre le problème inverse de façon similaire à celle utilisée pour la régularisation entropique. La solution du problème inverse doit exister et être unique pour toutes valeurs du paramètre de régularisation, mais la stabilité de la solution est primordiale. Nous nous sommes également intéressés au problème du calcul numérique de la solution.

Comme nous l'avons déjà mentionné, la détermination des directions de propagation des ondes est un problème crucial qui a de nombreuses applications. L'utilisation du concept de FDO permet une analyse correcte du phénomène dans le cas où la mesure est effectuée dans un milieu homogène. Ce milieu peut être le vide ou un plasma. Dans le cas du vide, l'analyse par FDO peut être utilisée pour interpréter les données mesurées par un radar. Cela avait d'ailleurs été envisagé [27] pour analyser les données du G.P.R. (Ground Penetrating Radar) de la mission martienne Netlander. Le G.P.R. est un radar dont le but était de sonder les premiers kilomètres du sous-sol de Mars afin de détecter la présence de réservoirs d'eau glacée.

Pour tester la méthode de régularisation entropique dans le vide, nous l'avons mis en œuvre sur des données synthétiques. Les données simulées ont été générées en supposant que l'onde est constituée d'une somme finie d'ondes planes, ou bien que ses directions de propagation forment un continuum. Dans le vide les ondes peuvent se propager dans différents modes de polarisation. Afin d'analyser l'onde à l'aide du concept de FDO, en utilisant l'équation

(1.1), il est primordial de supposer que l'onde n'a qu'un seul mode de polarisation. Mais le problème, plus complexe, où l'onde est constituée de différents modes est aussi plus réaliste. Un test numérique a ainsi été réalisé pour voir si la méthode de régularisation entropique permettait de conclure dans ce cas.

Nous avons également testé la méthode dans le cas où l'onde considéré est dans un plasma. Pour cela, les données du satellite de recherche magnétosphérique FREJA ont été utilisées, l'événement considéré étant un souffle ELF. Ces données ont déjà été traitées, nous avons pu comparer les résultats obtenus par régularisation entropique avec ceux d'autres méthodes. La stabilité de la solution étant primordiale pour l'interprétation, nous l'avons étudié à l'aide de tests statistiques.

Des tests numériques ont enfin été effectués pour illustrer la généralisation de la méthode. Le problème inverse considéré est celui de l'approximation d'une fonction, définie sur un compact, dont la valeur est connue sur un nombre fini de points. Différents opérateurs de régularisation ont été considérés pour cette détermination, afin de comparer les fonctions obtenues.

Ce mémoire est organisé en deux parties principales, la première traite du modèle et de l'approche théorique du problème inverse. Nous donnons une description de l'environnement spatial proche de la terre et posons le problème de l'analyse des ondes dans ce milieu au Chapitre 2. Ensuite, nous décrivons comment est obtenue l'équation (1.1) qui définit le problème inverse. Pour cela nous ferons quelques rappels à propos de l'analyse de Fourier qui est sous-jacente au modèle. Le chapitre 3 contient une description de la théorie de la régularisation. Nous décrivons également la méthode de Tikhonov de façon à introduire la régularisation entropique. Nous présentons ensuite les résultats connus sur cette méthode, particulièrement ceux d'*Amato et Hughes* [1] qui montrent que c'est bien une stratégie de régularisation. La méthode de régularisation entropique utilisée pour résoudre le problème mal posé du calcul de la FDO est ensuite présentée. Enfin, dans la dernière section de ce chapitre, nous étudions la dualité du problème de régularisation entropique. Nous y ferons le lien entre nos algorithmes et le problème dual. Le chapitre 4, présente la généralisation de la méthode d'inversion pour des espaces de Banach abstraits avec un opérateur de régularisation strictement convexe. Dans un premier temps, une étude théorique de la méthode est faite pour savoir sous quelles hypothèses peut-on trouver une solution unique et stable qui peut se déterminer comme un point fixe en dimension finie. Ensuite nous nous intéressons à la résolution numérique du problème, en utilisant le théorème de point fixe de Banach, comme ce qui à été fait pour montrer les résultats du Chapitre 3.

La seconde partie principale du mémoire traite de la résolution numérique du problème inverse. Le Chapitre 5 rassemble les tests numériques effectués sur le problème de la détermination de la fonction de distribution des ondes. Nous commençons par tester la méthode sur des données synthétiques simulées dans le cas d'ondes se propageant dans le vide. Nous considérons ensuite la résolution du problème dans un plasma magnétosphérique. Pour cela nous avons utilisé des données provenant du satellite FREJA. L'événement que nous étudions est un souffle ELF (Extremely Low Frequencies) enregistré par le satellite le 8 avril 1995 ; les données consistent en six matrices spectrales qui ont déjà été analysées par *Santolík et Parrot* dans l'article [53]. Nous avons strictement utilisé les mêmes données et le même modèle pour le calcul de la FDO, afin de comparer entre elles les diverses méthodes utilisées. Nous

avons plus particulièrement considéré la méthode de régularisation entropique, la méthode de maximum d'entropie de *Lefevre et Delannoy* [25, 26], et la méthode des pics gaussiens utilisée par *Santolík* [50]. Nous avons également fait différents tests statistiques pour vérifier et quantifier la stabilité de la méthode, ainsi que pour déterminer son aptitude à détecter correctement une onde plane.

Les Annexes A, B et C sont trois articles écrits pendant ma thèse. L'Annexe D contient les informations nécessaires pour faire fonctionner le programme d'inversion programmé en langage MATLAB. C'est, ce programme, qui a servi à faire tous les calculs numériques de ce mémoire.

Première partie

Modèle physique et Approche
théorique

Chapitre 2

Rappels : plasmas spatiaux et ondes électromagnétiques

2.1 Description de l'environnement Terrestre et analyse des ondes qui s'y propagent

La magnétosphère est une cavité entourant la planète Terre gouvernée par le champ magnétique terrestre. Sa découverte remonte en 1958 où la sonde Explorer 1 a détecté les ceintures de radiation (Van Allen) grâce à un capteur Geiger. La magnétosphère est constituée de plasma, un gaz ionisé conducteur. On décrit souvent le plasma comme le “quatrième état de la matière”, et on pense que 99% de la masse de l'univers existe sous cette forme. La physique des plasmas est de ce fait indispensable pour décrire les phénomènes qui régissent la magnétosphère.

D'une part, le mouvement des particules dans la magnétosphère est conditionné par le champ magnétique terrestre. D'autre part, la magnétosphère est soumise à un flux de plasma provenant du soleil : le *vent solaire*. L'existence du vent solaire a été proposée dans les années 50 par E. Parker pour expliquer la présence de la ceinture de radiation. Les premiers signes expérimentaux du vent solaire ont été observés quelques années plus tard. La magnétosphère agit comme un bouclier qui fait obstacle au vent solaire, la forme de ce bouclier changeant continuellement pour s'adapter à la pression du vent solaire. Le vent solaire comprime la cavité magnétosphérique du côté jour, et crée une importante queue magnétique du côté nuit. La rencontre entre le flux de plasma et la magnétosphère donne naissance à une onde de choc en avant de la magnétosphère. Lorsque le vent solaire a de grandes variations, par exemple après une éruption solaire ou une éjection de masse coronale, la variation de pression sur la magnétosphère terrestre peut déclencher des perturbations du magnétisme terrestre appelées les “orages magnétiques”. Les aurores polaires sont les conséquences de ces orages. Elles sont créées par l'injection de particules énergétiques dans l'atmosphère terrestre. Nous voyons donc que l'activité solaire et la magnétosphère terrestre sont intimement liées. Les phénomènes se produisant dans la magnétosphère étant très complexes, nous ne disposons encore d'aucun modèle complet qui les décrivent correctement. Les expériences de diagnostic de plasma, embarquées à bord de satellites sont donc d'une très grande importance pour l'étude des

phénomènes se produisant dans l'environnement terrestre.

Notons que la Terre n'est pas la seule planète à posséder une magnétosphère, plusieurs planètes du système solaire en possèdent également une. Par exemple les planètes géantes Jupiter et Saturne possèdent des magnétosphères très étendues, la taille de ces dernières étant liées à leurs champs magnétiques internes qui sont beaucoup plus intense que celui de la Terre. Comme sur la Terre, on observe également des aurores polaires sur ces planètes. De cette manière, la connaissance de la magnétosphère terrestre permet de servir de référence dans les études des phénomènes d'interaction entre les corps du système solaire et le vent solaire.

Dans le plasma magnétosphérique, les particules chargées sont soumises à l'action des champs électrique et magnétique, elles génèrent ainsi des ondes sous l'effet d'instabilités, ces ondes influencent à leur tour les particules du milieu. Les ondes jouent donc un rôle important dans le comportement et l'évolution de la magnétosphère terrestre. L'étude de ces ondes est ainsi un moyen puissant de diagnostiquer des plasmas, pour cela on utilise un des capteurs électriques et magnétiques afin de mesurer les formes d'ondes. Ce qui permet le calcul de la matrice spectrale avec laquelle les caractéristiques des ondes seront dérivé (mode de propagation, polarisation, direction de propagation).

2.2 Analyse des signaux et fonction de distribution des ondes

L'équation (1.1) à la base du problème inverse que nous souhaitons résoudre à été établie il y a environ 30 ans par *Storey et Lefeuvre* [61, 62, 63]. Avant de commencer l'étude de la résolution de ce problème inverse, nous allons donner une idée rapide de la construction de ce modèle. Nous commençons d'abord par quelques définitions utiles pour le traitement du signal.

2.2.1 Définitions et notations

Modélisons le champ électromagnétique mesuré par les antennes via une fonction (ou signal) $m : \mathbb{R} \rightarrow \mathbb{C}^n$, $t \mapsto X(t)$, où la variable t représente le temps.

Définition 1 *L'énergie E_m du signal m est définie par*

$$E_m = \int_{\mathbb{R}} |m(t)|^2 dt;$$

lorsque m est assez régulier, nous pouvons définir la puissance à l'instant t par $|m(t)|^2$.

Du point de vue de cette définition, l'espace $L^2(\mathbb{R})$ est l'espace des signaux définis sur \mathbb{R} , d'énergie finie. On notera $\hat{m}(\nu)$ la transformée de Fourier de m , i.e.

$$\hat{m}(\nu) = \int_{\mathbb{R}} m(t) \exp(-2\pi i t \nu) dt.$$

La transformée de Fourier permet de voir un signal par les fréquences qui le composent. Dans le cadre de la physique expérimentale, pour l'analyse des ondes, il est préférable de voir le

signal mesuré du côté des fréquences plutôt que du côté temporel. Car les phénomènes étudiés sont souvent visibles dans une certaine bande de fréquences plutôt que temporellement. Voici quelques propriétés de la transformée de Fourier sur l'espace $L^2(\mathbb{R})$ voir [55].

Théorème 1 *La transformée de Fourier est un isomorphisme de l'espace $L^2(\mathbb{R})$. C'est aussi une isométrie par l'égalité de Plancherel*

$$E_m = \|m\|_{L^2}^2 = \|\hat{m}\|_{L^2}^2 \quad (2.1)$$

Grâce à l'égalité (2.1), nous pouvons mesurer l'énergie du signal m soit du côté temporel, soit du côté fréquentiel. En conséquence, nous définissons la *densité de puissance spectrale* par la fonction $\nu \mapsto |\hat{m}(\nu)|^2$. La *matrice spectrale de puissance* à la fréquence ν est, quant à elle, définie par $\hat{m}(\nu)\hat{m}^*(\nu)$. C'est grâce à cette matrice spectrale que nous allons chercher à déterminer les caractéristiques de l'onde électromagnétique à la fréquence ν .

En réalité, le signal mesuré est toujours entâché d'un bruit. Il n'est donc pas déterministe. C'est pourquoi, des moments statistiques du signal mesuré doivent intervenir. Sous quelques hypothèses de stationnarité et d'ergodicité du signal aléatoire m , on définit la densité spectrale par $\nu \mapsto \mathbb{E}(|\hat{m}(\nu)|^2)$ où \mathbb{E} est l'espérance. De la même manière, la matrice spectrale à la puissance ν est définie par $\mathbb{E}(\hat{m}(\nu)\hat{m}^*(\nu))$. Pour plus de détails à propos des techniques du traitement du signal on peut se référer à [31, 36, 46, 66]

2.2.2 L'équation de propagation des ondes

Soit une fonction

$$\begin{cases} X : \mathbb{R} \times \mathbb{R}^n & \longrightarrow \mathbb{C}^n \\ (t, x) & \longmapsto X(t, x) \end{cases} ,$$

la variable t représentant le temps et x la variable d'espace. Pour que la fonction X modélise une onde elle doit vérifier l'équation de propagation des ondes :

$$\partial_t^2 X = c^{-1} \Delta X \quad (2.2)$$

où $c > 0$ est la célérité de l'onde et Δ le laplacien. Les équations de Maxwell [59] montrent que les champs d'une onde électromagnétique vérifient cette équation. Nous allons nous intéresser à une famille particulière de solutions de l'équation (2.2) : les ondes planes.

Définition 2 *On appelle ondes planes, les solutions $\{e_{w,k}\}$ de l'équation (2.2) définies par*

$$e_{w,k}(t, x) = A \exp(2\pi i [wt - \langle k, x \rangle_{\mathbb{R}^n}])$$

où $A \in \mathbb{C}^n$, $w \in \mathbb{R}$ et $k \in \mathbb{R}^n$. Comme ce sont des solutions de (2.2), on en déduit que w et k sont reliés par l'équation

$$cw^2 = \|k\|_{\mathbb{R}^n}^2 \quad (2.3)$$

Ces ondes sont appelées ondes planes car pour t fixé, elles sont constantes sur le plan

$$k^\perp = \{x \in \mathbb{R}^n \mid \langle k, x \rangle_{\mathbb{R}^n} = 0\} .$$

La relation (2.3) permet de voir que k est la direction de propagation de l'onde puisque

$$e_{w,k} \left(t, x + \frac{k}{cw} \right) = e_{w,k}(t-1, x).$$

Comme l'onde plane $e_{w,k}(t, x)$ a une unique direction de propagation, on peut définir sa FDO par une masse de Dirac $|A|^2 \delta_{w,k}$. En effet, toute l'énergie de l'onde $|A|^2$ est concentrée en une fréquence w et une normale d'onde k . Il faut remarquer que les ondes planes sont des objets théoriques puisqu'elles ont une énergie infinie, ce sont des éléments de $\mathcal{S}'(\mathbb{R}^{n+1})$.

Pour être plus réaliste, nous allons dans la suite essayer de définir la FDO d'une fonction d'énergie finie, i.e. de $L^2(\mathbb{R}^{n+1})$. Puis en passant à la limite, nous allons voir si cette définition est cohérente avec celle des ondes planes. Pour finir, remarquons que la notion d'onde à énergie finie est incompatible avec l'équation des ondes. En effet :

Proposition 1 *L'unique solution d'énergie finie $X : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{C}$ de l'équation (2.2), c'est-à-dire dans $L^2(\mathbb{R} \times \mathbb{R}^n)$, est la solution nulle.*

Démonstration. Supposons que $X \in L^2(\mathbb{R} \times \mathbb{R}^n)$ vérifie l'équation des ondes (2.2). En calculant la transformée de Fourier de X on obtient $(c\nu^2 - \|\xi\|^2)\hat{X}(\nu, \xi) = 0$, avec $(\nu, \xi) \in \mathbb{R} \times \mathbb{R}^n$. On en déduit que $\hat{X}(\nu, \xi) = 0$ en dehors du cône $\left\{ (\nu, \xi) \in \mathbb{R} \times \mathbb{R}^n \mid c\nu^2 = \|\xi\|^2 \right\}$. Comme ce cône est de mesure nulle dans $\mathbb{R} \times \mathbb{R}^n$, on obtient $\hat{X} = 0$ p.p. \square

2.2.3 La fonction de distribution des ondes dans L^2

Soit $X \in L^2(\mathbb{R}^4)$ le signal représentant l'onde que nous voulons analyser. Pour que X représente correctement une onde, il faudrait qu'elle vérifie l'équation de propagation (2.2). Ceci est impossible d'après la Proposition 1. Puisque la fonction X ne vérifie plus l'équation (2.2), il n'est plus du tout certain que les vecteurs k représentent des directions de propagation. Pour cela, il faudrait retrouver une version affaiblie de la relation (2.3). Nous allons modifier les ondes planes de manière à les rendre d'énergie finie et qu'elles soient "presque" des solutions de (2.2). Nous allons donc les multiplier par une fonction de la classe de Schwartz $\mathcal{S}(\mathbb{R} \times \mathbb{R}^n)$ qui est presque constante dans un voisinage de zéro. Plus précisément, si $g(t, x) \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^n)$, on considère que la fonction d'énergie finie

$$X(t, x) = A \exp(2\pi i [wt - \langle k, x \rangle_{\mathbb{R}^n}]) g(t, x) \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^n) \subset L^2(\mathbb{R} \times \mathbb{R}^n) \quad (2.4)$$

est une onde plane. On choisit la fonction g paire et de telle sorte qu'elle soit très mal localisée et presque constante dans un voisinage de 0. De cette manière l'équation de propagation des ondes est vérifiée sur ce voisinage. On remarque que

$$e_{w,k} \left(t, x + \frac{k}{cw} \right) g(t, x) = e_{w,k}(t-1, x) g(t, x).$$

Comme g est presque constante dans un voisinage de zéro, cela signifie que l'onde se propage bien dans la direction de la normale d'onde k pour x, t assez petits. Nous pouvons maintenant donner la définition de la FDO du signal X .

Définition 3 On définit la FDO F_X du signal X , par

$$\begin{cases} F_X : \mathbb{R}^{n+1} & \longrightarrow \mathbb{R}^+ \\ (\nu, \xi) & \longmapsto |\hat{X}(\nu, \xi)|^2 \end{cases}$$

où

$$\hat{X}(\nu, \xi) := \int_{\mathbb{R}} X(t, x) \exp(-2\pi i [t\nu - \langle \xi, x \rangle_{\mathbb{R}^n}]) dt.$$

Cette définition correspond bien à l'intuition que l'on avait dans le cas des ondes planes. Mais on ne pouvait pas donner cette définition dans l'espace des distributions tempérées car on ne peut définir le produit de deux mesures de Dirac [21, 54]. Par l'égalité de Parseval l'intégrale de la FDO est égale à l'énergie totale de l'onde. La définition de la FDO correspond donc bien au fait que F_X donne la répartition de l'énergie suivant la fréquence et la normale d'onde. Il faut maintenant vérifier que pour une fonction de la forme (2.4), sa FDO se concentre autour de w et k . Pour cela nous avons besoin du principe d'incertitude d'Heisenberg [17, 67].

Théorème 2 (Inégalité d'Heisenberg) Soit $f \in L^2(\mathbb{R})$

$$\int_{\mathbb{R}} x^2 |f(x)|^2 dx \int_{\mathbb{R}} \xi^2 |\hat{f}(\xi)|^2 d\xi \geq \frac{1}{4\pi^2} \|f\|_{L^2}^4 \quad (2.5)$$

Ce théorème signifie que si une fonction est très localisée alors sa transformée de Fourier est très mal localisée. En particulier on peut montrer (Théorème de Paley-Wiener) qu'une fonction à support compact a pour transformée de Fourier une fonction entière [17, 68]. Une illustration plus concrète de ce théorème est le cas de la fonction $x \mapsto e^{-\delta x^2}$ où $\delta > 0$. Sa transformée de Fourier est $\xi \mapsto \frac{\sqrt{\pi}}{2} e^{-\frac{\pi^2 \xi^2}{\delta}}$. Notons que le cas de la gaussienne est très particulier car elle minimise le principe d'incertitude (2.5).

Calculons maintenant la FDO de X : on obtient $F_X(\nu, \xi) = |\hat{g}(\nu - w, \xi - k)|^2$. Mais en vertu du principe d'incertitude de Heisenberg, \hat{g} est une fonction très localisée en 0. Donc la FDO F_X a ses valeurs concentrées autour de w et de k . On peut voir cela comme une version affaiblie de (2.3). En conclusion, X décrit bien une onde plane puisque elle satisfait bien l'équation de propagation des ondes sur un voisinage de 0 et puisque les valeurs de la FDO sont concentrées autour de k et de w .

Nous allons chercher à obtenir l'équation qui relie la FDO aux mesures du champ (1.1). Pour cela, nous allons considérer qu'avec une certaine polarisation, une onde plane de fréquence w et de normale d'onde k s'écrit

$$A(k, w) \exp(2\pi i [wt - \langle k, x \rangle_{\mathbb{R}^n}]) g(t, x) \quad (2.6)$$

où $A : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{C}^n$ est une fonction connue qui dépend du milieu de propagation et du type de polarisation choisi. Nous allons, pour une fréquence fixée w , faire une somme d'ondes planes de normales d'onde différentes. On considère donc que l'onde électromagnétique est donnée par

$$X(t, x) = \int_{\mathbb{R}^n} A(k, w) \rho(k) \exp(2\pi i [wt - \langle k, x \rangle_{\mathbb{R}^n}]) g(t, x) dk \quad (2.7)$$

on en déduit que $\hat{X}(\nu, \xi) = \int_{\mathbb{R}^n} A(k, w)\rho(k)\hat{g}(\nu - w, \xi - k)dk$. Supposons que la fonction \hat{g} est de la forme $\hat{g} = h \otimes u$ avec $h \in \mathcal{S}(\mathbb{R})$ pour la variable $\nu \in \mathbb{R}$ et $u \in \mathcal{S}(\mathbb{R}^n)$ pour $\xi \in \mathbb{R}^n$. Si u est presque une approximation de l'identité, alors

$$\hat{X}(\nu, \xi) \approx A(\xi, w)\rho(\xi)h(\nu - w)$$

et la FDO de X est

$$F_X(\nu, \xi) = |\hat{X}(\nu, \xi)|^2 \approx |\rho(\xi)h(\nu - w)|^2$$

puisque pour tout $\xi \in \mathbb{R}^n$ et $w \in \mathbb{R}$, $|A(\xi, w)| = 1$.

Afin d'obtenir la formule qui relie la FDO aux mesures du champ (1.1), nous devons faire des hypothèses sur les propriétés statistiques du signal mesuré.

Hypothèse 1 *La matrice spectrale V estimée à partir des mesures du champ, est aussi une estimation de l'intégrale*

$$\int_{\mathbb{R}^n} \hat{X}(\nu, \xi)\hat{X}^*(\nu, \xi)d\xi$$

Cette hypothèse peut être considérée comme une sorte d'ergodicité, on déduit que

$$V \approx \int_{\mathbb{R}^n} \hat{X}(\nu, \xi)\hat{X}^*(\nu, \xi)d\xi \approx \int_{\mathbb{R}^n} A(\xi, w)A^*(\xi, w)|\rho(\xi)h(\nu - w)|^2d\xi \quad (2.8)$$

$$V \approx \int_{\mathbb{R}^n} q(\xi, w)F_X(\xi, w)d\xi \quad (2.9)$$

où $q(\xi, w) = A(\xi, w)A^*(\xi, w)$ est le noyau d'intégration. Nous avons donc retrouvé la formule (1.1) qui relie la matrice spectrale des mesures à la FDO [61, 62, 63]. Toutefois, il faut bien avoir à l'esprit que sans l'hypothèse (1) le modèle comporte des termes croisés.

2.3 Les noyaux d'intégration du vide

Dans la section précédente nous avons retrouvé l'équation reliant la FDO à la matrice spectrale mesurée. Maintenant nous allons voir un exemple de résolution du problème direct : il s'agit de déterminer la matrice spectrale lorsque l'on connaît la FDO. Le problème direct revient tout simplement à déterminer les noyaux d'intégration selon les hypothèses sur le milieu de propagation. L'expression analytique des noyaux d'intégration du modèle plasma froid a été donnée par *Storey et Lefevre* [61, 62]. Ce modèle a ensuite été modifié par *Santolík* [50] en utilisant un modèle de plasma chaud et en prenant en compte l'effet Doppler. Dans l'approche de *Santolík*, les noyaux sont calculées numériquement en résolvant la relation de dispersion du plasma. Pour plus de simplicité, nous avons choisi de montrer comment sont obtenus les noyaux d'intégration du vide, qui ont une expression analytique [27]. Soulignons que ce sont ces noyaux que nous avons utilisés pour les exemples numériques de la Section 5.2.

Considérons l'espace euclidien \mathbb{R}^3 muni du produit scalaire habituel, et $\{x, y, z\}$ la base canonique que nous désignerons par \mathcal{B} . Nous allons considérer dans la suite une onde électromagnétique plane de champ électrique E et magnétique H définie par

$$E(t) = \text{Re} [\mathcal{E}(t)], H(t) = \text{Re} [\mathcal{H}(t)]$$

avec

$$\mathcal{E}(t) = A \exp(i(\omega t)), \mathcal{H}(t) = B \exp(i(\omega t))$$

où $A, B \in \mathbb{C}^3$ sont les amplitudes complexes du champ et $\omega \in \mathbb{R}$ la pulsation. Le vecteur d'onde $K \in \mathbb{R}^3$ de l'onde électromagnétique est tel que pour tout $t \in \mathbb{R}$, $\langle E(t), H(t) \rangle_{\mathbb{R}^3} = 0$, $\langle E(t), K \rangle_{\mathbb{R}^3} = 0$ et $\{K, E(t), H(t)\}$ forment un trièdre direct.

2.3.1 Une nouvelle base

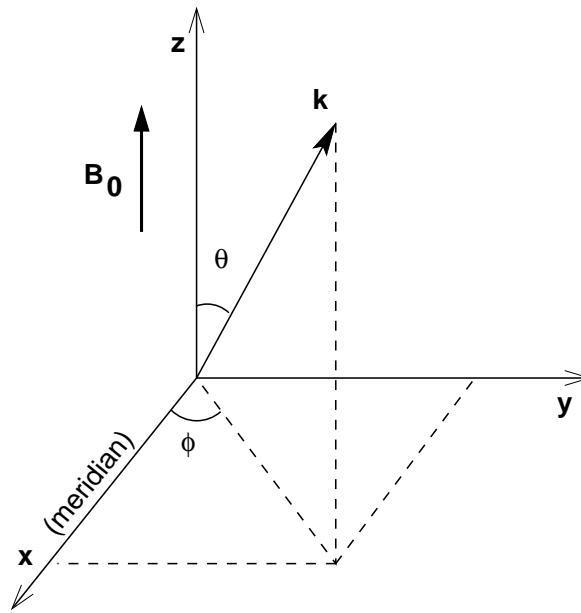


FIG. 2.1 – Système de coordonnées. Le champ magnétique B_0 et le méridien ont été indiqués afin de montrer dans quel repère on considère les mesures d'une expérience dans la magnétosphère

Nous allons écrire le vecteur $k = \frac{K}{\|K\|}$ en fonction de l'angle polaire θ et de l'angle azimutal ϕ (voir la figure 2.1); on a alors

$$k = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}.$$

Nous définissons ensuite le vecteur e_{\parallel} en ajoutant $\frac{\pi}{2}$ à θ dans l'expression précédente de sorte que $\langle e_{\parallel}, k \rangle_{\mathbb{R}^3} = 0$ et $\|e_{\parallel}\|_{\mathbb{R}^3} = 1$, on obtient

$$e_{\parallel} = \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ -\sin \theta \end{pmatrix};$$

nous pouvons maintenant définir un troisième vecteur e_{\perp} orthogonal aux deux précédents

$$e_{\perp} = \begin{pmatrix} \sin \phi \\ -\cos \phi \\ 0 \end{pmatrix}.$$

Les vecteurs $\{e_{\parallel}, e_{\perp}, k\}$ forment donc une base orthonormée directe de \mathbb{R}^3 que nous noterons B . Cette base va nous servir à écrire les expressions du champ électrique E et du champ magnétique H dans la base \mathcal{B} en fonction des angles polaire et azimutal. Notons P la matrice de passage de la base B à la base \mathcal{B}

$$P = \begin{pmatrix} \cos \theta \cos \phi & \sin \phi & \sin \theta \cos \phi \\ \cos \theta \sin \phi & -\cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}.$$

Nous appellerons plan d'incidence Π le plan contenant les vecteur k et e_{\parallel} , i.e

$$\Pi = \text{Vect}(\{k, e_{\parallel}\});$$

le front d'onde Φ est le plan perpendiculaire au vecteur k , i.e

$$\Phi = \text{Vect}(\{e_{\parallel}, e_{\perp}\});$$

les champs E et H varient dans le plan Φ .

2.3.2 Expression du champ électrique E

Comme $\langle E(t), k \rangle_{\mathbb{R}^3} = 0$ le champ électrique E s'écrit dans la base B sous la forme

$$E_B(t) = \begin{pmatrix} E_{\parallel}(t) \\ E_{\perp}(t) \\ 0 \end{pmatrix}$$

avec

$$\begin{aligned} E_{\perp}(t) &= \text{Re} [A_{\perp} \exp(i\omega t)] \\ E_{\parallel}(t) &= \text{Re} [A_{\parallel} \exp(i\omega t)] \end{aligned}$$

où $A_{\perp} = \langle A, e_{\perp} \rangle_{\mathbb{C}^3}$ et $A_{\parallel} = \langle A, e_{\parallel} \rangle_{\mathbb{C}^3}$. Définissons maintenant le paramètre de polarisation (en supposant que $A_{\parallel} \neq 0$) de l'onde p par

$$p = \frac{A_{\perp}}{A_{\parallel}} = |p| \exp(i\beta)$$

Nous allons maintenant exprimer E_B dans la nouvelle base grâce à la matrice de passage P , on obtient

$$\begin{aligned} E_x(t) &= E_{\perp}(t) \sin(\phi) + E_{\parallel}(t) \cos(\theta) \cos(\phi) \\ E_y(t) &= -E_{\perp}(t) \cos(\phi) + E_{\parallel}(t) \cos(\theta) \sin(\phi) \\ E_z(t) &= -E_{\parallel}(t) \sin(\theta) \end{aligned}$$

ce qui se traduit dans l'écriture complexe par

$$\begin{aligned}\mathcal{E}_x(t) &= \mathcal{E}_{\parallel}(t) (p \sin(\phi) + \cos(\theta) \cos(\phi)) \\ \mathcal{E}_y(t) &= \mathcal{E}_{\parallel}(t) (-p \cos(\phi) + \cos(\theta) \sin(\phi)) \\ \mathcal{E}_z(t) &= -\mathcal{E}_{\parallel}(t) \sin(\theta)\end{aligned}$$

2.3.3 Expression du champ magnétique H

Nous pouvons faire le même calcul pour le champ magnétique H ; on a

$$H_B(t) = \begin{pmatrix} H_{\parallel}(t) \\ H_{\perp}(t) \\ 0 \end{pmatrix}$$

avec

$$\begin{aligned}H_{\perp}(t) &= \operatorname{Re} [B_{\perp} \exp(i\omega t)] \\ H_{\parallel}(t) &= \operatorname{Re} [B_{\parallel} \exp(i\omega t)]\end{aligned}$$

où $B_{\perp} = \langle B, e_{\perp} \rangle_{\mathbb{C}^3}$ et $B_{\parallel} = \langle B, e_{\parallel} \rangle_{\mathbb{C}^3}$.

D'après les équations de Maxwell on a $E_{\parallel} H_{\parallel} + E_{\perp} H_{\perp} = 0$, d'où

$$\begin{aligned}H_{\parallel} &= -cE_{\perp} \\ H_{\perp} &= cE_{\parallel}\end{aligned}\tag{2.10}$$

où $c \in \mathbb{R}$ est une constante. Comme $\{k, E_B, H_B\}$ forment un trièdre direct on en déduit que $c \geq 0$; les équations de Maxwell donnent $c = \frac{1}{Z_0}$ où $Z_0 \in \mathbb{R}^{*+}$ est appelée l'impédance du vide. Finalement

$$\begin{aligned}H_x(t) &= cE_{\parallel}(t) \sin(\phi) - cE_{\perp}(t) \cos(\theta) \cos(\phi) \\ H_y(t) &= -cE_{\parallel}(t) \cos(\phi) - cE_{\perp}(t) \cos(\theta) \sin(\phi) \\ H_z(t) &= -cH_{\perp}(t) \sin(\theta)\end{aligned}$$

Ce qui s'écrit en notation complexe

$$\begin{aligned}\mathcal{H}_x &= c\mathcal{E}_{\parallel} (\sin(\phi) - p \cos(\theta) \cos(\phi)) \\ \mathcal{H}_y &= -c\mathcal{E}_{\parallel} (\cos(\phi) + p \cos(\theta) \sin(\phi)) \\ \mathcal{H}_z &= c\mathcal{E}_{\parallel} p \sin(\theta)\end{aligned}$$

2.3.4 Le vecteur champ électrique généralisé ε

Nous allons construire le vecteur champ électrique généralisé ε de l'onde [24, 50], dont les six composantes sont

$$\varepsilon_1 = \mathcal{E}_x; \varepsilon_2 = \mathcal{E}_y; \varepsilon_3 = \mathcal{E}_z; \varepsilon_4 = Z_0 \mathcal{H}_x; \varepsilon_5 = Z_0 \mathcal{H}_y; \varepsilon_6 = Z_0 \mathcal{H}_z;$$

d'après ce qui précède on a

$$\begin{aligned}
\varepsilon_1 &= \mathcal{E}_{//} (p \sin(\phi) + \cos(\theta) \cos(\phi)) \\
\varepsilon_2 &= \mathcal{E}_{//} (-p \cos(\phi) + \cos(\theta) \sin(\phi)) \\
\varepsilon_3 &= -\mathcal{E}_{//} \sin(\theta) \\
\varepsilon_4 &= \mathcal{E}_{//} (-p \cos(\phi) \cos(\theta) + \sin(\phi)) \\
\varepsilon_5 &= -\mathcal{E}_{//} (p \sin(\phi) \cos(\theta) + \cos(\phi)) \\
\varepsilon_6 &= \mathcal{E}_{//} (|p| \exp(i\beta) \sin(\theta))
\end{aligned}$$

Ce vecteur ε nous permet de définir les 36 noyaux d'intégration $(a_{ij})_{1 \leq i, j \leq 6}$ par la formule

$$\forall 1 \leq i, j \leq 6, \quad a_{ij} = \frac{\varepsilon_i \varepsilon_j^*}{\rho}$$

où ρ est une constante de normalisation. Nous avons donc établi l'expression des noyaux d'intégration du vide. Ce sont ces noyaux que nous allons utiliser pour les tests numériques à la Section 5.2, dans le cas du vide.

Nous ne décrivons pas plus le modèle physique. Dans le chapitre suivant nous allons nous concentrer sur la méthode à utiliser pour résoudre le problème inverse défini par l'équation (1.1).

Chapitre 3

Problème inverse et régularisation entropique

Dans le chapitre précédent, nous avons étudié le concept de fonction de distribution des ondes (FDO) et nous avons vu qu'il était relié aux mesures du champ électromagnétique par l'équation (1.1). Notre but, maintenant, est de déterminer la FDO F à partir de la matrice V et grâce à l'équation (1.1). C'est ce qu'on appelle un problème inverse. Typiquement, quand on a un opérateur $K : X \rightarrow Y$, où X et Y sont des espaces vectoriels normés, on appelle

- problème direct : connaissant $x \in X$ évaluer la valeur Kx .
- problème inverse : connaissant $y \in Y$ résoudre $Kx = y$.

Dans le cas de la détermination de la FDO à partir des mesures du champ, l'opérateur intégral qui relie la FDO aux mesures n'est pas bijectif. Il n'y a donc pas unicité de la FDO pour une valeur donnée des mesures : c'est un problème mal posé et même hautement indéterminé.

Avant de décrire la méthode de régularisation entropique que nous avons développée, faisons quelques rappels sur les problèmes inverses et la théorie de la régularisation.

3.1 Problème inverse mal posé

Définition 4 (Hadamard) Soient X et Y des espaces vectoriels normés et $K : X \rightarrow Y$ un opérateur. Le problème inverse $Kx = y$ est bien posé au sens de Hadamard si

1. **Existence** : Pour tout $y \in Y$ il existe $x \in X$ tel que $Kx = y$.
2. **Unicité** : Pour tout $y \in Y$, il y a au plus une solution $x \in X$.
3. **Stabilité** : La solution x dépend continûment de la donnée y .

Si au moins une de ces trois conditions n'est pas vérifiée, alors le problème est dit mal posé.

Le choix des espaces de départ et d'arrivée X et Y est bien sûr très important dans cette définition. La stabilité est une condition primordiale. En effet, s'il y a un problème de stabilité, le calcul numérique de la solution peut devenir impossible à cause des erreurs de mesures ou d'arrondis. Longtemps, on a pensé que l'étude des problèmes mal posés n'avait pas d'intérêt et que tout problème "intéressant" devait vérifier les trois conditions ci-dessus. Ce point de

vue est trop réducteur comme le montrent les exemples qui vont suivre. Les méthodes de régularisation que nous étudierons dans la section 3.2, sont les plus employées pour résoudre ces problèmes mal-posés.

3.1.1 Exemples

Systeme linéaire

On considère une matrice A sur \mathbb{R} qui a $p \in \mathbb{N}^*$ lignes et $n > p$ colonnes. Le problème du calcul de $x \in \mathbb{R}^n$ à partir de $y \in \mathbb{R}^p$, en résolvant l'équation $Ax = y$, est un problème mal posé. Premièrement, il n'est pas du tout sûr qu'il existe une solution à ce problème : pour garantir l'existence il faut supposer que $y \in \text{Im}(A)$. Ensuite, la solution n'est pas unique puisque l'application linéaire associée à la matrice A n'est pas une bijection. Dans ce cas, il y a une infinité de solutions puisqu'il suffit de rajouter un élément de $\text{Ker}(A)$ à une solution pour en obtenir une autre.

Dérivation Numérique

Le problème du calcul numérique de la dérivée d'une fonction que l'on mesure est un problème mal posé. Pour voir cela on considère l'espace de Banach X des fonctions C^∞ définies sur $[0, 1]$ à valeurs dans \mathbb{R} muni de la norme de la convergence uniforme. On définit l'opérateur linéaire $I : X \rightarrow X$ par

$$If(t) = \int_0^t f(s) ds, \quad \forall f \in X.$$

Le problème direct consiste donc à calculer la primitive de $f \in X$ qui vaut 0 en 0, ce problème est clairement bien posé. Pour résoudre le problème inverse il faut, connaissant $F \in X$ calculer sa dérivée $f \in X$. Le problème inverse est mal posé. En effet, nous savons qu'il existe une dérivée unique de $F \in X$ mais l'opérateur de dérivation n'est pas continu de X dans lui-même. La Figure 3.1 permet d'observer cette instabilité, nous avons calculé la dérivée de la fonction $t \mapsto \sin(3t) + p(t)$ où la fonction p est une petite perturbation (un bruit blanc). On observe une grande déviation entre la dérivée de la fonction F et celle de la fonction perturbée.

La Transformée de Radon

Un autre exemple de problème inverse classique est celui de la transformée de Radon. Nous allons la définir dans \mathbb{R}^2 . Notons que dans le cas de \mathbb{R}^2 la transformée de Radon s'appelle aussi transformée "X-ray". On considère une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ à un support compact $\Omega \subset \mathbb{R}^2$. Pour un angle $\theta \in [0, 2\pi[$ et $s > 0$ on définit la transformée de Radon Rf de f par

$$Rf(\theta, s) = \int_{\{x \in \mathbb{R}^2 | \langle x, \nu_\theta \rangle = s\}} f(x) dx \quad \text{où} \quad \nu_\theta = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

Ce qui correspond à intégrer la fonction f sur toutes les droites δ définies par l'angle θ et le nombre $s > 0$ (voir la figure 3.2).

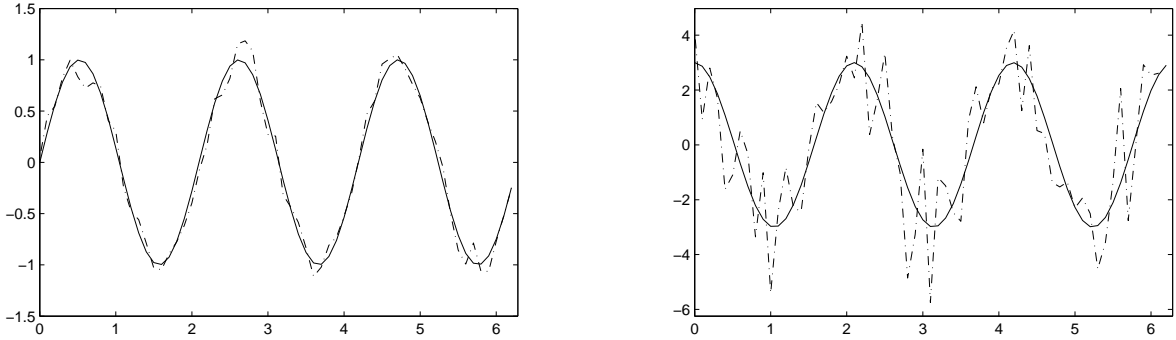


FIG. 3.1 – La dérivation est un problème inverse mal posé au sens de la stabilité. A gauche : la fonction $f : t \mapsto \sin(3t)$ et la même fonction perturbée par un bruit blanc. A droite : la dérivée de f et son évaluation numérique.

La transformée de Radon est utilisée en imagerie médicale. Le problème inverse est de déterminer la fonction f à partir de sa transformée de Radon Rf . C'est Radon [45] en 1917, qui fut le premier à donner une formule d'inversion explicite. Cette formule d'inversion s'écrit

$$f(x) = \int_0^{2\pi} \int_{\mathbb{R}^+} \frac{\partial_s Rf(\theta, s)}{\langle x, \nu_\theta \rangle_{\mathbb{R}^2} - s} ds d\theta,$$

où l'intégrale sur \mathbb{R}^+ est définie au sens des valeurs principales [54], et ∂_s est l'opérateur de dérivation par rapport à la variable s . Le problème est mal posé car la formule d'inversion fournit une solution instable. Comme dans l'exemple précédent, c'est l'opérateur de dérivation utilisé dans la formule d'inversion qui est la source de cette instabilité. Pour des détails complets sur la transformée de Radon nous renvoyons à [33].

3.2 Méthodes de Régularisation

Pour présenter le principe de la régularisation nous allons considérer un problème inverse $Kx = y$ où $K : X \rightarrow Y$ est un opérateur compact injectif. Le fait de choisir K injectif n'est pas très contraignant car on peut toujours restreindre l'espace X au complément orthogonal de $N(K)$, où N désigne le noyau. Les espaces X et Y sont des espaces de Hilbert. Nous supposons de plus que $y \in K(X)$, i.e. le problème inverse possède une solution unique. Ce qui rend le problème mal-posé est la non continuité de l'opérateur inverse.

Définition 5 Une famille d'opérateurs linéaires bornés $R_\mu : Y \rightarrow X$ est une "stratégie de régularisation" si

$$\forall x \in X, \quad \lim_{\mu \rightarrow 0} R_\mu Kx = x \quad (3.1)$$

i.e. l'opérateur $R_\mu K$ converge simplement vers l'identité.

Théorème 3 Soit R_μ une stratégie de régularisation pour l'opérateur $K : X \rightarrow Y$, où X est un espace de dimension infinie. Alors

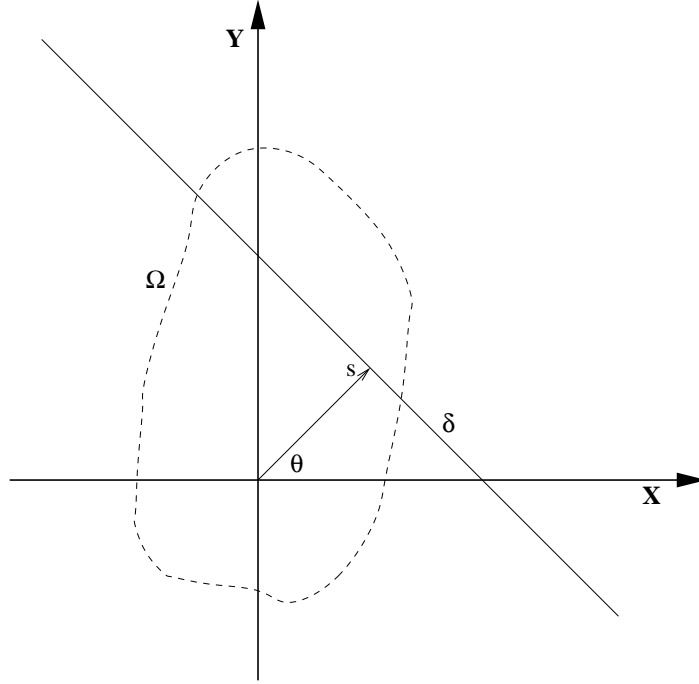


FIG. 3.2 – Transformée de Radon

1. Les opérateurs R_μ ne sont pas uniformément bornés : il existe une suite $\{\mu_j\}_{j \in \mathbb{N}} \subset \mathbb{R}^+$ telle que

$$\lim_j \|R_{\mu_j}\|_{\mathcal{L}(Y,X)} = \infty.$$

2. Il n'y a pas de convergence de $R_\mu K$ vers l'identité au sens de la norme d'opérateur.

Démonstration. Pour la démonstration de ce théorème voir [22] (page 24-25). Nous allons toutefois montrer le premier point. Supposons qu'il existe $c > 0$ tel que pour tout $\mu > 0$, $\|R_\mu\|_{\mathcal{L}(Y,X)} \leq c$. Donc $\|R_\mu y\|_X \leq c \|y\|_Y$. Comme de plus pour tout $y \in K(X)$, $R_\mu y \rightarrow K^{-1}y$ quand $\mu \rightarrow 0$ on en déduit que l'opérateur K^{-1} est borné. L'opérateur identité est donc compact puisque $I = K^{-1}K$ avec K compact et K^{-1} borné. C'est une contradiction avec le fait que X est de dimension infini par le Théorème de Riesz [11] (Th VI.5 p 92). \square

La donnée initiale $y \in Y$ n'est jamais connue exactement : il y a toujours un bruit qui vient la perturber. Notons y^δ la donnée perturbée où le nombre $\delta > 0$ est le niveau de bruit, i.e.

$$\|y - y^\delta\|_Y \leq \delta$$

Notons $x^{\mu,\delta} := R_\mu y^\delta$ l'approximation de la solution du problème inverse $Kx = y$ obtenue avec l'opérateur de régularisation et la donnée perturbée. En utilisant l'inégalité triangulaire sur $\|x^{\mu,\delta} - x\|_X$ on obtient

$$\|x^{\mu,\delta} - x\|_X \leq \delta \|R_\mu\|_{\mathcal{L}(Y,X)} + \|R_\mu Kx - x\|_X. \quad (3.2)$$

Le premier terme de droite de l'équation (3.2) représente la majoration de l'erreur due au niveau de bruit. Par le Théorème 3, nous avons vu que $\|R_\mu\| \rightarrow \infty$ quand $\mu \rightarrow 0$. Il ne faut donc pas choisir μ trop petit sinon l'erreur peut devenir très grande. Par contre le second terme de droite de (3.2) tend vers 0 quand μ tend vers 0 par définition de R_μ .

Nous allons faire tendre le niveau de bruit δ vers 0 et nous allons choisir une stratégie de régularisation de manière à ne pas commettre une trop grande erreur sur la vraie solution x .

Définition 6 Une stratégie de régularisation $\delta \mapsto \mu(\delta)$ est admissible si pour tout $x \in X$

$$\lim_{\delta \rightarrow 0} \mu(\delta) = 0 \text{ et } \lim_{\delta \rightarrow 0} \sup_{y^\delta \in Y} \left\{ \left\| R_{\mu(\delta)} y^\delta - x \right\|_X ; \text{ tel que } \left\| Kx - y^\delta \right\|_Y \leq \delta \right\} = 0 \quad (3.3)$$

Plusieurs exemples de stratégies de régularisation admissibles se trouvent dans [22]. Nous allons présenter dans la section suivante la méthode de Tikhonov pour approcher les problèmes inverses.

3.2.1 Méthode de Tikhonov

Le principe de la méthode de Tikhonov pour résoudre le problème inverse mal posé $Kx = y$ est de choisir comme solution l'élément $x^\mu \in X$ qui minimise la quantité

$$\|Kx - y\|_Y^2 + \mu \|x\|_X^2 \quad (3.4)$$

L'existence et l'unicité du minimum est immédiate par coercivité et stricte convexité de $x \mapsto \|x\|_X^2$. Le paramètre μ est appelé le paramètre de régularisation. Pour que l'élément x^μ qui réalise le minimum ait une faible erreur avec la donnée y , ce paramètre doit être choisi assez petit. Il doit également être choisi assez grand pour que la stricte convexité du terme $\|x\|_X$ corrige l'instabilité du problème posé. On appelle $x \mapsto \|x\|_X^2$ la fonctionnelle régularisante. On a le résultat suivant [22, 64, 65]

Théorème 4 Soit $\mu > 0$ et $K : X \rightarrow Y$ un opérateur linéaire borné de l'espace de Hilbert X vers l'espace de Hilbert Y . Alors la fonctionnelle de Tikhonov admet un unique minimum en $x^\mu \in X$. L'élément x^μ est la solution de l'équation normale

$$\alpha x^\mu + K^* K x^\mu = K^* y \quad (3.5)$$

Grâce à l'équation (3.5) nous pouvons définir l'opérateur de régularisation de Tikhonov par

$$R_\mu = (\mu I + K^* K)^{-1} K^* : Y \rightarrow X. \quad (3.6)$$

Il reste à démontrer que cet opérateur est bien un opérateur de régularisation et sous quelles conditions le choix de μ en fonction du niveau de bruit δ est admissible. C'est l'objet du théorème suivant.

Théorème 5 Soient $K : X \rightarrow Y$ un opérateur linéaire compact et $\mu > 0$. L'opérateur $\mu I + K^* K$ est inversible et l'opérateur $R_\mu : Y \rightarrow X$ défini par (3.6) est une stratégie de régularisation avec $\|R_\mu\|_{\mathcal{L}(Y,X)} \leq \frac{1}{2\sqrt{\mu}}$. Tout choix de $\mu(\delta) \rightarrow 0$ avec $\delta^2 \mu(\delta) \rightarrow 0$ est admissible.

Pour la démonstration de ce théorème et des résultats sur la vitesse de convergence quand $\delta \rightarrow 0$, voir [64, 22]

La méthode de régularisation de Tikhonov, i.e. la minimisation globale de la fonctionnelle (3.4), est en fait équivalente à un autre problème de minimisation avec contraintes. Ce résultat important permet de comprendre la méthode sous deux angles différents. Nous allons retrouver une propriété similaire dans la section suivante pour le cas de la régularisation entropique.

Théorème 6 *Soit x^μ la solution du problème*

$$\min_{x \in X} \|Kx - y\|_Y^2 + \mu \|x\|_X^2. \quad (3.7)$$

Posons $\epsilon = \|Kx^\mu - y\|_Y$ et $D_\epsilon = \{x \in X \mid \|Kx - y\|_Y \leq \epsilon\}$. Alors x^μ est aussi solution du problème

$$\min_{x \in D_\epsilon} \|x\|_X^2 \quad (3.8)$$

Réciproquement si x_ϵ est la solution de (3.8) et si $0 \notin D_\epsilon$, alors il existe $\mu_\epsilon > 0$ tel que x_ϵ , soit la solution de (3.7).

Démonstration. (\Rightarrow) Soit $x_\epsilon \in X$ l'unique solution de (3.8) (Il y a clairement existence et unicité de la solution) nous allons montrer que x_ϵ est solution de (3.7). Comme x^μ est solution de (3.7), on a

$$\|Kx^\mu - y\|_Y^2 + \mu \|x^\mu\|_X^2 \leq \|Kx_\epsilon - y\|_Y^2 + \mu \|x_\epsilon\|_X^2, \quad (3.9)$$

comme de plus $x^\mu \in D_\epsilon$ et x_ϵ est la solution de (3.8), on a

$$\|Kx_\epsilon - y\|_Y^2 + \mu \|x_\epsilon\|_X^2 \leq \|Kx_\epsilon - y\|_Y^2 + \mu \|x^\mu\|_X^2.$$

Il s'ensuit

$$\epsilon^2 \leq \|Kx_\epsilon - y\|_Y^2 \leq \epsilon^2,$$

et donc

$$\|Kx_\epsilon - y\|_Y = \|Kx^\mu - y\|_Y = \epsilon.$$

Finalement, puisque $x^\mu \in D_\epsilon$, on a

$$\|x_\epsilon\|_X \leq \|x^\mu\|_X \Rightarrow \|Kx_\epsilon - y\|_Y^2 + \mu \|x_\epsilon\|_X^2 \leq \|Kx^\mu - y\|_Y^2 + \mu \|x^\mu\|_X^2,$$

et on en déduit $x_\epsilon = x^\mu$ par unicité du minimum.

(\Leftarrow) Soit $\epsilon > 0$, comme x^μ est solution de (3.7), on a

$$\begin{aligned} \|Kx^\mu - y\|_Y^2 + \mu \|x^\mu\|_X^2 &\leq \|Kx_\epsilon - y\|_Y^2 + \mu \|x_\epsilon\|_X^2 \\ &\leq \epsilon^2 + \mu \|x_\epsilon\|_X^2, \end{aligned}$$

en passant à la limite lorsque $\mu \rightarrow 0$ on obtient

$$\forall \epsilon > 0, \lim_{\mu \rightarrow 0} \|Kx^\mu - y\|_Y^2 \leq \epsilon^2,$$

donc

$$\lim_{\mu \rightarrow 0} \|Kx^\mu - y\|_Y^2 = 0.$$

Soit $\epsilon > 0$, il existe donc $\mu_0 > 0$ tel que pour tout $\mu \leq \mu_0$ on a $x^\mu \in D_\epsilon$. Maintenant, s'il existe un $\mu_\epsilon > 0$ tel que

$$\|x_\epsilon\|_X = \|x^{\mu_\epsilon}\|_X,$$

d'après l'équation (3.9)

$$\|Kx^{\mu_\epsilon} - y\|_Y^2 \leq \|Kx_\epsilon - y\|_Y^2,$$

et on a la conclusion par unicité du minimum. Supposons donc que :

$$\forall \mu > 0 \quad \|x_\epsilon\| = r_\epsilon < \|x^\mu\|_X.$$

D'autre part, l'hypothèse $0 \notin D_\epsilon$ empêche 0 d'être solution, donc $r_\epsilon > 0$. La fonction $\mu \mapsto \|x^\mu\|_X$ est continue de

$$]0, +\infty[\rightarrow]0, m[$$

avec $m > \|x^{\mu_0}\|$, puisque $x^\mu : \mu \mapsto (K^*K + \mu I)^{-1}K^*y$ est continue. En utilisant le théorème des valeurs intermédiaires, on en déduit l'existence de $\mu_\epsilon > 0$ tel que $\|x^{\mu_\epsilon}\|_X = r_\epsilon$. ce qui termine la démonstration. \square

La méthode de régularisation de Tikhonov est une des méthodes les plus employées pour résoudre les problèmes mal posés. Par exemple elle est utilisée avec succès pour inverser des matrices mal conditionnées [41, 65]. Cependant le choix de la norme au carré comme opérateur de régularisation n'est pas une recette miracle et il faut sélectionner l'opérateur de régularisation en fonction de ce que l'on cherche à obtenir. Dans les exemples numériques qui vont suivre nous allons voir que le choix de l'opérateur de régularisation conduit à des solutions radicalement différentes.

3.2.2 Exemples numériques pour le calcul d'une dérivée

Nous allons appliquer la méthode de Tikhonov pour résoudre le problème de la dérivation numérique présentée dans la Section 3.1.1. Pour cette résolution nous allons nous placer dans l'espace $L^2([0, 2\pi], \mathbb{R})$ et nous calculons la dérivée en résolvant le problème

$$\min_{f \in L^2([0, 2\pi])} \|F_b - If\|_{L^2([0, 2\pi])}^2 + \mu \|f\|_{L^2([0, 2\pi])}^2$$

où $F_b : [0, 2\pi] \rightarrow \mathbb{R}$ est la fonction dont nous voulons calculer la dérivée. On suppose que cette fonction est une version perturbée de la fonction F et que l'erreur (le niveau de bruit) est $\|F - F_b\| = \delta$. L'opérateur d'intégration I est le même que celui de la section 3.1.1.

Sur la Figure 3.3.a) on distingue la fonction F et sa version perturbée F_b qui nous ont servi pour l'exemple. En utilisant la méthode ci dessus pour calculer f nous obtenons la fonction tracée sur la Figure 3.3.b), le paramètre de régularisation ayant été choisi convenablement par rapport au niveau de bruit. On remarque que cette solution n'est pas satisfaisante car la fonction obtenue ne ressemble pas beaucoup à un cosinus comme elle le devrait. Cela s'explique par le fait que la minimisation de la norme de f dans $L^2([0, 2\pi])$ n'oblige pas

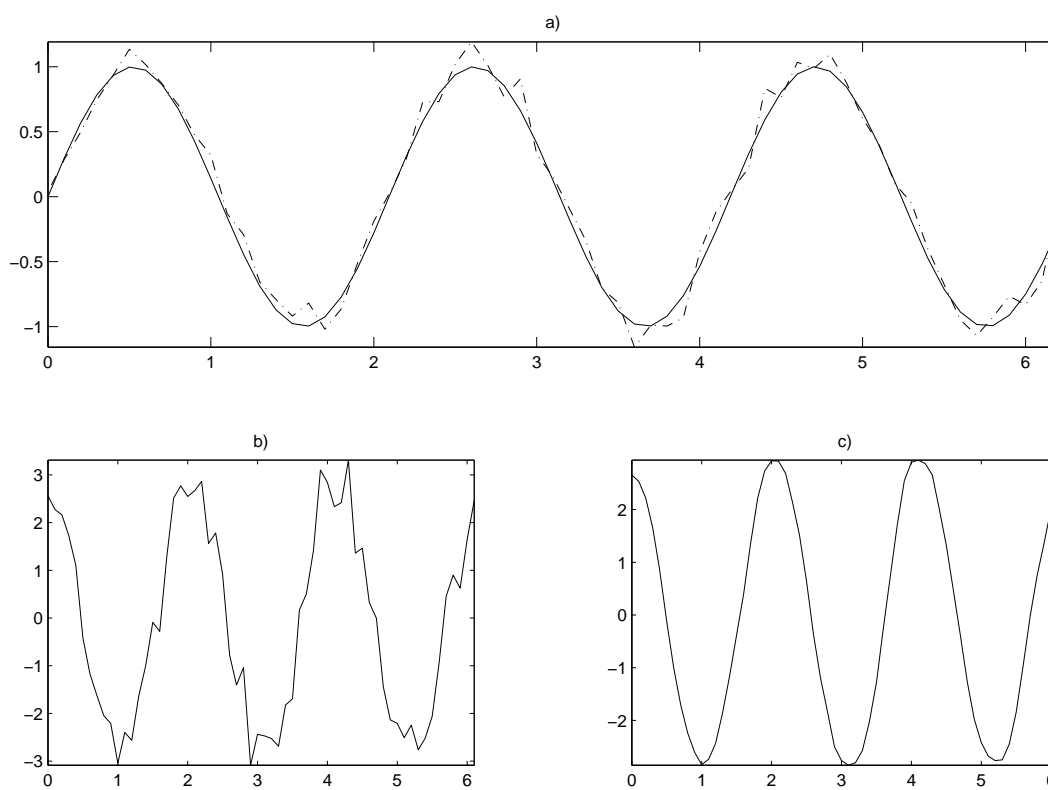


FIG. 3.3 – a) Graphe de la fonction initiale (trait continu) et de la fonction bruitée. b) Dérivée calculée grâce à la méthode de Tikhonov. c) Dérivée calculée par régularisation, l'opérateur de régularisation étant la norme de l'espace de Sobolev H_0^1 .

la fonction obtenue à être régulière. Une alternative pour apporter plus de régularité est d'utiliser la méthode de Tikhonov dans un espace de Sobolev, plutôt que dans $L^2([0, 2\pi])$. Calculons donc la solution du problème

$$\min_{f \in H_0^1([0, 2\pi])} \|F_b - If\|_{L^2([0, 2\pi])}^2 + \mu \|f\|_{H^1([0, 2\pi])}^2$$

La Figure 3.3.c) montre la dérivée calculée de cette manière. La fonction f est ici bien plus régulière que celle trouvée précédemment, de plus elle ressemble beaucoup à une fonction cosinus ce qui était le résultat souhaité. Cet exemple montre clairement que pour résoudre un problème inverse, il faut rajouter des hypothèses de manière à obtenir un résultat conforme à nos attentes.

3.3 Régularisation entropique

Notons E un sous ensemble compact de \mathbb{R}^n et σ une mesure sur E . Nous supposons pour simplifier que $\sigma(E) = 1$. Dans cette section nous allons étudier la régularisation entropique

dans l'espace des fonctions de E dans \mathbb{R} , intégrables pour la mesure σ , i.e. $L^1(E, \mathbb{R}, \sigma)$ comme dans [1, 16]. Toutefois, nous allons aussi regarder le problème dans $L^2(E, \mathbb{R}, \sigma)$ pour les applications aux calculs de la FDO.

L'entropie que nous allons considérer est celle de Boltzmann-Shannon [56, 7], nous allons la décrire ainsi que quelques unes de ses propriétés dans la sous section suivante.

3.3.1 L'entropie

Définition 7 Soit $f, g \in L^1(E, \mathbb{R}, \sigma)$ telles que f et g soient deux densités de probabilité pour la mesure σ . Si $f d\sigma \ll g d\sigma$ on définit le contenu d'information de f par rapport à g par

$$I(f, g) = \int_E f \ln \frac{f}{g} d\sigma$$

sinon on pose $I(f, g) = \infty$.

