



HAL
open science

DECOMPOSITIONS ET ALGORITHMES PROXIMAUX POUR L'ANALYSE ET LE TRAITEMENT ITERATIF DES SIGNAUX

Valérie Rozenbaum Wajs

► **To cite this version:**

Valérie Rozenbaum Wajs. DECOMPOSITIONS ET ALGORITHMES PROXIMAUX POUR L'ANALYSE ET LE TRAITEMENT ITERATIF DES SIGNAUX. Analyse numérique [math.NA]. Université Pierre et Marie Curie - Paris VI, 2007. Français. NNT: . tel-00935698

HAL Id: tel-00935698

<https://theses.hal.science/tel-00935698>

Submitted on 23 Jan 2014

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 DE L'UNIVERSITÉ PIERRE ET MARIE CURIE – PARIS VI

Spécialité :

MATHÉMATIQUES APPLIQUÉES

Présentée par :

Valérie ROZENBAUM WAJS

pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE – PARIS VI

Sujet :

**DÉCOMPOSITIONS ET ALGORITHMES PROXIMAUX POUR
L'ANALYSE ET LE TRAITEMENT ITÉRATIF DES SIGNAUX**

Soutenue le 2 juillet 2007 devant le jury composé de :

Gilles AUBERT	<i>Rapporteur</i>
Antonin CHAMBOLLE	<i>Rapporteur</i>
Albert COHEN	<i>Examineur</i>
Roberto COMINETTI	<i>Examineur</i>
Patrick Louis COMBETTES	<i>Directeur de thèse</i>
Myriam COMTE	<i>Examineur</i>
Jean-Michel POGGI	<i>Président</i>

Laboratoire Jacques-Louis Lions – UMR 7598

À Maman et Papa

*Il revoit, les yeux clos, son vert printemps actif,
Ses bras musclés toujours dompter le fer rétif,
Le tordre, le tailler, le sculpter – virtuose! –*

*Plongé dans son brûlant passé forgé d'ardeur,
Il entend crépiter dans une apothéose
D'étincelles d'or fin, l'hymne du fier labeur.*

Extrait du poème *Le Vieux Forgeron*
des *Échos du Cœur* d'Albert Beja.

Remerciements

Je remercie tout d'abord mes parents pour l'éducation qu'ils m'ont transmise, pour leur soutien durant toutes mes études, et pour leur dévouement sans limite. Je remercie mon mari pour son soutien, ses conseils et sa patience durant la préparation du doctorat. Je remercie également mon frère pour ses encouragements, ses conseils et sa complicité depuis notre enfance. Je remercie toute ma famille. Je rends un hommage très affectueux à mon grand-père Albert Beja qui, par des méthodes pédagogiques exceptionnelles, m'a donné le goût des mathématiques, de leur rigueur et de leur enseignement. Je rends un hommage très affectueux à mon grand-père Samuel Rozenbaum déporté à Auschwitz. Je rends un hommage très affectueux à mes grand-mères dont le souvenir de leur sagesse, de leur dynamisme et de leur tenacité m'aide chaque jour à concilier la vie familiale et professionnelle. Je remercie mon Maître J.-D. Frankforter et son épouse qui, par leurs enseignements et leurs conseils, me soutiennent dans cette même voie.

Cette thèse n'aurait pas vu le jour sans la direction de Monsieur P. L. Combettes (Université Paris VI) que je remercie pour m'avoir guidée dans mes recherches, pour m'avoir transmis des méthodes de travail, pour m'avoir aidée à mener de front la famille, la thèse et les enseignements. Ces travaux ont été effectués en partie en collaboration avec Monsieur J.-C. Pesquet et Caroline Chaux (Université Marne-La-Vallée). Je les remercie non seulement pour leurs contributions scientifiques et méthodiques mais aussi pour m'avoir motivée depuis le début de cette collaboration.

Je remercie messieurs G. Aubert (Université de Nice Sophia-Antipolis) et A. Chambolle (École Polytechnique CMAP) d'avoir rapporté sur ma thèse et participer au jury. Je remercie également le Président de jury Monsieur J.-M. Poggi (Université Paris XI) et les autres membres de jury, Monsieur A. Cohen (Université Paris VI), Monsieur R. Cominetti (Université du Chili) pour les multiples questions posées pour enrichir les perspectives de la thèse.

Je rends hommage à Monsieur T. Lachand-Robert, qui en dirigeant mon mémoire de DEA au Laboratoire Jacques-Louis Lions, m'a fortement motivée lors de mes premiers pas dans la recherche. Je remercie Madame M. Comte (Université Paris VI), Monsieur B. Maury (Université Paris-Sud) et Monsieur G. Carlier (Université Paris-Dauphine) pour m'avoir également guidée pendant mon mémoire de DEA. J'ai effectué la thèse dans le chaleureux Laboratoire Jacques-Louis Lions dans de très bonnes conditions. Je remercie tous les membres du laboratoire et les étudiants de mon bureau de thèse. J'ai aussi eu le plaisir d'enseigner en tant qu'ATER dans l'UFR de mathématiques de l'Université René Descartes. Je remercie tous les membres de cet UFR pour leur accueil chaleureux et tout particulièrement Monsieur B. Ycart, Monsieur G. Koepfler, Madame C. Graffigne pour leurs nombreux conseils et leur disponibilité. Je remercie également Monsieur L. D. Cohen (Université Paris-Dauphine) pour son soutien moral et ses précieux conseils. Je remercie tous mes amis.

Table des matières

Remerciements	vii
Résumé	xiii
Notations et Glossaire	xiv
1 Introduction	1
1.1 Le traitement du signal	1
1.1.1 Problèmes inverses	2
1.1.2 Décomposition et analyse de signaux	2
1.2 Un problème variationnel central	3
1.3 Des projecteurs à l'opérateur proximal	4
1.4 L'algorithme de décomposition explicite-implicite	6
1.5 Cadre et motivation de notre démarche	7
1.6 Plan de la thèse	8
1.7 Bibliographie	10
2 L'Opérateur Proximal	14
2.1 Une extension des projecteurs	14
2.1.1 Caractérisation de l'opérateur proximal	14
2.1.2 Quelques propriétés comparatives	15
2.2 Techniques de calcul de l'opérateur proximal	17
2.2.1 Changement de métrique	17
2.2.2 Opérations élémentaires	18
2.2.3 Composition et décomposition	19

2.3	Le principe de décomposition de Moreau	20
2.3.1	Énoncé	21
2.3.2	Une technique calculatoire	22
2.3.2.1	Fonction d'appui	22
2.3.2.2	Enveloppe de Pasch-Hausdorff	24
2.3.2.3	Fonctions réelles	27
2.3.3	Un outil d'analyse de signaux	30
2.3.3.1	Cas des projecteurs	30
2.3.3.2	Seuillage	31
2.3.3.3	Simulations numériques	32
2.3.3.3.1	Analyse M -bande d'un signal	32
2.3.3.3.2	Résultats	33
2.4	Bilan	34
2.5	Bibliographie	47
3	Décomposition, Restauration et Reconstruction de Signaux par la Méthode Proximale Explicite-Implicite	49
3.1	Présentation globale	49
3.2	Les résultats principaux	50
3.2.1	L'opérateur proximal	50
3.2.2	Le Problème 3.1.1 et la méthode de résolution de type explicite- implicite	51
3.2.3	Applications en traitement du signal	52
3.2.3.1	Problèmes impliquant des sommes d'enveloppes de Moreau	53
3.2.3.2	Problèmes inverses linéaires « régularisés »	54
3.2.3.3	Problèmes de débruitage	56
3.3	Article en anglais	57
3.3.1	Introduction	57
3.3.1.1	Notation	58
3.3.2	Proximity operators	59

3.3.2.1	Elements of convex analysis	59
3.3.2.2	Firmly nonexpansive operators	60
3.3.2.3	Proximity operators	61
3.3.2.4	Proximal calculus	62
3.3.2.5	Moreau's decomposition	63
3.3.2.6	Examples of proximity operators	65
3.3.3	Properties and numerical solution of Problem 3.3.1	69
3.3.4	Problems involving sums of Moreau envelopes	74
3.3.4.1	Problem statement	74
3.3.4.2	Properties and numerical solution of Problem 3.3.28	75
3.3.4.3	Proximal split feasibility problems	76
3.3.4.4	The $u + v$ signal decomposition model	77
3.3.4.5	Hard-constrained signal feasibility problems	79
3.3.5	Linear inverse problems	79
3.3.5.1	Problem statement	79
3.3.5.2	Properties and numerical solution of Problem 3.3.39	81
3.3.5.3	Constrained least-squares problems	84
3.3.5.4	Sparse regularization problems	86
3.3.6	Denoising problems	92
3.3.6.1	Problem statement and basic properties	92
3.3.6.2	A split denoising problem	94
3.4	Bibliographie	96

4 Formulation Variationnelle de Problèmes Inverses sur des Trames 101

4.1	Description générale	101
4.2	Les trames	103
4.3	Les principaux résultats	105
4.3.1	Quelques cas d'applications du Problème 4.1.2	105
4.3.1.1	Représentation parcimonieuse	105
4.3.1.2	Approche bayésienne	105
4.3.2	Construction de solutions au Problème 4.1.2	106

4.3.3	Simulations numériques	108
4.4	Article en anglais	108
4.4.1	Introduction	108
4.4.2	Basic tool : proximity operator	110
4.4.2.1	Notation	110
4.4.2.2	Background	110
4.4.2.3	Forward-backward splitting	112
4.4.2.4	Decomposition formula	114
4.4.3	Problem formulation	116
4.4.3.1	Assumptions and problem statement	116
4.4.3.2	Existence and characterization of solutions	117
4.4.3.3	Specific frameworks	119
4.4.3.3.1	Inverse problems with sparsity constraints	119
4.4.3.3.2	Bayesian statistical framework	120
4.4.4	Proximity operators associated with log-concave densities	122
4.4.5	Algorithm	128
4.4.6	Numerical results	131
4.4.6.1	Example 1	133
4.4.6.2	Example 2	133
4.5	Bibliographie	136
5	Bilan et Perspectives	139
5.1	Bilan	139
5.2	Perspectives	139
5.3	Bibliographie	141

Résumé

Décompositions et Algorithmes Proximaux pour l'Analyse et le Traitement Itératif des Signaux

Cette thèse est consacrée à l'étude et la résolution de certains problèmes non linéaires du traitement du signal et de l'image via l'analyse convexe. Nous proposons une étude variationnelle unifiée de problèmes inverses et de problèmes de décomposition de signaux qui ont, jusqu'à présent, été étudiés individuellement en raison de leur apparente disparité. Dans le modèle adopté, cette famille de problèmes est réduite génériquement à la minimisation d'une somme de deux fonctions soumises à certaines propriétés de régularité. Des résultats d'existence, d'unicité et de caractérisation du problème ainsi posé sont obtenus. L'opérateur proximal, introduit par Moreau en 1962 pour les besoins de la mécanique, joue un rôle essentiel dans notre analyse. Nous l'utilisons notamment pour obtenir de nouveaux schémas non linéaires de décomposition de signaux. Cet outil est par ailleurs au cœur de l'algorithme explicite-implicite que nous proposons pour la résolution du problème générique. Ce cadre théorique est appliqué à l'analyse de signaux et à la restauration d'images. Les problèmes de restauration que nous abordons sont posés sur des trames et notre approche permet de prendre en compte des contraintes de parcimonie ou de modéliser des formulations bayésiennes avec des connaissances a priori sur les lois des coefficients de la décomposition. Des résultats numériques sont fournis.

Abstract

Proximal Decompositions and Algorithms for Signal Analysis and Iterative Signal Processing

This thesis is devoted to the study and the resolution of certain nonlinear problems in signal and image processing via convex analysis. We propose a unified variational investigation of inverse problems and signal decomposition problems which have so far been studied individually, because of their apparent disparity. In the model we adopt, this family of problems is reduced generically to the minimization of the sum of two convex functions with certain regularity properties. Existence, uniqueness and characterization results are obtained for this problem. The proximity operator, introduced by Moreau in 1962 to study certain problems in mechanics, plays a basic role in our analysis. We apply it in particular to obtain new nonlinear signal decomposition schemes. Moreover, this tool is at the heart of the forward-backward algorithm which we propose to solve the generic problem. This theoretical framework is applied to signal analysis and to image restoration. The restoration problems under consideration are posed on frames and our approach makes it possible to take into account sparsity constraints or to model Bayesian formulations with a priori knowledge on the distribution of the coefficients of the decomposition. Numerical results are provided.

Notations et Glossaire

Les notations suivantes seront utilisées dans toute la thèse. De plus, nous rappelons quelques définitions de base en analyse convexe.

Notations générales

\mathcal{X}, \mathcal{Y}	Espaces de Hilbert réels pour la théorie
$\mathcal{H}, \mathcal{G}, \mathcal{K}$	Espaces de Hilbert réels pour les applications
\mathbb{K}	Sous-ensemble fini ou infini de \mathbb{N}
$l^2(\mathbb{K})$	Ensemble des suites réelles $(\xi_k)_{k \in \mathbb{K}}$ telles que $\sum_{k \in \mathbb{K}} \xi_k ^2 < +\infty$
$2^{\mathcal{X}}$	La famille des sous-ensembles de \mathcal{X}
Ω	Domaine ouvert de \mathbb{R}^m , $m \in \mathbb{N}^*$
$B(x; \rho)$	La boule centrée en $x \in \mathcal{X}$, de rayon $\rho > 0$
$\langle \cdot \cdot \rangle$	Produit scalaire hilbertien
$\ \cdot \ $	Norme hilbertienne
d	Distance hilbertienne
L^*	Adjoint de l'application linéaire bornée $L: \mathcal{X} \rightarrow \mathcal{Y}$
$\complement C$	Complémentaire dans \mathcal{X} de l'ensemble C de \mathcal{X}
C^\perp	Orthogonal de l'ensemble C de \mathcal{X}
K^\ominus	Polaire du cône convexe K de \mathcal{X}
$\text{int } C$	Intérieur de l'ensemble C de \mathcal{X}
$\text{conv } C$	Enveloppe convexe de l'ensemble C de \mathcal{X}
$x_n \rightarrow x$	La suite $(x_n)_{n \in \mathbb{N}}$ de \mathcal{X} converge fortement vers x
$x_n \rightharpoonup x$	La suite $(x_n)_{n \in \mathbb{N}}$ de \mathcal{X} converge faiblement vers x
$x = \underline{\lim} x_n$	x est la limite inférieure de la suite $(x_n)_{n \in \mathbb{N}}$ de \mathbb{R}

Espaces fonctionnels

$\Gamma_0(\mathcal{X})$	$\{ \varphi: \mathcal{X} \rightarrow]-\infty, +\infty] \mid \varphi \text{ convexe, propre, semi-continue inférieurement} \}$
$\mathcal{B}(\mathcal{X}, \mathcal{Y})$	$\{ L: \mathcal{X} \rightarrow \mathcal{Y} \mid L \text{ linéaire, borné} \}$
$\mathcal{B}(\mathcal{X})$	$\mathcal{B}(\mathcal{X}, \mathcal{X})$
$\mathcal{S}_{++}(\mathcal{X})$	$\{ M \in \mathcal{B}(\mathcal{X}) \mid M^* = M, M \text{ strictement positif} \}$
$L^p(\Omega)$	$\{ x: \Omega \rightarrow \mathbb{R} \text{ mesurable au sens de Lebesgue} \mid \int_{\Omega} x(\omega) ^p d\omega < +\infty \}$,
$W^{1,1}(\Omega)$	$\{ x \in L^1(\Omega) \mid \int_{\Omega} \ \nabla x(\omega)\ d\omega < +\infty \}$,
$H^1(\Omega)$	$\{ x \in L^2(\Omega) \mid \int_{\Omega} \ \nabla x(\omega)\ ^2 d\omega < +\infty \}$

Opérateurs particuliers de \mathcal{X} dans \mathcal{X}

P_C	Projecteur sur l'ensemble convexe fermé non vide C de \mathcal{X}
Id	Opérateur identité sur \mathcal{X}

Fonctions particulières de \mathcal{X} dans $[-\infty, +\infty]$

Soit C un ensemble convexe fermé non vide de \mathcal{X} .

$$\iota_C: x \mapsto \begin{cases} 0, & \text{si } x \in C; \\ +\infty, & \text{si } x \notin C. \end{cases} \quad \text{Fonction indicatrice de } C$$

$$1_C: x \mapsto \begin{cases} 1, & \text{si } x \in C; \\ 0, & \text{si } x \notin C. \end{cases} \quad \text{Fonction caractéristique de } C$$

$$\begin{aligned} d_C: x &\mapsto \inf \|x - C\| && \text{Distance à } C \\ \sigma_C: x &\mapsto \sup \langle x | C \rangle && \text{Fonction d'appui de } C \\ \text{tv}: W^{1,1}(\Omega) &\rightarrow \mathbb{R}: x \mapsto \int_{\Omega} \|\nabla x(\omega)\| d\omega && \text{Variation totale} \end{aligned}$$

Fonctions particulières de \mathbb{R} dans $[-\infty, +\infty]$

$$\text{soft}_{[\underline{\omega}, \bar{\omega}]}: \xi \mapsto \begin{cases} \xi - \underline{\omega}, & \text{si } \xi < \underline{\omega}; \\ 0, & \text{si } \underline{\omega} \leq \xi \leq \bar{\omega}; \\ \xi - \bar{\omega}, & \text{si } \xi > \bar{\omega}. \end{cases} \quad \text{Seuilleur doux sur } [\underline{\omega}, \bar{\omega}], \text{ avec } (\underline{\omega}, \bar{\omega}) \in \mathbb{R}^2$$

$$\text{sign}: \xi \mapsto \begin{cases} 1, & \text{si } \xi > 0; \\ 0, & \text{si } \xi = 0; \\ -1, & \text{si } \xi < 0. \end{cases} \quad \text{Fonction signe}$$

Notations et définitions relatives à l'opérateur multivoque $A: \mathcal{X} \rightarrow 2^{\mathcal{X}}$

$$\begin{aligned} \text{gr } A &= \{(x, u) \in \mathcal{X}^2 \mid u \in Ax\} && \text{Graphe de } A \\ A^{-1} &&& \text{Inverse de l'opérateur } A \\ \text{gr } A^{-1} &= \{(u, x) \in \mathcal{X}^2 \mid (x, u) \in \text{gr } A\} && \text{Graphe de } A^{-1} \\ J_{\gamma A} &= (\text{Id} + \gamma A)^{-1}, \gamma \in]0, +\infty[&& \text{Résolvante de l'opérateur } A \\ (\forall (x, u) \in \text{gr } A) (\forall (y, v) \in \text{gr } A) \langle x - y \mid u - v \rangle &\geq 0 && A \text{ est monotone} \\ (x, u) \in \text{gr } A \Leftrightarrow (\forall (y, v) \in \text{gr } A) \langle x - y \mid u - v \rangle &\geq 0 && A \text{ est maximal monotone} \end{aligned}$$

Quelques définitions relatives à un opérateur $T: \mathcal{X} \rightarrow \mathcal{X}$

- T est β -lipschitz, avec $\beta \in]0, +\infty[$:

$$(\forall (x, y) \in \mathcal{X}^2) \quad \|Tx - Ty\| \leq \beta \|x - y\|.$$

- T est une contraction :

$$(\forall (x, y) \in \mathcal{X}^2) \quad \|Tx - Ty\| \leq \|x - y\|.$$

- T est une contraction ferme : il vérifie l'une des conditions équivalentes suivantes :

- (i) $(\forall(x, y) \in \mathcal{X}^2) \quad \|Tx - Ty\|^2 \leq \langle Tx - Ty \mid x - y \rangle$
- (ii) $(\forall(x, y) \in \mathcal{X}^2) \quad \|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(\text{Id} - T)x - (\text{Id} - T)y\|^2.$

- T est α -moyenné, $\alpha \in]0, 1[$: il existe une contraction $R: \mathcal{X} \rightarrow \mathcal{X}$ telle que $T = (1 - \alpha)\text{Id} + \alpha R$. L'ensemble de opérateurs α -moyennés est noté $\mathcal{A}(\alpha)$.

Notations relatives à $\varphi \in \Gamma_0(\mathcal{X})$

$\text{dom } \varphi = \{x \in \mathcal{X} \mid \varphi(x) < +\infty\}$	Domaine de φ
$\text{epi } \varphi = \{(x, \lambda) \in \mathcal{X} \times \mathbb{R} \mid \varphi(x) \leq \lambda\}$	Épigraphe de φ
$\text{Argmin } \varphi$	Ensemble des minimiseurs de φ
$\text{argmin } \varphi$	Le minimiseur de φ en cas d'unicité
$\varphi^* = \sup_{x \in \mathcal{X}} \langle x \mid \cdot \rangle - \varphi(x)$	Conjuguée de φ
$\varphi 0^+ = \sigma_{\text{dom } \varphi^*}$	Fonction de récession de φ
$\varphi \square \psi = \inf_{y \in \mathcal{X}} \varphi(y) + \psi(y - \cdot)$	Convolution infinimale de φ et $\psi \in \Gamma_0(\mathcal{X})$
$\gamma\varphi = \varphi \square \left(\frac{\ \cdot\ ^2}{2\gamma} \right), \gamma > 0$	Enveloppe de Moreau d'indice γ de φ
$\text{ph}_{\varphi, \gamma} = \varphi \square \left(\frac{\ \cdot\ }{\gamma} \right), \gamma > 0$	Enveloppe de Pasch-Hausdorff d'indice γ de φ

Quelques définitions relatives à une fonction $\varphi \in \Gamma_0(\mathcal{X})$

- φ est fortement convexe de constante $\beta \in]0, +\infty[$: pour tout $(x, y) \in (\text{dom } \varphi)^2$ et tout $\alpha \in]0, 1[$,

$$\varphi(\alpha x + (1 - \alpha)y) \leq \alpha\varphi(x) + (1 - \alpha)\varphi(y) - \beta\alpha(1 - \alpha)\frac{\|x - y\|^2}{2}.$$

- φ est uniformément convexe : \mathcal{C} étant la classe des fonctions croissantes de $[0, +\infty[$ dans $[0, +\infty]$, s'annulant uniquement en 0, il existe $c \in \mathcal{C}$ telle que, pour tout $(x, y) \in (\text{dom } \varphi)^2$,

$$\varphi(\alpha x + (1 - \alpha)y) \leq \alpha\varphi(x) + (1 - \alpha)\varphi(y) - \alpha(1 - \alpha)c(\|x - y\|).$$

- φ est coercive :

$$\lim_{\|x\| \rightarrow +\infty} \varphi(x) = +\infty.$$

- Le sous-différentiel de φ en $x \in \mathcal{X}$:

$$\partial\varphi(x) = \{u \in \mathcal{X} \mid (\forall y \in \mathcal{X}) \langle y - x \mid u \rangle + \varphi(x) \leq \varphi(y)\}.$$

Si φ est Gâteaux différentiable en $x \in \mathcal{X}$, alors $\partial\varphi(x) = \{\nabla\varphi(x)\}$ où $\nabla\varphi(x)$ est le gradient de φ en $x \in \mathcal{X}$.

- L'opérateur proximal de φ :

$$\text{prox}_\varphi: \mathcal{X} \rightarrow \mathcal{X}: x \mapsto \underset{y \in \mathcal{X}}{\text{argmin}} \varphi(y) + \frac{1}{2}\|y - x\|^2.$$

Chapitre 1

Introduction

Cette thèse est consacrée à l'étude et la résolution de certains problèmes non linéaires du traitement du signal et de l'image, via l'analyse convexe et les méthodes variationnelles.

1.1 Le traitement du signal

Dans divers domaines d'applications (physique, biologie, médecine, économie, finance, par exemple), pour reprendre la définition de [18], un *signal* est une « représentation mathématique d'une grandeur mesurable, évoluant en fonction d'une ou plusieurs variables ». Un signal est donc un objet mathématique qui véhicule de l'information. Par exemple, la modélisation mathématique d'une image ou d'un son est un signal. Le traitement du signal a pour objectif d'entreprendre et de développer des techniques de transformation, d'analyse et d'interprétation des signaux. Par exemple, il s'intéresse à la détection, à l'estimation, au codage, à la compression, à la restitution et à la reconstruction d'un signal. Cette discipline repose sur la théorie du signal qui développe, d'après [18], « la modélisation fonctionnelle des signaux et de leurs propriétés, la modélisation opérationnelle de leurs transformations, l'étude des systèmes qui les génèrent ». Elle devient incontournable parce qu'elle constitue un outil indispensable dans d'innombrables domaines de la science et la technologie [25] tels que les appareils numériques (photos, vidéo, musique, télévision), les télécommunications, la musique, l'acoustique, la géophysique, l'imagerie médicale, militaire et spatiale, la météorologie, la sismologie, la géologie. Dans cette thèse, nous nous intéressons à des problèmes *inverses* du traitement du signal mais aussi à des problèmes d'*analyse et décomposition* de signaux.

1.1.1 Problèmes inverses

Le but d'un problème inverse est de restituer un signal originel, à partir d'informations a priori et de l'observation d'un ou plusieurs signaux physiquement associés [13, 40]. En particulier, dans certains problèmes, il existe une relation mathématique simple entre le signal originel \bar{x} et le signal observé z , tous deux dans des espaces de Hilbert réels, notés respectivement \mathcal{H} et \mathcal{G} , à savoir

$$z = T\bar{x} + u, \tag{1.1}$$

où $T: \mathcal{H} \rightarrow \mathcal{G}$ est un opérateur linéaire ou non, modélisant le processus de formation des observations, et où u modélise la mesure du bruit. Plus précisément, en traitement de l'image, domaine applicatif privilégié dans cette thèse, deux principaux types de problèmes inverses se dégagent [16] : la *restauration* de l'image correspondant à l'estimation de la forme originale d'une image à partir d'une ou plusieurs versions dégradées, et la *reconstruction* de l'image déterminant la forme originale d'une image à partir de données indirectement liées à cette image. L'impact de ces deux applications majeures peut s'illustrer par quelques exemples fréquents de problèmes d'estimation et de synthèse de signaux :

- le débruitage [28, 37] afin de réduire le bruit qui parasite le signal originel ;
- la déconvolution [45], de la forme de (1.1), où T est l'opérateur de convolution par un noyau t , représentation mathématique du filtre ou de l'instrument physique transformant le signal originel ;
- l'interpolation [30], opération de reconstruction d'un signal, à partir de sa connaissance sur une partie de son support ; cette technique est utilisée, en particulier, pour les prétraitements photométriques et colorimétriques des images, ou dans les problèmes d'occultation ;
- l'extrapolation, opération de reconstruction un signal au delà d'un support donné ;
- la tomographie [26, 36], application en plein essor depuis les années 1970, dans le but de reconstituer le volume d'un objet 3D ou 2D à partir de projections respectivement 2D ou 1D.

D'autres applications sont mentionnées dans [1, 8, 16, 25, 41] (voir les références mentionnées) concernant des types bien spécifiques d'images, en astronomie ou en optique ondulatoire.

1.1.2 Décomposition et analyse de signaux

Cette thèse se penche, tout particulièrement, sur les décompositions de signaux sous la forme

$$x = x^{\oplus} + x^{\ominus}, \tag{1.2}$$

où x^\oplus et x^\ominus sont alors respectivement interprétés comme la composante du signal possédant une certaine propriété et comme le signal résiduel. Par exemple, dans le cas du débruitage, x^\oplus représente le signal débruité. Ou encore, en traitement de l'image, un modèle d'image peut être créé en analysant la décomposition de l'image donnée selon deux composantes, l'une caractérisant les structures géométriques de l'image et l'autre sa texture [2, 3, 4, 32, 38, 43, 44].

1.2 Un problème variationnel central

La résolution des problèmes de restauration de signaux prend naissance dans un cadre linéaire. C'est par exemple le cas du filtrage de Wiener, tout d'abord, généralisé au problème de la déconvolution [45], ou encore, des techniques de projections sur des sous-espaces vectoriels fermés de l'espace de Hilbert réel \mathcal{H} , unifiées par Youla, dans la fin des années 1970 [46]. L'amélioration des techniques d'estimation jusqu'alors utilisées (par exemple, au sens des moindres carrés, du maximum de vraisemblance, du maximum d'entropie, du maximum a priori [1]) est alors un besoin réel car elles ne vérifient pas toutes les contraintes imposées a priori. Cette approche linéaire est généralisée dans le cas de contraintes convexes, par Youla [47], avec des projecteurs, cette fois, convexes, selon l'algorithme POCS (Projections Onto Convex Sets en anglais). Ce dernier algorithme, loin d'être parfait (en particulier, coûteux, convergence lente), a été amélioré grâce à des notions d'analyse convexe plus approfondies, par de nouveaux algorithmes [5, 6, 17, 19, 21], mais aussi par des formulations variationnelles.

Dans les années 1970, parallèlement aux méthodes de projections, une autre technique est introduite par Hunt [27], consistant à « régulariser » le problème des moindres carrés et résoudre le suivant.

Problème 1.2.1 L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad \|Tx - z\|^2 + \lambda \|Lx\|^2, \quad (1.3)$$

où $\lambda \in]0, +\infty[$, $T \in \mathcal{B}(\mathcal{H})$ et $L \in \mathcal{B}(\mathcal{H})$ est autoadjoint.

Une variante de ce problème est celui de Rudin-Osher-Fatemi [9, 39] étudié sur $\mathcal{H} = L^2(\Omega)$ ($\Omega \subset \mathbb{R}^m$, $m \in \mathbb{N}^*$) :

Problème 1.2.2 L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad \|Tx - z\|^2 + \lambda \text{tv}(x), \quad (1.4)$$

où $\lambda \in]0, +\infty[$, $T \in \mathcal{B}(\mathcal{H})$ et la variation totale en $x \in \mathcal{H}$ est définie par :

$$\text{tv}: x \mapsto \int_{\Omega} \|\nabla x(\omega)\| d\omega. \quad (1.5)$$

D'autres approches variationnelles ont été élaborées, entre autres [14, 21] le

Problème 1.2.3 Soient $(S_i)_{1 \leq i \leq m}$ des convexes fermés non vides de \mathcal{H} . L'objectif est de

$$\underset{x \in S_1}{\text{minimiser}} \sum_{i=2}^m d_{S_i}^2(x). \quad (1.6)$$

D'ores et déjà, nous pouvons remarquer que les fonctionnelles à minimiser dans ces problèmes s'écrivent sous la forme de la somme de deux fonctions $f_1 + f_2$:

- pour le Problème 1.2.1, $f_1 = \lambda \|L \cdot \|^2$ et $f_2 = \|T \cdot -z\|^2$ dont le gradient est lipschitz ;
- pour le Problème 1.2.2, $f_1 = \lambda \text{tv}$ et $f_2 = \|T \cdot -z\|^2$ dont le gradient est lipschitz ;
- pour le Problème 1.2.3, $f_1 = \iota_{S_1}$ et $f_2 = \sum_{i=2}^m d_{S_i}^2$ dont le gradient est lipschitz.

À la lumière de ces exemples, il devient intéressant de se pencher sur le problème de minimisation suivant :

Problème 1.2.4 Soient $f_1 : \mathcal{H} \rightarrow]-\infty, +\infty]$ et $f_2 : \mathcal{H} \rightarrow \mathbb{R}$ deux fonctions de $\Gamma_0(\mathcal{H})$, l'ensemble des fonctions propres, convexes, semi-continues inférieurement. On suppose que f_2 est différentiable sur \mathcal{H} , de gradient $1/\beta$ -lipschitz, pour un certain $\beta \in]0, +\infty[$. L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} f_1(x) + f_2(x). \quad (1.7)$$

Le Problème 1.2.4 impose, certes, des conditions relativement restrictives aux fonctions f_1 et f_2 , mais il a une grande ampleur au cœur de divers problèmes du traitement du signal, comme ceux mis en valeur dans les Chapitres 3 et 4.