La notation $f \ll g$ signifie que la mesure de densité f est absolument continue par rapport à la mesure de densité g . La justification du fait que $I(f, g)$ représente l'information contenue dans f par rapport à g se trouve dans l'article de *Shore et Johnson* [57]. Les auteurs démontrent que c'est l'unique façon de définir l'information par une approche axiomatique.

On appelle densité non-informative la densité de probabilité g qui n'apporte aucune information sur le système physique. On considère que cette densité est connue grâce au modèle. Dans ce cas, nous définissons la négentropie $H(f)$ d'une densité de probabilité f dans le système par $H(f) = I(f, g)$. Dans le cas de la détermination de la FDO, nous supposons que le système est isotrope, donc la densité non informative est donnée par la densité uniforme sur la sphère. Nous posons donc

$$H(f) = \int \phi \circ f d\sigma$$

où la fonction $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}$ est définie par

$$\phi(x) = \begin{cases} x \ln(x) & \text{if } x > 0 \\ 0 & \text{si } x = 0 \\ +\infty & \text{sinon} \end{cases}$$

Ainsi définie la fonctionnelle négentropie vérifie les propriétés suivantes.

Proposition 2 La fonctionnelle négentropie $H : L^2(E, \mathbb{R}, \sigma) \rightarrow \mathbb{R}$ est semi-continue inférieurement (s.c.i.) et strictement convexe. Si $f \in L^2(E, \mathbb{R}, \sigma)$, $f \geq 0$ alors

$$e^{-1} \leq H(f) \leq \|f\|_{L^2} < \infty.$$

Enfin si $f \in L^2(E, \mathbb{R}^+, \sigma)$ et $\|f\|_{L^1} = 1$ alors $H(f) \geq 0$.

Démonstration. Pour la démonstration complète de cette proposition voir le Lemme 1 page 107. La semi-continuité inférieure de H découle de l'inégalité de Fatou. Si $\{f_n\}_{n \in \mathbb{N}}$ est une

suite de $L^2(E, \mathbb{R}^+, \sigma)$ et $f_n \rightarrow f$ p.p. alors $f_n \ln f_n \rightarrow f \ln f$ p.p. par continuité la fonction ϕ sur \mathbb{R}^+ . Par le Lemme de Fatou

$$H(f) \leq \liminf H(f_n).$$

L'inégalité $H(f) \geq 0$ si f est une densité de probabilité, est une conséquence de l'inégalité de Jensen. \square

L'entropie a été largement utilisée dans de nombreux domaines, pour résoudre des problèmes mal-posés, par exemple dans le domaine de la reconstruction d'image [35], en astronomie ou pour l'estimation spectrale [58, 13]. *Borwein et Lewis* [7, 8, 9] ont étudié la dualité pour le problème du maximum d'entropie ainsi que l'utilisation de l'entropie pour déterminer une fonction par ses moments. Dans la sous-section suivante nous allons présenter la méthode que nous avons développée pour le calcul des directions de propagation d'une onde électromagnétique.

3.3.2 Principe de la méthode

Le principe de la régularisation entropique est le même que celui de la méthode de Tikhonov, la fonctionnelle régularisante étant à présent la négentropie. Plus précisément, on détermine la fonction f positive qui minimise

$$J_\mu := \|y - Kx\|^2 + \mu H(x) \quad (3.10)$$

Cette méthode est utilisée dans le cas où l'on recherche une fonction positive car l'entropie $-H$ se comporte comme une fonction barrière, ce qui assure la positivité de la solution.

Nous allons présenter quelques résultats dûs à *Amato et Hughes* [1] sur la régularisation entropique. Ces résultats sont obtenus pour l'espace L^1 , alors que nous nous sommes placés dans L^2 (cf section 3.4). Soit l'espace $X = L^1([0, 1], \mathbb{R}^+)$ et D l'ensemble où l'entropie est bien définie

$$D = \{f \in X \mid H(f) < \infty\}.$$

Nous allons remplacer la négentropie H par

$$E = H + \exp(-1)$$

afin d'obtenir une fonctionnelle à valeurs strictement positives. Soit G un espace de Hilbert et $K : D \rightarrow G$ un opérateur linéaire.

Définition 8 On définit l'opérateur $R^1 : G \times]0, \infty[\rightarrow D$ par

$$E(R^1(u, \delta)) = \inf\{E(f) \mid f \in D, \|Kf - u\| \leq \delta\}. \quad (3.11)$$

Définition 9 On définit l'opérateur $R^2 : G \times]0, \infty[\rightarrow D$ par

$$J_\mu(R^2(u, \delta)) = \inf_{f \in D} J_\mu(f). \quad (3.12)$$

Si pour un certain $u \in G$ et $\delta > 0$, $R^1(u, \delta)$ ou $R^2(u, \delta)$ existent, alors ils sont uniques par stricte convexité de la négentropie. Nous n'allons pas examiner plus en détail la question de l'existence (se référer à [15]). En revanche, de manière analogue au Théorème 6, nous allons voir que les opérateurs R^1 et R^2 sont les mêmes et qu'ils définissent bien une stratégie de régularisation.

Définition 10 *On appelle solution de maximum d'entropie du problème inverse $Kx = y$, l'unique élément x^* qui vérifie*

$$E(x^*) = \inf\{ E(x) | \forall x \in X, Kx = y \}$$

Théorème 7 *Soit $y \in K(D)$ et soit $\{y_n\}$ une suite de G telle que pour tout $n \in \mathbb{N}$, $\|y - y_n\| < \delta_n$ avec $\lim \delta_n = 0$. Alors $R^1(y_n, \delta_n)$ converge vers la solution de maximum d'entropie de $Kx = y$.*

Ce théorème montre que R^1 est bien un opérateur de régularisation. Le théorème qui va suivre montre que les deux opérateurs R^1 et R^2 sont équivalents. On en déduit donc que l'opérateur R^2 est également un opérateur de régularisation. Pour la preuve de ces théorèmes consulter [1].

Théorème 8 *Soient la fonction $e(x) := \exp(-1)$, $u \in G$ et $\mu_0 > 0$ tel que*

$$\|Ae - u\|^2 > \|AR(u, \mu_0) - u\|^2 =: \delta_0. \quad (3.13)$$

Alors il existe une fonction monotone continue $\mu_u : [0, \delta_0] \rightarrow [0, \lambda_0]$ avec $\mu_u(0) = 0, \mu_u(\delta_0) = \lambda_0$ telle que

$$R^2(u, \mu_u(\delta)) = R^1(u, \delta).$$

Le problème de la régularisation entropique a également été étudié par *Engl et Landl* [16]. Dans cet article, ils présentent une méthode pour transformer le problème de régularisation entropique en un problème de régularisation classique. Pour ce faire, ils utilisent un opérateur non-linéaire \mathfrak{F} qui agit sur E de la manière suivante

$$E(\mathfrak{F}x) = \|x\|^2 + c.$$

Ainsi la substitution de x par $\mathfrak{F}x$ dans la fonctionnelle J_μ revient à la minimisation de la fonctionnelle de Tikhonov. Toutefois une non-linéarité apparaît par le biais de l'opérateur \mathfrak{F} .

3.4 Formulation du problème

Revenons au problème du calcul de la FDO par régularisation entropique. Pour cela nous allons nous placer dans l'espace $L^2(E, \mathbb{R}, \sigma)$. Ici E est la sphère unité de \mathbb{R}^3 , mais nous garderons la notation E puisque les résultats qui vont suivre sont valables pour des ensembles E plus généraux. Notons $V \in \mathbb{C}^n$ la donnée provenant des mesures du champ électromagnétique ; c'est un vecteur formé des composantes de la matrice spectrale estimée. Notons $F \in L^2(E, \mathbb{R}, \sigma)$ la fonction de distribution des ondes que nous souhaitons calculer et

$\psi : L^2(E, \mathbb{R}, \sigma) \rightarrow \mathbb{C}^n$ l'opérateur intégral qui relie la FDO aux données (équation (1.1)). Le problème à résoudre est

$$(\mathcal{P}_\mu) \begin{cases} \min \|V - \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F), \\ F \in L^2(E, \mathbb{R}^+, \sigma) \end{cases}$$

où $\mu > 0$ est le paramètre de régularisation. Il faut remarquer que le problème de la détermination d'une fonction F telle que $\psi[F] = y$ est un problème hautement indéterminé, car l'opérateur linéaire ϕ envoie l'espace de dimension infinie $L^2(E, \mathbb{R}, \sigma)$ dans \mathbb{C}^n .

Dans le cas physique, comme les mesures sont toujours bruitées, nous avons seulement besoin de calculer une solution pour une valeur assez grande du paramètre de régularisation μ . Le théorème suivant est à la base des méthodes numériques que nous avons utilisées.

Théorème 9 *Il existe $\mu > 0$ (assez grand) tel que le problème (\mathcal{P}_μ) a une solution unique F_μ qui s'écrit sous la forme*

$$F_\mu = \exp\left(-1 + \frac{2}{\mu}\psi^*(l)\right)$$

avec

$$l = V - \psi[F_\mu].$$

La démonstration de ce théorème se trouve dans la section A.4, mais nous présentons les grandes lignes de la démonstration. Tout d'abord, la fonction coût J_μ est s.c.i. et strictement convexe (cf. Lemme (2) page 108) par stricte convexité de H . Elle n'est cependant ni coercive ni différentiable sur l'ensemble des contraintes K . C'est pourquoi on relaxe les contraintes en prenant un sous ensemble borné de K sur lequel J_μ a de "bonnes propriétés". On en déduit une condition suffisante d'optimalité pour le problème (\mathcal{P}_μ)

$$-2\psi^*(V - \psi F_\mu) + \mu \nabla H(F_\mu) = 0. \quad (3.14)$$

Comme $\nabla H : L^2(E, \mathbb{R}, \sigma) \rightarrow L^2(E, \mathbb{R}, \sigma)$ a un inverse qui est la fonction $F \mapsto \exp(-1 + F)$, on en déduit que la condition d'optimalité se réécrit sous la forme d'un point fixe :

$$F = \exp\left(-1 + \frac{2}{\mu}\psi^*(V - \psi[F])\right). \quad (3.15)$$

Pour finir, on montre que pour μ assez grand, il existe $F \in L^2(E, \mathbb{R}, \sigma)$ vérifiant (3.14) à l'aide du théorème du point fixe de Banach appliqué à la suite

$$\begin{cases} F_0 \in L^2(E, \mathbb{R}, \sigma) \\ F_{k+1} = \exp(-1 + \frac{2}{\mu}\psi^*(V - \psi(F_k))) \end{cases}$$

(cf. Théorème 2 page 112).

La suite $(F_k)_{k \in \mathbb{N}}$ est une suite dans un espace de dimension infinie, elle est donc difficilement utilisable pour calculer la solution numériquement. En utilisant des techniques de dualité, on peut établir un algorithme en dimension finie. Soit F_μ la solution du problème (\mathcal{P}_μ) obtenue comme limite de la suite $(F_k)_{k \in \mathbb{N}}$, et posons $l = V - \psi[F_\mu]$. On remarque que

F_μ est entièrement définie par la valeur de $l \in \mathbb{C}^n$ puisque $F_\mu = \exp(-1 + \frac{2}{\mu}\psi^*(l))$. Ainsi l vérifie l'équation

$$l = V - \psi \left[\exp \left(-1 + \frac{2}{\mu} \psi^*(l) \right) \right]. \quad (3.16)$$

Grâce à cette équation nous pouvons écrire un algorithme en dimension finie pour calculer la solution régularisée du problème inverse.

3.4.1 Algorithme 1

Le Théorème 9 indique que l'on peut rechercher la solution du problème (\mathcal{P}_μ) grâce à un algorithme de point fixe dans l'espace \mathbb{C}^n . En fait, il y a des problèmes de convergence si l'on applique l'algorithme tel quel. On observe numériquement que l'algorithme converge seulement pour des valeurs de μ trop grandes par rapport à l'erreur obtenue avec la donnée initiale. Pour résoudre ces problèmes nous introduisons un paramètre $\tau \in [0, 1]$ qui permet de rendre la fonction que l'on itère, contractante. L'introduction de ce paramètre va entraîner la convergence de l'algorithme, mais le nombre d'itérations va en être grandement augmenté et le temps de calcul beaucoup plus long. Voici cet algorithme

1. Initialisation

Choisir $V \in \mathbb{C}^n, l_0 \in \mathbb{C}^n, \mu > 0, \epsilon > 0, \tau \in]0, 1]$.

2. Itération k

(a) Calculer

$$\gamma_\mu(l_{k-1}) := V - \psi \left[\exp \left(-1 + \frac{2}{\mu} \psi^*[l_{k-1}](\sigma) \right) \right].$$

(b) Calculer $l_k = l_{k-1} - \tau[l_{k-1} - \gamma_\mu(l_{k-1})]$.

3. critère d'arrêt

Si $|l_k - l_{k-1}| < \epsilon$, alors STOP, sinon $k := k + 1$ et retour à 2.

Proposition 3 *On pose $l_0 = 0$. Si le paramètre μ est choisi assez grand et τ assez petit, alors l'algorithme 1 converge : $l_k \rightarrow l^*$. La solution du problème de régularisation (\mathcal{P}_μ) est donnée par*

$$F_\mu = \exp \left(-1 + \frac{2}{\mu} \psi^*[l^*] \right)$$

Pour plus de détails sur les conditions de convergence et pour la démonstration de cette proposition on peut voir la Proposition 1 page 114.

La proposition précédente garantit la convergence de l'algorithme si le paramètre de régularisation est choisi assez grand. Cela est satisfaisant pour les applications numériques car la donnée $V \in \mathbb{C}^n$ est toujours bruitée et donc μ doit être assez grand pour obtenir une solution stable. D'un autre côté, on aimerait savoir si on peut calculer la solution du problème (\mathcal{P}_μ) pour toutes valeurs de μ avec l'algorithme 1 : c'est l'objet du corollaire suivant.

Corollaire 1 *Il existe $\mu > 0$ et $\tau \in]0, 1]$ tels que pour tout $\alpha > 0$, la suite $(\ell_k)_{k \in \mathbb{N}}$ définie par (3.17) converge vers $\ell^* \in \mathbb{C}^n$.*

$$\begin{cases} \ell_0 = 0 \\ \ell_{k+1} = (1 - \frac{\tau}{\alpha})\ell_k + \tau \gamma_\mu(\ell_k) \end{cases} \quad (3.17)$$

La fonction $F = \exp\left(-1 + \frac{2}{\mu}\psi^*[\ell^*]\right)$ est la solution du problème $(\mathcal{P}_{\frac{\mu}{\alpha}})$.

Nous pouvons donc calculer la solution du problème de régularisation pour toutes les valeurs du paramètre de régularisation. En pratique, cela se révèle difficile car le nombre d'itérations devient très grand lorsque l'on cherche à calculer la solution pour des petites valeurs de μ . La Figure 5.4 page 74 est une illustration de ce phénomène.

3.4.2 Second modèle

La solution de (\mathcal{P}_μ) n'est pas une approximation totalement satisfaisante des solutions du problème inverse. Le premier défaut est qu'elle ne dépend pas linéairement de la donnée initiale. Cette dépendance non-linéaire de la solution s'explique par le fait que l'entropie, i.e. la fonctionnelle régularisante du problème (\mathcal{P}_μ) , n'est pas linéaire. De plus, il ne s'agit pas réellement de l'entropie au sens de la définition (7) puisque la FDO n'est pas une densité de probabilité. Pour pallier ces inconvénients, on considère une version modifiée du problème de régularisation $(\tilde{\mathcal{P}}_\mu)$.

$$(\tilde{\mathcal{P}}_\mu) \left\{ \begin{array}{l} \min \|V - \alpha\psi[F]\|^2 + \mu H(F) \\ (\alpha, F) \in \mathbb{R}^+ \times L^2(S, \mathbb{R}^+, \sigma) \cap \{\|F\|_{L^1} = 1\} \end{array} \right.$$

Dans le problème $(\tilde{\mathcal{P}}_\mu)$, une paramétrisation de la FDO a été faite en posant $F = \alpha f$ avec $\alpha = \|F\|$ et $f = \frac{F}{\|F\|}$. Le paramètre $\alpha \in \mathbb{R}^+$ représente la puissance de l'onde et f est une densité de probabilité. Nous minimisons donc vraiment l'information de la FDO, au sens de la définition (7). De plus, $(\tilde{\mathcal{P}}_\mu)$ permet de résoudre partiellement le problème de la dépendance non linéaire de la FDO par rapport aux mesures du champ. En effet, si on multiplie la donnée V par λ dans le problème $(\tilde{\mathcal{P}}_\mu)$, nous voyons que cela revient à multiplier α par λ et μ par λ^2 .

Le problème $(\tilde{\mathcal{P}}_\mu)$ est beaucoup plus complexe que (\mathcal{P}_μ) . En effet la fonction coût n'est pas convexe à cause de la paramétrisation de la FDO. i.e. du terme bilinéaire $(\alpha, F) \mapsto \alpha F$. La Figure 3.4 est une illustration de la non-convexité de la fonction coût. Nous n'avons donc pas de condition d'optimalité pour ce problème. Toutefois on peut utiliser le premier algorithme pour α *a priori*, puis on détermine le meilleur α vis à vis de la solution obtenue via l'équation

$$\bar{\alpha} = \operatorname{argmin}_{\alpha \in \mathbb{R}} \|V - \alpha\psi[F]\|$$

ce qui est équivalent à

$$\bar{\alpha} = \frac{\operatorname{Re}(\langle V, g(l_{k-1}) \rangle_{\mathbb{C}^n})}{\langle g(l_{k-1}), g(l_{k-1}) \rangle_{\mathbb{C}^n}}, \quad \text{où } g(l_{k-1}) = \psi \left[\exp \left(-1 + \frac{2}{\mu} \psi^*[l_{k-1}](\sigma) \right) \right].$$

Nous pouvons maintenant écrire le second algorithme.

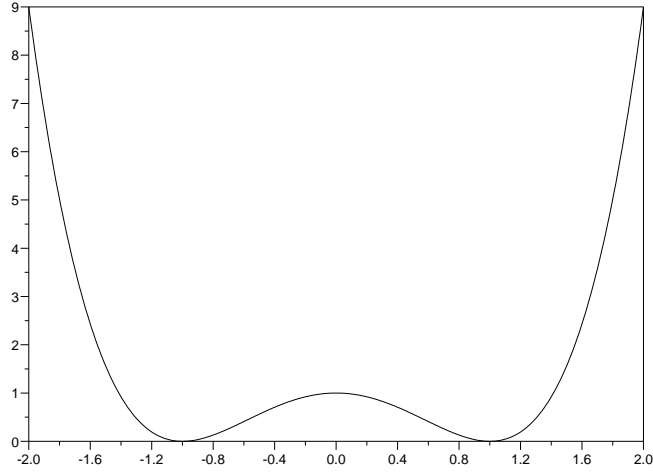


FIG. 3.4 – On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto (1 - xy)^2$. Le graphique ci-dessus représente la fonction $z \mapsto (1 - z^2)^2$ pour $z \in [-2, 2]$, c'est également le tracé de la fonction f pour $z = x = y$. On observe bien la non-convexité de la fonction f .

3.4.3 Algorithme 2

1. Initialisation

Choisir $V \in \mathbb{C}^n, l_0 \in \mathbb{C}^n, \mu > 0, \epsilon > 0$.

2. Itération k

(a) Calculer

$$g(l_{k-1}) = \psi \left[\exp \left(-1 + \frac{2}{\mu} \psi^*[l_{k-1}](\sigma) \right) \right].$$

(b) Calculer

$$\alpha_{k-1} = \frac{\operatorname{Re}(\langle V, g(l_{k-1}) \rangle_{\mathbb{C}^n})}{\langle g(l_{k-1}), g(l_{k-1}) \rangle_{\mathbb{C}^n}}.$$

(c) $l_k = V - \alpha_{k-1}g(l_{k-1})$.

3. Critère d'arrêt

Si $|l_k - l_{k-1}| < \epsilon$, alors STOP, sinon $k := k + 1$ et retour à 2.

Le théorème suivant permet de conclure à propos de la convergence de cet algorithme.

Théorème 10 *Il existe $\mu > 0$ assez grand et $\tau \in]0, 1[$ tels que l'algorithme 2 converge,*

c'est-à-dire $l_k \rightarrow l$ et $\alpha_k \rightarrow \alpha$. La limite vérifie $l = V - \alpha g_\mu(l)$ avec $\alpha > 0$ et on pose

$$F_l = \exp\left(-1 + \frac{2}{\mu}\psi^*[l]\right) \text{ et } f_l = \frac{F_l}{\|F_l\|_{L^1(E, \mathbb{R}, \sigma)}}.$$

Alors f_l est une densité de probabilité : c'est l'unique solution du problème

$$\begin{cases} \min \left\| \frac{V}{\alpha \|F_l\|_{L^1(E, \mathbb{R}, \sigma)}} - \psi(F) \right\|_{\mathbb{C}^n}^2 + \frac{\mu}{\alpha \|F_l\|_{L^1(E, \mathbb{R}, \sigma)}} \int_E \phi \circ F(\sigma) d\sigma \\ F \in \left\{ f \in \mathbb{H} \mid f \geq 0 \text{ } \sigma.p.p, \|f\|_{L^1(E, \mathbb{R}, \sigma)} = 1 \right\} \end{cases} \quad (3.18)$$

De plus, f_l vérifie

$$\forall \alpha' \in \mathbb{R}, \left\| V - \alpha \|F_l\|_{L^1(E, \mathbb{R}, \sigma)} \psi[f_l] \right\|_{\mathbb{C}^n}^2 \leq \|V - \alpha' \psi[f_l]\|_{\mathbb{C}^n}^2 \quad (3.19)$$

Pour la démonstration de ce théorème voir la Section A.5 page 116. Nous n'avons pas démontré la convergence de l'algorithme pour toute valeur de $\mu > 0$. Cependant nous n'avons pas rencontré de problèmes de convergence dans les applications numériques.

3.5 Dualité

La régularisation entropique est très utilisée en traitement du signal et en particulier pour la restauration des images [35, 34]. Pour beaucoup de ces problèmes le nombre de variables est très grand ; dans le cas de la restauration d'image ce nombre n'est rien d'autre que le nombre de pixels constituant l'image. La résolution se fait alors en passant par le problème dual qui est un problème sans contraintes. La solution est ensuite calculée numériquement via une méthode de descente [4] sur le problème dual.

Nous allons calculer le problème dual de (\mathcal{P}_μ) dans cette sous-section. Avant de présenter les calculs, nous faisons quelques rappels sur la dualité en programmation convexe.

3.5.1 Introduction et rappels

Pour une description complète de la théorie de la dualité en analyse convexe voir [15, 4, 47]. Soit V un espace de Banach et son dual V' . On considère une fonction $F : V \rightarrow \mathbb{R} \cup \{+\infty\}$ et on s'intéresse à la minimisation de cette fonction, ce qui constitue le problème *primal*

$$(\mathcal{P}) \quad \inf_{u \in V} F(u)$$

Soient Y un espace de Banach et son dual Y' . On considère une fonction $\Phi : V \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$, telle que

$$\Phi(u, 0) = F(u).$$

Nous allons utiliser la fonction Φ pour fabriquer des problèmes *perturbés* (\mathcal{P}_p) en posant

$$(\mathcal{P}_p) \quad \inf_{u \in V} \Phi(u, p).$$

Les problèmes perturbés (\mathcal{P}_p) vont permettre de définir un problème *dual* (\mathcal{D}). Pour cela on considère la fonction conjuguée ([15, 47]) $\Phi^* = V' \times Y' \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par :

$$\Phi^*(u', p') = \sup_{u \in V, p \in Y} [\langle p', p \rangle + \langle u', u \rangle - \Phi(u, p)].$$

L'expression du problème dual relativement aux perturbations considérées est alors :

$$(\mathcal{D}) \quad \sup_{p' \in Y'} \{-\Phi^*(0, p')\}$$

Définition 11 *Le Lagrangien $\mathcal{L} : V \times Y' \rightarrow \bar{\mathbb{R}}$ du problème primal relativement aux perturbations considérées est défini par*

$$-L(u, p') = \sup_{p \in Y} \{\langle p', p \rangle - \Phi(u, p)\}$$

On peut montrer que le Lagrangien est concave en p' et que si Φ est convexe alors le Lagrangien est convexe en u [15].

Définition 12 *On dit que (\bar{u}, \bar{p}') est un point selle du Lagrangien si*

$$L(\bar{u}, p') \leq L(\bar{u}, \bar{p}') \leq L(u, \bar{p}') \quad \forall u \in V, \forall p' \in Y'.$$

Proposition 4 *Sous l'hypothèse que Φ est propre, convexe et s.c.i, les deux conditions suivantes sont équivalentes*

1. (\bar{u}, \bar{p}') est un point selle de L .
2. \bar{u} est solution de (\mathcal{P}), \bar{p}' est solution de (\mathcal{D}) et $\inf(\mathcal{P}) = \sup(\mathcal{D})$

Cela signifie que pour résoudre les problèmes primal et dual, il suffit de trouver un point selle du Lagrangien.

3.5.2 Théorème de Kuhn-Tucker

Un cas particulier important de ce qui précède est la dualité en programmation convexe et le théorème de Kuhn-Tucker. On considère le problème de minimisation suivant

$$\min \{f(x) \mid f_i(x) = 0, i = 1, \dots, n\} \quad (3.20)$$

où $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe et où, pour $i \in \{1, \dots, n\}$, les $f_i : X \rightarrow \mathbb{R} \cup \{+\infty\}$ sont des fonctions convexes continues. Le problème (3.20) est le problème *primal*. Le *Lagrangien* du problème (3.20) est une fonction $\mathcal{L} : X \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par

$$\mathcal{L}(x, \lambda_1, \dots, \lambda_n) = f(x) + \sum_{i=1}^n \lambda_i f_i(x).$$

Théorème 11 (Kuhn-Tucker) *Supposons que f est propre et qu'il existe une solution réalisable dans l'ensemble $\text{ri}(\text{dom } \phi)$, où ri désigne l'intérieur relatif. Alors $\bar{x} \in X$ est solution du problème de minimisation (3.20) si et seulement si, il existe $\{\lambda_i \in \mathbb{R} \mid i = 1, \dots, n\}$ (les multiplicateurs de Lagrange) tels que*

$$\begin{cases} 0 \in \partial \mathcal{L}(\bar{x}, \lambda_1, \dots, \lambda_n) = \partial f(\bar{x}) + \sum_{i=1}^n \lambda_i \partial f_i(\bar{x}) \\ f_i(\bar{x}) = 0 \text{ pour tout } i \in \{1, \dots, n\}. \end{cases} \quad (3.21)$$

On remarque que si le vecteur de Kuhn-Tucker $(x, \lambda_1, \dots, \lambda_n)$ existe, c'est un point selle du Lagrangien. Cela signifie que, pour tout $x \in X$ et tout $h_i \in \mathbb{R}, i = 1, \dots, n$, on a

$$\mathcal{L}(\bar{x}, h_1, \dots, h_n) \leq \mathcal{L}(\bar{x}, \lambda_1, \dots, \lambda_n) \leq \mathcal{L}(x, h_1, \dots, h_n).$$

Le problème *primal* peut donc se réécrire, avec le Lagrangien, sous la forme

$$\inf_{x \in X} \sup_{h \in \mathbb{R}^n} \mathcal{L}(x, h_1, \dots, h_n)$$

On appelle problème *dual* le problème

$$\sup_{h \in \mathbb{R}^n} \inf_{x \in X} \mathcal{L}(x, h_1, \dots, h_n)$$

Les solutions du problème primal et du problème dual sont reliées par la *relation d'extrémalité*. Cette relation n'est rien d'autre que la relation (3.21) du le Théorème précédent.

Remarque 1 *Dans cette version du théorème de Kuhn-Tucker il n'y a que des contraintes d'égalité. Mais il y a aussi un résultat dans le cas où les contraintes sont constituées d'un mélange de contraintes d'inégalité et d'égalité, voir [47, 4, 14]*

3.5.3 Calcul du problème dual pour la méthode du maximum d'entropie

Nous allons maintenant montrer, de façon rapide, comment on obtient le dual du problème du maximum d'entropie (pour les détails de la théorie consulter [7, 10]). Le problème primal s'écrit

$$(\mathcal{P}) \begin{cases} \min H(f) \\ f \geq 0, \psi[f] = V \end{cases}$$

L'entropie $-H$ agit comme une fonction barrière, elle force la solution du problème de maximum d'entropie à être strictement positive. La contrainte $f \geq 0$ est donc toujours inactive et le Lagrangien est [10]

$$\mathcal{L}(f, \lambda) = H(f) + \langle \lambda, \psi[f] - V \rangle_{\mathbb{C}^n}.$$

Calculons le sous-différentiel de H : il faut tout d'abord calculer le sous-différentiel de la fonction ϕ définie en page 27. Rappelons que le sous-différentiel [4, 15] d'une fonction $f : X \rightarrow \mathbb{R}$ est défini en un point $x \in X$ par

$$\lambda \in \partial f(x) \subset X' \Leftrightarrow f(x) + \langle \lambda, y - x \rangle_{X', X} \leq f(y).$$

Pour la fonction ϕ , on obtient

$$\partial\phi(x) = \begin{cases} \{1 + \ln x\} & \text{si } x > 0 \\ \emptyset & \text{sinon.} \end{cases}$$

On en déduit que le sous différentiel de H est

$$\partial H(f) = \begin{cases} \{1 + \ln f\} & \text{si } f \in \{f \in L^2 \mid f > 0 \text{ p.p.}\} \\ \emptyset & \text{sinon.} \end{cases}$$

Le sous-différentiel du Lagrangien s'écrit donc sous la forme

$$\partial\mathcal{L}(f, \lambda) = \partial H(f) + \psi^*[\lambda],$$

on en déduit

$$\partial\mathcal{L}(f, \lambda) \ni 0 \Leftrightarrow 1 + \ln(f) + \psi^*[\lambda] = 0.$$

Le minimum est donc atteint pour $f = \exp(-1 - \psi^*[\lambda])$ et on obtient l'expression du problème dual en remplaçant f par cette expression dans le Lagrangien ; il vient

$$\begin{aligned} \mathcal{L}(\exp(-1 - \psi^*[\lambda]), \lambda) &= \langle \exp(-1 - \psi^*[\lambda]), -1 - \psi^*[\lambda] \rangle_{L^2} \\ &+ \langle \psi^*[\lambda], \exp(-1 - \psi^*[\lambda]) \rangle_{L^2} - \langle \lambda, V \rangle_{\mathbb{C}^n} \\ &= \langle \exp(-1 - \psi^*[\lambda]), -1 - \psi^*[\lambda] + \psi^*[\lambda] \rangle_{L^2} - \langle \lambda, V \rangle_{\mathbb{C}^n} \end{aligned}$$

Finalement (\mathcal{D}) s'écrit :

$$(\mathcal{D}) \left\{ \begin{array}{l} \max \\ \lambda \in \mathbb{C}^n \end{array} \{- \langle \exp(-1 - \psi^*[\lambda]), 1 \rangle_{L^2} - \langle \lambda, V \rangle_{\mathbb{C}^n}\} \right.$$

Le dual du problème de maximum d'entropie est donc un problème sans contrainte. Pour le résoudre il suffit de calculer l'unique $\lambda \in \mathbb{C}^n$ qui annule le gradient de la fonction

$$\Lambda : \lambda \mapsto - \langle \exp(-1 - \psi^*[\lambda]), 1 \rangle_{L^2} - \langle \lambda, V \rangle_{\mathbb{C}^n},$$

c'est à dire

$$\Lambda(\lambda) = - \int_E \exp(-1 - \psi^*[\lambda]) d\sigma - \sum_{i=1}^n \lambda_i V_i.$$

Le gradient de Λ est

$$\begin{aligned} \nabla\Lambda(\lambda). \delta\lambda &= - \int_E \exp(-1 - \psi^*[\lambda]) \psi^*[\delta\lambda] d\sigma - \langle \delta\lambda, V \rangle_{\mathbb{C}^n} \\ &= \langle \psi^*[\delta\lambda], \exp(-1 - \psi^*[\lambda]) \rangle_{L^2} - \langle \delta\lambda, V \rangle_{\mathbb{C}^n} \\ &= \langle \delta\lambda, \psi[\exp(-1 - \psi^*[\lambda])] \rangle_{\mathbb{C}^n} - \langle \delta\lambda, V \rangle_{\mathbb{C}^n} \\ &\Rightarrow \nabla\Lambda(\lambda) = \psi[\exp(-1 - \psi^*[\lambda])] - V \end{aligned}$$

L'optimum est donc atteint pour $\lambda \in \mathbb{C}^n$ tel que

$$\nabla\Lambda(\lambda) = 0 \Leftrightarrow \psi[\exp(-1 - \psi^*[\lambda])] = V,$$

i.e. la solution de maximum d'entropie est $f = \exp(-1 - \psi^*[\lambda])$ où λ est l'unique élément de \mathbb{C}^n rendant f réalisable.

3.5.4 Calcul du dual du problème (\mathcal{P}_μ)

Rappelons que le problème (\mathcal{P}_μ) s'écrit sous la forme

$$(\mathcal{P}_\mu) \begin{cases} \min \|V - \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F), \\ F \in L^2(E, \mathbb{R}^+, \sigma) \end{cases} .$$

Le dual (\mathcal{D}_μ) de (\mathcal{P}_μ) va être calculé grâce à la dualité de Fenchel-Rockafellar (voir [15] pages 57-61). La fonction que l'on veut minimiser s'écrit

$$J_\mu(f) = G(\psi[f]) + F(f)$$

où $F(f) = \mu H(f)$ et $G(x) = \|V - x\|_{\mathbb{C}^n}^2$. Le problème dual est [15] :

$$\sup_{\ell \in \mathbb{C}^n} [-G^*(-\ell) - F^*(\psi^*[\ell])]$$

Calculons la conjuguée de F ; le sous différentiel de $f \mapsto \langle \ell, f \rangle - \mu(H(f))$ en $f > 0$ p.p est $\ell - \mu(1 + \ln f)$. On en déduit

$$\begin{aligned} F^*(\ell) &= \left\langle \ell, \exp\left(-1 + \frac{\ell}{\mu}\right) \right\rangle_{L^2(E, \mathbb{R}, \sigma)} - \mu \left\langle \exp\left(-1 + \frac{\ell}{\mu}\right), -1 + \frac{\ell}{\mu} \right\rangle_{L^2(E, \mathbb{R}, \sigma)} \\ &= \mu \int_E \exp\left(-1 + \frac{\ell}{\mu}\right) d\sigma = \mu \left\langle \exp\left(-1 + \frac{\ell}{\mu}\right), 1 \right\rangle_{L^2(E, \mathbb{R}, \sigma)} . \end{aligned}$$

La fonction G est la norme de \mathbb{C}^n translatée par $V \in \mathbb{C}^n$, sa conjuguée est bien connue [47] et s'écrit

$$G^*(\ell) = \frac{1}{4} \|\ell\|_{\mathbb{C}^n}^2 + \langle \ell, V \rangle_{\mathbb{C}^n}$$

Le problème dual est donc

$$\sup_{\ell \in \mathbb{C}^n} [-G^*(-\ell) - F^*(\psi^*[\ell])] = \sup_{\ell \in \mathbb{C}^n} \left[-\frac{1}{4} \|\ell\|_{\mathbb{C}^n}^2 + \langle \ell, V \rangle_{\mathbb{C}^n} - \mu \int_E \exp\left(-1 + \frac{\psi^*[\ell]}{\mu}\right) d\sigma \right],$$

$$(\mathcal{D}_\mu) \quad \sup_{\ell \in \mathbb{C}^n} \left\{ -\mu \int_E \exp\left(-1 + \frac{\psi^*[\ell]}{\mu}\right) d\sigma + \langle \ell, V \rangle - \frac{1}{4} \|\ell\|^2 \right\}$$

Sans utiliser la théorie de la dualité de Fenchel-Rockafellar il est possible, grâce aux résultats précédents, de montrer que les problèmes (\mathcal{P}_μ) et (\mathcal{D}_μ) sont équivalents. La dérivée de la fonction coût du problème (\mathcal{D}_μ) , est

$$V - \frac{1}{2}\ell - \psi \left[\exp\left(-1 + \frac{\psi^*[\ell]}{\mu}\right) \right].$$

Il s'ensuit que l'optimum est atteint en $2\bar{\ell}$ si et seulement si

$$V - \bar{\ell} - \psi \left[\exp\left(-1 + \frac{2}{\mu} \psi^*[\bar{\ell}]\right) \right] = 0;$$

or d'après le Théorème 9, la dernière équation est une condition nécessaire et suffisante d'optimalité pour le problème (\mathcal{P}_μ) et la solution de ce problème s'écrit

$$F_\mu = \exp \left(-1 + \frac{2}{\mu} \psi^* [\bar{\ell}] \right)$$

Les problèmes (\mathcal{P}_μ) et (\mathcal{D}_μ) sont donc deux versions du même problème. Il est aisé de constater que l'on obtient le problème dual en remplaçant F par ℓ via les relations $F = \exp \left(-1 + \frac{\psi^*[\ell]}{\mu} \right)$ et $\frac{\ell}{2} = V - \psi[F]$ dans le problème primal et vice versa.

3.5.5 Dual du problème de maximum d'entropie relaxé

Le Théorème 8 page 8 montre que le problème de régularisation entropique (\mathcal{P}_μ) est semblable au problème

$$(P^\varepsilon) \left\{ \begin{array}{l} \min H(f) \\ f \in L^2, f \geq 0, \|V - \psi[f]\|^2 \leq \varepsilon^2 \end{array} \right. ,$$

Dans ce cas l'erreur commise avec la donnée $V \in \mathbb{C}^n$ est fixée *a priori* par $\varepsilon > 0$. L'existence d'une solution n'est pas évidente. Si nous étions dans l'espace $L^1(E, \mathbb{R}, \sigma)$, il n'y aurait pas de problème d'existence puisque la négentropie H est coercive. Cela n'est plus vrai dans $L^2(E, \mathbb{R}, \sigma)$ comme le montre l'exemple de la Section A.4. Toutefois nous savons qu'il existe une solution pour $\varepsilon > 0$ assez grand, car la solution F_μ de (\mathcal{P}_μ) est solution de (P^ε) pour $\varepsilon = \|V - \psi[F_\mu]\|^2$ (la preuve est similaire à celle du Théorème 6). Nous en concluons qu'il existe une solution au problème (P^ε) pour ε assez grand et que cette solution est strictement positive. Nous pouvons donc éliminer la contrainte de positivité $f \geq 0$. Ecrivons le Lagrangien du problème : soient $f \in L^2(E, \mathbb{R}, \sigma)$, $f \geq 0$ admissible et $\alpha \in \mathbb{R}^+$

$$\mathcal{L}(f, \alpha) = H(f) + \alpha \left[\|V - \psi[f]\|^2 - \varepsilon^2 \right].$$

Calculons le dual du problème. Le problème primal est

$$\inf_f \sup_{\alpha \geq 0} \mathcal{L}(f, \alpha),$$

et le dual est

$$\sup_{\alpha \geq 0} \inf_f \mathcal{L}(f, \alpha).$$

$$\begin{aligned} \inf_f \mathcal{L}(f, \alpha) &= \inf_f H(f) + \alpha \|V - \psi[f]\|^2 - \alpha \varepsilon^2 \\ &= \inf_f \left[H(f) + \alpha \|V - \psi[f]\|^2 \right] - \alpha \varepsilon^2 \\ &= \alpha \underbrace{\left[\inf_f \mu H(f) + \|V - \psi[f]\|^2 \right]}_{(\mathcal{P}_\mu)} - \alpha \varepsilon^2 \end{aligned}$$

où nous avons posé $\mu = \alpha^{-1}$. Nous reconnaissons le problème (\mathcal{P}_μ) , comme $\inf \mathcal{P}_\mu = \sup \mathcal{D}_\mu$ nous le remplaçons par l'expression de son dual calculé précédemment. On obtient

$$\begin{aligned} \sup_{\alpha \geq 0} \inf_f \mathcal{L}(f, \alpha) &= \sup_{\alpha \geq 0} \alpha \left[\max_{\ell} \frac{-1}{\alpha} \int_E \exp(-1 + \alpha \psi^*[\ell]) d\sigma + \langle \ell, V \rangle - \frac{\|\ell\|^2}{4} \right] - \alpha \epsilon^2 \\ &= \sup_{\alpha \geq 0} \max_{\ell} - \int_E \exp(-1 + \alpha \psi^*[\ell]) d\sigma + \alpha \langle \ell, V \rangle - \alpha \frac{\|\ell\|^2}{4} - \alpha \epsilon^2 \end{aligned}$$

posons $l = \frac{\ell}{\mu} = \alpha \ell$, il vient

$$\sup_{\alpha \geq 0} \inf_f \mathcal{L}(f, \alpha) = \sup_{\alpha \geq 0} \max_{\ell} - \int_E \exp(-1 + \psi^*[l]) d\sigma + \langle l, V \rangle - \frac{\|l\|^2}{4\alpha} - \alpha \epsilon^2.$$

On peut intervertir le sup et le max; en dérivant l'expression

$$- \int_E \exp(-1 - \psi^*[l]) d\sigma + \langle l, V \rangle - \frac{\|l\|^2}{4\alpha} - \alpha \epsilon^2,$$

nous voyons que le sup sur les α est atteint pour

$$\alpha = \frac{\|l\|}{2\epsilon}.$$

Finalement, l'expression du problème dual est

$$\sup_l \left\{ - \int_E \exp(-1 + \psi^*[l]) d\sigma + \langle l, V \rangle - \epsilon \|l\| \right\} \quad (3.22)$$

Chapitre 4

Généralisation de l'algorithme d'inversion à des opérateurs de régularisation strictement convexes

4.1 Introduction

Nous allons généraliser la méthode de régularisation entropique, introduite au chapitre précédent à d'autres opérateurs de régularisation. Le cadre de travail est le suivant : nous étudions un problème inverse $\psi[x] = y$ où ψ est un opérateur linéaire d'un espace de dimension infinie à valeurs dans un espace de dimension finie (typiquement \mathbb{C}^n). Dans le chapitre précédent nous étions dans le même cas puisque que nous devions déterminer une fonction $F : S \rightarrow \mathbb{R}^+$ à partir de la matrice V des mesures du champ. La méthode de régularisation entropique utilisée dans la section 3.4 a plusieurs avantages. Premièrement, la méthode fournit une expression analytique de la solution, ce qui permet de connaître la régularité de cette solution et surtout de ne pas utiliser de discrétisation du problème. Deuxièmement, la solution se calcule grâce à un algorithme de point fixe très simple à mettre en œuvre et les calculs sont rapides car on évite la discrétisation.

Dans ce chapitre, nous allons chercher à généraliser les résultats de la section 3.4, à des opérateurs de régularisation autres que l'entropie. Plus précisément nous allons chercher pour quels opérateurs de régularisation la solution a une expression analytique qui puisse se calculer sans passer par une discrétisation du problème.

4.2 Régularisation du problème inverse

On se place dans un espace de Banach **réflexif** X et on considère un opérateur linéaire $\psi \in \mathcal{L}(X, \mathbb{C}^n)$. Nous cherchons à déterminer $x \in X$ à partir de $y \in \mathbb{C}^n$ grâce à l'équation

$$\psi[x] = y. \tag{4.1}$$

C'est un problème inverse mal posé car ψ n'est clairement pas bijectif, le noyau $N(\psi)$ est un sous-espace de dimension infinie de X et si $x_0 \in X$ est solution de (4.1) alors

$$\forall x \in x_0 + N(\psi), \psi[x] = y.$$

Cette équation a donc *a priori* une infinité de solutions. De plus, il peut ne pas exister de solution $x \in X$ à l'équation (4.1). Pour garantir l'existence il faut supposer $y \in R(\psi)$ (image de ψ). Notons

$$S(y) = \{x \in X \mid \psi[x] = y\}$$

l'ensemble (vide ou non) des solutions de (4.1).

Pour trouver une solution unique et assurer la stabilité nous allons minimiser l'erreur $x \mapsto \|y - \psi[x]\|_{\mathbb{C}^n}^2$ pénalisée par une fonction $\phi : X \rightarrow \mathbb{R}$ semi-continue inférieurement (s.c.i.), Gâteaux-différentiable et convexe. La fonction ϕ sera l'opérateur de régularisation (ou fonctionnelle régularisante), i.e. un critère pour choisir une solution unique et stable. Le problème d'optimisation à résoudre s'écrit

$$(\mathcal{P}_\mu) \begin{cases} \min J_\mu(x) \\ x \in K \end{cases}$$

où la fonction coût est donnée par

$$J_\mu(x) := \|y - \psi[x]\|_{\mathbb{C}^n}^2 + \mu\phi(x).$$

L'ensemble K est un ensemble de contraintes. Dans la suite, nous étudierons surtout le cas sans contraintes $K = X$. La fonction J_μ est s.c.i., convexe et coercive sur $R(\psi^*)$. Pour garantir l'existence d'une solution de (\mathcal{P}_μ) , il faut supposer ϕ coercive sur $N(\psi)$. Il en résulte que J_μ est coercive sur $X = N(\psi) \oplus R(\psi^*)$ et pour montrer l'existence il suffit d'appliquer les théorèmes standard ([11] (Cor III.20 page 46)). Afin d'obtenir une unique solution, nous allons de plus supposer que la fonctionnelle ϕ est strictement convexe. Nous noterons x_μ l'unique solution du problème (\mathcal{P}_μ) , i.e. la solution du problème inverse (4.1) obtenue par régularisation.

Comme dans le cas de la régularisation entropique (chapitre précédent), le paramètre de régularisation permet de créer un équilibre entre l'erreur avec la donnée y et la stabilité de la solution. Quand μ est choisi très petit, l'erreur sera faible mais la solution obtenue sera très instable : une petite perturbation de la donnée y entraînera une grande variation de la solution du problème (\mathcal{P}_μ) . Par contre, pour des plus grandes valeurs de μ on obtient une solution plus stable mais avec une plus grande erreur. On peut se demander ce qui se passe pour les valeurs "limites" du paramètre de régularisation, c'est l'objet de la proposition suivante.

Proposition 5 Soient $\psi \in L(X, \mathbb{C}^n)$ et $\phi : X \rightarrow \mathbb{R}$ une fonctionnelle s.c.i, strictement convexe et coercive. Soit $y \in \mathbb{C}^n$ tel que $S(y) \neq \emptyset$ et notons x_∞ et x_0 les éléments de X qui vérifient respectivement

$$\phi(x_\infty) = \min_{x \in X} \phi(x) \tag{4.2}$$

$$\phi(x_0) = \min\{\phi(x) \mid \psi(x) = y\}. \tag{4.3}$$

Alors la solution x_μ du problème (\mathcal{P}_μ) vérifie

$$x_\mu \rightharpoonup x_\infty \text{ quand } \mu \rightarrow +\infty \quad (4.4)$$

$$x_\mu \rightarrow x_0 \text{ quand } \mu \rightarrow 0. \quad (4.5)$$

Démonstration. L'existence et l'unicité des éléments x_0 et x_∞ se déduisent de la coercivité et de la stricte convexité de ϕ .

1. On peut toujours supposer que $\phi(x_\infty) = 0$ quitte à ajouter une constante à la fonction ϕ , c'est à dire $\phi \geq 0$. Comme x_μ est la solution du problème (\mathcal{P}_μ) , on a pour tout $\mu > 0$

$$\|y - \psi[x_\mu]\|^2 + \mu\phi(x_\mu) \leq \|y - \psi[x_\infty]\|^2$$

donc

$$0 \leq \phi(x_\mu) \leq \frac{1}{\mu} \left[\|y - \psi[x_\infty]\|^2 - \|y - \psi[x_\mu]\|^2 \right] \leq \frac{1}{\mu} \|y - \psi[x_\infty]\|^2;$$

on en déduit que $\phi(x_\mu) \rightarrow 0 = \phi(x_\infty)$ quand $\mu \rightarrow +\infty$. Comme la fonction ϕ est coercive, l'ensemble $\{x_\mu ; \mu \geq A\}$ est borné dans X pour $A > 0$. Comme X est réflexif, on peut extraire une suite notée $\{x_{\mu_n}\}_{n \in \mathbb{N}}$, avec $\mu_n \rightarrow +\infty$, qui converge faiblement : $x_{\mu_n} \rightharpoonup \bar{x} \in X$. La semi-continuité inférieure de ϕ donne

$$\phi(\bar{x}) \leq \liminf_{n \in \mathbb{N}} \phi(x_{\mu_n}) = 0 = \phi(x_\infty).$$

Par unicité du minimum $\bar{x} = x_\infty$ est la limite de cette suite. Cela démontre que

$$x_\mu \rightharpoonup x_\infty \text{ quand } \mu \rightarrow +\infty.$$

2. Pour tout $\mu > 0$, on a

$$\mu\phi(x_\infty) = 0 \leq \mu\phi(x_\mu) \leq \|y - \psi[x_\mu]\|^2 + \mu\phi(x_\mu) \leq \mu\phi(x_0). \quad (4.6)$$

Donc, pour tout $\mu > 0$, $\phi(x_\mu) \leq \phi(x_0)$ et la famille $\phi(x_\mu)$ est bornée. Soit $\{\mu_n\}_{n \in \mathbb{N}}$ une suite qui tend vers 0. On peut extraire de la suite $\{x_{\mu_n}\}$ une sous-suite, toujours notée x_{μ_n} , qui converge faiblement vers $\tilde{x} \in X$. D'après l'équation (4.6) on a

$$0 \leq \|y - \psi(x_{\mu_n})\|^2 \leq \mu(\phi(x_0) - \phi(x_{\mu_n})) \leq \mu\phi(x_0);$$

d'où, en passant à la limite on obtient

$$\psi(\tilde{x}) = y \text{ et } \phi(\tilde{x}) \leq \phi(x_0).$$

Ainsi, toute la suite $\{x_{\mu_n}\}_{n \in \mathbb{N}}$ converge vers $\tilde{x} = x_0$ par unicité de x_0 .

□

Ainsi, lorsque μ tend vers 0, la solution du problème (\mathcal{P}_μ) converge faiblement vers l'unique x de $S(y)$ qui minimise la fonctionnelle ϕ . A l'inverse quand μ devient très grand, la solution tend vers l'élément x_∞ qui est le minimum global de ϕ . En particulier, dans le cas de la régularisation entropique, lorsque μ devient très grand on obtient une solution qui "s'aplatit"

de plus en plus car le minimiseur global de H est une fonction constante. Notons que l'on obtient des convergences fortes dans la proposition précédente, si on suppose que $\nabla\phi$ est un opérateur fortement monotone dans un voisinage de x_0 et de x_∞ .

Un parallèle peut être fait entre le problème (\mathcal{P}_μ) et la régularisation de Moreau-Yosida [5]. On peut considérer que l'opérateur $\psi : X \rightarrow \mathbb{C}^n$ est injectif, quite à restreindre l'espace X au complément orthogonal de $N(\psi)$. Dans ce cas on peut reformuler le problème (\mathcal{P}_μ) par

$$\min_{u \in \mathbb{C}^n} \frac{1}{\mu} \|y - u\|_{\mathbb{C}^n}^2 + \phi(\psi^{-1}[u]),$$

et on reconnaît la régularisée de Moreau-Yosida de la fonction $\phi \circ \psi^{-1}$. Cette régularisée est une fonction convexe, continue et Gâteaux-différentiable et elle converge quand $\mu \rightarrow 0$ vers $\phi \circ \psi^{-1}$ [5].

Sans supposer que ϕ est injective, on considère la fonction

$$\varphi(u) = \inf \{ \phi(\psi^{-1}[u]) \}$$

cette fonction est convexe et propre puisque ϕ est coercive sur les ensembles $\psi^{-1}[u]$. Le problème (\mathcal{P}_μ) est identique à

$$\min_{u \in \mathbb{C}^n} \frac{1}{\mu} \|y - u\|_{\mathbb{C}^n}^2 + \inf \{ \phi(\psi^{-1}[u]) \},$$

c'est à dire à la régularisée de Moreau-Yosida de φ . Lorsque $\mu \rightarrow 0$ on obtient que la solution du problème (\mathcal{P}_μ) est

$$\varphi(y) = \inf \{ \phi(\psi^{-1}[y]) \} = x_0,$$

et on retrouve le résultat de la proposition précédente.

4.2.1 Stabilité de la solution

Il y a trois notions de stabilité à considérer. La plus faible est celle où la solution x_μ du problème de régularisation dépend continûment de la donnée $y \in \mathbb{C}^n$. C'est d'ailleurs une condition pour que le problème soit bien posé au sens de Hadamard [22, 64]. Mais cela n'est pas satisfaisant car on aimerait avoir une dépendance plus fine. C'est pourquoi nous allons dans la suite chercher à obtenir une solution qui varie de façon lipschitz par rapport à la donnée y . Plus précisément, si on note $x_\mu(y)$ et $x_\mu(y + \delta y)$ les solutions correspondant respectivement aux données $y \in \mathbb{C}^n$ et $y + \delta y \in \mathbb{C}^n$, où δy représente une petite variation, nous voulons :

$$\|x_\mu(y) - x_\mu(y + \delta y)\| \leq c \|\delta y\| \text{ avec une constante } c > 0 \quad (4.7)$$

En physique, la dépendance lipschitzienne de la solution par rapport à la donnée y n'est pas suffisante. En effet, si la constante de lipschitz $c > 0$ est très grande, il est possible qu'une petite perturbation de la donnée initiale entraîne une grande variation de la solution. C'est pourquoi, nous considérons que la solution est stable pour des applications en physique expérimentale si la constante $c \approx 1$ ou même $c = o(1)$.

La stabilité des solutions est assurée quand $\mu \rightarrow +\infty$. En effet pour n'importe quelle valeur de $y \in \mathbb{C}^n$, on a $x_\mu(y) \rightarrow x_\infty$ d'après la Proposition 5. La solution n'est alors plus du tout influencée par le choix de $y \in \mathbb{C}^n$. A l'opposé, quand $\mu \rightarrow 0$, la solution n'a aucune raison d'être stable. Par exemple, dans l'exemple de la sous-section 3.2.2, l'opérateur de dérivation n'est pas continu : il n'y a donc pas de stabilité.

Dans la section suivante nous allons essayer de généraliser la méthode employée dans le chapitre 2 pour résoudre (\mathcal{P}_μ) dans le cas où $\phi : X \rightarrow \mathbb{R}$ est strictement convexe.

4.3 Cas où ϕ est strictement convexe

Comme nous l'avons vu dans la section précédente, si on suppose que $\phi : X \rightarrow \mathbb{R}$ est strictement convexe, alors il existe une unique solution au problème (\mathcal{P}_μ) . Si on suppose de plus que ϕ est Gâteaux-différentiable sur tout X , alors la fonction J_μ est Gâteaux-différentiable et une condition d'optimalité du premier ordre est $\nabla J_\mu(x_\mu) = 0$. On a pour tout $x, h \in X$

$$\begin{aligned} \langle \nabla J_\mu(x), h \rangle_{X', X} &= -2 \langle y - \psi[x], \psi[h] \rangle_{\mathbb{C}^n} + \mu \langle \nabla \phi(x), h \rangle_{X', X} \\ &= \langle -2\psi^*[y - \psi[x]], h \rangle_{X', X} + \mu \langle \nabla \phi(x), h \rangle_{X', X} \\ \Rightarrow \nabla J_\mu(x) &= -2\psi^*[y - \psi[x]] + \mu \nabla \phi(x), \end{aligned}$$

et la condition d'optimalité est

$$-2\psi^*[y - \psi[x_\mu]] + \mu \nabla \phi(x_\mu) = 0 \quad \text{dans } X'. \quad (4.8)$$

(X' désignant le dual de X). Cependant le cas où ϕ est définie sur tout X est très réducteur et la proposition suivante traite le cas où $\text{dom } \phi \subsetneq X$. Nous noterons $(\text{dom } \phi)^\circ$ l'intérieur du domaine de ϕ .

Proposition 6 *Soit $\psi \in \mathcal{L}(X, \mathbb{C}^n)$, et $\phi : X \rightarrow \mathbb{R}$ de domaine $\text{dom } \phi \subset X$, tel que $(\text{dom } \phi)^\circ \neq \emptyset$. On suppose ϕ est strictement convexe, coercive sur $N(\psi)$, Gâteaux-différentiable sur $(\text{dom } \phi)^\circ$ et*

$$\begin{aligned} (\mathcal{H}_1) \quad \forall \{x_n\}_{n \in \mathbb{N}} \subset (\text{dom } \phi)^\circ, \lim_{n \in \mathbb{N}} x_n = \bar{x} \in \text{Fr}(\text{dom } \phi) \\ \Rightarrow \exists x \in (\text{dom } \phi)^\circ, \lim_{n \in \mathbb{N}} \nabla \phi(x_n) \cdot (x - \bar{x}) = -\infty. \end{aligned}$$

Alors, la solution du problème (\mathcal{P}_μ) , $x_\mu \in (\text{dom } \phi)^\circ$ et (4.8) est une condition nécessaire et suffisante d'optimalité.

Démonstration. En utilisant le sous différentiel de ϕ une condition nécessaire d'optimalité du problème s'écrit

$$-2\psi^*[y - \psi[x]] + \mu \partial \phi(x) \ni 0.$$

Supposons que l'optimum soit atteint en un point \bar{x} du bord $\text{Fr}(\text{dom } \phi)$. Soit $x \in (\text{dom } \phi)^\circ$ donné par l'hypothèse (\mathcal{H}_1) . Construisons une suite $\{x_n\}_{n \in \mathbb{N}^*} \subset (\text{dom } \phi)^\circ$ qui converge vers \bar{x} , en posant

$$x_n = \bar{x} - \frac{1}{n}(\bar{x} - x),$$

et on a $n(x_n - \bar{x}) = (x - \bar{x})$. Par stricte monotonie du sous-différentiel $\partial\phi$ on a

$$\partial\phi(\bar{x}).(x_n - \bar{x}) < \partial\phi(x_n).(x_n - \bar{x}),$$

En multipliant par $n \in \mathbb{N}^*$ de chaque côté on obtient

$$\partial\phi(\bar{x}).(x - \bar{x}) \leq \lim_n \nabla\phi(x_n).(x - \bar{x}) = -\infty$$

Le sous-différentiel est donc vide au point \bar{x} , c'est une contradiction avec le fait que $\psi^*[y - \psi[\bar{x}]] \in X'$. L'optimum ne peut donc pas être atteint en un point du bord $\text{Fr}(\text{dom } \phi)$. Finalement, la condition d'optimalité est (4.8) puisque ϕ est Gâteaux-différentiable sur l'ensemble $(\text{dom } \phi)^\circ$. \square

Voici un exemple de fonctionnelle $\phi : L^2(\mathbb{R}^n) \rightarrow \mathbb{R}$ vérifiant l'hypothèse (\mathcal{H}_1) . On considère la fonctionnelle ϕ définie sur l'ensemble $\text{dom } \phi = \{x \in L^2(\mathbb{R}^n) \mid \|x\|_{L^2} \leq 1\}$ et qui, pour un x de cet ensemble, à pour expression

$$\phi(x) = -\sqrt{1 - \|x\|_{L^2}^2}$$

Cette fonction est continue sur son domaine, différentiable et pour $h \in L^2(\mathbb{R}^n)$

$$\langle \nabla\phi(x), h \rangle_{L^2} = \frac{\langle x, h \rangle_{L^2}}{\sqrt{1 - \|x\|_{L^2}^2}};$$

la fonctionnelle ϕ vérifie donc bien l'hypothèse (\mathcal{H}_1) .

Remarque 2 *On peut affaiblir un peu les hypothèses en supposant que ϕ n'est pas différentiable, mais que son sous-différentiel $\partial\phi$ est univoque sur $(\text{dom } \phi)^\circ$ et*

$$\begin{aligned} (\mathcal{H}'_1) \quad \forall \{x_n\}_{n \in \mathbb{N}} \subset (\text{dom } \phi)^\circ, \lim_{n \in \mathbb{N}} x_n = \bar{x} \in \text{Fr}(\text{dom } \phi) \\ \Rightarrow \exists x \in (\text{dom } \phi)^\circ, \lim_{n \in \mathbb{N}} \partial\phi(x_n).(x - \bar{x}) = -\infty. \end{aligned}$$

Sous cette hypothèse, la conclusion de la proposition ne change pas. Un exemple de fonctionnelle ϕ vérifiant cette hypothèse est donnée dans la sous-section 4.6.3. Notons que le fait que le sous-différentiel soit univoque n'entraîne pas que ϕ est gâteaux différentiable, à moins que ϕ soit continue [15].

Par la suite nous supposerons toujours que nous sommes soit dans le cas de la proposition précédente, soit dans le cas où ϕ est Gâteaux-différentiable sur tout X . Dans tout les cas, la condition d'optimalité est (4.8).