1.3 Des projecteurs à l'opérateur proximal

Les projecteurs sont des outils clés en traitement du signal, non seulement pour résoudre les problèmes inverses, mais aussi pour l'analyse des signaux [8, 13, 16, 18, 21, 26, 41, 47]. En effet, par le théorème de projection selon un sous-espace vectoriel fermé V de l'espace de Hilbert \mathcal{H} , nous obtenons la décomposition linéaire d'un signal sous la forme unique (1.2)

$$\begin{cases} x = x^\oplus + x^\ominus, \\ x^\oplus = P_V x, \\ x^\ominus = P_{V^\perp} x, \end{cases} \quad (1.8)$$

où V^\perp est l'orthogonal de V . Par exemple, nous obtenons la décomposition linéaire d'un signal x de $L^2(\Omega)$ suivant les basses fréquences (x^\oplus) et hautes fréquences (x^\ominus), en choisissant

$$V = \{z \in L^2(\mathbb{R}) \mid \hat{z}1_{\mathcal{C}B} = 0\}, \quad (1.9)$$

où :

- (i) $B = [-b/2, b/2]$, avec $b \in]0, +\infty[$, est une bande de fréquence limitée ;
- (ii) pour tout $z \in L^2(\mathbb{R})$, $\hat{z} \in L^2(\mathbb{R})$ est sa transformée de Fourier définie par

$$\hat{z}: \mathbb{R} \rightarrow \mathbb{R}: \xi \mapsto \hat{z}(\xi) = \int_{-\infty}^{+\infty} z(t) e^{-i2\pi\xi t} dt; \quad (1.10)$$

- (iii) $1_{\mathcal{C}B}$ est la fonction caractéristique de $\mathcal{C}B$ définie sur \mathbb{R} par

$$1_{\mathcal{C}B}: \mathbb{R} \rightarrow \mathbb{R}: \xi \mapsto \begin{cases} 0, & \text{si } \xi \in B; \\ 1, & \text{si } \xi \in \mathcal{C}B. \end{cases} \quad (1.11)$$

Cependant le théorème de projection classique est trop restrictif car il ne permet que des décompositions linéaires : par exemple, il ne traite pas le cas où V est un cône convexe fermé, qui a des applications notamment en holographie, en étudiant des signaux dont la phase est connue, ou en spectroscopie, en étudiant des signaux positifs. Dans de telles situations, la décomposition est le projecteur relatif à un cône convexe et au cône polaire associé [18, 40, 47].

De plus, les décompositions des signaux sont parfois plus complexes. Par exemple, les projecteurs ne peuvent modéliser la décomposition associée au seuillage doux d'indice $\gamma \in]0, +\infty[$

$$x^\oplus = \begin{cases} \left(1 - \frac{\gamma}{\|x\|}\right)x, & \text{si } \|x\| > \gamma; \\ 0, & \text{si } \|x\| \leq \gamma. \end{cases} \quad (1.12)$$

Comment apporter une formulation à la fois plus globale et plus synthétique des décompositions de signaux ? Une approche est proposée dans [18] par le biais des opérateurs monotones. Rappelons la décomposition non linéaire paramétrée par un opérateur multivoque maximal monotone $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$, de résolvante $J_{\gamma A} = (\text{Id} + \gamma A)^{-1}$ avec $\gamma \in]0, +\infty[$:

$$\begin{cases} x = x^\oplus + x^\ominus, \\ x^\oplus = J_{\gamma A}x, \\ x^\ominus = \gamma J_{A^{-1}/\gamma}(x/\gamma). \end{cases} \quad (1.13)$$

En particulier, soit $\varphi \in \Gamma_0(\mathcal{H})$. Alors le sous-différentiel

$$\partial\varphi: \mathcal{X} \rightarrow 2^{\mathcal{X}}: x \mapsto \{u \in \mathcal{X} \mid (\forall y \in \mathcal{X}) \langle y - x \mid u \rangle + \varphi(x) \leq \varphi(y)\} \quad (1.14)$$

est maximal monotone d'inverse $\partial\varphi^*$, où $\varphi^* = \sup_{x \in \mathcal{X}} \langle x \mid \cdot \rangle - \varphi(x)$ est la conjuguée de φ . Dans ce cas, $J_{\gamma\partial\varphi} = \text{prox}_{\gamma\varphi}$ est l'opérateur *proximal* de $\gamma\varphi$ introduit, en 1962, par J.-J. Moreau [33] de la manière suivante :

$$\text{prox}_{\gamma\varphi}: \mathcal{H} \rightarrow \mathcal{H}: x \mapsto \underset{y \in \mathcal{H}}{\text{argmin}} \varphi(y) + \frac{1}{2\gamma}\|x - y\|^2. \quad (1.15)$$

Ici, la décomposition (1.13) revient au principe de décomposition de Moreau [33, 34, 35] :

$$\begin{cases} x = x_{\gamma}^{\oplus} + x_{\gamma}^{\ominus}, \\ x_{\gamma}^{\oplus} = \text{prox}_{\gamma\varphi} x, \\ x_{\gamma}^{\ominus} = \gamma \text{prox}_{\varphi^*/\gamma}(x/\gamma). \end{cases} \quad (1.16)$$

C'est le principe de décomposition que nous utiliserons dans toute la thèse.

Prenons maintenant l'exemple où $\varphi = d_C$ où C est un convexe fermé non vide de \mathcal{H} . Alors [18]

$$\text{prox}_{\gamma\varphi} x = \begin{cases} x + \frac{\gamma}{d_C(x)}(P_C x - x), & \text{si } d_C(x) > \gamma; \\ P_C x, & \text{si } d_C(x) \leq \gamma. \end{cases} \quad (1.17)$$

Dans le cas particulier où $C = \{0\}$, $\text{prox}_{\gamma\varphi}$ n'est rien d'autre que le seuillage doux (1.12). Prenons un second exemple. Soit $\varphi = \iota_C$ où C est un convexe fermé non vide de \mathcal{H} . Alors $\text{prox}_{\varphi} = P_C$. L'opérateur proximal se positionne également comme une généralisation des projecteurs.

Le Chapitre 2 est entièrement consacré à l'étude de l'opérateur proximal, outil fondamental de la thèse.

1.4 L'algorithme de décomposition explicite-implicite

La résolution du Problème 1.2.4 est abordée, dans cette thèse, par un algorithme de type *explicite-implicite* (*forward-backward algorithm* en anglais), s'articulant en deux étapes : la première explicite n'implique que f_2 ; la seconde implicite est fonction de l'opérateur proximal de f_1 . Cet algorithme, qui apparaît sous des formes plus élémentaires dans [31, 42], est un cas particulier de l'algorithme à $m \in \mathbb{N}^*$ pas dans un espace de Hilbert réel \mathcal{H} , étudié dans [20], sous la forme :

Algorithme 1.4.1 Fixer $x_0 \in \mathcal{H}$ et construire, pour tout $n \in \mathbb{N}$,

$$x_{n+1} = x_n + \lambda_n \left(T_{1,n} \left(T_{2,n} (\cdots T_{m-1,n} (T_{m,n} x_n + e_{m,n}) + e_{m-1,n} + \cdots) + e_{2,n} \right) + e_{1,n} - x_n \right), \quad (1.18)$$

où :

- (i) pour tout $i \in \{1, \dots, m\}$, $T_{i,n} \in \mathcal{A}(\alpha_{i,n})$ où $\alpha_{i,n} \in]0, 1[$ et $\mathcal{A}(\alpha_{i,n})$ est la classe des opérateurs $T: \mathcal{H} \rightarrow \mathcal{H}$ qui sont $\alpha_{i,n}$ -moyennés, i.e., il existe une contraction $R: \mathcal{H} \rightarrow \mathcal{H}$ telle que $T = (1 - \alpha_{i,n}) \text{Id} + \alpha_{i,n} R$;
- (ii) pour tout $i \in \{1, \dots, m\}$, $e_{i,n} \in \mathcal{H}$ est un terme modélisant l'erreur autorisée à chaque étape i ;
- (iii) $\lambda_n \in]0, 1]$ est un paramètre de relaxation.

Cet algorithme a été élaboré pour résoudre, entre autres, des problèmes d'équilibre dans un espace de Hilbert \mathcal{H} , sous forme :

- de somme d'opérateurs maximaux monotones $(A_i)_{i \in \mathbb{K} \subset \mathbb{N}}$:

$$\text{Trouver } x \in \mathcal{H} \text{ tel que } 0 \in \sum_{i \in \mathbb{K}} A_i x; \quad (1.19)$$

- d'intersection d'opérateurs maximaux monotones $(A_i)_{i \in \mathbb{K} \subset \mathbb{N}}$:

$$\text{Trouver } x \in \mathcal{H} \text{ tel que } 0 \in \bigcap_{i \in \mathbb{K}} A_i x. \quad (1.20)$$

L'Algorithme 1.4.1 a, de plus, l'avantage d'être assez général. En effet, comme détaillé dans [20], il admet pour cas particuliers plusieurs algorithmes de base, entre autres, en choisissant $m = 2$, les erreurs $(e_{i,n})_{1 \leq i \leq m}$ nulles et $\lambda_n \equiv 1$, l'algorithme de Douglas-Rachford, celui de Peaceman-Rachford, celui des projections alternées de Von Neumann, celui des minimisations alternées d'Acker et Prestel. En introduisant un terme d'erreurs numériques, des améliorations de ces diverses méthodes sur les critères d'admissibilité et de convergence sont également des cas particuliers de l'Algorithme 1.4.1, comme l'algorithme du point proximal, où $T_{1,n} = J_{\gamma_n} A$, et la méthode de décomposition à deux étapes non stationnaire, où $T_{1,n} = J_{\gamma_n} A$ et $T_{2,n} = \text{Id} - \gamma_n B$ [29, 42].

1.5 Cadre et motivation de notre démarche

Dans cette thèse, nous étudions le Problème 1.2.4 ainsi que ses applications à divers problèmes de traitement du signal, et nous illustrons la portée de nos résultats par des simulations numériques en imagerie. Le Problème 1.2.4, s'il est restrictif,

est soumis à des hypothèses assez courantes dans divers champs d’applications en traitement du signal et se positionne ainsi comme un problème important. Ce point sera amplement illustré par les exemples concrets fournis dans les Chapitres 3 et 4.

Le second avantage du Problème 1.2.4 est de posséder des résultats d’existence, d’unicité, de caractérisation, obtenus par le biais de l’opérateur proximal, puissant outil d’analyse convexe dont les propriétés sont explorées et exploitées tout au long de la thèse. Ces résultats permettront d’analyser de manière unifiée divers problèmes du traitement de l’image qui, jusqu’à présent, ont été étudiés individuellement en raison de leur apparente disparité.

De plus, l’intérêt applicatif d’une telle démarche est confirmé par des simulations, obtenues grâce à l’algorithme explicite-implicite de résolution du Problème 1.2.4. Ces résultats numériques sont effectués sur quelques problèmes inverses posés sur des trames, outils assez récemment employés en traitement du signal.

Il est à remarquer que les mathématiques appliquées au traitement du signal constituent une technologie à part entière qui doit répondre aux exigences actuelles et futures [25] : recherche de modélisations et simulations, afin de concevoir de nouveaux *matériels virtuels* qui ont certaines propriétés requises pour obtenir de meilleurs résultats peu coûteux, tant en rapidité qu’en qualité d’image, ou pour extraire de l’information à partir de données et en s’appliquant au plus large spectre d’applications possibles.

Notre motivation principale est de répondre au mieux à ces attentes : nous apportons un modèle qui globalise une large classe de problèmes en traitement du signal et dont la résolution, grâce à des outils théoriques et numériques puissants, conduit à une analyse simplifiée de ces applications. Des simulations numériques sur des applications à la restauration d’images sont également fournies comme illustrations.

1.6 Plan de la thèse

Nous exposons maintenant l’organisation de la thèse, avec un aperçu des résultats essentiels obtenus.

Le Chapitre 2 présente une synthèse sur l’opérateur proximal introduit par J.-J. Moreau, en 1962, sur un espace de Hilbert réel \mathcal{H} . Il reprend les études menées dans les autres chapitres et approfondit certains résultats sur ce sujet. Par comparaison avec les projecteurs, dont il représente une généralisation, les premières caractérisations et propriétés essentielles sont dégagées. Pour ne nommer que les principales :

- c’est d’abord une contraction ferme, notion importante pour la convergence d’algorithmes basés sur cet outil ;
- dans beaucoup de cas utiles de fonctions $\varphi \in \Gamma_0(\mathcal{H})$, l’opérateur proximal

- prox_φ se détermine explicitement par des opérations élémentaires ;
- des propriétés de décomposition selon une base hilbertienne sont établies ;
- le principe de décomposition de Moreau (1.16) a un double intérêt : d'un point de vue théorique, il simplifie les calculs en exhibant une relation entre l'opérateur proximal de φ et celui de sa conjuguée ; d'un point de vue appliqué, il présente un puissant schéma de décomposition de signaux linéaire ou non linéaire.

Des exemples illustrent cette théorie ; en particulier, dans la section consacrée au principe de décomposition de Moreau, nous présentons des simulations numériques. L'article de conférence [22] reprend certaines de ces notions.

L'objectif du Chapitre 3 est de résoudre le Problème 1.2.4 à l'aide de l'opérateur proximal et de mettre en évidence son impact au cœur des problèmes de traitement du signal. Les résultats s'organisent en trois étapes :

- la première est consacrée à l'opérateur proximal ;
- la seconde étudie le Problème 1.2.4 en énonçant des propriétés d'existence, d'unicité et de caractérisation puis un algorithme de résolution, l'algorithme proximal explicite-implicite dont la convergence est établie ;
- la dernière partie illustre la seconde par des applications en traitement du signal, entre autres, le schéma de décomposition du signal $u + v$ pour analyser la texture et la géométrie d'une image, les problèmes d'admissibilité d'un signal soumis à des contraintes, les problèmes inverses linéaires (notamment la méthode proximale de Landweber, les problèmes de contraintes de moindres carrés, les problèmes de régularisation parcimonieuse) et les problèmes de débruitage.

Ces travaux ont fait l'objet de l'article de conférence [10] et de l'article de revue [23].

Le Chapitre 4 se focalise sur les applications du Problème 1.2.4 aux problèmes inverses en imagerie, soumis à des informations a priori sur la représentation de la solution idéale sur une trame. Ce travail s'organise en trois points majeurs.

- Tout d'abord, nous présentons une formulation du Problème 1.2.4 dans l'espace $l^2(\mathbb{K})$ des coefficients de la trame, modélisation, en particulier, des problèmes inverses linéaires soumis à des contraintes parcimonieuses, et des problèmes inverses linéaires de restauration d'images par l'approche statistique bayésienne.
- La seconde partie théorique rappelle brièvement et généralise quelques propriétés de calcul de l'outil principal qu'est l'opérateur proximal. Puis elle présente des exemples d'opérateurs proximaux de fonctions potentiels associées à des distributions de probabilité, à une seule variable, log-concaves, représentant chaque coefficient de la trame dans les formulations bayésiennes. Ces résultats ont pour intérêt de fournir des calculs explicites qui seront importants dans la troisième partie.

- La dernière partie traite le problème soulevé dans la première partie sous trois aspects : d'un point de vue théorique, d'abord, en prouvant l'existence de solutions, d'un point de vue algorithmique, en étudiant la convergence de l'algorithme de type explicite-implicite associé au problème posé, et d'un point de vue numérique par des simulations montrant l'intérêt de notre étude sur des tels problèmes de restauration d'images posés sur une trame, et ce, en utilisant les fonctions potentiels exhibées auparavant.

L'article de conférence [11] et l'article de revue [12] sont les fruits des travaux de ce chapitre.

Enfin, le Chapitre 5 vient conclure ces travaux de recherche et proposer de nouvelles perspectives.

1.7 Bibliographie

- [1] H. C. ANDREWS AND B. R. HUNT, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
- [2] J.-F. AUJOL, G. AUBERT, L. BLANC-FÉRAUD, AND A. CHAMBOLLE, *Image decomposition into a bounded variation component and an oscillating component*, Journal of Mathematical Imaging and Vision, 22 (2005), pp. 71–88.
- [3] J.-F. AUJOL AND A. CHAMBOLLE, *Dual norms and image decomposition models*, International Journal of Computer Vision, 63 (2005), pp. 85–104.
- [4] J.-F. AUJOL, G. GILBOA, T. CHAN, AND S. OSHER, *Structure-texture image decomposition - Modeling, algorithms, and parameter selection*, International Journal of Computer Vision, 67 (2006), pp. 111–136.
- [5] H. H. BAUSCHKE AND J. M. BORWEIN, *On projections algorithms for solving convex feasibility problems*, SIAM Review, 38 (1996), pp. 367–426.
- [6] H. H. BAUSCHKE, P. L. COMBETTES, AND S. G. KRUK, *Extrapolation algorithm for affine-convex feasibility problems*, Numerical Algorithms, 41 (2006), pp. 239–274.
- [7] H. BRÉZIS, *Analyse Fonctionnelle : Théorie et Applications*, Dunod, Seconde Édition, Paris, 1993.
- [8] C. L. BYRNE, *Signal Processing – A Mathematical Approach*, A. K. Peters, Welleseley, MA, 2005.
- [9] A. CHAMBOLLE AND P. L. LIONS, *Image recovery via total variation minimization and related problems*, Numerische Mathematik, 76 (1997), pp. 167–188.
- [10] C. CHAUX, P. L. COMBETTES, J.-C. PESQUET, AND V. R. WAJS, *A forward-backward algorithm for image restoration with sparse representations*, Proceedings of the International Conference on Signal Processing with Adaptive Sparse Structured Representations, Rennes, France (2005), pp. 49–52.

- [11] C. CHAUX, P. L. COMBETTES, J.-C. PESQUET, AND V. R. WAJS, *Iterative image deconvolution using overcomplete representations*, Proceedings of the 14th European Signal Processing Conference, Florence, Italy (2006), 4 pages.
- [12] C. CHAUX, P. L. COMBETTES, J.-C. PESQUET, AND V. R. WAJS, *A variational formulation for frame-based inverse problems*, Inverse Problems, 23 (2007), pp. 1495–1518.
- [13] P. L. COMBETTES, *The foundations of set theoretic estimation*, Proceedings of the IEEE, 81 (1993), pp. 182–208.
- [14] P. L. COMBETTES, *Inconsistent signal feasibility problems : Least-squares solutions in a product space*, IEEE Transactions on Signal Processing, 42 (1994), pp. 2955–2966.
- [15] P. L. COMBETTES, *Construction d'un point fixe commun à une famille de contractions fermes*, Comptes Rendus de l'Académie des Sciences de Paris, Série I (Mathématique), 320 (1995), pp. 1385–1390.
- [16] P. L. COMBETTES, *The convex feasibility problem in image recovery*, Advances in Imaging and Electron Physics, 95 (1996), (P. Hawkes, Ed.), Academic Press, New York, pp. 155–270.
- [17] P. L. COMBETTES, *Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections*, IEEE Transactions on Image Processing, 6 (1997), pp. 493–506.
- [18] P. L. COMBETTES, *Convexité et signal*, Actes du Congrès de Mathématiques Appliquées et Industrielles SMAI'01, Pompadour, France (2001), pp. 6–16.
- [19] P. L. COMBETTES, *Quasi-Fejerian analysis of some optimization algorithms*, in Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, (D. Butnariu, Y. Censor, and S. Reich, Eds.), (2001) Elsevier, New York, pp. 115–152.
- [20] P. L. COMBETTES, *Solving monotone inclusions via compositions of nonexpansive averaged operators*, Optimization, 53 (2004), pp. 475–504.
- [21] P. L. COMBETTES AND P. BONDON, *Hard-constrained inconsistent signal feasibility problems*, IEEE Transactions on Signal Processing, 47 (1999), pp. 2460–2468.
- [22] P. L. COMBETTES AND V. R. WAJS, *Theoretical analysis of some regularized image denoising methods*, Proceedings of the IEEE International Conference on Image Processing, 1, Singapore (2004), pp. 969–972.
- [23] P. L. COMBETTES AND V. R. WAJS, *Signal recovery by proximal forward-backward splitting*, Multiscale Modeling and Simulation, 4 (2005), pp. 1168–1200.
- [24] J. M. DYE AND S. REICH, *Unrestricted iterations of nonexpansive mappings in Hilbert space*, Nonlinear Analysis, Theory, Methods, and Applications, 18 (1992), pp. 199–207.

- [25] K. M. FURATI, Z. NASHED, AND A. H. SIDDIQI, *Mathematical Models and Methods for Real World Systems*, CRC, Boca Raton, FL, 2005.
- [26] G. T. HERMAN, *Image Reconstruction from Projections, the Fundamentals of Computerized Tomography*, Academic Press, New York, 1980.
- [27] B. R. HUNT, *The application of constrained least-squares estimation to image restoration by digital computer*, IEEE Transactions on Computers, 22 (1973), pp. 805–812.
- [28] M. KUNT, *Traitement Numérique des Signaux*, Presses Polytechniques et Universitaires Romandes, Troisième Édition, Lausanne, 1996.
- [29] B. LEMAIRE, *Which fixed point does the iteration method select ?*, Lecture Notes in Economics and Mathematical Systems, 452 (1997), pp. 154–167.
- [30] H. MAÎTRE, *Le Traitement des Images*, Hermès Science, Paris, 2003.
- [31] B. MERCIER, *Inéquations variationnelles de la mécanique*, Publications Mathématiques d’Orsay, 80.01 (1980), Université Paris 11, Orsay, France.
- [32] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS, Providence, RI, 2001.
- [33] J.-J. MOREAU, *Fonctions convexes duales et points proximaux dans un espace hilbertien*, Comptes-Rendus de l’Académie des Sciences, Paris, Série A. Mathématique, 255 (1962), pp. 2897–2899.
- [34] J.-J. MOREAU, *Propriétés des applications « prox »*. Comptes-Rendus de l’Académie des Sciences, Paris, Série A. Mathématique, 256 (1963), pp. 1069–1071.
- [35] J.-J. MOREAU, *Proximité et dualité dans un espace hilbertien*, Bulletin de la Société Mathématique de France, 93 (1965), pp. 273–299.
- [36] F. NATTERER, *The Mathematics of Computerized Tomography*, SIAM, Second Edition, Philadelphia, 2001.
- [37] A. V. OPPENHEIM AND R. W. SCHAFER, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
- [38] S. OSHER, A. SOLÉ, AND L. VESE, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Modeling and Simulation, 1 (2003), pp. 349–370.
- [39] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.
- [40] H. STARK, *Image Recovery : Theory and Application*, Academic Press, San Diego, CA, 1987.
- [41] H. STARK AND Y. YANG, *Vector Space Projections : A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics*, Wiley, New York, 1998.

- [42] P. TSENG, *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM Journal on Control and Optimization, 29 (1991), pp. 119–138.
- [43] L. A. VESE AND S. J. OSHER, *Modeling textures with total variation minimization and oscillating patterns in image processing*, Journal on Scientific Computing, 19 (2003), pp. 553–572.
- [44] L. A. VESE AND S. J. OSHER, *Image denoising and decomposition with total variation minimization and oscillatory functions*, Journal of Mathematical Imaging and Vision, 20 (2004), pp. 7–18.
- [45] N. WIENER, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*, The Technology Press of the Massachusetts Institute of Technology, Cambridge, 1949.
- [46] D. C. YOULA, *Generalized image restoration by the method of alternating orthogonal projections*, IEEE Transactions on Circuits and Systems, 25 (1978), pp. 694–702.
- [47] D. C. YOULA AND H. WEBB, *Image restoration by the method of convex projections : Part 1 – theory*, IEEE Transactions on Medical Imaging, 1 (1982), pp. 81–94.

Chapitre 2

L'Opérateur Proximal

En 1962, Jean-Jacques Moreau a introduit la notion d'*opérateur proximal*, dans un espace de Hilbert réel, pour répondre à des besoins de la mécanique, et tout particulièrement, de la dynamique et de la statique des systèmes matériels à liaisons unilatérales (cf. [17, 18, 19] et leurs références). En traitement du signal, l'intérêt de l'opérateur proximal a été mis en évidence dans [9].

Ce chapitre propose une synthèse et un approfondissement des propriétés de cet opérateur, outil clé de toute notre étude tant théorique qu'appliquée. Des exemples sont, de plus, exhibés pour éclairer sur l'utilité et l'efficacité d'un tel opérateur. Une majeure partie des résultats est la source des articles publiés et présentés au Chapitre 3 et au Chapitre 4.

L'opérateur proximal est une extension des projecteurs (Section 2.1), dont des techniques de calcul sont décrites à la Section 2.2. La Section 2.3 est consacrée au principe de décomposition de Moreau en tant qu'outil d'analyse de signaux, et donne lieu à des simulations numériques.

Dans ce Chapitre, \mathcal{X} et \mathcal{Y} sont des espaces de Hilbert réels.

2.1 Une extension des projecteurs

2.1.1 Caractérisation de l'opérateur proximal

Définition 2.1.1 [19] *Soit $\varphi \in \Gamma_0(\mathcal{X})$. L'opérateur proximal de φ , noté $\text{prox}_\varphi: \mathcal{X} \rightarrow \mathcal{X}$, associe, à tout $x \in \mathcal{X}$, l'unique minimiseur de la fonction $\varphi + \|\cdot - x\|^2/2$.*

Cette définition conduit à une double caractérisation de l'opérateur proximal : par des inégalités, tout d'abord, en utilisant la définition du sous-différentiel (cf.

Notations et Glossaire),

$$(\forall x \in \mathcal{X})(\forall p \in \mathcal{X}) \quad p = \text{prox}_\varphi x \Leftrightarrow x - p \in \partial\varphi(p), \quad (2.1)$$

$$\Leftrightarrow (\forall y \in \mathcal{X}) \quad \langle y - p \mid x - p \rangle + \varphi(p) \leq \varphi(y), \quad (2.2)$$

puis, à partir de la résolvante de l'opérateur maximal monotone $\partial\varphi$ (cf. Notations et Glossaire),

$$(\forall x \in \mathcal{X}) \quad \text{prox}_\varphi x = J_{\partial\varphi} x = (\text{Id} + \partial\varphi)^{-1}x. \quad (2.3)$$

Soulignons le lien avec l'enveloppe de Moreau de $\varphi \in \Gamma_0(\mathcal{X})$, indexée par $\gamma \in]0, +\infty[$, notée $\gamma\varphi$: cette convolution infimale, $\varphi \square (\|\cdot\|^2/(2\gamma))$, est une fonction convexe continue vérifiant,

$$(\forall x \in \mathcal{X}) \quad \gamma\varphi(x) = \varphi(\text{prox}_{\gamma\varphi} x) + \frac{1}{2\gamma} \|\text{prox}_{\gamma\varphi} x - x\|^2. \quad (2.4)$$

Proposition 2.1.2 [19] *Soit $\varphi \in \Gamma_0(\mathcal{X})$. Alors $\gamma\varphi$ est Fréchet-différentiable sur \mathcal{X} et $\nabla \gamma\varphi = (\text{Id} - \text{prox}_{\gamma\varphi})/\gamma$.*

Considérons, maintenant, le cas particulier où $\varphi = \iota_C$, C étant un ensemble convexe fermé non vide de \mathcal{X} . Alors

$$\gamma\varphi = \frac{1}{2\gamma} d_C^2 \quad \text{et} \quad \text{prox}_{\gamma\varphi} = P_C. \quad (2.5)$$

Considérer les projecteurs comme des opérateurs proximaux a une large portée en traitement du signal, notamment, dans le cas où C est un sous-espace vectoriel ou un cône convexe fermé de \mathcal{X} . En effet, grâce au principe de décomposition de Moreau, exposé à la Section 2.3, nous retrouvons des décompositions bien connues en traitement du signal [9].

2.1.2 Quelques propriétés comparatives

Nous procédons maintenant à une étude comparative de quelques caractéristiques des projecteurs et des opérateurs proximaux.

Rappelons, tout d'abord, que P_C est linéaire si et seulement si C est un sous-espace vectoriel fermé de \mathcal{X} . L'opérateur proximal, quant à lui, n'est en général pas linéaire :

Proposition 2.1.3 [19, Section 3] *$L \in \mathcal{B}(\mathcal{X})$ est un opérateur proximal si et seulement si $L = L^*$, $\|L\| \leq 1$, et $(\forall x \in \mathcal{X}) \langle Lx \mid x \rangle \geq 0$.*

L'une des propriétés, commune aux projecteurs et à l'opérateur proximal, primordiale pour l'étude algorithmique est le caractère fermement contractant :

Définition 2.1.4 $T: \mathcal{X} \rightarrow \mathcal{X}$ est une contraction ferme s'il vérifie l'une des conditions équivalentes suivantes :

- (i) $(\forall(x, y) \in \mathcal{X}^2) \quad \|Tx - Ty\|^2 \leq \langle Tx - Ty \mid x - y \rangle.$
- (ii) $(\forall(x, y) \in \mathcal{X}^2) \quad \|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(\text{Id} - T)x - (\text{Id} - T)y\|^2.$

Cette propriété permet d'obtenir la convergence de l'algorithme de projections alternées de Von Neumann [9] basé sur des projecteurs affines, et aussi, celle de l'algorithme de type explicite-implicite, exposé au Chapitre 3, basé sur des opérateurs proximaux. Cette propriété est étudiée à la Section 3.3.2 de l'article de revue rapporté au Chapitre 3.

Proposition 2.1.5 Soit $\varphi \in \Gamma_0(\mathcal{X})$. Alors prox_φ et $\text{Id} - \text{prox}_\varphi$ sont des contractions fermes.

La Proposition 2.1.5 a un second intérêt : par contraposée, elle détecte les opérateurs qui ne sont pas proximaux. En particulier, les opérateurs non lipschitz ne sont pas proximaux. Prenons un exemple bien connu en traitement du signal :

Exemple 2.1.6 La méthode de seuillage permet de séparer le signal du bruit. Le paramètre $\gamma \in]0, +\infty[$ étant lié au niveau de bruit, le seuillage dur correspond à un opérateur S_γ défini sur \mathcal{X} par :

$$(\forall x \in \mathcal{X}) \quad S_\gamma x = \begin{cases} x, & \text{si } \|x\| > \gamma; \\ 0, & \text{si } \|x\| \leq \gamma. \end{cases} \quad (2.6)$$

Les résultats par seuillage dur sont souvent moins bons lorsque le nombre d'observations est modéré car le seuillage est trop brutal puisqu'il crée une discontinuité. La Section 2.3.3.2 est d'ailleurs consacrée aux relations entre opérateur proximal et seuillage.

Étudions maintenant les points fixes de l'opérateur proximal (i.e., les éléments $x \in \mathcal{X}$ vérifiant $\text{prox}_\varphi x = x$) :

Proposition 2.1.7 Soit $\varphi \in \Gamma_0(\mathcal{X})$. L'ensemble des points fixes de l'opérateur proximal de φ est $\text{Argmin } \varphi$.

Démonstration. Soit $x \in \mathcal{X}$ vérifiant $\text{prox}_\varphi x = x$. Par (2.1),

$$\text{prox}_\varphi x = x \Leftrightarrow 0 \in \partial\varphi(x). \quad (2.7)$$

D'après la règle de Fermat [11], x est un minimiseur de φ . \square

Remarque 2.1.8 L'ensemble des points fixes de l'opérateur proximal peut être vide (si φ n'admet pas de borne inférieure finie), tandis que celui d'un projecteur sur un ensemble non vide convexe fermé C de \mathcal{X} est toujours non vide, c'est C tout entier.

Une autre propriété des projecteurs est l'involution : l'opérateur proximal ne l'est pas, en général.

2.2 Techniques de calcul de l'opérateur proximal

L'objectif de cette section est d'établir une liste de propriétés de l'opérateur proximal fournissant, de manière succincte, une expression simplifiée voire explicite de cet opérateur. De tels résultats ont un large impact sur les méthodes d'optimisation pour établir la convergence de certains algorithmes ou pour faciliter leur mise en œuvre numérique.

2.2.1 Changement de métrique

Il est parfois nécessaire de calculer l'opérateur proximal dans une métrique différente de la métrique canonique. Pour établir une telle expression, nous introduisons tout d'abord de nouvelles notations.