Proposition 7 *Supposons que la fonctionnelle régularisante $\phi : X \rightarrow \mathbb{R}$ est strictement convexe. Alors l'opérateur $\nabla\phi : X \rightarrow X'$ a un inverse noté $\gamma : R(\nabla\phi) \subset X' \rightarrow X$. Il existe une unique solution $x_\mu \in X$ au problème (\mathcal{P}_μ) caractérisée par*

$$x_\mu = \gamma \left(\frac{2}{\mu} \psi^*[y - \psi[x_\mu]] \right). \quad (4.9)$$

De plus il existe $\ell_\mu \in \mathbb{C}^n$ unique point fixe de la fonction $p_\mu : l \mapsto y - \psi \left(\gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right)$. La solution x_μ du problème (\mathcal{P}_μ) est donnée par

$$\begin{cases} \ell_\mu = y - \psi \left(\gamma \left(\frac{2}{\mu} \psi^*[\ell_\mu] \right) \right) \\ x_\mu = \gamma \left(\frac{2}{\mu} \psi^*[\ell_\mu] \right) \end{cases}.$$

Démonstration. Comme ϕ est strictement convexe, son gradient $\nabla\phi : X \rightarrow X'$ est un opérateur strictement monotone. Donc pour tout $x, y \in X$, avec $x \neq y$

$$\langle \nabla\phi(x) - \nabla\phi(y), x - y \rangle_{X', X} > 0,$$

on en déduit que $\nabla\phi$ est une bijection de X dans $R(\nabla\phi)$ et nous pouvons définir l'inverse $\gamma : R(\nabla\phi) \rightarrow X$ de $\nabla\phi$. Cela permet de conclure que la condition d'optimalité (4.8) est équivalente à (4.9). En effet

$$(4.8) \Rightarrow x_\mu = \gamma(\nabla\phi(x_\mu)) = \gamma \left(\frac{2}{\mu} \psi^*[y - \psi[x_\mu]] \right),$$

ce qui montre que la condition est nécessaire. Pour montrer que la condition est suffisante, il suffit d'appliquer $\nabla\phi$ sur l'équation (4.9), on obtient

$$\nabla\phi(x_\mu) = \nabla\phi \circ \gamma \left(\frac{2}{\mu} \psi^*[y - \psi[x_\mu]] \right) = \frac{2}{\mu} \psi^*[y - \psi[x_\mu]].$$

Finalement, le fait que la fonction p_μ a un unique point fixe découle de l'unicité de la solution x_μ du problème, et on pose $\ell_\mu = y - \psi[x_\mu]$. \square

Corollaire 2 Soient un opérateur linéaire $\psi : X \rightarrow \mathbb{C}^n$, $\phi : X \rightarrow \mathbb{R}$ une fonction convexe, s.c.i, Gâteaux-différentiable et coercive sur $N(\psi)$ et $y \in \mathbb{C}^n$. Alors toutes les fonctions de $\mathbb{C}^n \rightarrow \mathbb{C}^n$ définies par

$$l \mapsto y - \psi[\gamma(\alpha\psi^*[l])],$$

avec $\alpha > 0$, admettent au moins un point fixe. Si ϕ est strictement convexe, ce point fixe est unique.

Ce corollaire permet de construire des fonctions qui ont un unique point fixe mais qui ne sont pas continues. Mais l'utilité de ce corollaire est limitée par le fait que nous ne pouvons pas voir facilement si une fonction de \mathbb{C}^n dans lui même peut s'écrire sous cette forme.

La Proposition 7 montre que la solution du problème (\mathcal{P}_μ) s'écrit sous la forme

$$x_\mu = \gamma \left(\frac{2}{\mu} \psi^*[y - \psi[x_\mu]] \right) \quad (4.10)$$

où la fonction $\gamma : X' \rightarrow X$ est l'inverse du gradient de ϕ . Comme l'opérateur adjoint ψ^* est défini sur \mathbb{C}^n la solution est caractérisée par un élément de \mathbb{C}^n . Nous n'avons donc pas besoin de discrétiser le problème, pour calculer la solution : il suffit de connaître les n nombres

complexes qui la caractérisent. Nous avons donc une expression analytique de la solution et nous pouvons calculer la solution grâce au même algorithme de point fixe que celui du chapitre 3. Remarquons que dans le cas de la régularisation entropique nous avons utilisé le même argument : le gradient de l'entropie est inversible et l'inverse s'identifie à la fonction exponentielle. Remarquons aussi que (comme pour le cas de la régularisation entropique) nous pouvons calculer la solution grâce à un point fixe dans l'espace de Banach X . En pratique, un tel algorithme a peu d'intérêt car numériquement, il est difficile et coûteux de le mettre en œuvre. Dans la Section 4.3.1 nous allons chercher des conditions pour que l'opérateur ϕ ait un gradient inversible. Nous allons également chercher des conditions pour que la solution soit stable, c'est à dire vérifie l'équation (4.7).

4.3.1 Régularité de l'inverse γ

Etudions la stabilité de la solution obtenue en résolvant le problème (\mathcal{P}_μ) . Nous allons chercher des conditions sur l'opérateur de régularisation $\phi : X \rightarrow \mathbb{R}$ pour que la solution soit stable au sens de l'équation (4.7). La condition d'optimalité (4.8) peut se réécrire

$$\psi^*\psi[x_\mu] + \frac{\mu}{2}\nabla\phi(x_\mu) = \psi^*[y]$$

Posons

$$A : X \rightarrow X', x \mapsto \psi^*\psi[x] + \frac{\mu}{2}\nabla\phi(x).$$

Une condition suffisante de stabilité est que l'opérateur A^{-1} existe et soit localement lipschitzien. Cela est clairement lié à la régularité de l'opérateur inverse γ . La proposition qui va suivre, va nous donner des conditions sur ϕ pour avoir cette propriété.

Proposition 8 *Si $\nabla\phi$ vérifie la condition suivante (locale forte monotonie)*

$$\forall x \in \text{dom } \phi \subset X, \exists r > 0, \exists c(x) > 0, \forall y \in B(x, r),$$

$$\langle \nabla\phi(x) - \nabla\phi(y), x - y \rangle \geq c(x) \|x - y\|^2, \quad (4.11)$$

alors γ est localement lipschitzien. La solution x_μ est stable au sens de l'équation (4.7) car A^{-1} est localement Lipschitzienne. Si on suppose de plus que $\nabla\phi$ est C^1 (au sens Fréchet) alors pour tout $x' \in R(\nabla\phi)$, il existe $\epsilon > 0$, pour tout $y' \in B(x', \epsilon) \cap R(\nabla\phi)$ on a

$$\|\nabla^2\phi(x)^{-1} \cdot (y' - x') - \gamma(y') + \gamma(x')\|_X = o(\|y' - x'\|).$$

Démonstration. Si nous réécrivons (4.11) avec $x = \gamma(x'), y = \gamma(y')$ où $x', y' \in R(\nabla\phi)$, on obtient : $\forall x \in \text{dom } \phi, \exists r > 0, \exists c(x) > 0, \forall y \in B(x, r),$

$$\langle x' - y', \gamma(x') - \gamma(y') \rangle \geq c(\gamma(x')) \|\gamma(x') - \gamma(y')\|^2. \quad (4.12)$$

d'où

$$\|\gamma(x') - \gamma(y')\| \leq (1/c(\gamma(x'))) \|x' - y'\|.$$

La fonction γ est donc continue et même localement lipschitzienne. Montrons, maintenant que l'opérateur A^{-1} est localement lipschitzien. Comme $A = \psi^*\psi + \frac{\mu}{2}\nabla\phi$ et que $\psi^*\psi$ est un opérateur positif, on a

$$\begin{aligned}\langle Ax - Ay, x - y \rangle_{X',X} &\geq \langle \psi^*\psi[x - y], x - y \rangle_{X',X} + \frac{\mu}{2} \langle \nabla\phi(x) - \nabla\phi(y), x - y \rangle_{X',X} \\ &\geq \frac{c(x)\mu}{2} \|x - y\|^2;\end{aligned}$$

par l'inégalité de Cauchy-Schwartz on déduit que $\|Ax - Ay\| \geq c(x)\frac{\mu}{2} \|x - y\|$. L'opérateur A^{-1} est donc localement lipschitzien.

Si $\nabla\phi$ est C^1 , elle est a fortiori continue. Ainsi $R(\nabla\phi)$ est connexe puisque c'est l'image d'un connexe par une application continue. Soient $\varepsilon > 0$ et $x' \in R(\nabla\phi)$ et $y' \in B(x', \varepsilon) \cap R(\nabla\phi)$. Remarquons que $B(x', \varepsilon) \cap R(\nabla\phi) \neq \{x'\}$ par connexité. Notons $y = \gamma(y')$ et $x = \gamma(x')$ alors

$$\gamma(y') - \gamma(x') = y - x,$$

Il s'ensuit

$$\begin{aligned}\|\nabla\phi(y) - \nabla\phi(x) - \nabla^2\phi(x).(y - x)\|_{X'} &= o(\|y - x\|) \\ \|\nabla^2\phi(x)^{-1}.(\nabla\phi(y) - \nabla\phi(x)) - y + x\|_X &= o(\|y - x\|)\end{aligned}$$

finalement

$$\|\nabla^2\phi(x)^{-1}.(y' - x') - \gamma(y') + \gamma(x')\|_X = o(\|\gamma(y') - \gamma(x')\|).$$

Comme γ est lipschitzienne, on a

$$o(\|\gamma(y') - \gamma(x')\|) = o(\|y' - x'\|)$$

ce qui termine la preuve. \square

Corollaire 3 *Sous les mêmes hypothèses que dans la Proposition 8, si $\nabla\phi$ est fortement monotone, son inverse γ existe et est lipschitzien.*

Démonstration. Si $\nabla\phi$ est fortement monotone, il existe $\alpha > 0$ tel que la condition (4.11) soit vérifiée avec $c(x) = \alpha$ pour tout $x \in X$. L'inverse γ est donc α^{-1} -Lipschitzien. \square

Remarque 3 *Supposons qu'il existe un ensemble $I \subset \text{dom } \phi$ fermé telle que la condition (4.11) de la Proposition 8 soit vérifiée seulement pour les éléments de $\text{dom } \phi - I$. Nous dirons dans ce cas que l'ensemble I est l'ensemble des solutions "instables". En effet pour un point $x \in I$ nous ne pouvons pas dire si γ est lipschitzienne en x et donc nous ne savons pas si la solution est stable. Le fait de connaître l'ensemble I permet de caractériser les données y ne convenant pas au problème. En effet, si nous choisissons $y \in \mathbb{C}^n$ qui est proche de l'ensemble $\psi[I] \subset \mathbb{C}^n$ il est possible que cela entraîne un défaut de stabilité. Nous donnerons dans la sous-section 4.6.3, un exemple de fonctionnelle ϕ qui n'est pas stable partout sur son domaine de définition.*

Proposition 9 *On considère l'ensemble Φ des opérateurs de régularisation $\phi : X \rightarrow \mathbb{R}$ strictement convexes et qui vérifient la condition de stabilité (4.11). Si $\beta : X \rightarrow \mathbb{R}$ est une fonction convexe et $a_1, a_2 \in \mathbb{R}^+$ alors*

$$a_1\beta + a_2\phi \in \Phi.$$

Démonstration. Il suffit de voir que pour $x, y \in X$ on a

$$\langle a_1\beta + a_2\phi(x - y), x - y \rangle_{X', X} = a_1 \langle \beta(x - y), x - y \rangle_{X', X} + a_2 \langle \phi(x - y), x - y \rangle_{X', X}$$

Comme β est convexe on a $\langle \beta(x - y), x - y \rangle_{X', X} \geq 0$, d'où la conclusion. \square
 Cette proposition montre que l'ensemble des fonctions Φ qui ont un gradient inversible et qui vérifient la condition de stabilité (4.11) est un cône. C'est même un peu plus qu'un cône puisque la somme d'une fonction convexe et d'un élément de Φ est un élément de Φ . Nous allons présenter quelques exemples de fonctionnelles ϕ qui permettent d'obtenir des solutions stables.

4.4 Exemples dans le cas strictement convexe

4.4.1 Cas fortement convexe

Considérons le cas où $\nabla\phi$ est fortement monotone de constante $\alpha > 0$, et donc γ est α^{-1} -lipschitzien. Par l'inégalité de Cauchy-Schwartz on a $\|Ax - Ay\| \geq \alpha \frac{\mu}{2} \|x - y\|$. L'inverse de l'opérateur A , qui sert à calculer la solution est donc lipschitzien de constante $\frac{2}{\mu\alpha}$. Remarquons que la constante de lipschitz décroît comme $\frac{1}{\mu}$: donc plus μ est grand plus la solution obtenue est stable. Inversement cette constante explose quand $\mu \rightarrow 0$, c'est une illustration du Théorème 3 page 21. Nous allons donner quelques exemples de telles fonctionnelles.

1. Soit X un espace de Banach de dual X' strictement convexe. On considère l'opérateur de régularisation donné par la norme de X , i.e. $\phi = \frac{1}{2} \|\cdot\|_X^2$. Alors l'inverse de $\nabla\phi$ existe et est lipschitzien. En effet $\nabla\phi$ est l'application de dualité $\Lambda : X \rightarrow X'$ qui est univoque sous l'hypothèse X' strictement convexe. De plus Λ est fortement monotone de constante $\alpha = 1$. Ainsi pour un espace de Banach réflexif X , la norme est un opérateur de régularisation qui permet d'obtenir une solution stable qui s'écrit sous la forme (4.9). En particulier l'application norme d'un espace de Hilbert est un "bon" opérateur de régularisation dans notre cadre de travail.
2. Considérons le cas où l'opérateur $\phi : X \rightarrow \mathbb{R}$ est donnée par $\phi(x) = \|Ax\|_Y^2$, où Y est un espace de Hilbert et $A : X \rightarrow Y$ est un opérateur linéaire. Notons que ce cas à été étudié dans le cas où A est un opérateur fermé [23]. Si

$$\langle A^*A(x - y), x - y \rangle_{X', X} \geq \alpha \|x - y\|_X^2,$$

l'opérateur $\phi : x \mapsto \|Ax\|_Y^2$ est fortement convexe. La constante $\alpha > 0$ est ici la plus petite valeur propre de l'opérateur A^*A . Dans ce cas, $\nabla\phi$ est un C^1 -difféomorphisme. L'inverse γ de $\nabla\phi$ est lipschitzien de constante $1/\alpha$. $R(\nabla\phi) = A^*AX$ est un sous-espace vectoriel de X' , il est fermé puisque γ est continue et $R(\nabla\phi) = \gamma^{-1}(X)$.

Nous pouvons toutefois remarquer que l'exemple 2 peut se généraliser en prenant un espace de Banach Y et en utilisant l'application de dualité $\Lambda : Y \rightarrow Y'$ pour définir la fonctionnelle ϕ par $\phi(x) = \langle \Lambda(Ax), Ax \rangle_{Y', Y}$.

4.4.2 Cas particuliers où ϕ est non local

Un cas particulier intéressant est celui où l'opérateur ϕ est défini à l'aide d'une intégrale. Le cas de l'entropie en est un exemple particulier. Nous allons voir sous quelles conditions un tel opérateur a un gradient inversible et conduit à une solution stable.

Proposition 10 *On considère des opérateurs $\phi : L^2(E, \mathbb{R}^q, \lambda) \rightarrow \mathbb{R}$ qui s'écrivent sous la forme*

$$\phi(x) = \int_E j \circ x \, d\lambda \quad (4.13)$$

où $E \subset \mathbb{R}^p$ est un pavé, i.e. $E = \prod_{i=1}^p [a_i, b_i]$, λ est la mesure de Lebesgue et $j : \mathbb{R}^q \rightarrow \mathbb{R}$ est une fonction $L_{loc}^\infty(\mathbb{R}^q)$ telle que

$$\beta = \lim_{t \rightarrow +\infty} \sup_{\|x\| \geq t} \frac{|j(x)|}{\|x\|^2} < +\infty$$

On suppose que $\lambda(E) < \infty$. Il y a équivalence entre

1. ϕ est un opérateur strictement convexe et s.c.i.
2. La fonction j est strictement convexe.

De plus, il existe un inverse à gauche $\gamma_g : R(\nabla\phi) \rightarrow X$ de $\partial\phi$. Il coïncide avec l'inverse de la dérivée de la fonction j là où elle existe.

Démonstration. L'opérateur ϕ est bien défini car d'après la définition de β

$$\forall \varepsilon > 0, \exists \tau > 0, \forall x \in \mathbb{R}^q, \|x\| \geq \tau \Rightarrow |j(x)| \leq (\beta + \varepsilon) \|x\|^2.$$

Il s'ensuit

$$\begin{aligned} |\phi(x)| &\leq \int_{\{\|x(\lambda)\| < \tau\}} |j(x(\lambda))| \, d\lambda + (\beta + \varepsilon) \int_{\{\|x(\lambda)\| \geq \tau\}} \|x(\lambda)\|^2 \, d\lambda < +\infty \\ &\leq \int_{\{\|x(\lambda)\| < \tau\}} |j(x(\lambda))| \, d\lambda + (\beta + \varepsilon) \|x\|_{L^2(E, \mathbb{R}^q)}^2 < +\infty. \end{aligned}$$

L'équivalence entre la stricte convexité de j et celle de ϕ est immédiate par positivité de l'intégrale dans un sens et en utilisant des fonctions constantes dans l'autre sens (hypothèse $\lambda(E) < \infty$). En utilisant le Théorème de Rademacher [4, 47] on en déduit que j est de classe C^1 presque partout et continue. En utilisant la même démonstration que pour l'entropie (lemme de Fatou), la continuité de j implique que ϕ est s.c.i.

Calculons le sous différentiel de ϕ , notons \mathcal{N} l'ensemble des points (de mesure nulle) où j n'est pas C^1 . On a

$$\partial\phi(x) = \partial j \circ x$$

et la fonction $s \mapsto \partial j \circ x(s)$ est multivaluée aux points de l'ensemble $\{s \in E; x(s) \in \mathcal{N}\}$. Comme ∂j est monotone univoque p.p, il existe un inverse à gauche univoque γ_g de ∂j et donc de $\partial\phi$. \square

Nous avons montré que si ϕ est convexe et s.c.i s'écrit sous la forme (4.13), alors on peut définir un inverse à gauche de son sous-différentiel. Remarquons que nous n'avons pas un opérateur ϕ Gâteaux-différentiable. On peut maintenant se demander si l'équation

$$x_\mu = \gamma_g\left(\frac{2}{\mu}\psi^*[\ell_\mu]\right) \text{ avec } \ell_\mu = y - \psi[x_\mu] \quad (4.14)$$

est une condition d'optimalité pour le problème (\mathcal{P}_μ) . C'est ce que nous allons voir dans la proposition suivante.

Proposition 11 *L'équation (4.14) est une condition nécessaire et suffisante d'optimalité pour le problème (\mathcal{P}_μ) .*

Démonstration. La condition d'optimalité du problème (\mathcal{P}_μ) s'écrit

$$-\psi^*[y - \psi[x]] + \mu\partial\phi(x) \ni 0.$$

Comme γ_g est un inverse à gauche de $\partial\phi$ on en déduit

$$x = \gamma_g(\partial\phi(x)) = \gamma_g\left(\frac{1}{\mu}\psi^*[y - \psi[x]]\right).$$

La relation (4.14) est donc une condition nécessaire d'optimalité. Montrons que c'est une condition suffisante. Pour cela réécrivons la condition d'optimalité à l'aide de l'équation (4.14)

$$-\psi^*[y - \psi[x]] + \mu\partial\phi\left(\gamma_g\left(\frac{1}{\mu}\psi^*[y - \psi[x]]\right)\right) \ni 0.$$

cela est vérifié puisque l'on a toujours $\partial\phi(\gamma_g(x)) \ni x$, donc

$$\mu\partial\phi\left(\gamma_g\left(\frac{1}{\mu}\psi^*[y - \psi[x]]\right)\right) \ni \mu\frac{1}{\mu}\psi^*[y - \psi[x]].$$

\square

4.5 Problèmes avec contraintes

Nous considérons désormais que l'ensemble convexe fermé des contraintes K du problème (\mathcal{P}_μ) n'est plus égal à X . Pour assurer l'existence d'une solution dans ce cas il suffit que la fonction ϕ soit coercive sur $N(\psi)$ ou que K soit borné. L'unicité de la solution dépend encore

une fois de la stricte convexité de ϕ . La condition d'optimalité du problème avec contraintes est

$$\forall h \in K, \langle -2\psi^*[y - \psi[x_\mu]] + \mu\nabla\phi(x_\mu), h - x_\mu \rangle_{X',X} \geq 0 \quad (4.15)$$

Si nous supposons que X est un espace de Hilbert, la condition d'optimalité (4.15) s'écrit sous la forme d'une projection

$$x_\mu = \pi_K(x_\mu - 2\psi^*[y - \psi[x_\mu]] + \mu\nabla\phi(x_\mu)).$$

On peut aussi formuler la condition d'optimalité, dans un Banach, comme une projection grâce à l'application de dualité Λ . On réécrit (4.15) sous la forme

$$\forall h \in K, \langle \Lambda(\Lambda^{-1}(-2\psi^*[y - \psi[x_\mu]] + \mu\nabla\phi(x_\mu)) + x_\mu - x_\mu), h - x_\mu \rangle_{X',X} \geq 0$$

d'où

$$x_\mu = \pi_K(x_\mu - \Lambda^{-1}(2\psi^*[y - \psi[x_\mu]] + \mu\nabla\phi(x_\mu))).$$

Mais en utilisant cette équation, nous ne pouvons pas espérer obtenir une solution de la forme (4.10). C'est pourquoi nous allons modifier le problème (\mathcal{P}_μ) pour revenir à une minimisation globale en posant

$$(\mathcal{P}_\mu^\varepsilon) \begin{cases} \min \|y - \psi[x]\|_X^2 + \mu\phi(x) + \phi_K^\varepsilon(x) \\ x \in X \end{cases}$$

où $\phi_K^\varepsilon : X \rightarrow \mathbb{R}$ est une approximation régulière de la fonction identité 1_K de l'ensemble K .

Une première façon de définir ϕ_K^ε est de poser

$$\phi_K^\varepsilon(x) = \frac{1}{\varepsilon} \text{dist}(x, K)^2$$

où $\text{dist}(x, K)$ désigne la distance de x à K , i.e. $\text{dist}(x, K) = \min\{\|x - y\|_X ; y \in K\}$. Cette distance est bien définie par convexité de K . Comme $x \mapsto \text{dist}(x, K)$ est continue et convexe $x \mapsto \mu\phi(x) + \phi_K^\varepsilon$ l'est également. On a le résultat suivant sur la convergence des solutions du problème $(\mathcal{P}_\mu^\varepsilon)$ vers les solutions de (\mathcal{P}_μ) ; c'est un résultat classique.

Proposition 12 Soit $\{x_\mu^\varepsilon\}$ une famille de solutions du problème $(\mathcal{P}_\mu^\varepsilon)$.

1. Si ϕ est strictement convexe, alors x_μ^ε converge faiblement dans X vers x_μ , la solution de (\mathcal{P}_μ) , quand $\varepsilon \rightarrow 0$.
2. Si ϕ est seulement convexe, la solution x_μ^ε converge faiblement dans X , à une sous suite près, vers une solution de (\mathcal{P}_μ)

Démonstration. Donnons une idée de la preuve dans le cas où K est borné et ϕ strictement convexe. Pour tout $\varepsilon > 0$, on a

$$\|y - \psi[x_\mu^\varepsilon]\|_X^2 + \mu\phi(x_\mu^\varepsilon) + \frac{1}{\varepsilon} \text{dist}(x_\mu^\varepsilon, K)^2 \leq \|y - \psi[x_\mu]\|_X^2 + \mu\phi(x_\mu)$$

d'où

$$0 \leq \text{dist}(x_\mu^\varepsilon, K)^2 \leq \varepsilon \left[\|y - \psi[x_\mu]\|_X^2 - \|y - \psi[x_\mu^\varepsilon]\|_X^2 + \mu\phi(x_\mu) - \mu\phi(x_\mu^\varepsilon) \right]$$

Le second membre tend vers 0. On en déduit, puisque K est borné, que la famille $\{x_\mu^\varepsilon\}_{\varepsilon>0}$ est bornée. Comme X est réflexif, on peut donc extraire de la famille $\{x_\mu^\varepsilon\}_{\varepsilon>0}$ une suite qui converge faiblement vers $\bar{x} \in K$, puisque $\text{dist}(x_\mu^\varepsilon, K) \rightarrow 0$ quand $\varepsilon \rightarrow 0$. Par semi-continuité inférieure de la fonction coût, on déduit que

$$\|y - \psi[\bar{x}]\|_X^2 + \mu\phi(\bar{x}) \leq \|y - \psi[x_\mu]\|_X^2 + \mu\phi(x_\mu)$$

ce qui montre que $\bar{x} = x_\mu$ par unicité. \square

Une alternative au choix de la fonction $\phi_K^\varepsilon : X \rightarrow \mathbb{R}$ est de choisir une fonction qui, pour tout $\varepsilon > 0$, a pour domaine de définition $\text{dom } \phi_K^\varepsilon = K$ et qui vérifie l'hypothèse (\mathcal{H}_1) . Par exemple :

$$\phi_K^\varepsilon(x) = \begin{cases} -\varepsilon \ln(\text{dist}(x, \text{Fr}(K))) & \text{si } x \in K \\ +\infty & \text{sinon} \end{cases},$$

en supposant que K est assez régulier pour qu'il n'y ait pas de problème de définition et que $\overset{\circ}{K} \neq \emptyset$. L'utilisation d'une telle fonctionnelle va éliminer la contrainte et la solution va rester dans l'intérieur de K (contrairement au cas précédent), quelle que soit la valeur du paramètre $\varepsilon > 0$.

On peut démontrer le même résultat de convergence vers la solution de (\mathcal{P}_μ) quand $\varepsilon \rightarrow 0$ que dans la proposition précédente. Voici une brève description de la preuve. Avec les mêmes notations que pour la démonstration de la proposition, supposons que K est un ensemble borné et que pour $\mu > 0$ fixé la solution x_μ de (\mathcal{P}_μ) est atteinte au bord de K . Soit une suite $\{x_n\}_{n \in \mathbb{N}} \subset \overset{\circ}{K}$ qui converge fortement vers x_μ et telle que pour tout $n \in \mathbb{N}$, $\text{dist}(x_n, \text{Fr}(K)) > 0$. Pour tout $\varepsilon > 0$, on a

$$\|y - \psi[x_\mu^\varepsilon]\|_X^2 + \mu\phi(x_\mu^\varepsilon) + \phi_K^\varepsilon(x_\mu^\varepsilon) \leq \|y - \psi[x_n]\|_X^2 + \mu\phi(x_n) + \phi_K^\varepsilon(x_n)$$

Posons $\varepsilon_n = \text{dist}(x_n, \text{Fr}(K)) > 0$, comme $x_n \rightarrow x_\mu$ et que x_μ est sur le bord de K , on en déduit que $\varepsilon_n \rightarrow 0$, et

$$\lim_{n \in \mathbb{N}} \phi_K^{\varepsilon_n}(x_n) = \lim_{n \in \mathbb{N}} -\varepsilon_n \ln(\varepsilon_n) = 0.$$

Comme la famille $\{x_\mu^{\varepsilon_n}\}_{n \in \mathbb{N}} \subset K$ et que K est borné et fermé on peut en extraire une sous-suite qui converge faiblement vers $\bar{x} \in K$, d'où

$$\|y - \psi[\bar{x}]\|_X^2 + \mu\phi(\bar{x}) + \lim_{n \in \mathbb{N}} \phi_K^{\varepsilon_n}(x_\mu^{\varepsilon_n}) \leq \|y - \psi[x_\mu]\|_X^2 + \mu\phi(x_\mu)$$

Comme x_μ est la solution du problème avec contraintes, on a

$$\lim_{n \in \mathbb{N}} \phi_K^{\varepsilon_n}(x_\mu^{\varepsilon_n}) \leq \|y - \psi[x_\mu]\|_X^2 + \mu\phi(x_\mu) - \|y - \psi[\bar{x}]\|_X^2 - \mu\phi(\bar{x}) \leq 0.$$

Le domaine K étant borné, nous avons la majoration

$$\forall x \in \overset{\circ}{K}, \phi_K^\varepsilon(x) \geq -\varepsilon \ln(\text{diam } K)$$

d'où

$$\forall x \in \overset{\circ}{K}, \lim_{\varepsilon \rightarrow 0} \phi_K^\varepsilon(x) \geq 0$$

Finalement, la suite $x_\mu^{\varepsilon_n}$ converge faiblement vers $\bar{x} \in K$ et

$$\|y - \psi[\bar{x}]\|_X^2 + \mu\phi(\bar{x}) \leq \|y - \psi[x_\mu]\|_X^2 + \mu\phi(x_\mu),$$

par unicité du minimum on en déduit que $\bar{x} = x_\mu$.

4.6 Résolution

Comme nous l'avons dit, notre but est de déterminer une solution du problème d'optimisation qui s'écrit sous la forme (4.10). Dans les sections précédentes, nous avons regardé les différents cas possibles pour lesquels il est possible de rechercher une telle solution. Dans cette section nous allons simplement nous concentrer sur la résolution du problème de point fixe défini par l'équation (4.10). Nous allons envisager la résolution sous deux points de vue différents, celui des théorèmes de point fixe de Schauder et de Banach.

L'équation (4.10) est un problème de point fixe en dimension infinie. Nous pouvons toutefois revenir à un problème en dimension finie en utilisant l'opérateur adjoint ψ^* [42]. Plus précisément, nous allons chercher la solution x sous la forme

$$x_\mu = \gamma \left(\frac{2}{\mu} \psi^*[\ell_\mu] \right)$$

où $\ell_\mu \in \mathbb{C}^n$ est à déterminer comme la solution de

$$\ell_\mu = y - \psi \left[\gamma \left(\frac{2}{\mu} \psi^*[\ell_\mu] \right) \right] =: p_\mu(\ell_\mu) \quad (4.16)$$

La fonction p_μ est bien une fonction de \mathbb{C}^n dans \mathbb{C}^n . Pour ne pas avoir de problème de définition, nous supposons qu'il existe $\mu_0 > 0$ et $R > \|y\|_{\mathbb{C}^n}$

$$\frac{2}{\mu_0} \psi^*[B(0, R)] \subset R(\nabla\phi) \quad (4.17)$$

où $B(0, R)$ est la boule de rayon R . Remarquons que pour tout $\mu > 0$, sous les hypothèses de la Proposition (7), la fonction p_μ a toujours un point fixe et donc $R(\nabla\phi) \cap B(0, R) \neq \emptyset$. Remarquons également que si l'hypothèse (4.17) est vérifiée, alors elle l'est aussi pour tout $\mu > \mu_0$.

Il paraît naturel que la fonction p_μ ait un point fixe pour $\mu > 0$ assez grand, ce qui permet de trouver une solution au problème sous la forme (4.10). Il faut donc trouver un moyen pour calculer ce point fixe. On peut donc s'intéresser à la suite définie par approximations successives

$$l_0 \in B(0, R), \text{ et } l_{k+1} = p_\mu(l_k). \quad (4.18)$$

Si cette suite converge, sa limite est nécessairement un point fixe de p_μ . Cependant, avec cette suite, deux termes consécutifs peuvent être relativement éloignés. C'est pourquoi nous allons surtout étudier la suite "relaxée" $\{\lambda_k\}_{k \in \mathbb{N}}$ définie pour $\tau \in]0, 1]$ par

$$\lambda_0 \in B(0, R), \text{ et } \lambda_{k+1} = (1 - \tau)\lambda_k + \tau p_\mu(\lambda_k). \quad (4.19)$$

4.6.1 Lien avec la dualité

Proposition 13 *Le problème dual (\mathcal{D}_μ) de (\mathcal{P}_μ) est*

$$\sup_{\ell \in \mathbb{C}^n} \left\{ - \left\langle \psi^*[\ell], \gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right\rangle_X + \mu \phi \left(\gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right) - \frac{1}{4} \|\ell\|_{\mathbb{C}^n}^2 + \langle \ell, y \rangle_{\mathbb{C}^n} \right\}.$$

La preuve de cette proposition est identique à celle que nous avons faite pour la régularisation entropique. A l'aide de la dualité de Fenchel-Rockafellar. La fonction que nous souhaitons minimiser peut s'écrire sous la forme :

$$J_\mu(x) = G(\psi[x]) + F(x),$$

où $G(x) = \|y - x\|_{\mathbb{C}^n}^2$ et $F(x) = \mu\phi(x)$. Dans ce cas, le problème dual est [15] :

$$\sup_{\ell \in \mathbb{C}^n} [-G^*(-\ell) - F^*(\psi^*[\ell])].$$

La conjuguée de G à été calculée dans le chapitre 3, et s'écrit

$$G^*(\ell) = \frac{1}{4} \|\ell\|_{\mathbb{C}^n}^2 + \langle \ell, y \rangle_{\mathbb{C}^n}.$$

Calculons la conjuguée de F ; en dérivant $x \mapsto \langle \ell, x \rangle - \mu\phi(x)$ on en déduit que cette fonction atteint son maximum pour

$$\mu \nabla \phi(x) = \ell \Leftrightarrow x = \gamma \left(\frac{\ell}{\mu} \right),$$

d'où

$$F^*(\ell) = \left\langle \ell, \gamma \left(\frac{\ell}{\mu} \right) \right\rangle_X - \mu \phi \left(\gamma \left(\frac{\ell}{\mu} \right) \right).$$

Le problème dual s'écrit donc sous la forme

$$(\mathcal{D}_\mu) \quad \sup_{\ell \in \mathbb{C}^n} \left\{ - \left\langle \psi^*[\ell], \gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right\rangle_X + \mu \phi \left(\gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right) - \frac{1}{4} \|\ell\|_{\mathbb{C}^n}^2 + \langle \ell, y \rangle_{\mathbb{C}^n} \right\}.$$

Soit la fonction $m : \mathbb{C}^n \rightarrow \mathbb{R}$ définie par

$$m : \ell \mapsto - \left\langle \psi^*[\ell], \gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right\rangle_X + \mu \phi \left(\gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right) - \frac{1}{4} \|\ell\|_{\mathbb{C}^n}^2 + \langle \ell, y \rangle_{\mathbb{C}^n}$$

Nous voulons déterminer l'unique $\ell \in \mathbb{C}^n$ tel que $\nabla m(\ell) = 0$; en calculant $\nabla m(\ell)$ on obtient

$$\nabla m(\ell) = -\frac{1}{2}\ell + y - \psi \left[\gamma \left(\frac{\psi^*[\ell]}{\mu} \right) \right].$$

D'après le calcul précédent et par un changement de variable, le calcul de la solution de $\nabla m(\ell) = 0$ revient à calculer la solution de l'équation

$$\ell - y + \psi \left[\gamma \left(\frac{2}{\mu} \psi^*[\ell] \right) \right] = 0. \quad (4.20)$$

Ecrivons la méthode du gradient pour résoudre (4.20), on calcule les termes de la suite $\{l_k\}_{k \in \mathbb{N}}$ avec $l_0 \in \mathbb{C}^n$ et

$$\begin{aligned} l_{k+1} &= l_k - \beta \left(l_k - y = \psi \left[\frac{2}{\mu} \psi^*[l_k] \right] \right) \\ l_{k+1} &= l_k - \beta l_k + \beta (y - \psi \left[\gamma \left(\frac{2}{\mu} \psi^*[l_k] \right) \right]) \\ l_{k+1} &= (1 - \beta)l_k + \beta p_\mu(l_k) \end{aligned}$$

Nous en concluons que l'algorithme définissant la suite $\{\lambda_k\}_{k \in \mathbb{N}}$, définie par (4.19), n'est rien d'autre qu'une méthode de descente sur le problème dual (à une constante multiplicative près).

4.6.2 Utilisation du théorème de Banach

Nous voulons calculer le point fixe de la fonction

$$p_\mu : l \mapsto y - \psi \left[\gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right].$$

Pour cela, nous allons utiliser le théorème de point fixe de Banach afin de montrer la convergence de la suite (4.19) vers ce point fixe. En conséquence, l'existence et l'unicité du point fixe de la fonction p_μ pour toutes valeurs de $\mu > 0$, seront à nouveau démontrés. Pour utiliser ce théorème il faut que la fonction $p_\mu : \mathbb{C}^n \rightarrow \mathbb{C}^n$ soit assez régulière. Nous supposons donc que l'inverse $\gamma : X' \rightarrow X$ est C^1 . La différentielle de la fonction p_μ en $l \in \mathbb{C}^n$ est

$$\forall h \in \mathbb{C}^n, \nabla p_\mu(l).h = -\frac{2}{\mu} \psi \left[\nabla \gamma \left(\frac{2}{\mu} \psi^*[l] \right). \psi^*[h] \right].$$

Proposition 14 *Supposons que l'hypothèse (4.17) est vérifiée et qu'il existe une constante C telle que pour tout $l \in B(0, R)$*

$$\left\| \psi \left[\gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right] \right\|_{\mathbb{C}^n} \leq \frac{C}{\mu} \|l\|_{\mathbb{C}^n}. \quad (4.21)$$

Si $\|l\|_{\mathbb{C}^n} \leq R$ et si

$$\mu \geq \frac{CR}{R - \|y\|_{\mathbb{C}^n}} \quad (4.22)$$

alors $\|p_\mu(l)\|_{\mathbb{C}^n} \leq R$; il existe $\tau > 0$ assez petit tel que la suite définie par (4.19) converge vers le point fixe de p_μ ce qui permet de trouver la solution de (\mathcal{P}_μ) grâce à (4.10).

Démonstration. En utilisant (4.21) on obtient

$$\begin{aligned} \|p_\mu(l)\|_{\mathbb{C}^n} &\leq \|y\|_{\mathbb{C}^n} + \left\| \psi \left[\gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right] \right\|_{\mathbb{C}^n} \\ &\leq \|y\|_{\mathbb{C}^n} + \frac{1}{\mu} C \|l\|_{\mathbb{C}^n} \leq \|y\|_{\mathbb{C}^n} + \frac{1}{\mu} CR \end{aligned}$$

La majoration de $p_\mu(l)$ par R se déduit de la dernière équation grâce à l'inégalité (4.22). Ainsi la boule de rayon R est stable par p_μ . Nous allons montrer que l'on peut choisir $0 < \tau < 1$ tel que la fonction qui définit la suite (4.19) est une contraction, ce qui prouvera la convergence vers un unique point fixe par le théorème de Banach. Pour tout $l \in \mathbb{C}^n$, $\nabla p_\mu(l)$ est une matrice négative. En effet par positivité de $\nabla \gamma$ (puisque ϕ est convexe), on a

$$\left\langle \psi \left[\nabla \gamma \left(\frac{2}{\mu} \psi^*[l] \right) \cdot \psi^*[h] \right], h \right\rangle_{\mathbb{C}^n} = \left\langle \nabla \gamma \left(\frac{2}{\mu} \psi^*[l] \right) \cdot \psi^*[h], \psi^*[h] \right\rangle_{\mathbb{C}^n} \geq 0.$$

En procédant comme dans [42], on majore le rayon spectral ρ de $-\nabla p$ à l'aide de la norme de Frobenius

$$\|\nabla p_\mu(l)\|_F = \frac{2}{\mu} \sqrt{\sum_{1 \leq i, j, n} \left| \left\langle \nabla \gamma \left(\frac{2}{\mu} \psi^*[l] \right) \cdot \psi^*[e_i], \psi^*[e_j] \right\rangle_{\mathbb{C}^n} \right|^2}$$

où $\{e_i\}_{1 \leq i \leq n}$ désigne la base canonique de \mathbb{C}^n . Enfin, puisque $-\nabla p_\mu$ est une matrice positive et qu'on dispose d'une minoration de sa plus petite valeur propre, il existe $0 < \tau < 1$ tel que $l \mapsto (1 - \tau)l + \tau p_\mu(l)$ soit une contraction. La suite (4.19) converge donc vers l'unique point fixe de la fonction p_μ . \square

Cette proposition montre l'existence du point fixe de la fonction p_μ ainsi que la convergence de la suite $\{\lambda_k\}_{k \in \mathbb{N}}$ vers ce point fixe si le paramètre μ est choisi assez grand. Toutefois, d'après la Proposition 7, nous savons que la fonction p_μ a un point fixe pour toute valeur de $\mu > 0$. Mais nous pouvons également le démontrer en modifiant légèrement la suite $\{\lambda\}_{k \in \mathbb{N}}$. En effet d'après la proposition précédente, il existe $\mu > 0$ assez grand, et $\tau > 0$ assez petit tels que $l \mapsto (1 - \tau)l + \tau p_\mu(l)$ soit une contraction. La matrice $(1 - \tau)I + \tau \nabla p_\mu(l)$ a donc des valeurs propres plus petites que 1 en valeur absolue. Or comme la matrice $\nabla p_\mu(l)$ est négative, pour $\alpha > 1$ les valeurs propres de $(1 - \frac{\tau}{\alpha})I + \tau \nabla p_\mu(l)$ sont également inférieures à 1. La fonction $(1 - \frac{\tau}{\alpha})l + \tau \nabla p_\mu(l)$ est donc une contraction, elle a un point fixe $l \in \mathbb{C}^n$ et

$$\begin{aligned} l &= \left(1 - \frac{\tau}{\alpha}\right)l + \tau p_\mu(l) \\ \frac{l}{\alpha} &= p_\mu \left(\frac{l}{\alpha}\right) \end{aligned}$$

ce n'est rien d'autre que la condition d'optimalité du problème (\mathcal{P}_μ) et on pose $\ell_\mu = \frac{l}{\alpha}$.

4.6.3 Premières applications

La régularisation entropique

Dans [42] nous avons utilisé l'entropie comme opérateur de régularisation. Plus précisément, on pose $E = [0, 1]^2$, et on considère $\phi : L^2(E, \mathbb{R}, \lambda) \rightarrow \mathbb{R}$ défini par

$$\phi(x) = \int_E j \circ x \, d\lambda, \text{ avec } \begin{cases} j : \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto x \ln x \end{cases} ;$$

son domaine de définition est

$$\text{dom } \phi = \{x \in L^2(E, \mathbb{R}, \lambda) \mid x \geq 0 \text{ } \lambda.p.p.\}.$$

L'opérateur d'entropie $\phi : L^2(E, \mathbb{R}, \lambda) \rightarrow \mathbb{R}$ est strictement convexe et s.c.i. [1, 16, 42]. Il vérifie de plus l'hypothèse (\mathcal{H}'_1) : son sous-différentiel est univoque pour $x \in (\text{dom } \phi)^\circ$ et $\partial\phi(x) = \{1 + \ln x\}$. Les solutions sont donc positives. Par contre, l'entropie n'est pas coercive sur $L^2(E, \mathbb{R}, \lambda)$, mais on peut tout de même montrer que (4.9) est une condition nécessaire et suffisante d'optimalité sous quelques hypothèses sur ψ [42].

L'inverse $\gamma : L^2(E, \mathbb{R}, \lambda) \rightarrow \text{dom } \phi$ est localement lipschitzienne puisque c'est la fonction $x \mapsto \exp(-1 + x)$ et la condition (4.17) est clairement vérifiée. Nous pouvons donc utiliser la Proposition 14, la suite $\{\lambda_k\}_{k \in \mathbb{N}}$ converge vers la solution du problème. Cette méthode de régularisation entropique a été utilisée pour résoudre un problème inverse en physique spatiale, afin d'interpréter correctement des résultats expérimentaux [44, 43]

Minimisation des dérivées partielles

Soit l'espace de Hilbert $X = H_0^1(\Omega)$ avec Ω borné et assez régulier. Nous considérons le problème d'optimisation suivant :

$$(P_\mu^2) \begin{cases} \min \|y - \psi[f]\|_{C^n}^2 + \mu \sum_{j=1}^n \|\partial_{x_j} f\|_{L^2}^2 \\ f \in X \end{cases} .$$

La fonction coût est continue, strictement convexe, coercive et différentiable. Il existe donc une unique solution à ce problème. Comme c'est une minimisation globale la condition d'optimalité est

$$-\psi^*[y - \psi[f]] + \mu \Delta f = 0$$

Le laplacien est l'injection canonique de X dans son dual $X' = H^{-1}(\Omega)$. Son inverse est continu et s'écrit

$$\gamma : f' \mapsto \int_{\Omega} f'(x) G(x, y) dy$$

où $G : \Omega \times \Omega \rightarrow \mathbb{R}$ est la fonction de Green de l'ensemble Ω [18], i.e. $\Delta G(x, y) = \delta_0(x - y)$ au sens des distributions et sa trace est nulle sur le bord de $\Omega \times \Omega$. L'application γ est clairement lipschitzienne de constante C_γ : l'inverse de la première valeur propre du Laplacien sur Ω . Les hypothèses des Propositions 8 et 14 sont donc vérifiées, et on peut déterminer la solution du problème (P_μ^2) en utilisant la suite (4.19).

On peut remarquer que si on prend l'espace $H^2(\Omega)$ à la place de $H_0^1(\Omega)$, la fonction régularisante n'est plus strictement convexe et l'opérateur inverse est multivoque. On peut toutefois rendre l'opérateur univoque en fixant la trace de la solution aux bords.

Limitations : la méthode de Tikhonov dans les espaces L^p

Nous allons regarder la méthode de Tikhonov dans les espaces $X = L^p(E, \mathbb{R}, \lambda)$. C'est à dire

$$(P_2) \begin{cases} \min \|y - \psi[x]\|_X^2 + \mu \|x\|_p \\ x \in X \end{cases} .$$

Nous supposons que $p > 2$ est un nombre pair, afin que $x \mapsto x^p$ soit strictement convexe. Ce problème est équivalent à

$$(P'_2) \left\{ \begin{array}{l} \min \|y - \psi[x]\|_X^{2p} + \mu \int x^p d\sigma \\ x \in X \end{array} \right. .$$

La fonction coût est continue, strictement convexe, Gâteaux différentiable. La condition d'optimalité du problème est

$$-2 \|y - \psi[x]\|_{\mathbb{C}^n}^{2(p-1)} \psi^*[y - \psi[x]] + \mu x^{p-1} = 0 \text{ dans } X'$$

Comme la fonction $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ définie par $\gamma(x) = \sqrt[p-1]{x}$ est un inverse de x^{p-1} , la condition d'optimalité est équivalente à

$$x = \sqrt[p-1]{\frac{2 \|y - \psi[x]\|_{\mathbb{C}^n}^{2(p-1)} \psi^*[y - \psi[x]]}{\mu}} \quad p.p$$

Toutefois les hypothèses de la Proposition 8 ne sont pas vérifiées car la fonction $x \mapsto x^{p-1}$ ne vérifie pas la condition de stabilité (4.11). En effet $x \mapsto \sqrt[p-1]{x}$ n'est pas Lipschitzienne en $x = 0$, l'ensemble des solutions instables est donc $I = \{0\}$.

Nous ne pouvons donc pas utiliser la Proposition 14 pour montrer que la suite 4.19 converge vers une solution. Il faut donc trouver un équivalent de la Proposition 14 sous des hypothèses plus faibles sur la régularité de γ . Pour cela nous allons utiliser le théorème de Schauder.

4.6.4 Utilisation du théorème de Schauder

Nous savons déjà que la fonction p_μ définie par (4.16) a un point fixe sous les hypothèses de la Proposition 7. Nous allons affiner ce résultat en montrant plus précisément que p_μ laisse la boule fermée de rayon R , $\bar{B}(0, R)$ stable pour R et μ assez grands, dans une perspective numérique (utilisation d'une suite restant dans cette boule pour le calcul du point fixe).

On définit la fonction $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ par

$$f(\delta) = \sup_{\bar{B}(0, \delta)} \|\gamma(x)\|_X .$$

Pour que cette fonction soit bien définie il faut supposer que

$$\|\gamma(x)\|_X \geq \limsup_{c \rightarrow x} \|\gamma(c)\|_X \tag{4.23}$$

Une condition suffisante pour cela est que $x \mapsto \|\gamma(x)\|_X$ soit s.c.s. Si la fonction γ est continue, alors cette hypothèse est clairement vérifiée. Nous allons maintenant définir une fonction qui va jouer le rôle d'un "inverse" de f . On définit $f^\dagger : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \cup +\infty$ par

$$\forall R \in \mathbb{R}^+, \quad f^\dagger(R) = \sup\{\delta \geq 0 \mid f(\delta) \leq R\}$$

Lemme 1 *Supposons que $\gamma : X \rightarrow X'$ est continue. Alors f est continue, croissante et*

$$\delta \leq f^\dagger(R) \Leftrightarrow f(\delta) \leq R$$

Démonstration. La fonction f est croissante puisque qu'on prend le sup sur des ensembles de plus en plus grands. Supposons qu'elle n'est pas continue : il existe $\varepsilon > 0$ tel que pour tout $\eta > 0$ on peut trouver $0 \leq \delta \leq \delta'$ avec $|\delta - \delta'| \leq \eta$ et $f(\delta') - f(\delta) \geq \varepsilon$. Comme γ est continue, le sup est atteint, il existe $x_1 \in \bar{B}(0, \delta)$ et $x_2 \in \bar{B}(0, \delta')$ tel que

$$\begin{aligned} \varepsilon &\leq f(\delta') - f(\delta) \leq \|\gamma(x_2)\|_X - \|\gamma(x_1)\|_X \\ &\leq \|\gamma(x_2)\|_X - \left\| \gamma\left(\frac{\delta}{\delta'}x_2\right) \right\|_X \\ &\leq \left\| \gamma(x_2) - \gamma\left(\frac{\delta}{\delta'}x_2\right) \right\|_X \end{aligned}$$

Cela signifie qu'il existe ε , tel que pour tout $\eta > 0$, il existe $x \in \bar{B}(x_2, \eta)$ et

$$\|\gamma(x_2) - \gamma(x)\|_X \geq \varepsilon.$$

Ce qui contredit la continuité de γ au point x_2 . La fonction f est donc continue.

Le sens (\Leftarrow) est immédiat. Montrons la réciproque. Si $f^\dagger(R) = +\infty$ le résultat est trivial. Si $f^\dagger(R) < +\infty$, il existe δ_0 tel que $f(\delta_0) > R$ d'où, par croissance de f

$$f^{-1}(]-\infty, R]) \subset [0, \delta_0[.$$

L'ensemble $f^{-1}(]-\infty, R])$ est donc borné; comme f est continue, c'est aussi un ensemble fermé : il est donc compact. L'ensemble $\{\delta > 0 \mid f(\delta) \leq R\}$ est compact, et donc son sup est atteint. Il s'ensuit $f(f^\dagger(R)) = R$ par croissance de f . \square

Lemme 2 *Soient $C_\psi > 0$ et $C_{\psi^*} > 0$ deux constantes définies par*

$$C_\psi = \|\psi\|_{\mathcal{L}(X, \mathbb{C}^n)} \quad \text{et} \quad C_{\psi^*} = \|\psi^*\|_{\mathcal{L}(\mathbb{C}^n, X')}$$

Supposons la fonction f bien définie. Si $\|l\|_{\mathbb{C}^n} \leq R$ et si

$$\mu \geq 2C_{\psi^*}R \left[f^\dagger\left(\frac{R - \|y\|}{C_\psi}\right) \right]^{-1} \quad (4.24)$$

alors $\|p_\mu(l)\|_{\mathbb{C}^n} \leq R$

Démonstration. On a la majoration suivante

$$\left\| \frac{2}{\mu} \psi^*[l] \right\|_{X'} \leq \frac{2}{\mu} C_{\psi^*} \|l\|_{\mathbb{C}^n} \leq \frac{2}{\mu} C_{\psi^*} R.$$

L'inégalité précédente et (4.24) impliquent que

$$\left\| \frac{2}{\mu} \psi^*[l] \right\| \leq f^\dagger\left(\frac{R - \|y\|}{C_\psi}\right);$$

en appliquant le lemme précédent on obtient

$$\begin{aligned} \left\| \gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right\|_X &\leq f \left(\left\| \frac{2}{\mu} \psi^*[l] \right\|_{X'} \right) \leq f \left(f^\dagger \left(\frac{R - \|y\|}{C_\psi} \right) \right) = \left(\frac{R - \|y\|}{C_\psi} \right) \\ \left\| \gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right\|_X &\leq \left(\frac{R - \|y\|}{C_\psi} \right) \end{aligned}$$

donc

$$\|p_\mu(l)\| \leq \|y\| + C_\psi \left\| \gamma \left(\frac{2}{\mu} \psi^*[l] \right) \right\| \leq R$$

□

Un corollaire immédiat de ce Lemme et du Théorème de Schauder est le théorème suivant

Théorème 12 *Sous les conditions du Lemme 2, la fonction $p_\mu : \bar{B}(0, R) \rightarrow \bar{B}(0, R)$, définie par (4.16), a un point fixe $\ell_\mu \in B(0, R)$. La solution du problème (\mathcal{P}_μ) est donc*

$$x_\mu = \gamma \left(\frac{2}{\mu} \psi^*[\ell_\mu] \right)$$

Démonstration. On utilise le Lemme 2 avec le théorème de Schauder que nous rappelons ci-dessous :

Théorème 13 (Schauder) *Soit M un convexe fermé non vide d'un espace de Banach X . On suppose que $T : M \rightarrow M$ est une fonction continue. Alors T a un point fixe.*

Pour la démonstration de ce théorème voir [69].

□

Pour faire le calcul effectif de ce point fixe, nous allons utiliser la fonction $g_\mu : \bar{B}(0, R) \rightarrow \bar{B}(0, R)$ et la suite $\{\lambda_k\}_{k \in \mathbb{N}}$ définie par (4.19), i.e.

$$\lambda_0 \in B(0, R), \text{ et } \lambda_{k+1} = g_\mu(\lambda_k).$$

Il est clair que les fonctions g_μ et p_μ ont les mêmes points fixes. Il reste à montrer la convergence de cette suite. Ce point est pour l'instant encore ouvert.

4.6.5 La méthode de Tikhonov dans L^p : Conclusion

Revenons à la méthode de Tikhonov dans les L^p , on pose

$$\begin{cases} \xi_\mu : \mathbb{C}^n &\longrightarrow L^p \\ l &\longmapsto p^{-1} \sqrt{\frac{2\|l\|^{2(p-1)}}{\mu}} \psi^*[l] \end{cases}$$

et

$$\begin{cases} p_\mu : \mathbb{C}^n &\longrightarrow \mathbb{C}^n \\ l &\longmapsto y - \psi[\xi_\mu(l)] \end{cases} .$$

Nous allons montrer l'existence d'un point fixe de p_μ , pour μ assez grand, par le Théorème de Schauder. Pour cela, nous allons montrer que la boule $\bar{B}(0, R)$ est stable par la fonction p_μ , ce qui est une première étape vers les calcul numérique du point fixe.

Proposition 15 *Si $\|l\| \leq R$ et si*

$$\mu \geq 2C_{\psi^*} R^{2p-1} \left(\frac{R - \|y\|}{C_{\psi}} \right)^{1-p},$$

alors $\|p(l)\| \leq R$. La solution du problème (\mathcal{P}_2) est dans ce cas donnée par

$$x = \sqrt[p-1]{\frac{2 \|\ell_{\mu}\|^{2(p-1)}}{\mu} \psi^*[\ell_{\mu}]} \quad p.p.$$

où ℓ_{μ} est l'unique point fixe de p_{μ} .

Démonstration. Le fait que \bar{B} soit stable par p_{μ} si μ est assez grand se démontre de manière identique à ce qui précède. Il suffit de voir que f^{\dagger} est la fonction puissance $p-1$. \square

Nous savons que la fonction p_{μ} a au moins un point fixe dans la boule $\bar{B}(0, R)$, ce point fixe ℓ_{μ} est unique par stricte convexité de la fonction coût. Voyons s'il est possible de revenir au cas localement lipschitzien pour calculer ℓ_{μ} . Nous allons supposer qu'il existe un ensemble compact $\mathcal{C} \subset \bar{B}(0, R)$ et un $\rho > 0$ tel que

$$\forall \ell \in \mathcal{C}, \psi^*[\ell] \geq \rho p.p.,$$

et un μ_0 tel que

$$\{\ell_{\mu} \mid \mu \geq \mu_0 > 0\} \subset \mathcal{C}.$$

Pour $p > 2$ on en déduit

$$\forall \ell \in \mathcal{C}, (\psi^*[\ell])^{\frac{2-p}{p-1}} \leq \rho^{\frac{2-p}{p-1}}. \quad (4.25)$$

La différentielle de ξ_{μ} s'écrit

$$\nabla \xi_{\mu}[\ell].h = 4 \langle \ell, h \rangle \sqrt[p-1]{\frac{\psi^*[\ell]}{\mu}} + 2 \|\ell\|^2 \left(\frac{\psi^*[\ell]}{\mu} \right)^{\frac{2-p}{p-1}}.$$

En supposant que pour tout $\ell \in \mathcal{C}$ on a $\psi^*[\ell] \in L^{\infty}$ et d'après (4.25), on en déduit que ξ_{μ} est lipschitzienne sur \mathcal{C} :

$$|\nabla \xi_{\mu}[\ell].h| \leq 4 \|l\| \|h\| \sqrt[p-1]{\frac{\|\psi^*[\ell]\|_{L^{\infty}}}{\mu}} + 2 \|\ell\|^2 \left(\frac{\rho}{\mu} \right)^{\frac{2-p}{p-1}}.$$

Sous ces hypothèses, la convergence de la suite $\{\lambda_k\}_{k \in \mathbb{N}}$ définie par (4.19), se montre de la même manière que dans la Proposition 14. On montre que $g_{\mu}(\ell) = (1 - \tau)\ell + \tau p_{\mu}(\ell)$ est contractante pour $\mu \geq \mu_0$ et pour τ assez petit, en majorant la plus grande valeur propre de la matrice positive $-\nabla p_{\mu}$. Un exemple de résolution numérique de ce problème sera donné au Chapitre 6.

4.6.6 Approximation de L^∞ par des normes L^p

Pour finir ce chapitre, on étudie une variante du problème (\mathcal{P}_2) étudiée par *Attouch et Cominetti* [3], défini par

$$(\mathcal{P}_3) \begin{cases} \min \|y - \psi[x]\|^2 + \frac{1}{p} \|x\|_p^p \\ x \in L^p \end{cases}$$

où $p \geq 2$ est un nombre pair. Ici il n'y a pas de paramètre de régularisation μ car nous voulons approcher un problème variationnel dans L^∞ . Montrons que nous pouvons trouver une solution de ce problème sous la forme (4.10) pour p assez grand.

Lemme 3 *Il existe un p tel que la fonction $\beta_p : l \mapsto y - \psi(\sqrt[p-1]{2\psi^*(l)})$ ait un point fixe l_p .*

Démonstration. supposons que $\|l\| \leq R$ et $R > \|y\|$, alors

$$\|\beta_p\| \leq \|y\| + C_\psi \sqrt[p-1]{2C_{\psi^*}R}$$

donc une condition pour que la boule $\bar{B}(0, R)$ soit stable par β_p est

$$(p-1) \geq \ln(2C_{\psi^*}R) \left(\ln \left(\frac{R - \|y\|}{C_\psi} \right) \right)^{-1}$$

(R est indépendant de p). □

Nous avons montré que pour p assez grand, la solution du problème s'écrivait sous la forme

$$\sqrt[p-1]{2\psi^*[l_p]} \tag{4.26}$$

où les points fixes l_p des fonction β_p sont dans la boule $\bar{B}(0, R)$. La suite $\{l_p\}$ est donc bornée on en déduit par continuité de ψ^* que la suite $\{\sqrt[p-1]{2\psi^*[l_p]}\}$ est bornée dans L^p . Nous pouvons en extraire une sous-suite qui converge faiblement dans tous les $L^p, p > 1$ vers $\bar{u} \in L^\infty$.

En supposant qu'il existe un $c > 0$ tel que pour tout $p > 2$,

$$|\psi^*[l_p]| < c, \quad p.p.$$

l'élément \bar{u} est admissible pour le problème dans L^∞ , i.e. $\|\bar{u}\|_{L^\infty} \leq 1$, puisque

$$\limsup |\sqrt[p-1]{2\psi^*[l_p]}| \leq 1 \quad p.p.$$

. De plus

$$\forall v \in L^\infty, \quad \|y - \psi[\bar{u}]\| \leq \liminf J_p(l_p) \leq \liminf J_p(v) = \|y - \psi[v]\|$$

Cela correspond au corollaire 3.6 du papier d'*Attouch et Cominetti* [3]. Nous ne pouvons cependant pas passer à la limite quand $p \rightarrow +\infty$ dans l'expression (4.26). En effet la suite de fonction $\{\sqrt[p-1]{\cdot}\}$ converge simplement vers la fonction sign, mais elle ne converge pas uniformément puisque la fonction sign est discontinue.

Deuxième partie

**Applications et simulation
Numérique**

Chapitre 5

Calcul numérique de la fonction de distribution des ondes : cas du vide et des plasmas

5.1 Introduction

Dans ce chapitre, nous allons mettre en oeuvre les algorithmes de régularisation entropique décrits dans le chapitre 3 et dans [42]. Le problème considéré est celui de la détermination de la fonction de distribution des ondes (FDO), à partir de la matrice spectrale mesurée du champ électromagnétique. Rappelons que la FDO est une fonction F définie sur la sphère unité S à valeurs positives. La matrice spectrale est reliée à la FDO par l'équation (1.1) donnée en page 2. Dans la suite nous identifierons toujours la matrice spectrale à un vecteur de \mathbb{C}^n où n est le nombre de coefficients de la matrice. La raison de cette identification vient tout simplement du fait que nous n'utiliserons jamais d'opérations matricielles. Pour parcourir la sphère unité S de \mathbb{R}^3 , nous utiliserons l'homéomorphisme :

$$\left\{ \begin{array}{l} \mathcal{F} : [0, \pi[\times [0, 2\pi[\longrightarrow S \\ (\theta, \phi) \longmapsto \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ \sin \theta \end{pmatrix} \end{array} \right. ,$$

où les angles θ et ϕ sont respectivement appelés angle polaire et azimutal (voir la Figure 2.1, page 15). Rappelons que l'équation (1.1) définissant le problème inverse à résoudre est :

$$V = \int_0^\pi \int_0^{2\pi} q(\theta, \phi) F(\theta, \phi) \sin \theta \, d\theta d\phi \quad (5.1)$$

où $V \in \mathbb{C}^n$ est la matrice spectrale mesurée, $q : S \rightarrow \mathbb{C}^n$ est le noyau d'intégration et $F : S \rightarrow \mathbb{R}^+$ est la FDO.

Les premiers tests numériques, ont été effectués sur des données synthétiques, dans le cas du vide. Nous avons utilisé des données synthétiques afin d'avoir une première idée du

comportement des algorithmes 1 et 2 et les valider. Ces tests sont présentés dans la Section 5.2. Dans un premier temps la matrice spectrale a été simulée en prenant une FDO qui s'écrit comme une somme de masses de Dirac, i.e. correspondant à une somme d'ondes planes. Ensuite, nous avons étudié le cas, plus réaliste, où la FDO servant à simuler la matrice spectrale est une fonction bornée et continue. Pour chacun de ces tests, nous avons fabriqué une matrice de bruit afin de perturber la matrice spectrale simulée et d'observer la stabilité *a posteriori* des solutions.

Dans la Section 5.3, nous appliquons les algorithmes de régularisation entropique sur des données enregistrées par le satellite FREJA. Le satellite FREJA est un satellite de recherche magnétosphérique Suédois. L'événement étudié est un souffle ELF, nous allons appliquer nos algorithmes, sur six points temps-fréquences de cet événement. Ces données ont déjà été étudiées par *Santolik et Parrot* [53], le calcul des FDO étant fait par d'autres méthodes. Nous pourrions donc comparer les résultats obtenus par régularisation entropique avec les résultats déjà publiés. Des tests pour quantifier la stabilité des solutions obtenues sont également présentés. Nous regarderons d'abord la dépendance de la solution par rapport au paramètre de régularisation, puis en perturbant la matrice spectrale nous allons quantifier la dépendance par rapport à la donnée. Enfin nous présenterons une étude statistique pour savoir si la méthode utilisée est capable de déterminer avec précision la position d'une ou de deux ondes planes.

5.2 Cas du vide : données synthétiques

Les tests présentés dans cette section, ont été faits pour valider (dans le cas du vide et sur des données artificielles) la méthode de régularisation entropique développée dans le chapitre 3. Ces tests nous ont permis de vérifier le bon fonctionnement des algorithmes 1 et 2 avant de passer aux tests sur des données d'expériences. C'est la seule section où nous présenterons des résultats obtenus grâce à l'algorithme 1. En effet, comme l'algorithme 2 fonctionne très bien dans tous nos essais numériques et qu'il n'est pas perturbé par un changement d'échelle de la donnée, nous l'utiliserons préférentiellement dans les sections suivantes.

5.2.1 Mise en oeuvre numérique

D'après le modèle, la matrice spectrale mesurée est reliée à la FDO par l'équation intégrale (5.1) où le noyau d'intégration $q : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{C}^n$ est une fonction connue qui dépend du milieu de propagation. Dans le cadre de ces tests, nous avons supposé que l'onde observée se déplace dans le vide. Les noyaux correspondants sont bien connus [27], ils sont explicités dans la Section 2.3. Ces noyaux dépendent de deux paramètres $|p|$ et β qui indiquent le mode de polarisation de l'onde. Pour faire les calculs, il est très important de supposer que les ondes étudiées ont la même polarisation, sinon le problème est difficile à résoudre. Nous avons pris arbitrairement $|p| = 1$ et $\beta = 1$. Nous avons également décidé d'utiliser toutes les composantes du champ électromagnétique : la donnée est (dans ce cas) un vecteur $V \in \mathbb{C}^{36}$.

Calcul des intégrales

Pour tous les exemples numériques de ce chapitre, nous avons utilisé une méthode de quadrature de Legendre-Gauss pour calculer les intégrales. Cette méthode est bien adaptée à la nature du problème puisque, les solutions obtenues par régularisation entropique ont la même régularité que les noyaux d'intégration. Or ceux-ci sont lisses dans les milieux considérés

Rappelons la méthode de quadrature de Legendre-Gauss, à $m \in \mathbb{N}^*$ points, pour calculer l'intégrale une fonction $f : [-1, 1] \rightarrow \mathbb{R}$:

$$\int_{-1}^1 f(x) dx = \sum_{k=1}^m H_k f(x_k) + \mathcal{E}$$

où les $\{x_k\}_{1 \leq k \leq m}$ sont les racines du polynôme de Legendre P_m de degré m [19, 41] et pour tout $1 \leq k \leq m$

$$H_k = -\frac{2}{(m+1)P_{m+1}(x_k)P'_m(x_k)} = \frac{2}{mP_{m-1}(x_k)P'_m(x_k)}.$$

La quantité \mathcal{E} représente l'erreur de la formule de quadrature; si la fonction f est C^{2m} , il existe $\xi \in [-1, 1]$ tel que

$$\mathcal{E} = \frac{2^{2m+1}(m!)^4}{(2m+1)[(2m)!]^3} f^{(2m)}(\xi).$$

Comme les noyaux d'intégration sont de classe C^∞ , cette méthode de calcul d'intégrale semble tout à fait appropriée.

Simulation de la matrice spectrale

Pour simuler la matrice spectrale, nous avons calculé l'intégrale de l'équation (5.1) pour une FDO initiale F_I . Dans le cas où F_I est une somme de masses de Dirac, l'évaluation de cette intégrale se fait simplement en sommant les valeurs du noyau q sur le support de la distribution F_I . Nous ne considérerons au maximum qu'une somme de trois masses de Dirac. On note $(\theta_i, \phi_i)_{i=1,2,3}$ les positions des masses de Dirac sur $[0, \pi] \times [0, 2\pi]$ et $(P_i)_{i=1,2,3}$ leurs puissances. En accord avec l'équation (5.1) on synthétise V par

$$V = \sum_{i=1,2,3} P_i q(\theta_i, \phi_i)$$

Notons que le cas d'une FDO constituée d'une somme de masses de Dirac est un cas théorique qui est *a priori* mal adapté pour une telle analyse. En effet, si la FDO est une somme de $m \in \mathbb{N}^*$ ondes planes cela signifie que l'onde se propage dans exactement m directions, ce qui est en désaccord avec le modèle construit où nous avons supposé que la FDO était une fonction de L^2 , décrivant la densité d'énergie dans toutes les directions (voir la Section 2.2). Dans la réalité, l'onde n'est de toute façon jamais concentrée sur une unique direction de propagation. L'analyse FDO, dans ce cas, peut donc être vue comme une approximation L^∞ de masses de Dirac. Notons également, que le problème de la détermination de la FDO sous la forme de $m \in \mathbb{N}^*$ masses de Dirac, n'est pas un problème mal posé si m est assez petit.

Dans le cas, plus classique, où la fonction F_I est lisse, nous avons utilisé la méthode de quadrature de Legendre-Gauss. Toutefois, le nombre de points servant à calculer l'intégrale est choisi plus grand que dans l'algorithme de résolution, afin d'éviter le "crime d'inversion". Nous avons choisi 64×128 points sur $[0, \pi[\times [0, 2\pi[$ pour les simulations alors que dans l'algorithme, la méthode utilise 10×20 ou (plus rarement) 20×40 sur le même pavé $[0, \pi[\times [0, 2\pi[$.

La matrice V correspondant à la FDO initiale, F_I étant désormais simulée, il faut maintenant y rajouter un bruit pour vérifier la stabilité de la reconstruction. Comme les matrices étudiées sont hermitiennes par construction, nous allons construire une matrice de bruit qui est elle aussi hermitienne. Ainsi, la matrice de bruit N est fabriquée en générant un vecteur gaussien b de taille p (et on a $p^2 = n$) et en posant $N = bb^*$. On met ensuite la matrice N sous forme de vecteur de \mathbb{C}^n , de la même manière que pour la matrice spectrale V . Nous choisissons le vecteur gaussien b de manière à ce que chacune de ses composantes ait la même loi. La variance de cette loi est ensuite adaptée au niveau de bruit souhaité pour l'expérience.

Initialisation des algorithmes

Pour initialiser les algorithmes 1 et 2 décrits à la Section 3.4, nous avons posé $l_0 = 0, \epsilon = 10^{-7}$. Le paramètre τ a été choisi de façon automatique pendant les calculs. Rappelons que le paramètre $\tau \in [0, 1]$ doit être choisi assez petit pour que l'algorithme converge. Mais s'il est trop petit, l'algorithme va converger vers l'unique solution du problème (\mathcal{P}_μ) de façon très lente. Pour choisir τ nous avons utilisé la règle suivante

1. Nous choisissons $\tau_0 \in [0, 1]$, $r \in]0, 1[$ et $N \in \mathbb{N}^*$.
2. Si l'algorithme ne descend pas avec τ_k à l'itération κ , i.e. $\|l_{\kappa+1} - l_\kappa\| > \|l_\kappa - l_{\kappa-1}\|$, on pose

$$\tau_{k+1} = r\tau.$$

3. Ayant trouvé τ_k tel que l'algorithme converge, on cherche le meilleur τ possible pour les calculs à venir. Pour cela, on prend le plus grand $\tau \in \{(N-j)r\tau_{k-1} + j\tau_{k-1} | j = 0, \dots, N-1\}$ tel que l'algorithme continue à converger.

Dans nos programmes, nous avons posé $r = 2$ et $N = 10$.

Le paramètre de régularisation $\mu > 0$ est choisi en vertu du principe de décalage de Morozov [22], ce principe consiste à fixer le paramètre μ tel que la solution correspondante ait une erreur égale au niveau de bruit. Dans la pratique nous supposons qu'un paramètre μ est valable si l'erreur appartient à un petit intervalle contenant la valeur du niveau de bruit $\delta > 0$. Comme nous traitons ici de données synthétiques, nous ne serons pas trop exigeants sur le choix de μ , c'est pourquoi nous considérerons que ce paramètre est valide si

$$\delta \leq \|V - \psi(F)\|_{\mathbb{C}^n} \leq 2\delta. \quad (5.2)$$

Ce choix est fait *a posteriori* en utilisant le fait que la fonction $\mu \mapsto \|V - \psi[F_\mu]\|_{\mathbb{C}^n}$ où F_μ est la solution du problème, est presque linéaire quand μ est petit, voir la Figure 5.1.

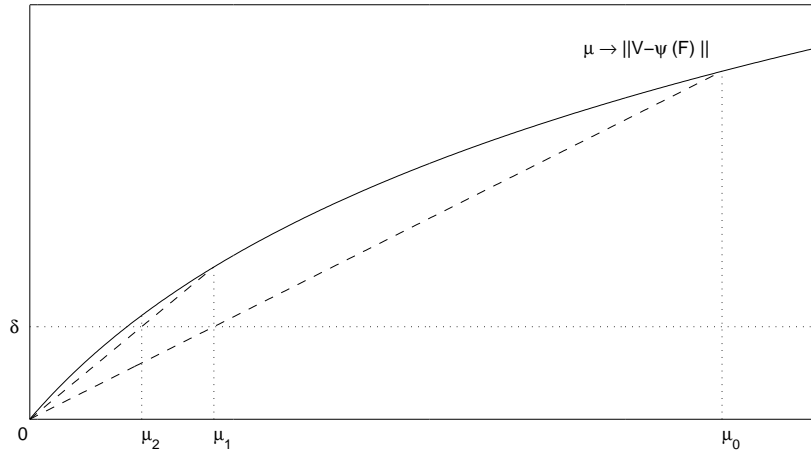


FIG. 5.1 – Méthode utilisée pour déterminer un μ tel que la solution F_μ du problème vérifie le principe de décalage de Morozov.

Notations et représentation des solutions obtenues

Avant de donner les résultats, voyons comment la solution est habituellement représentée. La fonction de distribution des ondes est une fonction définie sur la sphère unité de \mathbb{R}^3 , la Figure 5.5 montre la façon dont elle est habituellement représentée. La sphère est divisée en deux hémisphères : l’hémisphère nord pour $0 \leq \theta \leq \frac{\pi}{2}$ et l’hémisphère sud pour $\frac{\pi}{2} \leq \theta \leq \pi$. Sur chaque hémisphère, on représente la valeur de F par un code de couleur linéaire. Nous avons parfois utilisé une représentation en 3D ou avec des contours.

La solution F_μ du problème inverse étant déterminée par un des deux algorithmes (avec μ fixé), nous pouvons étudier ses caractéristiques. La valeur optimale de la fonction coût sera notée J_μ et H sera la valeur de l’entropie de la solution. Nous noterons ϵ_r l’erreur de la solution F_μ , c’est à dire

$$\epsilon_r = \|V - \psi[F_\mu]\|.$$

L’énergie de la solution calculée sera notée α . L’erreur relative entre l’énergie de la solution F_μ et la “véritable” valeur de l’énergie de l’onde sera notée ϵ_α . Nous avons également, pour chaque solution, déterminé sa valeur maximale, sa valeur minimale ainsi que le nombre d’itérations nécessaire pour le calcul.

5.2.2 Test avec trois masses de Dirac - Algorithme 1

Dans [42], l’algorithme 1 a été testé pour une donnée correspondant à une FDO constituée d’une masse de Dirac. Les calculs ont été faits pour différentes positions de la masse de Dirac, du pôle nord au pôle sud. Nous ne présenterons pas ici ces résultats, mais on peut les consulter en Annexe A page 103. Nous constatons sur cet exemple numérique que la méthode est capable de détecter de façon assez précise la position de la masse de Dirac.

Dans cette sous-section, nous allons tester l’algorithme 1 sur une donnée fabriquée à partir

de trois masses de Dirac réparties aux sommets d'un triangle équilatéral sur le rectangle $[0, \pi] \times [0, 2\pi]$. Nous avons bien entendu rajouté un bruit sur la donnée de manière à tester la stabilité de la solution. Nous avons utilisé l'algorithme 1 pour calculer la solution avec cette donnée. Cette solution est montrée en Figure 5.2 avec une représentation des contours de la fonction. On distingue bien sur le graphe, trois pics situés aux emplacements attendus. Le niveau de bruit pour cet exemple est de 3.1% et l'erreur de la solution obtenue est 5.5% pour $\mu = 0.5$. Le choix de μ est correct puisque la solution vérifie bien l'équation (5.2), i.e. l'erreur est supérieure au niveau de bruit mais est tout de même assez faible. La solution obtenue correspond tout à fait à la distribution F_I avec laquelle V a été simulée, F_μ a trois pics qui sont très pointus et bien positionnés. La solution du problème inverse obtenue par l'algorithme 1 est donc satisfaisante sur cet exemple. Dans la Sous-section suivante, nous allons regarder le comportement des solutions quand le paramètre de régularisation varie.

5.2.3 Effet de L'entropie

Afin d'observer les variations de la solution par rapport au paramètre de régularisation μ , nous avons refait le calcul avec la donnée V précédente, i.e. simulé à l'aide de trois masses de Dirac, mais avec différentes valeurs de μ . Les résultats obtenus, pour quatre valeurs de μ et par l'algorithme 1, sont résumés dans le Tableau 5.1. Comme l'entropie de la solution croît avec μ , nous nous attendions à ce que les trois pics disparaissent et que la solution soit très plate pour un μ choisi assez grand. Cela est confirmé par le graphe de la Figure 5.3 où l'on distingue la solution obtenue pour μ égal à 0.5 à gauche et 50 à droite. Le graphe de gauche qui est seulement une représentation différente de la Figure 5.2, est conforme aux résultats attendus et vérifie le principe de décalage de Morozov. Le graphe de droite ne vérifie pas ce principe, il est très plat et on ne distingue plus les trois pics. Comme cela était prévu, il y a perte d'information.

Cet exemple montre l'importance du niveau de bruit. En effet si le niveau de bruit est trop élevé, le paramètre μ devra être choisi assez grand et la solution obtenue contiendra très peu d'information. En choisissant μ plus petit, nous aurons plus d'information mais nous ne pourrons pas dire si cette information est "réelle" ou non.

Le Tableau 5.1 permet de voir l'évolution des caractéristiques de la solution pour les quatre valeurs de μ choisies. Nous avons refait les mêmes calculs sur un plus grand nombre de μ afin de tracer des courbes montrant l'évolution de ces diverses caractéristiques. La Figure 5.4 montre l'évolution du coût, de l'entropie, de l'erreur, de la puissance et du nombre d'itération pour les valeurs de μ comprises entre 0.5 et 50 avec un pas de 0.1. Nous observons sur cette figure (conformément à nos attentes) que la fonction coût est majorée et que l'entropie décroît contrairement à l'erreur.

5.2.4 Test avec une FDO continue - Algorithme 2

La simulation des données avec une FDO, sous la forme d'une masse de Dirac est un cas particulier assez limité. Comme le montrent les résultats du Chapitre 3, la solution obtenue par régularisation entropique a la même régularité que les noyaux d'intégrations et ceux-ci sont lisses dans nos exemples. C'est pourquoi, nous allons maintenant faire un test pour une

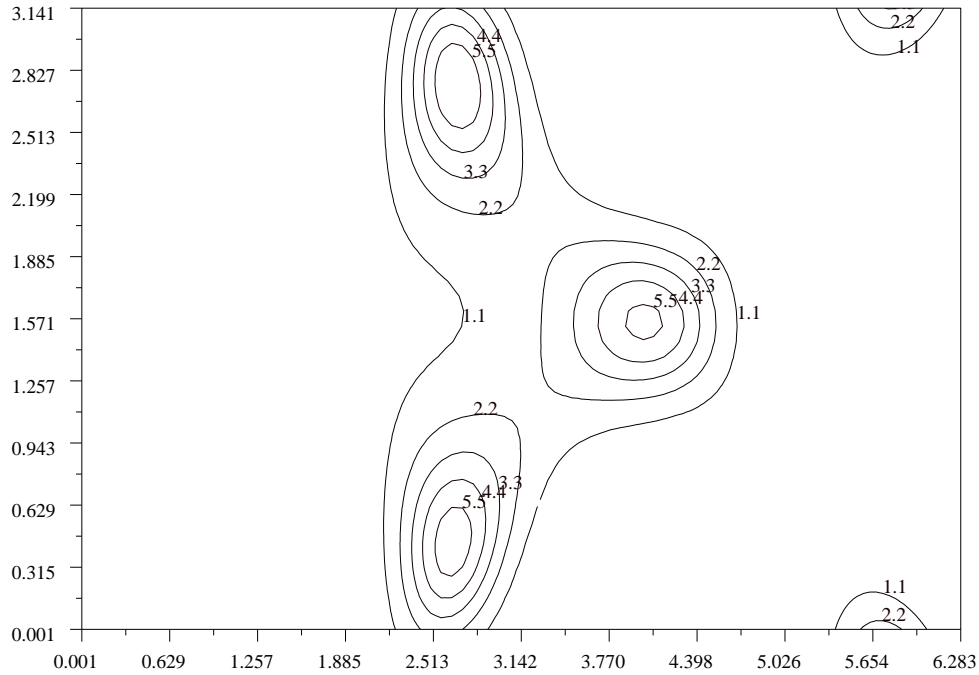


FIG. 5.2 – Tracé en contours de la solution obtenue par l’algorithme 1 pour la matrice spectrale V simulée à l’aide de trois masses de Dirac. La solution a également trois pics situés sur le support des trois Dirac.

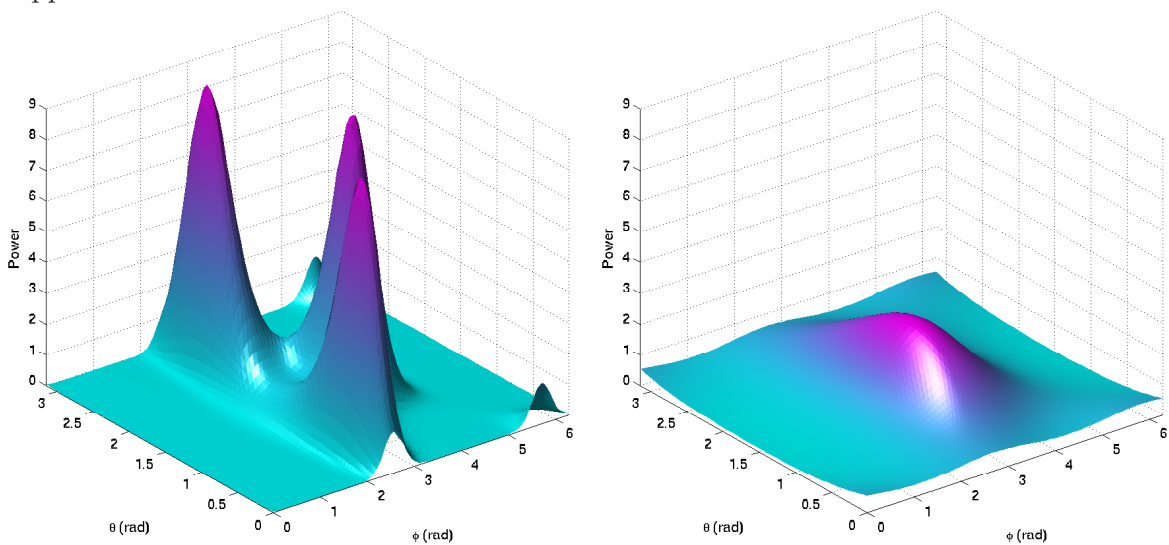


FIG. 5.3 – Calcul de la FDO dans le cas où V a été simulé à l’aide de trois masses de Dirac. A gauche : solution calculée pour $\mu = 0.5$. A droite : solution calculée avec $\mu = 50$.

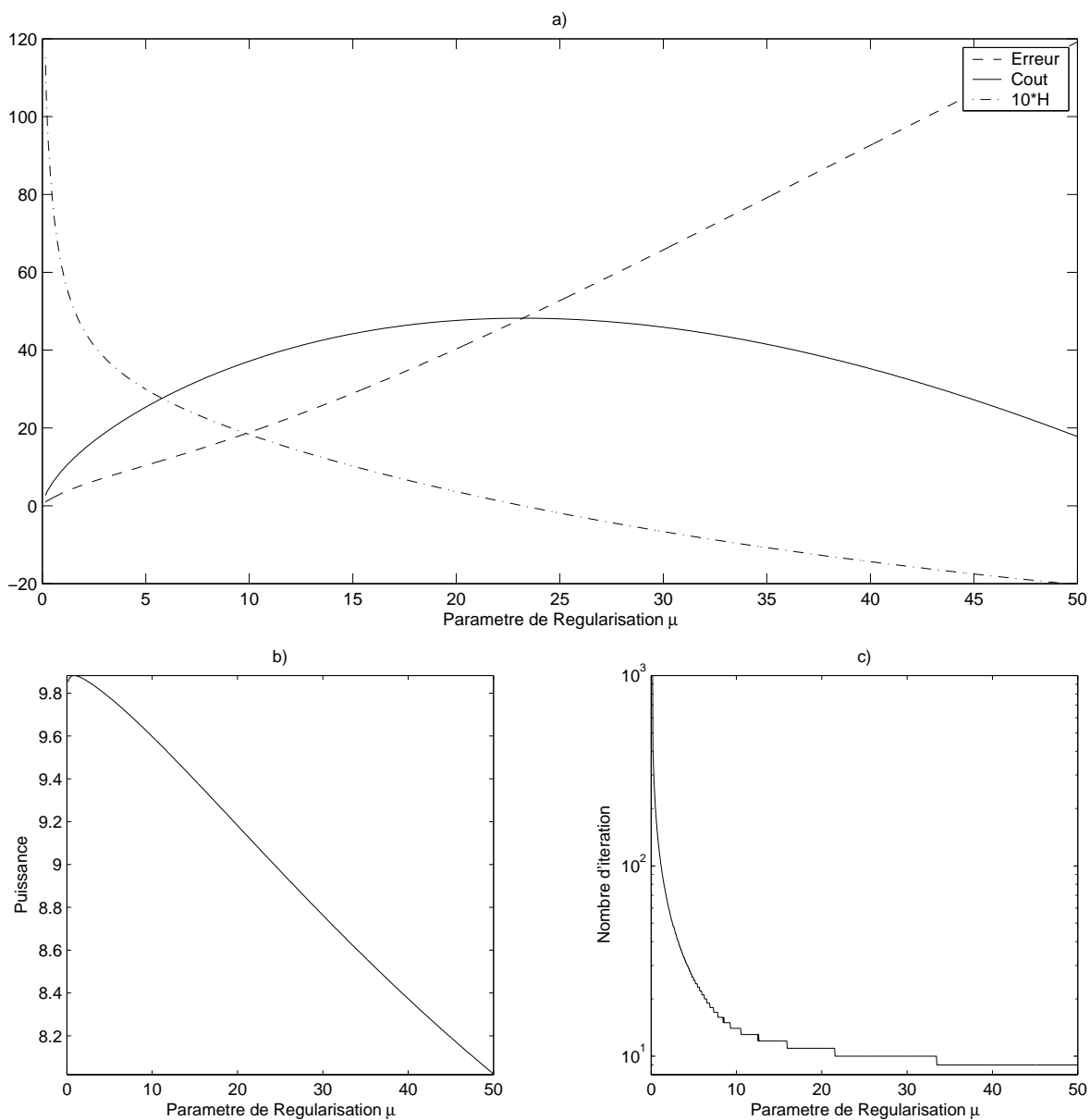


FIG. 5.4 – a) Graphe de l'erreur, de l'entropie, et du coût de la FDO en fonction de μ . b) Variation de la puissance de la FDO en fonction de μ . c) Nombre d'itérations nécessaires au calcul de la FDO en fonction de μ .

μ	α	ϵ_α	ϵ_r	H	J_μ	Min	Max	#It.	τ
0.5	9.891	2.07%	1.842	-1.452	5.741	2.200E-6	6.947	8377	2.5E-3
1	9.868	1.83%	3.159	-1.660	9.360	1.521E-4	4.597	3317	6.25E-2
5	9.814	1.27%	10.27	-1.990	24.691	2.678E-2	3.136	465	0.05
50	8.008	17.35%	117.633	-2.338	17.862	0.195	1.725	7	1

TAB. 5.1 – Résultats obtenus pour une matrice simulée V à partir de trois masses de Dirac. Nous avons fait les calculs pour quatre valeurs différentes du paramètre μ . Min et Max sont respectivement le minimum et la maximum de la solution sur la sphère unité S .

matrice spectrale simulée avec une FDO continue. Nous avons choisi de prendre comme FDO initiale la fonction :

$$F_I(\theta, \phi) := 4 (\cos^2(\theta) + \sin^2(\phi)).$$

Afin d'éviter le "crime d'inversion", nous avons utilisé pour calculer la matrice spectrale, un pas d'intégration plus fin que celui des algorithmes de résolution. Ensuite nous avons bruité la matrice spectrale, le niveau de bruit étant de 5.1%.

La Figure 5.5 représente la fonction initiale, ainsi que la fonction obtenue par l'algorithme 2. La solution a une erreur de 8% avec la donnée, elle n'est donc pas plus précise que le niveau de bruit. Nous n'avons pas voulu tracer la solution obtenue par l'algorithme 1 puisque celle-ci est pratiquement identique à celle présentée. Nous voyons sur ces graphiques, que la solution calculée par régularisation entropique possède la même géométrie que la FDO initiale, elle est juste un peu plus aplatie, ce qui est certainement un effet de la minimisation de l'entropie.

5.2.5 Test avec deux modes de polarisation

On suppose que l'onde étudiée est constituée de deux modes de polarisation superposés. Dans ce cas, la matrice spectrale V est reliée aux FDO F_1 et F_2 de chacun des modes, par l'équation

$$V = \sum_{j=1,2} \int_0^{2\pi} \int_0^\pi q_j(\theta, \phi) F_j(\theta, \phi) \sin \theta d\theta d\phi$$

L'équation qui relie V aux FDO F_j , est toujours linéaire, nous pouvons donc appliquer la méthode de régularisation entropique comme nous l'avons fait précédemment. Toutefois la taille de la matrice V n'ayant pas changée on peut estimer qu'il sera plus difficile de déterminer deux FDO à partir de V .

Nous avons fait un essai avec les noyaux du vide, en prenant une onde constituée de deux ondes planes ayant des modes de polarisation différents. Nous avons utilisé l'algorithme 1 pour faire les calculs, et nous constatons que les solutions obtenues ont une très grande erreur par rapport au niveau de bruit. Nous n'avons pas pu calculer à l'aide de cet algorithme, une solution telle que $\epsilon_r > 500$. Cela signifie que nous ne pouvons pas raisonnablement utiliser ces solutions pour l'analyse de l'onde, l'erreur étant beaucoup trop grande. La figure 5.6 montre

le résultat obtenu par l'algorithme 1 avec $\epsilon_r \approx 600$. Nous voyons tout de même sur cette figure, que la solution a bien une direction privilégiée pour chacun des modes de polarisation, et que les directions de ces ondes sont assez bien déterminées bien que légèrement décalées.

5.2.6 A propos de la dualité

Comme nous l'avons vu dans la Section 3.5, l'algorithme 1 est une méthode de gradient sur le dual du problème de régularisation entropique (\mathcal{P}_μ). Nous constatons également, que cette descente est de plus en plus difficile, au fur et à mesure que μ décroît, puisque nous sommes obligés de choisir τ de plus en plus petit. Le Tableau 5.1 et la Figure 5.4 montrent bien ce phénomène, le nombre d'itérations décroît et la valeur de τ augmente lorsque μ croît.

Nous pouvons nous demander s'il serait possible, d'utiliser un algorithme quasi-Newton afin de calculer la solution plus rapidement pour les petites valeurs de μ . Nous avons donc fait un test avec l'algorithme de Levenberg-Marquardt [41, 6] pour comparer les vitesses de convergence entre les algorithmes. De fait, on observe que les méthodes quasi-Newton ne sont pas du tout adaptées à ce problème. En effet, comme ψ n'est pas bijective, lorsque nous choisissons μ très petit, la stricte-convexité de la fonction coût n'est numériquement plus visible. La matrice hessienne, de la fonction coût est presque singulière et donc les méthodes quasi-Newton convergent très mal.

Numériquement nous constatons qu'il est très difficile (voire impossible) de faire converger l'algorithme de Levenberg-Marquardt pour des valeurs de μ inférieures à 2. Lorsque nous prenons μ plus grand, l'algorithme 1 converge plus vite que l'algorithme de Levenberg-Marquardt, ce qui peut être expliqué par le temps de calcul du Hessien qui est très coûteux dans la méthode de Levenberg-Marquardt. Par exemple, lorsque $\mu = 8$ nous avons constaté que l'algorithme 1 était environ 3 fois plus rapide. Il apparaît donc assez mal adapté d'utiliser cette méthode pour calculer la solution et nous allons considérer une autre approche.

Dans le Chapitre 3, nous avons vu que le problème de régularisation entropique (\mathcal{P}_μ) était équivalent au problème

$$(P^\varepsilon) \left\{ \begin{array}{l} \min H(f) \\ f \in L^2, f \geq 0, \|V - \psi[f]\|_{\mathbb{C}^n}^2 \leq \varepsilon^2 \end{array} \right. ,$$

sauf que pour ce problème, nous n'avons pas de paramètre de régularisation μ à fixer, mais seulement l'erreur ε voulue. Si nous connaissons le niveau de bruit δ , nous pouvons donc résoudre le problème sans avoir aucun paramètre à fixer pour l'inversion. Le problème (P^ε) est assez complexe à résoudre sous cette forme. En effet si nous utilisons une méthode de gradient pour calculer l'optimum du problème, nous sommes obligés de faire des projections pour que la solution vérifie les contraintes. Cependant, le dual du problème (P^ε) a été calculé dans la sous-section 3.5.5, et ce problème dual s'écrit

$$\sup_{l \in \mathbb{C}^n} \left\{ - \int_E \exp(-1 + \psi^*[l]) d\sigma + \langle l, V \rangle_{\mathbb{C}^n} - \varepsilon \|l\|_{\mathbb{C}^n} \right\}. \quad (5.3)$$

Le problème dual est un problème sans contrainte, une méthode de descente sans projection peut donc être utilisée pour en calculer l'optimum. Il y a cependant une nette différence

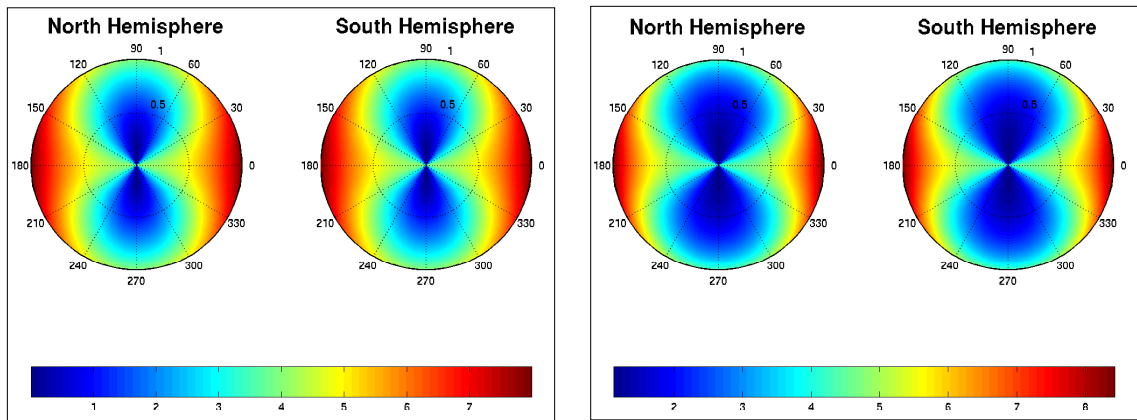


FIG. 5.5 – À gauche, tracé de la fonction initiale F_I qui a été utilisée pour simuler la matrice spectrale. À droite, solution obtenue par l'algorithme 1 avec $\mu = 3$

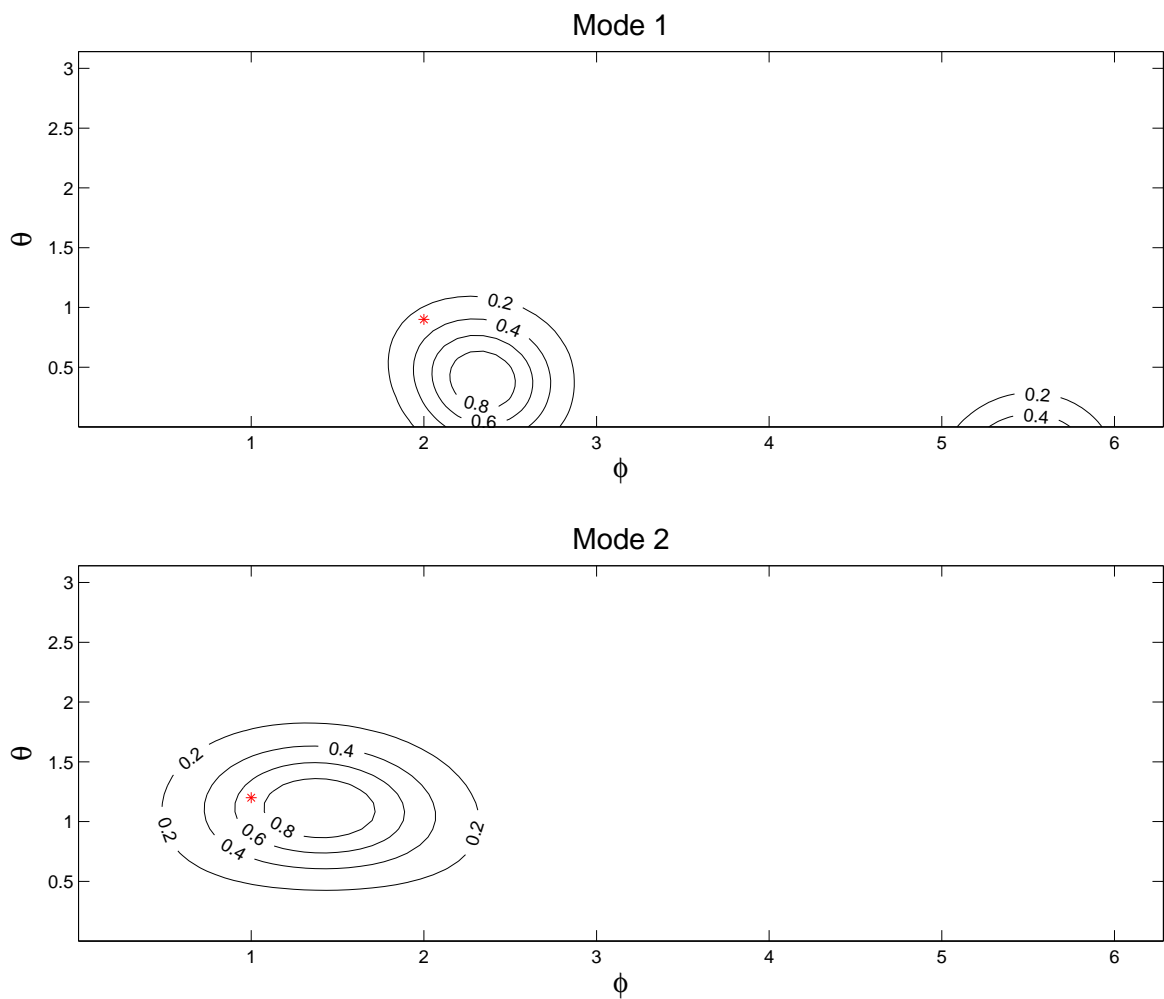


FIG. 5.6 – Calcul des FDO dans chacun des modes de polarisation. Les astérisques indiquent la véritable direction des ondes

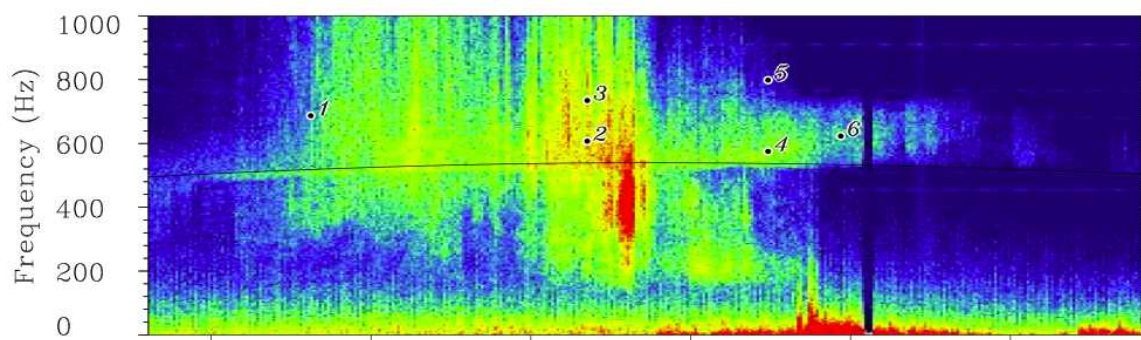


FIG. 5.7 – Spectrogramme du souffle ELF enregistré par le satellite FREJA (figure tirée de [53]).

entre ce problème et le problème dual (\mathcal{D}_μ) de (\mathcal{P}_μ) : la fonction coût de (\mathcal{D}_μ) est Gâteaux-différentiable alors que la fonction coût de (5.3) n'est pas différentiable en 0. Nous pouvons donc nous attendre à ce que le calcul numérique de la solution à l'aide d'une méthode de descente sur cette fonction coût, soit plus délicat qu'avec les algorithmes 1 et 2.

Nous avons testé cette méthode dans le cas du vide, sur les exemples où la matrice spectrale est simulée à partir de 3 masses de Dirac et à partir d'une FDO continue. Nous avons utilisé une méthode de gradient pour calculer l'optimum de (5.3). Il faut choisir, afin d'initialiser l'algorithme, un point de départ différent de 0 pour éviter la singularité du gradient de la fonction coût. Nous avons utilisé un pas de descente τ constant comme pour l'algorithme 1. Sur les deux exemples considérés, nous avons retrouvé les solutions calculées précédemment. La convergence de la méthode a lieu avec une vitesse comparable à celle de l'algorithme 1. Toutefois, nous n'avons pas besoin de chercher un μ convenable puisque l'erreur a été fixée à la valeur ε souhaitée avant les calculs.

5.3 Cas des plasmas : données expérimentales

Le satellite FREJA a enregistré un intense souffle ELF le 8 Avril 1995 entre 15h48 et 15h58 UT. Nous avons fait une analyse FDO de cet événement en utilisant la régularisation entropique. Le spectrogramme de l'enregistrement est montré en Figure 5.7, les six points notés sur le spectrogramme sont les six points temps-fréquences sur lesquels l'analyse va être effectuée. L'analyse onde plane et le calcul des FDO ont été effectués précédemment par *Santolík et Parrot* [53]. Pour la calcul de la FDO, diverses techniques ont été utilisées : la discrétisation, la méthode des pics gaussiens et la méthode du maximum d'entropie de *Lefevre et Delannoy* [26]. Les principes de chacune de ces méthodes sont les suivants.

- La méthode de la FDO discrète : il s'agit d'une méthode de régularisation où le paramètre de régularisation est la taille du maillage utilisé pour la discrétisation. Afin de discrétiser correctement la sphère unité S , *Santolík* [50] a utilisé un maillage en forme d'igloo afin d'obtenir une discrétisation assez homogène sur la sphère. Nous ne ferons

cependant pas de comparaison avec les résultats obtenus par cette méthode.

- La méthode des Pics Gaussiens (GP) : on cherche F sous la forme

$$F(\theta, \phi) = \sum_{k=1}^N \frac{\rho_k}{\pi \Delta_k^2} \exp \left[\frac{2d_S^2 \{(\theta, \phi), (\theta_k, \phi_k)\}}{\Delta_k^2} \right]$$

où N est le nombre de pics choisis, (θ_k, ϕ_k) , ρ_k et Δ_k sont respectivement la direction, l'énergie et la largeur angulaire du pic k . La fonction $d_S : S \times S \rightarrow \mathbb{R}$ est la distance sur la sphère S , les points de la sphère étant décrits par l'angle polaire $\theta \in [0, \pi]$ et l'angle azimutal $\phi \in [0, 2\pi]$.

- La méthode de Maximum d'Entropie (ME), développée par *Lefevre et Delannoy* [24] : il s'agit de calculer la solution de maximum d'entropie (voir la Définition 10 page 29) de l'équation (1.1). Pratiquement cette solution est déterminée en prenant la FDO F sous la forme

$$F(\theta, \phi) = \exp \left[-1 + \sum_{k=1}^n \lambda_k q_k(\theta, \phi) \right],$$

et les multiplicateurs de Lagrange $\{\lambda_k\}_{1 \leq k \leq n}$ sont identifiés à l'aide de la méthode de Levenberg-Marquardt [6, 41].

Pour plus de précision sur ces différentes méthodes, on peut consulter [50, 53, 24, 25]. Ces données ayant déjà été analysées [53], il est commode de comparer les solutions obtenues par régularisation entropique avec celles obtenues par les méthodes GP et ME.

Contrairement aux données synthétiques utilisées dans les tests précédents, les données que nous avons à traiter ici sont des données d'expérience enregistrées par l'instrument F4 [20] à bord de FREJA. Nous n'avons pas la même précision de mesure sur toutes les composantes du champ, c'est pourquoi il faut préalablement préparer les données pour l'analyse.

5.3.1 Préparation des données

Le champ électromagnétique est enregistré par le satellite FREJA à l'aide de 4 antennes [20]. Il y a trois antennes magnétiques dans des directions orthogonales et une antenne électrique. La position des antennes magnétiques est fixe dans le repère du satellite, alors que l'antenne électrique a un mouvement giratoire. Le champ mesuré par cette antenne, est la moyenne des mesures pendant une rotation d'un quart de tour. La précision des mesures du champ électrique est donc bien moindre que celle du champ magnétique. Nous allons devoir effectuer une normalisation sur la matrice spectrale et sur les noyaux d'intégration afin d'éliminer ces différences de précision.

Pour chacun des six points $k = 1, \dots, 6$, nous disposons de la matrice spectrale estimée $\tilde{V}^{(k)} \in \mathbb{C}^n$, de la matrice des erreurs de mesure $\epsilon^{(k)} \in \mathbb{C}^n$ ainsi que du calcul des noyaux d'intégration du plasma à la fréquence considérée $\tilde{q}^{(k)}$. Avant de commencer les calculs nous effectuons une normalisation de ces données, en posant

$$\forall k \in \{1, \dots, 6\}, \forall i \in \{1, \dots, n\}, V_i^{(k)} = \frac{\tilde{V}_i^{(k)}}{\epsilon_i^{(k)}} \text{ et } q_i^{(k)}(\theta, \phi) = \frac{\tilde{q}_i^{(k)}}{\epsilon_i^{(k)}}$$

Après cette normalisation, tous les éléments de la matrice spectrale ont une erreur du même ordre. Nous supposons donc que le niveau de bruit, après normalisation, est $\delta^2 = 16$.

Pour les calculs, nous avons utilisé les mêmes noyaux d'intégration que dans [53]. Ils ont été calculés à partir d'un modèle de plasma froid, avec l'hypothèse d'un plasma hydrogène-oxygène, simplement ionisé et sans collision. *Santolík et Parrot* [53] ont utilisé une fraction de 5% d'hydrogène. La composition du plasma est importante car la fréquence des ondes est inférieure à deux fois la gyro-fréquence f_{H^+} des électrons. Les noyaux ont été calculés en utilisant le mode droit, car c'est le seul mode présent dans un plasma froid pour une fréquence supérieure à f_{H^+} [59].

Nous devons aussi mettre les noyaux à l'échelle. En effet les noyaux calculés à l'aide de la relation de dispersion admettent de grandes variations, ce qui est très gênant du point de vue numérique. Nous les avons donc multipliés par une constante de la forme $10^{-\beta}$. La valeur de β utilisée pour chacun des six points est donnée au Tableau 5.2.

5.3.2 Représentation des solutions obtenues

Pour représenter graphiquement les FDO, calculées par régularisation entropique, nous avons utilisé les mêmes conventions que dans l'article de *Santolík et Parrot* [53]. Les FDO sont normalisées de manière à ce qu'elles aient pour valeurs maximales 1 et elles sont tracées en niveau de gris avec une échelle linéaire. Des contours indiquant les niveaux $k/6$ pour $k = 1, \dots, 5$ sont aussi tracés. La Figure 5.8 est un exemple de la représentation graphique que nous utiliserons par la suite, avec les graduations des angles θ et ϕ ainsi que le code de niveau de gris utilisé.

5.3.3 Résolution avec le champ magnétique

Les mesures du champ magnétique sont beaucoup plus fiables que les mesures faites à l'aide de l'antenne électrique, cette dernière n'ayant pas de position bien définie à l'instant des mesures à cause de sa rotation. C'est pourquoi, nous avons d'abord fait les calculs de la FDO pour les six points, en ne prenant en compte que les trois composantes magnétiques. La Figure 5.9 montre les résultats obtenus par la méthode de régularisation entropique, pour les points 1,2 et 4. En l'absence de composantes électriques du champ électromagnétique, les solutions présentent une symétrie nord-sud. Cela était prévu par la théorie, car d'après les équations de Maxwell, il faut au moins quatre composantes du champ pour déterminer la direction et le sens d'une onde. Avec trois composantes nous pouvons seulement calculer la direction.

Les résultats de la Figure 5.9 sont tout à fait en accord avec ceux obtenus dans [53], ils permettent également de valider les solutions obtenues avec toutes les composantes du champ, comme nous allons le voir dans le paragraphe suivant.

5.3.4 Calcul de la FDO avec toutes les composantes

La Figure 5.10 montre les solutions calculées par l'algorithme de régularisation entropique (ERA), par la méthode des pics gaussiens (GP) et par la méthode du maximum d'entropie (ME). Les solutions ERA ont été obtenues à partir de l'algorithme 2 qui produit des solutions

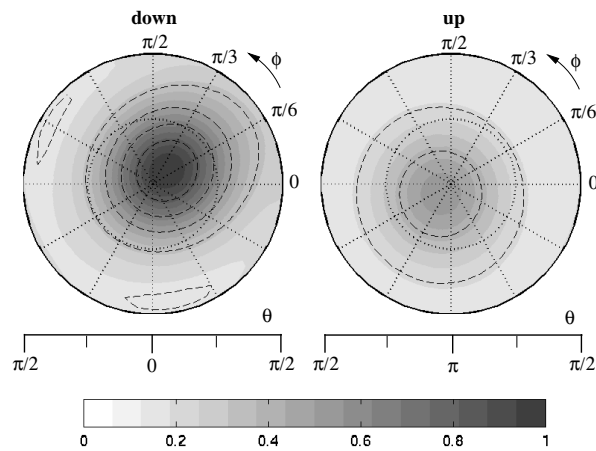


FIG. 5.8 – Représentation de la FDO utilisée afin de faire des comparaisons avec les résultats de [53]

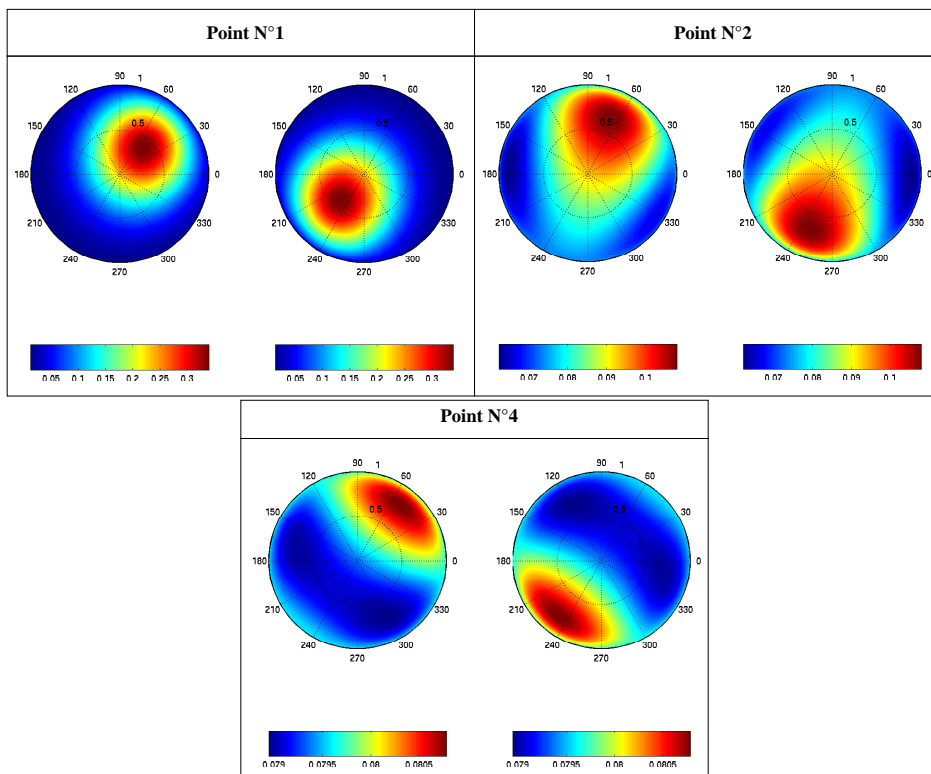


FIG. 5.9 – Solutions calculées par l’algorithme 2, pour les points 1,2 et 4, en ne prenant en compte que les composantes du champ magnétique. Le fait que les solutions soient symétriques est normal puisque d’après les équations de Maxwell, sans composante électrique on ne peut pas déterminer le sens de propagation.

qui dépendent de façon positivement homogène de la donnée. Les solutions GP et ME ont été calculées par *Santolík et Parrot* [53].

De façon générale, nous n’observons pas de grandes différences entre toutes ces solutions. Les solutions ERA sont plus étalées que les autres, ce qui est simplement l’effet de la maximisation de l’entropie. Les solutions ME sont moins étalées car l’entropie est maximisée avec des contraintes beaucoup plus fortes. Nous constatons aussi, que les solutions obtenues avec toutes les composantes du champ, s’accordent bien avec les solutions de la Figure 5.9 déterminée seulement à partir des mesures du champ magnétique.

Pour le point n°1 (panneaux en haut à gauche sur la Figure 5.10), l’analyse onde plane effectuée dans [53] révèle une onde se déplaçant vers le bas à $\theta \approx 30^\circ$, $\phi \approx 45^\circ$ avec un degré de polarisation élevé. Les solutions GP et ME ont bien un pic assez fin, qui se situe, à peu près à la position prédite par l’analyse onde-plane. Pour la solution ERA, nous observons bien un pic principal à $\theta \approx 14^\circ$, $\phi \approx 43^\circ$ dans l’hémisphère nord (ondes descendantes), mais il y a aussi un pic secondaire dans la direction opposé (hémisphère sud), celui-ci étant moins puissant.

Pour le point n°2, l’analyse onde plane détermine une direction de propagation vers le bas, l’angle polaire est très petit. Cependant, le degré de polarisation est aussi très faible. La solution ERA montre un pic très large en direction du bas, à $\theta \approx 33^\circ$, $\phi \approx 63^\circ$. Ce pic correspond bien au pic primaire obtenu par les deux autres méthodes. Toutefois les deux autres méthodes détectent un pic secondaire, bien défini pour GP et plus critiquable pour ME, dans l’autre hémisphère. Dans la solution ERA, un tel pic n’est pas observé, un accroissement de l’énergie est cependant clairement identifié dans la même région. Nous remarquons que la solution ERA reflète bien le faible degré de polarisation trouvé lors de l’analyse onde plane [53], ce qui signifie que l’onde est censée se propager dans un grand nombre de directions.

Le point n°3 est presque similaire au point n°1. Les trois méthodes détectent un pic principal dans l’hémisphère nord, mais, comme pour le point un, la solution ERA a un pic secondaire dans la direction opposée au premier. Ce pic secondaire n’est pas détecté par les deux autres méthodes. Le point n°4 est quant à lui presque le symétrique du point n°2, en échangeant les hémisphères. Le pic principal est détecté dans l’hémisphère sud par les trois méthodes. Notons que la solution ERA est, comme pour le point n°2, très diffuse.

Pour les points n°5 et n°6 il n’y a pas de différence marquée entre les trois solutions. En effet, les solutions obtenues par ME et GP sont très proches de celles données par ERA, à savoir un pic dans la direction $\theta \approx 60^\circ$, $\phi \approx 201^\circ$ pour le point n°5 et $\theta \approx 108^\circ$, $\phi \approx 200^\circ$ pour le point n°6.

5.3.5 Tests de stabilité I

Afin de vérifier et de quantifier la stabilité des solutions de façon globale, nous avons fait une étude statistique des variations de l’énergie des solutions. Nous avons donc créé des matrices spectrales perturbées V_N où N et une matrice de bruit simulée de la même manière que dans la sous-section 5.2.1. Puis nous avons calculé la déviation d’énergie $C_1 \in \mathbb{R}^+$ de la solution perturbée F_N par

$$C_1 := \frac{\int_0^\pi \int_0^{2\pi} |F(\theta, \phi) - F_N(\theta, \phi)| \sin \theta \, d\phi \, d\theta}{\|V - V_N\|}$$

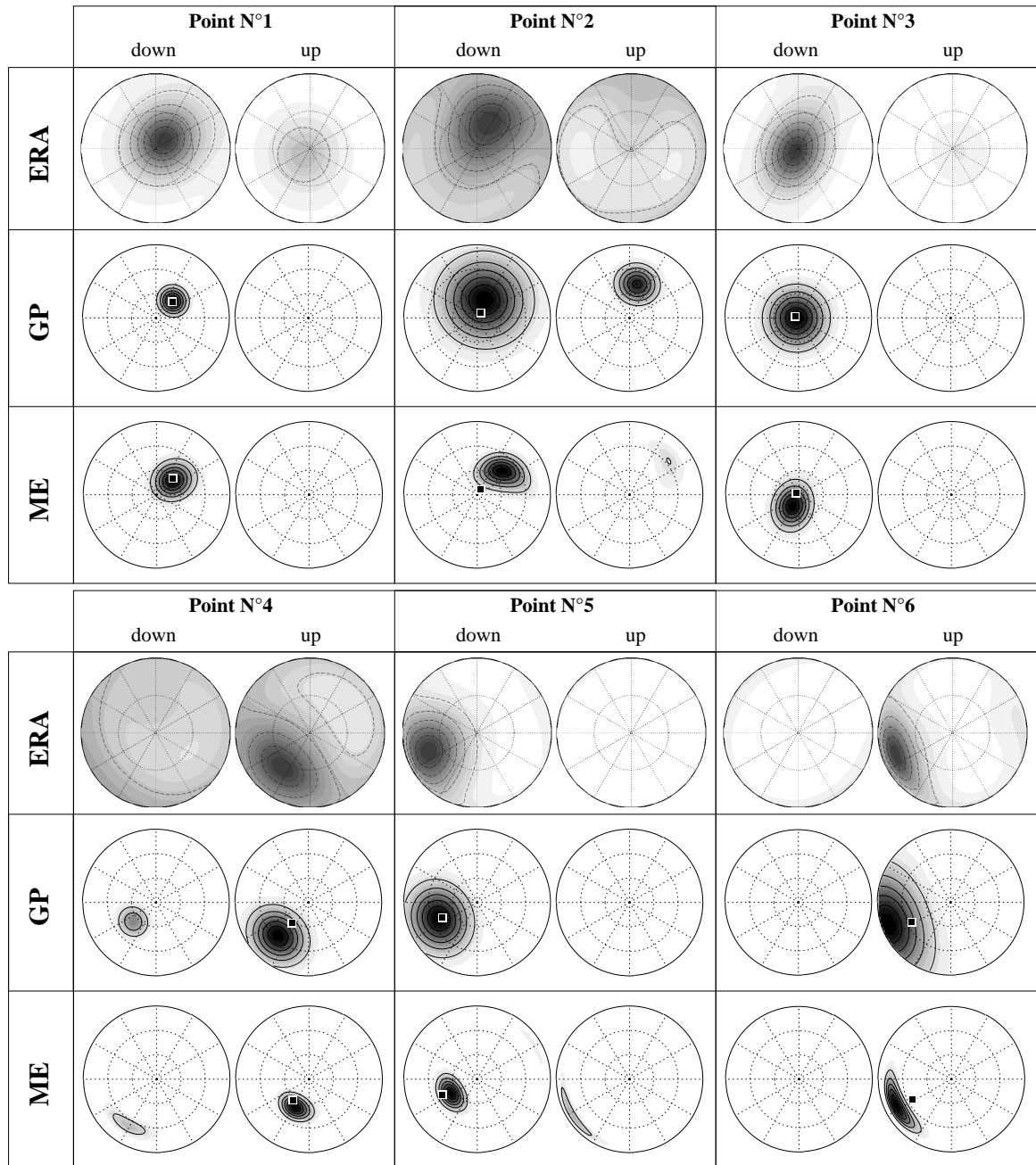


FIG. 5.10 – FDO calculées pour les six points du spectrogramme du souffle ELF. En haut on distingue les solutions déterminées par régularisation entropique (ERA), au milieu ce sont les solutions de la méthode des pics gaussiens et en bas les solutions de maximum d'entropie (ME). Les figures des solutions (ME) et (GP) proviennent de [53].

Nous avons répété ce processus 1000 fois, afin d'obtenir un échantillon assez important de la variable C_1 . Ensuite nous avons calculé la moyenne $\mathbb{E}(C_1)$ et la variance $\text{Var}(C_1)$ de l'échantillon. Nous avons aussi déterminé l'erreur relative moyenne de la solution perturbée pour un niveau de bruit δ :

$$V_1^\delta := \frac{\mathbb{E}(C_1)\delta}{\int_0^\pi \int_0^{2\pi} |F(\theta, \phi)| \sin \theta \, d\phi \, d\theta}.$$

Le Tableau 5.2 donne les valeurs obtenues pour chacun des six points. Nous observons qu'en moyenne la variation d'énergie V_1^δ est 21%. La variation de l'énergie est la plus faible pour le point 2, alors que la plus grande variation est observée pour le point 1. Nous en concluons que les solutions sont bien stables d'un point de vue global.

Point N°	1	2	3	4	5	6
Temps UT	15 :49 :15	15 :52 :43	15 :52 :43	15 :54 :59	15 :54 :59	15 :55 :53
Fréquences [H_z]	688	608	736	576	800	624
FDO						
β	13	11	11	12	15	12
Energie	2.98	2.83	0.58	9.90	3.62e-1	9.98e-1
Minimum	4.61e-2	7.19e-2	3.56e-3	3.01e-1	1.72e-3	6.30e-3
Maximum	1.38	0.61	3.64e-1	2.35	2.42e-1	8.61e-1
Statistiques						
$\mathbb{E}(C_1)$	1.97e-1	1.14e-1	2.93e-2	5.01e-1	1.97e-2	5.01e-2
$\text{Var}(C_1)$	8.76e-4	3.54e-4	2.17e-5	8.92e-3	1.25e-5	1.02e-4
V_1^δ	26.4%	16.1%	20.2%	20.2%	21.7%	20.08%
ζ	0.0694	0.0695	0.0353	0.0372	0.0332	0.026

TAB. 5.2 – Résultats obtenus pour chacun des points. le nombre $10^{-\beta}$ est le facteur d'échelle utilisé pour réduire les noyaux d'intégration. Pour chaque point, nous avons calculé l'énergie le minimum et le maximum de la FDO. Il y a également les résultats des tests de stabilité $\mathbb{E}(C_1)$, V_1^δ et le facteur ζ entre la FDO et $\mathbb{E}(C_\infty)$

5.3.6 Tests de stabilité II

Le test de stabilité précédent, nous donne une estimation de la variation d'énergie moyenne de la solution perturbée. Cela donne une idée globale de la stabilité. Nous allons maintenant quantifier la stabilité des solutions, de façon locale. On considère, comme précédemment, la solution F_N relativement à la matrice perturbée V_N et on définit la fonction

$$C_\infty = \frac{|F(\theta, \phi) - F_N(\theta, \phi)|}{\|V - V_N\|}.$$

Cette fonction, représente les variations de la solution perturbée F_N par rapport à F , que nous avons normalisée, par la norme de la perturbation $\|V - V_N\|$. Comme précédemment,

nous avons réitéré ce processus 1000 fois, afin d'avoir un échantillon assez large, puis nous avons calculé la moyenne empirique $\mathbb{E}(C_\infty)$ de cet échantillon. La fonction $\mathbb{E}(C_\infty)$ représente ainsi la variation moyenne de la solution, pour une perturbation de norme 1, contrairement à $\mathbb{E}(C_1)$ elle permet d'étudier la stabilité de façon ponctuelle. Pour représenter ces fonctions nous avons utilisé les mêmes conventions que pour tracer une FDO (voir la Figure 5.8). Nous avons donc normalisé la fonction $\mathbb{E}(C_\infty)$ pour cette représentation, et nous avons calculé son échelle ζ par rapport à la FDO.