Notation 2.2.1 Soient $\varphi \in \Gamma_0(\mathcal{X})$ et $\mathcal{S}_{++}(\mathcal{X})$ l'ensemble des opérateurs de $\mathcal{B}(\mathcal{X})$ autoadjoints et strictement positifs. Soit $M \in \mathcal{S}_{++}(\mathcal{X})$. Alors M admet une unique racine carrée, notée \sqrt{M} , i.e., il existe un unique opérateur $R \in \mathcal{S}_{++}(\mathcal{X})$ tel que $R^2 = M$ [20]. Le produit scalaire relatif à M sur \mathcal{X} est $\langle \cdot | \cdot \rangle_M = \langle \cdot | M \cdot \rangle$ et $\| \cdot \|_M$ est la norme associée. Le sous-différentiel de φ relatif à M est, en tout $x \in \mathcal{X}$,

$$\partial_M \varphi(x) = \{u \in \mathcal{X} \mid (\forall y \in \mathcal{X}) \quad \langle y - x | u \rangle_M + \varphi(x) \leq \varphi(y)\}. \quad (2.8)$$

Dans le cas où φ est différentiable en $x \in \mathcal{X}$,

$$\partial_M \varphi(x) = \{\nabla_M \varphi(x)\} = \{M^{-1} \nabla \varphi(x)\}. \quad (2.9)$$

L'opérateur proximal de φ relatif à M est défini pour tout $x \in \mathcal{X}$ par :

$$\text{prox}_\varphi^M x = \underset{y \in \mathcal{X}}{\text{argmin}} \quad \varphi(y) + \frac{1}{2} \|y - x\|_M^2. \quad (2.10)$$

Proposition 2.2.2 (Formule de changement de métrique.)

Soient $M \in \mathcal{S}_{++}(\mathcal{X})$ et $\varphi \in \Gamma_0(\mathcal{X})$. Alors

$$(\forall x \in \mathcal{X}) \quad \text{prox}_\varphi^M x = \sqrt{M}^{-1} \text{prox}_{\varphi \circ \sqrt{M}^{-1}}(\sqrt{M}x). \quad (2.11)$$

Démonstration. D'après (2.1),

$$\begin{aligned}
(\forall x \in \mathcal{X}) \quad p = \text{prox}_\varphi^M x &\Leftrightarrow x - p \in \partial_M \varphi(p) \\
&\Leftrightarrow M(x - p) \in \partial \varphi(p) \\
&\Leftrightarrow \sqrt{M}(x - p) \in \sqrt{M}^{-1} \partial \varphi(p) \\
&\Leftrightarrow \sqrt{M}x - \sqrt{M}p \in \partial(\varphi \circ \sqrt{M}^{-1})(\sqrt{M}p). \tag{2.12}
\end{aligned}$$

□

Certaines méthodes d'optimisation, basées sur l'opérateur proximal et les approximations quasi-newtoniennes de fonctions régulières, en travaillant sur des métriques variables [3, 14] peuvent être formulées de manière plus synthétique par la Proposition 2.2.2.

2.2.2 Opérations élémentaires

L'idée est de déterminer, par des opérations élémentaires, l'opérateur proximal d'une fonction à partir de celui d'une autre, comme étudié dans la Section 3.3.2 de l'article de revue rapporté au Chapitre 3.

Proposition 2.2.3 *Soit $\varphi \in \Gamma_0(\mathcal{X})$ et $x \in \mathcal{X}$. Nous avons les propriétés de calcul suivantes :*

- (i) *Perturbation quadratique : Soit $\psi = \varphi + \alpha \|\cdot\|^2/2 + \langle \cdot | u \rangle + \beta$, où $u \in \mathcal{X}$, $\alpha \in [0, +\infty[$ et $\beta \in \mathbb{R}$. Alors, $\text{prox}_\psi x = \text{prox}_{\varphi/(\alpha+1)}((x - u)/(\alpha + 1))$.*
- (ii) *Translation : Soit $\psi = \varphi(\cdot - z)$, où $z \in \mathcal{X}$. Alors, $\text{prox}_\psi x = z + \text{prox}_\varphi(x - z)$.*
- (iii) *Dilatation : Soit $\psi = \varphi(\cdot/\rho)$, où $\rho \in \mathbb{R} \setminus \{0\}$. Alors, $\text{prox}_\psi x = \rho \text{prox}_{\varphi/\rho^2}(x/\rho)$.*
- (iv) *Réflexion : Soit $\psi: y \mapsto \varphi(-y)$. Alors, $\text{prox}_\psi x = -\text{prox}_\varphi(-x)$.*
- (v) *Enveloppe de Moreau : Soit $\psi = \gamma\varphi$, où $\gamma \in]0, +\infty[$. Alors,*

$$\text{prox}_\psi x = x + \frac{1}{\gamma + 1} (\text{prox}_{(\gamma+1)\varphi} x - x). \tag{2.13}$$

- (vi) *Soit $\psi = \|\cdot\|^2/(2\gamma) - \gamma\varphi$, où $\gamma \in]0, +\infty[$ et $\varphi \in \Gamma_0(\mathcal{X})$, et soit $x \in \mathcal{X}$. Alors $\psi \in \Gamma_0(\mathcal{X})$ et*

$$\text{prox}_\psi x = x - \frac{1}{\gamma} \text{prox}_{\frac{\gamma^2}{\gamma+1}\varphi} \left(\frac{\gamma x}{\gamma + 1} \right). \tag{2.14}$$

Cette proposition a de nombreuses applications pour les fonctions réelles, c'est-à-dire, lorsque $\mathcal{X} = \mathbb{R}$. En effet, dans le Chapitre 4, par le biais de la Proposition 2.2.3(i), nous déterminons, sans calcul, les opérateurs proximaux des fonctions potentiels de densités de probabilité log-concaves, de type $\exp(-\phi)$:

(i) la distribution Gaussienne, i.e., pour $\tau \in]0, +\infty[$ fixé,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \tau|\xi|^2; \quad (2.15)$$

(ii) la distribution maximum d'entropie, obtenue en maximisant l'entropie dont les moments de premier et de second ordre sont connus, ainsi que celui d'ordre p avec $2 \neq p \in]1, +\infty[$: soient $\omega \in]0, +\infty[, \tau \in [0, +\infty[, \kappa \in]0, +\infty[$ fixés,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \omega|\xi| + \tau|\xi|^2 + \kappa|\xi|^p; \quad (2.16)$$

(iii) la distribution exponentielle, i.e., pour $\omega \in]0, +\infty[$ fixé,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} \omega\xi, & \text{si } \xi \geq 0; \\ +\infty, & \text{si } \xi < 0; \end{cases} \quad (2.17)$$

(iv) la distribution gamma, i.e., pour $\omega \in]0, +\infty[, \kappa \in]0, +\infty[$ fixés,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} -\kappa \ln(\xi) + \omega\xi, & \text{si } \xi > 0; \\ +\infty, & \text{si } \xi \leq 0; \end{cases} \quad (2.18)$$

(v) la distribution chi, i.e., pour $\kappa \in]0, +\infty[$ fixé,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} -\kappa \ln(\xi) + \xi^2/2, & \text{si } \xi > 0; \\ +\infty, & \text{si } \xi \leq 0. \end{cases} \quad (2.19)$$

De plus, de la Proposition 2.2.3(v), nous déduisons (voir la Section 3.3.2.6 du Chapitre 3) l'opérateur proximal de $d_C^2/(2\gamma)$, où C est un convexe fermé non vide de \mathcal{X} .

2.2.3 Composition et décomposition

Dans la Section 3.3.2 de l'article rapporté au Chapitre 3, nous démontrons la proposition suivante :

Proposition 2.2.4 *Soit $\psi = \varphi \circ L$ où $\varphi \in \Gamma_0(\mathcal{Y})$ et où $L \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ est bijectif avec $L^{-1} = L^*$. Alors, $\text{prox}_\psi = L^* \circ \text{prox}_\varphi \circ L$.*

De cette proposition, nous déduisons l'opérateur proximal d'une fonction décomposée selon une base hilbertienne de \mathcal{X} et dont les composantes sont toutes positives, de minimum nul, atteint en zéro. La démonstration est fournie dans l'Exemple 3.3.20 du Chapitre 3. L'article exposé au Chapitre 4 en propose une généralisation à la Section 4.4.2 :

Proposition 2.2.5 Soit $\Upsilon: \mathcal{X} \rightarrow]-\infty, +\infty]: x \mapsto \sum_{i \in \mathbb{I}} \psi_i(\langle x | o_i \rangle)$, où :

- (i) $\mathbb{I} \subset \mathbb{N}$;
- (ii) $(o_i)_{i \in \mathbb{I}}$ est une base hilbertienne de \mathcal{X} ;
- (iii) $(\psi_i)_{i \in \mathbb{I}}$ sont des fonctions dans $\Gamma_0(\mathbb{R})$;
- (iv) ou bien \mathbb{I} est fini, ou bien il existe un sous-ensemble \mathbb{J} de \mathbb{I} tel que :
 - (a) $\mathbb{I} \setminus \mathbb{J}$ est fini;
 - (b) $(\forall i \in \mathbb{J}) \psi_i \geq 0$;
 - (c) il existe une suite $(\zeta_i)_{i \in \mathbb{J}}$ de \mathbb{R} telle que :

$$\sum_{i \in \mathbb{J}} |\zeta_i|^2 < +\infty, \sum_{i \in \mathbb{J}} |\text{prox}_{\psi_i} \zeta_i|^2 < +\infty, \sum_{i \in \mathbb{J}} \psi_i(\zeta_i) < +\infty. \quad (2.20)$$

Alors $\Upsilon \in \Gamma_0(\mathcal{X})$ et $(\forall x \in \mathcal{X}) \text{prox}_{\Upsilon} x = \sum_{i \in \mathbb{I}} (\text{prox}_{\psi_i} \langle x | o_i \rangle) o_i$.

Cette proposition est l'outil théorique de base du Chapitre 4 pour analyser et résoudre les problèmes inverses, de régularisation parcimonieuse et par approche bayésienne, traités sur des trames.

Il est à noter que dans le cas particulier où l'espace de Hilbert réel d'étude est une somme hilbertienne directe, le résultat se montre directement et succinctement, comme exposé à la Section 3.3.2 du Chapitre 4 :

Proposition 2.2.6 Soient $(\phi_k)_{1 \leq k \leq m}$ des fonctions dans $\Gamma_0(\mathcal{X})$, \mathcal{X}^m la somme hilbertienne directe, et soit $\varphi: \mathcal{X}^m \rightarrow]-\infty, +\infty]: (x_k)_{1 \leq k \leq m} \mapsto \sum_{k=1}^m \phi_k(x_k)$. Alors, $\text{prox}_{\varphi} = (\text{prox}_{\phi_k})_{1 \leq k \leq m}$.

Cette propriété permet notamment de déterminer l'opérateur proximal de certaines entropies définies sur l'espace euclidien \mathbb{R}^m . Par exemple (voir l'Exemple 3.3.19 du Chapitre 3), il est possible de déterminer l'opérateur proximal de l'entropie de Burg à partir de celui de

$$\phi: \xi \mapsto \begin{cases} -\ln(\xi), & \text{si } \xi > 0; \\ +\infty, & \text{si } \xi \leq 0; \end{cases} \quad (2.21)$$

qui est connu de manière explicite. En effet, l'entropie de Burg est définie sur \mathbb{R}^m par $\varphi: x = (\xi_k)_{1 \leq k \leq m} \mapsto \sum_{k=1}^m \phi(\xi_k)$.

2.3 Le principe de décomposition de Moreau

Nous consacrons cette section au principe de décomposition de Moreau parce qu'il a une très grande portée scientifique. En effet, c'est non seulement une technique de calcul donnant, très succinctement, l'opérateur proximal d'une fonction à

partir de celui de sa conjuguée, mais aussi c'est un outil puissant contribuant à l'analyse des signaux en traitement du signal [9], par exemple, pour le débruitage. L'objectif de cette section est justement de mettre en valeur ces diverses facettes de la décomposition de Moreau [18], en traitement du signal, en exposant notamment les résultats de l'article intégré au Chapitre 3.

2.3.1 Énoncé

Le principe de décomposition de Moreau est une généralisation du théorème de projection orthogonale qui consiste à décomposer tout élément $x \in \mathcal{X}$ en la somme de son projeté selon un sous-espace vectoriel fermé V de \mathcal{X} et de son projeté selon l'orthogonal de V , noté V^\perp :

$$(\forall x \in \mathcal{X}) \quad \begin{cases} \|x\|^2 = d_V^2(x) + d_{V^\perp}^2(x), \\ x = P_V x + P_{V^\perp} x, \\ \langle P_V x \mid P_{V^\perp} x \rangle = 0. \end{cases} \quad (2.22)$$

En posant $\varphi = \iota_V$, comme $\varphi^* = \iota_{V^\perp}$, nous obtenons alors :

$$(\forall x \in \mathcal{X}) \quad \begin{cases} \|x\|^2 = 2(\langle \varphi(x) \mid \varphi^*(x) \rangle), \\ x = \text{prox}_\varphi x + \text{prox}_{\varphi^*} x, \\ \varphi(\text{prox}_\varphi x) + \varphi^*(\text{prox}_{\varphi^*} x) = \langle \text{prox}_\varphi x \mid \text{prox}_{\varphi^*} x \rangle. \end{cases} \quad (2.23)$$

Moreau a généralisé ce principe à toute fonction $\varphi \in \Gamma_0(\mathcal{X})$:

Proposition 2.3.1 *Soient $\varphi \in \Gamma_0(\mathcal{X})$, $\gamma \in]0, +\infty[$, et $x \in \mathcal{X}$. Alors*

$$\|x\|^2 = 2\gamma(\gamma\varphi(x) + \langle \varphi^*(x/\gamma) \mid x/\gamma \rangle) \quad (2.24)$$

et

$$x = x_\gamma^\oplus + x_\gamma^\ominus, \quad \text{où} \quad \begin{cases} x_\gamma^\oplus = \text{prox}_{\gamma\varphi} x \\ x_\gamma^\ominus = \gamma \text{prox}_{\varphi^*/\gamma}(x/\gamma). \end{cases} \quad (2.25)$$

De plus,

$$\varphi(x_\gamma^\oplus) + \varphi^*(x_\gamma^\ominus/\gamma) = \langle x_\gamma^\oplus \mid x_\gamma^\ominus \rangle / \gamma. \quad (2.26)$$

Une démonstration de ce principe est proposée à la Section 3.3.2.5 du Chapitre 3.

2.3.2 Une technique calculatoire

Un des intérêts du principe de décomposition de Moreau est de fournir directement l'opérateur proximal d'une fonction de $\Gamma_0(\mathcal{X})$ en fonction de celui de sa conjuguée. En pratique, détenir une valeur explicite de l'opérateur proximal d'une fonction de $\Gamma_0(\mathcal{X})$, permet de déduire, de façon systématique, celui de sa conjuguée. Cette méthode est d'autant plus intéressante que cette conjuguée a une expression trop lourde pour calculer son opérateur proximal directement. En pratique, nous utilisons comme technique de calcul l'écriture suivante de la Proposition 2.3.1.

Corollaire 2.3.2 *Soient $\varphi \in \Gamma_0(\mathcal{X})$, $\gamma \in]0, +\infty[$ et $x \in \mathcal{X}$. Alors*

$$\text{prox}_{\gamma\varphi^*} x = x - \gamma \text{prox}_{\varphi/\gamma}(x/\gamma). \quad (2.27)$$

Donnons maintenant des exemples illustrant son efficacité.

2.3.2.1 Fonction d'appui

Dans l'article du Chapitre 3, nous déterminons l'opérateur proximal de la fonction $\varphi: \mathcal{X} \rightarrow]-\infty, +\infty]$ de $\Gamma_0(\mathcal{X})$ définie par

$$\varphi: x \mapsto \sup_{y \in D} \langle x \mid Ly \rangle, \quad (2.28)$$

où $L \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ et D est un ensemble non vide de \mathcal{Y} .

Dans le cas particulier où C est l'enveloppe convexe fermée de $L(D)$, $\varphi = \sigma_C$. Du Corollaire 2.3.2, nous obtenons, pour tout $x \in \mathcal{X}$, et pour tout $\gamma \in]0, +\infty[$,

$$\text{prox}_{\gamma\varphi} x = x - P_{\gamma C}x. \quad (2.29)$$

Mis à part la comparaison, soulignée dans l'article, avec l'approche de [7] (cas particulier où φ est la variation totale discrète), il existe d'autres illustrations de cet exemple, entre autres le cas où φ est une perspective. La perspective, introduite dans [21], est une fonction approfondie dans [16] pour généraliser les fonctions barrières, outils des problèmes de programmation linéaire :

Définition 2.3.3 *Soit $\psi \in \Gamma_0(\mathbb{R}^m)$, $m \in \mathbb{N}^*$. La perspective φ associée à ψ est la fonction d'appui de l'hypographe $\text{hypo}(-\psi^*)$ de la fonction concave $-\psi^*$, i.e.,*

$$\varphi = \sigma_{\text{hypo}(-\psi^*)} \quad \text{où} \quad \text{hypo}(-\psi^*) = \{(y, \zeta) \in \mathbb{R}^{m+1} \mid \zeta \leq -\psi^*(y)\}. \quad (2.30)$$

De (2.29), nous déduisons que

$$(\forall (x, \xi) \in \mathbb{R}^{m+1}) \quad \text{prox}_{\gamma\varphi}(x, \xi) = (x, \xi) - P_{\gamma \text{hypo}(-\psi^*)}(x, \xi). \quad (2.31)$$

Une écriture plus classique de la fonction perspective est la suivante [12, Chapitre IV, Section 2.2] :

Proposition 2.3.4 Soit $\psi \in \Gamma_0(\mathbb{R}^m)$, $m \in \mathbb{N}^*$. La perspective associée à ψ , est la fonction $\varphi: \mathbb{R}^{m+1} \rightarrow]-\infty, +\infty]$ vérifiant :

$$(\forall x \in \mathbb{R}^m)(\forall \xi \in \mathbb{R}) \quad \varphi(x, \xi) = \begin{cases} \xi \psi\left(\frac{x}{\xi}\right), & \text{si } \xi > 0; \\ \psi 0^+(x), & \text{si } \xi = 0; \\ +\infty, & \text{si } \xi < 0; \end{cases} \quad (2.32)$$

où $\psi 0^+$ désigne la fonction de récession de ψ , i.e., la fonction d'appui de $\text{dom } \psi^*$:

$$\psi 0^+(x) = \sigma_{\text{dom } \psi^*}(x) = \sup_{y \in \text{dom } \psi^*} \langle x \mid y \rangle. \quad (2.33)$$

Remarquons que sous cette forme, le calcul de l'opérateur proximal nécessite une étude par cas assez longue.

Prenons maintenant un exemple de perspective :

Exemple 2.3.5 Soient $p \geq 1$ et $q \geq 1$ tels que $1/p + 1/q = 1$. Soit $\psi = \|\cdot\|^p$. La perspective φ associée à ψ , est une fonction de $\Gamma_0(\mathbb{R}^{m+1})$ telle que pour tout $\gamma \in]0, +\infty[$ et pour tout $(x, \xi) \in \mathbb{R}^{m+1}$,

$$\text{prox}_{\gamma\varphi}(x, \xi) = \begin{cases} (0, 0), & \text{si } \xi < 0 \text{ et } \|x\| \leq (-qp^{q-1}\xi)^{1/q}; \\ \left(\left(1 - \frac{(-q(\gamma p)^{q-1}\xi)^{1/q}}{\|x\|} \right) x, 0 \right), & \text{si } \xi < 0 \text{ et } \|x\| > (-qp^{q-1}\xi)^{1/q}; \\ (x, \xi), & \text{si } \xi \geq 0 \text{ et } x \in \mathbb{R}^m. \end{cases} \quad (2.34)$$

Démonstration. Déterminons, tout d'abord, $\text{hypo}(-\psi^*)$. D'après [4], $\psi^* = p\|\cdot/p\|^q/q$. Soit $(y, \zeta) \in \text{hypo}(-\psi^*)$. De la définition de $\text{hypo}(-\psi^*)$ par (2.30),

$$\zeta \leq -\frac{p\|y/p\|^q}{q} \leq 0 \quad \text{et} \quad y \in B(0; p(-q\zeta/p)^{1/q}). \quad (2.35)$$

Réciproquement, soit $(y, \zeta) \in \mathbb{R}^m \times \mathbb{R}_-$ tel que $y \in B(0; p(-q\zeta/p)^{1/q})$. Alors, $\zeta \leq -p\|y/p\|^q/q$ et $(y, \zeta) \in \text{hypo}(-\psi^*)$. D'où,

$$\text{hypo}(-\psi^*) = \{(y, \zeta) \in \mathbb{R}^m \times \mathbb{R}_- \mid y \in B(0; (-qp^{q-1}\zeta)^{1/q})\}. \quad (2.36)$$

Maintenant, d'après (2.31),

$$(\forall (x, \xi) \in \mathbb{R}^{m+1}) \quad \text{prox}_{\gamma\varphi}(x, \xi) = (x, \xi) - P_{\gamma \text{hypo}(-\psi^*)}(x, \xi). \quad (2.37)$$

Il suffit donc de déterminer une valeur explicite de $P_{\gamma \text{hypo}(-\psi^*)}$ pour obtenir celle de $\text{prox}_{\gamma\varphi}$. D'après (2.36),

$$\gamma \text{hypo}(-\psi^*) = \{(\gamma y, \gamma \zeta) \in \mathbb{R}^m \times \mathbb{R}_- \mid y \in B(0; (-qp^{q-1}\zeta)^{1/q})\}. \quad (2.38)$$

Remarquons que

$$\begin{aligned} (x, \xi) \in \gamma \text{ hypo}(-\psi^*) &\Leftrightarrow (x/\gamma, \xi/\gamma) \in \text{hypo}(-\psi^*) \\ &\Leftrightarrow x \in B(0; (-q(\gamma p)^{q-1}\xi)^{1/q}). \end{aligned} \quad (2.39)$$

Pour calculer $P_{\gamma \text{ hypo}(-\psi^*)}$, trois cas sont alors à distinguer.

- Supposons $(x, \xi) \in \gamma \text{ hypo}(-\psi^*)$. Alors $P_{\gamma \text{ hypo}(-\psi^*)}(x, \xi) = (x, \xi)$.
- Supposons $(x, \xi) \in \mathbb{R}^m \times \mathbb{R}_-$ et $(x, \xi) \notin \gamma \text{ hypo}(-\psi^*)$. Alors

$$\|x\| > (-q(\gamma p)^{q-1}\xi)^{1/q} \quad \text{et} \quad (z, \xi) \in \gamma \text{ hypo}(-\psi^*), \quad (2.40)$$

où z est défini par :

$$z = (-q(\gamma p)^{q-1}\xi)^{1/q} \frac{x}{\|x\|}. \quad (2.41)$$

De plus, pour tout $(y, \theta) \in \gamma \text{ hypo}(-\psi^*)$,

$$\|(y, \theta) - (x, \xi)\| \geq \|(y, \xi) - (x, \xi)\| \geq \|(z, \xi) - (x, \xi)\|. \quad (2.42)$$

Nous déduisons que (z, ξ) est le projeté de (x, ξ) sur $\gamma \text{ hypo}(-\psi^*)$.

- Supposons que $(x, \xi) \in \mathbb{R}^m \times \mathbb{R}_+$. Soit $(y, \theta) \in \gamma \text{ hypo}(-\psi^*)$. Alors (y, θ) vérifie :

$$\begin{aligned} \|(y, \theta) - (x, \xi)\|^{m+1} &\geq \|(y, 0) - (x, \xi)\|^{m+1} = \|y - x\|^{m+1} + \xi^{m+1} \\ &\geq \|x\|^{m+1} + \xi^{m+1} = \|(x, \xi)\|^{m+1}. \end{aligned} \quad (2.43)$$

Ainsi, $(0, 0)$ est le projeté de (x, ξ) sur $\gamma \text{ hypo}(-\psi^*)$.

□

2.3.2.2 Enveloppe de Pasch-Hausdorff

L'enveloppe de Pasch-Hausdorff est une convolution infimale particulière étudiée dans un cas plus général dans [2, 5]. Nous étudions l'opérateur proximal de l'enveloppe de Pasch-Hausdorff d'une fonction soumise à la

Condition 2.3.6 *La fonction $\varphi \in \Gamma_0(\mathcal{X})$ est telle que $\inf \varphi(\mathcal{X}) > -\infty$.*

Définition 2.3.7 *Soient $\gamma \in]0, +\infty[$ et φ une fonction vérifiant la Condition 2.3.6. L'enveloppe de Pasch-Hausdorff de φ , d'indice γ , est la convolution infimale, notée $\text{ph}_{\varphi, \gamma}$, définie par :*

$$(\forall x \in \mathcal{X}) \quad \text{ph}_{\varphi, \gamma}(x) = \left(\varphi \square \left(\frac{\|\cdot\|}{\gamma} \right) \right)(x) = \inf_{y \in \mathcal{X}} \varphi(y) + \frac{\|y - x\|}{\gamma}. \quad (2.44)$$

Proposition 2.3.8 Soient $(\gamma, \mu) \in]0, +\infty[^2$ et φ vérifiant la Condition 2.3.6. Alors les propriétés suivantes sont vérifiées :

- (i) $\text{ph}_{\varphi, \gamma} \in \Gamma_0(\mathcal{X})$.
- (ii) Pour tout $x \in \mathcal{X}$, en notant $\bar{x} = \text{argmin}_{B(0; \mu/\gamma)} \psi$ où $\psi = \mu\varphi^*(\cdot/\mu) + \|\cdot - x\|^2/2$,

$$\text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = x - \mu \text{prox}_{\psi} x = x - \mu \bar{x}. \quad (2.45)$$

- (iii) $(\forall x \in \mathcal{X}) \quad \|x - \text{prox}_{\mu\varphi} x\| \leq \frac{\mu^2}{\gamma} \Leftrightarrow \text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = \text{prox}_{\mu\varphi} x$.

Démonstration. (i) : C'est une conséquence de [13, Corollaire 6.5.3] et de [13, Proposition 6.5.5], $\|\cdot\|$ étant bornée, et φ vérifiant la Condition 2.3.6.

(ii) : La conjuguée est obtenue par [13, Théorème 6.5.2], sachant que $(\|\cdot\|/\gamma)^* = \iota_{B(0; 1/\gamma)}$, $\text{ph}_{\varphi, \gamma}^* = \varphi^* + (\|\cdot\|/\gamma)^*$. Par le Corollaire 2.3.2, comme pour tout $x \in \mathcal{X}$

$$\text{prox}_{\varphi^*/\mu}(x/\mu) = \text{argmin}_{y \in B(0; 1/\gamma)} \frac{1}{\mu} \varphi^*(y) + \|y - \frac{x}{\mu}\|^2 = \bar{x}, \quad (2.46)$$

nous obtenons l'expression (2.45) de l'opérateur proximal de $\mu \text{ph}_{\varphi, \gamma}$.

(iii) : Soit $x \in \mathcal{X}$. Supposons que $\|x - \text{prox}_{\mu\varphi} x\| \leq \mu^2/\gamma$. D'après le Corollaire 2.3.2,

$$\|\text{prox}_{\varphi^*/\mu}(x/\mu)\| = \frac{1}{\mu} \|x - \text{prox}_{\mu\varphi} x\| \leq \frac{\mu}{\gamma}. \quad (2.47)$$

Comme, de plus, $\text{prox}_{\varphi^*/\mu}(x/\mu)$ minimise ψ sur \mathcal{X} , il vient $\bar{x} = \text{prox}_{\varphi^*/\mu}(x/\mu)$ puis $\text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = \text{prox}_{\mu\varphi} x$. Réciproquement, si $\text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = \text{prox}_{\mu\varphi} x$, d'après (2.45),

$$\|x - \text{prox}_{\mu\varphi} x\| = \mu \|\bar{x}\| \leq \mu^2/\gamma. \quad (2.48)$$

□

Donnons maintenant des exemples de fonctions φ , vérifiant la Condition 2.3.6, et dont l'enveloppe de Pasch-Hausdorff a un opérateur proximal qui se détermine de manière explicite sur \mathcal{X} tout entier.

Exemple 2.3.9 Soient $(\mu, \gamma) \in]0, +\infty[^2$ et C un ensemble convexe fermé non vide de \mathcal{X} .

- (i) Si $\varphi = \iota_C$, alors $\text{ph}_{\varphi, \gamma} = d_C$. L'opérateur proximal est calculé au Chapitre 3, dans la Section 3.3.2.6 : pour tout $x \in \mathcal{X}$,

$$\text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = \begin{cases} x + \frac{\mu}{d_C(x)} (P_C x - x), & \text{si } d_C(x) > \mu; \\ P_C x, & \text{si } d_C(x) \leq \mu. \end{cases} \quad (2.49)$$

- (ii) Si $\varphi = \sigma_C$, alors $\text{ph}_{\varphi, \gamma}^* = \iota_C + \iota_{B(0;1/\gamma)} = \iota_{C \cap B(0;1/\gamma)}$, et d'après (2.5), pour tout $x \in \mathcal{X}$,

$$\text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = x - \mu P_{C \cap B(0; \mu/\gamma)} x. \quad (2.50)$$

Exemple 2.3.10 Soit $(\mu, \gamma) \in]0, +\infty[^2$. Posons $\varphi = \|\cdot\|^2/2$. Alors,

$$(\forall x \in \mathcal{X}) \quad \text{prox}_{\mu \text{ph}_{\varphi, \gamma}} x = x - \mu P_{B(0; \mu/\gamma)} \left(\frac{x}{2} \right). \quad (2.51)$$

Démonstration. La fonction $\text{ph}_{\varphi, \gamma}^* = \|\cdot\|^2/2 + \iota_{B(0;1/\gamma)}$ est une perturbation quadratique de $\iota_{B(0;1/\gamma)}$. La conclusion vient de la Proposition 2.2.3(i), avec $\alpha = 1$ et $u = 0$, et de la Proposition 2.3.8. \square

Exemple 2.3.11 (Fonctions réelles.) Soient $(\mu, \gamma) \in]0, +\infty[^2$ et $\phi: \mathbb{R} \rightarrow \mathbb{R}$ vérifiant la Condition 2.3.6 avec $\mathcal{X} = \mathbb{R}$. Alors, pour tout $\xi \in \mathbb{R}$,

$$\text{prox}_{\mu \text{ph}_{\phi, \gamma}} \xi = \text{prox}_{\mu \phi} \xi + \text{soft}_{[-\mu^2/\gamma, \mu^2/\gamma]} (\xi - \text{prox}_{\mu \phi} \xi), \quad (2.52)$$

où le seuillage doux sur $[-\mu^2/\gamma, \mu^2/\gamma]$ est défini par

$$\text{soft}_{[-\mu^2/\gamma, \mu^2/\gamma]} : \xi \mapsto \begin{cases} \xi + \frac{\mu^2}{\gamma}, & \text{si } \xi < -\frac{\mu^2}{\gamma}; \\ 0, & \text{si } \xi \in \left[-\frac{\mu^2}{\gamma}, \frac{\mu^2}{\gamma} \right]; \\ \xi - \frac{\mu^2}{\gamma}, & \text{si } \xi > \frac{\mu^2}{\gamma}. \end{cases} \quad (2.53)$$

Nous donnons une démonstration où le lemme élémentaire qui suit est nécessaire.

Lemme 2.3.12

- (i) Soit ψ une fonction convexe admettant un minimum en $\xi_0 \in \mathbb{R}$. Alors ψ est décroissante sur $] -\infty, \xi_0]$ et croissante sur $[\xi_0, +\infty[$.
- (ii) Soient $\gamma \in]0, +\infty[$, $\phi \in \Gamma_0(\mathbb{R})$ et $\xi \in \mathbb{R}$. La fonction $\gamma\phi + |\cdot - \xi|^2/2$ est décroissante sur $] -\infty, \text{prox}_{\gamma\phi} \xi]$ et croissante sur $[\text{prox}_{\gamma\phi} \xi, +\infty[$.

Démonstration. Soit $\xi \in \mathbb{R}$. De la Proposition 2.3.8(iii), si $|\xi - \text{prox}_{\mu\phi} \xi| \leq \mu^2/\gamma$, $\text{prox}_{\mu \text{ph}_{\phi, \gamma}} \xi = \text{prox}_{\mu\phi} \xi$.

Maintenant, supposons que $|\xi - \text{prox}_{\mu\phi} \xi| > \mu^2/\gamma$, par exemple, $\text{prox}_{\phi^*/\mu} \xi > \mu/\gamma$. Comme $\phi^*/\mu + |\cdot - \xi|^2/2$ est convexe sur \mathbb{R} et atteint son minimum en $\text{prox}_{\phi^*/\mu} \xi$, d'après le lemme ci-dessus, elle est décroissante sur $] -\infty, \text{prox}_{\phi^*/\mu} \xi]$, puis μ/γ minimise $\phi^*/\mu + |\cdot - \xi|^2/2$ sur $[-\mu/\gamma, \mu/\gamma]$. Si $\text{prox}_{\phi^*/\mu} \xi < -\mu/\gamma$, le même raisonnement

s'applique. Nous déduisons de cette étude de cas :

$$\text{prox}_{\mu \text{ph}_{\phi, \gamma}} \xi = \begin{cases} \xi + \frac{\mu^2}{\gamma}, & \text{si } \xi - \text{prox}_{\mu\phi} \xi < -\frac{\mu^2}{\gamma}; \\ \text{prox}_{\mu\phi} \xi, & \text{si } |\xi - \text{prox}_{\mu\phi} \xi| \leq \frac{\mu^2}{\gamma}; \\ \xi - \frac{\mu^2}{\gamma}, & \text{si } \xi - \text{prox}_{\mu\phi} \xi > \frac{\mu^2}{\gamma}. \end{cases} \quad (2.54)$$

□

2.3.2.3 Fonctions réelles

Nous débutons cette section par quelques propriétés propres au fait que $\mathcal{X} = \mathbb{R}$, permettant de simplifier les calculs déterminant l'opérateur proximal. Le Lemme 2.3.13 et la Remarque 2.3.14 sont détaillés à la Section 4.4.4 du Chapitre 4.

Lemme 2.3.13 [10, Section 2] *Soit $\phi \in \Gamma_0(\mathbb{R})$.*

- (i) *La fonction $\text{prox}_{\phi} : \mathbb{R} \rightarrow \mathbb{R}$ est croissante.*
- (ii) *Supposons que ϕ admet 0 comme minimiseur. Alors,*

$$(\forall \xi \in \mathbb{R}) \quad \begin{cases} 0 \leq \text{prox}_{\phi} \xi \leq \xi, & \text{si } \xi > 0; \\ \text{prox}_{\phi} \xi = 0, & \text{si } \xi = 0; \\ \xi \leq \text{prox}_{\phi} \xi \leq 0, & \text{si } \xi < 0. \end{cases} \quad (2.55)$$

Ceci s'applique, en particulier, quand ϕ est paire.

Remarque 2.3.14 *Soit $\phi \in \Gamma_0(\mathbb{R})$.*

- (i) $(\forall \xi \in \mathbb{R}) \phi(\text{prox}_{\phi} \xi) < +\infty$.
- (ii) Si ϕ est paire, alors prox_{ϕ} est impaire. Dans un tel cas, il suffit de déterminer $\text{prox}_{\phi} \xi$ pour $\xi \geq 0$ et d'étendre le résultat à $\xi < 0$ par antisymétrie.
- (iii) Soit $\xi \in \mathbb{R}$. Si ϕ est différentiable en $\text{prox}_{\phi} \xi$, alors

$$(\forall \pi \in \mathbb{R}) \quad \pi = \text{prox}_{\phi} \xi \quad \Leftrightarrow \quad \pi + \phi'(\pi) = \xi. \quad (2.56)$$

Du Lemme 2.3.13, de la Remarque 2.3.14 et du Corollaire 2.3.2, nous déduisons à la Section 4.4.4 du Chapitre 4, une expression explicite des opérateurs proximaux des fonctions potentiels de densités de probabilité log-concaves, de type $\exp(-\phi)$:

- (i) la distribution de Laplace, i.e., pour $\omega \in]0, +\infty[$ fixé,

$$\phi : \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \omega|\xi|; \quad (2.57)$$

(ii) la distribution gaussienne généralisée, i.e., pour $\kappa \in]0, +\infty[$ et $p \in \{4/3, 3/2, 3, 4\}$ fixés,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \kappa|\xi|^p; \quad (2.58)$$

(iii) la distribution de Laplace lissée, i.e., pour $\omega \in]0, +\infty[$ fixé,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \omega|\xi| - \ln(1 + \omega|\xi|); \quad (2.59)$$

(iv) la distribution de Huber, i.e., pour $\omega \in]0, +\infty[$ et $\tau \in]0, +\infty[$ fixés,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} \tau\xi^2, & \text{si } |\xi| \leq \omega/\sqrt{2\tau}; \\ \omega\sqrt{2\tau}|\xi| - \omega^2/2, & \text{sinon;} \end{cases} \quad (2.60)$$

(v) la distribution triangulaire, i.e., pour $\underline{\omega} \in]-\infty, 0[$ et $\bar{\omega} \in]0, +\infty[$ fixés,

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} -\ln(\xi - \underline{\omega}) + \ln(-\underline{\omega}), & \text{si } \xi \in]\underline{\omega}, 0]; \\ -\ln(\bar{\omega} - \xi) + \ln(\bar{\omega}), & \text{si } \xi \in]0, \bar{\omega}]; \\ +\infty, & \text{sinon.} \end{cases} \quad (2.61)$$

Donnons un autre exemple de fonctions réelles dont l'opérateur proximal se calcule à l'aide du Corollaire 2.3.2.

Exemple 2.3.15 Soit $\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \kappa|\xi|^p$, où $p \in \{5, 5/4\}$ et $\kappa \in]0, +\infty[$. Pour tout $\xi \in \mathbb{R}$,

(i) pour $p = 5$:

$$\text{prox}_\phi \xi = \text{sign}(\xi) \frac{-\sqrt{\chi} + \sqrt{-\chi + 2\sqrt{\chi^2 + \frac{4|\xi|}{5\kappa}}}}{2} \quad (2.62)$$

où,

$$\chi = \frac{1}{(50\kappa^2)^{1/3}} \left(\left(1 + \sqrt{1 + \frac{1280\kappa|\xi|^3}{27}} \right)^{1/3} + \left(1 - \sqrt{1 + \frac{1280\kappa|\xi|^3}{27}} \right)^{1/3} \right). \quad (2.63)$$

(ii) pour $p = 5/4$:

$$\text{prox}_\phi \xi = \xi - \kappa \text{sign}(\xi) \frac{-\sqrt{\chi} + \sqrt{-\chi + 2\sqrt{\chi^2 + \frac{4|\xi|}{5}}}}{2} \quad (2.64)$$

où,

$$\chi = \frac{\kappa^{2/3}}{50^{1/3}} \left(\left(1 + \sqrt{1 + \frac{1280|\xi|^3}{27\kappa^4}} \right)^{1/3} + \left(1 - \sqrt{1 + \frac{1280|\xi|^3}{27\kappa^4}} \right)^{1/3} \right). \quad (2.65)$$

Démonstration. (i) : Soient $\xi \in \mathbb{R}$ et $\pi = \text{prox}_\phi \xi$. De la Remarque 2.3.14(ii), comme ϕ est paire, il suffit de traiter le cas $\xi \geq 0$. Comme ϕ est différentiable, il suit de (2.55) et (2.56) que π est l'unique solution positive telle que

$$\pi + 5\kappa\pi^4 = \xi. \quad (2.66)$$

Cette solution est racine réelle positive de $P = 5\kappa X^4 + X - \xi$. Comme P admet seulement deux racines réelles de signe opposé, P se factorise sous la forme :

$$P = 5\kappa(X^2 + aX + b)(X^2 + cX + d). \quad (2.67)$$

où $(a, b, c, d) \in \mathbb{R}^4$ vérifie $\delta = a^2 - 4b \geq 0$ et $c^2 - 4d < 0$. Par identification, (a, b, c, d) est l'unique solution du système

$$\begin{cases} 5\kappa(a + c) = 0 \\ 5\kappa(b + d + ac) = 0 \\ 5\kappa(bc + ad) = 1 \\ 5\kappa bd = -\xi. \end{cases} \Leftrightarrow \begin{cases} c = -a \\ b + d - a^2 = 0 \\ a(d - b) = 1/(5\kappa) \\ 5\kappa bd = -\xi. \end{cases} \Leftrightarrow \begin{cases} c = -a \\ b = (a^2 - 1/(5\kappa a))/2 \\ d = (a^2 + 1/(5\kappa a))/2 \\ 5\kappa(a^4 - 1/(25\kappa^2 a^2)) = -4\xi. \end{cases} \quad (2.68)$$

Notons que si $a < 0$, alors $\delta = -a^2 + 2/(5\kappa a) < 0$. Nous en déduisons que $a > 0$. De là, la racine positive de P est :

$$\pi = \frac{-a + \sqrt{\delta}}{2} \quad \text{avec} \quad \delta = -a^2 + 2\sqrt{a^4 + \frac{4\xi}{5\kappa}}. \quad (2.69)$$

Pour conclure, tout revient à résoudre :

$$5\kappa\left(a^4 - \frac{1}{25\kappa^2 a^2}\right) = -4\xi \quad \Leftrightarrow \quad 25\kappa^2 a^6 + 20\kappa\xi a^2 - 1 = 0. \quad (2.70)$$

En posant $\chi = a^2$, cette dernière équation équivaut à

$$\begin{cases} \chi^3 + \frac{4\xi}{5\kappa}\chi - \frac{1}{25\kappa^2} = 0, \\ \chi \geq 0. \end{cases} \quad (2.71)$$

Cherchons pour cela la solution de (2.71), $\chi \geq 0$, de la forme $u + v$ avec $uv = -4\xi/(15\kappa)$:

$$\begin{cases} uv = -\frac{4}{15\kappa}\xi \\ 25\kappa^2(u^3 + v^3) = 1 \\ u + v \geq 0. \end{cases} \quad (2.72)$$

Alors u^3 et v^3 sont racines du polynôme

$$25\kappa^2 T^2 - T - \frac{4^3 \xi^3}{3^3 5 \kappa} \quad \text{de discriminant} \quad 1 + \frac{1280 \kappa \xi^3}{27}. \quad (2.73)$$

D'où,

$$\chi = \frac{1}{(50\kappa^2)^{1/3}} \left(\left(1 + \sqrt{1 + \frac{1280 \kappa \xi^3}{27}} \right)^{1/3} + \left(1 - \sqrt{1 + \frac{1280 \kappa \xi^3}{27}} \right)^{1/3} \right), \quad (2.74)$$

et de (2.69),

$$\pi = \frac{-\sqrt{\chi} + \sqrt{-\chi + 2\sqrt{\chi^2 + \frac{4\xi}{5\kappa}}}}{2}. \quad (2.75)$$

(ii) : Soient $\xi \in \mathbb{R}$ et $\pi = \text{prox}_\phi \xi$. D'après le Corollaire 2.3.2

$$\pi = \xi - \kappa \text{prox}_{\phi^*/\kappa}(\xi/\kappa). \quad (2.76)$$

□

2.3.3 Un outil d'analyse de signaux

Le second intérêt de la décomposition proximal est de fournir un outil d'analyse des signaux [9]. Concrètement, dans la Proposition 2.3.1, x^\oplus représente une partie du signal possédant les caractéristiques recherchées et x^\ominus la partie résiduelle correspondante. L'opérateur proximal joue en quelque sorte le rôle de filtre.

2.3.3.1 Cas des projecteurs

L'équation (2.5) de la Section 2.1 montre que les projecteurs sont des opérateurs proximaux particuliers : si $\varphi = \iota_C$, où C est une partie convexe fermée non vide d'un espace de Hilbert \mathcal{H} , alors $\text{prox}_\varphi = P_C$.

Tout d'abord, si $C = V$ est un sous-espace vectoriel de \mathcal{H} , nous retrouvons les applications à la décomposition d'un signal x de $L^2(\Omega)$ suivant les basses fréquences (x^\oplus) et hautes fréquences (x^\ominus), en choisissant

$$V = \{z \in L^2(\mathbb{R}) \mid \hat{z} 1_B = 0\}, \quad (2.77)$$

où

- (i) $B = [-b/2, b/2]$, avec $b \in]0, +\infty[$, est une bande de fréquences ;

(ii) pour tout $z \in L^2(\mathbb{R})$, $\hat{z} \in L^2(\mathbb{R})$ est sa transformée de Fourier définie par

$$(\forall \xi \in \mathbb{R}) \quad \hat{z}(\xi) = \int_{-\infty}^{+\infty} z(t) e^{-i2\pi\xi t} dt; \quad (2.78)$$

(iii) $1_{\mathcal{C}B}$ est la fonction caractéristique de $\mathcal{C}B$ définie sur \mathbb{R} par

$$(\forall \xi \in \mathbb{R}) \quad 1_{\mathcal{C}B}(\xi) = \begin{cases} 0, & \text{si } \xi \in B; \\ 1, & \text{si } \xi \in \mathcal{C}B. \end{cases} \quad (2.79)$$

Nous obtenons ainsi la décomposition

$$\begin{cases} x = x^\oplus + x^\ominus, \\ x^\oplus = P_V x, \\ x^\ominus = P_{V^\perp} x, \end{cases} \quad (2.80)$$

où V^\perp est l'orthogonal de V .

Les Figures 2.2 et 2.3 illustrent une telle décomposition de l'image satellitaire \bar{x} de $\mathcal{H} = \mathbb{R}^{256 \times 256}$ de la Figure 2.1.

Plus encore, si $C = K$ est un cône convexe fermé non vide de \mathcal{H} (i.e., $K + K \subset K$ et $(\forall \alpha \in]0, +\infty[) \alpha K \subset K$), alors $\varphi^* = \iota_{K^\ominus}$, où

$$K^\ominus = \{u \in \mathcal{X} \mid (\forall x \in K) \langle x \mid u \rangle \leq 0\} \quad (2.81)$$

est le cône polaire de K . Le Corollaire 2.3.2 se formule dans ce cas

$$\begin{cases} \|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x), \\ x = P_K x + P_{K^\ominus} x, \\ \langle P_K x \mid P_{K^\ominus} x \rangle = 0. \end{cases} \quad (2.82)$$

De telles décompositions coniques ont une large portée en traitement du signal [6, 9, 23, 25]. Par exemple, en holographie pour reconstruire les signaux dont la phase de Fourier est connue, ou en imagerie pour les signaux positifs.

2.3.3.2 Seuillage

Le but de cette section est d'exposer des opérateurs proximaux qui sont des seuilleurs. L'Exemple 2.1.6 montre que le seuillage dur n'est pas un opérateur proximal. Qu'en est-il du seuillage doux ? L'Exemple 2.3.9 donne un exemple de seuillage : l'opérateur proximal de d_C , où C est un ensemble convexe fermé non vide, présente un seuil de proximité γ déterminé en fonction de la contrainte imposée et modélisée

par C . En particulier, si $C = \{0\}$, $d_C = \|\cdot\|$, l'opérateur proximal correspond à un seuillage sur l'énergie :

$$\text{prox}_{\gamma\|\cdot\|} x = \begin{cases} \left(1 - \frac{\gamma}{\|x\|}\right)x, & \text{si } \|x\| > \gamma; \\ 0, & \text{si } \|x\| \leq \gamma. \end{cases} \quad (2.83)$$

Si $\mathcal{X} = \mathbb{R}$, nous retrouvons le seuillage doux scalaire

$$\text{prox}_{\gamma|\cdot|} x = \text{sign}(x) \max\{|x| - \gamma, 0\}. \quad (2.84)$$

De manière générale, les seuillages doux scalaires utilisés dans les méthodes itératives sont caractérisés par la proposition suivante :

Proposition 2.3.16 [10, Section 2] *Supposons que $\phi = \psi + \sigma_\Omega$, où $\psi \in \Gamma_0(\mathbb{R})$ est différentiable en 0 avec $\psi'(0) = 0$, et où $\Omega \subset \mathbb{R}$ est un intervalle fermé non vide tel que $\underline{\omega} = \inf \Omega$ et $\bar{\omega} = \sup \Omega$. Alors $\text{prox}_\phi = \text{prox}_\psi \circ \text{soft}_\Omega$, où soft_Ω est le seuillage doux sur Ω défini par*

$$\text{soft}_{[\underline{\omega}, \bar{\omega}]} : \xi \mapsto \begin{cases} \xi - \underline{\omega}, & \text{si } \xi < \underline{\omega}; \\ 0, & \text{si } \underline{\omega} \leq \xi \leq \bar{\omega}; \\ \xi - \bar{\omega}, & \text{si } \xi > \bar{\omega}. \end{cases} \quad (2.85)$$

Cette proposition permet de plus de déduire l'opérateur proximal de la fonction potentiel de la distribution du maximum d'entropie, dans la Section 4.4.4 du Chapitre 4.

2.3.3.3 Simulations numériques

Le but de cette section est de fournir des illustrations de nos résultats théoriques obtenus et du principe de décomposition en tant qu'outil d'analyse des images. Nous décrivons tout d'abord l'analyse multi-résolution M -bande de $L^2(\mathbb{R})$, que nous utilisons pour obtenir les résultats exposés ensuite.

2.3.3.3.1 Analyse M -bande d'un signal L'analyse multi-résolution M -bande est une extension de l'analyse multi-résolution classique. Cette dernière consiste à construire une suite d'approximations d'un signal x donné. Pour ce faire, on introduit une suite décroissante de sous-espaces vectoriels fermés $(V_j)_{j \in \mathbb{Z}}$, possédant certaines propriétés [15]. Pour tout $j \in \mathbb{Z}$, l'approximation au niveau de résolution j est définie par $P_{V_j}x$. Le « détail » du signal perdu entre deux résolutions j et $j + 1$, par le théorème de projection orthogonale, correspond au projeté de x sur $W_{j+1} = V_{j+1}^\perp$.

L'analyse multi-résolution M - bande, où $M \in \mathbb{N}$ est supérieur ou égal à 2, consiste alors à définir pour chaque niveau de résolution $j \in \mathbb{Z}$, non pas les sous-espaces vectoriels V_j et W_j , mais les sous-espaces vectoriels V_j et $(W_j^m)_{1 \leq m < M}$ tels que

$$V_{j-1} = V_j \oplus \bigoplus_{m=1}^{M-1} W_j^m. \quad (2.86)$$

La décomposition du signal en ondelettes M -bandes est obtenue en décomposant le signal sur chacun des sous-espaces vectoriels V_j et W_j^m , sachant que V_j et W_j^m admettent des bases orthonormales fonctions respectivement d'une fonction échelle ou ondelette père $\psi_0 \in L^2(\mathbb{R})$ et d'une ondelette mère $\psi_m \in L^2(\mathbb{R})$ [24]. Rappelons, à ce propos, qu'une ondelette est une fonction $\psi \in L^2(\mathbb{R})$ (à valeurs réelles dans notre étude) de moyenne nulle, i.e.,

$$\int_{-\infty}^{+\infty} \psi(\xi) d\xi = 0. \quad (2.87)$$

De plus, par dilatation d'un facteur d'échelle $e \in]0, +\infty[$ et par translation d'un paramètre $\theta \in \mathbb{R}$, la fonction

$$\psi_{\theta,e}: \mathbb{R} \rightarrow \mathbb{R}: \xi \mapsto \frac{1}{\sqrt{e}} \psi\left(\frac{\xi - \theta}{e}\right) \quad (2.88)$$

est encore une ondelette.

2.3.3.3.2 Résultats Les simulations numériques ont été effectuées en collaboration avec Caroline Chau. L'espace de Hilbert est $\mathcal{H} = \mathbb{R}^{256 \times 256}$. L'image traitée $\bar{x} \in \mathcal{H}$ est l'image satellitaire de la Figure 2.1. Pour obtenir l'analyse de cette image, nous la décomposons en ondelettes M -bande ($M = 4$ bandes, 2 niveaux de résolution) puis nous lui appliquons $\text{prox}_{\gamma\varphi}$, pour différents $\gamma \in]0, +\infty[$ et différentes fonctions φ vérifiant les hypothèses suivantes.

Hypothèse 2.3.17

$$\varphi: \mathcal{H} \rightarrow]-\infty, +\infty]: x \mapsto \sum_{k \in \mathbb{K}} \phi_k(\langle x | e_k \rangle), \quad (2.89)$$

où :

- (i) $\mathbb{K} \subset \mathbb{N}$;
- (ii) $(e_k)_{k \in \mathbb{K}}$ est une base hilbertienne de \mathcal{H} ;
- (iii) $(\forall k \in \mathbb{K}) \phi_k = \phi \in \Gamma_0(\mathbb{R})$;
- (iv) il existe $\xi_0 \in \mathbb{R}$ tel que $\phi \geq \phi(\xi_0) = 0$.

Corollaire 2.3.18 Soit φ une fonction soumise à l'Hypothèse 2.3.17. Alors $\varphi \in \Gamma_0(\mathcal{H})$ et les propriétés suivantes sont vérifiées :

- (i) $(\forall x \in \mathcal{H}) \quad \text{prox}_\varphi x = \sum_{k \in \mathbb{K}} (\text{prox}_\phi \langle x | e_k \rangle) e_k.$
- (ii) $(\forall x \in \mathcal{H}) \quad \text{prox}_{\varphi^*} x = \sum_{k \in \mathbb{K}} (\text{prox}_{\phi^*} \langle x | e_k \rangle) e_k.$

Démonstration. (i) : Il suffit d'appliquer la Proposition 2.2.5 dans la base hilbertienne $(e_k)_{k \in \mathbb{K}}$, avec pour tout $i \in \mathbb{I}$, $\psi_i = \phi - \phi(\xi_0)$.

(ii) : Soit $x \in \mathcal{H}$. Pour tout $k \in \mathbb{K}$, d'après le Corollaire 2.3.2,

$$\text{prox}_{\phi^*} \langle x | e_k \rangle = \langle x | e_k \rangle - \text{prox}_\phi \langle x | e_k \rangle. \quad (2.90)$$

Puis,

$$\sum_{k \in \mathbb{K}} (\text{prox}_{\phi^*} \langle x | e_k \rangle) e_k = x - \sum_{k \in \mathbb{K}} (\text{prox}_\phi \langle x | e_k \rangle) e_k = x - \text{prox}_\varphi x \quad (2.91)$$

La conclusion provient à nouveau du Corollaire 2.3.2. \square

Comme expliqué au paragraphe précédent, nous décomposons en fait l'image sur chacun des sous-espaces vectoriels V_j et W_j^m , avec $j \in \{1, 2\}$ et $m \in \{1, 2, 3\}$, selon $(e_k)_{k \in \mathbb{K}}$ le banc de filtres d'Alkin et Caglar 4-bandes [1]. L'intérêt du principe de décomposition de Moreau est de fournir, en une unique équation, pour un paramètre $\gamma \in]0, +\infty[$ fixé, à partir du signal original \bar{x} , le signal transformé \bar{x}_γ^\oplus et le signal résiduel \bar{x}_γ^\ominus , selon l'écriture suivante : pour tout φ vérifiant l'Hypothèse 2.3.17, d'après le Corollaire 2.3.18,

$$\begin{cases} \bar{x} = \bar{x}_\gamma^\oplus + \bar{x}_\gamma^\ominus, \\ \bar{x} = \sum_{k \in \mathbb{K}} \langle \bar{x} | e_k \rangle e_k = \sum_{k \in \mathbb{K}} \bar{\xi}_k e_k, \\ \bar{x}_\gamma^\oplus = \text{prox}_{\gamma\varphi} \bar{x} = \sum_{k \in \mathbb{K}} (\text{prox}_{\gamma\phi} \bar{\xi}_k) e_k, \\ \bar{x}_\gamma^\ominus = \gamma \text{prox}_{\varphi^*/\gamma}(\bar{x}/\gamma) = \gamma \sum_{k \in \mathbb{K}} (\text{prox}_{\phi^*/\gamma}(\bar{\xi}_k/\gamma)) e_k. \end{cases} \quad (2.92)$$

Divers exemples décompositions proximales sont donnés dans les Figures 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11 et 2.12 qui permettent d'évaluer les divers types d'analyses possibles.

2.4 Bilan

Ce chapitre vient mettre en valeur l'opérateur proximal, véritable moteur de cette thèse, établissant, par ses puissantes propriétés, des interactions entre les différents chapitres, sous divers angles d'étude :

- d'un point de vue théorique, c'est une technique de calcul donnant, pour certaines fonctions utiles en traitement du signal, une expression explicite. Lorsque $\mathcal{X} = \mathbb{R}$, en particulier, nous obtenons des résultats exacts explicites ou numériques simples ;



FIG. 2.1 – Image originale 256×256 SPOT5, \bar{x} .

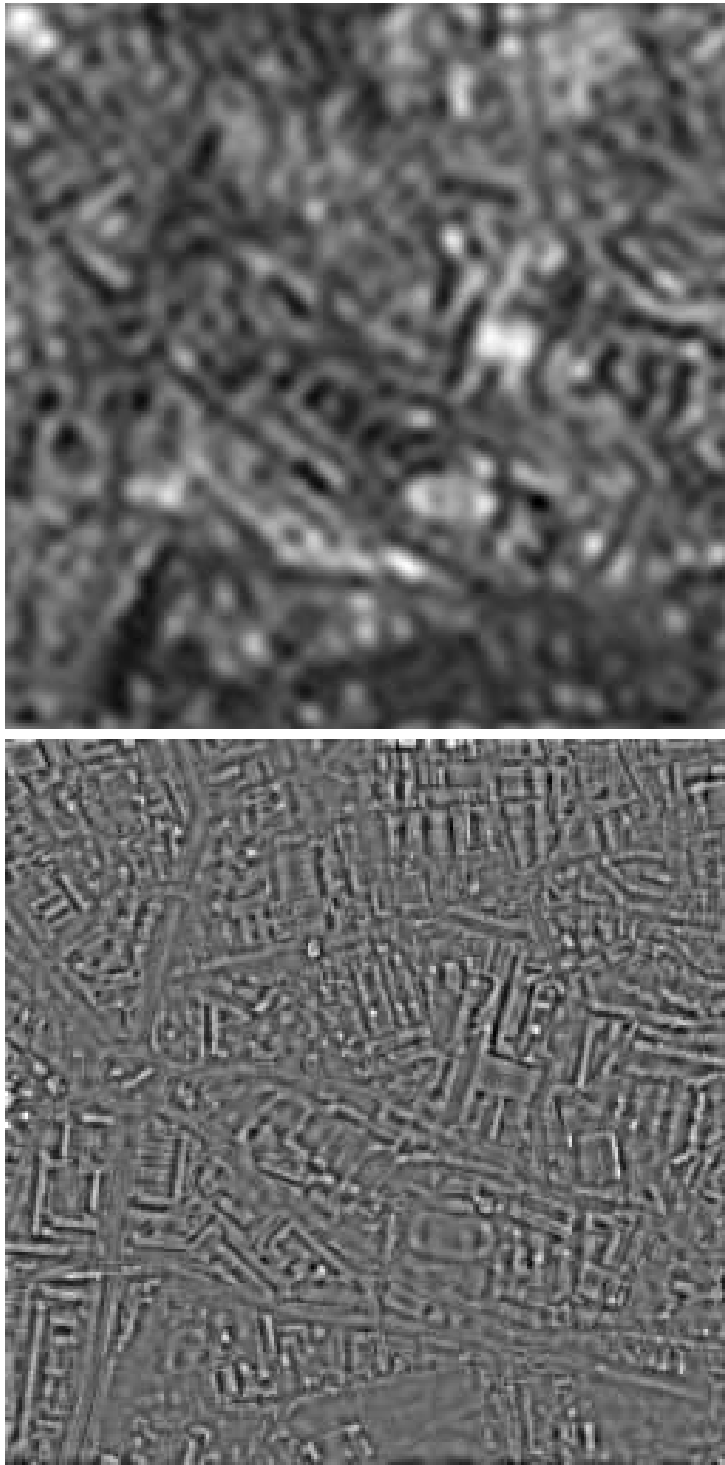


FIG. 2.2 – Décomposition suivant les basses fréquences \bar{x}_γ^\oplus , en haut, et les hautes fréquences \bar{x}_γ^\ominus , en bas, selon (2.80), avec $b = 50$.

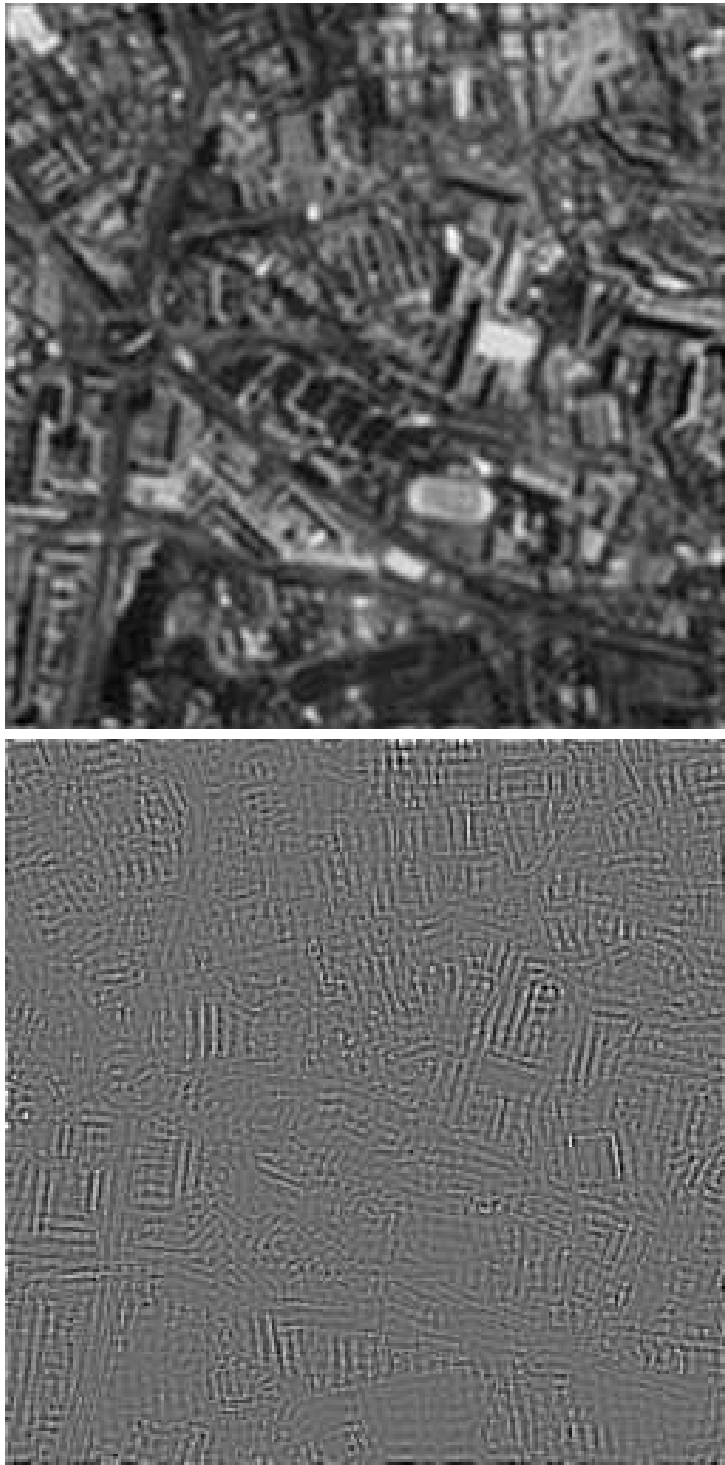


FIG. 2.3 – Décomposition suivant les basses fréquences \bar{x}_γ^\oplus , en haut, et les hautes fréquences \bar{x}_γ^\ominus , en bas, selon (2.80), avec $b = 100$.



FIG. 2.4 - \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour $\phi: \xi \mapsto |\xi|^{3/2}$ et $\gamma = 7$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d'Alkin et Caglar [1].



FIG. 2.5 - \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour $\phi: \xi \mapsto |\xi|^{3/2}$ et $\gamma = 2$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d'Alkin et Caglar [1].