Nous avons tracé sur la Figure 5.11 les FDO et les fonction $\mathbb{E}(C_\infty)$, pour chacun des six points. Nous observons sur cette figure que la stabilité des solutions est convenable, la valeur de ζ est inférieure à 7% pour tous les points.

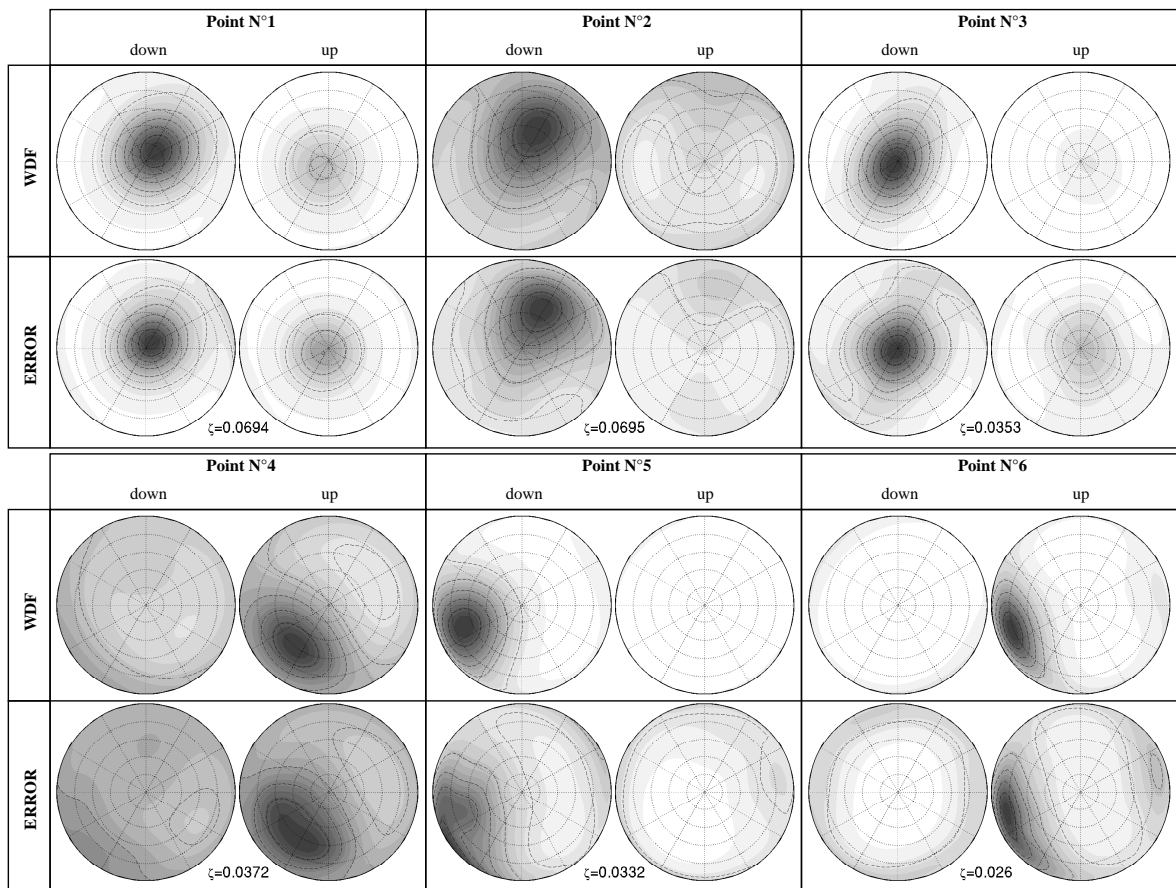


FIG. 5.11 – Les FDO pour chacun des six points, comparées avec les fonction $\mathbb{E}(C_\infty)$ donnant l'erreur moyenne pour une perturbation de norme 1.

5.3.7 Détection d'ondes planes

Nous avons fait un test statistique, pour vérifier si l'algorithme d'inversion est capable de détecter correctement la direction d'une onde plane. Contrairement à l'étude faite dans l'Annexe A qui utilisait les noyaux du vide, nous allons utiliser les noyaux du plasma froid, plus précisément ceux du point N°2 du souffle ELF enregistré par FREJA. Pour ce test, nous avons choisi 1000 directions aléatoires $(\theta_i, \phi_i)_{1 \leq i \leq 1000}$ et puissances aléatoires $(\rho_i)_{1 \leq i \leq 1000}$ avec lesquelles nous avons simulé 1000 matrices spectrales. Ensuite, nous avons utilisé l'algorithme 2 pour calculer la FDO correspondant à chacune de ces matrices, la direction de propagation calculée $(\theta'_i, \phi'_i)_{1 \leq i \leq 1000}$ a été déterminée en dérivant les FDO obtenues. Enfin, nous avons déterminé les distances angulaires $(\epsilon_i)_{1 \leq i \leq 1000}$ entre la vraie direction de propagation et celle calculée par

$$\epsilon_i = \arccos[\cos(\theta_i - \theta'_i) \cos(\phi_i - \phi'_i) + \cos(\theta_i) \cos(\theta'_i)(1 - \cos(\phi_i - \phi'_i))] \quad (5.4)$$

Les résultats obtenus sont montrés en Figure 5.12.

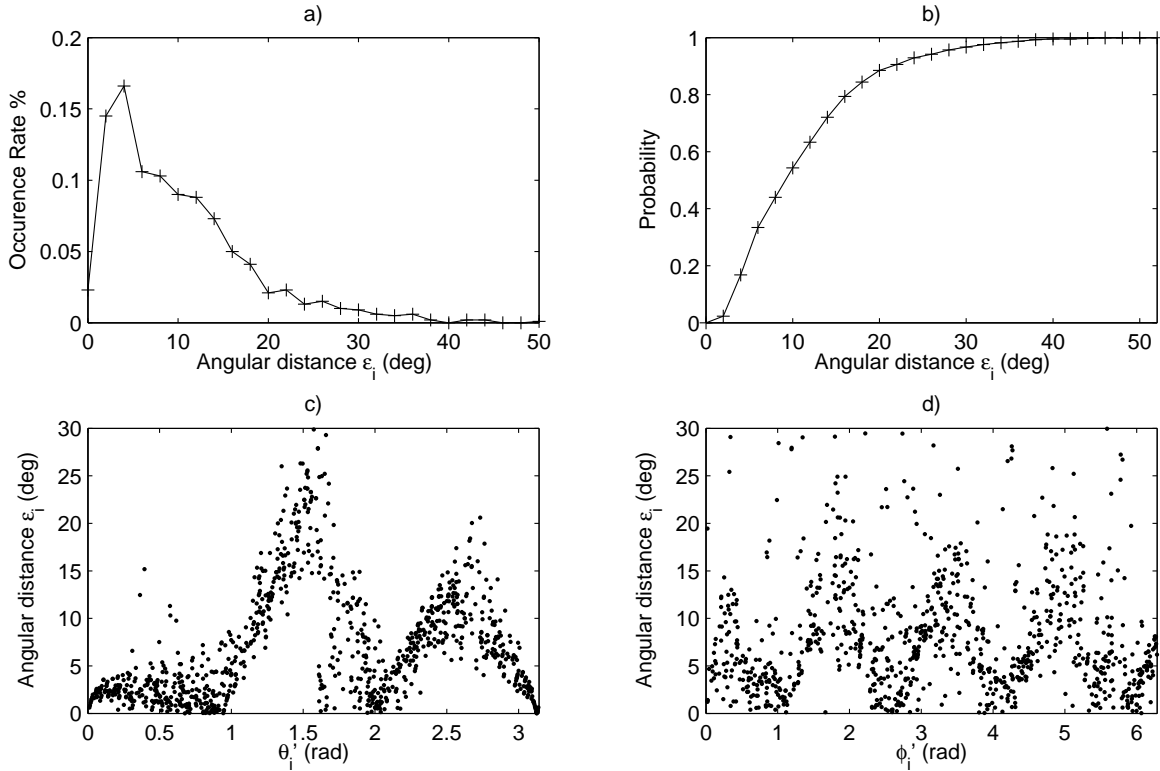


FIG. 5.12 – Résultats obtenus pour la détection d'une onde plane de direction aléatoire, avec 1000 expériences. a) représente la fréquence en fonction de l'angle d'erreur, et b) montre la fonction de répartition. Les figures c) et d) représentent l'erreur en fonction des angles polaire θ et azimutal ϕ .

Nous avons également fait un autre test, en simulant la matrice spectrale à l'aide de deux ondes planes de direction et de distance aléatoire. Nous allons calculer la FDO F pour chacune des matrices, avec une valeur de μ fixée. Puis, la direction des pics est déterminée. Pour cela, le programme calcule le gradient de la FDO ∇F , et détermine toutes les directions (θ, ϕ) tel que $\nabla F(\theta, \phi) = 0$. On ne conserve que les deux directions (θ, ϕ) telles que le gradient de F s'annule, et telles que $F(\theta, \phi)$ soit maximal. Il est possible que les directions calculées ne correspondent pas vraiment à des pics : c'est pourquoi nous avons déterminé à l'aide de la matrice hessienne, si la fonction était concave au voisinage de cette direction. Cela nous à permis de calculer le nombre de pics correctement déterminés. Nous avons, pour chacun des tests i , déterminé les angles d'erreurs $\epsilon_i^1, \epsilon_i^2$ avec les deux véritables directions de propagation. La moyenne $\bar{\epsilon}$ et l'écart type σ_ϵ de l'erreur ont été calculés.

$$\bar{\epsilon} = \frac{1}{N} \sum_{i=1}^N \frac{\epsilon_i^1 + \epsilon_i^2}{2} \text{ et } \sigma_\epsilon = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{(\epsilon_i^1 - \bar{\epsilon})^2 + (\epsilon_i^2 - \bar{\epsilon})^2}{2}}$$

La Table 5.3 permet d'analyser les résultats obtenus. En prenant tout l'échantillon on observe une erreur moyenne de 18.7° , avec un écart-type de 23.2° . L'erreur et l'écart-type diminue lorsque l'on ne regarde que les cas où 2 pics ont été correctement détectés, ce qui correspond à environ 60% de l'échantillon total. Nous voyons également sur le Tableau 5.3 que l'erreur moyenne est bien plus grande lorsque la distance entre les deux directions est faible ($a < 0.3$, 15.9% des cas) ou lorsque la direction se trouve près des pôles (28.5% des cas).

Sélection	$\bar{\epsilon}$	σ_ϵ	Nombre	%
Tout l'échantillon	18.7°	23.2°	10601	100 %
2 pics correctement détectés	16.9°	21.5°	6274	59.2 %
$a < 0.3$	25.2°	28.1°	1694	15.9 %
$\sin(\theta_i) < \sin(\pi - 0.3)$	26.1°	34.6°	3019	28.5 %

TAB. 5.3 – Résultat du test de détection de 2 ondes planes de directions aléatoires. La moyenne $\bar{\epsilon}$ et l'écart-type σ_ϵ de l'angle d'erreur sont calculés sur différents sous-ensembles de l'échantillon total.

5.4 Discussion

En résumé dans ce chapitre nous avons étudié la résolution du problème de calcul de la fonction de distribution des ondes, par régularisation entropique. Dans un premier temps les algorithmes que nous avons développés ont été testés dans le cas du vide, sur des données synthétiques. Les résultats obtenus par les algorithmes 1 et 2 sur ces données, sont satisfaisants dans tous les exemples que nous avons présentés (somme de masses de Dirac ou fonction continue). Nous avons également essayé de résoudre le problème lorsque plusieurs modes de polarisation sont présents : les résultats ne sont pas totalement satisfaisants car on obtient

une erreur bien trop grande par rapport au niveau de bruit. Cela peut s'expliquer par le fait, qu'avec la même donnée, nous essayons de calculer 2 FDO au lieu d'une.

Une fois ces tests effectués, nous sommes passés à des tests numériques dans le cas du plasma, pour des données mesurées à bord du satellite FREJA. Nous avons traité ces données avec la méthode de régularisation entropique, et nous avons comparé les résultats avec ceux obtenus par *Santolík et Parrot* [53] sur les mêmes données, via les méthodes des pics gaussiens et du maximum d'entropie [25, 26]. Il n'apparaît pas de grand décalage entre toutes les solutions calculées. Les FDO calculée par régularisation entropique sont cependant plus plates que les autres, ce qui signifie que leur entropie est plus grande. Les solutions obtenues à l'aide des méthodes de maximum d'entropie et des pics gaussiens, semblent fournir des informations plus précises sur la FDO. Nous n'avons cependant aucun moyen de vérifier si cette information est "réelle" ou non, puisque le problème est hautement indéterminé. C'est pourquoi il est préférable d'obtenir des solutions contenant peu d'informations.

Le choix du paramètre μ a été fait en fonction du principe de décalage de Morozov [22], afin d'obtenir des solutions stables. Nous avons vérifié cette stabilité à l'aide de tests statistiques, de façon globale et locale. De façon globale, la variation moyenne de l'énergie des solutions soumises à une perturbation est de l'ordre de 21%, ce qui montre que les solutions sont bien stables. Cette stabilité est également vérifiée de façon locale, où pour une perturbation de norme 1 l'erreur moyenne maximale observée est de 7%. Nous observons, que les fonctions donnant l'erreur moyenne dans toutes les directions ont globalement la même forme que la FDO.

Chapitre 6

Résultats numériques avec d'autres opérateurs de régularisation

Dans le chapitre précédent, nous avons étudié la résolution du problème inverse de l'identification des directions de propagation d'une onde électromagnétique. Pour cela, nous avons utilisé des méthodes basées sur la régularisation entropique, qui ont été développées à la Section 3.4 et dans [42]. Ces méthodes ont été généralisées dans le chapitre 4, pour des espaces de Banach, en considérant un opérateur de régularisation s.c.i., Gâteaux-différentiable et strictement convexe. Nous allons dans ce chapitre présenter quelques exemples numériques pour illustrer les résultats de cette généralisation.

Le problème considéré est celui de la détermination d'une fonction, définie sur un compact, à partir de ses valeurs en un nombre fini de points. C'est clairement un problème inverse hautement indéterminé puisque nous voulons déterminer un élément d'un espace vectoriel de dimension infinie à partir d'un vecteur de dimension finie. Pour faire l'inversion nous allons utiliser différents opérateurs de régularisation afin d'observer les différences obtenues entre les solutions. Nous considérerons dans un premier temps des opérateurs construits à partir de la norme de l'espace de Sobolev H_0^1 . Ensuite, nous utiliserons une norme L^p . Le choix de l'opérateur de régularisation induit des différences importantes au niveau de la résolution du problème. Par exemple, avec la norme H_0^1 , la solution du problème est le point fixe d'une application linéaire, alors qu'avec la norme L^p ($p \neq 2$) nous devons trouver le point fixe d'une fonction non-linéaire.

6.1 Inversion dans $H_0^2(E)$

Soit une fonction $f : E \rightarrow \mathbb{R}$ où E est un compact à bord régulier et $\{x_i\}_{1 \leq i \leq n} \subset E$ un ensemble de points. Nous allons nous intéresser aux problèmes de la détermination de la fonction f à partir des mesures de cette fonction sur l'ensemble $\{x_i\}_{1 \leq i \leq n}$. Ce problème d'approximation est bien connu et de nombreuses méthodes ont été développées pour le résoudre. On peut, par exemple, utiliser une base (tronquée) pour décomposer f (méthode de Galerkin [70]) et la calculer par identification. Dans ce chapitre, nous allons utiliser un opérateur de régularisation pour calculer la fonction, notre but étant plutôt de souligner

l'importance du choix de l'opérateur par rapport au résultat souhaité.

En supposant que $f \in \mathcal{C}^\infty(E)$, on peut définir l'opérateur $\psi : \mathcal{C}^\infty(E) \rightarrow \mathbb{R}^n$ par

$$\mathbb{R}^n \ni y = \psi[f], \text{ avec } \forall i \in \{1, \dots, n\}, y_j = f(x_i) = \int_E f(x) \delta_0(x - x_i).$$

Cette définition ne nous satisfait pas, car nous préférons éviter d'utiliser les mesures de Dirac. Nous préférons travailler dans l'espace de Hilbert $L^2(E)$ ou dans un espace de Sobolev. Pour remplacer ces mesures de Dirac nous allons utiliser une approximation de l'unité, que nous allons construire à l'aide d'une fonction gaussienne. On pose $q : E \rightarrow \mathbb{R}^n$ où

$$\forall i \in \{1, \dots, n\}, q_i(x) = C_i \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right);$$

les C_i sont des constantes de normalisation et le paramètre $\sigma \in \mathbb{R}^+$ est choisi assez petit de manière à ce que pour $f \in \mathcal{C}^\infty(E)$ on ait

$$\forall i \in \{1, \dots, n\}, \int_E q_i(x) f(x) dx \approx f(x_i).$$

On définit l'opérateur linéaire $\psi : L^2(E) \rightarrow \mathbb{R}^n$ par

$$\psi[f] = \int_E q(x) f(x) dx.$$

Le problème inverse à résoudre est celui de la détermination de f , à partir de la donnée $y \in \mathbb{R}^n$ via l'équation $y = \psi[f]$. Ce qui revient à dire, que nous voulons déterminer la fonction f à partir d'un nombre fini de mesures, le problème est donc hautement indéterminé.

Pour résoudre le problème inverse, nous allons utiliser l'opérateur de régularisation décrit à la page 59 et nous placer dans l'espace de Sobolev $H_0^2(E)$ [11]. On détermine donc la solution du problème inverse, en résolvant

$$\min_{f \in H_0^2(E)} \|y - \psi[f]\|_{\mathbb{R}^n}^2 + \mu \sum_{j=1}^n \|\partial_{x_j} f\|_{L^2}^2.$$

D'après la Proposition 8 et les calculs effectués page 59, la condition d'optimalité du problème est :

$$\Delta f = \frac{-1}{\mu} \psi^*(\ell) \text{ p.p., avec } \ell = y - \psi[f].$$

La fonction f est donc la solution du problème de Poisson

$$\Delta f = u, \text{ sur } E, \text{ avec } u = \frac{-1}{\mu} \psi^*(\ell) \in L^2(E) \quad (6.1)$$

Pour calculer la solution du problème de Poisson sur E , nous devons déterminer la fonction de Green $G : E \times E \rightarrow \mathbb{R}$ de l'ensemble E , i.e.

$$\forall x, \xi \in E, \Delta G(x, \xi) = \delta_0(x - \xi) \text{ dans } H^{-2}(E)$$

Grâce à G , la solution $f \in H_0^2(E)$ du problème de Poisson $\Delta f = u$ s'écrit

$$\forall \xi \in E, \quad f(\xi) = \int_E G(x, \xi) u(x) dx.$$

Nous allons nous placer sur le compact $E = [0, 1]^2$ pour faire un exemple numérique. Nous allons d'abord déterminer la fonction de Green de l'ensemble E , c'est le but de la proposition suivante.

Proposition 16 *La fonction de Green $G : E \times E \rightarrow \mathbb{R}$ de l'ensemble $E = [0, 1]^2$ associée au problème de Poisson est*

$$G(x, y, \xi, \nu) = 4 \sum_{n, m \in \mathbb{Z}} \frac{\varphi_{mn}(x, y) \varphi_{mn}(\xi, \nu)}{\pi^2(m^2 + n^2)}$$

où $x, y, \xi, \nu \in [0, 1]$ et

$$\varphi_{mn}(x, y) = \sin(n\pi x) \sin(m\pi y)$$

La solution f de l'équation (6.1) est donnée par

$$f(\xi, \nu) = \int_E G(x, y, \xi, \nu) u(x, y) dx dy.$$

Démonstration.

1. Les fonctions φ_{mn} sont périodiques : on les identifie avec des fonction de $L^2(E')$ avec $E' = [0, 2]^2$, pour calculer leurs coefficients de Fourier $\mathcal{F}(\varphi_{mn})$. On obtient

$$\mathcal{F}(\varphi_{mn}) = \delta_{mn},$$

et on déduit que

$$\sum_{(m, n) \in \mathbb{Z}^2} |\mathcal{F}(G)_{mn}|^2 \leq \sum_{(m, n) \in \mathbb{Z}^2} \frac{1}{(1 + m^2 + n^2)^2} < +\infty.$$

Donc $G \in L^2(E)$ et $T := \Delta G \in H^{-2}(E)$.

2. Soit $f \in H_0^2(E)$ on peut identifier f par périodicité, avec une fonction $f' \in H_0^2(E')$ où $E' = [0, 2]^2$, comme nous l'avons fait pour la fonction φ_{mn} . Dans ce cas les coefficients de Fourier a_{mn} , $(m, n) \in \mathbb{Z}^2$, de la fonction f' vérifient

$$a_{mn} = \int_{E'} f'(x, y) \exp(2\pi i[mx + ny]) dx dy = \int_{E'} f'(x, y) \varphi_{mn}(xy) dx dy \quad (6.2)$$

puisque l'intégrale de f' contre la fonction $x \mapsto \cos(n\pi x)$ est nulle par périodicité :

$$\begin{aligned} \int_0^2 f'(x, y) \cos(n\pi x) dx &= \int_0^1 f'(x, y) \cos(n\pi x) dx + \int_1^2 f'(x, y) \cos(n\pi x) dx \\ &= \int_0^1 f'(x, y) \cos(n\pi x) dx - \int_0^1 f'(x, y) \cos(n\pi x) dx = 0, \end{aligned}$$

et on obtient le même résultat si l'on intègre sur la variable y . Cela justifie l'équation (6.2), Les coefficients de Fourier a_{mn} peuvent donc être calculés en intégrant f' contre φ_{mn} et on a la formule de représentation :

$$f'(x, y) = \sum_{(m,n) \in \mathbb{Z}^2} a_{mn} \varphi_{mn}(x, y).$$

Par périodicité de f' on obtient

$$\int_E f(x, y) \varphi_{mn}(x, y) dx dy = \frac{1}{4} \int_{E'} f'(x, y) \varphi_{mn}(x, y) dx dy.$$

3. On en déduit

$$\begin{aligned} \langle T, f \rangle_{H^{-2}, H_0^2} &= \frac{1}{4} \sum_{(n,m) \in \mathbb{Z}^2} \left(\int_{E'} 4f'(x, y) \varphi_{mn}(x, y) \varphi_{mn}(\xi, \nu) dx dy \right) \\ &= \sum_{(n,m) \in \mathbb{Z}^2} a_{mn} \varphi_{mn}(\xi, \nu) = f(\xi, \nu). \end{aligned}$$

□

Remarque 4 On montre en dérivant deux fois la fonction φ_{mn} que pour $m \neq 0$ ou $n \neq 0$ on a

$$\Delta \varphi_{mn} = -\pi^2(m^2 + n^2) \varphi_{mn}$$

Les fonctions φ_{mn} sont donc des fonctions propres du Laplacien associées aux valeurs propres $-\pi^2(m^2 + n^2)$.

Réécrivons la condition d'optimalité du problème, on a

$$f(\xi, \nu) = \int_E G(x, y, \xi, \nu) \frac{-1}{\mu} \psi^*[\ell](x, y) dx dy, \text{ avec } \ell = y - \psi[f].$$

Comme f dépend de façon linéaire de ℓ on peut réécrire la condition d'optimalité sous la forme

$$\ell - \frac{1}{\mu} M.\ell = y \text{ où } M.\ell = \psi \left[\int_E G(x, y, \xi, \nu) \psi^*[\ell](x, y) dx dy \right]$$

La solution du problème peut donc se calculer en résolvant un système linéaire de dimension n , et la solution est

$$\ell = \left(I - \frac{1}{\mu} M \right)^{-1} y$$

Pour calculer numériquement la solution nous allons résoudre ce système linéaire. Nous avons choisi d'utiliser 5 points $\{x_i\}_{1 \leq i \leq 5} \subset E$ pour les exemples numériques, ces 5 points sont :

$$x_1 = (0.5, 0.5), x_2 = (0.2, 0.5), x_3 = (0.8, 0.5), x_4 = (0.5, 0.2), x_5 = (0.5, 0.8).$$

La largeur des gaussiennes a été fixée en prenant $\sigma^2 = 0.005$, de manière à obtenir des pics assez abrupts. La donnée $y \in \mathbb{R}^5$ que nous avons choisie est

$$y_1 = 1, y_2 = 0.8, y_3 = 0.6, y_4 = 0.8, y_5 = 0.6;$$

cette donnée a ensuite été perturbée par un bruit blanc, le rapport signal sur bruit étant de 2.9%.

Pour déterminer une solution satisfaisante, nous avons choisi une valeur de μ telle que la solution obtenue ait une erreur de 3.5% (l'erreur est supérieure au niveau de bruit mais quand même assez faible). La Figure 6.1 représente la fonction obtenue, nous voyons que cette fonction a une forme correspondant bien à la donnée y . Nous voyons également que la fonction obtenue vérifie bien la contrainte de nullité sur le bord de E . Nous allons dans la sous-section suivante, tâcher d'éliminer cette contrainte en supposant que la fonction que f est périodique et en changeant l'opérateur de régularisation.

6.2 Inversion dans $H^2(E)$ avec des conditions périodiques

Nous supposons maintenant que f est une fonction périodique, et nous nous placerons dans l'espace $H^2(E)$ pour résoudre le problème inverse. Une idée qui apparaît naturellement pour cette résolution, est de déterminer la fonction f par transformation de Fourier, en supposant qu'elle est la somme de n fréquences. Ce n'est pas l'approche que nous avons choisie : nous allons utiliser un opérateur de régularisation. Nous ne pouvons plus utiliser l'opérateur de l'exemple précédent, car il n'est pas strictement convexe dans l'espace $H^2(E)$, nous allons donc utiliser la norme de $H^1(E)$

$$\|f\|_{H^1}^2 = \|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2.$$

Pour obtenir f , il faut résoudre le problème d'optimisation :

$$\min_{f \in H^2(E)} \|y - \psi[f]\|_{\mathbb{R}^n}^2 + \mu \|f\|_{H^1(E)}^2.$$

et la condition d'optimalité est

$$(I + \Delta)f = \frac{-1}{\mu} \psi^*[\ell] \text{ p.p., avec } \ell = y - \psi[f].$$

En utilisant la transformée de Fourier cette condition devient

$$f = \frac{-1}{\mu} \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\psi^*[\ell])_{mn}}{1 + m^2 + n^2} \right)$$

La linéarité nous permet d'écrire la condition d'optimalité sous la forme

$$(I - N).\ell = y, \text{ avec } N.\ell = \psi \left[\frac{-1}{\mu} \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\psi^*[\ell])_{mn}}{1 + m^2 + n^2} \right) \right],$$

il suffit de résoudre ce système linéaire pour obtenir la solution.

Nous avons pris pour cet exemple le E de l'exemple translaté de $-\frac{1}{2}$ sur les deux axes. Nous avons en conséquence translaté de la même façon les points x_i :

$$x_1 = (0.0, 0.0), x_2 = (-0.3, 0.0), x_3 = (0.3, 0.0), x_4 = (0.0, -0.3), x_5 = (0.0, 0.3).$$

Nous avons utilisé la même donnée y que dans l'exemple précédent, et après perturbation le niveau de bruit est de 3.7%. La Figure 6.2 montre la fonction obtenue par cette méthode, avec une erreur quasiment égale au niveau de bruit (3.7%). Nous observons bien sur cette figure que la contrainte de nullité sur le bord de E a été remplacée par une contrainte de périodicité. Toutefois la solution a tendance à s'aplatir loin des pics, ce qui résulte de la minimisation de la norme H^1 .

Afin de rechercher une solution qui ne tend pas à s'annuler loin des pics, nous allons modifier le problème de minimisation à résoudre. Nous allons imposer le fait que la solution recherchée doit avoir une valeur moyenne très proche de la moyenne de ses valeurs mesurées y . Nous allons résoudre pour cela, le problème

$$\min_{f \in H^2(E)} \|y - \psi[f]\|_{\mathbb{R}^n}^2 + \mu \|f - \bar{y}\|_{H^1}^2.$$

où

$$\bar{y} = \frac{1}{5} \sum_{1 \leq i \leq 5} y_i$$

la condition d'optimalité est

$$(I + \Delta)f = \bar{y} + \frac{-1}{\mu} \psi^*[\ell] \text{ p.p., avec } \ell = y - \psi[f];$$

le point fixe ℓ_μ peut être calculé à l'aide du même système linéaire que dans l'exemple précédent. Nous avons utilisé la même donnée, avec un niveau de bruit de 2.6%. La Figure 6.3 montre la solution obtenue à l'aide de cette méthode, l'erreur étant de 3.1%. Nous voyons bien sur cette figure que la fonction est beaucoup plus plate que les précédentes, ce qui est normal puisqu'elle tend vers la valeur de la moyenne \bar{y} , loin des points $\{x_i\}$.

6.3 Inversion dans $L^p(E)$

Les trois exemples que nous venons de présenter ont tous en commun le fait que la solution puisse se calculer en résolvant un système linéaire. Nous allons maintenant étudier un opérateur de régularisation, tel que la solution du problème s'obtient comme un point fixe en dimension finie, d'une fonction non-linéaire. Pour déterminer ce point fixe nous allons utiliser la suite (4.19) définie page 55.

Le méthode que nous allons utiliser est la méthode de Tikhonov dans L^p avec p pair. Le problème à résoudre est

$$(P_2) \begin{cases} \min \|y - \psi[f]\|_{\mathbb{R}^n}^2 + \mu \|x\|_p \\ x \in L^p(E) \end{cases} .$$

d'après les calculs effectués dans la sous-section 4.6.3, et d'après la Proposition 8, la condition d'optimalité est

$$f = \sqrt[p-1]{\frac{2\|\ell\|^{2(p-1)}}{\mu}} \psi^*[\ell] \quad p.p, \text{ avec } \ell = y - \psi[f].$$

Nous allons utiliser la suite (4.19) pour calculer la solution du problème. Pour l'exemple nous allons prendre $E = [0, 1]^2$, comme précédemment, avec les mêmes points $\{x_i\}_{1 \leq i \leq 5}$ et la même donnée y . Nous allons également choisir $p = 20$.

Nous avons observé une bonne convergence de la suite (4.19) pour τ choisi assez petit. Cependant, pour obtenir une erreur comparable au niveau de bruit, nous avons dû choisir μ très petit et le calcul de la solution a été beaucoup plus long que dans les exemples précédents. La Figure 6.4 montre la fonction obtenue par cette méthode avec une erreur à peu près égale à deux fois le niveau de bruit. Les pics sont beaucoup plus arrondis que dans les exemples précédents. ce qui s'explique par l'utilisation de la racine p -ième qui favorise les valeurs différentes de 0.

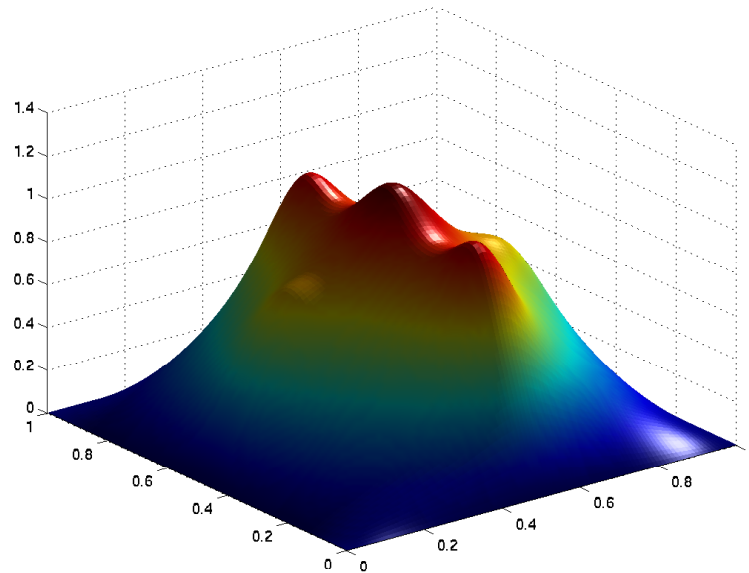


FIG. 6.1 – Solution obtenue en utilisant la norme du gradient comme opérateur de régularisation, avec une contrainte de nullité au bord.

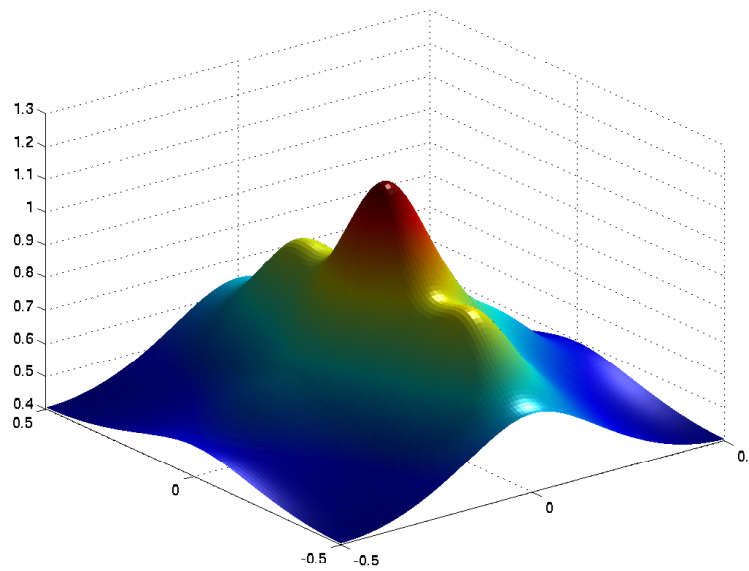


FIG. 6.2 – Solution obtenue en utilisant la norme H^1 comme opérateur de régularisation, sous une condition de périodicité.

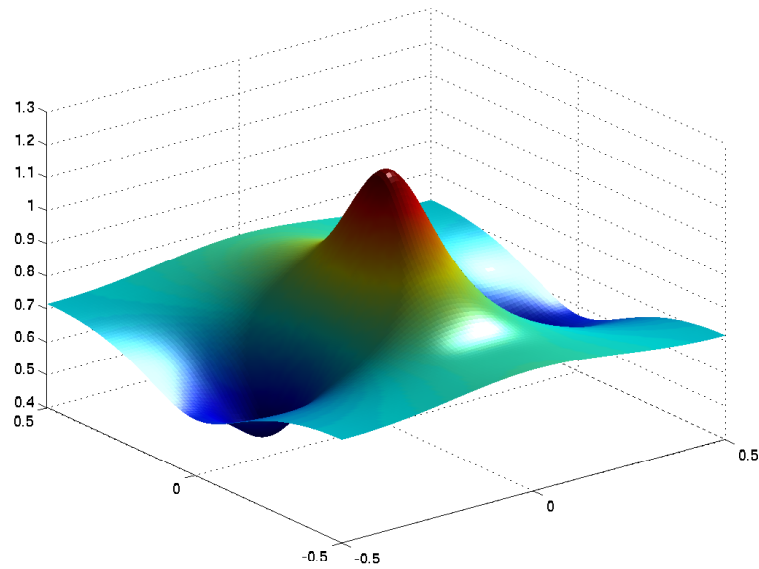


FIG. 6.3 – Cette fonction a été obtenue en utilisant comme opérateur de régularisation, la norme H^1 translatée par la valeur moyenne de y .

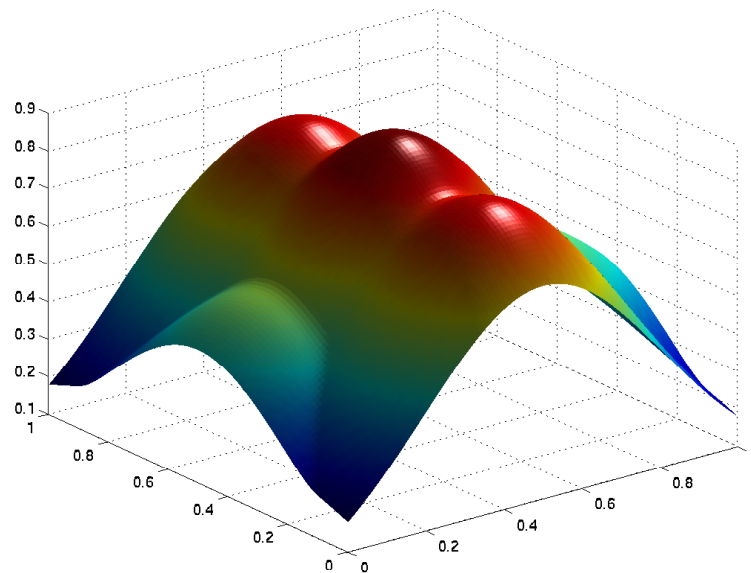


FIG. 6.4 – Solution obtenue à l'aide de la méthode de Tikhonov dans L^p .

Chapitre 7

Conclusion et perspectives

La détermination de directions de propagation d'une onde électromagnétique, à partir de la mesure locale du champ, est d'une importance cruciale dans de nombreux domaines d'applications. Lorsque l'onde étudiée a une direction de propagation unique (i.e. l'onde est considérée plane), le problème de l'identification de cette direction de propagation à partir de la matrice spectrale mesurée est un problème bien posé. Diverses méthodes basées sur l'analyse des vecteurs propres de la matrice spectrale ont été développées dans ce but [32]. Toutefois, ces techniques ne permettent pas toujours d'analyser de façon réaliste les ondes électromagnétiques puisqu'en pratique ces ondes ne sont généralement pas planes à cause des multiples réflexions et réfractions dans le milieu de propagation.

Il y a une trentaine d'années *Storey et Lefevre* [61] ont introduit le concept de fonction de distribution des ondes (FDO). Cette fonction décrit la densité d'énergie de l'onde suivant toutes les directions. Ce concept permet la manipulation d'ondes électromagnétiques constituées d'une somme continue d'ondes planes. *Storey et Lefevre* [61, 62, 63] ont également décrit l'équation qui relie la FDO à la matrice spectrale estimée à partir des mesures du champ. La détermination, par cette équation, de la FDO à partir de la matrice spectrale est un problème mal posé : on peut trouver un ensemble de fonctions de dimension infinie qui vérifient cette équation.

Afin d'interpréter les données d'expériences magnétosphériques, de nombreux auteurs ont utilisé un modèle de FDO paramétrique pour résoudre ce problème inverse. Le calcul de la solution revient alors à l'identification des paramètres du modèle, ce qui est souvent fait dans la pratique grâce à une minimisation au sens des moindres carrés. Par exemple pour la méthode des pics gaussiens, on considère que la FDO est constituée d'une somme de gaussiennes et les paramètres à identifier sont la direction, la puissance et la largeur de chaque pic. Lorsque le nombre de pics gaussiens est correctement fixé, cette méthode fournit des solutions stables. Pour la méthode du maximum d'entropie [25], le modèle est choisi de manière à maximiser l'entropie de la FDO. Cependant l'utilisation de ce modèle ne garantit pas la stabilité de la solution. Pour l'assurer il faut éliminer les noyaux les "moins significatifs" [24]. De façon générale lorsque l'on utilise un modèle de FDO, il faut fixer un certain nombre de paramètres avant le calcul. C'est pour éviter de faire un tel choix et aussi pour ne pas avoir à choisir à un modèle de FDO *a priori* que nous préconisons une méthode de régularisation.

Pour résoudre un problème mal-posé, la théorie de la régularisation consiste à utiliser une famille d'opérateurs indexés par un paramètre $\mu > 0$ appelé paramètre de régularisation. Lorsque le paramètre μ est choisi très petit, on obtient une solution dont l'erreur avec la donnée est faible, mais qui est très instable. A l'inverse, si μ est choisi plus grand, l'erreur sera plus grande mais la solution sera plus stable. Un compromis entre la stabilité et l'erreur doit donc être trouvé. En pratique on ajuste ce paramètre en fonction du niveau de bruit.

L'opérateur de régularisation doit être bien choisi afin que les solutions soient convables vis à vis du problème considéré. Nous avons utilisé l'opérateur d'entropie car il offre plusieurs avantages. Premièrement, la FDO calculée doit être positive puisqu'elle représente une énergie. Or, l'entropie agit comme une fonction barrière, elle force les solutions obtenues à être positives, ce qui nous permet d'éliminer cette contrainte. Deuxièmement, dans un problème mal posé tel que (1.1), il vaut mieux chercher une solution contenant le moins d'information possible afin de préserver l'information essentielle. Il existe beaucoup de solutions du problème qui apportent un grand nombre d'information, mais nous n'avons aucun moyen de dire si toutes ces informations sont bien "réelles". Il a été montré de façon axiomatique [57] que l'entropie est un bon moyen pour quantifier l'information d'une fonction ; en conséquence, la régularisation entropique permet d'obtenir des solutions contenant peu d'information, ce qui est préférable.

La régularisation entropique est souvent utilisée dans l'espace L^1 pour résoudre les problèmes mal-posés. Nous avons cependant préféré travailler dans l'espace L^2 . En effet, dans L^1 , l'entropie permet de définir une stratégie de régularisation [1, 16], mais le modèle de FDO autorisant l'utilisation de L^2 , l'étude mathématique se simplifie en raison des bonnes propriétés géométriques de l'espace L^2 . Nous avons montré dans ce cas que pour toutes valeurs de $\mu > 0$, la solution du problème vérifiait une équation de point fixe en dimension finie. Cette condition d'optimalité étant très simple, nous avons cherché à l'utiliser dans le calcul numérique des solutions. Un premier algorithme a ainsi été établi pour calculer le point fixe recherché, et nous avons montré la convergence de cet algorithme pour toute valeur du paramètre μ . Ce qui montre l'existence de la solution de régularisation entropique pour tout $\mu > 0$. Toutefois la solution obtenue n'est pas totalement satisfaisante car elle ne dépend pas linéairement de la donnée. Pour pallier ce problème, nous avons modifié la méthode de résolution.

Un nouveau problème d'optimisation a été établi pour résoudre le problème inverse avec pour objectif de vraiment minimiser l'entropie au sens des probabilités et de corriger la non-linéarité de la solution. Ce nouveau problème est bien plus complexe que le précédent car il contient un terme bilinéaire qui rend la fonction coût non-convexe. Ce terme bilinéaire provient de la paramétrisation de la FDO $F = \alpha f$, où α est l'énergie et f une densité de probabilité. Nous avons donc établi un second algorithme pour résoudre ce nouveau problème. Cet algorithme dérivé du premier permet de déterminer un minimum local. Nous avons montré que cet algorithme convergeait si le paramètre μ était assez grand. Dans le cadre de la physique expérimentale, le fait que μ soit grand n'est pas très contraignant puisque la donnée est toujours entachée d'un bruit : nous sommes ainsi obligés de choisir μ assez grand pour assurer la stabilité de la solution. L'algorithme permet d'obtenir une solution qui dépend de façon positivement homogène de la donnée initiale, ce qui est très important, car les données d'expériences sont généralement mises à l'échelle pour les applications numériques. L'écriture

du problème dual du problème de régularisation entropique permet de voir que l'algorithme 1 n'est rien d'autre qu'une méthode de gradient sur ce dual. Nous voyons également qu'il est possible de s'affranchir du paramètre de régularisation μ en le remplaçant par le niveau de bruit.

La solution du problème de régularisation entropique s'obtenant très simplement à l'aide d'un point fixe en dimension finie, nous avons cherché à généraliser ce résultat dans un cadre plus général. Nous avons donc remplacé l'espace de Hilbert L^2 par un espace de Banach réflexif et nous avons pris un opérateur de régularisation ϕ s.c.i., Gâteaux-différentiable convexe. Nous avons établi (Proposition 7) sous quelles hypothèses la solution du problème pouvait se calculer à l'aide d'un point fixe en dimension finie. Ensuite nous avons donné une condition sur ϕ pour que la solution obtenue soit stable (Proposition 8). Ici, nous ne considérons pas la stabilité au sens de Hadamard (c'est à dire la continuité de la solution par rapport à la donnée) mais nous voulons que la solution dépende de façon Lipschitz par rapport à la donnée. La résolution du problème a été étudiée à l'aide de la condition d'optimalité et une suite a été établie pour trouver l'optimum. Nous avons cherché à démontrer la convergence de cette suite. Le théorème de point fixe de Banach a été utilisé, de manière analogue au cas de la régularisation entropique, pour montrer la convergence de la suite. Nous avons également, dans le cas de fonctions moins régulières, cherché des conditions pour que la suite reste bornée. Cela nous a permis de redémontrer l'existence du point fixe à l'aide du théorème de Schauder. C'est également une première étape pour montrer la convergence de la suite dans ce cas.

Les algorithmes 1 et 2 ont été construits pour le problème de l'identification de la fonction de distribution des ondes. Nous avons effectué différents tests. Dans un premier temps le cas du vide est considéré avec des données simulées. Les résultats obtenus par les deux algorithmes sont tout à fait satisfaisants pour détecter la direction et la puissance d'une onde plane ainsi que pour identifier une onde constituée d'une somme de trois ondes planes. Nous avons également fait un test dans le cas où l'onde électromagnétique est constituée d'une somme de deux ondes planes de polarisations différentes. Dans ce cas, la solution déterminée par la méthode de régularisation entropique a une erreur très grande par rapport au niveau de bruit et elle n'est donc pas utilisable. Toutefois, nous observons que malgré cette erreur, la solution localise assez bien la direction des ondes dans chacun des modes (Figure 5.6).

Nous avons ensuite testé l'algorithme 2 pour le calcul de la FDO dans un plasma magnétosphérique. Les données que nous avons utilisées sont six matrices spectrales enregistrées par le satellite FREJA, ces six matrices correspondent à six point temps-fréquence du spectrogramme d'un souffle ELF. L'analyse de ces données via la FDO ayant déjà été faite par *Santolík et Parrot* [53], en utilisant diverses méthodes (maximum d'entropie, pic gaussiens,...etc), cela nous a permis de confronter nos résultats avec ceux déjà obtenus. Cette comparaison montre qu'il n'y a globalement pas de contradiction entre les solutions obtenues par régularisation entropique et celles obtenues par les autres méthodes. Les solutions de régularisation entropique sont cependant plus "aplaties" que les autres, ce qui s'explique par la maximisation de l'entropie. Il faut également remarquer qu'avec les algorithmes 1 et 2, la règle de choix de μ et τ étant implantée, les solutions sont déterminées de façon automatique ce qui n'est pas le cas pour les méthodes utilisées dans [53]. En comparant les solutions obtenues (Figure 5.10), on peut penser que les solutions calculées par la méthode des pics gaussiens ou

du maximum d'entropie sont plus précises et qu'elles apportent plus d'information que les solutions de régularisation entropique. Cependant il n'y a aucun moyen de vérifier si l'information supplémentaire apportée par ces solutions est vraiment fiable. Dans le cadre d'un problème hautement indéterminé tel que (1.1), il est préférable de rechercher des solutions qui n'apportent pas trop d'information, comme les solutions que nous avons déterminées par régularisation entropique. Ces solutions peuvent également servir comme point de départ, pour avoir une idée globale de la FDO, ensuite une analyse plus fine est possible via d'autres méthodes.

Nous avons ensuite effectué des tests numériques pour illustrer les résultats théorique sur la généralisation de la méthode. Nous avons cherché à déterminer une fonction, définie sur un compact, à partir de sa mesure sur un nombre fini de points. Pour l'inversion nous avons utilisé différents opérateurs de régularisation ce qui nous a permis d'illustrer l'importance du choix de cet opérateur.

Il reste des points à démontrer sur la méthode de régularisation entropique pour le calcul de la FDO. Le premier point est la démonstration de la convergence de l'algorithme 2 pour toutes valeurs de μ . Du point de vue numérique nous n'avons pas observé de problème de convergence pour des petites valeurs de μ , mais dans la démonstration nous n'avons pas pu nous passer du fait que μ devait être assez grand. L'unicité de la solution est aussi un point à étudier.

Nous avons généralisé la méthode pour d'autres opérateurs de régularisation et pour des espaces de Banach abstraits. Beaucoup de choses sont encore à étudier dans ce domaine : nous nous sommes par exemple restreints aux espace de Banach réflexifs et aux opérateurs ϕ strictement convexes. Dans le cas où ϕ est convexe (pas strictement) la solution du problème n'est pas forcément unique, et l'inverse de $\nabla\phi$ devient multivoque. Le cas avec contraintes n'a pas été étudié en détails, nous nous sommes contentés de revenir à un problème sans contraintes en modifiant l'opérateur de régularisation. Enfin la convergence de la suite sous des hypothèses de régularité plus faible reste à étudier. Les tests numériques, dans ce cas, n'ont pas révélé de problème de convergence. Numériquement, on peut étendre ces méthodes à d'autres types de problèmes comme par exemple au traitement de l'image.

Du point de vue de la physique, la méthode n'a pas encore été testée sur un nombre suffisant de données. Comme la méthode est automatique, il faudrait en effet voir si elle ne pourrait pas être utilisée de façon systématique, ce qui autoriserait la création de bases de données en vue d'études statistiques. Le problème de la détermination de la FDO lorsque deux modes d'ondes sont présents n'est pas résolu de façon correcte par la méthode. Enfin, il faudrait quantifier l'impact de la non-linéarité du modèle sous-jacent sur les solutions.

Annexe A

Determination of a Power Density by an Entropy Regularization Method

Olivier PROT^{*†}, Maïtine BERGOUNIOUX^{*}, Jean Gabriel TROTIGNON[†]

[†]*LPCE/CNRS, UMR 6115, 3A, Avenue de la Recherche Scientifique 45071 Orléans Cedex
2, France*

^{*}*MAPMO, UMR 6628 Université d'Orléans-CNRS, BP. 6759 45067 Orléans Cedex 2,
France*

prot@cnrs-orleans.fr, maitine.bergounioux@labomath.univ-orleans.fr

Abstract. The determination of directional power density distribution of an electromagnetic wave from the electromagnetic field measurement can be expressed as an ill-posed inverse problem. We consider the resolution of this inverse problem via a maximum entropy regularization method. A finite dimensional algorithm is derived from optimality conditions, and we prove its convergence. A variant of this algorithm is also studied. This second one leads to a solution which maximizes entropy in the probabilistic sense. Some numerical examples are given.

KEYWORDS : *Inverse problem, regularization methods.*

AMS : 49M30, 65R32.

A.1 Introduction

In this paper, we present entropy regularization algorithms in order to determine an electromagnetic wave propagation directions, from the measurement of the six components of the electromagnetic field. Most of existing methods assume that the electromagnetic wave is similar to a single plane wave. In this case, for a fixed frequency, the electromagnetic field is fully described by the wave normal direction vector k . Nevertheless, this assumption is generally too restrictive. For a more realistic analysis of an electromagnetic wave in a plasma, Storey and Lefeuvre have introduced the concept of wave distribution function (WDF) [24]. This function describes the wave energy density distribution for every frequency and propagation modes. The WDF f is related to the spectral matrix V of the electromagnetic field

component by a fredholm integral of the first kind [24]

$$V(w) = \int q(k, w) f(k, w) dk \quad (\text{A.1})$$

where k is the wave normal vector and w the frequency. In practice the WDF is computed for the most significant frequency (often there is only one significant frequency) of the spectrogram. The integrating kernel q is a known function depending of the propagation media. The spectral matrix V is defined with the measured electromagnetic field $e : \mathbb{R}^+ \rightarrow \mathbb{R}^s$ by $V = \hat{e}(w)\hat{e}^*(w)$ where \hat{e} denotes the Fourier Transform of e and $s \in \mathbb{N}^*$ is the number of field components. It is an Hermitian matrix. This definition of the spectral matrix is an approximation of the real spectral matrix. Indeed, e is a random signal and we should use mean values of e . Here we assume e is deterministic. In what follows, we will identify the spectral matrix to vector of \mathbb{C}^n where $n = s^2$. Solutions of the inverse problem of determining f from measurements of V were proposed by Lefeuvre using a maximum entropy method [25]. However, for a fixed frequency, we have to solve the integral equation problem

$$V = \int q(x) f(x) dx$$

where the unknown function f is nonnegative. This inverse problem is known to be ill-posed.

The concept of WDF can be transposed to the case of electromagnetic wave propagating in vacuum. Equation (A.1) remains valid if we use the vacuum kernels instead of the plasma kernels and if the electromagnetic wave has a single polarization mode [27]. The use of the WDF concept for electromagnetic wave propagating in vacuum has been studied for the interpretation of ground penetrating radar investigations, and in particular the one proposed for the NetLander mission to Mars [27]. The aim of this instrument was to explore the first kilometers of Mars subsurface to study its geological structure and look for water.

Here we use a maximum entropy regularization method to solve this problem. We minimize the quantity

$$\left\| V - \int q(x) f(x) dx \right\|_{\mathbb{C}^n}^2 + \mu H(f)$$

with the constraint $f \geq 0$, where H is an negentropic term (to be defined later) and μ a regularization parameter. In fact, H is not the negentropy in the probabilistic sense since f is not a density. But the minimization of H leads to a smooth solution. The main disadvantage of the maximum entropy solution of [25], is that the constraints on the solution are too strong. The regularization process provides a relaxed problem and the error we introduce allows to search a solution in a much wider domain. Moreover, we obtain a solution not far from the data that looks like a minimizer of H .

The maximum entropy regularization method is a useful tool to solve such inverse problems. Amato et al. [1] have studied the convergence of this method to show that it is a correct regularization process. This convergence is also studied in [16] by making a link with the classical Tikhonov method [64, 65]. A generalization is investigated in [30].

The mathematical model of the problem is described in section A.2. We define the negentropy in section A.3. In section A.4 we set the optimization problem : the feasible domain of

this problem has to be relaxed to find optimality conditions. We present two algorithms, but the solutions we obtain do not minimize negentropy. So we modify the optimization problem section A.5. Finally we present numerical tests in section A.6.

A.2 Mathematical model

In this section we present the mathematical model. We consider (E, \mathcal{A}, σ) a measured space, where E is a compact subset of \mathbb{R}^p , $p \geq 1$ and the measure σ verifies $\sigma(E) < +\infty$. The power density on (E, \mathcal{A}) can be defined as :

Definition 1 *The Power density is a couple (α, m) where $\alpha \in \mathbb{R}^+$ and m is a probability measure on (E, \mathcal{A}) . Let $A \subset E$: the power π_A of the subset A is given by*

$$\pi_A = \alpha m(A).$$

The aim of this paper is to determine a power density (α, m) which verifies the equation

$$\alpha \int_E q dm = V$$

where $V \in \mathbb{C}^n$ is known and $q \in L^2(E, \mathbb{C}^n, \sigma)$ is the integration kernel. In plasma physics, we have to solve this kind of problem to determine the power density distribution of an electromagnetic wave m and the total power α from the measurement of the electromagnetic field components. In this case we typically have $n = 36$.

The set of probability measures is too large and we will only consider measures that are continuous with respect to the measure σ . We denote $\mathbb{H} = L^2(E, \mathbb{R}, \sigma)$ ($\subset L^1(E, \mathbb{R}, \sigma)$ thanks to our assumptions). For all $F \in \mathbb{H}$ and $F \geq 0$ *s.a.e* we can define a power density $(\alpha, \frac{F}{\alpha} d\sigma)$ where $\alpha = \|F\|_{L^1(E, \mathbb{R}, \sigma)}$ and $\frac{F}{\alpha} d\sigma$ is the measure of density $\frac{F}{\alpha}$ with respect to σ . Note that : $F \mapsto \int_E \alpha q dm$ where $\alpha = \|F\|_{L^1(E, \mathbb{R}, \sigma)}$ and $m = \frac{F d\sigma}{\alpha}$ is a linear bounded operator from \mathbb{H} to \mathbb{C}^n .

More generally, we consider a linear bounded operator $\psi = (\psi_1, \dots, \psi_n) : \mathbb{H} \rightarrow \mathbb{C}^n$, $\psi^* : \mathbb{C}^n \rightarrow \mathbb{H}$ its adjoint operator and we assume that $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$ (R denotes the range). We have to solve

$$\psi[F] = V \tag{A.2}$$

From the Riesz Theorem, we deduce there exist n functions $q_i \in L^2(E, \mathbb{C}, \sigma)$, $i = 1, \dots, n$ such that $\psi_i[F] = \langle q_i, F \rangle_{\mathbb{H}}$. These functions are integration kernels, and we have

$$\forall l \in \mathbb{C}^n, \quad \psi^*(l) = \text{Re} \left(\sum_{i=1}^n l_i \bar{q}_i \right)$$

where \bar{q}_i denotes the conjugate complex of q_i . The condition $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$ is equivalent to

$$\forall i = 1, \dots, n, \quad q_i \in L^\infty(E, \mathbb{C}, \sigma).$$

The problem of solving equation (A.2) is an ill-posed problem. Indeed ψ is an operator from \mathbb{H} (an infinite dimensional Hilbert space) to the finite dimensional Hilbert space \mathbb{C}^n .

So, the operator ψ is not injective and there is no uniqueness of the solution (if it exists). In addition, We want to determine a solution which is also stable, that is continuous with respect to the data V . This will be crucial for the physical and numerical stability. To deal with this challenge, we use a maximum entropy regularization method.

The principle of Tikhonov's regularization method is to minimize the quantity

$$\|\psi[F] - V\|_{\mathbb{C}^n}^2 + \mu \Omega(F) \quad (\text{A.3})$$

where μ is called the regularization parameter and Ω is a suitable regularizing functional. This method is equivalent to minimize the functional Ω on the set $\{\|\psi[F] - V\|_{\mathbb{C}^n} \leq \delta(\mu)\}$ [1]. In this paper we use an entropic term as the regularizing functional (see section A.3) and we restrict the domain to the Hilbert space \mathbb{H} whereas usually, maximum entropy regularization is performed in L^1 . We will see in the following section that there is no problem to define the entropy. First we recall some definitions.

A.3 About entropy

Let be the function $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}$ defined by

$$\phi(x) = \begin{cases} x \ln(x) & \text{if } x > 0 \\ 0 & \text{else.} \end{cases}$$

The notation $m_1 \ll m_2$ means that the measure m_1 is absolutely continuous with respect to the measure m_2 .

Definition 2 Let $f, g \in L^1(E, \mathbb{R}, \sigma)$ such that $f d\sigma$ and $g d\sigma$ are two probability measures and $f d\sigma \ll g d\sigma$. We define the relative information content of f with respect to g by

$$I(f, g) = \int_E f \ln \frac{f}{g} d\sigma \quad (\text{A.4})$$

If the condition $f d\sigma \ll g d\sigma$ is not verified then $I(f, g) = +\infty$. If g is the non informative probability density (see the definition below) then $I(f, g)$ is called the information content of f . The negentropy (negative entropy) of f is then defined by

$$H(f) = I(f, g)$$

The non-informative probability density is a known function of the model. Physically, it is the probability density of a noise measured in the system. For example, in the case of an electromagnetic wave in vacuum WDF, E is the unit sphere. Now, since isotropy occurs (there is no privileged direction of propagation for a plane wave in vacuum) we deduce that the non informative density probability is constant over the unit sphere. So, we may assume that the non-informative probability density is given by

$$\forall x \in E, \quad g(x) = \frac{1}{\sigma(E)} \quad (\text{A.5})$$

For the sake of simplicity, we will suppose that $\sigma(E) = 1$, so the negentropy of f is

$$H(f) = \int_E \phi \circ f \, d\sigma \quad (\text{A.6})$$

Entropy of a probability density can be seen as a “distance” between f and the density of the non-informative probability g . To calculate the solution of the inverse problem we minimize a functional involving negentropy. Thus, we determine the solution which contains the less information with respect to the density g . From the physical point of view, this allows to preserve the physical significant information. So, it is primordial to know the non-informative density g quiet accurately. One can refer to [57] for the axiomatic derivation of the maximum entropy principle. Next lemma gives some properties of H .

Lemma 1 *In the sequel we denote $K = \{f \in L^2(E, \mathbb{R}, \sigma) | f \geq 0 \text{ } \sigma.a.e.\}$. For all $f \in K$, $-e^{-1} \leq H(f) < +\infty$. The functional $H : L^2(E, \mathbb{R}, \sigma) \rightarrow \mathbb{R}$ is lower semi-continuous (l.s.c), strictly convex on K and verifies : $\forall \varepsilon > 0, \forall f \in K$ such that $f \geq \varepsilon \text{ } \sigma.a.e.$, $\forall g \in K, \forall \lambda \in [0, 1]$*

$$H(f + \lambda(g - f)) - H(f) = \lambda \int_E (1 + \ln f(x))(f - g)(x) \, d\sigma(x) + o(\lambda) \quad (\text{A.7})$$

Proof. Let be $f \in L^2(E, \mathbb{R}, \sigma)$,

$$H(f) = \int_{\{x \in E | f(x) < 1\}} \phi \circ f(x) \, d\sigma(x) + \int_{\{x \in E | f(x) \geq 1\}} \phi \circ f(x) \, d\sigma(x);$$

as

$$\forall x \in [0, 1], \quad -e^{-1} \leq x \ln x \leq 0$$

and

$$\forall x \geq 1, \quad 0 \leq x \ln x < x^2$$

so we deduce

$$H(f) \leq \|f\|_{L^2(E, \mathbb{R}, \sigma)} < +\infty \text{ and } H(f) \geq -e^{-1}$$

The proof of the lower semi-continuity of functional H can be found in [30, 1].