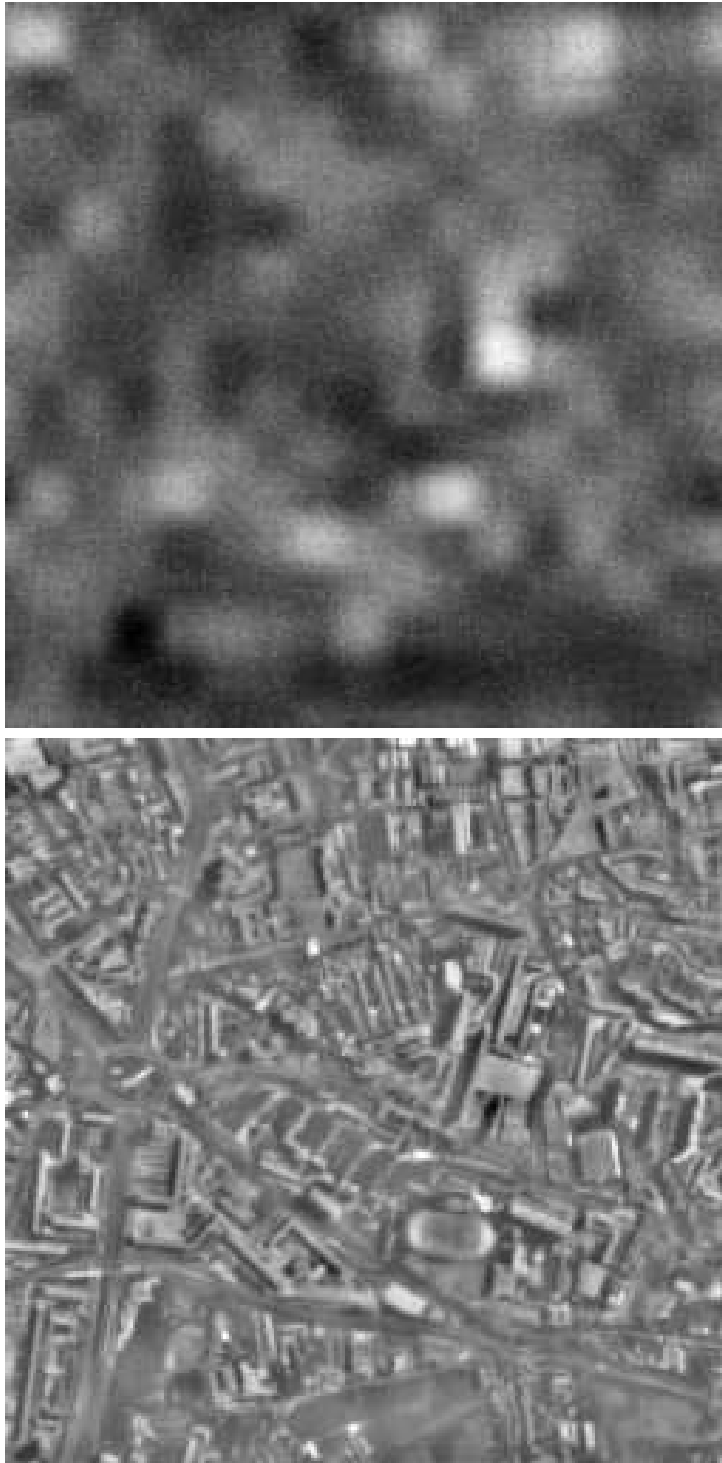


FIG. 2.6 – \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution triangulaire définie par (2.61), avec $-\underline{\omega} = \bar{\omega} = 7$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d’Alkin et Caglar [1].

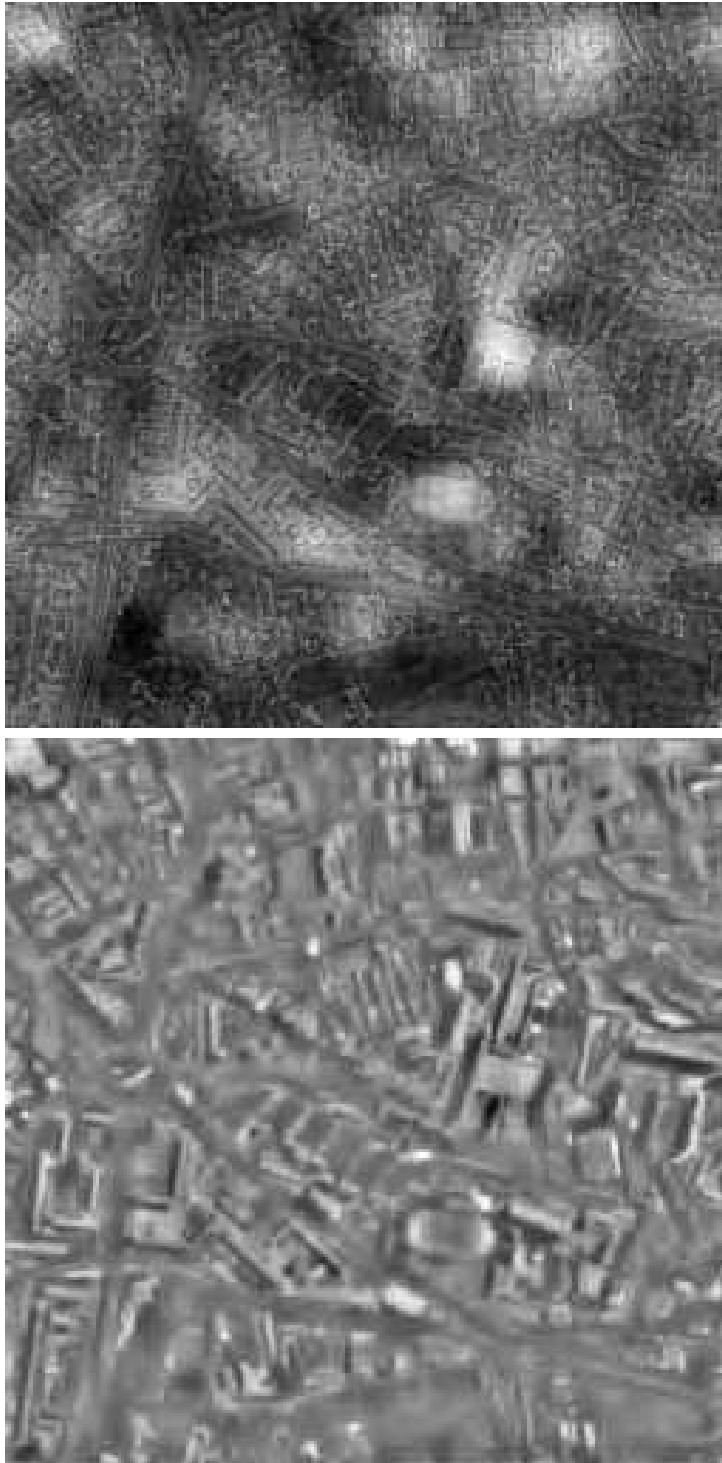


FIG. 2.7 - \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution triangulaire définie par (2.61) avec $-\underline{\omega} = \bar{\omega} = 30$ et $\gamma = 7$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d'Alkin et Caglar [1].

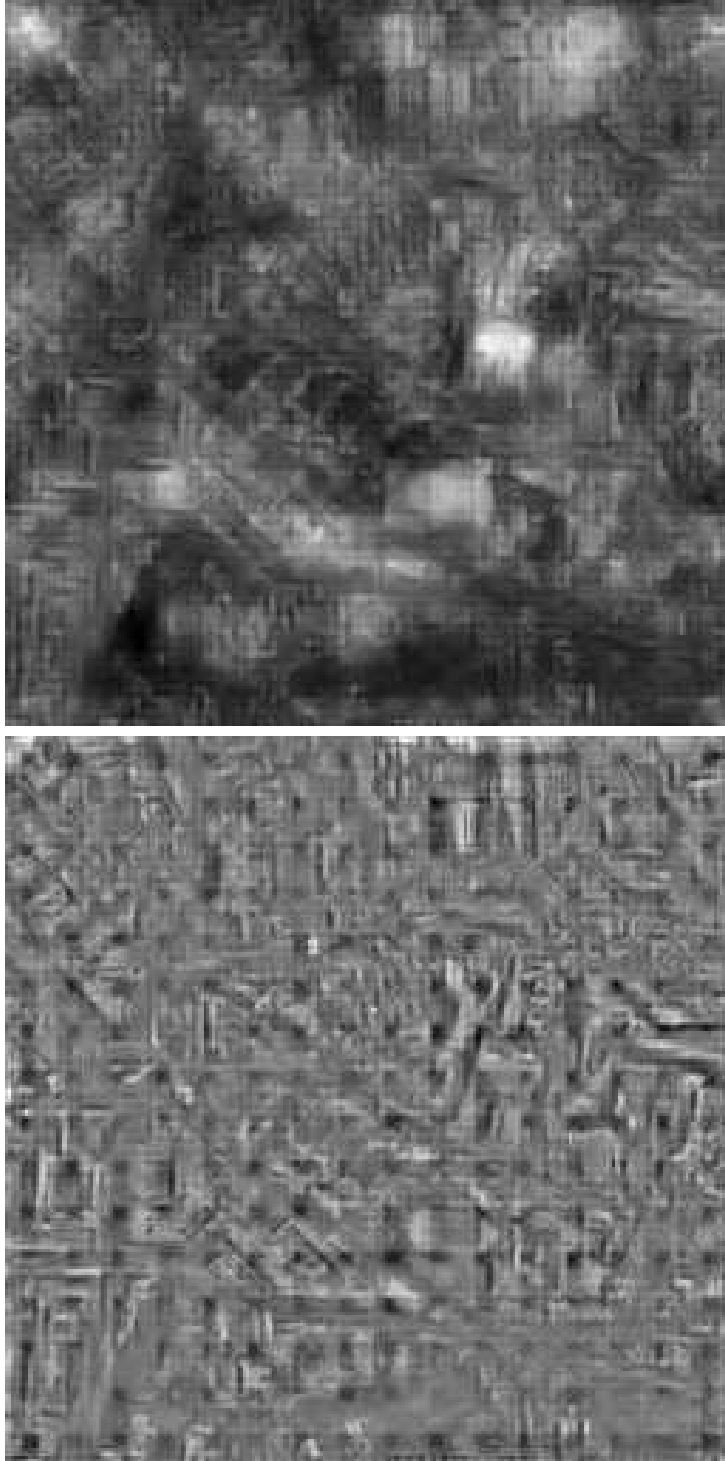


FIG. 2.8 – $\overline{x}_\gamma^\oplus$, en haut, $\overline{x}_\gamma^\ominus$, en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution chi définie par (2.19), avec $\kappa = 7$ et $\gamma = 1$ et sont décomposés en ondelettes M -bandes ($M=4$ bandes, 2 niveaux de résolution) selon le banc de filtre d'Alkin et Caglar [1].

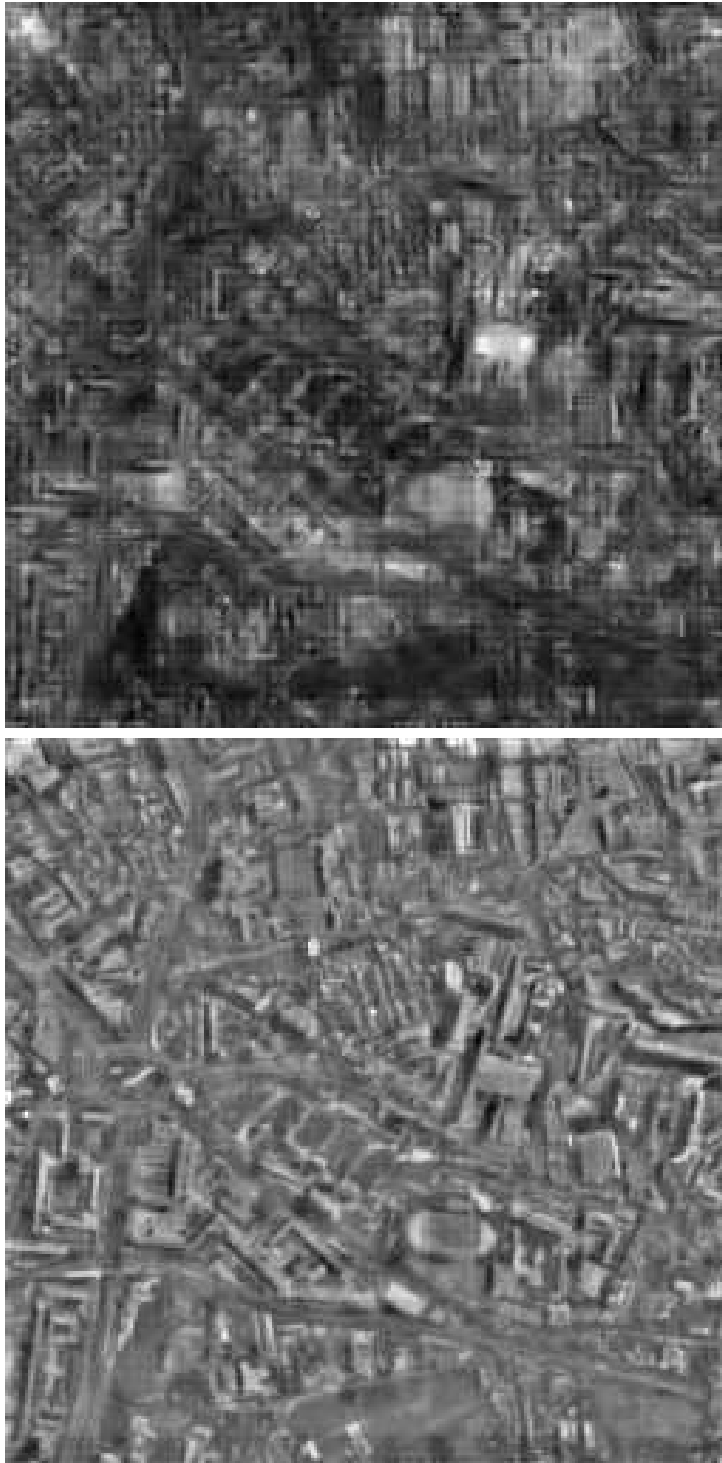


FIG. 2.9 – \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution gamma définie par (2.18), avec $\kappa = 7$, $\omega = 1$ et $\gamma = 1$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d'Alkin et Caglar [1].

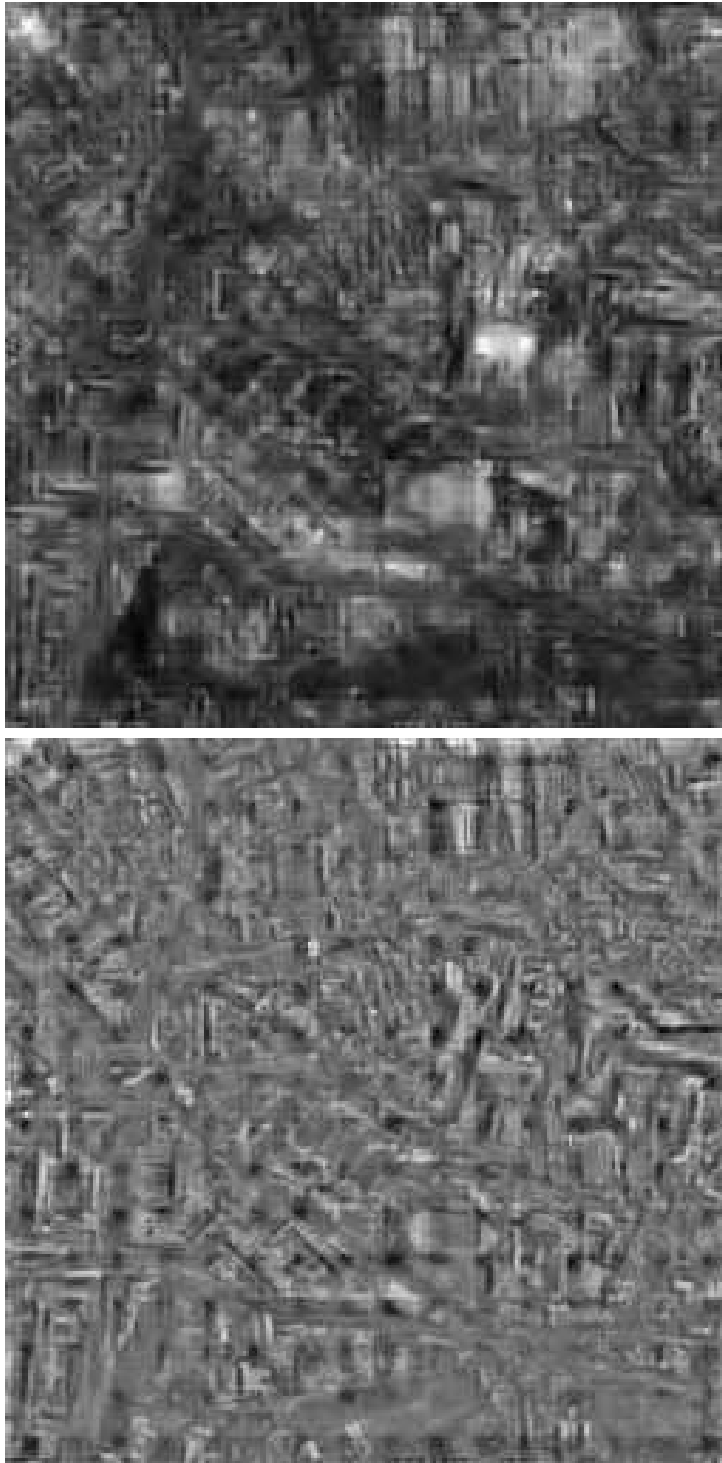


FIG. 2.10 – \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution gamma définie par (2.18), avec $\kappa = 7$, $\omega = 10$ et $\gamma = 1$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d’Alkin et Caglar [1].

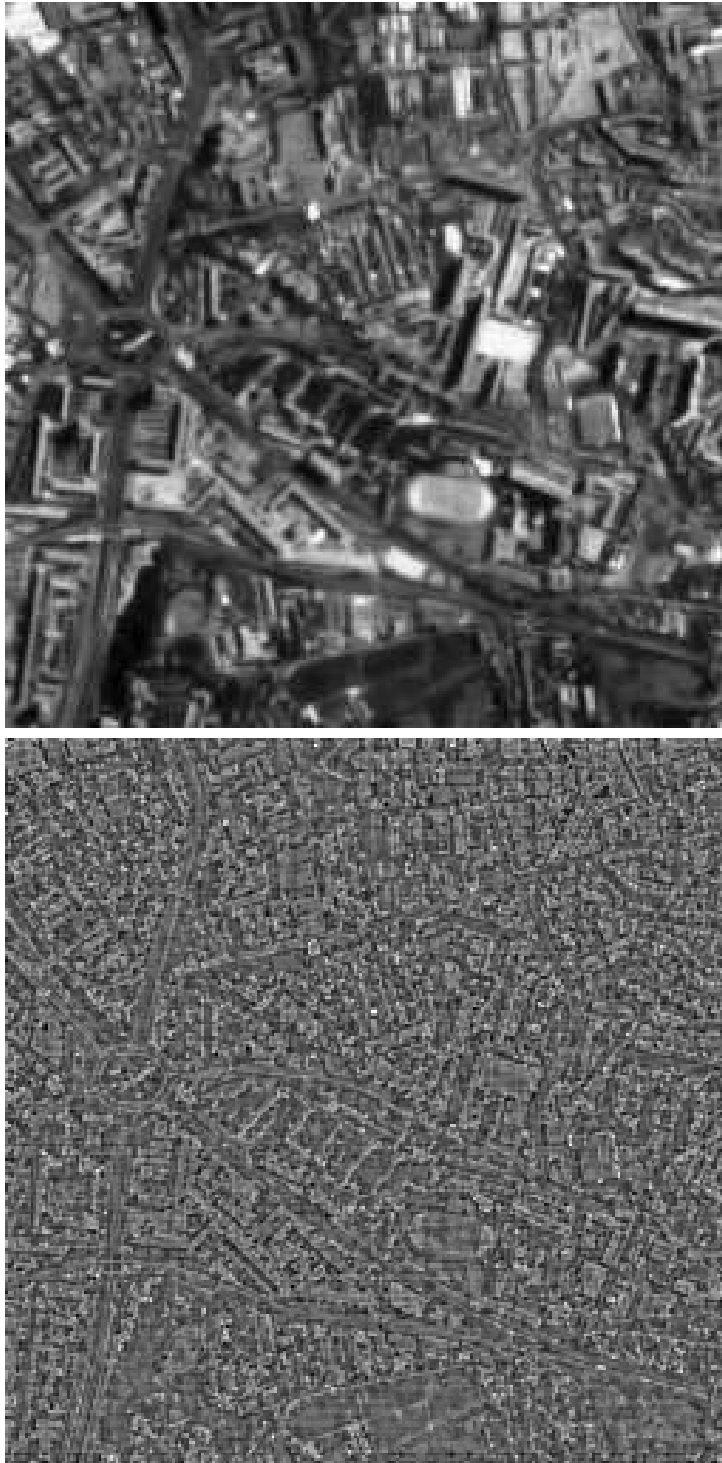


FIG. 2.11 – \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution de Huber définie par (2.60), avec $\omega = 7$, $\tau = 1$ et $\gamma = 1$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d’Alkin et Caglar [1].

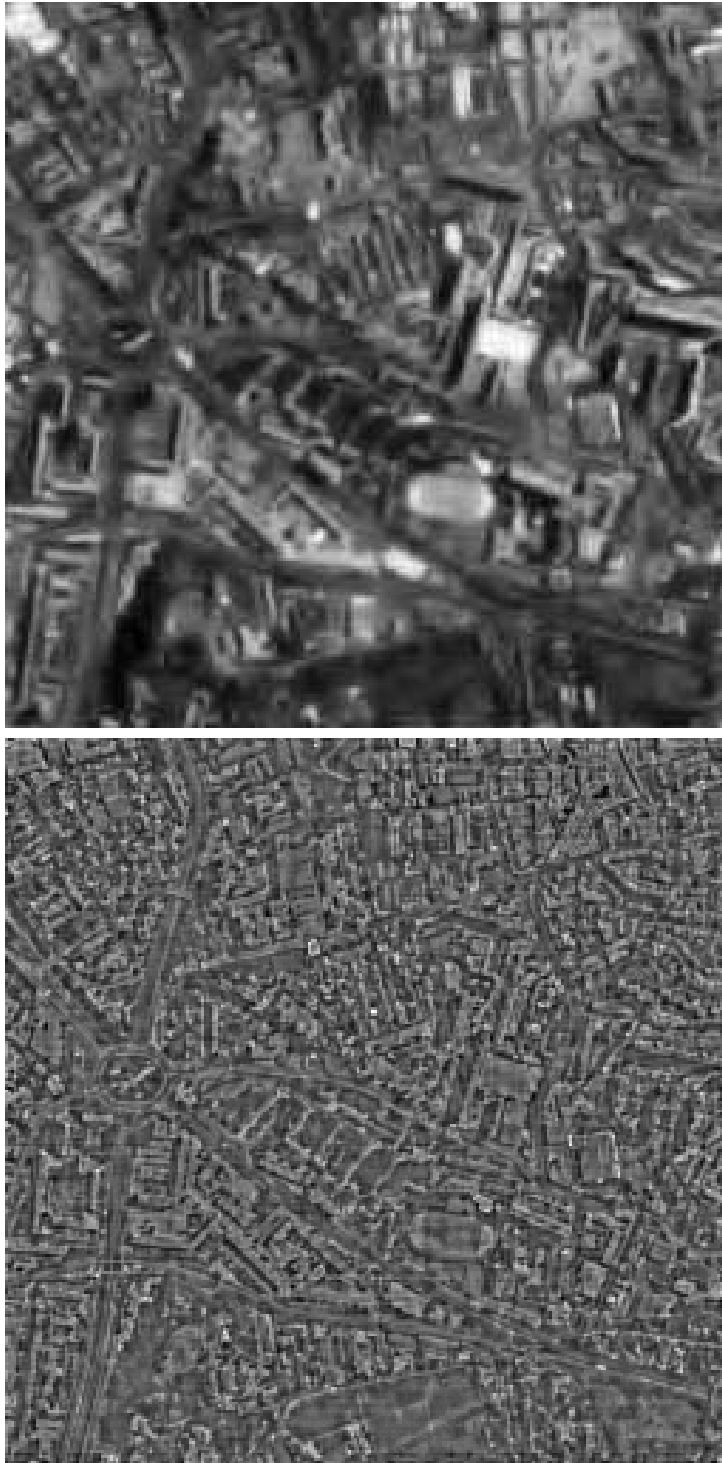


FIG. 2.12 – \bar{x}_γ^\oplus , en haut, \bar{x}_γ^\ominus , en bas, vérifient (2.92) pour ϕ la fonction potentiel de la distribution de Huber définie par (2.60), avec $\omega = 7$, $\tau = 10$ et $\gamma = 1$ et sont décomposés en ondelettes M -bandes ($M = 4$ bandes, 2 niveaux de résolution) selon le banc de filtre d’Alkin et Caglar [1].

- d'un point de vue applicatif, c'est en quelque sorte un filtre permettant d'analyser les signaux et tout particulièrement les images.

2.5 Bibliographie

- [1] O. ALKIN AND H. CAGLAR, *Design of efficient M-band coders with linear-phase and perfect-reconstruction properties*, IEEE Transactions on Signal Processing, 43 (1995), pp. 1579–1590.
- [2] H. ATTOUCH AND R. J.-B. WETS, *Epigraphical analysis*, Analyse Non Linéaire, (H. Attouch, J.-P. Aubin, F. Clarke, and I. Ekeland Eds.) Gauthier-Villars, Paris (1989), pp. 73–100.
- [3] J. F. BONNANS, J. CH. GILBERT, C. LEMARÉCHAL, AND C. A. SAGASTIZÁBAL, *A family of variable metric proximal methods*, Mathematical Programming, 68 (1995), pp. 15–47.
- [4] J. M. BORWEIN AND A. S. LEWIS, *Convex Analysis and Nonlinear Optimization. Theory and Examples*, CMS Books in Mathematics, Springer-Verlag, New York, Second Edition, 2006.
- [5] J. M. BORWEIN AND J. D. VANDERWERFF, *Convergence of Lipschitz regularisations of convex functions*, Journal of Functional Analysis, 128 (1995), pp. 139–162.
- [6] J. A. CADZOW AND Y. SUN, *Sequences with positive semidefinite Fourier transforms*, IEEE Transactions on Acoustics, Speech, and Signal Processing, 34 (1986), pp. 1502–1510.
- [7] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision, 20 (2004), pp. 89–97.
- [8] P. L. COMBETTES, *The convex feasibility problem in image recovery*, Advances in Imaging and Electron Physics, 95, (P. Hawkes, Ed.), Academic Press, New York (1996), pp. 155–270.
- [9] P. L. COMBETTES, *Convexité et signal*, Actes du Congrès de Mathématiques Appliquées et Industrielles SMAI'01, Pompadour, France (2001), pp. 6–16.
- [10] P. L. Combettes and J.-C. Pesquet, *Proximal thresholding algorithm for minimization over orthonormal bases*, *SIAM Journal on Optimization*, to appear.
- [11] I. EKELAND ET R. TEMAM, *Analyse Convexe et Problèmes Variationnels*, Dunod, Paris, 1974.
- [12] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, Heidelberg, 1993.
- [13] P. J. LAURENT, *Approximation et Optimisation*, Hermann, Paris, 1972.

- [14] C. LEMARÉCHAL AND C. SAGASTIZÁBAL, *Variable metric bundle methods : from conceptual to implementable forms*, Mathematical Programming, 76 (1997), pp. 393–410.
- [15] S. G. MALLAT, *A Wavelet Tour of Signal Processing*, Second Edition, Academic Press, New York, 1999.
- [16] P. MARÉCHAL, *On the convexity of the multiplicative potential and penalty functions and related topics*, Mathematical Programming, 89 (2001), pp. 505–516.
- [17] J.-J. MOREAU, *Fonctions convexes duales et points proximaux dans un espace hilbertien*, Comptes-Rendus de l'Académie des Sciences, Paris, Série A. Mathématique, 255 (1962), pp. 2897–2899.
- [18] J.-J. MOREAU, *Propriétés des applications « prox »*. Comptes-Rendus de l'Académie des Sciences, Paris, Série A. Mathématique, 256 (1963), pp. 1069–1071.
- [19] J.-J. MOREAU, *Proximité et dualité dans un espace hilbertien*, Bulletin de la Société Mathématique de France, 93 (1965), pp. 273–299.
- [20] F. RIESZ AND B. S. NAGY, *Functional Analysis*, Dover Publications, New York, 1990.
- [21] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton N.J, 1970.
- [22] R. T. ROCKAFELLAR AND R. J. B. WETS, *Variational Analysis*, Springer-Verlag, New York, 1998.
- [23] H. STARK, *Image Recovery : Theory and Application*, Academic Press, San Diego, CA, 1987.
- [24] P. STEFFEN, P. HELLER, R. A. GOPINATH, AND C. S. BURRUS, *Theory of regular M-band wavelet bases*, IEEE Transactions on Signal Processing, 41 (1993), pp. 3497–3511.
- [25] D. C. YOULA AND H. WEBB, *Image restoration by the method of convex projections : Part 1 – Theory*, IEEE Transactions on Medical Imaging, 1 (1982), pp. 81–94.

Chapitre 3

Décomposition, Restauration et Reconstruction de Signaux par la Méthode Proximale Explicite-Implicite

Ce chapitre présente les travaux de recherche qui ont abouti à la publication rapportée à la Section 3.3. Après un résumé de cette dernière, nous en exposons les résultats majeurs, pour finir par l'article publié en anglais.

3.1 Présentation globale

L'idée première est d'apporter une étude *synthétique* de différents types de problèmes variationnels inverses en traitement du signal, qui sont à première vue dissociés. Pour les unifier, nous nous proposons de résoudre, comme suit, le problème de minimisation, sur un espace de Hilbert réel \mathcal{H} , de deux fonctions convexes possédant certaines propriétés de régularité :

Problème 3.1.1 Soient $f_1 : \mathcal{H} \rightarrow]-\infty, +\infty]$ et $f_2 : \mathcal{H} \rightarrow \mathbb{R}$ deux fonctions propres, convexes, semi-continues inférieurement telles que f_2 soit différentiable sur tout \mathcal{H} , de gradient $1/\beta$ -lipschitz, pour un certain $\beta \in]0, +\infty[$. L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad f_1(x) + f_2(x). \tag{3.1}$$

On note G l'ensemble des solutions de ce problème.

Parmi les applications directes au Problème 3.1.1, nous pouvons relever les problèmes soumis à des contraintes de moindre carré [35, 48, 63], les problèmes

de régularisation parcimonieuse [10, 30, 31, 36], ou de régularisation de Fourier [46, 50], les problèmes de décomposition en composantes de texture ou de géométrie [5, 6, 7, 57, 71], les problèmes d’admissibilité soumis à des contraintes fortes [26], les problèmes de synthèse de signaux par projections alternées [26, 38], les problèmes de distance aux moindres carrés [22], les problèmes d’admissibilité décomposés [13, 15], les problèmes de variation totale [19, 62], quelques problèmes d’estimation par maximum a posteriori [68, 69].

Pour résoudre le Problème 3.1.1, nous utilisons un algorithme de type *explicite-implicite* dont la convergence est établie à l’aide de récentes méthodes de décomposition basées sur des opérateurs monotones [25]. Cette technique a l’avantage d’étendre et d’apporter une analyse simplifiée de diverses méthodes itératives déjà existantes. Une nouveauté de notre travail est d’utiliser la notion d’*opérateur proximal*, exposée et approfondie au Chapitre 2.

L’article de revue, rapporté à la Section 3.3, présente trois majeures parties : la première (Section 3.3.2) est consacrée aux opérateurs proximaux, la seconde (Section 3.3.3) à l’analyse théorique et algorithmique du Problème 3.1.1, la dernière (Section 3.3.4 à Section 3.3.6) aux applications en traitement du signal.

3.2 Les résultats principaux

3.2.1 L’opérateur proximal

Tous les résultats de la Section 3.3.2 de l’article sont détaillés et approfondis dans le Chapitre 2, consacré à l’opérateur proximal.

Pour synthétiser, après une présentation de cet opérateur dans l’espace de Hilbert réel \mathcal{X} , nous exposons ses propriétés essentielles : l’opérateur proximal est une contraction ferme qu’il est possible de déterminer, entre autres, par des opérations élémentaires sur la fonction à laquelle il est associé (perturbation quadratique, translation, dilatation, réflexion, enveloppe de Moreau, composition avec certaines applications linéaires bijectives, décomposition sur la somme hilbertienne directe \mathcal{X}^m , $m \in \mathbb{N}$). La dernière propriété énoncée concerne la décomposition de Moreau.