We now prove (A.7) : let $f \in K$ such that $f \geq \varepsilon$ let $g \in K$. We have

$$\text{For a.e. } x \in E, \quad \phi(f(x) + \lambda(g(x) - f(x))) - \phi(f(x)) = \lambda(1 + \ln f(x))(g(x) - f(x)) + o(\lambda)$$

since ϕ is derivable on \mathbb{R}^{+*} and $\phi'(x) = 1 + \ln x$.

The functional H is strictly convex by the strict convexity of ϕ on \mathbb{R}^+ . □

A.4 A penalized problem

We now define the penalized cost functional or smoothing functional J_μ we want to minimize

$$\begin{cases} J_\mu : \mathbb{H} & \longrightarrow \mathbb{R} \\ F & \longmapsto \|V - \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F) \end{cases} \quad (\text{A.8})$$

where $\mu > 0$ is a regularization parameter. We have

Lemma 2 *The functional J_μ is l.s.c and strictly convex on K . In addition, if $\varepsilon > 0$, $f, g \in K$ with $f \geq \varepsilon$ σ .a.e, then*

$$\forall \lambda \in]0, 1[, \quad J_\mu(f + \lambda(g - f)) - J_\mu(f) = \lambda \langle DJ_\mu(f), g - f \rangle_{\mathbb{H}} + o(\lambda) \quad (\text{A.9})$$

where

$$DJ_\mu(f) = \mu(1 + \ln f) - 2\psi^*[V - \psi[f]] \quad (\text{A.10})$$

and

$$J_\mu(g) \geq J_\mu(f) + \langle DJ_\mu(f), g - f \rangle_{\mathbb{H}} \quad (\text{A.11})$$

Proof. The functional $F \mapsto \|V - \psi[F]\|_{\mathbb{C}^n}^2$ is continuous on \mathbb{H} by continuity of ψ , hence it is l.s.c. We conclude that J_μ is l.s.c on K by Lemma 1. Furthermore, $F \mapsto \|V - \psi[F]\|_{\mathbb{C}^n}^2$ is Fréchet-differentiable and its gradient is $-2\psi^*[V - \psi[F]]$, so equation (A.9) is proved by Lemma 1 as well.

The functional J_μ is strictly convex on K by strict convexity of H and by convexity of the term $\|V - \psi[F]\|_{\mathbb{C}^n}^2$. So, we can write

$$\begin{aligned} J_\mu(f + \lambda(g - f)) &< (1 - \lambda)J_\mu(f) + \lambda J_\mu(g) \\ \frac{J_\mu(f + \lambda(g - f)) - J_\mu(f)}{\lambda} &< J_\mu(g) - J_\mu(f) \end{aligned}$$

Equation (A.11) follows by taking the limit of the last equation when $\lambda \rightarrow 0$. □

The penalized optimization problem stands

$$(\mathcal{P}_\mu) \begin{cases} \min J_\mu(F) \\ F \in K = \{f \in \mathbb{H} \mid f \geq 0 \text{ } \sigma\text{-a.e}\} \end{cases} .$$

The existence of a solution to problem (\mathcal{P}_μ) is not obvious since the cost functional is not coercive in \mathbb{H} . To illustrate this fact let us give a simple counter-example : we set $E = [0, 1]$, σ is the Lebesgue's measure on $[0, 1]$ and $f : E \rightarrow \mathbb{R}$, $x \mapsto x^{-\frac{1}{2}}$. Then $H(f) = 2 < \infty$ and we can build a sequence $\{f_k\}_{k \in \mathbb{N}} \subset \mathbb{H}$ such that $f_k \rightarrow f$ *a.e.*, $\|f_k\|_{\mathbb{H}} \rightarrow +\infty$ and $H(f_k) \rightarrow H(f) < +\infty$.

Nevertheless, since J_μ is strictly convex, the solution to (\mathcal{P}_μ) is unique if it exists. It is a function of $\mathbb{H} : F_{\mu, V}$. The power density can be obtained by setting $\alpha = \|F_{\mu, V}\|_{L^1(E, \mathbb{R}, \sigma)}$ and $m = \frac{F_{\mu, V}}{\alpha} d\sigma$.

Therefore, we do not minimize the negentropy of definition 2 since $F_{\mu, V}$ is not a probability density. Rigorously the negentropy of the solution is $H(\frac{F_{\mu, V}}{\alpha})$. Moreover, the cost functional does not verify equation (A.9) on the whole set K because ϕ is not derivable at 0. So, we have to modify this problem, taking a smaller feasible set. We study the modified problem in next section to determine approximate first order optimality conditions.

A.4.1 Relaxation of the feasible set

We just mentioned that (\mathcal{P}_μ) cannot be directly solved because J_μ is not coercive and (A.9) is not satisfied. So we choose a smaller feasible set to ensure (A.9). To deal with the lack of coercivity, we bound this new domain. For $0 < \varepsilon < T < +\infty$, we set

$$K_{\varepsilon,T} = \{f \in \mathbb{H} \mid \varepsilon \leq f \text{ } \sigma.a.e \text{ and } \|f\|_{\mathbb{H}} \leq T\}. \quad (\text{A.12})$$

It is a closed, convex subset of \mathbb{H} . The “relaxed” problem reads

$$(\mathcal{P}_\mu^{\varepsilon,T}) \begin{cases} \min J_\mu(F) \\ F \in K_{\varepsilon,T} \end{cases}$$

Theorem 1 *Problem $(\mathcal{P}_\mu^{\varepsilon,T})$ has a unique solution $F_\mu^{\varepsilon,T} \in K_{\varepsilon,T}$. A necessary and sufficient condition of optimality is*

$$\forall f \in K_{\varepsilon,T}, \quad \langle DJ_\mu(F_\mu^{\varepsilon,T}), f - F_\mu^{\varepsilon,T} \rangle_{\mathbb{H}} \geq 0 \quad (\text{A.13})$$

where

$$DJ_\mu(F_\mu^{\varepsilon,T}) = -2\psi^*[V - \psi[F_\mu^{\varepsilon,T}]] + \mu(1 + \ln F_\mu^{\varepsilon,T})$$

Proof. The existence and uniqueness of the solution is standard, see [11] (Cor.III.20 p.46). We call it $F_\mu^{\varepsilon,T}$. Then

$$\forall g \in K_{\varepsilon,T}, \forall \lambda \in [0, 1], \quad J_\mu(F_\mu^{\varepsilon,T} + \lambda(g - F_\mu^{\varepsilon,T})) - J_\mu(F_\mu^{\varepsilon,T}) \geq 0.$$

With Lemmas 1 and 2 this is equivalent to

$$\forall g \in K_{\varepsilon,T}, \quad \langle DJ_\mu(F_\mu^{\varepsilon,T}), g - F_\mu^{\varepsilon,T} \rangle_{\mathbb{H}} \geq 0.$$

□ With the optimality condition (A.13), we may now construct the solution to (\mathcal{P}_μ) .

Lemma 3 *If there exists $F_\mu^{\varepsilon,T} \in K_{\varepsilon,T}$ such that*

$$F_\mu^{\varepsilon,T} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_\mu^{\varepsilon,T}]]\right), \quad (\text{A.14})$$

i.e. $F_\mu^{\varepsilon,T}$ is a fixed point of a functional

$$\begin{cases} \Gamma_\mu : \mathbb{H} \longrightarrow \mathbb{H} \\ F \longmapsto \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F]]\right) \end{cases},$$

then for all $0 < \varepsilon' \leq \varepsilon$ and $T' \geq T$, $F_\mu^{\varepsilon',T'}$ is the unique solution of $(\mathcal{P}_\mu^{\varepsilon',T'})$. Furthermore $F_\mu^{\varepsilon,T}$ is the unique solution of problem (\mathcal{P}_μ) and we note it $F_{\mu,V}$.

Proof. Let $F_\mu^{\varepsilon,T} \in K_{\varepsilon,T}$, such that

$$F_\mu^{\varepsilon,T} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_\mu^{\varepsilon,T}]]\right) \quad \sigma.a.e.$$

We get

$$\mu \ln(F_\mu^{\varepsilon,T}) = -\mu + 2\psi^*[V - \psi[F_\mu^{\varepsilon,T}]] \quad \sigma.a.e.$$

So

$$\forall f \in K_\varepsilon, \langle -2\psi^*[V - \psi[F_\mu^{\varepsilon,T}]] + \mu(1 + \ln(F_\mu^{\varepsilon,T})), f - F_\mu^{\varepsilon,T} \rangle_{\mathbb{H}} = 0,$$

and we see that $F_\mu^{\varepsilon,T}$ verifies (A.13). Therefore $F_\mu^{\varepsilon,T}$ is the solution of $(\mathcal{P}_\mu^{\varepsilon,T})$. As $K_{\varepsilon,T} \subset K_{\varepsilon',T'}$ for all $0 < \varepsilon' \leq \varepsilon$ and $T' \geq T$, we conclude that $F_\mu^{\varepsilon,T}$ is the solution of problem $(\mathcal{P}_\mu^{\varepsilon',T'})$. It is also the solution of problem (\mathcal{P}_μ) : suppose that $F' \in K$ exists, such that $J_\mu(F') \leq J_\mu(F_\mu^{\varepsilon,T})$ then the function

$$\tilde{F} = \frac{F_\mu^{\varepsilon,T} + F'}{2}$$

verifies $\tilde{F} \geq \frac{\varepsilon}{2} = \tilde{\varepsilon}$ and $\|\tilde{F}\|_{\mathbb{H}} \leq \frac{1}{2}(\|F_\mu^{\varepsilon,T}\|_{\mathbb{H}} + \|F'\|_{\mathbb{H}}) = \tilde{T}$, so $\tilde{F} \in K_{\tilde{\varepsilon},\tilde{T}}$. Since $F_\mu^{\varepsilon,T}$ is the solution of $(\mathcal{P}_\mu^{\tilde{\varepsilon},\tilde{T}})$ we have $J(F_\mu^{\varepsilon,T}) \leq J(F')$ and we deduce that $F_\mu^{\varepsilon,T}$ is a solution of (\mathcal{P}_μ) . Moreover, it is the unique solution of (\mathcal{P}_μ) since J_μ is strictly convex. \square Lemma 3 also shows that if the functional Γ_μ has a fixed point, it is unique.

We are now able to find the solution as a fixed point. In next subsection we study the existence of a sequence that converges to this fixed point. That will be the essential tool to set an infinite dimensional algorithm.

A.4.2 An infinite dimensional algorithm

Let us define the sequence $\{F_k\}_{k \in \mathbb{N}}$ of \mathbb{H} by

$$\begin{cases} F_0 \in \mathbb{H} \\ F_{k+1} = \exp(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_k]]) \end{cases} .$$

If it converges, the limit is a fixed point of the functional Γ_μ . It is also the solution to (\mathcal{P}_μ) .

Lemma 4 *The functional Γ_μ is continuous from \mathbb{H} to $L^\infty(E, \mathbb{R}, \sigma)$. Furthermore we have the inequality*

$$\|\Gamma_\mu(F)\|_{L^\infty(E, \mathbb{R}, \sigma)} \leq \exp\left(-1 + \frac{2}{\mu}C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi\|F\|_{\mathbb{H}}]\right). \quad (\text{A.15})$$

Proof. We remark that $F \mapsto \exp(F)$ is continuous from $L^\infty(E, \mathbb{R}, \sigma)$ to itself and ψ^* is continuous from \mathbb{C}^n to $L^\infty(E, \mathbb{R}, \sigma)$.

Inequalities are obtained with the continuity of operators ψ and ψ^* : there exist two constants C_ψ and C_{ψ^*} such that

$$\begin{aligned} \|\psi(F)\|_{\mathbb{C}^n} &\leq C_\psi\|F\|_{\mathbb{H}} \\ \|\psi^*(x)\|_{L^\infty(E, \mathbb{R}, \sigma)} &\leq C_{\psi^*}\|x\|_{\mathbb{C}^n} \end{aligned}$$

Since exponential function is non decreasing one obtains the results by injecting the last inequalities in the expression of Γ_μ . \square

Now we show the convergence of the sequence $\{F_k\}_{k \in \mathbb{N}}$. The following lemma gives a condition on the regularization parameter μ which implies that the sequence $\{F_k\}_{k \in \mathbb{N}}$ stays in a ball of fixed radius.

Lemma 5 *Let F be such that $\|F\|_{\mathbb{H}} \leq R$; if*

$$\mu \geq \frac{2C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R]}{1 + \ln(R)} \quad (\text{A.16})$$

then $\|\Gamma_\mu(F)\|_{\mathbb{H}} \leq R$ and $\|\Gamma_\mu(F)\|_{L^\infty(E, \mathbb{R}, \sigma)} \leq R$

Proof. We have

$$\begin{aligned} \mu &\geq \frac{2C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R]}{1 + \ln(R)} \\ \mu[1 + \ln(R)] &\geq 2C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R] \end{aligned}$$

so

$$R \geq \exp\left(-1 + \frac{2}{\mu}C_{\psi^*}[\|V\|_{\mathbb{C}^n} + C_\psi R]\right)$$

and we deduce the two inequalities. \square

We will use a fixed point criterion : if the functional Γ_μ is a contraction, then the sequence is converging. In the following lemma, we give a condition on μ for the sequence to converge.

Lemma 6 *Γ_μ is Fréchet-differentiable on \mathbb{H} and its derivative is*

$$d\Gamma_\mu(F).h = -\frac{2}{\mu}\Gamma_\mu(F)\psi^* \circ \psi[h]. \quad (\text{A.17})$$

Let $\|F_0\|_{\mathbb{H}} \leq R$, if μ verifies (A.16) and if

$$\mu > 2R^2C_{\psi^*}C_\psi \quad (\text{A.18})$$

then the sequence $(F_k)_{k \in \mathbb{N}}$ converges in \mathbb{H} and in $L^\infty(E, \mathbb{R}, \sigma)$ to the unique fixed point $F_{\mu, V}$ of Γ_μ .

Proof. Functional Γ_μ is differentiable since $F \mapsto \exp(F)$ is differentiable from $L^\infty(E, \mathbb{R}, \sigma)$ to itself, and $F \mapsto \psi^*[\psi[F]]$ is linear continuous from \mathbb{H} to $L^\infty(E, \mathbb{R}, \sigma)$. By Lemma 5, we deduce that for all $k \in \mathbb{N}$, $0 \leq F_k \leq R$. Furthermore, equation (A.18) leads to

$$\sup_{\{F \in \mathbb{H} | 0 \leq F \leq R\}} \|d\Gamma_\mu(F)\|_{\mathcal{L}(L^\infty, L^\infty)} < 1$$

We conclude using the Banach fixed point theorem on the complete set $\{F \in \mathbb{H} | 0 \leq F \leq R\}$ with the distance induced by L^∞ . The sequence converges to the unique fixed point $F_{\mu, V} \in \{F \in \mathbb{H} | 0 \leq F \leq R\}$ of Γ_μ strongly in $L^\infty(E, \mathbb{R}, \sigma)$ and in \mathbb{H} (by compactness). \square

We may summarize in :

Theorem 2 *If μ verifies (A.16) and (A.18) (i.e. μ large enough), then problem (\mathcal{P}_μ) has a unique solution $F_{\mu,V}$ limit of the sequence $\{F_k\}_{k \in \mathbb{N}}$ defined by*

$$\begin{cases} F_0 \in \{F \in \mathbb{H} \mid 0 \leq F \leq R\} \\ F_{k+1} = \exp\left(-1 + \frac{2}{\mu}\psi^*[V - \psi[F_k]]\right) = \Gamma_\mu(F_k) \end{cases}$$

The convergence stands in L^∞ and in \mathbb{H} .

Shortly speaking, we have an infinite dimensional algorithm that converges to the solution of the maximum entropy regularization problem for a fixed parameter μ great enough. Theorem 2 shows that the solution $F_{\mu,V}$ of (\mathcal{P}_μ) is obtained as the limit of the sequence $\{F_k\}_{k \in \mathbb{N}}$ and belongs to $L^\infty(E, \mathbb{R}, \sigma)$.

However this algorithm is not useful from the numerical point of view. Indeed, it is an infinite dimensional one and an inappropriate discretization process may lead to slow computations. Nevertheless, we are able to derive a finite dimensional algorithm from the optimality condition (A.13). This is the aim of next subsection.

A.4.3 A finite dimensional algorithm

Lemma 3 suggests to look for the solution $F_{\mu,V}$ of problem (\mathcal{P}_μ) as $F_{\mu,V} = G_{\mu,V}$ where

$$G_{\mu,V} = \exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda]\right) \quad (\text{A.19})$$

where $\lambda \in \mathbb{C}^n$ has to be determined. Next Lemma gives a sufficient optimality condition on λ to solve (\mathcal{P}_μ) . In addition, we have an analytic expression for the solution.

Lemma 7 *Let $\lambda \in \mathbb{C}^n$ such that*

$$\lambda = V - \int_E q(\sigma) \exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda](\sigma)\right) d\sigma \quad (\text{A.20})$$

then the function $G_{\mu,V} \in \mathbb{H}$, defined by (A.19), is the unique solution of (\mathcal{P}_μ)

Proof. By definition of $G_{\mu,V}$ and with the assumption $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$, there exist $\varepsilon > 0$ and $T > \varepsilon$ such that $G_{\mu,V} \in K_{\varepsilon,T}$. Writing the expression of $F_{\mu,V}$ in (A.13) we can see that it is verified. So $F_{\mu,V}$ is the unique solution of problem $(\mathcal{P}_\mu^{\varepsilon,T})$. With Lemma 3 and the strict convexity of J_μ it follows that it is the unique solution of (\mathcal{P}_μ) . \square

Therefore we only need to find the value of $\lambda \in \mathbb{C}^n$ to determine $G_{\mu,V}$. So the problem turns to be a finite dimensional one. Let us define the sequence $\{\lambda_k\}_{k \in \mathbb{N}}$ of \mathbb{C}^n as

$$\begin{cases} \lambda_0 \in \mathbb{C}^n \\ \lambda_{k+1} = \gamma_\mu(\lambda_k) \stackrel{\text{def}}{=} V - \int_E q(\sigma) \exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda_k](\sigma)\right) d\sigma \end{cases} \quad (\text{A.21})$$

The function $\gamma_\mu : \mathbb{C}^n \rightarrow \mathbb{C}^n$ defined in (A.21) is differentiable and its derivative is

$$d\gamma_\mu(\lambda).h = -\frac{2}{\mu} \int_E q(\sigma) \bar{q}^t(\sigma) \exp\left(-1 + \frac{2}{\mu}\psi^*[\lambda](\sigma)\right) .h d\sigma; \quad (\text{A.22})$$

thus

$$\|d\gamma_\mu(\lambda_k).h\|_{\mathbb{C}^n} \leq \frac{2}{\mu} \rho_\mu^{(k)} \|h\|_{\mathbb{C}^n} \quad (\text{A.23})$$

where $\rho_\mu^{(k)}$ is the spectral radius of matrix $M_\mu^{(k)}$ of dimension (n, n) , such that for all $1 \leq i, j \leq n$ (we use euclidean norm on \mathbb{C}^n)

$$M_{\mu,ij}^{(k)} = \int_E q_i(\sigma) \bar{q}_j(\sigma) \exp\left(-1 + \frac{2}{\mu} \psi^*[\lambda_k](\sigma)\right) d\sigma$$

Using the Frobenius norm of matrix $M_\mu^{(k)}$ we have the inequality

$$\rho_\mu^{(k)} \leq \left\| M_\mu^{(k)} \right\|_{F\tau} \leq \exp\left(-1 + \frac{2}{\mu} C_{\psi^*} \|\lambda_k\|_{\mathbb{C}^n}\right) \sqrt{\sum_{1 \leq i, j \leq n} \left| \int_E q_i(\sigma) \bar{q}_j(\sigma) d\sigma \right|^2} \quad (\text{A.24})$$

The sequence $\{\lambda_k\}_{k \in \mathbb{N}}$ converges to the fixed point of γ_μ only if $\rho_\mu^{(k)}$ is small enough for any k . So we cannot use it to calculate λ . However, we are able to construct another sequence converging to λ noting that for all $\tau > 0$, $\lambda = \gamma_\mu(\lambda)$ is equivalent to $\lambda = \lambda - \tau[\lambda - \gamma_\mu(\lambda)]$. So we can obtain λ as limit of the sequence $\{l_k\}_{k \in \mathbb{N}}$

$$\begin{cases} l_0 \in \mathbb{C}^n \\ l_{k+1} = l_k - \tau[l_k - \gamma_\mu(l_k)] \end{cases} \quad (\text{A.25})$$

This sequence will be used to determinate the solution practically. If τ is chosen small enough and μ is great enough it converges.

Lemma 8 *Assume that for every $k \in \mathbb{N}$ the spectral radius of $M_\mu^{(k)}$ is less than $m_\mu \in \mathbb{R}^{+*}$; then the sequence l_k converges if $0 < \tau < \frac{1}{1 + \frac{2}{\mu} m_\mu}$*

Proof. The matrix $M_\mu^{(k)}$ is hermitian of nonnegative type since for all $\sigma \in E$, $q(\sigma) \bar{q}^t(\sigma)$ is nonnegative hermitian and $\exp\left(-1 + \frac{2}{\mu} \psi^*[\lambda_k](\sigma)\right) > 0$. Let $h_\tau : \mathbb{C}^n \rightarrow \mathbb{C}^n$ be the function $\lambda \mapsto \lambda - \tau[\lambda - \gamma_\mu(\lambda)]$. The derivative of h_τ is

$$\forall \lambda, v \in \mathbb{C}^n, \quad dh_\tau(\lambda).v = ([1 - \tau]I - \frac{2}{\mu} \tau M_\mu)v$$

Since $M_\mu^{(k)}$ is hermitian, there exists an orthogonal basis of eigenvectors and we call $B^{(k)}$ the transition matrix. So $B_*^{(k)}([1 - \tau]I - \frac{2}{\mu} \tau M_\mu^{(k)})B^{(k)} = [1 - \tau]I - \frac{2}{\mu} \tau \Delta_\mu^{(k)}$, where $\frac{2}{\mu} \tau \Delta_\mu^{(k)}$ is a diagonal matrix with positive elements. Thus, the spectral radius of $[1 - \tau]I - \frac{2}{\mu} \tau M_\mu^{(k)}$ is strictly less than $1 - \tau$ since the spectral radius of $M_\mu^{(k)}$ is less than to m_μ and $0 < \tau < \frac{1}{1 + \frac{2}{\mu} m_\mu}$.

We deduce

$$\begin{aligned} \|l_{k+1} - l_k\|_{\mathbb{C}^n} &\leq \left\| ([1 - \tau]I - \tau M_\mu^{(k)}) \right\| \|l_k - l_{k-1}\|_{\mathbb{C}^n} \\ &\leq (1 - \tau) \|l_k - l_{k-1}\|_{\mathbb{C}^n} \end{aligned}$$

because $\|A\|^2 = \rho(A^*A)$. So the sequence converges. \square

We have a proof of the convergency of sequence $\{l_k\}_{k \in \mathbb{N}}$ if the spectral radius of the matrix $M^{(k)\mu}$ is uniformly bounded with respect to k . In next Lemma, we give an estimate of the spectral radius of $M^{(k)\mu}$ with the Frobenius norm.

Lemma 9 *Let $\|l_0\|_{\mathbb{C}^n} \leq R$; if*

$$\mu \geq 2C_{\psi^*}R \left(\log \left(\frac{R - \|V\|_{\mathbb{C}^n}}{C'_{\psi}} \right) + 1 \right)^{-1} \quad (\text{A.26})$$

where $C'_{\psi} > 0$ verifies

$$\|\psi(F)\|_{\mathbb{C}^n} \leq C'_{\psi} \|F\|_{L^\infty(E, \mathbb{R}, \sigma)},$$

then the whole sequence is bounded by R , and the spectral radius $\rho_\mu^{(k)}$ of $M_\mu^{(k)}$ satisfies

$$\rho_\mu^{(k)} \leq \exp \left(-1 + \frac{2}{\mu} C_{\psi^*} R \right) \sqrt{\sum_{1 \leq i, j \leq n} \left| \int_E q_i(\sigma) \bar{q}_j(\sigma) d\sigma \right|^2} \quad (\text{A.27})$$

Proof. We prove the result by induction. Let us assume there exists $k \in \mathbb{N}^*$ such that $\forall j \leq k$, $\|l_j\|_{\mathbb{C}^n} \leq R$, then

$$\|l_{k+1}\|_{\mathbb{C}^n} \leq (1 - \tau)R + \tau \left(\|V\|_{\mathbb{C}^n} + C'_{\psi} \exp \left(-1 + \frac{2}{\mu} C_{\psi^*} R \right) \right)$$

and condition (A.26) implies

$$\|V\|_{\mathbb{C}^n} + C'_{\psi} \exp \left(-1 + \frac{2}{\mu} C_{\psi^*} R \right) \leq R.$$

For all $k \in \mathbb{N}$ we have $\|l_k\| \leq R$ and (A.27) is a direct consequence of (A.24). \square

Proposition 1 *If (μ, R) satisfies condition (A.26) then γ_μ has an unique fixed point λ in the closed ball $B(0, R)$. Moreover if $\|l_0\|_{\mathbb{C}^n} \leq R$ and τ is small enough, then λ is the limit of the sequence $\{l_k\}_{k \in \mathbb{N}}$.*

We now give a more precise description of the algorithm defined by (A.25) :

Algorithm A1

1. Initialization

Given $V \in \mathbb{C}^n$, choose $l_0 \in \mathbb{C}^n, \mu > 0, \epsilon > 0, \tau \in]0, 1]$.

2. Iteration k

(a) Compute

$$\gamma_\mu(l_{k-1}) = V - \int_E q(\sigma) \exp \left(-1 + \frac{2}{\mu} \psi^*[l_{k-1}](\sigma) \right) d\sigma$$

(b) Compute $l_k = l_{k-1} - \tau[l_{k-1} - \gamma_\mu(l_{k-1})]$

3. Stopping criterion

If $|l_k - l_{k-1}| < \epsilon$, then STOP, else $k := k + 1$ and go to 2.

The algorithm converges if the regularization parameter is great enough. The main advantage of this method is that it determines a vector of \mathbb{C}^n which is the fixed point of a functional. Moreover we have an analytic expression for this solution. The convergence of the algorithm is linear since we have shown that $\|l_{k+1} - l_k\|_{\mathbb{C}^n} \leq (1 - \tau) \|l_k - l_{k-1}\|_{\mathbb{C}^n}$ in Lemma 8. The number τ has to be chosen as great as possible for a faster convergence.

Now we perform a sensitivity analysis of the optimal value function with respect to μ and V . Let $V \in \mathbb{C}^n$ be fixed. We suppose that for the data V the sequence (A.21) converges for all $\mu \geq \mu_0 > 0$. We define the function \mathfrak{C}_1 by

$$\begin{cases} \mathfrak{C}_1 : & [\mu_0, +\infty[\longrightarrow \mathbb{R} \\ & \mu \longmapsto J_\mu(F_{\mu,V}) + \mu e^{-1} \end{cases} \quad (\text{A.28})$$

Similarly for any $\lambda > 0$ fixed, we suppose that the sequence (A.21) converges for all $V \in B(0, R)$ where $R > 0$, and we define the function \mathfrak{C}_2 by

$$\begin{cases} \mathfrak{C}_2 : & B(0, R) \subset \mathbb{C}^n \longrightarrow \mathbb{R} \\ & V \longmapsto J_\mu(F_{\mu,V}) + \mu e^{-1} \end{cases} \quad (\text{A.29})$$

Proposition 2 *Let \mathfrak{C}_1 and \mathfrak{C}_2 be the functions defined by (A.28) and (A.29).*

1. \mathfrak{C}_1 is continuous and increasing.
2. \mathfrak{C}_2 is continuous and verifies

$$\mathfrak{C}_2 \leq \|V\|_{\mathbb{C}^n}^2 + \frac{\mu}{e} \quad (\text{A.30})$$

Proof. Let be $\mu_1 > \mu_2 \geq \mu_0$. For all $F \in K$, $J_{\mu_1}(F) > J_{\mu_2}(F)$ since $(H(F) + e^{-1}) \geq 0$, hence \mathfrak{C}_1 is increasing. To prove continuity we suppose there exists a sequence $\{\mu_k\}_{k \in \mathbb{N}}$ and a $\delta > 0$ such that $\mu_k \rightarrow \mu$ and for all k , $|\mathfrak{C}_1(\mu_k) - \mathfrak{C}_1(\mu)| > \delta$. Let $\varepsilon > 0$ be small enough, so $\exists N \in \mathbb{N}$ such that $|\mu - \mu_N| < \frac{\delta(\mu - \varepsilon)}{2\mathfrak{C}_1(\mu)}$ and $|\mu - \mu_k| < \varepsilon$. Assume that $\mu_N < \mu$ then

$$J_\mu(F_{\mu_N,V}) - \mathfrak{C}_1(\mu_N) = (\mu - \mu_N)H(F_{\mu_N,V})$$

so $|J_\mu(F_{\mu_N,V}) - \mathfrak{C}_1(\mu_N)| < \delta/2$ since $H(F_{\mu_N,V}) \leq \frac{\mathfrak{C}_1(\mu)}{\mu_N}$. Hence $J_\mu(F_{\mu_N,V}) < \mathfrak{C}_1(\mu)$. This contradiction proves the result. The case $\mu_N \geq \mu$ can be shown similarly.

We can show the continuity of \mathfrak{C}_2 by the same way. The inequality (A.30) is obtained by taking $F = 0$. \square

We just proved the fixed point existence if μ is large enough; in fact, it is true for any $\mu > 0$: we use a scaling method to prove it :

Proposition 3 *For every $\mu > 0$, the solution F_μ of problem (\mathcal{P}_μ) exists and verifies (A.14). Moreover, F_μ can be computed with the sequence $\{\ell'_k\}_{k \in \mathbb{N}}$ defined by*

$$\begin{cases} \ell'_o \in \mathbb{C}^n \\ \ell'_{k+1} = (1 - \frac{\tau_o}{\alpha})\ell'_k + \tau_o \gamma_{\mu_o}(\ell'_k) \end{cases}$$

for some $\mu_o > \mu > 0$, $\tau_o \in]0, 1[$ and $\alpha = \frac{\mu_o}{\mu}$.

Proof. With proposition 1, we know that there exist $\ell_o \in \mathbb{C}^n, \mu_o > 0$ and $0 < \tau_o < 1$ such that the sequence $\{\ell_k\}_{k \in \mathbb{N}}$ defined by (A.25) converges. Then the solution F_μ of problem (\mathcal{P}_μ) exists for every $\mu \geq \mu_o$.

Assume now that $\mu < \mu_o$ and set $\alpha := \frac{\mu_o}{\mu} > 1$. As shown in the proof of Lemma 8, the eigenvalues of the gradient of the function $\ell \mapsto (1 - \tau_o)\ell + \tau_o\gamma_{\mu_o}(\ell)$ are nonnegative, and the spectral radius is bounded by $1 - \tau_o$. As $-d\gamma_{\mu_o}$ is nonnegative and $\alpha > 1$, we deduce as before that the sequence $\{\ell'_k\}$ defined by

$$\begin{cases} \ell'_o \in \mathbb{C}^n \\ \ell'_{k+1} = (1 - \frac{\tau_o}{\alpha})\ell'_k + \tau_o\gamma_{\mu_o}(\ell'_k) \end{cases}$$

is converging to some ℓ' that verifies

$$\begin{aligned} \ell' &= (1 - \frac{\tau_o}{\alpha})\ell' + \tau_o\gamma_{\mu_o}(\ell') , \\ \frac{\tau_o}{\alpha}\ell' &= \tau_o\gamma_{\mu_o}(\ell') ; \end{aligned}$$

Setting $\lambda = \frac{\ell'}{\alpha}$ yields

$$\lambda = \gamma_{\mu_o}(\alpha\lambda) = V - \psi[\exp(-1 + \frac{2}{\mu_o}\psi^*(\alpha\lambda))]$$

$$\lambda = V - \psi[\exp(-1 + \frac{2\alpha}{\mu_o}\psi^*(\lambda))] = V - \psi[\exp(-1 + 2\mu\psi^*(\lambda))] = \gamma_\mu(\lambda) .$$

We conclude with Lemma 7, that (\mathcal{P}_μ) has a (unique) solution. \square

In summary, we have a finite dimensional algorithm that converges to the solution of maximum entropy regularization for all $\mu > 0$. However, last proposition show that the number of iteration increase when μ decrease since the speed of convergency is $(1 - \frac{\tau_o\mu}{\mu_o})$.

We know the analytic expression of the solution and the problem is now a finite dimensional one. However, the entropy involved in the functional we have minimized is not the entropy in the sense of definition 2 because the solution is not a probability density. The algorithm we present in the next section allows to find a solution which maximize the true entropy. It is a variant of the previous one.

A.5 Computing the probability density

We look for a solution that maximizes entropy in the sense of definition 2. This leads to the problem

$$(\tilde{\mathcal{P}}_\mu) \left\{ \begin{array}{l} \min \|V - \alpha\psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F) \\ (\alpha, F) \in \mathbb{R}^+ \times \left\{ f \in K \mid \|f\|_{L^1(E, \mathbb{R}, \sigma)} = 1 \right\} \end{array} \right\}$$

Here, $\alpha \in \mathbb{R}^+$ is the total power and F is a probability density because of the constraint $F \in \left\{ f \in K \mid \|f\|_{L^1(E, \mathbb{R}, \sigma)} = 1 \right\}$. So, if the solution exists, it minimizes negentropy as defined in definition 2.

We cannot solve $(\tilde{\mathcal{P}}_\mu)$ in the same way as problem (\mathcal{P}_μ) : indeed the cost functional is no more convex because of the bilinear term $\alpha \psi[F]$. On the other hand, the feasible domain is convex since $\forall f, g \in \left\{ f \in K \mid \|f\|_{L^1(E, \mathbb{R}, \sigma)} = 1 \right\}$, $\forall \lambda \in [0, 1]$, we have $\lambda f + (1 - \lambda)g \in K$, and

$$\|\lambda f + (1 - \lambda)g\|_{L^1(E, \mathbb{R}, \sigma)} = \lambda \|f\|_{L^1(E, \mathbb{R}, \sigma)} + (1 - \lambda) \|g\|_{L^1(E, \mathbb{R}, \sigma)} = 1$$

by positivity of f, g . Anyway, if F is solution of (\mathcal{P}_μ) , we can view $\|F\|_{L^1(E, \mathbb{R}, \sigma)}$ as an approximation of the power. So we use the previous section results on (\mathcal{P}_μ) . Let us define the sequence $\{l_k\}_{k \in \mathbb{N}}$ by

$$\begin{cases} l_0 \in \mathbb{C}^n \\ \delta_{k, \mu} = \frac{\operatorname{Re}(\langle V, g_\mu(l_k) \rangle_{\mathbb{C}^n})}{\langle g_\mu(l_k), g_\mu(l_k) \rangle_{\mathbb{C}^n}} \\ l_{k+1} = V - \delta_{k, \mu} g_\mu(l_k) \end{cases} \quad (\text{A.31})$$

where $g_\mu : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is defined by

$$g_\mu(l) = V - \gamma_\mu(l) = \int_E q(\sigma) \exp\left(-1 + \frac{2}{\mu} \psi^*[l](\sigma)\right) d\sigma. \quad (\text{A.32})$$

We assume that

$$\forall l \in \mathbb{C}^n, \quad \operatorname{Re}(\langle V, g_\mu(l) \rangle_{\mathbb{C}^n}) \geq 0;$$

since $\delta_{k, \mu}$ is an estimation of the power, it has to be positive. If $\delta_{k, \mu} \rightarrow \delta^*$ and $l_k \rightarrow l^*$ then

$$l^* = V - \delta^* g_\mu(l^*) \quad (\text{A.33})$$

Last equation is quite similar to the necessary and sufficient condition of optimality (A.13) of problem (\mathcal{P}_μ) .

Proposition 4 *Let be $l^* = V - \delta^* g_\mu(l^*)$ with $\delta^* > 0$; we define*

$$F_{l^*} = \exp\left(-1 + \frac{2}{\mu} \psi^*[l^*]\right) \text{ and } f_{l^*} = \frac{F_{l^*}}{\|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)}}$$

Then f_{l^} is a probability density and it is the unique solution of problem*

$$\begin{cases} \min \left\| \frac{V}{\delta^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)}} - \psi(F) \right\|_{\mathbb{C}^n}^2 + \frac{\mu}{\delta^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)}} \int_E \phi \circ F(\sigma) d\sigma \\ F \in \left\{ f \in \mathbb{H} \mid f \geq 0 \text{ } \sigma.a.e., \|f\|_{L^1(E, \mathbb{R}, \sigma)} = 1 \right\} \end{cases} \quad (\text{A.34})$$

Moreover, f_{l^} verifies*

$$\forall \alpha \in \mathbb{R}, \quad \left\| V - \delta^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)} \psi[f_{l^*}] \right\|_{\mathbb{C}^n}^2 \leq \|V - \alpha \psi[f_{l^*}]\|_{\mathbb{C}^n}^2 \quad (\text{A.35})$$

Proof. We have

$$\begin{aligned} l^* = V - \delta^* g_\mu(l^*) &\Rightarrow -2\psi^*[V - \delta^* g_\mu(l^*) - l^*] = 0 \\ &\Rightarrow -2\psi^*[V - \delta^* g_\mu(l^*)] + \mu \left(\frac{2}{\mu} \psi^*[l^*] \right) = 0 \\ &\Rightarrow -2\psi^*[V - \delta^* \psi[F_{l^*}]] + \mu(1 + \ln F_{l^*}) = 0 \end{aligned}$$

Dividing the last equation by δ^* gives a sufficient optimality condition for problem (A.34). So F_{l^*} is the unique solution to this problem. As $\{f \in \mathbb{H} \mid f \geq 0 \text{ } \sigma.a.e., \|f\|_{L^1} = 1\} \subset K$, we get the conclusion.

To show (A.35) we suppose that $\psi[f_{l^*}] \neq 0$ (otherwise the result is obvious). Let us define

$$\begin{cases} \theta : \mathbb{R} \longrightarrow \mathbb{R} \\ \alpha \longmapsto \left\| V - \alpha \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)} \psi[f_{l^*}] \right\|_{\mathbb{C}^n}^2 \end{cases} .$$

It is a continuous and strictly convex function. The unique minimizer of θ (denoted α^*) verifies

$$\theta'(\alpha) = -2 \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)} \operatorname{Re} \left(\left\langle V - \alpha^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)} \psi[f_{l^*}], \psi[f_{l^*}] \right\rangle_{\mathbb{C}^n} \right) = 0$$

that is

$$\alpha^* = \frac{\operatorname{Re} \left(\langle V, \psi[f_{l^*}] \rangle_{\mathbb{C}^n} \right)}{\|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)} \|\psi[f_{l^*}]\|_{\mathbb{C}^n}^2} = \delta^*$$

□

The above proposition shows that if the sequence given by (A.31) converges, then its limit provides a power density $(\delta^*, f_{l^*} d\sigma)$ which minimizes the error with the data V . Moreover f_{l^*} has a minimal negentropy in the probabilistic sense. Now, we show that, the obtained solution is the unique solution of problem $(\tilde{\mathcal{P}}_{\delta^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)}})$. Let us call

$$\mathfrak{J}_\mu(\alpha, F) = \|V - \alpha \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F),$$

we have $\mathfrak{J}_\mu(1, F) = J_\mu(F)$.

The previous result proves that if the sequence defined by (A.31) converges then its limit is a solution of problem $(\tilde{\mathcal{P}}_{\mu \delta^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)}})$. We have a result of convergence :

Proposition 5 *Assume there exists $c > 0$ such that*

$$\forall \mu \geq 0, \forall l \in \mathbb{C}^n, \quad \|g_\mu(l)\|_{\mathbb{C}^n} \geq c ; \quad (\text{A.36})$$

then there exists a μ_0 such that sequence defined by (A.31) converges for every $\mu \geq \mu_0$ large enough.

Proof. We have to compute the derivative of the function defining the sequence (A.31) :

$$\varphi_\mu : l \mapsto V - \frac{\operatorname{Re} \left(\langle V, g_\mu(l) \rangle_{\mathbb{C}^n} \right)}{\langle g_\mu(l), g_\mu(l) \rangle_{\mathbb{C}^n}} g_\mu(l) .$$

We get

$$\begin{aligned} \varphi'_\mu(l) &= -\frac{\operatorname{Re}(\langle V, g_\mu(l) \rangle_{\mathbb{C}^n})}{\|g_\mu(l)\|_{\mathbb{C}^n}^2} \nabla g_\mu(l) \\ &\quad - g_\mu(l) \left[\frac{\operatorname{Re}(\langle V, \nabla g_\mu(l) \rangle_{\mathbb{C}^n})}{\|g_\mu(l)\|_{\mathbb{C}^n}^2} - 2 \frac{\operatorname{Re}(\langle V, \nabla g_\mu(l) \rangle_{\mathbb{C}^n})}{\|g_\mu(l)\|_{\mathbb{C}^n}^2} \frac{\operatorname{Re}(\langle g_\mu(l), \nabla g_\mu(l) \rangle_{\mathbb{C}^n})}{\|g_\mu(l)\|_{\mathbb{C}^n}^2} \right] \end{aligned}$$

A short computation gives

$$\sup_{l \in \mathbb{C}^n} \|\varphi'_\mu(l)\| \leq 2\|V\| \kappa_\mu (1 + 2\kappa_\mu),$$

where

$$\kappa_\mu \stackrel{\text{def}}{=} \sup_{l \in \mathbb{C}^n} \frac{\|\nabla g_\mu(l)\|}{\|g_\mu(l)\|}.$$

Thanks to assumption (A.36) and equation (A.23), there exists a constant $c' > 0$ such that $\kappa_\mu < \frac{c'}{\mu}$. It is clear there exists $\mu_0 > 0$ such that φ_μ is contractive for every $\mu \geq \mu_0$. \square

Remark 1 Assumption (A.36) seems to be surprising, but in fact it is primordial. This assumption can easily be derived from the physical model : $g_\mu(l)$ is the spectral matrix corresponding to the WDF $\exp(-1 + \frac{2}{\mu}\psi^*[l]) > 0$. Since the WDF is positive, it means that the electromagnetic wave has a non-zero energy and thus the spectral matrix cannot be zero. We can see (A.36) from the model noting that the spectral matrix trace is the constant function 1 on the sphere.

Remark 2 We do not have proved the convergency for all $\mu > 0$. Anyway, in practice $V \ll \psi$. So we have to perform a (physical) normalization process : this means that we use the operator $\varepsilon\psi$ instead of ψ where $\varepsilon > 0$ is a small number (usually $\varepsilon \simeq \frac{\|V\|}{\|\psi\|}$). Hence, we have to solve the problem $(\tilde{\mathcal{P}}_{\mu/\varepsilon^2})$ (according to the proof of Proposition 3), and the new regularization parameter $\tilde{\mu} = \frac{\mu}{\varepsilon^2}$ is large.

We now can write the algorithm more precisely and we introduce a relaxation parameter τ useful for numerical computations :

Algorithm A2

1. Initialization

Given $V \in \mathbb{C}^n$, choose $\ell_o \in \mathbb{C}^n, \mu > 0, \tau \in]0, 1[, \epsilon > 0$.

2. Iteration k

(a) Compute

$$g_\mu(\ell_{k-1}) = \int_E q(\sigma) \exp\left(-1 + \frac{2}{\mu}\psi^*[\ell_{k-1}](\sigma)\right) d\sigma.$$

(b) Compute

$$\delta_{k-1,\mu} = \frac{\operatorname{Re}(\langle V, g_\mu(\ell_{k-1}) \rangle_{\mathbb{C}^n})}{\langle g_\mu(\ell_{k-1}), g_\mu(\ell_{k-1}) \rangle_{\mathbb{C}^n}}.$$

(c) Compute $\ell_k = (1 - \tau)\ell_{k-1} + \tau[V - \delta_{k-1,\mu}g_\mu(\ell_{k-1})]$.

3. Stopping criterion

If $|\ell_k - \ell_{k-1}| < \epsilon$, then STOP, else $k := k + 1$ and go to 2.

Remarque 5 *In this section we have supposed that $R(\psi^*) \subset L^\infty(E, \mathbb{R}, \sigma)$ to established the algorithms. If this condition is not verified we can slightly modify the operator ψ to overcome this problem. More precisely if $\psi : F \mapsto \int_E q(\sigma)F(\sigma) d\sigma$ with $q \notin L^\infty(E, \mathbb{C}^n, \sigma)$, we can find a $q^\epsilon \in L^\infty(E, \mathbb{C}^n, \sigma)$ such that $\|q - q^\epsilon\|_{L^2(E, \mathbb{C}^n, \sigma)} \leq \epsilon$. The function q^ϵ allows to define a new operator $\psi_\epsilon : F \mapsto \int_E q^\epsilon(\sigma)f(\sigma) d\sigma$ which verifies $R(\psi_\epsilon) \subset L^\infty(E, \mathbb{R}, \sigma)$ and $\|\psi_\epsilon[F] - \psi[F]\|_{\mathbb{C}^n} \leq \epsilon\|F\|_{\mathbb{H}}$. This new operator can be used for the computation instead of ψ , and it verifies the desired condition.*

A.6 Numerical tests

In this section we perform some numerical tests for the resolution of (A.2) with the two algorithms previously described. For these tests we consider the propagation of an electromagnetic wave in vacuum. We give below the corresponding expression of operator ψ . We denote $\epsilon_r = \|V - \psi[F_{\mu,V}]\|_{\mathbb{C}^n}^2$ the error and α is the computed approximation of total power P .

For the numerical computation of the solution, we need to find $\tau > 0$ small enough for the algorithms to converge. The number τ must be great enough for a fast computation. To determine this parameter we have used the following rule :

- let $\tau_o > 0$, $k = 0$ and $\ell_o = 0$.
- Iteration k : τ_k and ℓ_k are known ; compute ℓ_{k+1}
- if $\|\ell_{k+1} - \ell_k\|_{\mathbb{C}^n} \geq \|\ell_k - \ell_{k-1}\|_{\mathbb{C}^n}$ then $\tau_{k+1} = \tau_k r$ where $0 < r < 1$.

In the computation we used $r = 0.5$. We can also utilize a linear search strategy for the best value of τ . We have chosen to stop the algorithms when $\|\ell_{k+1} - \ell_k\|_{\mathbb{C}^n} < \epsilon$ where $\epsilon > 0$ is chosen small enough (here $\epsilon = 10^{-12}$).

All numerical tests were performed on a Power Mac 2×G4 with the scientific software Scilab 2.6. For these examples the average time for one iteration was 28.5 ms. This is corresponding to the average number of 35 iterations per second.

A.6.1 Wave distribution function in vacuum

In the case of an electromagnetic wave propagating in vacuum, we have the relation

$$V = \int_0^\pi \int_0^{2\pi} q(\theta, \phi)F(\theta, \phi) \sin \theta d\phi d\theta \quad (\text{A.37})$$

where V is the data vector, F is the WDF of the electromagnetic wave and q is the integrating kernel of vacuum. Writing the components of the spectral matrix as a vector provides the data

V . We see with (A.37) that we integrate over the unit sphere : $\theta \in [0, \pi]$ denotes the polar angle and $\phi \in [0, 2\pi]$ the azimuthal angle. This is quite clear because WDF is the directional distribution of the power of electromagnetic wave. The expression of q is analytically known [27].

According to (A.37) we put $E = [0, \pi] \times [0, 2\pi]$, $d\sigma = \frac{\sin\theta}{2\pi} d\theta d\phi$ and we define the operator ψ by

$$\begin{cases} \psi : \mathbb{H} & \longrightarrow \mathbb{C}^n \\ F & \longmapsto \int_0^\pi \int_0^{2\pi} q(\theta, \phi) F(\theta, \phi) \sin\theta d\phi d\theta \end{cases} \quad (\text{A.38})$$

This operator is continuous on \mathbb{H} and the adjoint operator verifies $R(\psi^*) \subset L^\infty(E, \mathbb{C}^n, \sigma)$ [27]. From the expression of ψ , we see that we have to compute a double integral. For this computation we have used the Gauss-Legendre quadrature method with 20 points on the interval $[0, \pi]$ and 40 points on $[0, 2\pi]$. In the following tests, this integration method seems to be accurate enough because the results don't change dramatically if the number of Gauss points is larger.

Numerical validation of algorithms was done with simulated data. In the first tests V was obtained for a WDF that was a sum of dirac measures. The data vector of a dirac in (θ_d, ϕ_d) of power P is simulated by taking $V = P/\sin(\theta_d) q(\theta_d, \phi_d)$. The case of a dirac is important because it physically corresponds to an electromagnetic plane wave. Second tests are performed with data corresponding to an almost everywhere continuous WDF. The simulation has been made by computing the integral (A.37). For some of these examples, a noisy data V^δ was introduced with $\|V - V^\delta\|_{\mathbb{C}^n} \leq \delta$ where $\delta > 0$ is the noise level.

The choice of the regularization parameter is a difficult task in the case of noisy data. It must be chosen small enough to make a small error, and large enough to guarantee stability. A commonly used rule is the discrepancy principle of Morozov [22]. We choose μ such that $\|V - \psi[F_\mu]\|_{\mathbb{C}^n} = \delta$ i.e. we don't make a smaller error than the noise level. In the numerical tests, since we used simulated data, we decided to choose μ such that $\delta \leq \|V - \psi[F_\mu]\|_{\mathbb{C}^n} \leq 2\delta$. This choice is made *a posteriori* using the fact that the function $\mu \mapsto \|V - \psi[F_{\mu,V}]\|_{\mathbb{C}^n}$ is quite linear when μ is small (see Figure 1.). Note that the noise level δ is known from the experience.

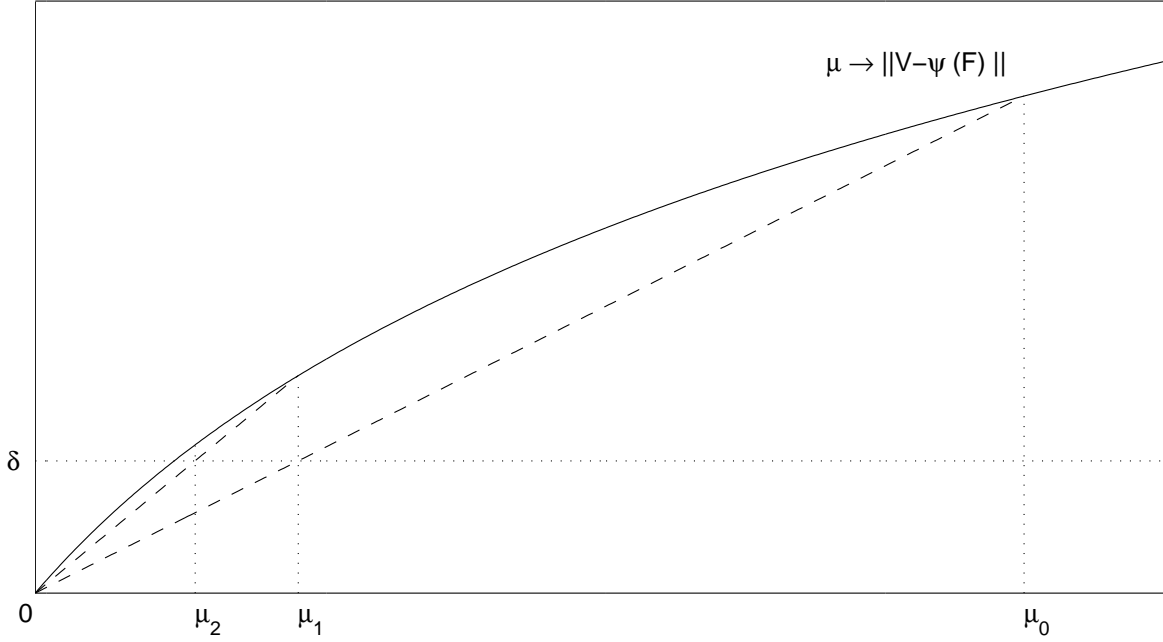


Figure 1. Iterative search of the regularization parameter μ that fits the Morozov discrepancy principle

A.6.2 Test of A1

We first used algorithm A1 for data corresponding to a Dirac measure or a sum of Dirac measures. We present the results in the case of one single Dirac and the sum of three Dirac measures to illustrate the behaviour of the algorithm that is able to detect more than one direction.

Case of one Dirac measure

We have made computations with $\theta_d = 1.5$, $\phi_d = 2$ and for $P_i = i/2$ with $i = 1, \dots, 50$. For each value of P_i we have computed the relative power error given by

$$\epsilon_\alpha^i = \frac{|\alpha_i - P_i|}{P_i}$$

where α_i is the computed approximation of total power P_i . For this example we set $\mu = 1$. The function $\epsilon_\alpha(P)$ is plotted in figure 1. We see that the error increases when P decreases; it is less than 10 % for a power $P \geq 4$. We have also computed the mean direction $(\bar{\theta}, \bar{\phi})$ given by

$$\bar{\theta} = \frac{1}{\|F\|_{L^1(E, \mathbb{R}, \sigma)}} \int_E \theta F(\theta, \phi) d\sigma(\theta, \phi), \quad \bar{\phi} = \frac{1}{\|F\|_{L^1(E, \mathbb{R}, \sigma)}} \int_E \phi F(\theta, \phi) d\sigma(\theta, \phi). \quad (\text{A.39})$$

of a Dirac with $\phi_d = 2, P = 8$ and for θ_d varying from 0 to π . From these values we can compute the angles errors ϵ_θ and ϵ_ϕ by

$$\epsilon_\theta = \frac{|\theta_d - \bar{\theta}|}{\pi}, \epsilon_\phi = \frac{|\phi_d - \bar{\phi}|}{2\pi}$$

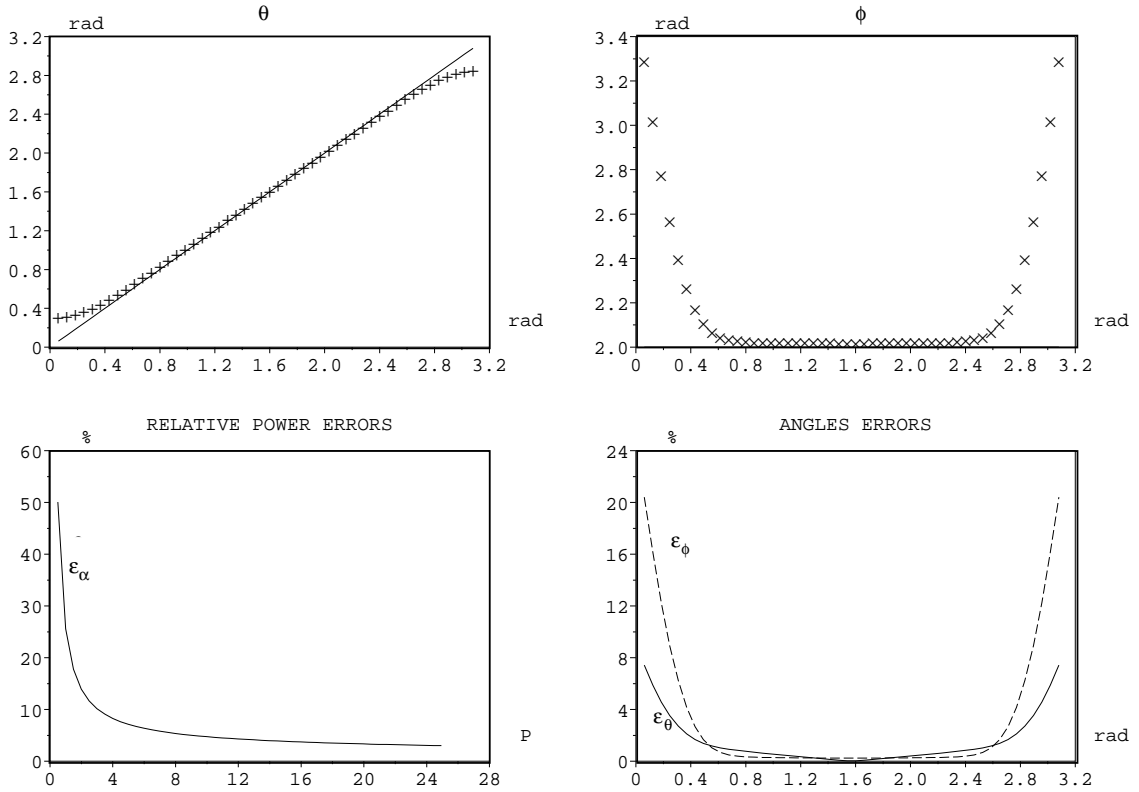


Figure 2. Variation of the average direction $\bar{\theta}$ (plotted with +, the true value is the solid line), $\bar{\phi}$ (plotted with x) and errors $\epsilon_\theta, \epsilon_\phi$ for a dirac with $P = 8, \phi_d = 2$ and θ_d varying from 0 to π . The plot named “relative power error” show the variation of ϵ_α for a Dirac with $\theta_d = 1.5, \phi_d = 2$ and P varying from 0.5 to 25.

We can see how the average direction changes with θ_d . We see that the average direction is close the true one when it is far enough from the poles. The “large” error on ϕ_d near the poles can be explained by the “bad” representation of the unit sphere in (A.39). Similarly, we obtain a large error if ϕ_d is chosen near 0 or 2π . Another value for ϕ_d far enough from the edge of E gives similar result as plotted in Figure 2. We conclude that the results obtained in this case are satisfactory. The method is able to detect the direction and the power of a dirac with small errors if the direction is not too close to the edge of E and if the power is great enough.

Case of the sum of three Dirac measures

We have built an example for a sum of three Dirac measures which support is vertices of an equilateral triangle on E . For this example we have added a noise to the simulated data V . It verifies $\|V - V^\delta\|_{\mathbb{C}^n} < 1$ (precisely = 0.764, this correspond to a relative noise level of 3.1%). The regularization parameter μ was chosen according to the noise level ; we have taken $\mu = 0.5$ because the corresponding solution has a small error and $1 = \delta < \|V - \psi[F]\|_{\mathbb{C}^n}$, more precisely the relative error is 5.5% (see Table A.1 for the results). We plot the contour of the solution on Figure 3. We observe three peaks centered on the directions of the three Dirac measures. The solution is in this way satisfactory.

We have performed a sensitivity analysis with respect to μ . Indeed, in section A.3 we gave an interpretation of entropy as the “distance” to the non-informative probability density g . To see the numerical effects of entropy, we computed the solution for the same data with different values for μ . Theoretically, a great value of μ leads to a solution with a great entropy, that is a solution not too “far” from g . In Figure 3 we see the result obtain for $\mu = 50$ and the different results are reported in Table 1. We see that the error ϵ_r quickly increases with μ while H is decreasing. The number of iteration is dramatically increasing when μ becomes small, so that we are not able to compute the μ -solution for all μ with this algorithm. We observe that we need more and more iteration when $\mu \rightarrow 0$, to converge. The solution obtained for $\mu = 50$ is very flat and has a very large error $\epsilon_r = 117.6$. We remark that the optimal value function is not increasing but there is no contradiction with proposition 2 since we have to add $\frac{\mu}{e}$ to obtain an increasing function.

μ	α	ϵ_α	ϵ_r	H	J_μ	Min	Max	#It.	τ
0.5	9.891	2.07%	1.842	-1.452	5.741	2.200E-6	6.947	8377	2.5E-3
1	9.868	1.83%	3.159	-1.660	9.360	1.521E-4	4.597	3317	6.25E-2
5	9.814	1.27%	10.27	-1.990	24.691	2.678E-2	3.136	465	0.05
50	8.008	17.35%	117.633	-2.338	17.862	0.195	1.725	7	1

TAB. A.1 – Results obtained for a simulated noisy data V of a sum of three Dirac measures. We have made four computation for different values of the regularization parameter μ . The Min and Max are respectively the minimum and the maximum of the solution on E

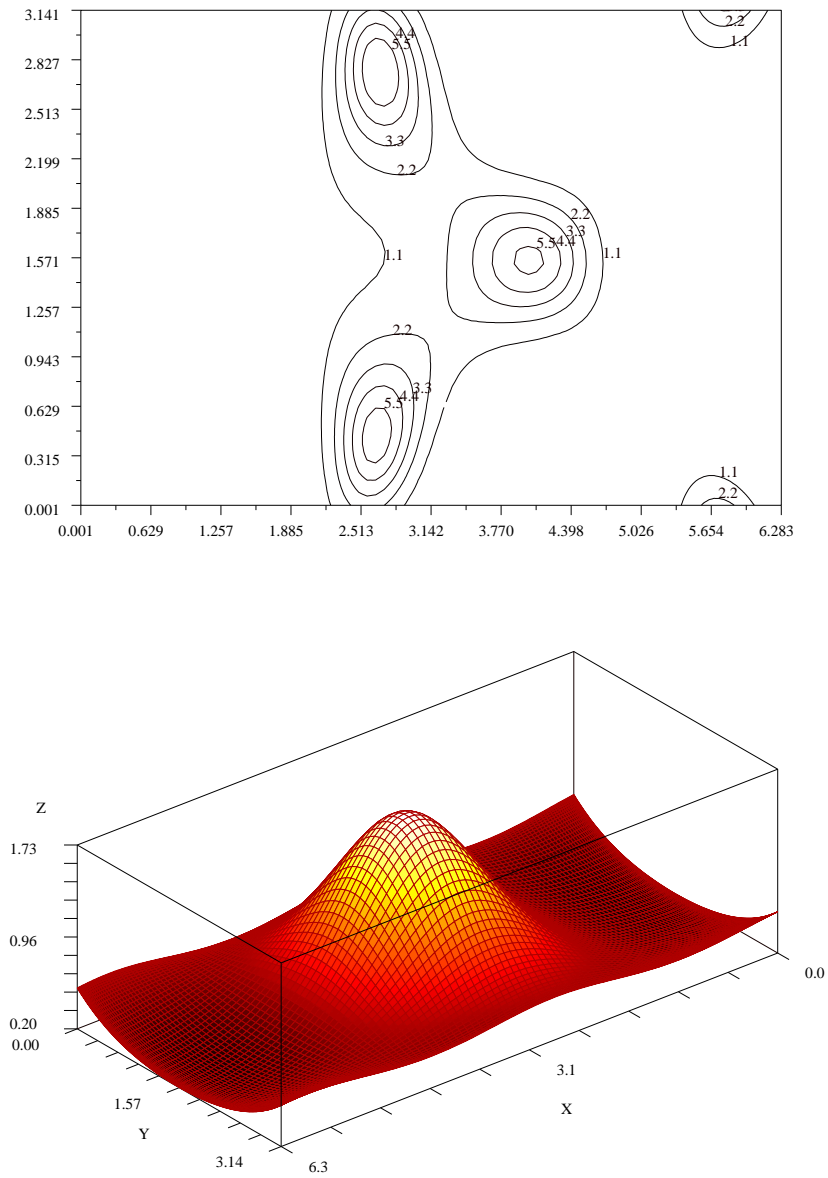


Figure 3 :Up : contour plot of the solution obtained by algorithm (A.25) for three Dirac measures and $\mu = 0.5$. Down : solution obtained for three dirac, with $\mu = 50$.