Enfin, des exemples de calcul explicite d’opérateur proximal sont donnés pour illustrer cette théorie : opérateurs proximaux linéaires, opérateur proximal du carré de la distance à un ensemble convexe fermé non vide de \mathcal{X} , seuillage doux scalaire, opérateur proximal de la fonction d’appui de l’enveloppe convexe fermée de $L(D)$ où $L \in \mathcal{B}(\mathcal{Y}, \mathcal{X})$ et D est un sous-ensemble non vide de l’espace de Hilbert réel \mathcal{Y} , opérateur proximal d’une fonction définie sur une base hilbertienne de \mathcal{X} , dont les composantes sont positives, de minimum nul en 0.

3.2.2 Le Problème 3.1.1 et la méthode de résolution de type explicite-implicite

La Section 3.3.3 de l'article étudie le Problème 3.1.1. D'un point de vue théorique, d'abord, nous exposons ses propriétés principales d'existence, d'unicité et de caractérisation des solutions.

Proposition 3.2.1

(i) *Existence* : Le Problème 3.1.1 possède au moins une solution si $f_1 + f_2$ est coercive, i.e.,

$$\lim_{\|x\| \rightarrow +\infty} f_1(x) + f_2(x) = +\infty. \quad (3.2)$$

(ii) *Unicité* : Le Problème 3.1.1 possède au plus une solution si $f_1 + f_2$ est strictement convexe, en particulier, si f_1 ou f_2 l'est.

(iii) *Caractérisation* : Soient $x \in \mathcal{H}$ et $\gamma \in]0, +\infty[$. Les conditions suivantes sont équivalentes :

(a) x est solution du Problème 3.1.1.

(b) $x = \text{prox}_{\gamma f_1}(x - \gamma \nabla f_2(x))$.

(c) $(\forall y \in \mathcal{H}) \langle x - y \mid \nabla f_2(x) \rangle + f_1(x) \leq f_1(y)$.

Comment résoudre numériquement une solution du Problème 3.1.1 ? Reprenons la caractérisation énoncée dans la Proposition 3.2.1(iii)(b). Elle suggère naturellement de considérer les itérations du point fixe $x_{n+1} = \text{prox}_{\gamma f_1}(x_n - \gamma \nabla f_2(x_n))$, construites en deux étapes, la première *explicite*, $x_{n+\frac{1}{2}} = x_n - \gamma \nabla f_2(x_n)$, la seconde *implicite*, $x_{n+1} = \text{prox}_{\gamma f_1} x_{n+\frac{1}{2}}$.

Plus généralement, il est intéressant de prendre en compte les erreurs tolérées dans la mise en œuvre numérique de l'algorithme. La convergence se justifie alors par certains résultats de [25].

Condition 3.2.2 Soit X un sous-ensemble non vide de l'espace de Hilbert réel \mathcal{X} . Une fonction $\varphi \in \Gamma_0(\mathcal{X})$ vérifie cette Condition sur X si pour toutes suites $(y_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ de \mathcal{X} et pour tous points $y \in \mathcal{X}$ et $v \in \partial\varphi(y)$, nous avons

$$[y_n \rightharpoonup y, v_n \rightharpoonup v, v_n \in \partial\varphi(y_n)] \Rightarrow y \text{ est un point d'accumulation fort de } (y_n)_{n \in \mathbb{N}}. \quad (3.3)$$

Le principal résultat de convergence est le

Théorème 3.2.3 Supposons que G soit non vide. Soit $(\gamma_n)_{n \in \mathbb{N}}$ une suite de $]0, +\infty[$ telle que $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2\beta$, soit $(\lambda_n)_{n \in \mathbb{N}}$ une suite de $]0, 1[$ telle que

$\inf_{n \in \mathbb{N}} \lambda_n > 0$, et soient $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ des suites de \mathcal{H} telles que $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ et $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fixons $x_0 \in \mathcal{H}$, et pour tout $n \in \mathbb{N}$, posons

$$x_{n+1} = x_n + \lambda_n \left(\text{prox}_{\gamma_n f_1} (x_n - \gamma_n (\nabla f_2(x_n) + b_n)) + a_n - x_n \right). \quad (3.4)$$

Alors,

- (i) $(x_n)_{n \in \mathbb{N}}$ converge faiblement vers un point $x \in G$.
- (ii) $\sum_{n \in \mathbb{N}} \|\nabla f_2(x_n) - \nabla f_2(x)\|^2 < +\infty$.
- (iii) $\sum_{n \in \mathbb{N}} \|\text{prox}_{\gamma_n f_1}(x_n - \gamma_n \nabla f_2(x_n)) - x_n\|^2 < +\infty$.
- (iv) $(x_n)_{n \in \mathbb{N}}$ converge fortement vers x si et seulement si $\liminf d_G(x_n) = 0$. En particulier, la convergence forte est obtenue dans chacun des cas suivants :
 - (a) $\text{int } G \neq \emptyset$.
 - (b) f_1 vérifie la Condition 3.2.2 sur G .
 - (c) f_2 vérifie la Condition 3.2.2 sur G .

Voici quelques cas où la Condition 3.2.2 est vérifiée :

Proposition 3.2.4 Soit \mathcal{X} un espace de Hilbert réel. Soient $\varphi \in \Gamma_0(\mathcal{X})$ et $\emptyset \neq X \subset D$, où $D = \text{dom } \varphi$. Soit \mathcal{C} l'ensemble des fonctions croissantes de $[0, +\infty[$ dans $[0, +\infty]$, qui s'annulent uniquement en 0. Alors φ vérifie la Condition 3.2.2 sur X dans chacun des cas suivants :

- (i) D est borné, relativement compact.
- (ii) φ est différentiable sur \mathcal{X} et $\text{Id} - \nabla \varphi$ est demi-compact [77, Section 10.4].
- (iii) Pour tout $y \in X$ et $v \in \partial \varphi(y)$, il existe une fonction $c \in \mathcal{C}$ telle que

$$(\forall x \in D) \quad \langle x - y \mid v \rangle + \varphi(y) + c(\|x - y\|) \leq \varphi(x). \quad (3.5)$$

- (iv) φ est uniformément convexe en tout point de \mathcal{X} .
- (v) φ est uniformément convexe.
- (vi) φ est uniformément convexe sur tout ensemble borné.
- (vii) φ est fortement convexe.

3.2.3 Applications en traitement du signal

La dernière partie, composée des trois dernières sections de l'article, met en évidence les applications en traitement du signal, qui se modélisent sous la forme du Problème 3.1.1.

3.2.3.1 Problèmes impliquant des sommes d'enveloppes de Moreau

Avant d'énoncer le problème, rappelons la définition de l'enveloppe de Moreau d'une fonction $\varphi \in \Gamma_0(\mathcal{H})$:

Définition 3.2.5 Soient $\varphi \in \Gamma_0(\mathcal{H})$ et $\gamma \in]0, +\infty[$. L'enveloppe de Moreau de φ d'indice γ , notée $\gamma\varphi$, est la convolution infimale $\varphi \square (\|\cdot\|^2/(2\gamma))$.

Problème 3.2.6 Soient

- (i) $(\mathcal{K}_i)_{1 \leq i \leq m}$ des espaces de Hilbert réels ;
- (ii) Pour tout $i \in \{1, \dots, m\}$, $L_i \in \mathcal{B}(\mathcal{H}, \mathcal{K}_i)$ non nul, $\varphi_i \in \Gamma_0(\mathcal{K}_i)$ et $\rho_i \in]0, +\infty[$;
- (iii) $f_1 \in \Gamma_0(\mathcal{H})$.

L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad f_1(x) + \sum_{i=1}^m \rho_i \varphi_i(L_i x). \quad (3.6)$$

L'ensemble des solutions de ce problème est noté G .

Proposition 3.2.7 Le Problème 3.2.6 est un cas particulier du Problème 3.1.1, avec $f_2 = \sum_{i=1}^m \rho_i \varphi_i \circ L_i$ et $\beta = (\sum_{i=1}^m \|L_i\|^2 / \rho_i)^{-1}$.

Du Théorème 3.2.3, nous déduisons l'algorithme qui numérise la solution du Problème 3.2.6 ainsi que sa convergence. Ces résultats font l'objet du Théorème 3.3.30.

Présentons quelques applications du Problème 3.2.6 en traitement du signal :

- problèmes d'admissibilité décomposés : ils correspondent au cas $m = 1$; par exemple lorsque $\rho_1 = 1$, $f_1 = \iota_C$ et $\varphi_1 = \iota_Q$ où $C \subset \mathcal{H}$ et $Q \subset \mathcal{K}_1$ sont deux convexes non vides fermés, le problème a été introduit dans [15]. Au passage, nous retrouvons les schémas d'extrapolation de signaux classiques de Gerchberg [37] et Papoulis [59] ;
- problèmes de texture/géométrie : soit $x \in \mathcal{H}$ un signal décomposé sous la forme $x = u + v$, où u caractérise les composantes géométriques du signal et v sa texture. Le signal recherché, à partir d'une observation bruitée $z \in \mathcal{H}$, est solution du problème variationnel [5, 6, 7, 71, 72] :

$$\underset{(u,v) \in \mathcal{H} \times \mathcal{H}}{\text{minimiser}} \quad \psi(u) + \phi(v) + \frac{1}{2\rho} \|u + v - z\|^2, \quad (3.7)$$

où ψ et ϕ sont dans $\Gamma_0(\mathcal{H})$ et $\rho \in]0, +\infty[$. Ce problème se ramène au Problème 3.2.6, avec $m = 1$, $f_1 = \psi$, $\mathcal{K}_1 = \mathcal{H}$, $L = \text{Id}$, $\rho_1 = \rho$ et $\varphi_1 = \varphi = \phi(z - \cdot)$;

- problèmes d'admissibilité de signaux soumis à des contraintes fortes [26] dont l'objectif est de

$$\text{minimiser}_{x \in C} \frac{1}{2} \sum_{i=1}^m \omega_i d_{C_i}^2(x), \quad (3.8)$$

avec $\omega_i \in]0, +\infty[$: ce problème se ramène au Problème 3.2.6, avec $f_1 = \iota_C$ et pour tout $i \in \{1, \dots, m\}$, $\mathcal{K}_i = \mathcal{H}$, $L_i = \text{Id}$, $\rho_i = 1/\omega_i$ et $\varphi_i = \iota_{C_i}$, où C et $(C_i)_{1 \leq i \leq m}$ sont des convexes non vides fermés de \mathcal{H} .

3.2.3.2 Problèmes inverses linéaires « régularisés »

On considère le problème inverse où le signal observé z , dans un espace de Hilbert réel \mathcal{G} , est lié au signal étudié \bar{x} par le modèle suivant :

$$z = T\bar{x} + w, \quad (3.9)$$

où $T \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ et $w \in \mathcal{G}$ représente la réalisation du bruit superposé au signal. Le problème considéré est alors le suivant (cf. [2, 16, 23, 39, 66, 67] pour des cas particuliers) :

Problème 3.2.8 Soient

- (i) \mathcal{K} un espace de Hilbert réel ;
- (ii) $T : \mathcal{H} \rightarrow \mathcal{G}$ un opérateur linéaire borné non nul ;
- (iii) $L : \mathcal{H} \rightarrow \mathcal{K}$ un opérateur linéaire borné bijectif tel que $L^{-1} = L^*$;
- (iv) $f \in \Gamma_0(\mathcal{K})$.

L'objectif est de

$$\text{minimiser}_{x \in \mathcal{H}} f(Lx) + \frac{1}{2} \|Tx - z\|^2. \quad (3.10)$$

L'ensemble des solutions de ce problème est noté G .

Dans la fonctionnelle à minimiser, $\|Tx - z\|^2/2$ symbolise le terme de *fidélité* du modèle aux données, tandis que le terme $f(Lx)$ correspond aux informations a priori sur le signal original \bar{x} .

Proposition 3.2.9 *Le Problème 3.2.8 est un cas particulier du Problème 3.1.1, avec $f_1 = f \circ L$, $f_2 : x \mapsto \|Tx - z\|^2/2$ et $\beta = 1/\|T\|^2$.*

Proposition 3.2.10

- (i) *Le Problème 3.2.8 admet au moins une solution si f est coercive.*

- (ii) Le Problème 3.2.8 admet au plus une solution si l'une des conditions est vérifiée :
- (a) f est strictement convexe.
 - (b) T est injectif.
- (iii) Le Problème 3.2.8 admet exactement une solution si T est borné inférieurement, i.e.,

$$(\exists \kappa \in]0, +\infty[)(\forall x \in \mathcal{H}) \quad \|Tx\| \geq \kappa \|x\|. \quad (3.11)$$

- (iv) Soient $x \in \mathcal{H}$ et $\gamma \in]0, +\infty[$. Alors les conditions suivantes sont équivalentes :
- (a) x est solution du Problème 3.2.8.
 - (b) $x = (L^* \circ \text{prox}_{\gamma f} \circ L)(x + \gamma T^*(z - Tx))$.
 - (c) $(\forall y \in \mathcal{H}) \langle Ty - Tx \mid z - Tx \rangle + f(Lx) \leq f(Ly)$.

Là encore, le Problème 3.2.8 donne lieu à de nombreuses applications conséquentes en traitement du signal :

- approche statistique [68, 69] : $\mathcal{K} = \mathcal{H}$, $L = \text{Id}$ et w est la réalisation d'un bruit Gaussien, modèles discrets étudiés par le maximum a posteriori (MAP), avec une densité proportionnelle à $\exp(-f)$;
- méthodes entropiques, variation totale, information de Fisher [3, 19, 32, 40, 45] : $\mathcal{K} = \mathcal{H} = H^1(\Omega)$, où Ω est un domaine ouvert de \mathbb{R}^m , $L = \text{Id}$ et f est une fonctionnelle intégrale de la forme

$$f : x \mapsto \gamma \int_{\Omega} \varphi(\omega, x(\omega), \nabla x(\omega)) d\omega, \quad \text{où } \gamma \in]0, +\infty[; \quad (3.12)$$

- problèmes de moindres carrés : $f = 0$, plus généralement, en choisissant $\mathcal{K} = \mathcal{H}$, $L = \text{Id}$ et $f = \iota_C$ où C est un convexe non vide fermé de \mathcal{H} , l'objectif est de

$$\underset{x \in C}{\text{minimiser}} \quad \frac{1}{2} \|Tx - z\|^2; \quad (3.13)$$

l'algorithme du Théorème 3.2.3 se particularise à l'itération classique de Landweber ;

- régularisation parcimonieuse : l'idée est de décomposer le signal selon une base hilbertienne et de transformer les coefficients de la décomposition pour construire des approximations parcimonieuses ou des estimateurs [18, 20, 30, 31, 33, 49]. Le problème suivant modélise, de manière plus générale, une telle situation :

Problème 3.2.11 Soient

- (i) $T : \mathcal{H} \rightarrow \mathcal{G}$ un opérateur linéaire borné non nul ;

- (ii) $(e_k)_{k \in \mathbb{N}}$ une base hilbertienne de \mathcal{H} ;
- (iii) $(\phi_k)_{k \in \mathbb{N}}$ des fonctions de $\Gamma_0(\mathbb{R})$ telles que $(\forall k \in \mathbb{N}) \phi_k \geq \phi_k(0) = 0$.

L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad \frac{1}{2} \|Tx - z\|^2 + \sum_{k \in \mathbb{N}} \phi_k(\langle x | e_k \rangle). \quad (3.14)$$

Ce problème est un cas particulier du Problème 3.2.8, avec $\mathcal{K} = l^2(\mathbb{N})$, $L : x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{N}}$ et $f : (\xi_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \phi_k(\xi_k)$. La Proposition 3.3.53 donne des résultats d'existence, d'unicité et de caractérisation de toute solution, le Corollaire 3.3.54, sa résolution numérique.

3.2.3.3 Problèmes de débruitage

Les problèmes de débruitage reviennent au Problème 3.2.8 avec $\mathcal{G} = \mathcal{H}$ et $T = \text{Id}$:

Problème 3.2.12 Soient

- (i) \mathcal{K} un espace de Hilbert réel;
- (ii) $L \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ bijectif tel que $L^{-1} = L^*$;
- (iii) $f \in \Gamma_0(\mathcal{K})$.

L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \quad f(Lx) + \frac{1}{2} \|x - z\|^2. \quad (3.15)$$

Proposition 3.2.13 *Le Problème 3.2.12 possède exactement une solution z^\oplus qui est caractérisée par l'une des conditions équivalentes suivantes :*

- (i) $z^\oplus = \text{prox}_{f \circ L} z = (L^* \circ \text{prox}_f \circ L)z$.
- (ii) $(\forall y \in \mathcal{H}) \langle y - z^\oplus | z - z^\oplus \rangle + f(Lz^\oplus) \leq f(Ly)$.

La décomposition de Moreau fournit alors le signal résiduel $z^\ominus = z - z^\oplus$.

Un cas particulier intéressant consiste à décomposer f , si possible, en la somme de deux fonctions $f = \varphi + \psi$, où $\varphi \in \Gamma_0(\mathcal{K})$ a un opérateur proximal relativement facile à implémenter et $\psi \in \Gamma_0(\mathcal{K})$ différentiable de gradient lipschitz sur \mathcal{K} . Par exemple, si $\varphi = \iota_C$, où C est un convexe non vide fermé de \mathcal{K} , nous obtenons une solution numérique du problème de débruitage, associé à $\psi \circ L$, sur l'ensemble d'admissibilité $L^{-1}(C)$. Ainsi, le bruit superposé à un signal est supprimé à partir de la connaissance d'informations a priori.

L'article fournit une étude précise sur un plus vaste champ d'exemples.

3.3 Article en anglais

SIGNAL RECOVERY BY PROXIMAL FORWARD-BACKWARD SPLITTING¹

Abstract : We show that various inverse problems in signal recovery can be formulated as the generic problem of minimizing the sum of two convex functions with certain regularity properties. This formulation makes it possible to derive existence, uniqueness, characterization, and stability results in a unified and standardized fashion for a large class of apparently disparate problems. Recent results on monotone operator splitting methods are applied to establish the convergence of a forward-backward algorithm to solve the generic problem. In turn, we recover, extend, and provide a simplified analysis for a variety of existing iterative methods. Applications to geometry/texture image decomposition schemes are also discussed. A novelty of our framework is to use extensively the notion of a proximity operator, which was introduced by Moreau in the 1960s.

3.3.1 Introduction

Signal recovery encompasses the large body of inverse problems in which a multi-dimensional signal \bar{x} is to be inferred from the observation of data z consisting of signals physically or mathematically related to it [23, 66]. The original signal \bar{x} and the observation z are typically assumed to lie in some real Hilbert spaces \mathcal{H} and \mathcal{G} , respectively. For instance, in image restoration [2], the objective is to recover the original form of an image \bar{x} from the observation of a blurred and noise-corrupted version z , and therefore $\mathcal{H} = \mathcal{G}$. On the other hand, in signal reconstruction, the data z are indirectly related to \bar{x} and therefore \mathcal{H} and \mathcal{G} are often different spaces. Thus, in tomography [39], a signal must be recovered from a collection of measurements of lower dimensional signals; in phase retrieval, holography, or band-limited extrapolation [44, 66], a signal must be recovered from partial measurements of its Fourier transform.

Mathematically, signal recovery problems are most conveniently formulated as variational problems, the ultimate goal of which is to incorporate various forms of *a priori* information and impose some degree of consistency with the measured data z . The objective of the present paper is to investigate in a unified fashion the properties and the numerical solution of a variety of variational formulations which arise in the following format.

Problem 3.3.1 Let $f_1: \mathcal{H} \rightarrow]-\infty, +\infty]$ and $f_2: \mathcal{H} \rightarrow \mathbb{R}$ be two proper lower

¹P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *SIAM Journal on Multiscale Modeling and Simulation*, vol. 4, no. 4, pp. 1168-1200, November 2005.

semicontinuous convex functions such that f_2 is differentiable on \mathcal{H} with a $1/\beta$ -Lipschitz continuous gradient for some $\beta \in]0, +\infty[$. The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f_1(x) + f_2(x). \quad (3.16)$$

The set of solutions to this problem is denoted by G .

Despite its simplicity, Problem 3.3.1 will be shown to cover a wide range of apparently unrelated signal recovery formulations, including constrained least-squares problems [35, 48, 63], multiresolution sparse regularization problems [10, 30, 31, 36], Fourier regularization problems [46, 50], geometry/texture image decomposition problems [5, 6, 7, 57, 71], hard-constrained inconsistent feasibility problems [26], alternating projection signal synthesis problems [38, 60], least square-distance problems [22], split feasibility problems [13, 15], total variation problems [19, 62], as well as certain maximum *a posteriori* problems [68, 69]. Thus, our study of Problem 3.3.1 will not only bring together these and other signal recovery approaches within a common simple framework, but it will also capture and extend scattered results pertaining to their properties (existence, uniqueness, characterization, and stability of solutions) and to the convergence of associated numerical methods.

Our investigation relies to a large extent on convex analysis and, in particular, on the notion of a *proximity operator*, which was introduced by Moreau in [53]. Section 3.3.2 will provide an account of the main properties of these operators, together with specific examples. In Section 3.3.3, we study the properties of Problem 3.3.1 and analyze the convergence of a general *forward-backward splitting* algorithm to solve it. The principle of this algorithm is to use at every iteration the functions f_1 and f_2 separately; more specifically the core of an iteration consists of a forward (explicit) gradient step on f_2 , followed by a backward (implicit) step on f_1 . In the remaining Sections 3.3.4–3.3.6, the general results of Section 3.3.3 are specialized to various settings and the forward-backward splitting scheme is shown to reduce to familiar signal recovery algorithms, which were obtained and analyzed by different means in the literature. Section 3.3.4 is devoted to problems involving sums of Moreau envelopes, Section 3.3.5 to problems with linear data formation models, and Section 3.3.6 to denoising problems.

3.3.1.1 Notation

Let \mathcal{X} be a real Hilbert space. We denote by $\langle \cdot | \cdot \rangle$ its scalar product, by $\| \cdot \|$ the associated norm, and by d the associated distance; Id denotes the identity operator on \mathcal{X} and $B(x; \rho)$ the closed ball of center $x \in \mathcal{X}$ and radius $\rho \in]0, +\infty[$. The expressions $x_n \rightharpoonup x$ and $x_n \rightarrow x$ denote, respectively, the weak and the strong convergence to x of a sequence $(x_n)_{n \in \mathbb{N}}$ in \mathcal{X} .

Let $\varphi: \mathcal{X} \rightarrow [-\infty, +\infty]$ be a function. The domain and the epigraph of φ are $\text{dom } \varphi = \{x \in \mathcal{X} \mid \varphi(x) < +\infty\}$ and $\text{epi } \varphi = \{(x, \eta) \in \mathcal{X} \times \mathbb{R} \mid \varphi(x) \leq \eta\}$, respectively; φ is lower semicontinuous if $\text{epi } \varphi$ is closed in $\mathcal{X} \times \mathbb{R}$, and convex if $\text{epi } \varphi$ is convex in $\mathcal{X} \times \mathbb{R}$. $\Gamma_0(\mathcal{X})$ is the class of all lower semicontinuous convex functions from \mathcal{X} to $]-\infty, +\infty]$ that are not identically $+\infty$.

Let C be a subset of \mathcal{X} . The interior of C is denoted by $\text{int } C$ and its closure by \overline{C} . If C is nonempty, the distance from a point $x \in \mathcal{X}$ to C is $d_C(x) = \inf \|x - C\|$; if C is also closed and convex then, for every $x \in \mathcal{X}$, there exists a unique point $P_C x \in C$ such that $\|x - P_C x\| = d_C(x)$. The point $P_C x$ is the projection of x onto C and it is characterized by the relations

$$P_C x \in C \quad \text{and} \quad (\forall z \in C) \quad \langle z - P_C x \mid x - P_C x \rangle \leq 0. \quad (3.17)$$

3.3.2 Proximity operators

This section is devoted to the notion of a proximity operator, which was introduced by Moreau in 1962 [53] and further investigated in [54, 55] as a generalization of the notion of a convex projection operator. Though convex projection operators have been used extensively in nonlinear signal recovery (see [21, 23, 66, 67, 74] and the references therein), the use of proximity operators seems to have been initiated in [24]. Throughout, \mathcal{X} and \mathcal{Y} are real Hilbert spaces.

3.3.2.1 Elements of convex analysis

We recall key facts in convex analysis. Details and further results will be found in [76].

Let $\varphi \in \Gamma_0(\mathcal{X})$. The conjugate of φ is the function $\varphi^* \in \Gamma_0(\mathcal{X})$ defined by

$$(\forall u \in \mathcal{X}) \quad \varphi^*(u) = \sup_{x \in \mathcal{X}} \langle x \mid u \rangle - \varphi(x). \quad (3.18)$$

Moreover, $\varphi^{**} = \varphi$. For instance, the conjugate of the indicator function of a nonempty closed convex set C , i.e.,

$$\iota_C: x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{if } x \notin C, \end{cases} \quad (3.19)$$

is the support function of C , i.e.,

$$\iota_C^* = \sigma_C: u \mapsto \sup_{x \in C} \langle x \mid u \rangle. \quad (3.20)$$

Consequently,

$$\sigma_C^* = \iota_C^{**} = \iota_C. \quad (3.21)$$

The subdifferential of φ is the set-valued operator $\partial\varphi: \mathcal{X} \rightarrow 2^{\mathcal{X}}$ the value of which at $x \in \mathcal{X}$ is

$$\partial\varphi(x) = \{u \in \mathcal{X} \mid (\forall y \in \mathcal{X}) \langle y - x \mid u \rangle + \varphi(x) \leq \varphi(y)\} \quad (3.22)$$

or, equivalently,

$$\partial\varphi(x) = \{u \in \mathcal{X} \mid \varphi(x) + \varphi^*(u) = \langle x \mid u \rangle\}. \quad (3.23)$$

Accordingly (Fermat's rule),

$$(\forall x \in \mathcal{X}) \varphi(x) = \inf \varphi(\mathcal{X}) \Leftrightarrow 0 \in \partial\varphi(x). \quad (3.24)$$

Moreover, if φ is (Gâteaux) differentiable at x with gradient $\nabla\varphi(x)$, then $\partial\varphi(x) = \{\nabla\varphi(x)\}$. Now, let C be a nonempty closed convex subset of \mathcal{X} . Then the normal cone operator of C is

$$N_C = \partial\iota_C: x \mapsto \begin{cases} \{u \in \mathcal{X} \mid (\forall y \in C) \langle y - x \mid u \rangle \leq 0\}, & \text{if } x \in C; \\ \emptyset, & \text{otherwise.} \end{cases} \quad (3.25)$$

Furthermore,

$$(\forall x \in \mathcal{X}) \partial d_C(x) = \begin{cases} \left\{ \frac{x - P_C x}{d_C(x)} \right\}, & \text{if } x \notin C; \\ N_C(x) \cap B(0; 1), & \text{if } x \in C. \end{cases} \quad (3.26)$$

Lemma 3.3.2 [76, Corollary 2.4.5] *Let $(\phi_k)_{1 \leq k \leq m}$ be functions in $\Gamma_0(\mathcal{X})$, let \mathcal{X}^m be the standard Hilbert product space, and let $\varphi: \mathcal{X}^m \rightarrow]-\infty, +\infty]: (x_k)_{1 \leq k \leq m} \mapsto \sum_{k=1}^m \phi_k(x_k)$. Then $\partial\varphi = \times_{k=1}^m \partial\phi_k$.*

Lemma 3.3.3 *Let $\varphi \in \Gamma_0(\mathcal{Y})$, let $\psi \in \Gamma_0(\mathcal{X})$, and let $L: \mathcal{X} \rightarrow \mathcal{Y}$ be a bounded linear operator such that $0 \in \text{int}(\text{dom } \varphi - L(\text{dom } \psi))$. Then*

- (i) $\partial(\varphi \circ L + \psi) = L^* \circ (\partial\varphi) \circ L + \partial\psi$ [76, Theorem 2.8.3].
- (ii) $\inf_{x \in \mathcal{X}} (\varphi(Lx) + \psi(x)) = -\min_{v \in \mathcal{Y}} (\varphi^*(v) + \psi^*(-L^*v))$ (Fenchel-Rockafellar duality formula) [76, Corollary 2.8.5].

3.3.2.2 Firmly nonexpansive operators

Definition 3.3.4 An operator $T: \mathcal{X} \rightarrow \mathcal{X}$ is firmly nonexpansive if it satisfies one of the following equivalent conditions :

- (i) $(\forall (x, y) \in \mathcal{X}^2) \|Tx - Ty\|^2 \leq \langle Tx - Ty \mid x - y \rangle$.
- (ii) $(\forall (x, y) \in \mathcal{X}^2) \|Tx - Ty\|^2 \leq \|x - y\|^2 - \|(\text{Id} - T)x - (\text{Id} - T)y\|^2$.

It follows immediately that a firmly nonexpansive operator $T: \mathcal{X} \rightarrow \mathcal{X}$ is nonexpansive, i.e.,

$$(\forall (x, y) \in \mathcal{X}^2) \|Tx - Ty\| \leq \|x - y\|. \quad (3.27)$$

3.3.2.3 Proximity operators

The Moreau envelope of index $\gamma \in]0, +\infty[$ of a function $\varphi \in \Gamma_0(\mathcal{X})$ is the continuous convex function

$$\gamma\varphi: \mathcal{X} \rightarrow \mathbb{R}: x \mapsto \inf_{y \in \mathcal{X}} \varphi(y) + \frac{1}{2\gamma} \|x - y\|^2. \quad (3.28)$$

For every $x \in \mathcal{X}$, the infimum in (3.28) is achieved at a unique point $\text{prox}_{\gamma\varphi} x$ which is characterized by the inclusion

$$x - \text{prox}_{\gamma\varphi} x \in \gamma \partial\varphi(\text{prox}_{\gamma\varphi} x). \quad (3.29)$$

The operator

$$\text{prox}_\varphi: \mathcal{X} \rightarrow \mathcal{X}: x \mapsto \arg \min_{y \in \mathcal{X}} \varphi(y) + \frac{1}{2} \|x - y\|^2 \quad (3.30)$$

thus defined is called the proximity operator of φ . Let us note that, if $\varphi = \iota_C$, then

$$\gamma\varphi = \frac{1}{2\gamma} d_C^2 \quad \text{and} \quad \text{prox}_{\gamma\varphi} = P_C. \quad (3.31)$$

Proximity operators are therefore a generalization of projection operators.

Lemma 3.3.5 *Let $\varphi \in \Gamma_0(\mathcal{X})$. Then prox_φ and $\text{Id} - \text{prox}_\varphi$ are firmly nonexpansive.*

Proof. The first assertion appears implicitly in [55], we detail the argument for completeness. Take x and y in \mathcal{X} . Then (3.29) and (3.22) yield

$$\begin{cases} \langle \text{prox}_\varphi y - \text{prox}_\varphi x \mid x - \text{prox}_\varphi x \rangle + \varphi(\text{prox}_\varphi x) \leq \varphi(\text{prox}_\varphi y) \\ \langle \text{prox}_\varphi x - \text{prox}_\varphi y \mid y - \text{prox}_\varphi y \rangle + \varphi(\text{prox}_\varphi y) \leq \varphi(\text{prox}_\varphi x). \end{cases} \quad (3.32)$$

Adding these two inequalities, we obtain

$$\| \text{prox}_\varphi x - \text{prox}_\varphi y \|^2 \leq \langle \text{prox}_\varphi x - \text{prox}_\varphi y \mid x - y \rangle. \quad (3.33)$$

The second assertion follows at once from the symmetry between T and $\text{Id} - T$ in Definition 3.3.4(ii). \square

Lemma 3.3.6 *Let $\varphi \in \Gamma_0(\mathcal{X})$ and $\gamma \in]0, +\infty[$. Then $\gamma\varphi$ is Fréchet-differentiable on \mathcal{X} and $\nabla(\gamma\varphi) = (\text{Id} - \text{prox}_{\gamma\varphi})/\gamma$.*

Proof. A routine extension of [55, Proposition 7.d], where $\gamma = 1$. \square

3.3.2.4 Proximal calculus

Lemma 3.3.7 *Let $\varphi \in \Gamma_0(\mathcal{X})$ and $x \in \mathcal{X}$. Then we have the following.*

- (i) *Quadratic perturbation : Let $\psi = \varphi + \alpha \|\cdot\|^2/2 + \langle \cdot | u \rangle + \beta$, where $u \in \mathcal{X}$, $\alpha \in [0, +\infty[$, and $\beta \in \mathbb{R}$. Then $\text{prox}_\psi x = \text{prox}_{\varphi/(\alpha+1)}((x-u)/(\alpha+1))$.*
- (ii) *Translation : Let $\psi = \varphi(\cdot - z)$, where $z \in \mathcal{X}$. Then $\text{prox}_\psi x = z + \text{prox}_\varphi(x - z)$.*
- (iii) *Scaling : Let $\psi = \varphi(\cdot/\rho)$, where $\rho \in \mathbb{R} \setminus \{0\}$. Then $\text{prox}_\psi x = \rho \text{prox}_{\varphi/\rho^2}(x/\rho)$.*
- (iv) *Reflection : Let $\psi: y \mapsto \varphi(-y)$. Then $\text{prox}_\psi x = -\text{prox}_\varphi(-x)$.*
- (v) *Moreau envelope : Let $\psi = \gamma\varphi$, where $\gamma \in]0, +\infty[$. Then*

$$\text{prox}_\psi x = x + \frac{1}{\gamma + 1} (\text{prox}_{(\gamma+1)\varphi} x - x). \quad (3.34)$$

Proof. We observe that in all cases $\psi \in \Gamma_0(\mathcal{X})$. Now set $p = \text{prox}_\psi x$. As seen in (3.29), this is equivalent to $x - p \in \partial\psi(p)$.

(i) : It follows from Lemma 3.3.3(i) and (3.29) that $x - p \in \partial\psi(p) \Leftrightarrow x - p \in \partial\varphi(p) + \alpha p + u \Leftrightarrow (x - u)/(\alpha + 1) - p \in \partial(\varphi/(\alpha + 1))(p) \Leftrightarrow p = \text{prox}_{\varphi/(\alpha+1)}((x - u)/(\alpha + 1))$.

(ii) : It follows from (3.29) that $x - p \in \partial\psi(p) \Leftrightarrow x - p \in \partial\varphi(p - z) \Leftrightarrow (x - z) - (p - z) \in \partial\varphi(p - z) \Leftrightarrow p - z = \text{prox}_\varphi(x - z)$.

(iii) : It follows from Lemma 3.3.3(i) and (3.29) that $x - p \in \partial\psi(p) \Leftrightarrow x - p \in \rho^{-1}\partial\varphi(p/\rho) \Leftrightarrow x/\rho - p/\rho \in \partial(\varphi/\rho^2)(p/\rho) \Leftrightarrow p = \rho \text{prox}_{\varphi/\rho^2}(x/\rho)$.

(iv) : Set $\rho = -1$ in (iii).

(v) : See [27, Lemma 2.2]. \square

Lemma 3.3.8 *Let $\psi = \|\cdot\|^2/(2\gamma) - \gamma\varphi$, where $\gamma \in]0, +\infty[$ and $\varphi \in \Gamma_0(\mathcal{X})$, and let $x \in \mathcal{X}$. Then $\psi \in \Gamma_0(\mathcal{X})$ and*

$$\text{prox}_\psi x = x - \frac{1}{\gamma} \text{prox}_{\frac{\gamma^2}{\gamma+1}\varphi} \left(\frac{\gamma x}{\gamma + 1} \right). \quad (3.35)$$

Proof. Let $\varrho = \gamma\varphi + \|\cdot\|^2/2$. Then clearly $\varrho \in \Gamma_0(\mathcal{X})$ and hence $\varrho^* \in \Gamma_0(\mathcal{X})$. However, since (3.18) and (3.28) imply that $\psi = \varrho^*/\gamma$, we obtain $\psi \in \Gamma_0(\mathcal{X})$. Let us also observe that Lemma 3.3.6 asserts that ψ is differentiable with gradient $\nabla\psi = \text{prox}_{\gamma\varphi}/\gamma$.

Consequently, it follows from (3.29) that

$$\begin{aligned}
p = \text{prox}_\psi x &\Leftrightarrow x - p = (\text{prox}_{\gamma\varphi} p)/\gamma \\
&\Leftrightarrow p - \gamma(x - p) \in \gamma\partial\varphi(\gamma(x - p)) \\
&\Leftrightarrow \frac{\gamma x}{\gamma + 1} - \gamma(x - p) \in \frac{\gamma^2}{\gamma + 1}\partial\varphi(\gamma(x - p)) \\
&\Leftrightarrow \gamma(x - p) = \text{prox}_{\frac{\gamma^2}{\gamma+1}\varphi} \left(\frac{\gamma x}{\gamma + 1} \right) \\
&\Leftrightarrow p = x - \frac{1}{\gamma} \text{prox}_{\frac{\gamma^2}{\gamma+1}\varphi} \left(\frac{\gamma x}{\gamma + 1} \right). \tag{3.36}
\end{aligned}$$

□

Lemma 3.3.9 *Let $\psi = \varphi \circ L$, where $\varphi \in \Gamma_0(\mathcal{Y})$ and where $L: \mathcal{X} \rightarrow \mathcal{Y}$ is a bijective bounded linear operator such that $L^{-1} = L^*$. Then $\text{prox}_\psi = L^* \circ \text{prox}_\varphi \circ L$.*

Proof. It follows from the assumptions that $\psi \in \Gamma_0(\mathcal{X})$. Now let $(x, p) \in \mathcal{X}^2$. Since L is surjective, Lemma 3.3.3(i) asserts that $\partial\psi = L^* \circ (\partial\varphi) \circ L$. Therefore, it follows from (3.29) that $p = \text{prox}_\psi x \Leftrightarrow x - p \in L^*(\partial\varphi(Lp)) \Leftrightarrow Lx - Lp \in \partial\varphi(Lp) \Leftrightarrow Lp = \text{prox}_\varphi(Lx) \Leftrightarrow p = L^*(\text{prox}_\varphi(Lx))$. □

Lemma 3.3.10 *Let $(\phi_k)_{1 \leq k \leq m}$ be functions in $\Gamma_0(\mathcal{X})$, let \mathcal{X}^m be the standard Hilbert product space, and let $\varphi: \mathcal{X}^m \rightarrow]-\infty, +\infty] : (x_k)_{1 \leq k \leq m} \mapsto \sum_{k=1}^m \phi_k(x_k)$. Then $\text{prox}_\varphi = (\text{prox}_{\phi_k})_{1 \leq k \leq m}$.*

Proof. It is clear that $\varphi \in \Gamma_0(\mathcal{X}^m)$. Now take $(x_k)_{1 \leq k \leq m}$ and $(p_k)_{1 \leq k \leq m}$ in \mathcal{X}^m . Then it follows from (3.29) and Lemma 3.3.2 that $(p_k)_{1 \leq k \leq m} = \text{prox}_\varphi(x_k)_{1 \leq k \leq m} \Leftrightarrow (x_k - p_k)_{1 \leq k \leq m} \in \partial\varphi(p_k)_{1 \leq k \leq m} = \times_{k=1}^m \partial\phi_k(p_k) \Leftrightarrow (p_k)_{1 \leq k \leq m} = (\text{prox}_{\phi_k} x_k)_{1 \leq k \leq m}$. □

3.3.2.5 Moreau's decomposition

Let V be a closed vector subspace of \mathcal{X} with orthogonal complement V^\perp . The standard orthogonal projection theorem, which has far reaching applications in signal theory, states that the energy of a signal $x \in \mathcal{X}$ can be decomposed as $\|x\|^2 = d_V^2(x) + d_{V^\perp}^2(x)$ and that x itself can be written as $x = P_V x + P_{V^\perp} x$, where $\langle P_V x \mid P_{V^\perp} x \rangle = 0$. If we set $\varphi = \iota_V$, then $\varphi^* = \iota_{V^\perp}$ and it follows from (3.31) that these identities become $\|x\|^2 = 2(\iota_V(x) + \iota_{V^\perp}(x))$ and $x = \text{prox}_\varphi x + \text{prox}_{\varphi^*} x$. Moreau has shown that, remarkably, this decomposition principle holds true for any $\varphi \in \Gamma_0(\mathcal{X})$.

Lemma 3.3.11 *Let $\varphi \in \Gamma_0(\mathcal{X})$, $\gamma \in]0, +\infty[$, and $x \in \mathcal{X}$. Then*

$$\|x\|^2 = 2\gamma(\gamma\varphi(x) + \iota_{V^\perp}(x/\gamma)) \tag{3.37}$$

and

$$x = x_\gamma^\oplus + x_\gamma^\ominus, \quad \text{where} \quad \begin{cases} x_\gamma^\oplus = \text{prox}_{\gamma\varphi} x \\ x_\gamma^\ominus = \gamma \text{prox}_{\varphi^*/\gamma}(x/\gamma). \end{cases} \quad (3.38)$$

Moreover,

$$\varphi(x_\gamma^\oplus) + \varphi^*(x_\gamma^\ominus/\gamma) = \langle x_\gamma^\oplus \mid x_\gamma^\ominus \rangle / \gamma. \quad (3.39)$$

Proof. Using (3.28) and applying Lemma 3.3.3(ii) with $\mathcal{Y} = \mathcal{X}$, $L = \text{Id}$, and $\psi: y \mapsto \|x - y\|^2/(2\gamma)$ (hence $\psi^*: v \mapsto \gamma\|v\|^2/2 + \langle x \mid v \rangle$ by (3.18)), we obtain

$$\begin{aligned} \gamma\varphi(x) &= \inf_{y \in \mathcal{X}} \varphi(y) + \psi(y) \\ &= -\min_{v \in \mathcal{X}} \varphi^*(v) + \psi^*(-v) \\ &= -\min_{v \in \mathcal{X}} \varphi^*(v) + \frac{\gamma}{2}\|v\|^2 - \langle x \mid v \rangle \\ &= \frac{1}{2\gamma}\|x\|^2 - \min_{v \in \mathcal{X}} \varphi^*(v) + \frac{\gamma}{2}\|(x/\gamma) - v\|^2 \\ &= \frac{1}{2\gamma}\|x\|^2 - {}^{1/\gamma}(\varphi^*)(x/\gamma), \end{aligned} \quad (3.40)$$

which establishes (3.37). Next, we obtain (3.38) by differentiating (3.37) using Lemma 3.3.6. Finally, we observe that (3.29) and (3.23) yield

$$\begin{aligned} x_\gamma^\oplus = \text{prox}_{\gamma\varphi} x &\Leftrightarrow x - x_\gamma^\oplus \in \gamma\partial\varphi(x_\gamma^\oplus) \\ &\Leftrightarrow x_\gamma^\ominus/\gamma \in \partial\varphi(x_\gamma^\oplus) \\ &\Leftrightarrow \varphi(x_\gamma^\oplus) + \varphi^*(x_\gamma^\ominus/\gamma) = \langle x_\gamma^\oplus \mid x_\gamma^\ominus/\gamma \rangle, \end{aligned} \quad (3.41)$$

which establishes (3.39). \square

Remark 3.3.12 Let us make a few remarks concerning Moreau's decomposition.

(i) For $\gamma = 1$, Lemma 3.3.11 provides the nicely symmetric formulas

$$\begin{cases} \|x\|^2 = 2({}^1\varphi(x) + {}^1(\varphi^*)(x)) \\ x = \text{prox}_\varphi x + \text{prox}_{\varphi^*} x \\ \varphi(\text{prox}_\varphi x) + \varphi^*(\text{prox}_{\varphi^*} x) = \langle \text{prox}_\varphi x \mid \text{prox}_{\varphi^*} x \rangle, \end{cases} \quad (3.42)$$

which correspond to Moreau's original setting; see [53, 55], where alternate proofs are given.

(ii) Let $\varphi = \iota_K$, where K is a closed convex cone in \mathcal{X} (recall that $K \subset \mathcal{X}$ is a convex cone if $K + K \subset K$ and $(\forall \alpha \in]0, +\infty[) \alpha K \subset K$). Then $\varphi^* = \iota_{K^\ominus}$,

where $K^\ominus = \{u \in \mathcal{X} \mid (\forall x \in K) \langle x \mid u \rangle \leq 0\}$ is the polar cone of K . In this case (3.42) becomes

$$\begin{cases} \|x\|^2 = d_K^2(x) + d_{K^\ominus}^2(x) \\ x = P_K x + P_{K^\ominus} x \\ \langle P_K x \mid P_{K^\ominus} x \rangle = 0. \end{cases} \quad (3.43)$$

We thus obtain a decomposition of x into two orthogonal signals $P_K x$ and $P_{K^\ominus} x$. In signal theory, such conical decompositions appear for instance in [14, 66, 74]. They of course subsume the usual linear orthogonal decompositions discussed at the beginning of this section. Moreau established (3.43) prior to (3.42) in [52].

(iii) We have derived (3.38) from the energy decomposition principle (3.37). An alternate derivation can be made using the theory of maximal monotone operators [24].

(iv) Using Lemma 3.3.7(iii), we can rewrite (3.38) as

$$x = x_\gamma^\oplus + x_\gamma^\ominus, \text{ where } x_\gamma^\oplus = \text{prox}_{\gamma\varphi} x \text{ and } x_\gamma^\ominus = \text{prox}_{\gamma\varphi^*(\cdot/\gamma)} x. \quad (3.44)$$

(v) Equation (3.38) describes a powerful (generally nonlinear) signal decomposition scheme parameterized by a function $\varphi \in \Gamma_0(\mathcal{X})$ and a scalar $\gamma \in]0, +\infty[$. Signal denoising applications of this result will be discussed in Section 3.3.6.

3.3.2.6 Examples of proximity operators

We provide a few examples of proximity operators that are of interest in signal recovery.

Example 3.3.13 Suppose that $\varphi = 0$ in Lemma 3.3.7(i). Then taking $\alpha = 0$ shows that the translation $x \mapsto x - u$ is a proximity operator, while taking $u = 0$ shows that the transformation $x \mapsto \kappa x$ is also a proximity operator for $\kappa \in]0, 1]$.

More generally, linear proximity operators are characterized as follows.

Example 3.3.14 [55, Section 3] Let $L: \mathcal{X} \rightarrow \mathcal{X}$ be a bounded linear operator. Then L is a proximity operator if and only if $L = L^*$, $\|L\| \leq 1$, and $(\forall x \in \mathcal{X}) \langle Lx \mid x \rangle \geq 0$.

We have already seen in (3.31) that convex projection operators are proximity operators. More generally, the following example states that underrelaxed convex projection operators are proximity operators.

Example 3.3.15 Let C be a nonempty closed convex subset of \mathcal{X} , let $\gamma \in]0, +\infty[$, and let $x \in \mathcal{X}$. Then $\text{prox}_{d_C^2/(2\gamma)} x = x + \frac{1}{\gamma+1}(P_C x - x)$.

Proof. The proof is a direct consequence of (3.31) and Lemma 3.3.7(v). \square

A hard-thresholding transformation with respect to set distance, i.e.,

$$x \mapsto \begin{cases} x, & \text{if } d_C(x) > \gamma; \\ P_C x, & \text{if } d_C(x) \leq \gamma, \end{cases} \quad (3.45)$$

is not continuous and can therefore not be performed via a proximity operator (see Lemma 3.3.5). However, as our next example shows, soft-thresholding transformations can.

Example 3.3.16 Let C be a nonempty closed convex subset of \mathcal{X} , let $\gamma \in]0, +\infty[$, and let $x \in \mathcal{X}$. Then

$$\text{prox}_{\gamma d_C} x = \begin{cases} x + \frac{\gamma}{d_C(x)}(P_C x - x), & \text{if } d_C(x) > \gamma; \\ P_C x, & \text{if } d_C(x) \leq \gamma. \end{cases} \quad (3.46)$$

Proof. Suppose that $p = \text{prox}_{\gamma d_C} x$ or, equivalently, that $x - p \in \gamma \partial d_C(p)$. Then, in view of (3.17) and (3.25), it follows from (3.26) that

$$p \in C \Rightarrow x - p \in N_C(p) \cap B(0; \gamma) \Rightarrow \begin{cases} p = P_C x \\ d_C(x) \leq \gamma \end{cases} \quad (3.47)$$

and, on the other hand, that

$$\begin{aligned} p \notin C &\Rightarrow x - p = \gamma \left(\frac{p - P_C p}{d_C(p)} \right) \\ &\Rightarrow x - P_C p = \left(1 + \frac{\gamma}{d_C(p)} \right) (p - P_C p) \in N_C(P_C p) \end{aligned} \quad (3.48)$$

$$\Rightarrow P_C x = P_C p. \quad (3.49)$$

Consequently, we rewrite (3.48) as

$$\begin{aligned} p \notin C &\Rightarrow x - P_C x = \left(1 + \frac{\gamma}{d_C(p)} \right) (p - P_C p) \\ &\Rightarrow \begin{cases} d_C(x) = d_C(p) + \gamma \\ p = x + \frac{\gamma}{d_C(x)} (P_C x - x). \end{cases} \end{aligned} \quad (3.50)$$

Now suppose that $d_C(x) > \gamma$. Then $p \notin C$ since otherwise (3.47) would yield $d_C(x) \leq \gamma$, which is absurd. The expression of p is then supplied by (3.50). Next, suppose

that $d_C(x) \leq \gamma$. Then $p \in C$ since (3.50) yields $p \notin C \Rightarrow d_C(p) = d_C(x) - \gamma \leq 0 \Rightarrow p \in \overline{C} = C$, which is absurd. The expression of p is then supplied by (3.47). \square

In the above example, C can be thought of as a set of signals possessing a certain property (see [21, 23, 29, 67, 74] for examples of closed convex sets modeling pertinent constraints in signal recovery). If the signal x is close enough to satisfying the property in question, then $\text{prox}_{\gamma d_C} x$ is simply the projection of x onto C ; otherwise, $\text{prox}_{\gamma d_C} x$ is obtained through a nonstationary underrelaxation of this projection. Here is an important special case.

Example 3.3.17 Suppose that $C = \{0\}$ in Example 3.3.16. Then (3.46) becomes

$$\text{prox}_{\gamma \|\cdot\|} x = \begin{cases} \left(1 - \frac{\gamma}{\|x\|}\right)x, & \text{if } \|x\| > \gamma; \\ 0, & \text{if } \|x\| \leq \gamma. \end{cases} \quad (3.51)$$

In particular, if $\mathcal{X} = \mathbb{R}$, it reduces to the well-known scalar soft-thresholding (also known as a shrinkage) operation

$$\text{prox}_{\gamma|\cdot|} x = \text{sign}(x) \max\{|x| - \gamma, 0\}. \quad (3.52)$$

From a numerical standpoint, Moreau's decomposition (3.38) provides an alternative means to compute $x_\gamma^\oplus = \text{prox}_{\gamma\varphi} x$. This is especially important in situations when it may be difficult to obtain x_γ^\oplus directly but when the dual problem of applying $\text{prox}_{\varphi^*/\gamma}$ is easier. We can then compute $x_\gamma^\oplus = x - \gamma \text{prox}_{\varphi^*/\gamma}(x/\gamma)$ or, using (3.44),

$$x_\gamma^\oplus = x - \text{prox}_{\gamma\varphi^*(\cdot/\gamma)} x. \quad (3.53)$$

The following example illustrates this point.

Example 3.3.18 Suppose that $\varphi: \mathcal{X} \rightarrow]-\infty, +\infty]$ is defined as

$$\varphi: x \mapsto \sup_{y \in D} \langle x | Ly \rangle, \quad (3.54)$$

where $L: \mathcal{Y} \rightarrow \mathcal{X}$ is a bounded linear operator and where D is a nonempty subset of \mathcal{Y} . Then $\varphi \in \Gamma_0(\mathcal{X})$. Now let C be the closed convex hull of $L(D)$. Then, using (3.20), we can write (more generally, any positively homogeneous function φ in $\Gamma_0(\mathcal{X})$ assumes this form [4, Theorem 2.4.2])

$$\varphi: x \mapsto \sup_{u \in C} \langle x | u \rangle = \sigma_C(x). \quad (3.55)$$

In turn, (3.21) yields $\varphi^* = \sigma_C^* = \iota_C$ and (3.31) asserts that, for every $x \in \mathcal{X}$, we can calculate x_γ^\oplus through a projection operation, since (3.53) becomes

$$x_\gamma^\oplus = x - \text{prox}_{\gamma\iota_C(\cdot/\gamma)} x = x - P_{\gamma C} x. \quad (3.56)$$

In the case when φ is the discrete total variation functional, this approach is used implicitly in [17].

We now provide an application of the product space setting described in Lemma 3.3.10.

Example 3.3.19 Let $\gamma \in]0, +\infty[$ and define a function $\phi \in \Gamma_0(\mathbb{R})$ by

$$\phi: \xi \mapsto \begin{cases} -\ln(\xi), & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \leq 0. \end{cases} \quad (3.57)$$

Then a straightforward calculation gives $(\forall \xi \in \mathbb{R}) \operatorname{prox}_{\gamma\phi} \xi = (\xi + \sqrt{\xi^2 + 4\gamma})/2$. Now let φ be the Burg entropy function on the Euclidean space \mathbb{R}^m , i.e., $\varphi: x = (\xi_k)_{1 \leq k \leq m} \mapsto \sum_{k=1}^m \phi(\xi_k)$. Then it follows from Lemma 3.3.10 that

$$(\forall x \in \mathbb{R}^m) \operatorname{prox}_{\gamma\varphi} x = \frac{1}{2} \left(\xi_k + \sqrt{\xi_k^2 + 4\gamma} \right)_{1 \leq k \leq m}. \quad (3.58)$$

Our last two examples will play a central role in Section 3.3.5.4.

Example 3.3.20 Let $(e_k)_{k \in \mathbb{N}}$ be an orthonormal basis of \mathcal{X} , let $(\phi_k)_{k \in \mathbb{N}}$ be functions in $\Gamma_0(\mathbb{R})$ such that

$$(\forall k \in \mathbb{N}) \phi_k \geq 0 \quad \text{and} \quad \phi_k(0) = 0, \quad (3.59)$$

and let $\psi: \mathcal{X} \rightarrow]-\infty, +\infty] : x \mapsto \sum_{k \in \mathbb{N}} \phi_k(\langle x | e_k \rangle)$. Then :

- (i) $\psi \in \Gamma_0(\mathcal{X})$.
- (ii) $(\forall x \in \mathcal{X}) \operatorname{prox}_{\psi} x = \sum_{k \in \mathbb{N}} (\operatorname{prox}_{\phi_k} \langle x | e_k \rangle) e_k$.

Proof. Let us introduce an operator

$$L: \mathcal{X} \rightarrow \ell^2(\mathbb{N}): x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{N}} \quad (3.60)$$

and a function

$$\varphi: \ell^2(\mathbb{N}) \rightarrow]-\infty, +\infty] : (\xi_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \phi_k(\xi_k). \quad (3.61)$$

From standard Hilbertian analysis, L is an invertible bounded linear operator with

$$L^{-1} = L^*: \ell^2(\mathbb{N}) \rightarrow \mathcal{X}: (\xi_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \xi_k e_k. \quad (3.62)$$

(i) : In view of the properties of L , since $\psi = \varphi \circ L$, it is enough to show that $\varphi \in \Gamma_0(\ell^2(\mathbb{N}))$. To this end, define, for every $K \in \mathbb{N}$, $\varphi_K = \sum_{k=0}^K \varrho_k$, where $\varrho_k: (\xi_l)_{l \in \mathbb{N}} \mapsto \phi_k(\xi_k)$. Then it follows from the assumptions that φ_K is lower semi-continuous and convex on $\ell^2(\mathbb{N})$ as a finite sum of such functions. Consequently (see

Section 3.3.1.1), the sets $(\text{epi } \varphi_K)_{K \in \mathbb{N}}$ are closed and convex in $\ell^2(\mathbb{N}) \times \mathbb{R}$. Therefore, since by assumption (3.59) the functions $(\varphi_K)_{K \in \mathbb{N}}$ are nonnegative, the set

$$\text{epi } \varphi = \text{epi} \left(\sup_{K \in \mathbb{N}} \varphi_K \right) = \bigcap_{K \in \mathbb{N}} \text{epi } \varphi_K \quad (3.63)$$

is also closed and convex as an intersection of closed convex sets. This shows that φ is lower semicontinuous and convex. Finally, since (3.59) implies that $\varphi(0) = 0$, we conclude that $\varphi \in \Gamma_0(\ell^2(\mathbb{N}))$.

(ii) : Fix $\mathbf{x} = (\xi_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})$. Now set $\mathbf{p} = \text{prox}_\varphi \mathbf{x}$ and $\mathbf{q} = (\pi_k)_{k \in \mathbb{N}}$, where $(\forall k \in \mathbb{N}) \pi_k = \text{prox}_{\phi_k} \xi_k$. Then, in view of Lemma 3.3.9 and (3.62), it suffices to show that $\mathbf{p} = \mathbf{q}$. Let us first observe that, for every $k \in \mathbb{N}$, (3.59) implies that 0 minimizes ϕ_k and therefore that $\text{prox}_{\phi_k} 0 = 0$. Consequently, it follows from the nonexpansivity of the operators $(\text{prox}_{\phi_k})_{k \in \mathbb{N}}$ (see Lemma 3.3.5) that

$$\sum_{k \in \mathbb{N}} |\pi_k|^2 = \sum_{k \in \mathbb{N}} |\text{prox}_{\phi_k} \xi_k - \text{prox}_{\phi_k} 0|^2 \leq \sum_{k \in \mathbb{N}} |\xi_k - 0|^2 = \|\mathbf{x}\|^2. \quad (3.64)$$

Hence $\mathbf{q} \in \ell^2(\mathbb{N})$. Now let $\mathbf{y} = (\eta_k)_{k \in \mathbb{N}}$ be an arbitrary point in $\ell^2(\mathbb{N})$. It follows from (3.29) and (3.22) that \mathbf{p} is the unique point in $\ell^2(\mathbb{N})$ that satisfies

$$\langle \mathbf{y} - \mathbf{p} \mid \mathbf{x} - \mathbf{p} \rangle + \varphi(\mathbf{p}) \leq \varphi(\mathbf{y}). \quad (3.65)$$

On the other hand, the same characterization for each point in $(\pi_k)_{k \in \mathbb{N}}$ yields

$$(\forall k \in \mathbb{N}) \quad (\eta_k - \pi_k)(\xi_k - \pi_k) + \phi_k(\pi_k) \leq \phi_k(\eta_k). \quad (3.66)$$

Summing these last inequalities over $k \in \mathbb{N}$, we obtain $\langle \mathbf{y} - \mathbf{q} \mid \mathbf{x} - \mathbf{q} \rangle + \varphi(\mathbf{q}) \leq \varphi(\mathbf{y})$. In view of the characterization (3.65), we conclude that $\mathbf{p} = \mathbf{q}$. \square

The following special case is the widely used soft-thresholder that will be discussed in Problem 3.3.56 and Example 3.3.61.

Example 3.3.21 Let $(e_k)_{k \in \mathbb{N}}$ be an orthonormal basis of \mathcal{X} , let $(\omega_k)_{k \in \mathbb{N}}$ be a sequence in $]0, +\infty[$, let $\psi: \mathcal{X} \rightarrow]-\infty, +\infty]: x \mapsto \sum_{k \in \mathbb{N}} \omega_k |\langle x \mid e_k \rangle|$, and let $x \in \mathcal{X}$. Then $\text{prox}_\psi x = \sum_{k \in \mathbb{N}} \pi_k e_k$, where

$$(\forall k \in \mathbb{N}) \quad \pi_k = \text{sign}(\langle x \mid e_k \rangle) \max\{|\langle x \mid e_k \rangle| - \omega_k, 0\}. \quad (3.67)$$

Proof. Set $\phi_k = \omega_k |\cdot|$ in Example 3.3.20 and use (3.52). \square

3.3.3 Properties and numerical solution of Problem 3.3.1

We begin with some basic properties of Problem 3.3.1. Recall that the set of solutions to this problem is denoted by G .

Proposition 3.3.22

- (i) *Existence* : Problem 3.3.1 possesses at least one solution if $f_1 + f_2$ is coercive, i.e.,

$$\lim_{\|x\| \rightarrow +\infty} f_1(x) + f_2(x) = +\infty. \quad (3.68)$$

- (ii) *Uniqueness* : Problem 3.3.1 possesses at most one solution if $f_1 + f_2$ is strictly convex. This occurs in particular when f_1 or f_2 is strictly convex. [26].
- (iii) *Characterization* : Let $x \in \mathcal{H}$ and $\gamma \in]0, +\infty[$. Then the following statements are equivalent :
- (a) x solves Problem 3.3.1.
 - (b) $x = \text{prox}_{\gamma f_1} (x - \gamma \nabla f_2(x))$.
 - (c) $(\forall y \in \mathcal{H}) \langle x - y \mid \nabla f_2(x) \rangle + f_1(x) \leq f_1(y)$.

Proof. (i) : The assumptions on Problem 3.3.1 and (3.68) imply that $f_1 + f_2$ lies in $\Gamma_0(\mathcal{H})$ and that it is coercive. Hence the claim follows from [76, Theorem 2.5.1(ii)].

(ii) : See [76, Proposition 2.5.6].

(iii) : It follows from Fermat's rule (3.24), Lemma 3.3.3(i), and (3.29) that

$$x \in G \Leftrightarrow 0 \in \partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) = \partial f_1(x) + \{\nabla f_2(x)\} \quad (3.69)$$

$$\Leftrightarrow -\nabla f_2(x) \in \partial f_1(x) \quad (3.70)$$

$$\Leftrightarrow (x - \gamma \nabla f_2(x)) - x \in \gamma \partial f_1(x)$$

$$\Leftrightarrow x = \text{prox}_{\gamma f_1} (x - \gamma \nabla f_2(x)). \quad (3.71)$$

Using (3.70) and (3.22), we see that $x \in G \Leftrightarrow (\forall y \in \mathcal{H}) \langle y - x \mid -\nabla f_2(x) \rangle + f_1(x) \leq f_1(y)$. \square

The fixed point characterization provided by Proposition 3.3.22(iii)(b) suggests solving Problem 3.3.1 via the fixed point iteration $x_{n+1} = \text{prox}_{\gamma f_1} (x_n - \gamma \nabla f_2(x_n))$ for a suitable value of the parameter γ . This iteration, which is referred to as a forward-backward splitting process in optimization, consists of two separate steps. First one performs a forward (explicit) step involving only f_2 to compute $x_{n+\frac{1}{2}} = x_n - \gamma \nabla f_2(x_n)$; then one performs a backward (implicit) step involving only f_1 to compute $x_{n+1} = \text{prox}_{\gamma f_1} x_{n+\frac{1}{2}}$. Formally, this second step amounts to solving the inclusion (3.29), hence its implicit nature. The following theorem is an adaption of some results from [25], which provides a more general iteration in which the coefficient γ is made iteration-dependent, errors are allowed in the evaluation of the operators $\text{prox}_{\gamma f_1}$ and ∇f_2 , and a relaxation sequence $(\lambda_n)_{n \in \mathbb{N}}$ is introduced. The errors allow for some tolerance in the numerical implementation of the algorithm, while the flexibility introduced by the iteration-dependent parameters γ_n and λ_n can be used to improve its convergence pattern.

First, we need to introduce the following condition.

Condition 3.3.23 Let X be a nonempty subset of a real Hilbert space \mathcal{X} . We say that a function $\varphi \in \Gamma_0(\mathcal{X})$ satisfies this condition on X if for all sequences $(y_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$ in \mathcal{X} and points $y \in X$ and $v \in \partial\varphi(y)$, we have

$$\left[y_n \rightharpoonup y, v_n \rightarrow v, (\forall n \in \mathbb{N}) v_n \in \partial\varphi(y_n) \right] \Rightarrow y \text{ is a strong cluster point of } (y_n)_{n \in \mathbb{N}}. \quad (3.72)$$

Remark 3.3.24 In Condition 3.3.23, the inclusion $v \in \partial\varphi(y)$ is redundant and stated only for the sake of clarity. Indeed, since $\varphi \in \Gamma_0(\mathcal{X})$, $\partial\varphi$ is maximal monotone [76, Theorem 3.1.11] and its graph is therefore sequentially weakly-strongly closed in $\mathcal{X} \times \mathcal{X}$ [4, Proposition 3.5.6.2]. Accordingly, the statements $y_n \rightharpoonup y$, $v_n \rightarrow v$, and $(\forall n \in \mathbb{N}) v_n \in \partial\varphi(y_n)$ imply that $v \in \partial\varphi(y)$.