A.6.3 Test of A2 on a continuous density

In the previous section we presented tests of the first algorithm to identify a three-Dirac distribution. Now we give an example of reconstruction of an a.e. continuous distribution with algorithm A2. Let be $F_t(\theta, \phi) = 4(\cos^2 \theta + \sin^2 \phi)$, the function used to simulate the data vector V . In this example we use a noisy data V^δ with $\|V - V^\delta\| = 0.670$ which is about 5.1% compared with $\|V\|_{\mathbb{C}^n}$. To avoid a so-called inverse-crime we have used a finer quadrature-rule to compute the data V : more precisely we used 64 points on the interval $[0, \pi]$ and 128 points on $[0, 2\pi]$.

The function F_t is represented in Figure 4. We use a spherical plot : the value of the function is described by a gray level code on each hemisphere. We see that F_t is not continuous at the two poles of the sphere : this discontinuity is explained by the “bad” representation of the unit sphere : $E = [0, \pi] \times [0, 2\pi]$. The solution obtained by algorithm A2 is also plotted in Figure 4.

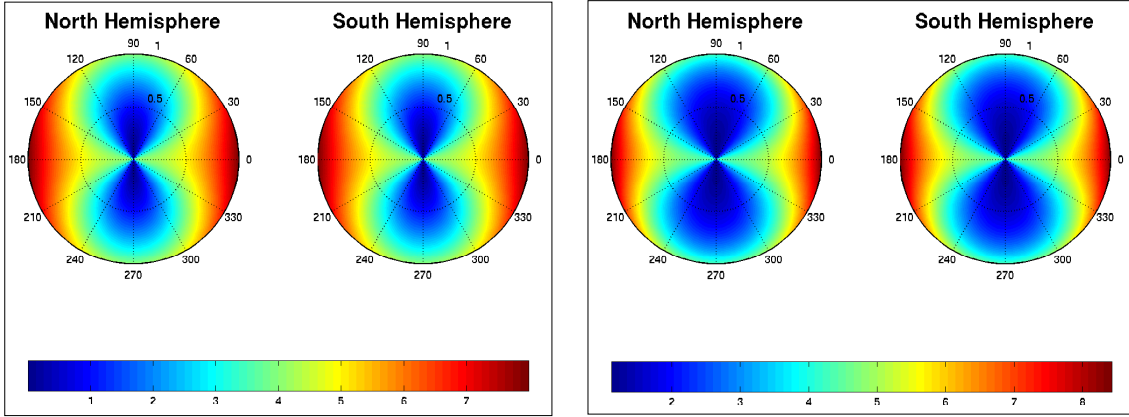


Figure 4. *Left* : spherical plot of the initial function F_t which provides the data vector V . *Right* : solution obtained by algorithm A2 corresponding to the data V of the function F_t

We see that the computed solution looks like F_t . We used the following parameters : $\epsilon = 10^{-12}$, $\mu = 3$, $\tau = 0.008$. The algorithm stopped after 1818 iterations, the error with respect to the data ϵ_r and the negentropy H are respectively 1.127 and -2.460 . (The error is about 8.0% compared to $\|V\|_{\mathbb{C}^n}$). We see that this solution is satisfying because ϵ_r is greater than the noise level but it is small enough anyway.

This example allows to compare problems (\mathcal{P}_μ) and $(\tilde{\mathcal{P}}_\mu)$. Let μ_e be defined by $\mu_e = \mu \delta^* \|F_{l^*}\|_{L^1(E, \mathbb{R}, \sigma)}$ with the notations of proposition 4. With this proposition we know that f_{l^*} achieves the minimum of functional

$$F \mapsto \left\| V - \frac{\mu_e}{\mu} \psi[F] \right\|_{\mathbb{C}^n}^2 + \mu_e H(F)$$

over K . So f_{l^*} is the solution of $(\tilde{\mathcal{P}}_{\mu_e})$. We can now compare the two problems by computing the solution of (\mathcal{P}_{μ_e}) .

We found $\mu_e = 176.35$ with the computation of the solution of $(\tilde{\mathcal{P}}_{\mu_e})$. The computed solution of problem (\mathcal{P}_{μ_e}) is very far from the data ($\epsilon_r = 5133.6$). So, for this example, the solution of $(\tilde{\mathcal{P}}_{\mu_e})$ is much better than the solution of (\mathcal{P}_{μ_e}) . For $(\tilde{\mathcal{P}}_{\mu})$ we get a much smaller error for a much larger μ .

A.7 Conclusion

In plasma physics, the determination of the directional power density distribution of an electromagnetic wave from the measurement of the field components is an inverse ill-posed problem. This problem can be written as

$$\psi[F] = V$$

where ψ is a linear bounded operator from $\mathbb{H} = L^2(E, \mathbb{R}, \sigma)$ to \mathbb{C}^n , V the spectral matrix and F the wave distribution function (WDF). Lefeuvre et al. [25] have proposed to solve this problem by maximizing an entropic term $-H(F) = -\int_E F \ln F d\sigma$ under the constraint $\psi[F] = V$. However this constraint is too “restrictive”, it indeed limits the feasible domain to a linear subspace of \mathbb{H} . That’s why we have studied the relaxed problem

$$(\mathcal{P}_{\mu}) \left\{ \begin{array}{l} \min J_{\mu}(F) = \|V - \psi[F]\|_{\mathbb{C}^n}^2 + \mu H(F) \\ F \in K = \{f \in \mathbb{H} | f \geq 0 \text{ } \sigma.a.e\} \end{array} \right. ,$$

where μ is a regularization parameter. The latter parameter has to be chosen small enough to allow a solution with a small error thanks to the data V , and large enough for stability. Solving (\mathcal{P}_{μ}) permit us to search the solution in a much larger domain. More precisely the obtained solution verifies $\|\psi[F] - V\|_{\mathbb{C}^n} \leq \varepsilon$ with $\varepsilon > 0$, this inequality is clearly more realistic from the numerical point of view. The regularization by the negentropy functional is also important from the physical point of view. Instead of using standard method to solve the problem (\mathcal{P}_{μ}) , we have first built a fixed point algorithm in \mathbb{C}^n (algorithm A1) thanks to a sufficient condition of optimality. Moreover, the uniqueness of the obtained solution has been proved. We have shown the existence of the solution for every value of the regularization parameter. Nevertheless in the numerical computation we have seen that it is not possible to compute the solution for too small value since the computational time goes to infinity.

As both this above method and the Lefeuvre’s one give solutions which don’t maximize entropy in the probabilistic sense, since they are not probability densities, we have built a second algorithm (algorithm A2) derived from the first one which allows the “true” entropy to be maximized.

We have performed numerical tests on experimental data from a satellite that is devoted to the study of magnetosphere. We have compared the solution given by the algorithms described in the present paper to the ones obtained with two different methods that are commonly used by physicists. Results show that the method we described is more accurate and stable. They are reported in [43].

Acknowledgement : We would like to thank the anonymous referee who carefully read the paper and made a lot of corrections and suggestions to improve it.

Annexe B

Ill-posed Problem Solving by an Entropy Regularization Method. Application to the Propagation Analysis of Electromagnetic Waves

O. Prot[†]

Laboratoire de Mathématique MAPMO, université d'Orléans, France

O. Santolík

Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic.

J-G. Trotignon

[†]Laboratoire de Physique et Chimie de l'Environnement, CNRS, France

Abstract. A new algorithm based on entropy regularization has been developed to solve ill-posed problems. Using a modelisation derived from the Wave Distribution Function (WDF) concept it is particularly well suited to determine the direction of propagation of electromagnetic waves. The Morozov's discrepancy principle has been used to fix the value of the regularization parameter. To assess the capabilities of this new method, data associated with an Electron Low Frequency (ELF) hiss event detected by the Freja satellite have been processed. The obtained results are compared to the results coming from a previous study dedicated to the very same data set and for which both the Gaussian Peaks (GP) and the Maximum Entropy (ME) methods were used to derive the wave vector direction. The results are presented as well as a statistical study of the ability of our method to detect plane waves.

B.1 Introduction

In this paper we present an entropy regularization method to solve linear inverse problems. This method has been developed to determine the direction of propagation of electromagnetic waves in homogeneous media [42]. However the method can be applied to other linear inverse problems with finite dimensional data and with a positivity constraint for the solution.

In any media, including space plasmas, the derivation of the wave normal directions of an electromagnetic wave from the measurement of the wave field components is a crucial problem

[61]. Most authors solve the problem by assuming the plane wave approximation. In such a case the electromagnetic field is fully described by its frequency, its wave normal direction k and its polarization characteristics. The problem is well-posed and several methods have been developed to derive k from the estimated spectral matrix of the field components [32].

The hypothesis of a single-plane wave is generally too restrictive. For a more realistic analysis of electromagnetic wave propagations, the concept of the wave distribution function (WDF) has to be used. This concept has been introduced by *Storey and Lefeuvre* [61] in the case of electromagnetic waves in homogeneous plasmas. The WDF is a function $F : S \rightarrow \mathbb{R}^+$ which specifies the distribution of the wave energy density on the unit sphere S for a fixed frequency ω and a given propagation mode, this is a positive function. We denote the field measurement by the function $e : \mathbb{R} \rightarrow \mathbb{R}^s$, where $s \in \mathbb{N}^*$ is the number of field components. Being the spectral matrix of auto and cross-spectrum of function e , at frequency ω , V is an hermitian matrices of size $s \times s$. The WDF F is related to the spectral matrix element V_{ij} of electromagnetic field components by the integral equation

$$V_{ij} = \int_0^\pi \int_0^{2\pi} q_{ij}(\theta, \phi) F(\theta, \phi) \sin \theta d\phi d\theta, \quad (\text{B.1})$$

where θ and ϕ are respectively the polar and azimuthal angles of k with respect to the static magnetic field B_0 (see Figure B.1). The integrating kernels $q_{ij} : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{C}$ are known functions depending on the propagation medium. The general expression of the integrating kernels for a cold plasma is given in *Storey and Lefeuvre* [61, 62]. In his thesis *Santolik* [50] studied the WDF in the case of a hot plasma including the Doppler effect. The WDF concept introduced by *Storey and Lefeuvre* [61] is not the only one, a similar concept has been introduced by *Rönmark* [48] using the Wiener function.

The problem of the determination of the WDF F from the spectral matrix V through equation (B.1) is highly undetermined. We need to make further assumptions on the WDF to determine a unique and stable solution. To solve (B.1) the most generally used technic consist in choosing a model for the WDF then the solution's parameter are computed by a least square minimization. For some WDF models, the minimization is not easy because we have to deal with a non-convex cost function. It is the case for the method of Gaussian Peaks (GP). Moreover, the choice of a model for the WDF implies that we have a priori information about it. For example, with the GP method we must choose the number of peaks.

One of the best known methods to solve (B.1) is the *Lefeuvre et al.* [29] Maximum Entropy (ME) method. In this method the integration kernels are first orthogonalized and then truncated to eliminate the most dependent kernels. Before the computation the number of kernels to be kept during the truncation are chosen adequately. Finally, the solution is computed by maximizing the entropy.

By truncating the kernels a lot of data are lost, consequently we decided to use a regularization method to keep the whole initial spectral matrix. Regularization methods are widely used to solve such inverse problems [22, 64]. Using an entropy regularization method, we determine a unique and stable solution. Moreover, we can adjust the stability of the solution to noise level via Morozov's discrepancy principle (see section B.2.2 for more details). *Prot et al.* [42] have studied the entropy regularization method and a finite dimensional algorithm

has been derived. One advantage of this algorithm is that we do not need to make a discretization of the WDF. This algorithm has also been tested on simulated data in the case of an electromagnetic wave in vacuum, and the results are encouraging. In this paper we present a numerical test on experimental data coming from the magnetospheric research satellite Freja.

B.2 Method

The principle of entropy regularization method to solve an ill-posed inverse problem will be presented below. But, let's first recall the definition of a well-posed problem.

Definition 3 *Let $f : X \rightarrow Y$ where X and Y are two Hilbert spaces. The problem of finding x from y by resolving equation $f(x) = y$ is well-posed if*

1. *There exists a unique solution.*
2. *The solution depends continuously on y (stability).*

otherwise the problem is said ill-posed.

Let \mathbb{H} be the Hilbert space $L^2(E, \mathbb{R}, \sigma)$, by identifying the spectral matrix as a vector $V \in \mathbb{C}^n$, with $n := s^2$, we can reformulate equation (B.1) by

$$\psi(F) = V, \tag{B.2}$$

where $F \in \mathbb{H}$ and $\psi : \mathbb{H} \rightarrow \mathbb{C}^n$ is a continuous linear operator corresponding to the integral equation (B.1). The problem of finding F from V through this equation is ill-posed and even highly undetermined, since ψ is a linear operator from an infinite dimensional space to a finite one.

Regularization methods are widely used to solve such kind of ill-posed problems. A classical example is the Tikhonov's regularization method [64], where a solution is computed by solving the minimization problem

$$\begin{cases} \min \|V - \psi(F)\|_{\mathbb{C}^n}^2 + \mu \|F\|_{\mathbb{H}}^2 \\ F \in K \end{cases} \tag{B.3}$$

where K is the constraint set, μ the regularization parameter, and $F \mapsto \|F\|_{\mathbb{H}}^2$ is the regularizing functional. The parameter μ must be chosen with caution to make a small error with the data and to obtain a stable solution. For the entropy regularization method, an entropic term H as regularizing functional is chosen. The entropy of a probability density F on S is given by $-H(F) := -\int_0^\pi \int_0^{2\pi} F(\theta, \phi) \ln F(\theta, \phi) \sin \theta d\phi d\theta$. Shore and Johnson [57] have shown that this is the only way to characterize the amount of information of the density F under very general axioms. The maximization of the entropy leads to a density F which contains the least amount of information. Moreover, this maximization is made easily since $-H$ is a strictly concave functional.

B.2.1 First minimization problem

We obtain the entropy regularization solution of problem (B.2) by solving problem

$$(\mathcal{P}_\mu) \begin{cases} \min \|V - \psi(F)\|_{\mathbb{C}^n}^2 + \mu H(F) \\ F \in K := \{f \in \mathbb{H} | f \geq 0 \text{ a.e}\} \end{cases} .$$

For all $\mu > 0$ there exists a unique solution for this problem [1, 16]. However, the next proposition gives a way to compute the solution numerically.

Proposition 6 For all $\mu > 0$, (\mathcal{P}_μ) has a unique solution F_μ that verifies

$$\begin{cases} F_\mu = \exp\left(-1 + \frac{2}{\mu}\psi^*(l)\right) \\ l = V - \psi(F_\mu) \end{cases} . \quad (\text{B.4})$$

The proof of this proposition can be found in *Prot et al.* [42]. From this proposition we see that we can compute the solution with a fixed point algorithm. Moreover, the use of the adjoint operator $\psi^* : \mathbb{C}^n \rightarrow \mathbb{H}$ permits to have a finite dimensional algorithm. This is very important since we do not need to make a discretization of the function F for the computation. The analytical expression of the solution and thus the regularity of the WDF are also known. More precisely the regularity of the solution is related to the regularity of the adjoint operator ψ^* .

In fact, the solution obtained by resolving (\mathcal{P}_μ) is not totally satisfying. The first problem is that $H(F)$ does not represent the entropy of the WDF since F is not a probability density. So we do not compute the WDF that contains the least amount of information. However by minimizing H we flatten the WDF, and intuitively, a flat solution corresponds well with a low amount of information. Another problem in this method is the non-linearity introduced by H , this is very annoying because the solution will depends of the scale of the data. More precisely, if the solution of (\mathcal{P}_μ) is F , the solution for the data λV is not λF . To suppress these undesired effects, a modified version of the minimization problem (\mathcal{P}_μ) has been introduced. It is presented in the next subsection.

B.2.2 Second optimization problem

The first optimization problem (\mathcal{P}_μ) , permits to obtain a solution of the inverse problem (B.2) via a finite dimensional algorithm. However, in (\mathcal{P}_μ) we do not maximize the entropy of the solution and the solution is not linearly dependent from the data. To overcome theses problems we can consider a new optimization problem

$$(\mathcal{P}_\mu^2) \begin{cases} \min \|V - \alpha\psi(F)\|^2 + \mu H(F) \\ (\alpha, F) \in \mathbb{R}^+ \times K \cap \{\|F\|_{L^1} = 1\} \end{cases}$$

In this minimization problem, a parametrization of the WDF αF , where α is the value of the integral of the WDF and F the probability density of the wave, is made. The constraints on F permit to minimize the entropy in the sense of the probability. If the solution (α, F) of (\mathcal{P}_μ^2) exists then for all $\lambda > 0$, $(\lambda\alpha, F)$ is the solution of $(\mathcal{P}_{\mu\lambda^2}^2)$ for the spectral matrix λV . This partly fixes the nonlinearity of the solution encountered in (\mathcal{P}_μ) , however any change of scale will have an effect on the regularization parameter.

Problem (\mathcal{P}_μ^2) is much more complicated than (\mathcal{P}_μ) , we have no optimality condition because of the bilinear term $(\alpha, F) \mapsto \alpha\psi(F)$. Nevertheless, it is possible to obtain a solution of (\mathcal{P}_μ^2) by using a modified version of the algorithm given by Proposition 6. This Entropy Regularization Algorithm (ERA) is

1. Initialization

Choose $V \in \mathbb{C}^n, l_0 \in \mathbb{C}^n, \mu > 0, \epsilon > 0, \tau \in]0, 1]$.

2. Iteration κ

(a) Compute $g(l_{\kappa-1}) = \psi[\Gamma(l_{\kappa-1})]$, where

$$\Gamma(l_{\kappa-1}) : (\theta, \phi) \mapsto \exp\left(-1 + \frac{2}{\mu}\psi^*[l_{\kappa-1}](\theta, \phi)\right).$$

(b) Compute $\delta_{\kappa-1} = \frac{Re(\langle V, g(l_{\kappa-1}) \rangle_{\mathbb{C}^n})}{\langle g(l_{\kappa-1}), g(l_{\kappa-1}) \rangle_{\mathbb{C}^n}}$.

(c) Compute $l_\kappa = (1 - \tau)l_{\kappa-1} + \tau(V - \delta_{\kappa-1}g(l_{\kappa-1}))$.

3. Stopping criterion

If $|l_\kappa - l_{\kappa-1}| < \epsilon$, then STOP, else $\kappa := \kappa + 1$ and go to 2.

The symbols Re and $\langle \cdot, \cdot \rangle$ denotes respectively the real part and the classical scalar product. *Prot et al.* [42] have demonstrated the convergence of this algorithm whenever the regularization parameter μ is chosen large enough. Moreover, if it converges to l^* and δ^* then $(c\delta^*, \Gamma(l^*)/c)$ is the unique solution of $(\mathcal{P}_{\mu c \delta^*}^2)$ where $c = \|\Gamma(l^*)\|_{L^1(E, \mathbb{R}, \sigma)}$. This algorithm works perfectly well on numerical examples. Nevertheless we have not yet proved the convergence for small values of μ [42].

We observed numerically that τ must be chosen small enough for the algorithm to converge. Nevertheless since the speed of the algorithm is related to τ , we must choose it as large as possible. The experience shows that a τ value smaller than 0.1 is sufficient enough for the algorithm to converge rapidly. The choice of the regularization parameter μ is done using the Morozov's discrepancy principle [22], i.e. we choose μ in such a way that the corresponding error in the data is equal to the noise level $N > 0$. Numerically, it means that μ fit the Morozov's discrepancy principle if the limit $(\delta_\mu, \Gamma(l_\mu))$ of the ERA verify $0.9N \leq \|V - \delta_\mu\psi[\Gamma(l_\mu)]\| \leq 1.1N$. An iterative search is used to derive the μ value [42]. Hence we do not have to tune any parameter during the computation. The other initialization parameters are set to $\epsilon = 10^{-7}$ and $l_0 = 0$.

B.3 Results

B.3.1 Analysis of an ELF hiss event recorded by Freja

The entropy regularization algorithm has been applied to experimental data recorded by the F4 wave experiment of the magnetospheric research satellite Freja. The F4 experiment measures waveforms of three orthogonal magnetic components and one electric component. The reader is referred to *Holback et al.* [20] for more details. The data were measured during an intense emission of ELF hiss, on april 8, 1995, between 1548 and 1558 UT. *Santolík*

and Parrot [53] have made a plane wave analysis of this event and a WDF analysis on 6 points inside the ELF hiss spectrogram, from which the 6 spectral matrices we have used are calculated. This event has been analyzed with the cold plasma approximation. We have used the entropy regularization algorithm on strictly the same data to avoid any ambiguity in the scientific interpretation of results. More precision about the integration kernel model can be found in *Santolík and Parrot* [53]. As the spectral matrices are noisy and the level of the noise is not the same on each component, a normalization of the spectral matrices has been made to obtain the same order of magnitude of the error on each component. Figure B.1 shows a typical representation of a WDF. It is the same representation used by *Santolík and Parrot* [53].

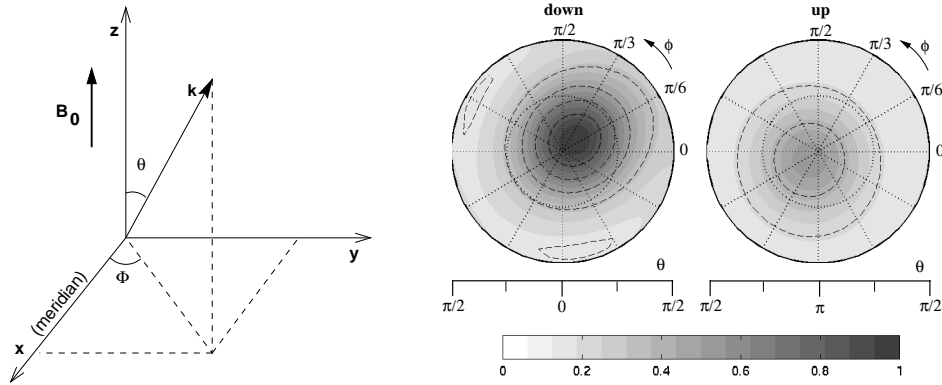


FIG. B.1 – Left : the polar θ and azimuthal angles ϕ in the coordinate system. Right : representation of a WDF with a polar plot, the sphere is divided into two hemispheres, the down hemisphere is for $0 \leq \theta \leq \pi/2$ whereas in the up hemisphere $\pi/2 \leq \theta \leq \pi$. The distribution of the energy on each hemisphere is represented with a gray color code, and with contour lines. The WDF is normalized and the contour lines represent the levels $\frac{\ell}{6}, \ell = 1, \dots, 5$.

We present the results obtained for 3 of the 6 spectral matrices. Figure B.2 shows the results obtained with the ERA and the ones obtained by *Santolík and Parrot* [53] with the GP and ME methods. For point 1 the three methods detect a peak in the downgoing direction. The location of this peak is the same for the three methods. However the ERA gives a very small peak in the upgoing direction. For point 2, the GP and ME solutions exhibit two peaks, the primary one is in the downgoing direction and the other is less powerful and located in the upgoing hemisphere. In the ERA solution, only one diffuse peak is actually obtained, in the downgoing direction. In the other hemisphere the ERA solution is almost flat. Nevertheless, it seems that at the location of the GP and ME secondary peak the ERA solution is also a little bit more powerful. For point 5, the three methods detect a peak in the downgoing direction. The ME peak is very sharp compared with the GP and ERA ones.

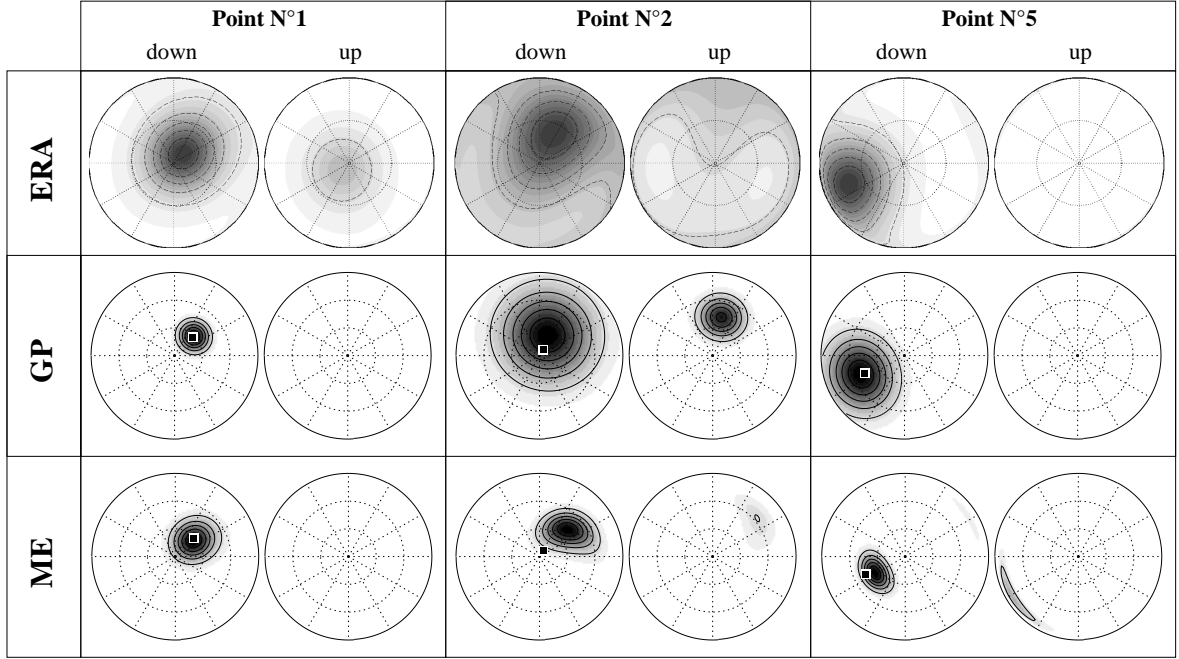


FIG. B.2 – Direction of the k-wave vector derived from the Entropy Regularization Algorithm, ERA (top), the method of gaussian peaks, GP (middle), and the maximum entropy method, ME (bottom). The GP and ME Figures are borrowed from *Santolík and Parrot* [53]. The square in GP and ME Figures represents the direction computed under the plane wave approximation.

B.3.2 Study of the ability to detect a plane wave

A statistical test has been made to see if the method is able to detect the direction in which a single-plane wave propagates. For this, we have taken a random spectral matrix corresponding to a plane wave of random direction (θ_i, ϕ_i) uniformly distributed on the sphere and of random power. Then ERA has been applied. Finally, the direction of the most powerful peak (θ'_i, ϕ'_i) is computed by derivation. To have a good statistical sample, this is done 1000 times. So $1 \leq i \leq 1000$, and for each i the angular distance ϵ_i between the true and the estimated direction is computed by

$$\epsilon_i = \arccos [\cos(\theta_i - \theta'_i) \cos(\phi_i - \phi'_i) + \cos(\theta_i) \cos(\theta'_i)(1 - \cos(\phi_i - \phi'_i))]. \quad (\text{B.5})$$

The results are shown in Figure B.3. The frequency of the measured angle error ϵ_i , and the repartition function are plotted in panels a) and b), respectively. As shown in b), the direction error is lower than 30° in almost 97% of the case.

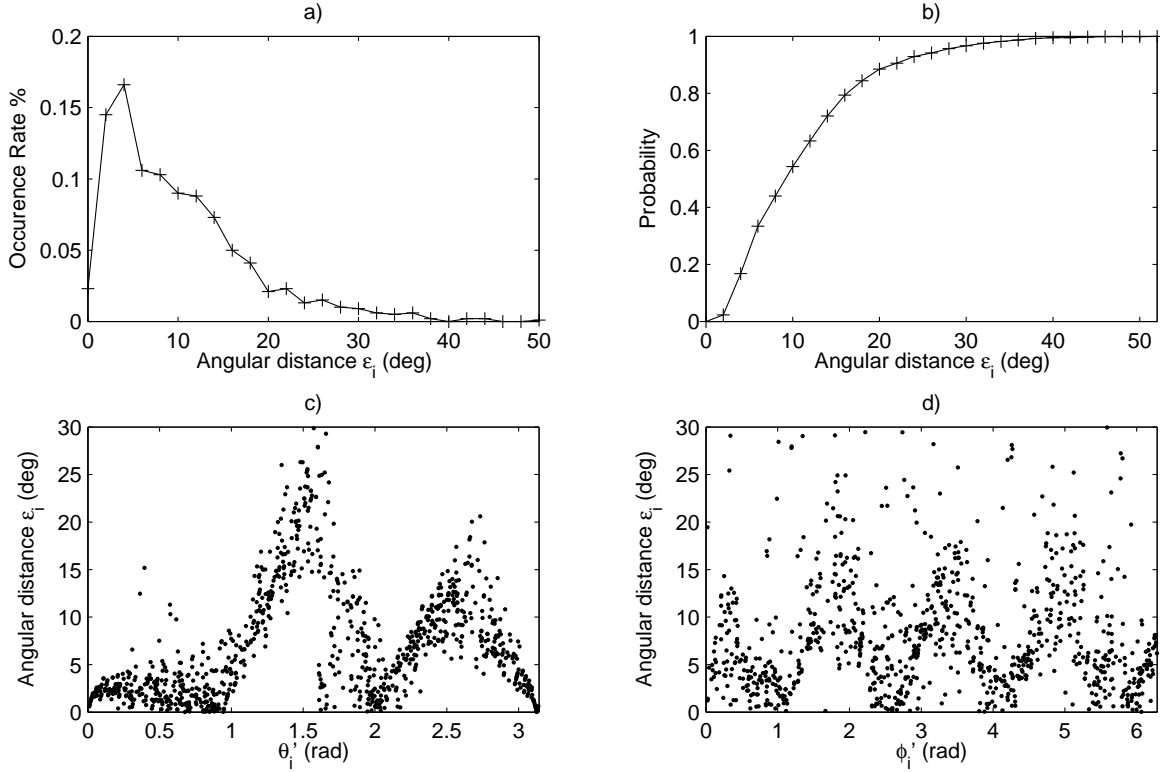


FIG. B.3 – Statistical results of the detection of 1000 random plane waves : a) is the distribution of the angle error ϵ_i , b) represents the repartition function of the angle error ϵ_i , c) and d) are the plots of the errors ϵ_i as functions of angles θ'_i and ϕ'_i .

B.4 Discussion

In this paper we have presented the entropy regularization method developed by *Prot et al.* [42] to compute the WDF from field measurements. We have done numerical tests on experimental data coming from the Freja satellite, which have already been analyzed by *Santolík and Parrot* [53].

Numerically, the entropy regularization algorithm, ERA, works satisfactorily for all the spectral matrices computed on these data. Moreover, the process for computing the solution, which is unique, is entirely automatic unlike both the gaussian peaks, GP, and maximum entropy, ME, methods. The main role of the μ parameter is to ensure that the provided solution remains stable with respect to a disturbance of the order of the noise level. The Morozov's discrepancy principle allow such a μ determination. This permits to choose μ quickly and we obtain solution which are compatible with the other solutions displayed in Figure B.2

The wave distribution function, WDF, obtained with the ERA fits well the solutions of the

GP and ME methods. However, they are more diffuse. This can be explained by the entropy maximization principle : the ERA solution is indeed known to contain less information. Figure B.2 shows that the GP and ME methods give more precise information on the WDF. However the initial problem being highly undetermined, a WDF with a low information is always preferable. This means that the ERA should be first used to give the global geometry of the WDF, because no parameter adjustment and no truncation of the spectral matrix elements are required, which are great advantages for statistical studies. Moreover, the results of section B.3.2 show that globally the method is able to detect quite accurately ($\epsilon_i \leq 30^\circ$) the direction of a plane wave. However, Figures B.3.c) and B.3.d) show that the accuracy of the detection depends on the direction of the plane wave. We see that the method is less accurate when the polar angle θ is near $\pi/2$. To perform more precise analyses, other methods such as the GP and ME methods have then to be used provided enough information about the nature of observed waves are available : for example, the expected number of wave components, the energy distribution between these components, and the noise level.

I would like to thanks M. Bergounioux (University of Orléans, France), F. Lefeuvre and J.L. Pinçon (CNRS, LPCE, France) for their very helpful remarks on this work.

Annexe C

An Entropy Regularization Method Applied to the Wave Distribution Function Identification for an ELF Hiss Event

Olivier Prot[†]

MAPMO, Université d'Orléans, France.

Ondřej Santolík

Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

Jean-Gabriel Trotignon

[†]Laboratoire de Physique et Chimie de l'Environnement, Centre National de la Recherche Scientifique, Orléans, France.

Hervé Deferaudy

Centre d'étude des Environnements Terrestre et Planétaires, Institut Pierre Simon Laplace, Vélizy, France.

*prot@cnrs-orleans.fr, ondrej.santolik@mff.cuni.cz, jean-gabriel.trotignon@cnrs-orleans.fr,
herve.deferaudy@cetp.ipsl.fr*

Abstract An ERA (entropy regularization algorithm) has been developed to compute the wave-energy density from electromagnetic field measurements. It is based on the WDF (wave distribution function) concept. To assess its suitability and efficiency, the algorithm is applied to experimental data that has already been analyzed using other inversion techniques. The FREJA satellite data that is used consists of six spectral matrices corresponding to six time-frequency points of an ELF hiss-event spectrogram. The WDF analysis is performed on these six points and the results are compared with those obtained previously. A statistical stability analysis confirms the stability of the solutions. The WDF computation is fast and without any pre-specified parameters. Although the entropy regularization leads to solutions that agree fairly well with those already published, some differences are observed, and these are discussed in detail. The main advantage of the ERA is to return the WDF that exhibits the largest entropy and to avoid the use of *a priori* models, which sometimes seem to be more accurate but without any justification.

C.1 Introduction

The propagation direction of an electromagnetic wave in a homogeneous medium can be modelled by the wave distribution function (WDF). The concept of WDF was introduced in [61, 62, 63] to study the propagation of electromagnetic waves in plasma. It is a positive function which represents the directional distribution of the electromagnetic wave energy. The WDF $F(\theta, \phi)$ is related to the estimated spectral matrix components V_{ij} by the integral equation over the unit sphere $[0, \pi] \times [0, 2\pi]$:

$$V_{ij} = \int_0^\pi \int_0^{2\pi} Q_{ij}(\theta, \phi) F(\theta, \phi) \sin \theta d\theta d\phi, \quad (\text{C.1})$$

where $Q_{ij} : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{C}$ are the integration kernels. They correspond to the auto- and cross-power spectra of the electromagnetic components of a plane wave in a given polarization mode [28]. The kernels Q_{ij} are known analytical functions that depend on the ambient plasma parameters. Their algebraic expressions in the cold plasma case are given in *Storey and Lefeuvre* [63]. The angles θ and ϕ are defined in Figure C.1.

We want to compute the WDF from the spectral matrix V through equation (C.1). This is clearly an ill-posed inverse problem, hence we must make further assumptions on F to get a stable solution. There are many methods to solve this kind of inverse problem, for example the method of maximum entropy (ME) of *Lefeuvre and Delannoy* [25, 26] in the cold plasma case. This method was modified in *Santolík* [50] and *Santolík and Parrot* [52] for a hot plasma, including Doppler effect. Another way to solve this inversion problem is to choose a parametric model for the WDF and then to identify the parameters that best fit the data. This identification is often made using a least-square minimization, for example, in the Levenberg-Marquardt algorithm [41, 6]. With this approach, *Santolík* used a Gaussian peaks (GP) model [53] to solve equation (C.1). Nevertheless, some problems arise because several parameters of the WDF models have to be fixed before starting computations. As a consequence, numerous solutions are obtained and there is no way to determine which one is the best.

The key step of *Lefeuvre and Delannoy* [26] maximum entropy method, is to orthogonalize the kernels Q_{ij} and then to make a truncation in order to retain the dominant linearly independent kernels. The discarded kernels are assumed “unimportant” for the wave analysis due to their low eigenvalues. In this process, many components of the initial spectral matrix are unfortunately lost.

In this paper, we use a new algorithm, called ERA (entropy regularization algorithm), which is based on the concept of entropy minimization to find the propagation direction of an electromagnetic wave. This method was developed in *Prot et al.* [42] for the case of a linear inverse problem with finite dimensional data and a positivity constraint. The algorithm is easy to use, the computations are fast and the resulting solutions minimize the entropy in the probabilistic sense. Therefore, it partially fixes the non-linear dependence of the solution introduced by entropic inversion methods. *Prot et al.* [42] tested the ERA in the case of waves propagating in vacuum. Satisfactory results have been obtained for simulated noisy spectral matrices. The ability of the ERA technique to detect a single plane-wave has also been studied in the case of a cold plasma model *Prot et al.* [44].

Regularization methods are widely used to solve ill-posed problems [65, 22]. The WDF obtained by using entropy as the regularizing functional represents the minimum information required to fit/explain the data [57, 37]. To ensure the stability of the solution, a regularization parameter has to be fixed, and, unlike in the ME method, no data of the spectral matrix needs to be rejected. In our computation, the regularization parameter is automatically selected in accordance with the Morozov's discrepancy principle [22].

The objective of the paper is to show the results obtained with the entropy regularization method, ERA, applied for the first time to experimental data. The ERA method, as well as the maximum entropy (ME) and Gaussian peaks (GP) methods are recalled in Section C.2, for comparison. Section C.3 starts with a brief description of the FREJA magnetospheric mission which provided the data used here. Then, a description of the ERA algorithm is presented at the end of Section C.3. Finally, before the discussions and conclusions, the FREJA data-analysis results are presented in Section C.4. In particular, the stability of the resulting ERA solutions is statistically evaluated.

C.2 Methods

C.2.1 Description of two other WDF methods

The methods for computing the WDF considered here [50, 53], are based on least square fit techniques. As parametric models of the WDF are used, the problem amounts to determining the parameter values that best fit the data.

In his thesis, *Santolík* [50] used a model of Gaussian peaks (GP) to compute the WDF. The main advantage of the GP is that it returns the best solution when the number of peaks, i.e. the number of wave propagation directions is known in advance. For each peak, the energy, the beam width, and the angles (θ, ϕ) that describe the propagation direction are computed. It is worth noting that, with many (> 3) peaks, the obtained solution becomes unstable and sometimes no solution at all is found. As the expected number of peaks is usually unknown, the GP can not be used automatically in statistical analysis.

In the Lefeuvre and Delannoy method (ME) [24, 29, 26], the model for the WDF is chosen to maximize the entropy. The only parameters to compute are the Lagrange multipliers of the maximization problem. However, since the integration kernels are not linearly independent, they have to be orthogonalized, and thus many dependent kernels have to be discarded to obtain a stable solution. In this process, it is assumed that the discarded kernels are insignificant for the wave analysis. In practice, the number of discarded kernels is fixed before the computation, and the geometry of the solution turns out to depend on it. *Storey* [60] has shown that some spectral matrices cannot be analyzed with this method because the discarding process leads to WDF solutions that do not fit the measured spectral matrix. This limitation has encouraged us to use a regularization technique that allows the whole initial spectral matrix to be kept, with the solution regularized for stability.

In the literature, the GP and ME methods have not been used automatically, mainly because some parameters have to be fixed before starting the computation. The ERA method presented here, should, in principle, be well-suited for this kind of analysis. The details of GP and ME methods are described in [26, 50, 53]. We can also mention the method of *Oscarsson*

et al. [38], where an alternative form for the WDF is used, and is not comparable with the one proposed here. The entropy regularization method presented here starts from a different point of view. The main advantage of this point of view is that we do not have to choose any model for the WDF or use a least square minimization. The solution is computed by solving an optimization problem via an iterative algorithm.

C.2.2 Entropy Regularization

The regularization theory, used to solve ill-posed problems, is based on the use of a family of operators indexed by a free positive parameter μ . When μ is small, the error in the solution is small but the solution becomes strongly unstable. Conversely, the error increases with larger μ but the stability is assured. So, a compromise between error and stability must be found. A most often used method of regularization is the one of Tikhonov [65, 64]. For the sake of simplicity, the spectral matrix V and the integration kernels are identified respectively as a vector in \mathbb{C}^n and a function $Q : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{C}^n$. Tikhonov's regularization principle is to solve

$$\min_F \varepsilon_r(V, F) + \mu \|F\|^2, \quad F \geq 0 \quad (\text{C.2})$$

where $\mu > 0$ is the regularization parameter, $\varepsilon_r(V, F)$ is the error in the data

$$\varepsilon_r(V, F) := \left\| V - \int_0^\pi \int_0^{2\pi} Q(\theta, \phi) F(\theta, \phi) \sin \theta \, d\phi \, d\theta \right\|_{\mathbb{C}^n}^2.$$

The strict convexity of the functional norm $\|F\|^2 := \int_0^\pi \int_0^{2\pi} |F(\theta, \phi)|^2 \sin \theta \, d\phi \, d\theta$ will ensure the stability of the solution. Unfortunately, it does not guarantee that the resulting WDF is positive, therefore it cannot be used for our purpose.

The entropy regularization principle is quite similar, but here the entropy, instead of the norm, is used as a regularizing functional. The entropy is based on the concept of the relative information content. *Shore and Johnson* [57] have shown that for two probability densities f and g , the information content $I(f, g)$ of f relative to g can be quantified by

$$I(f, g) := \int_0^\pi \int_0^{2\pi} f(\theta, \phi) \ln \frac{f(\theta, \phi)}{g(\theta, \phi)} \sin \theta \, d\phi \, d\theta. \quad (\text{C.3})$$

The entropy of a probability density is then defined by $-H := -I(f, \bar{g})$, where \bar{g} is the non-informative probability density, i.e. \bar{g} is the density of a the noise in the physical system. Since isotropy is assumed for the determination of the wave-energy density, \bar{g} is assumed to be constant. Hence the entropy is defined by

$$-H(F) := - \int_0^\pi \int_0^{2\pi} F(\theta, \phi) \ln F(\theta, \phi) \sin \theta \, d\phi \, d\theta. \quad (\text{C.4})$$

Using H as a regularizing functional has advantages; it leads to a smooth and positive solution, which is well suited in our framework. Moreover, the minimization of H flattens the solution. In fact, when $\mu \rightarrow +\infty$ the resulting solution becomes constant, which means that we have only trivial information about the variations in the wave-energy density.

The use of the entropy to solve an ill-posed problem is now a standard method in image restoration [41], radio astronomy, seismology, and many other applications [58]. According to [37], the minimization of entropy functional can be used in the WDF computation to reproduce the structures required by the data in the wave-energy density, with suppressed spurious peaks. Thus, we need to solve

$$\min_F \varepsilon(V, F) + \mu H(F), \quad F \geq 0. \quad (\text{C.5})$$

This optimization problem has a unique solution for all μ [1], denoted by F_μ . *Amato and Hughes* [1] and *Engl and Landl* [16] have shown the convergence of F_μ to the solution of maximum entropy, as $\mu \rightarrow 0$. In our case, to ensure the stability of solution, the regularization parameter μ cannot be chosen too small. A fixed-point algorithm has been investigated in *Prot et al.* [42] to solve (C.5) for large μ . This algorithm leads to a smooth solution.

As already stated, the regularization parameter $\mu > 0$ must be chosen sufficiently small to get a low error $\varepsilon_r(V, F_\mu)$, but it should be chosen large enough for the stability of the solution F_μ . The method used to choose μ will be described in Subsection C.2.3.

In fact, $-H$ is not the entropy in the probabilistic sense since the WDF F is not a probability density on $[0, \pi] \times [0, 2\pi]$. This implies that the solution F_μ is not linearly dependent on the data V . Indeed, for the spectral matrix $V' = \lambda V$, the solution of (C.5) is not λF_μ . This is very annoying since the data are always scaled for numerical computations. That is why we have modified the optimization problem to minimize the entropy in the probabilistic sense. Thus, the new optimization problem becomes

$$\begin{cases} \min_{(\alpha, f)} \varepsilon(V, \alpha f) + \mu H(f) \\ f \geq 0, \int_0^\pi \int_0^{2\pi} f(\theta, \phi) \sin \theta \, d\phi \, d\theta = 1, \alpha > 0 \end{cases} \quad (\text{C.6})$$

In this problem α represents the power of the electromagnetic wave and f the probability distribution of the wave direction. The function f is a probability density because of the integral constraint. From this point of view, the WDF of the wave is actually $F = \alpha f$.

The minimization problem (C.6) is much more difficult to solve than (C.5). We do not have a guaranteed optimality condition for this problem since the cost function is not convex and the solution may not be unique. Nevertheless [42] have established an iterative algorithm to compute a local solution of (C.6). This entropy regularization algorithm is as follows.

1. Initialization

Choose $V \in \mathbb{C}^n, l_0 \in \mathbb{C}^n, \mu > 0, \epsilon > 0, \tau \in [0, 1]$.

2. Iteration κ

(a) Compute

$$g(l_{\kappa-1}) := \int_0^\pi \int_0^{2\pi} Q(\theta, \phi) \Gamma(l_{\kappa-1})(\theta, \phi) \sin \theta \, d\phi \, d\theta$$

where

$$\Gamma(l_{\kappa-1}) := \exp \left(-1 + \frac{2}{\mu} \psi^*[l_{\kappa-1}](\theta, \phi) \right)$$

and

$$\psi^*[l](\theta, \phi) := \sum_{j=1}^n l_j Q_j(\theta, \phi).$$

(b) Compute $\delta_{\kappa-1} := \frac{\text{Re}(\langle V, g(l_{\kappa-1}) \rangle_{\mathbb{C}^n})}{\langle g(l_{\kappa-1}), g(l_{\kappa-1}) \rangle_{\mathbb{C}^n}}$.

(c) $l_{\kappa} = (1 - \tau)l_{\kappa-1} + \tau(V - \delta_{\kappa-1}g(l_{\kappa-1}))$.

3. Stopping criterion

If $|l_{\kappa} - l_{\kappa-1}| < \epsilon$, then STOP, else $\kappa := \kappa + 1$ and go to 2.

Suppose that $\alpha_{\kappa} \rightarrow \alpha$ and $l_{\kappa} \rightarrow l$ then (α^*, F^*) is a local solution of (C.6) [42], where

$$\alpha^* = \frac{\alpha}{\int_S \Gamma(l) d\sigma}$$

and

$$F^* = \frac{\Gamma(l)}{\int_S \Gamma(l) d\sigma}$$

The convergence of this algorithm has been proved when the regularization parameter is chosen large enough [42]. However, we do not know whether it converges for every values of $\mu > 0$, but it does work well numerically. We need a small parameter τ for the convergence of the algorithm, but it should not be too small to ensure fast computations. The choice of τ does not affect the solution but small values slow down the computation. In our algorithm, the parameter τ is computed automatically by the following rule : if ERA diverges, i.e. if $\|l_{\kappa+1} - l_{\kappa}\| > \|l_{\kappa} - l_{\kappa-1}\|$, we let $\tau_{n+1} = r\tau_n$ where $0 < r < 1$ and then restart the computation until it converges. After this, to refine the value of τ , we use a linear search, choosing the largest parameter $\{[(N-k)r\tau + k\tau]/N; k = 0, \dots, N-1\}$ for which ERA continues to converge.

We must now check if the nonlinearity problem introduced by the entropy is actually solved. By taking the data $V' = \lambda V$ for $\lambda > 0$ we see from the minimization problem (C.6), that the corresponding solution is nothing but $(\lambda\alpha^*, F^*)$ for the regularization parameter $\mu\lambda^2$. [42] have shown that this remains true for the local solution computed by the ERA. In the next section, this algorithm will be used to compute the WDF using the measurements from the FREJA satellite. However, a rule to find an appropriate regularization parameter has to be defined.

C.2.3 Choice of the regularization parameter

The choice of the regularization parameter is an intricate problem since we must trade off between the error and stability of the solution. For the stability of the solution μ must not be too small, whereas for a small error of fit this parameter must not be too large. In our computation, the Morozov's discrepancy principle will be used to automatically choose this parameter. Essentially, solutions whose errors of fit are smaller than the measured noise level cannot in any way be considered as valid.

Let $\mu > 0$, and denote by (α_μ, F_μ) the obtained solution of (C.6). Suppose that the noise level in the data is $\delta > 0$, then choose the regularization parameter $\mu > 0$ such that $\varepsilon_r(V, F_\mu) = \delta$. However, numerically, an equality is too restrictive. A μ that satisfies $\varepsilon_r(V, F_\mu) \leq \delta \pm 10\%$ is acceptable. To find such a μ we use the fact that $\mu \mapsto \varepsilon_r(V, F_\mu)$ is quite linear for small μ . More precisely, our rule will be

-
1. Choose $\mu_0 > 0$
 2. Compute F_{μ_k}
 3. If $\varepsilon_r(V, F_{\mu_k}) = \delta \pm 10\%$ STOP else

$$\mu_{k+1} = \varepsilon_r(V, F_{\mu_k}) \frac{\delta}{\mu_k}$$

and go to 1.

In fact, δ is not exactly the noise level, it is the estimation of the statistical error in the spectrum values. The number δ is computed when the estimation of the spectral matrix is done.

C.2.4 Stability Analysis

In Subsection 2.3 we have described how the regularization parameter μ in the algorithm is determined. In fact, μ is chosen such that the error in the data is equal to $\delta \pm 10\%$. We know that for all $\mu > 0$ the solution varies continuously as a function of the data, but a small perturbation of the spectral matrix could lead to a huge modification of the solution. To verify and quantify this, some statistical tests have to be done.

For this purpose, perturbed spectral matrix $V_N := V + N$ must be created, where V is the measured spectral matrix, N is the perturbation matrix ($N = bb^*$), with b a random Gaussian vector. Then the energy perturbation, C_1 , and the variation for each points of the perturbed solution, $C_\infty : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{R}^+$, can be computed from the perturbed solution F_N . C_1 and C_∞ are given by :

$$C_1 := \frac{\int_0^\pi \int_0^{2\pi} |F(\theta, \phi) - F_N(\theta, \phi)| \sin \theta \, d\phi \, d\theta}{\|V - V_N\|}, \quad (\text{C.7})$$

$$C_\infty(\theta, \phi) := \frac{|F(\theta, \phi) - F_N(\theta, \phi)|}{\|V - V_N\|}. \quad (\text{C.8})$$

The process is then repeated for a large number of random Gaussian vectors b . From the C_1 sample thus obtained, the expectation, the variance, and the quantity $V_1(\delta)$ that represents the relative power error of the perturbed solution are computed. $V_1(\delta)$ is defined as

$$V_1(\delta) := \frac{E(C_1)\sqrt{\delta}}{\int_0^\pi \int_0^{2\pi} |F(\theta, \phi)| \sin \theta \, d\phi \, d\theta}, \quad (\text{C.9})$$

and expressed in percent. In addition, from the expectation of the C_∞ sample, a map of the averaged variation of the perturbed WDF can be plotted on the sphere $[0, \pi] \times [0, 2\pi]$.

In *Lefevre and Delannoy* [26], the stability of the solution is expressed in terms of the ratio of the mean-square error in the solution to the mean-square value of the solution itself. Instead of this stability parameter, we have chosen to use C_1 . The main reason for this choice is that C_1 has the dimension of an energy, while the stability parameter in *Lefevre and Delannoy* [26] is related to the square of the energy, which is less convenient. The number C_1 and the function $C_\infty : [0, \pi] \times [0, 2\pi] \rightarrow \mathbb{R}^+$ allow the stability of the solution, namely the WDF, to be quantified from two opposite points of view. Unlike C_1 , which is a global estimator of the stability, the function C_∞ indeed measures the local deviation of the perturbed WDF compared with the non-perturbed solution. It is noteworthy that it becomes possible to highlight the place where the solution is the most unstable from the expectation of C_∞ , as we will see in Section C.4.2.

C.3 ERA Implementation on the FREJA Data

FREJA is a low-altitude magnetospheric research satellite. The F4 onboard wave experiment [20] measured waveforms of three orthogonal magnetic components and one electric component received by a spinning antenna. The F4 experiment has observed an intense emission of ELF hiss on April 8, 1995, between 1548 and 1558 UT. *Santolík and Parrot* [53] have made a plane-wave analysis of this event and a WDF analysis on six different points of the ELF hiss spectrogram.

For our WDF analysis we use the same data as *Santolík and Parrot* [53], that is, the six spectral matrices and the six error matrices corresponding to the six points in the spectrogram. Since FREJA measured three magnetic and one electric components, the spectral and the error matrices are 4×4 . These matrices are identified as vectors in \mathbb{C}^{16} and denoted by $v^k \in \mathbb{C}^{16}$ for the spectral matrices and by $e^k \in \mathbb{C}^{16}$ for the measurement-error matrices for points $k = 1, \dots, 6$. The measurement-error matrix is needed since we do not have the same quality of measurement on each of the antennae. The error matrix is computed using the statistical standard deviation of the spectral values.

We also use the same integration kernels as in *Santolík and Parrot* [53]. These kernels have been computed according to a assumed plasma model. They use the cold plasma approximation with the hypothesis of a collisionless, singly-ionized oxygen-hydrogen plasma with 5% fraction for the hydrogen. The composition of the plasma is rather important since the frequency of the waves is lower than $2f_{H^+}$, where f_{H^+} is the proton gyrofrequency. The right-hand polarization mode has been used to compute these kernels for frequencies above f_{H^+} because this mode is the only one expected in a cold plasma at these frequencies [59]. We denote by $q^k : S \rightarrow \mathbb{C}^{16}$ the computed integration kernels corresponding to points $k = 1, \dots, 6$.

To take into account the measurement-error matrix, the data and the integration kernels are normalized. More precisely, for all $i = 1, \dots, 16$, we let

$$V^k = \begin{pmatrix} v_i^k \\ e_i^k \end{pmatrix} \in \mathbb{C}^{16} \text{ and } \bar{Q}_k = \begin{pmatrix} q_i^k \\ e_i^k \end{pmatrix}$$

be the normalized spectral matrix and the normalized integration kernel. The goal of this normalization is to have the same order of error in each of the spectral matrix components. As a consequence, the noise level of the normalized data becomes $\delta = 16$.

As the kernel values often vary over a large range, for good computational performance it is necessary to scale them by letting $Q^k = 10^{-\beta} \bar{Q}^k$ such that $1 \leq \max |Q_i^k| \leq 50$ for all $i = 1, \dots, 16$. The β values for each of the six points are given in Table C.1.

The FREJA F4 data is first transformed into the same coordinate system as used in *Santolík and Parrot* [53]; the z axis is parallel to the earth magnetic field, and the x axis points toward the direction of decreasing magnetic latitude in the magnetic meridian plane (see Figure C.1). The k wave-normal direction is determined by the angles θ (polar) and ϕ (azimuthal).

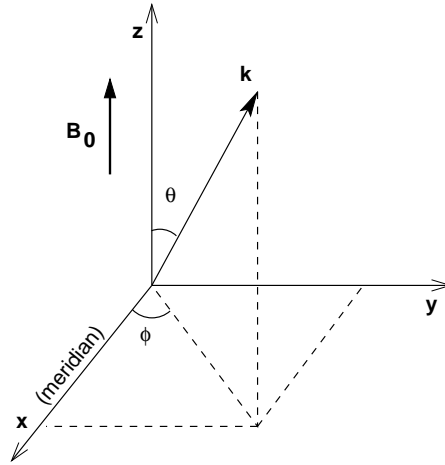


FIG. C.1 – The polar θ and azimuthal ϕ angles in the spacecraft coordinate system.

C.4 FREJA data analysis results

As mentioned before, the ERA algorithm has been applied to the same six time-frequency points as in *Santolík and Parrot* [53]. The same graphical representation is also used, for easy comparison. An example of this is given in Figure C.2, which is a spherical plot of the WDF value as a function of $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. In the northern hemisphere (left plot) θ varies from 0 to $\pi/2$, while in the southern one (right plot) it runs over $[\pi/2, \pi]$.

In our computations, the parameter ϵ used in the stopping criterion and the first step of the algorithm are set at $1e - 7$ and 0. All the computation have been made with the scientific software MATLAB[®] on an Apple ibook G3 700MHz. The computation times are variable depending of the integration kernel, the data and the algorithm parameters. Generally, computations are quite fast and the solutions are obtained in a few seconds. For example, the stability computations described in the previous section took approximately one hour for 1000 computed solutions.

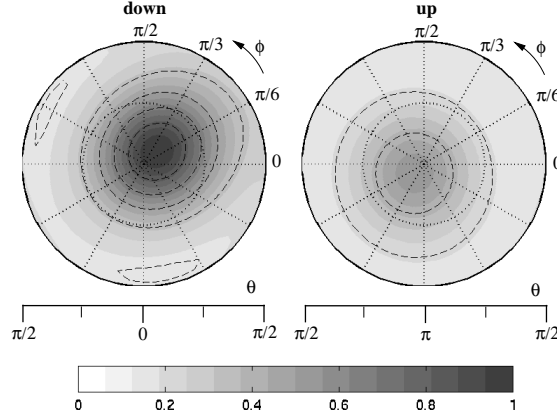


FIG. C.2 – Example of WDF graphical representation. θ is the polar angle and ϕ the azimuthal angle. The left hemisphere is for downgoing waves and the other for upgoing ones. Contours are shown with the dashed line, which correspond to the values of $\{1/6, 2/6, \dots, 5/6\}$. This example shows the obtained WDF for point 1 with $\mu = 20$.

C.4.1 Computed WDF

In Figure C.3, are displayed the WDFs for each of the six FREJA time-frequency points, computed using the following three methods (from top to bottom) : the entropy regularization algorithm (ERA), the Gaussian peaks method (GP) and the maximum entropy technique (ME). The level contours around the WDF peak are superimposed, and are shown as dashed lines in the ERA plots and as solid lines in the GP and ME plots.

On the whole, the ERA solutions turn out to be more spread out than both the ME and GP solutions, nevertheless there are no large discrepancies among all these solutions. If we now look at the various solutions in greater detail, there are actually some significant differences.

For the first point, the plane wave analysis made by *Santolík and Parrot* [53], returns a downgoing wave at $\theta \approx 30^\circ$, $\phi \approx 45^\circ$ with a high degree of polarization. As shown in the top-left panel of Figure C.3, the GP and ME solutions exhibit a sharp peak in almost the same position which is also in fairly good agreement with the position obtained by the plane-wave analysis. If we now look at the ERA solution, it becomes clear that the main peak occurs at the right place ($\theta \approx 14^\circ$, $\phi \approx 43^\circ$) in the northern hemisphere (downgoing waves). But there is also another peak in the opposite direction (southern hemisphere) although of much lower power.

For point 2, the plane wave analysis [53] gives a nearly field aligned downward propagation, with a low degree of polarization. The ERA predicts a very broad peak in the downgoing direction, at $\theta \approx 33^\circ$, $\phi \approx 63^\circ$. This peak fits well the primary peak of the other methods. However, in the GP and ME solutions, a secondary peak, well defined in the GP and quite questionable in the ME solution, is observed in the southern hemisphere. In the ERA, no such peak is predicted, but at most an increase of the energy occurs almost at the same

place ($90^\circ \leq \theta \leq 140^\circ$, $30^\circ \leq \phi \leq 120^\circ$). Let us note that the ERA solution reflects the low degree of polarization found in the plane-wave analysis [53], which means that the associated electromagnetic wave is supposed to propagate in many directions.

Point 3 turns out to be quite similar to point 1, and the results obtained by the three methods confirm this similarity. Again, a secondary peak is only seen in the ERA solution in the opposite direction of the main peak. Point 4 looks like point 2, where the downgoing and upgoing waves play symmetrical roles. In other words, the main peak is observed in the southern hemisphere for point 4 while it is in the northern hemisphere for point 2.

For points 5 and 6, a main peak is predicted by the three methods ($\theta \approx 60^\circ$ and $\phi \approx 201^\circ$ for point 5, and $\theta \approx 108^\circ$ and $\phi \approx 200^\circ$ for point 6, from the ERA solutions). There are no major discrepancies among the results obtained from the ERA, GP and ME techniques.

C.4.2 Stability Analysis

The results of stability test described in Section C.2.4 are as follows. The 1000 computations of perturbed solutions for all the six spectral matrices corresponding to the six points have been done in order to get a significant sample without excessive computation time. It took about one hour of computation for each point. In order to have perturbations of the same order as the noise level, which is about $\delta = 16$ according to the data normalization (see Section C.3), the variance of the Gaussian vector b has been fixed such that $E(\|V - V_N\|) = E(\|bb^*\|) \approx 10$. Nevertheless, as the variance is quite large, perturbations of the order of 60 instead of 10 are often observed.

For all of the six points, an averaged energy perturbation $V_1(\delta)$ of about 21% is observed (see Table C.1), which is a good result. The least energy perturbation is observed for point 2 whereas the most intense energy perturbation arises for point 1. The plots of the expectation of C_∞ for each of the six points are given in Figure C.4. In this Figure, the WDF and the expectation of C_∞ , called ERROR, are plotted using the usual spherical representation. In addition, the scale factor ζ between the WDF and the ERROR is given in each of the six panels corresponding to the six analyzed points. The ζ values are also given in Table C.1.

As seen in Figure C.4 as well as in Table C.1, the local stability of the solutions is quite acceptable (ζ less than 7%) and the ERROR increases with the WDF value. As shown in Table C.1, the highest value of $V_1(\delta)$ is the one for point 1. As 26% is not a high value, the solution remains stable. This is confirmed by choosing a higher value of μ and looking at the solution. Figure C.2 shows the WDF for point 1 for a doubled value of μ . The peak is again at the same place and, as expected, the solution does not differ significantly from the one in Figure C.3 (top-left panel).

C.5 Discussion and Conclusions

In a previous paper *Prot et al.* [42], a new inversion procedure was proposed to determine the wave energy density of electromagnetic fields. The wave distribution function, WDF, a concept introduced thirty years ago in *Storey and Lefevre* [61] was chosen due to its ability to handle continuum plane-wave distributions. The WDF determination from field measurements implies that a highly indeterminate and ill-posed problem has to be solved.

Several inversion techniques have been developed in the past, but very few of them may be applied generally. Often, some parameters have to be chosen quite arbitrarily and some a priori information has to be introduced. The entropy regularization algorithm, ERA, was therefore developed to overcome some of these limitations. In this paper, the ERA has been applied to the FREJA F4 experiment electromagnetic field measurements and the results compared with those produced by two other techniques, the maximum entropy, ME, and the Gaussian peak, GP, already published in *Santolík and Parrot* [53].

Overall, there is no contradiction between the ERA solutions and those computed in *Santolík and Parrot* [53]. Globally, ERA solutions are more spread out which is expected for entropy maximization procedures. This is indeed compatible with the idea of a solution with minimal “information”. Let us also note that a solution with high variations (a very sharp peak) and thus large information is in contradiction with a problem which is, originally, highly indeterminate.

As shown in Figure C.3, for points 2 and 4, the ERA solutions exhibit a fairly constant energy in all directions, while with the two other methods, two better defined and stronger peaks are found. According to *Santolík and Parrot* [53] the polarization degree is rather low, which is in favour of the ERA solution. Moreover, in the GP method, the a priori existence of two peaks has been imposed, which implies a lower entropy of the solution than the one for the ERA. This is also the case for the ME method. Although the existence of two peaks is not imposed in the ME, the entropy increase is due to the fact that the entropy minimization is done with a higher level of constraints than for the ERA.

From this we conclude that the ME and GP solutions, which appear to be more accurate, cannot be considered as better solutions than the ERA without any extra information about the shape of the wave energy-density. This is because the ERA maximizes the entropy and therefore minimizes the information content in the solution. We should also note that the secondary peak that is observed in the GP and, to a lesser degree, in the ME methods is also detected by the ERA as an increase of the energy density. The total energy of this wave is of the same order as the one given by the ME and GP, the only difference with the ERA solution is that its energy density appears as more concentrated.

At this stage of the discussion, it is worth noting that in regularization theories, a family of operators with a free parameter, $\mu > 0$, is used. For a small μ the solution produced has a small error of fit but is highly unstable. Conversely with a larger μ , the solution will deviate from the original data, but its stability will be ensured. In other words, a compromise has to be found between the error of fit and stability, by selecting a suitable value of μ . This is done automatically in the ERA, according to the Morozov’s discrepancy principle, which merely requires that a solution whose error of fit is smaller than the noise level has to be rejected. As a consequence, it is always possible to choose μ values lower than the one given by the Morozov’s principle in order to get solutions quite similar to the ones given by the GP and the ME techniques, knowing that their stability will be worse.

Although the stability of the solutions obtained to fit the FREJA data is controlled by the ERA algorithm, a statistical study has been performed to quantify it as a whole and point by point. To do this, the spectral matrices have been perturbed and the resulting energy deviation and the variation for each points of the WDF have been estimated. On the average, and for the six events considered in this study, the energy deviation turns out to be near

20%, while the variation for each points of the WDF is less than 7%.

As a final conclusion, we have shown that the entropy regularization algorithm, ERA, is well suited to analyze wave characteristics in space plasmas. It allows the wave distribution function, WDF, to be computed and, in particular, the wave propagation direction to be determined. Unlike other methods, such as the maximum entropy, ME, and Gaussian peaks, GP, the ERA does not need any assumptions, for example, about the shape of the WDF and/or the significance of the integration kernels. Therefore, a good strategy in analyzing a large amount of data, is first to apply the ERA to detect the wave events we are interested in, and then to use other techniques, for example the GP and/or ME, to refine the diagnostics of these waves. Alternatively, the ERA can also be used in the diagnostics refinement step by choosing the appropriate value of the regularization parameter μ .

Acknowledgments. The authors wish to thank M. Bergounioux (CNRS-MAPMO, Orléans, FRANCE), J.L. Pinçon and F. Lefeuvre (CNRS-LPCE, Orléans, FRANCE) for their very helpful suggestions.

Point N°	1	2	3	4	5	6
Time UT	15 :49 :15	15 :52 :43	15 :52 :43	15 :54 :59	15 :54 :59	15 :55 :53
Frequency[Hz]	688	608	736	576	800	624
WDF						
β	13	11	11	12	15	12
Energy	2.98	2.83	0.58	9.90	3.62e-1	9.98e-1
Min	4.61e-2	7.19e-2	3.56e-3	3.01e-1	1.72e-3	6.30e-3
Max	1.38	0.61	3.64e-1	2.35	2.42e-1	8.61e-1
Statistics						
$E(C_1)$	1.97e-1	1.14e-1	2.93e-2	5.01e-1	1.97e-2	5.01e-2
$Var(C_1)$	8.76e-4	3.54e-4	2.17e-5	8.92e-3	1.25e-5	1.02e-4
V_1^δ	26.4%	16.1%	20.2%	20.2%	21.7%	20.08%
ζ	0.0694	0.0695	0.0353	0.0372	0.0332	0.026

TAB. C.1 – Results obtained for each point. $10^{-\beta}$ is the scale factor used for the integration kernel. The Power, Min and Max are respectively the energy, the minimum and the maximum of the WDF. For each point we have the results of the stability analysis : the expectation and the variance of C_1 , the relative energy error V_1^δ and the factor ζ between the WDF and the expectation of C_∞

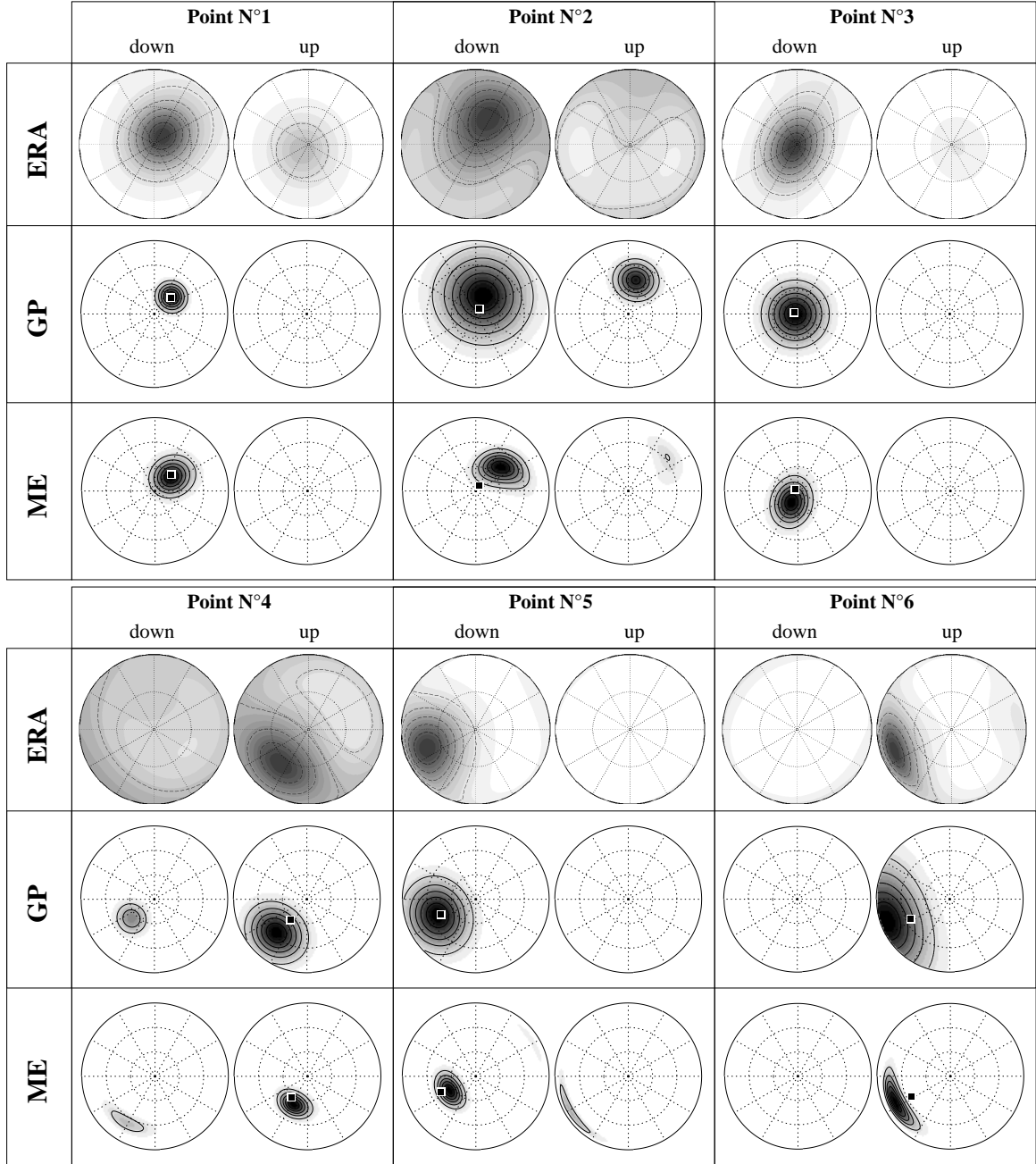


FIG. C.3 – Results obtained by ERA (top) for the six spectral matrices. The graphical representation is similar to that in [50]. The wave-energy density is coded as in Figure B.1. The WDF computed by the Maximum Entropy (bottom) and Gaussian Peaks (middle) methods for the points 1, 2 and 4. Borrowed from [53].

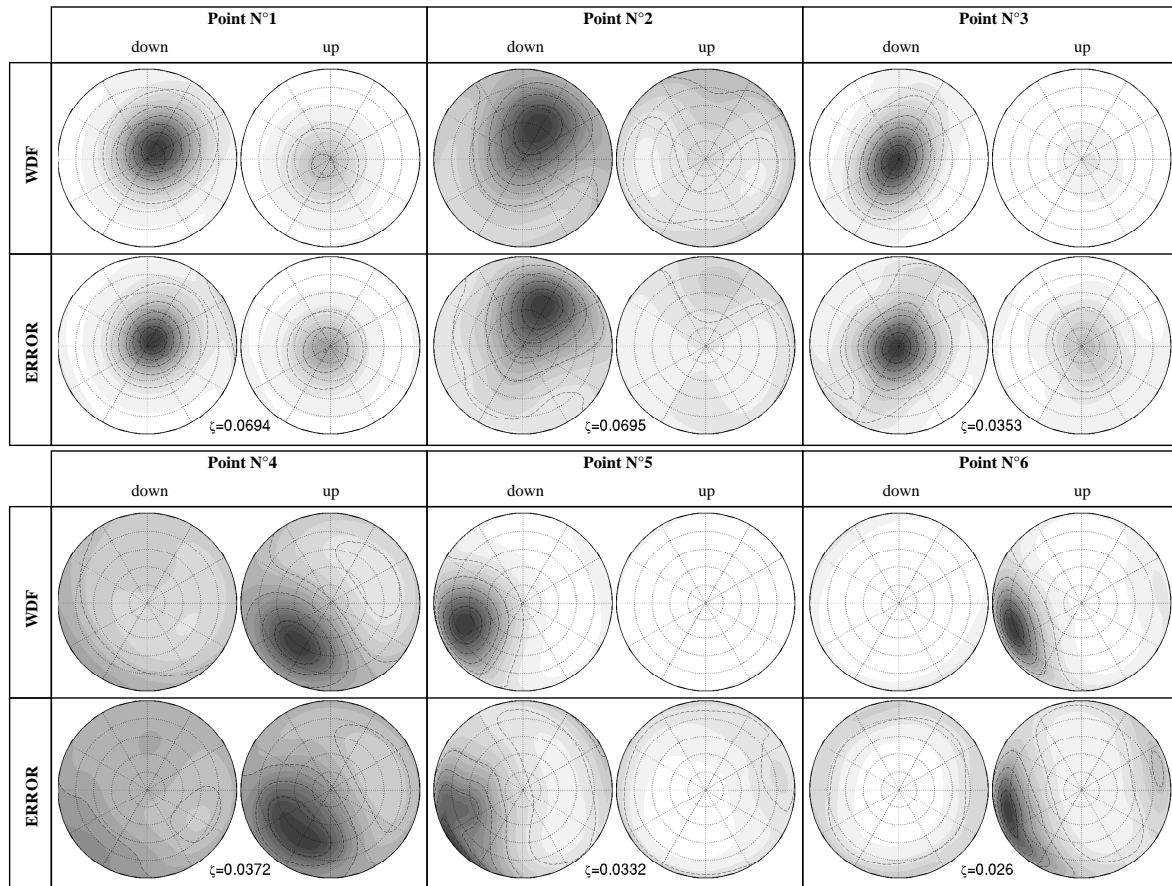


FIG. C.4 – Results of the stability analysis. For each point the plots show the WDF and the local error, which is the expectation of C_∞ . The scale factor between the WDF and the ERROR is given by the number ζ .

Annexe D

Programmes

Les algorithmes de régularisation entropique développés dans la Section 3.4, ont été programmés en langage MATLAB. Nous avons fait un programme par algorithme, ces programmes ont été appelés ERA I et ERA II. La description complète de ces programmes sera donnée dans la section suivante. Ce sont ces programmes que nous avons utilisé pour calculer la fonction de distribution des ondes dans le cas des données synthétiques et des données FREJA au Chapitre 5. Ces programmes ont été développés pour le calcul de la FDO, la procédure d'inversion est donc prévue pour déterminer une fonction définie sur la sphère unité de \mathbb{R}^3 .

Dans la Section D.1 nous décrirons les programmes qui ont permis d'adapter les données de FREJA aux programmes ERA I et ERA II. Les données de l'expérience sont stockées dans des fichiers textes formatés pour être utilisé dans le programme de *Santolík* [50]. Nous avons donc fait des procédures MATLAB afin de lire les données dans ces fichiers et de les mettre en forme pour les utiliser dans les programmes ERA I et ERA II.

D.1 Programme ERA I et ERA II

Les programmes ERA I et ERA II sont quasiment identiques, la Figure D.2 montre comment sont organisé les fichiers du programmes. Ces programmes ont été déposés sur le serveur CIEL :

`http :/ciel.ccsd.cnrs.fr/`

D.1.1 Description des fichiers

La Figure D.2 montre le diagramme de fonctionnement des programmes ERA I et ERA II. La procédure à suivre pour faire fonctionner le programme est la suivante :

- Le premier programme à lancer est `LEG_RUN.m`, il fait appel à `LEG_GAUSS.m` qui est une fonction servant pour calculer les poids et les points d'intégration de Legendre Gauss [19]. Ces points sont calculés pour quatre niveaux de discrétisation. Ceux qui

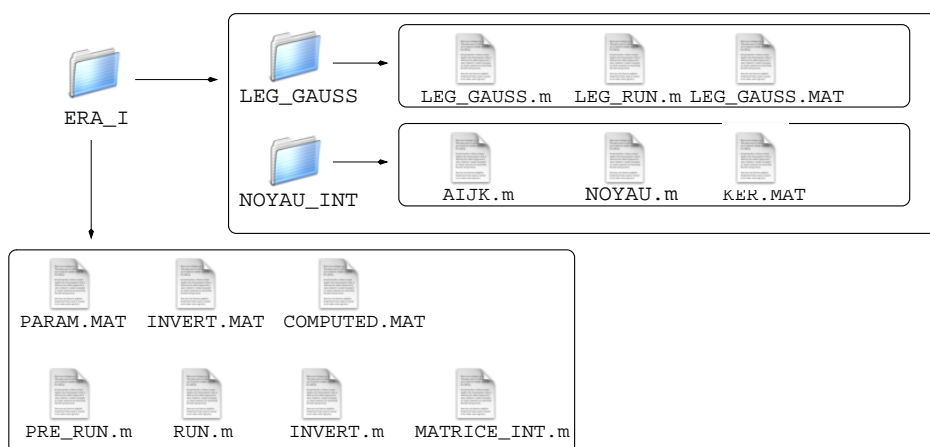


FIG. D.1 – Contenus du dossier ERA_I, avec la disposition des différents fichiers du programme.

ont moins de points serviront pendant les calculs alors que les grilles plus fines seront utilisées pour les représentations graphiques. Ces données sont ensuite sauvegardées dans le fichier LEG_GAUSS.MAT.

- Le second programme à exécuter est NOYAU.m, il utilise les points d'intégrations du fichier LEG_GAUSS.MAT Pour calculer les noyaux d'intégrations du vide sur les quatre niveaux de discrétisation. La fonction AIJK.m calcule ces noyaux d'intégration et les retourne au programme qui sauve ces données dans le fichier KERN.MAT.
- Ensuite le programme PRE_RUN.m collecte les données des fichiers LEG_GAUSS.MAT et KERN.MAT et les met en forme pour la procédure d'inversion pour le niveau de discrétisation choisi. Toutes ces données sont enregistrées dans le fichier INVERT.MAT.
- Enfin le programme RUN.m lit les données préparées, contenant les noyaux d'intégration, dans le fichier INVERT.MAT. La matrice spectrale et les paramètres d'initialisations de l'algorithme sont lus dans le fichier PARAM.MAT. La fonction MATRIX_INT fournit la matrice utilisée pour le calcul des intégrales dans le programme. La fonction INVERT.m n'est rien d'autre que l'algorithme d'inversion. La solution du problème inverse est sauvée dans le fichier COMPUTED.MAT.

D'une manière générale, dans tous les programmes, les fonctions de deux variables, comme la FDO où les noyaux d'intégration, sont mis en forme de vecteurs plutôt que comme une matrice. Pour cela on met bout à bout les lignes de la matrice. Il n'y a qu'à la fin du programme RUN.m que la FDO est remise sous la forme d'une matrice pour tracer les graphiques.

Dans le fichier LEG_GAUSS.MAT les poids et les abscisses d'intégrations sont stockés respectivement dans les variables X_{ij} et W_{ij} où $i = 1, \dots, 4$ correspond à la finesse de la grille de discrétisation, $j = 1$ pour l'intégration de l'angle polaire $\theta \in [0, \pi]$ et $j = 2$ pour l'intégration de l'angle azimutal $\phi \in [0, 2\pi]$.

La matrice spectrale V est stockée dans le fichier PARAM.MAT ainsi que les paramètres d'initialisations de l'algorithme : `epsil` pour la variable ε , `mu` pour le paramètre de régularisation μ , `factk` pour τ^{-1} , `DELT` pour le niveau de bruit $\delta > 0$ et `NUM` correspond au nombre de

“meilleurs” paramètres τ testés.

Le fichier `INVERT.MAT` contient les noyaux d’intégration pour deux niveaux de discrétisation : `Ker` et `Ker2`. Ce sont des matrices de taille $D_\phi D_\theta \times n$ où D_ϕ et D_θ sont le nombre de points de discrétisation utilisés dans chacune des directions et n est juste la dimension du vecteur V .

Pour utiliser le programme avec un modèle autre que les noyaux du vide, il faut simplement fournir le fichier `KERN.MAT` contenant les noyaux adaptés aux points de discrétisation de Legendre-Gauss.

D.1.2 Listings des programmes

```

%%%%%%%%% LEG_GAUSS.m %%%%%%%%%%%%%%
function [w,x]=LEG_GAUSS(n,a,b)
x1=a;x2=b;
x=[];
w=[];
EPSIL=10^(-11);
m=(n+1)/2;
xm=(x2+x1)/2;
x1=(x2-x1)/2;
for ii=1:floor(m)
z=cos(pi*(ii-0.25)/(n+0.5));
z1=z+1;
while abs(z-z1)>EPSIL
p1=1;
p2=0;
for jj=1:n
p3=p2;
p2=p1;
p1=((2*jj-1)*z*p2-(jj-1)*p3)/jj;
end;
pp=n*(z*p1-p2)/(z*z-1);
z1=z;
z=z1-p1/pp;
end;
x(ii)=xm-x1*z;
x(n+1-ii)=xm+x1*z;
w(ii)=2*x1/((1-z*z)*pp*pp);
w(n+1-ii)=w(ii);
end

%%%%%%%%% LEG_RUN.m %%%%%%%%%%%%%%
%% pas d'integration 10 20
display('* Abscisses et poids pour N1=10, N2=20 *');

```

PROGRAMME ERA I & II

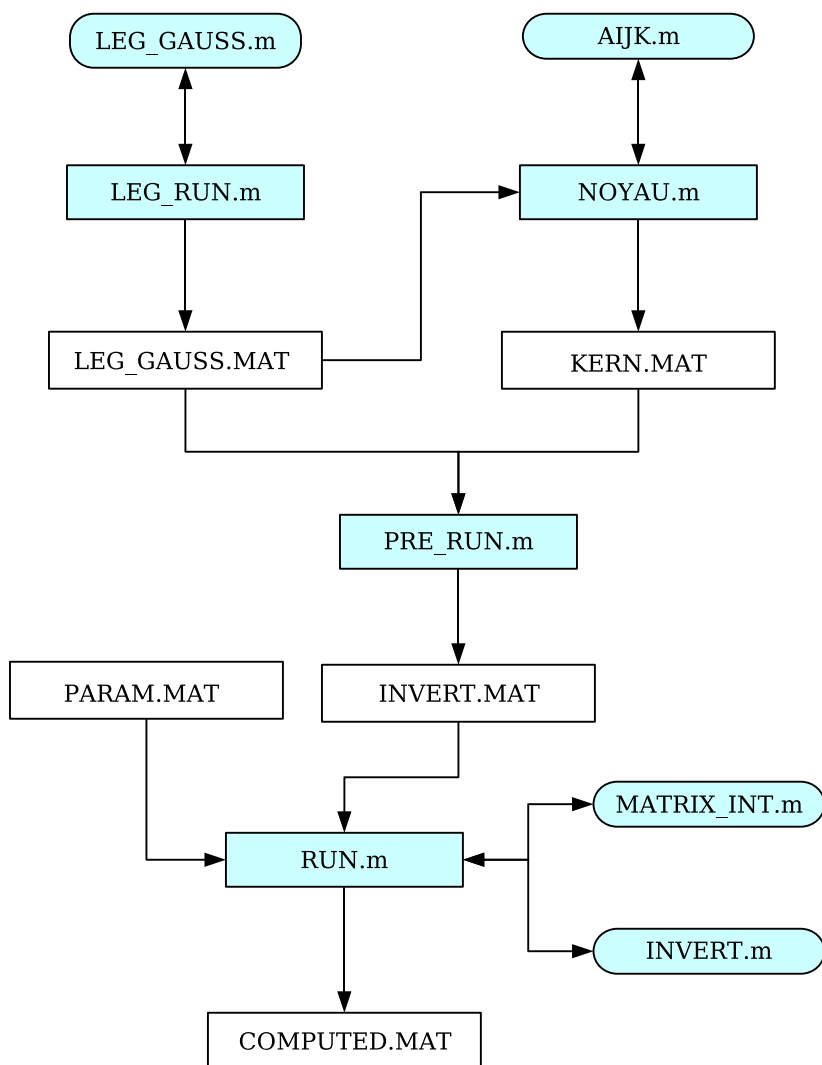


FIG. D.2 – Diagramme du programme ERA

```

[w,x]=LEG_GAUSS(10,0,pi);
W11=w;
X11=x;
[w,x]=LEG_GAUSS(20,0,2*pi);
W12=w;
X12=x;
%% pas d'integration 20 40
display('* Abscisses et poids pour N1=20, N2=40 *');
[w,x]=LEG_GAUSS(20,0,pi);
W21=w;
X21=x;
[w,x]=LEG_GAUSS(40,0,2*pi);
W22=w;
X22=x;
%% pas d'integration 40 80
display('* Abscisses et poids pour N1=40, N2=80 *');
[w,x]=LEG_GAUSS(40,0,pi);
W31=w;
X31=x;
[w,x]=LEG_GAUSS(80,0,2*pi);
W32=w;
X32=x;
%% pas d'integration 64 128
display('* Abscisses et poids pour N1=64, N2=128 *');
[w,x]=LEG_GAUSS(64,0,pi);
W41=w;
X41=x;
[w,x]=LEG_GAUSS(128,0,2*pi);
W42=w;
X42=x;
display('* termine *');
save 'LEG_GAUSS.mat' W11 W12 W21 W22 W31 W32 W41
...W42 X11 X12 X21 X22 X31 X32 X41 X42;
display('* Sauvergarde dans le fichier LEG_GAUSS.mat *')

%%%%%% aijk.m %%%%%%%%%%
function [V]=aijk(theta,phi,beta,modp)
e(1)=modp*exp(i*beta)*sin(phi)+cos(theta)*cos(phi);
e(2)=-modp*exp(i*beta)*cos(phi)+cos(theta)*sin(phi);
e(3)=-sin(theta);
e(4)=-modp*exp(i*beta)*cos(phi)*cos(theta)+sin(phi);
e(5)=-modp*exp(i*beta)*cos(theta)*sin(phi)+cos(phi);
e(6)=modp*exp(i*beta)*sin(theta);
for ii=1:6

```

```

for jj=1:6
V((ii-1)*6+jj)=e(ii)*conj(e(jj));
end;
end;

%%%%%%%% NOYAU.m %%%%%%%%%%%%%%%
display('* Calcul des noyaux *');
beta=1;modp=1;
load('../LEGENDRE_GAUSS/LEG_GAUSS.mat');
display('* Noyau Ker1 (N1=10,N2=20) *');
Ker=[];
for ii=1:10
for jj=1:20
theta=X11(ii);
phi=X12(jj);
[V]=aijk(theta,phi,beta,modp);
Ker=[Ker,V.'];
end;
end;
Ker1=Ker;
display('* Noyau Ker2 (N1=20,N2=40) *');
Ker=[];
for ii=1:20
for jj=1:40
theta=X21(ii);
phi=X22(jj);
[V]=aijk(theta,phi,beta,modp);
Ker=[Ker,V.'];
end;
end;
Ker2=Ker;
display('* Noyau Ker3 (N1=40,N2=80) *');
Ker=[];
for ii=1:40
for jj=1:80
theta=X31(ii);
phi=X32(jj);
[V]=aijk(theta,phi,beta,modp);
Ker=[Ker,V.'];
end;
end;
Ker3=Ker;
display('* Noyau Ker4 (N1=64,N2=128) *');
Ker=[];

```

```

for ii=1:64
for jj=1:128
theta=X41(ii);
phi=X42(jj);
[V]=aijk(theta,phi,beta,modp);
Ker=[Ker,V.'];
end;
end;
Ker4=Ker;
display('* Fin des calculs *');
save 'KERN.mat' Ker1 Ker2 Ker3 Ker4;
display('* Donnees sauvegardees dans le fichier KERN.mat *')

```

```

%%%%%%%%%% PRE_RUN.m %%%%%%%%%%%
load('NOYAU_INT/KERN.mat');
disp('* Lecture des Noyaux *');
Ker=Ker2;
Ker2=Ker4;
load('LEGENDRE_GAUSS/LEG_GAUSS.mat');
W2=W21;
W4=W22;
L=length(W2);
disp('* Modifications des noyaux *')
for ii=1:L
W2(ii)=W2(ii)*sin(X21(ii));
end;
disp('* Fin des calculs *');
save 'INVERT.mat' W2 W4 W41 W42 Ker Ker2;
disp('* Donnees sauvees dans INVERT.mat *');

```

```

%%%%%%%%%% MATRIX_INT.m %%%%%%%%%%%
function [Matr]=MATRIX_INT(W2,W4)
N1=length(W2);
N2=length(W4);
Matr=[];
for i=1:N1
Lig=[zeros(N2*(i-1),1);W4';zeros(N2*(N1-i),1)];
Matr=[Matr,Lig];
end;
Matr=sparse(Matr);

```

```

%%%%%%%%%% INVERT.m %%%%%%%%%%%
function [lambda,nbiter,NNN]=INVERT(lambda
...,X,mu,Matr,factk,epsil,W2,Ker)

```

```

nbiter=0;
komp=1000000;
nnn=epsil+1000000;
nlast=nnn+1;
NNN=[];
while (epsil<nnn)&(nlast>nnn)
in=2/mu*real(Ker'*lambda)-1;
wdf=exp(in);
s=(wdf'*Matr)*W2';
BB=ones(1,length(X));
II=kron(wdf,BB);
WWW=Ker.*(II');
s=(WWW*Matr)*W2';
lambdak=lambda+(X.'-s-lambda)/factk;
nlast=nnn;
nnn=norm(lambdak-lambda);
if nnn<(epsil*komp);
disp(nnn);
komp=komp/10;
end;
NNN=[NNN;nnn];
lambda=lambdak;
nbiter=nbiter+1;
end;

%%%%%%%%% RUN.m %%%%%%%%%%%%%%%
disp('* ERA *');
disp('* Chargement des données *');
load('INVERT.mat');
load('LEGENDRE_GAUSS/LEG_GAUSS.mat');
disp('* Chargement des paramètres *');
load('PARAM.mat');
lambda=zeros(length(V),1);
disp('* Calcul de la solution *');
[Matr]=MATRIX_INT(W2,W4);
mu2=mu;
while DELT*0.9>norm(lambda)^2 | norm(lambda)^2>(DELT*1.1)
mu=mu2;
zet=epsil+1;
factk=factk/2;
while zet>epsil
factk=factk*2;
[lambda,nbiter,NNN]=INVERT(lambda,

```

```

...V,mu,Matr,factk,epsil,W2,Ker);
zet=NNN(length(NNN));
end;
disp('* linear search of tau *')
ko=0;
while ko<NUM
ko=ko+1;
fnum=factk/NUM;
[lambda,nbiter,NNN]=INVERT(lambda,V,mu,Matr,
...factk+(NUM-ko)*fnum,epsil,W2,Ker);
zet=NNN(length(NNN));
vectorfactk(ko)=factk+(NUM-ko)*fnum;
if zet<epsil
ko=NUM;
end;
end;
if DELT<norm(lambda)^2
mu2=DELT*mu/norm(lambda)^2;
else
mu2=(DELT*1.1)*mu/norm(lambda)^2;
end;
AA1=abs(DELT*0.9-norm(lambda)^2);
AA2=abs(DELT*1.1-norm(lambda)^2);
AA=max(AA1,AA2);
disp(['* Décalage , ',num2str(floor(AA/DELT*1000)/10),' % *']);
end;
disp('* Analyse de la solution *');
wdf=[];
in=2/mu*real(Ker'*lambda)-1;
wdf=exp(in);
s=(wdf'*Matr)*W2';
II=[];
for compj=1:length(V)
II=[II,wdf];
end;
WWW=Ker.*(II');
s=(WWW*Matr)*W2';
ERR=norm(V.'-s)^2;
alpha=(wdf'*Matr)*W2';
flnf=wdf.*(log(wdf));
E1=(flnf'*Matr)*W2';
wdf2=wdf/alpha;
flnf=wdf2.*(log(wdf2));
E2=(flnf'*Matr)*W2';

```



```

Cout=ERR+mu*E1;
%%
in=2/mu*real(Ker2'*lambda)-1;
wdf=exp(in);
wvf=[];
for ii=1:64
for jj=1:128
wvf(ii,jj)=wdf(128*(ii-1)+jj);
end;
end;
%figure;
%set(gcf,'Renderer','Zbuffer');
%pcolor(X42,X41,wvf);
%shading interp;
%figure;
%contour(X42,X41,wvf,6);
%figure;
%set(gcf,'Renderer','Zbuffer');
%surf(X42,X41,wvf);
%light;
%shading interp;
%print -depsc '3d.eps';
figure;
set(gcf,'Renderer','Zbuffer');
XX=X41(1:32);
[th,r] = meshgrid(X42,sin(XX));
Wvf=wvf;
[X,Y]=pol2cart(th,r);
W1=[];
W2=[];
for ii=1:32
for jj=1:128
W1(ii,jj)=Wvf(ii,jj);
W2(ii,jj)=Wvf(65-ii,jj);
end;
end;
subplot(121);
numm=10;
h=polar([0 2*pi], [0 1]);
delete(h)
hold on
pcolor(X,Y,W1);
TTAB=caxis;
colorbar('horiz');

```

```

shading interp;
subplot(122);
h = polar([0 2*pi], [0 1]);
delete(h);
hold on;
pcolor(X,Y,W2);
colorbar('horiz');
shading interp;
disp('* Termine *');
save 'COMPUTED.mat' V lambda NNN alpha
...E1 E2 ERR Cout nbiter mu epsilon factk;
disp('* Donnees sauve dans le fichier COMPUTED.mat *');

```

D.2 Conversion des données de FREJA

Dans sa thèse O. Santolík a fait un programme pour résoudre le problème de la détermination de la FDO. Pour fonctionner son programme a besoin que les données sur l'évènement étudié soient mises dans des fichiers texte que le programme va lire pendant l'exécution. Nous avons fait un programme permettant de lire les données dans ces fichiers textes afin de les convertir en tableau MATLAB et de générer les fichiers KERN.MAT et PARAM.MAT utilisés par nos programmes.

Les fichiers utilisés par le programme de Santolík sont les suivants :

- Les fichiers `DispF.#` qui contiennent la solution de la relation de dispersion. C'est à partir de ce fichier que les noyaux d'intégration du milieu considéré sont calculés.
- Les informations sur l'évènement étudié (fréquence de l'onde, caractéristiques du plasma) sont stockées dans les fichiers `EventF.#`.
- La direction, la longueur, la fonction de transfert de chacune des antennes sont données dans le fichier `DEVICEF.#`. Le fait de connaître la direction des antennes dans le repère permet de déterminer les noyaux d'intégration du modèle de façon appropriés.
- Enfin, les fichiers `sm.#` contiennent la matrice spectrale estimée ainsi que la matrice des erreurs.

La figure D.3 illustre le fonctionnement de nos procédures permettant de traduire les données de FREJA en données utilisables par les programmes `ERA_1` et `ERA_2`.

- Le programme `CONVERTSM.m` va lire la matrice spectrale et la matrice des erreurs dans le fichier `sm.#`, il les sauve ensuite, sous forme de variable MATLAB, dans le fichier `SPECTRAL.MAT`.
- Ensuite le programme `SPECTRAL.m` normalise la matrice spectrale par les éléments de la matrice des erreurs et la sauve sous la forme d'un vecteur dans le fichier `PARAM.MAT`
- La solution de la relation de dispersion est lue et sauvé dans le fichier `DISP.MAT` au format MATLAB par le programme `CONVERT_DISP.m`.
- Le calcul des noyaux d'intégration se fait en deux étapes. Premièrement le programme `COMPUTE_KERN.m` calcule les noyaux d'intégration à partir de la solution de la relation de dispersion et des différentes caractéristiques des antennes. Ces noyaux sont sauvés

dans le fichier PRE_KERN.MAT. Deuxièmement, le programme COMPUTE_KERN2.m effectue la normalisation des noyaux relativement à la matrice des erreurs de mesures et sauve les noyaux obtenus dans le fichier KERN.mat. Ces noyaux sont discrétisés suivant les abscisses données par le fichier LEG_GAUSS.MAT, pour cela une interpolation est effectuée.

D.2.1 Listing des programmes

```

%%%%%%%%% CONVERTSM.m %%%%%%%%%%
N=4;
filename='sm.sm.2';
A=textread(filename,'%f','commentstyle','matlab');
MAT=zeros(N,N);
p=1;
for c=1:N
    for l=c:N
        if c==1
            MAT(l,l)=A(p);
            p=p+1;
        else
            MAT(l,c)=A(p)-i*A(p+1);
            p=p+2;
        end;
    end;
end;
de=diag(MAT);
de=diag(de);
MAT=MAT+MAT'-de;
for c=1:N
    for l=c:N
        if c==1
            ERRMAT(l,l)=A(p);
            p=p+1;
        else
            ERRMAT(l,c)=A(p)-i*A(p+1);
            p=p+2;
        end;
    end;
end;
de=diag(ERRMAT);
de=diag(de);
ERRMAT=ERRMAT+ERRMAT'-de;
save 'SPECTRAL.mat' MAT ERRMAT;
save '../NOYaux/SPECTRAL.mat' MAT ERRMAT;

```

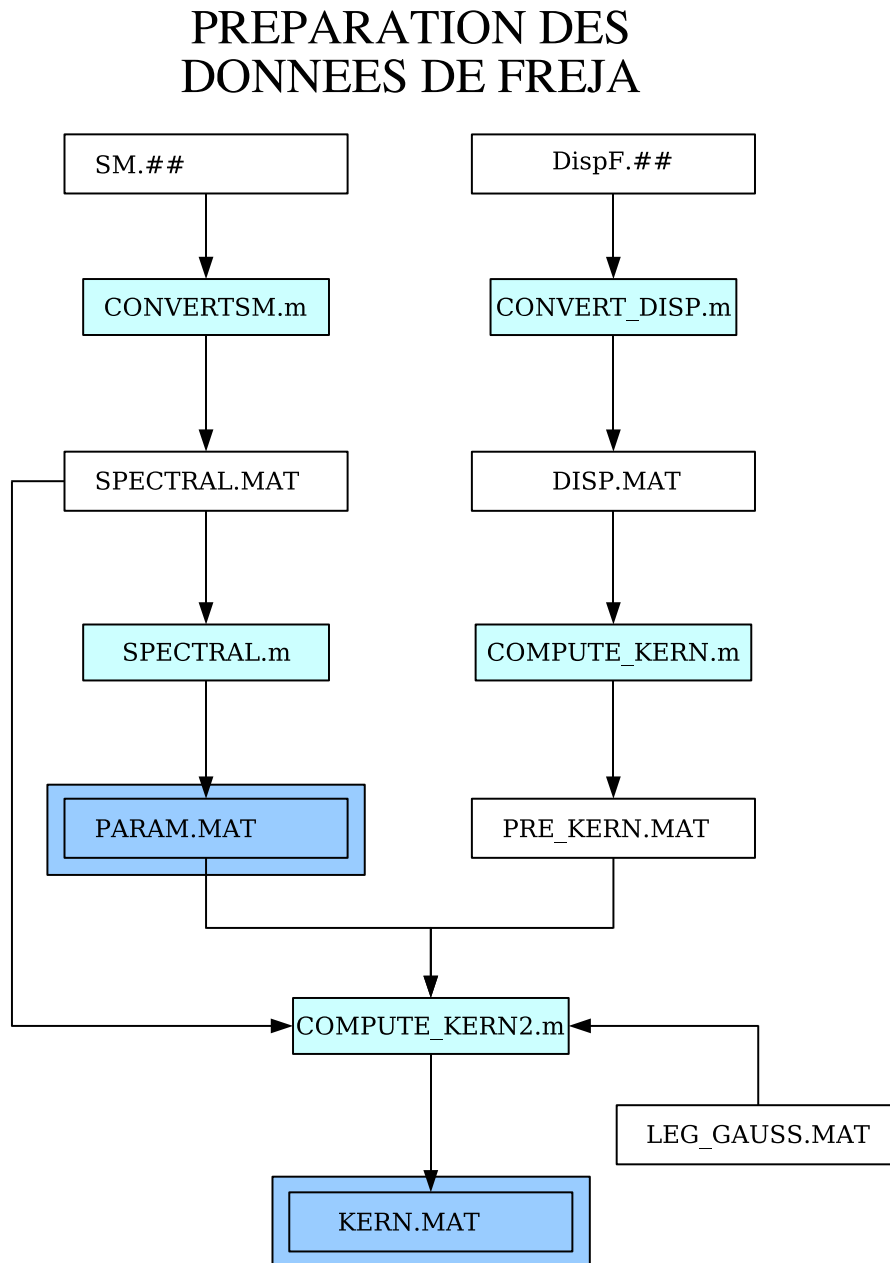


FIG. D.3 – Conversion des données FREJA

```

%%%%%%%%% SPECTRAL.m %%%%%%%%%%%%%%
load 'SPECTRAL.mat'
t=size(MAT);
V=[];
for ii=1:t(1)
    for jj=1:t(2)
        if ii==jj
            V(t(2)*(ii-1)+jj)=MAT(ii,jj)/ERRMAT(ii,jj);
        else
            V(t(2)*(ii-1)+jj)=real(MAT(ii,jj))/real(ERRMAT(ii,jj))
            ...+i*imag(MAT(ii,jj))/imag(ERRMAT(ii,jj));
        end;
    end;
end;
factk=20;
epsil=1E-6;
mu=5;
NUM=10;
DELT=16;
save 'PARAM.mat' V factk epsil mu NUM DELT;
save '../MIPE/PARAM.mat' V factk epsil mu NUM DELT;
save '../MIPE2/PARAM.mat' V factk epsil mu NUM DELT;

%%%%%%%%% CONVERT_DISP.m %%%%%%%%%%%%%%
filename='dispF.2';
A=textread(filename,'%f','commentstyle','matlab','headerlines',20);
p=1;
THETA=[];
PHI=[];
N=[];
THG=[];
VG=[];
EX=[];
EY=[];
EZ=[];
while p<length(A)
    THETA=[THETA;A(p)];
    p=p+1;
    PHI=[PHI;A(p)];
    p=p+1;
    N=[N;A(p)+i*A(p+1)];
    p=p+2;
end;

```

```

    VG=[VG;A(p)];
    p=p+1;
    THG=[THG;A(p)];
    p=p+1;
    EX=[EX;A(p)];
    p=p+1;
    EY=[EY;A(p)+i*A(p+1)];
    p=p+2;
    EZ=[EZ;A(p)+i*A(p+1)];
    p=p+2;
end;
save 'disp.mat' THETA PHI N THG VG EX EY EZ;
save '../NOYAUX/disp.mat'

%%%%%%%%%% COMPUTE_KERN.m %%%%%%%%%%
% charge le contenu du fichier dispF
load 'disp.mat'
% * paramètres *
% largeur de la bande (fichier sm.sm)
DWA=16;
% Matrice d'orientation 0
O=eye(3,3);
% Direction du champ magnétique
BO=[0;0;1];
% vitesse de l'observateur
v0=[0;0;0];
% Antenne (dans le même ordre que dans le fichier device)
ga1=[1;0;0];
ga2=[0;1;0];
ga3=[0;0;1];
ga4=[0.0104419999;-0.9402300119;-0.2571800053];
ga4=ga4/norm(ga4); %normalisation
% Fonction de transfert des antennes
ta1=1;
ta2=1;
ta3=1;
ta4=0.9644-i*0.1039;
% Matrice de passage
% Pour FREJA il n'y en a pas besoin car P(C-B) P(C-P) sont égale à Id !!!
% * Calcul des noyaux *
th=THETA/180*pi;
ph=0:pi/180:2*pi-pi/180;
ZZZ=zeros(length(EX),length(EX));
for tour=1:length(EX)

```

```

        ZZZ(tour,length(EX)-tour+1)=1;
    end;
    % rotation
    ph2=ph;
    ph=ph-(117.8/180)*pi;
    % Calcul des champs dans le systeme (p)
    EN1=EX*cos(ph)-EY*sin(ph);
    EN2=EX*sin(ph)+EY*cos(ph);
    EN3=EZ*ones(size(ph));
    BX=-(N.*EY.*cos(th));
    BY=N.*(EX.*cos(th)-EZ.*sin(th));
    BZ=N.*EY.*sin(th);
    BN1=BX*cos(ph)-BY*sin(ph);
    BN2=BX*sin(ph)+BY*cos(ph);
    BN3=BZ*ones(size(ph));
    EN1=[EN1;ZZZ*EN1];
    EN2=[EN2;ZZZ*EN2];
    EN3=[EN3;-ZZZ*EN3];
    BN1=[BN1;-ZZZ*BN1];
    BN2=[BN2;-ZZZ*BN2];
    BN3=[BN3;ZZZ*BN3];
    th1=[th;th+th(length(th))];
    th=th1;
    % Calcul des signaux élémentaire normalisé eta
    etaB3=ta4*(ga4(1)*EN1+ga4(2)*EN2+ga4(3)*EN3);
    etaE=BN1;
    etaB1=BN2;
    etaB2=BN3;
    % calcul des noyaux
    A=[];
    A(:, :, 1)=etaE.*conj(etaE)*DWA;
    A(:, :, 2)=etaE.*conj(etaB1)*DWA;
    A(:, :, 3)=etaE.*conj(etaB2)*DWA;
    A(:, :, 4)=etaE.*conj(etaB3)*DWA;
    A(:, :, 5)=conj(A(:, :, 2));
    A(:, :, 6)=etaB1.*conj(etaB1)*DWA;
    A(:, :, 7)=etaB1.*conj(etaB2)*DWA;
    A(:, :, 8)=etaB1.*conj(etaB3)*DWA;
    A(:, :, 9)=conj(A(:, :, 3));
    A(:, :, 10)=conj(A(:, :, 7));
    A(:, :, 11)=etaB2.*conj(etaB2)*DWA;
    A(:, :, 12)=etaB2.*conj(etaB3)*DWA;
    A(:, :, 13)=conj(A(:, :, 4));
    A(:, :, 14)=conj(A(:, :, 8));

```

```

A(:, :, 15)=conj(A(:, :, 12));
A(:, :, 16)=etaB3.*conj(etaB3)*DWA;
% contre rotation
ph=ph2;
save 'PRE_KERN.mat' A th ph;

%%%%%%%%% COMPUTE_KERN2.m %%%%%%%%%%
load 'LEG_GAUSS.mat'
load 'PRE_KERN.mat'
load 'SPECTRAL.mat'
Ker1=[];
Ker2=[];
Ker3=[];
Ker4=[];
n1=length(th);
n2=length(ph);
tinc=th(2)-th(1);
pinc=ph(2)-ph(1);
th1=[th(1)-tinc;th;th(n1)+tinc];
ph1=[ph(1)-pinc,ph,ph(n2)+pinc];
for p=1:16
    A1=A(:, :, p)/(10E+11);
    if p==1 | p==6 | p==11 | p==16
        cox=ceil(p/4);
        A1=real(A1)/real(ERRMAT(cox,cox));
    else
        cox=ceil(p/4);
        coy=p-4*(cox-1);
        A1=real(A1)/real(ERRMAT(cox,coy))+i*imag(A1)
            .../imag(ERRMAT(cox,coy));
    end;
    si=size(A1);
    s1=si(1);
    s2=si(2);
    A3=[A1(:, s2), A1, A1(:, 1)];
    A1=[A3(1, :); A3; A3(s1, :)];
    A2=[];
    A2p=[];
    A2=interp2(th1, ph1, A1.', X11, X12');
    A2=A2.';
    for ii=1:10
        for jj=1:20
            A2p(20*(ii-1)+jj)=A2(ii, jj);

```



```

        end;
    end;
    Ker1=[Ker1;A2p];
    A2=[];
    A2=interp2(th1,ph1,A1.',X21,X22');
    A2=A2.';
    for ii=1:20
        for jj=1:40
            A2p(40*(ii-1)+jj)=A2(ii,jj);
        end;
    end;
    Ker2=[Ker2;A2p];
    A2=[];
    A2=interp2(th1,ph1,A1.',X31,X32');
    A2=A2.';
    for ii=1:40
        for jj=1:80
            A2p(80*(ii-1)+jj)=A2(ii,jj);
        end;
    end;
    Ker3=[Ker3;A2p];
    A2=[];
    A2=interp2(th1,ph1,A1.',X41,X42');
    A2=A2.';
    for ii=1:64
        for jj=1:128
            A2p(128*(ii-1)+jj)=A2(ii,jj);
        end;
    end;
    Ker4=[Ker4;A2p];
    disp(['Composante ',num2str(p),' calculée']);
end;
disp('* Termine *');
save 'KERN.mat' Ker1 Ker2 Ker3 Ker4;
save '../MIPE/NOYAU_INT/KERN.mat' Ker1 Ker2 Ker3 Ker4;
save '../MIPE2/NOYAU_INT/KERN.mat' Ker1 Ker2 Ker3 Ker4;

```

Bibliographie

- [1] U. Amato , W. Hughes, *Maximum entropy regularization of Fredholm integral equations of the first kind*, Inverse Problems **7** (1991), 739–808.
- [2] C.W Arthur, R.L. McPherron , J.D. Means, *A comparative study of three techniques for using the spectral matrix in wave analysis*, Radio Science **11** (1976), 833–845.
- [3] H. Attouch , R. Cominetti, *L^p approximation of variational problem in L^1 and L^∞* , Nonlinear Analysis **36** (1999), 373–399.
- [4] D. Azé, *Eléments d'analyse convexe et variationnelle*, ellipses, 1997.
- [5] V. Barbu, *Analysis and control of nonlinear infinite dimensional systems*, Academic Press, New-York, 1993.
- [6] M. Bergounioux, *Optimisation et contrôle des systèmes linéaires*, Dunod, 2002.
- [7] J.M. Borwein , A.S. Lewis, *Duality relationships for entropy-like minimization problems*, SIAM J. Control Optim. **29** (1991), 325–338.
- [8] J.M. Borwein , A.S. Lewis, *Convergence of best entropy estimates*, SIAM J. Optim. **1** (1991), 191–205.
- [9] J.M. Borwein , A.S. Lewis, *Partially-finite programming in L_1 and the existence of maximum entropy estimates*, SIAM J. Optim. **3** (1993), 248–267.
- [10] J.M. Borwein, A.S. Lewis , D. Noll, *Maximum entropy reconstruction using derivative information part 1 : Fisher information and convex duality*, Math. of Operations Research **21** (1996), 442–468.
- [11] H. Brezis, *Analyse fonctionnelle*, Dunod, France, 1999.
- [12] L.J. Buchalet , F. Lefeuvre, *One and two direction models for VLF electromagnetic waves observed on-board GEOS 1*, J. Geophys. Res. **86** (1981), 2377–2383.
- [13] B. Buck , V. A. Macaulay, *Maximum entropy in action*, Oxford University Press, 1991.
- [14] G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Dunod, 1998.
- [15] I. Ekeland , R. Temam, *Analyse convexe et problèmes variationnels*, Dunod, Paris, 1974.
- [16] H. W. Engl , G. Landl, *Convergence rates for maximum entropy regularization*, SIAM J. Numer. Anal. **30** (1993), 1509–1536.
- [17] C. Gasquet , P. Witomski, *Analyse de fourier et applications*, Dunod, 2000.

- [18] D. Gilbarg , N. Trudinger, *Elliptic partial differential equation of the second order*, Springer-verlag, Berlin, 2001.
- [19] J.B. Hildebrand, *Introduction to numerical analysis*, McGraw-Hill, 1956.
- [20] B. Holback, S. E. Jansson, L. Ahlen, G. Lundgren, L. Lyngdal, S. Powell, , A. Meyer, *The FREJA wave and plasma density experiment*, Space Sci. Rev. **70** (1994), 577–592.
- [21] L. Hörmander, *The analysis of linear partial differential operators I*, Springer-Verlag, New-York, 1990.
- [22] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Springer-verlag, New-york, 1996.
- [23] K. Kunisch , W. Ring, *Regularization of nonlinear ill-posed problems with closed operators*, Numer. Funct. Anal. and Optimiz. **14** (1993), 389–404.
- [24] F. Lefeuvre, *Analyse de champs d'ondes électromagnétiques aléatoires observées dans la magnétosphère à partir de la mesure simultanée de leur six composantes*, Thèse d'état, Université d'Orléans, Château de la source BP 6749 45067 Orléans Cedex 2, 1977.
- [25] F. Lefeuvre , C. Delannoy, *Analysis of random electromagnetic wave field by a maximum entropy method*, Ann. Télécom. **34** (1979), 204–213.
- [26] ———, *A computer program for the maximum entropy estimation of wave distribution function*, Comp. Phy. Comm. **40** (1986), 389–419.
- [27] F. Lefeuvre, D. Lagoutte , M. Menvielle, *On use of the wave distribution concept to determine the directions of arrivals of radar echoes*, Planet. Space Sci. **48** (2000), 1321–1328.
- [28] F. Lefeuvre , J. L. Pinçon, *Determination of the wave-vector spectrum for plasma waves and turbulence observed in space plasmas*, J. Atm. Terr. Phys. **54** (1992), 1227–1235.
- [29] M. Lefeuvre, F. Parrot , C. Delannoy, *Wave distribution function estimation of VLF electromagnetic waves observed on board GEOS 1*, J. Geophys. Res. **86** (1981), 2359–2375.
- [30] A. S. Leonov, *A generalization of the maximal entropy method for solving ill-posed problems*, Siberian Math. J. **41** (2000), 716–721.
- [31] J. Max , J.L. Lacoume, *Méthodes et techniques du traitements de signal*, Dunod, 1996.
- [32] J.D. Means, *Use of the thee-dimensionnal covariance matrix in analysing the polarisation properties of plane waves*, J. Geophys. Res. **77** (1972), no. 28, 5551–5559.
- [33] F. Natterer, *The mathematics of computerized tomography*, Teubner-Verlag, Stuttgart, 1986.
- [34] D. Noll, *Consistency of a nonlinear deconvolution method with applications to image restoration*, Advances in Mathematical Sciences and Applications **7** (1997), 798–808.
- [35] ———, *Restoration of degraded images with maximum entropy*, J. Glob. Optim. **10** (1997), 91–103.
- [36] A.L. Oppenheim , R.W. Schafer, *Digital signal processing*, Prentice Hall, 1976.

- [37] T. Oscarsson, *Dual principles in maximum entropy reconstruction of the wave distribution function*, J. Comput. Phys. **110** (1994), 221–233.
- [38] T. Oscarsson, G. Stenberg , O. Santolík, *Wave mode identification via wave distribution function analysis*, Phys. Chem. Earth **26** (2001), 229–235.
- [39] M. Parrot , F. Lefeuvre, *Statistical study of the propagation characteristics of ELF hiss observed on GEOS-1, inside and outside the plasmasphere*, Ann. Geophys. **4** (1986), 363–384.
- [40] J.L. Pinçon, Y. Marouan , F. Lefeuvre, *Interpretation of measurements of the polarization percentage for plasma waves*, Ann. Geophys. **10** (1992), 82–95.
- [41] W.H. Press, S.A. Teutolsky, W.T. Vetterling , B.P. Flannery, *Numerical recipes*, Cambridge University Press, 1986.
- [42] O. Prot, M. Bergounioux , J. G. Trotignon, *Determination of a power density by an entropy regularization method*, J. Appl. Math. **2** (2005), 127–152.
- [43] O. Prot, O. Santolík , J. G. Trotignon, *An entropy regularization method applied to an ELF hiss event wave distribution function identification*, (submitted to J. Geophys. Res.) **0** (2004), 0.
- [44] O. Prot, O. Santolík , J.G. Trotignon, *Ill-posed problem solving by an entropy regularization method. application to the propagation analysis of electromagnetic waves*, WDS'04 Proceedings of Contributed Papers : Part III - Physics (J. Safrankova, ed.), Matfyzpress, Prague, 2004, pp. 593–599.
- [45] J. Radon, *Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten*, Ber. Ver. Sächs. Akad. **69** (1917), 262–277.
- [46] H. Reinhard, *Éléments de mathématiques du signal*, Dunod, 2002.
- [47] R.T. Rockafellar, *Convex analysis*, Princeton University Press, 1970.
- [48] K. Ronnmark , J. Larson, *Local spectra and wave distribution functions*, J. Geophys. Res. **93** (1988), 1809–1815.
- [49] J.C. Samson , V. Olson, *Some comments on the descriptions of the polarisation states of waves*, Geophys. J. R. Astron. Soc. **61** (1980), 115–129.
- [50] O. Santolík, *Etude de la fonction de distribution des ondes dans un plasma chaud*, Thèse, Université d'Orléans, 1995.
- [51] O. Santolík , M. Parrot, *Direct problem of the wave distribution function in a hot magnetospheric plasma*, Proceeding of WDS 94, 1994, pp. 19–23.
- [52] O. Santolík , M. Parrot, *The wave distribution function in a hot magnetospheric plasma : The direct problem*, J. Geophys. Res. **101** (1996), no. A5, 10639–10651.
- [53] O. Santolík , M. Parrot, *Application of wave distribution function methods to an ELF hiss event at high latitudes*, J. Geophys. Res. **105** (2000), 18,885–18,894.
- [54] L. Schwartz, *Théorie des distributions*, Hermann, Paris, 1966.
- [55] ———, *Analyse IV - applications à la théorie de la mesure*, Hermann, 1993.

- [56] C. E. Shannon, *A mathematical theory of communication*, The Bell System Technical Journal **27** (1948), 379–423, 623–656.
- [57] J. E. Shore , R. W. Johnson, *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, IEEE Trans. Inf. Theory **26** (1980), 26–36.
- [58] C.R. Smith , W.T. Jr. Grandy, *Maximum-entropy and bayesian spectral analysis and estimations problems*, Reidel, Dordrecht, 1987.
- [59] T.H. Stix, *Waves in plasmas*, Am. Inst. of Phys., New York, 1992.
- [60] L. R. O. Storey, *Wave distribution function analysis of plasmaspheric hiss*, Tech. Report D-711-2, STAR Laboratory Stanford University, 1987.
- [61] L. R. O. Storey , F. Lefeuvre, *Theory for the interpretation of measurements of a random electromagnetic wave field in space*, Space Res. **14** (1974), 381–386.
- [62] ———, *The analysis of 6-component measurement of a random electromagnetic wave field in a magnetoplasma 1, the direct problem*, Geophys. J. R. Astron. Soc. **56** (1979), 255–270.
- [63] ———, *The analysis of 6-component measurement of a random electromagnetic wave field in a magnetoplasma 2, the integration kernels*, Geophys. J. R. Astron. Soc. **62** (1980), 173–194.
- [64] A. N. Tikhonov, A. V. Goncharsky, V.V. Stepanov , A.G. Yagola, *Numerical methods for the solution of ill-posed problems*, Kluwer Academic Publishers, Dordrecht, 1995.
- [65] A.N. Tikhonov , V.Y. Arsenin, *Solutions of ill-posed problems*, J. Wiley and sons, 1977.
- [66] P.D. Welch, *The use of fast fourier transform for the estimation of power spectra : A method on time averaging over short modified periodogramms*, IEEE Trans. Audio and Electroacoust. **15** (1967), 70–73.
- [67] M. Willem, *Analyse harmonique réelle*, Hermann, Paris, 1995.
- [68] K. Yosida, *Functional analysis*, Springer-Verlag, New-York, 1980.
- [69] E. Zeidler, *Nonlinear functional analysis and its applications, part I : Fixed point theorems*, Springer-Verlag, Berlin, 1984.
- [70] ———, *Nonlinear functional analysis and its applications, part II : Linear monotones operators*, Springer-Verlag, Berlin, 1990.

MÉTHODE DE RÉGULARISATION ENTROPIQUE ET APPLICATION AU CALCUL DE LA FONCTION DE DISTRIBUTION DES ONDES

RÉSUMÉ : La détermination des directions de propagation d'une onde électromagnétique, à partir des mesures du champ, est un problème "mal-posé". En utilisant le concept de fonction de distribution des ondes (FDO), cela revient à "inverser" un opérateur linéaire non-bijectif. Nous avons développé une méthode de régularisation entropique pour résoudre ce problème. L'utilisation de l'entropie est numériquement avantageuse et permet de déterminer une solution contenant l'information minimale requise par la donnée. Une généralisation de la méthode a ensuite été étudiée. Le calcul effectif de la FDO a été effectué d'abord dans le cas du vide sur des données synthétiques, puis sur des données provenant du satellite FREJA. La méthode est automatique, robuste et permet de déterminer des solutions stables. Les résultats obtenus sont en accord avec ceux obtenus par d'autres méthodes. Ils sont toutefois plus diffus, ce qui est préférable dans la situation considérée.

ENTROPY REGULARIZATION METHOD AND APPLICATION TO THE IDENTIFICATION OF THE WAVE DISTRIBUTION FUNCTION

ABSTRACT : The identification of directional energy distribution from the field measurement is an "ill-posed" problem. Using the wave distribution function (WDF) concept, this identification is done by finding an "inverse" of a non-bijective linear operator. We have developed an entropy regularization method to solve this problem. The use of entropy has many advantages : in particular, it allows to determine a solution with minimal information required by the data. A theoretical generalization of the method has been studied, as well. This method has been used to compute the WDF in the case of vacuum for simulated data and on experimental data from the FREJA satellite. The method is automatic, robust, and we obtain stable solutions. The obtained results are in accordance with those from other methods. However they are more diffuse, which seems preferable in this context.

DISCIPLINE : Mathématiques Appliquées

MOTS CLÉS : Problème inverse, Régularisation, Traitement du signal, Simulation numérique, Propagation d'ondes.

MAPMO

Université d'Orléans
45067 Orléans Cedex 2

LPCE

3A, Avenue de la recherche Scientifique
45071 Orléans Cedex 2