Here is our main convergence result (recall that f_1 , f_2 , β , and G are defined in Problem 3.3.1).

Theorem 3.3.25 Suppose that $G \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2\beta$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set

$$x_{n+1} = x_n + \lambda_n \left(\text{prox}_{\gamma_n f_1} (x_n - \gamma_n (\nabla f_2(x_n) + b_n)) + a_n - x_n \right). \quad (3.73)$$

Then the following hold.

- (i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.
- (ii) $\sum_{n \in \mathbb{N}} \|\nabla f_2(x_n) - \nabla f_2(x)\|^2 < +\infty$.
- (iii) $\sum_{n \in \mathbb{N}} \|\text{prox}_{\gamma_n f_1} (x_n - \gamma_n \nabla f_2(x_n)) - x_n\|^2 < +\infty$.
- (iv) $(x_n)_{n \in \mathbb{N}}$ converges strongly to x if and only if $\underline{\lim} d_G(x_n) = 0$. In particular, strong convergence occurs in each of the following cases :
 - (a) $\text{int } G \neq \emptyset$.
 - (b) f_1 satisfies Condition 3.3.23 on G .
 - (c) f_2 satisfies Condition 3.3.23 on G .

Proof. It follows from (3.69) that

$$G = \{x \in \mathcal{H} \mid 0 \in \partial f_1(x) + \{\nabla f_2(x)\}\}. \quad (3.74)$$

Now let $A = \partial f_1$ and $B = \nabla f_2$. Since $f_1 \in \Gamma_0(\mathcal{H})$, [76, Theorem 3.1.11] asserts that A is maximal monotone. On the other hand since, by assumption, ∇f_2 is $1/\beta$ -Lipschitz continuous, it follows from [8, Corollaire 10] that βB is firmly nonexpansive.

(i) : Applying [25, Corollary 6.5], we obtain that $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in (A + B)^{-1}(0) = G$.

(ii)&(iii) : As in [25, Eq. (6.4)] set, for every $n \in \mathbb{N}$, $T_{1,n} = \text{prox}_{\gamma_n f_1}$, $\alpha_{1,n} = 1/2$, $T_{2,n} = \text{Id} - \gamma_n \nabla f_2$, and $\alpha_{2,n} = \gamma_n / (2\beta)$. Then [25, Remark 3.4] with $m = 2$ yields

$$\begin{cases} \sum_{n \in \mathbb{N}} \|(\text{Id} - T_{2,n})x_n - (\text{Id} - T_{2,n})x\|^2 < +\infty \\ \sum_{n \in \mathbb{N}} \|(T_{1,n} \circ T_{2,n})x_n - x_n\|^2 < +\infty. \end{cases} \quad (3.75)$$

The assumptions on $(\gamma_n)_{n \in \mathbb{N}}$ then provide the desired summability results.

(iv) : The characterization of strong convergence follows from [25, Theorem 3.3].

(iv)(a) : This is shown in [25, Remark 6.6].

(iv)(b) : Set $v = -\nabla f_2(x)$ and

$$(\forall n \in \mathbb{N}) \quad \begin{cases} y_n = \text{prox}_{\gamma_n f_1}(x_n - \gamma_n \nabla f_2(x_n)) \\ v_n = (x_n - y_n) / \gamma_n - \nabla f_2(x_n). \end{cases} \quad (3.76)$$

Then (3.29) yields $(\forall n \in \mathbb{N}) v_n \in \partial f_1(y_n)$. On the other hand, we derive from (i) and (iii) that $y_n \rightharpoonup x \in G$. Furthermore, since

$$(\forall n \in \mathbb{N}) \quad \|v_n - v\| \leq \frac{\|x_n - y_n\|}{\gamma_n} + \|\nabla f_2(x_n) - \nabla f_2(x)\|, \quad (3.77)$$

it follows from (ii), (iii), and the condition $\inf_{n \in \mathbb{N}} \gamma_n > 0$ that $v_n \rightarrow v$. It then results from Condition 3.3.23 that we can extract a subsequence $(y_{k_n})_{n \in \mathbb{N}}$ such that $y_{k_n} \rightarrow x$ and, in turn, from (iii) that $x_{k_n} \rightarrow x$. Accordingly, since $x \in G$, we have $d_G(x_{k_n}) \rightarrow 0$ and therefore $\underline{\lim} d_G(x_n) = 0$.

(iv)(c) : Set $v = \nabla f_2(x)$ and $(\forall n \in \mathbb{N}) v_n = \nabla f_2(x_n)$ (so certainly $v_n \in \partial f_2(x_n) = \{\nabla f_2(x_n)\}$). Then (i) yields $x_n \rightharpoonup x$ while (ii) yields $v_n \rightarrow v$. Therefore Condition 3.3.23 implies that $x \in G$ is a strong cluster point of $(x_n)_{n \in \mathbb{N}}$ and we conclude that $\underline{\lim} d_G(x_n) = 0$. \square

Remark 3.3.26 If $f_2 = 0$, $\lambda_n \equiv 1$, and $b_n \equiv 0$ in Theorem 3.3.25, we recover the *proximal point algorithm* and item (i), which states that $(x_n)_{n \in \mathbb{N}}$ converges weakly to a minimizer of f_1 , follows from [61, Theorem 1].

Further special cases of Theorem 3.3.25(iv)(b)&(iv)(c) can be constructed from the following proposition.

Proposition 3.3.27 *Let \mathcal{X} be a real Hilbert space. Suppose that $\varphi \in \Gamma_0(\mathcal{X})$ and that $\emptyset \neq X \subset D$, where $D = \text{dom } \varphi$. Let \mathcal{C} be the set of all nondecreasing functions from $[0, +\infty[$ to $[0, +\infty[$ that vanish only at 0. Then φ satisfies Condition 3.3.23 on X in each of the following cases :*

- (i) D is boundedly relatively compact (the closure of its intersection with any closed ball is compact).
- (ii) φ is differentiable on \mathcal{X} and $\text{Id} - \nabla\varphi$ is demicompact [77, Section 10.4] : for every bounded sequence $(y_n)_{n \in \mathbb{N}}$ in \mathcal{X} such that $(\nabla\varphi(y_n))_{n \in \mathbb{N}}$ converges strongly, $(y_n)_{n \in \mathbb{N}}$ admits a strong cluster point.
- (iii) For every $y \in X$ and $v \in \partial\varphi(y)$ there exists a function $c \in \mathcal{C}$ such that

$$(\forall x \in D) \quad \langle x - y \mid v \rangle + \varphi(y) + c(\|x - y\|) \leq \varphi(x). \quad (3.78)$$

- (iv) φ is uniformly convex at every point in X : for every $y \in X$ there exists a function $c \in \mathcal{C}$ such that, for every $x \in D$,

$$(\forall \alpha \in]0, 1[) \quad \varphi(\alpha x + (1 - \alpha)y) + \alpha(1 - \alpha)c(\|x - y\|) \leq \alpha\varphi(x) + (1 - \alpha)\varphi(y). \quad (3.79)$$

- (v) φ is uniformly convex : there exists a function $c \in \mathcal{C}$ such that, for every x and y in D , (3.79) holds.
- (vi) φ is uniformly convex on bounded sets : for every bounded convex set $C \subset \mathcal{X}$, $\varphi + \iota_C$ is uniformly convex, i.e., there exists a function $c \in \mathcal{C}$ such that, for every x and y in $C \cap D$, (3.79) holds.
- (vii) φ is strongly convex.

Proof. Take sequences $(y_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$ in \mathcal{X} and points $y \in X$ and $v \in \partial\varphi(y)$ such that $y_n \rightharpoonup y$, $v_n \rightarrow v$, and $(\forall n \in \mathbb{N}) v_n \in \partial\varphi(y_n)$.

(i) : The sequence $(y_n)_{n \in \mathbb{N}}$ is bounded (since it converges weakly) and lies in $\text{dom } \partial\varphi \subset D$. It therefore lies in a compact set and, as a result, y must be a strong cluster point.

(ii) : The sequence $(y_n)_{n \in \mathbb{N}}$ is bounded and, since φ is differentiable, $(\forall n \in \mathbb{N}) \nabla\varphi(y_n) = v_n \rightarrow v$. Hence the demicompactness assumption implies that we can extract a subsequence $(y_{k_n})_{n \in \mathbb{N}}$ that converges strongly. Since $y_n \rightharpoonup y$, we conclude that $y_{k_n} \rightarrow y$.

(iii) : It follows from (3.78) that

$$(\forall n \in \mathbb{N}) \quad \langle y_n - y \mid v \rangle + \varphi(y) + c(\|y_n - y\|) \leq \varphi(y_n). \quad (3.80)$$

On the other hand, it follows from (3.22) that

$$(\forall n \in \mathbb{N}) \quad \langle y - y_n \mid v_n \rangle + \varphi(y_n) \leq \varphi(y). \quad (3.81)$$

Adding these two inequalities, we obtain

$$(\forall n \in \mathbb{N}) \quad c(\|y_n - y\|) \leq \langle y_n - y \mid v_n - v \rangle. \quad (3.82)$$

However, since $y_n \rightharpoonup y$ and $v_n \rightarrow v$, we have $\langle y_n - y \mid v_n - v \rangle \rightarrow 0$. Therefore the assumptions on c and (3.82) yield $\|y_n - y\| \rightarrow 0$.

(iv) : For every x in D , we have [76, Section 3.5]

$$(3.79) \quad \Rightarrow \quad \langle x - y \mid v \rangle + \varphi(y) + c(\|x - y\|) \leq \varphi(x). \quad (3.83)$$

Hence (iv) is a special case of (iii).

(v) : This is a special case of (iv).

(vi) : Since $y_n \rightharpoonup y$, $(y_n)_{n \in \mathbb{N}}$ and y lie in some closed ball C . However since $f + \iota_C$ is uniformly convex, there exists $c \in \mathcal{C}$ such that (3.79) holds true for every $x \in C \cap D$. Thus, we deduce from (3.83) that (3.80) is satisfied, and we conclude as in (iii).

(vii) : This is a special case of (v) with $c: t \mapsto \rho t^2/2$ for some $\rho \in]0, +\infty[$ [76, Section 3.5]. \square

Examples of functions satisfying the various types of uniform convexity defined above can be found in [12, 75].

3.3.4 Problems involving sums of Moreau envelopes

3.3.4.1 Problem statement

We consider the following formulation, which is based on the notion of a Moreau envelope defined in (3.28).

Problem 3.3.28 Let

- (i) $(\mathcal{K}_i)_{1 \leq i \leq m}$ be real Hilbert spaces ;
- (ii) for every $i \in \{1, \dots, m\}$, $L_i: \mathcal{H} \rightarrow \mathcal{K}_i$ be a nonzero bounded linear operator, $\varphi_i \in \Gamma_0(\mathcal{K}_i)$, and $\rho_i \in]0, +\infty[$;
- (iii) $f_1 \in \Gamma_0(\mathcal{H})$.

The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f_1(x) + \sum_{i=1}^m \rho_i \varphi_i(L_i x). \quad (3.84)$$

The set of solutions to this problem is denoted by G .

Proposition 3.3.29 *Problem 3.3.28 is a special case of Problem 3.3.1 with $f_2 = \sum_{i=1}^m (\rho_i \varphi_i) \circ L_i$ and $\beta = (\sum_{i=1}^m \|L_i\|^2 / \rho_i)^{-1}$.*

Proof. Set

$$f_2 = \sum_{i=1}^m (\rho_i \varphi_i) \circ L_i \quad \text{and} \quad \beta = \left(\sum_{i=1}^m \|L_i\|^2 / \rho_i \right)^{-1}. \quad (3.85)$$

Since, for every $i \in \{1, \dots, m\}$, the function $\rho^i \varphi_i$ is finite, continuous, and convex, it belongs to $\Gamma_0(\mathcal{K}_i)$ and therefore $(\rho^i \varphi_i) \circ L_i \in \Gamma_0(\mathcal{H})$. Consequently, f_2 belongs to $\Gamma_0(\mathcal{H})$. Now, set $(\forall i \in \{1, \dots, m\}) T_i = \text{Id} - \text{prox}_{\rho^i \varphi_i}$. As seen in Lemma 3.3.5, the operators $(T_i)_{1 \leq i \leq m}$ are (firmly) nonexpansive. Therefore, for every $i \in \{1, \dots, m\}$, we obtain

$$\begin{aligned} (\forall (x, y) \in \mathcal{H}^2) \quad \|(L_i^* \circ T_i \circ L_i)x - (L_i^* \circ T_i \circ L_i)y\| &\leq \|L_i^*\| \cdot \|T_i(L_i x) - T_i(L_i y)\| \\ &\leq \|L_i^*\| \cdot \|L_i x - L_i y\| \\ &\leq \|L_i^*\| \cdot \|L_i\| \cdot \|x - y\| \\ &= \|L_i\|^2 \cdot \|x - y\|. \end{aligned} \quad (3.86)$$

On the other hand, we derive from Lemma 3.3.6 that

$$\nabla f_2 = \sum_{i=1}^m \nabla((\rho^i \varphi_i) \circ L_i) = \sum_{i=1}^m L_i^* \circ \left(\frac{T_i}{\rho_i}\right) \circ L_i = \sum_{i=1}^m \frac{1}{\rho_i} L_i^* \circ T_i \circ L_i. \quad (3.87)$$

Since (3.86) states that each operator $L_i^* \circ T_i \circ L_i$ is Lipschitz continuous with constant $\|L_i\|^2$, it ensues that ∇f_2 is Lipschitz continuous with constant $\sum_{i=1}^m \|L_i\|^2 / \rho_i$. We conclude that ∇f_2 is $1/\beta$ -Lipschitz continuous. \square

3.3.4.2 Properties and numerical solution of Problem 3.3.28

The following is a specialization of Theorem 3.3.25, in which we omit items (ii) and (iii) for the sake of brevity (special cases of item (ii) below can be derived from Theorem 3.3.25 and Proposition 3.3.27). The algorithm allows for the inexact computation of each proximity operator.

Theorem 3.3.30 *Suppose that $G \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2(\sum_{i=1}^m \|L_i\|^2 / \rho_i)^{-1}$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $((b_{i,n})_{n \in \mathbb{N}})_{1 \leq i \leq m}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\max_{1 \leq i \leq m} \sum_{n \in \mathbb{N}} \|b_{i,n}\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$\begin{aligned} x_{n+1} &= x_n + \\ &\lambda_n \left(\text{prox}_{\gamma_n f_1} \left(x_n + \gamma_n \left(\sum_{i=1}^m \frac{1}{\rho_i} ((L_i^* \circ (\text{prox}_{\rho^i \varphi_i} - \text{Id}) \circ L_i) x_n + b_{i,n}) \right) \right) + a_n - x_n \right). \end{aligned} \quad (3.88)$$

Then :

- (i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.
- (ii) $(x_n)_{n \in \mathbb{N}}$ converges strongly to x if and only if $\liminf d_G(x_n) = 0$.

Proof. The proof is a consequence of Proposition 3.3.29 and Theorem 3.3.25(i)&(iv) with $b_n = -\sum_{i=1}^m b_{i,n} / \rho_i$ and ∇f_2 given by (3.87). \square

3.3.4.3 Proximal split feasibility problems

We shall call the special case of Problem 3.3.28 when $m = 1$ a *proximal split feasibility problem*. In other words, we are given a real Hilbert space \mathcal{H} , a nonzero bounded linear operator $L: \mathcal{H} \rightarrow \mathcal{K}$, a function $f_1 \in \Gamma_0(\mathcal{H})$, a function $\varphi \in \Gamma_0(\mathcal{K})$, and a real number $\rho \in]0, +\infty[$. The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f_1(x) + \rho\varphi(Lx). \quad (3.89)$$

We denote by G the set of solutions to this problem.

Applying Theorem 3.3.30 with $m = 1$, we obtain at once the following convergence result.

Corollary 3.3.31 *Suppose that $G \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2\rho/\|L\|^2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1[$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(\text{prox}_{\gamma_n f_1} \left(x_n + \frac{\gamma_n}{\rho} ((L^* \circ (\text{prox}_{\rho\varphi} - \text{Id}) \circ L)x_n + b_n) \right) + a_n - x_n \right). \quad (3.90)$$

Then :

- (i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.
- (ii) $(x_n)_{n \in \mathbb{N}}$ converges strongly to x if and only if $\underline{\lim} d_G(x_n) = 0$.

Now, let us specialize the above setting to the case when $\rho = 1$, $f_1 = \iota_C$ and $\varphi = \iota_Q$, where $C \subset \mathcal{H}$ and $Q \subset \mathcal{K}$ are two nonempty closed convex sets. Then, in view of (3.31), (3.89) becomes

$$\underset{x \in C}{\text{minimize}} \quad d_Q(Lx). \quad (3.91)$$

In other words, one seeks a signal $x \in C$ such that the signal Lx is at minimal distance from Q ; in particular, when $C \cap L^{-1}(Q) \neq \emptyset$, one seeks a signal in $x \in C$ such that $Lx \in Q$. This is the so-called split feasibility problem introduced in [15] and further discussed in [13]. Let us observe that one of the earliest occurrence of this formulation is actually that provided by Youla in [73]. In that paper, the problem was to find a signal x in a closed vector subspace C , knowing its projection p onto a closed vector subspace V (hence $L = P_V$ and $Q = \{p\}$); it was also observed that the standard signal extrapolation schemes of Gerchberg [37] and Papoulis [59] fitted this framework.

In the present setting, Corollary 3.3.31(i) reduces to the following corollary.

Corollary 3.3.32 *Suppose that the set G of solutions to (3.91) is nonempty. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/\|L\|^2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(P_C \left(x_n + \gamma_n ((L^* \circ (P_Q - \text{Id}) \circ L)x_n + b_n) \right) + a_n - x_n \right). \quad (3.92)$$

Then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.

Remark 3.3.33 Corollary 3.3.32 improves upon [13, Theorem 2.1], where the additional assumptions $\dim \mathcal{H} < +\infty$, $\dim \mathcal{K} < +\infty$, $\lambda_n \equiv 1$, $\gamma_n \equiv \gamma \in]0, 2/\|L\|^2[$, $a_n \equiv 0$, and $b_n \equiv 0$ were made.

3.3.4.4 The $u + v$ signal decomposition model

Underlying many signal recovery problems is the decomposition of a signal $x \in \mathcal{H}$ as $x = u + v$, where u captures the geometric components of the signal (typically a function with bounded variations) and v models texture (typically an oscillatory function), e.g., [5, 6, 7, 51, 57, 71, 72]. The variational formulations proposed in [5, 6, 7, 71, 72] to achieve this decomposition based on a noisy observation $z \in \mathcal{H}$ of the signal of interest are of the general form

$$\underset{(u,v) \in \mathcal{H} \times \mathcal{H}}{\text{minimize}} \quad \psi(u) + \phi(v) + \frac{1}{2\rho} \|u + v - z\|^2, \quad (3.93)$$

where ψ and ϕ are in $\Gamma_0(\mathcal{H})$ and $\rho \in]0, +\infty[$. In order to cast this problem in our framework, let us introduce the function

$$\varphi: \mathcal{H} \rightarrow]-\infty, +\infty]: w \mapsto \phi(z - w). \quad (3.94)$$

Then $\varphi \in \Gamma_0(\mathcal{H})$ and the change of variable

$$w = z - v \quad (3.95)$$

in (3.93) yields

$$\underset{(u,w) \in \mathcal{H} \times \mathcal{H}}{\text{minimize}} \quad \psi(u) + \varphi(w) + \frac{1}{2\rho} \|u - w\|^2. \quad (3.96)$$

In view of (3.28), this problem can be rewritten in terms of the variable u as

$$\underset{u \in \mathcal{H}}{\text{minimize}} \quad \psi(u) + {}^{\rho}\varphi(u). \quad (3.97)$$

In other words, we obtain precisely the formulation (3.89) with $f_1 = \psi$, $\mathcal{K} = \mathcal{H}$, and $L = \text{Id}$.

We now derive from Corollary 3.3.31 and some facts from [9] the following result.

Corollary 3.3.34 *Suppose that (3.93) has at least one solution. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2\rho$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $u_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$u_{n+1} = u_n + \lambda_n \left(\text{prox}_{\gamma_n \psi} \left(u_n + \frac{\gamma_n}{\rho} (z - \text{prox}_{\rho\phi}(z - u_n) - u_n + b_n) \right) + a_n - u_n \right). \quad (3.98)$$

Then $(u_n)_{n \in \mathbb{N}}$ converges weakly to a solution u to (3.97) and $(u, \text{prox}_{\rho\phi}(z - u))$ is a solution to (3.93).

Proof. By assumption, the set G of solutions to (3.97) is nonempty. As noted above, (3.97) is a special case of (3.89) with $f_1 = \psi$, $\mathcal{K} = \mathcal{H}$, and $L = \text{Id}$. Moreover, in this case, (3.90) reduces to

$$u_{n+1} = u_n + \lambda_n \left(\text{prox}_{\gamma_n \psi} \left(u_n + \frac{\gamma_n}{\rho} (\text{prox}_{\rho\phi} u_n - u_n + b_n) \right) + a_n - u_n \right). \quad (3.99)$$

However, using (3.94) and Lemma 3.3.7(ii)&(iv), we obtain

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\rho\phi} x = z - \text{prox}_{\rho\phi}(z - x). \quad (3.100)$$

Therefore, (3.99) coincides with (3.98). Hence, since $\|L\| = 1$, we derive from Corollary 3.3.31 that the sequence $(u_n)_{n \in \mathbb{N}}$ converges weakly to a point $u \in G$. It then follows from [9, Propositions 3.2 and 4.1] that $(u, \text{prox}_{\rho\phi} u)$ is a solution to (3.96). In view of (3.100), this means that (u, w) is a solution to (3.96), where $w = z - \text{prox}_{\rho\phi}(z - u)$. Upon invoking the change of variable (3.95), we conclude that (u, v) is a solution to (3.93), where $v = z - w = \text{prox}_{\rho\phi}(z - u)$. \square

Remark 3.3.35 Consider the particular case when $\lambda_n \equiv 1$, $\gamma_n \equiv \rho$, $a_n \equiv 0$, and $b_n \equiv 0$. Then (3.98) becomes

$$u_{n+1} = \text{prox}_{\rho\psi} (z - \text{prox}_{\rho\phi}(z - u_n)). \quad (3.101)$$

Let us further assume, as in [5], that ψ is the support function of some nonempty closed convex set $K \subset \mathcal{H}$ and that ϕ is the indicator function of μK for some $\mu \in]0, +\infty[$. Then, since $\psi = \sigma_K$, it follows from (3.56) that $\text{prox}_{\rho\psi} = \text{Id} - P_{\rho K}$. On the other hand, since $\phi = \iota_{\mu K}$, (3.31) asserts that $\text{prox}_{\rho\phi} = P_{\mu K}$. Altogether, (3.101) becomes

$$u_{n+1} = z - P_{\mu K}(z - u_n) - P_{\rho K}(z - P_{\mu K}(z - u_n)). \quad (3.102)$$

This is precisely the iteration proposed in [5].

Remark 3.3.36 Problem (3.96) was originally studied in [1] and recently revisited in a broader context in [9]. The reader will find in the latter further properties, in particular from the viewpoint of duality.

3.3.4.5 Hard-constrained signal feasibility problems

Suppose that in Problem 3.3.28 we set $\mathcal{K}_i \equiv \mathcal{H}$, $L_i \equiv \text{Id}$, $f_1 = \iota_C$, and, for every $i \in \{1, \dots, m\}$, $\omega_i = 1/\rho_i$ and $\varphi_i = \iota_{C_i}$, where C and $(C_i)_{1 \leq i \leq m}$ are nonempty closed convex subsets of \mathcal{H} . Then, in view of (3.31), we obtain the so-called hard-constrained signal feasibility problem proposed in [26] to deal with inconsistent signal feasibility problems, namely

$$\underset{x \in C}{\text{minimize}} \quad \frac{1}{2} \sum_{i=1}^m \omega_i d_{C_i}^2(x). \quad (3.103)$$

We shall assume, without loss of generality, that $\sum_{i=1}^m \omega_i = 1$. In other words, (3.103) aims at producing a signal that satisfies the hard constraint modeled by C and that is closest, in a least-square distance sense, to satisfying the remaining constraints modeled by $(C_i)_{1 \leq i \leq m}$. In particular, if $C = \mathcal{H}$, one recovers the framework discussed in [22], where $x \mapsto \sum_{i=1}^m \omega_i d_{C_i}^2(x)/2$ was called a proximity function. Another example is when $m = 1$, i.e., when one seeks a signal $x \in C$ at minimal distance from C_1 . This setting is discussed in [38, 60]. Let us now specialize Theorem 3.3.30(i) (strong convergence follows as in Theorem 3.3.30(ii)) to the current hypotheses.

Corollary 3.3.37 *Suppose that the set G of solutions to (3.103) is nonempty. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1[$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $((b_{i,n})_{n \in \mathbb{N}})_{1 \leq i \leq m}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\max_{1 \leq i \leq m} \sum_{n \in \mathbb{N}} \|b_{i,n}\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(P_C \left(x_n + \gamma_n \left(\sum_{i=1}^m \omega_i (P_i x_n + b_{i,n}) - x_n \right) \right) + a_n - x_n \right). \quad (3.104)$$

Then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.

Remark 3.3.38 When $\gamma_n \equiv \gamma \in]0, 2[$, $b_{i,n} \equiv 0$, and $a_n \equiv 0$, Corollary 3.3.37 captures the scenario of [26, Proposition 9], which itself contains [22, Theorem 4] (where $C = \mathcal{H}$), and the convergence result of [38] (where $m = 1$).

3.3.5 Linear inverse problems

3.3.5.1 Problem statement

In Section 3.3.1, we have described the signal recovery problem as that of inferring a signal \bar{x} in a real Hilbert space \mathcal{H} from the observation of a signal z in a

real Hilbert space \mathcal{G} . In this section, we consider the standard linear data formation model in which z is related to \bar{x} via the model

$$z = T\bar{x} + w, \tag{3.105}$$

where $T: \mathcal{H} \rightarrow \mathcal{G}$ is a linear operator and where $w \in \mathcal{G}$ stands for an additive noise perturbation. This model covers numerous signal and image restoration and reconstruction prescriptions [2, 16, 23, 39, 66, 67]. The problem under consideration will be the following.

Problem 3.3.39 Let

- (i) \mathcal{K} be a real Hilbert space;
- (ii) $T: \mathcal{H} \rightarrow \mathcal{G}$ be a nonzero bounded linear operator;
- (iii) $L: \mathcal{H} \rightarrow \mathcal{K}$ be a bijective bounded linear operator such that $L^{-1} = L^*$;
- (iv) $f \in \Gamma_0(\mathcal{K})$.

The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(Lx) + \frac{1}{2} \|Tx - z\|^2. \tag{3.106}$$

The set of solutions to this problem is denoted by G .

In Problem 3.3.39, the term $\|Tx - z\|^2/2$ is a so-called data fidelity term which attempts to reflect the contribution of the data formation model (3.105), while the term $f(Lx)$ promotes prior knowledge about the original signal \bar{x} . This formulation covers various instances of linear inverse problems in signal recovery. Two specific frameworks will be discussed in Sections 3.3.5.3 and 3.3.5.4; other important examples are the following :

- In discrete models, the underlying Hilbert spaces are Euclidean spaces. If $\mathcal{K} = \mathcal{H}$, $L = \text{Id}$, and w is a realization of a multivariate zero mean Gaussian noise, then (3.106) with a suitable norm covers maximum *a posteriori* models with an *a priori* Gibbs density $p \propto \exp(-f)$. This setting is discussed in [68, 69].
- Let $\mathcal{K} = \mathcal{H} = H^1(\Omega)$, where Ω is an open domain of \mathbb{R}^m , let $L = \text{Id}$, and let f be an integral functional of the form

$$f: x \mapsto \gamma \int_{\Omega} \varphi(\omega, x(\omega), \nabla x(\omega)) d\omega, \tag{3.107}$$

where $\gamma \in]0, +\infty[$. Then (3.106) covers a variety of formulations, including total variation, least-squares, Fisher information, and entropic methods, e.g., [3, 19, 32, 40, 45]. Let us add that this framework also corresponds to the

Lagrangian formulation of the problems of [2, 42, 43, 56, 62, 70], the original form of which is

$$\underset{\|Tx-z\|^2 \leq \eta}{\text{minimize}} \int_{\Omega} \varphi(\omega, x(\omega), \nabla x(\omega)) d\omega, \quad (3.108)$$

where $\eta \in]0, +\infty[$. In this case, the parameter γ in (3.107) is the reciprocal of the Lagrange multiplier.

- In the Fourier regularization methods of [46, 50], $\mathcal{H} = L^2(\mathbb{R}^2)$, $\mathcal{K} = \mathcal{H} \times \mathcal{H}$, L is the Fourier transform, and $f: y \mapsto \gamma \|yh\|^2$, where h is the frequency response of a filter and $\gamma \in]0, +\infty[$.

3.3.5.2 Properties and numerical solution of Problem 3.3.39

Our analysis will be greatly simplified by the following observation.

Proposition 3.3.40 *Problem 3.3.39 is a special case of Problem 3.3.1 with $f_1 = f \circ L$, $f_2: x \mapsto \|Tx - z\|^2/2$, and $\beta = 1/\|T\|^2$.*

Proof. Set $f_1 = f \circ L$ and $f_2: x \mapsto \|Tx - z\|^2/2$. Then it follows from assumptions (i)–(iv) above that f_1 and f_2 are in $\Gamma_0(\mathcal{H})$, and that f_2 is differentiable on \mathcal{H} with $\nabla f_2: x \mapsto T^*(Tx - z)$. Consequently,

$$(\forall (x, y) \in \mathcal{H}^2) \quad \|\nabla f_2(x) - \nabla f_2(y)\| = \|T^*T(x - y)\| \leq \|T\|^2 \|x - y\|, \quad (3.109)$$

and ∇f_2 is therefore Lipschitz continuous with constant $\|T\|^2$. \square

Let us first provide existence and uniqueness conditions for Problem 3.3.39, as well as characterizations for its solutions.

Proposition 3.3.41

- (i) *Problem 3.3.39 possesses at least one solution if f is coercive.*
- (ii) *Problem 3.3.39 possesses at most one solution if one of the following conditions is satisfied :*
 - (a) *f is strictly convex.*
 - (b) *T is injective.*
- (iii) *Problem 3.3.39 possesses exactly one solution if T is bounded below, i.e.,*

$$(\exists \kappa \in]0, +\infty[)(\forall x \in \mathcal{H}) \quad \|Tx\| \geq \kappa \|x\|. \quad (3.110)$$

- (iv) *Let $x \in \mathcal{H}$ and $\gamma \in]0, +\infty[$. Then the following statements are equivalent :*
 - (a) *x solves Problem 3.3.39.*
 - (b) *$x = (L^* \circ \text{prox}_{\gamma f} \circ L)(x + \gamma T^*(z - Tx))$.*

$$(c) \quad (\forall y \in \mathcal{H}) \quad \langle Ty - Tx \mid z - Tx \rangle + f(Lx) \leq f(Ly).$$

Proof. Let f_1 and f_2 be as in Proposition 3.3.40.

(i) : In view of Proposition 3.3.22(i), it is enough to show that $f_1 + f_2$ is coercive. We have $f_1 + f_2 \geq f \circ L$. Moreover, since f is coercive, it follows from assumption (iii) in Problem 3.3.39 that $f \circ L$ is likewise. This shows the coercivity of $f_1 + f_2$.

(ii) : This follows from Proposition 3.3.22(ii) since, in item (ii)(a), f_1 is strictly convex by injectivity of L and, in item (ii)(b), f_2 is strictly convex. To show the latter, consider two distinct points x and y in \mathcal{H} and let $\alpha \in]0, 1[$. Then, by (3.110),

$$\begin{aligned} f_2(\alpha x + (1 - \alpha)y) &= \|\alpha(Tx - z) + (1 - \alpha)(Ty - z)\|^2/2 \\ &= \alpha\|Tx - z\|^2/2 + (1 - \alpha)\|Ty - z\|^2/2 \\ &\quad - \alpha(1 - \alpha)\|T(x - y)\|^2/2 \\ &\leq \alpha f_2(x) + (1 - \alpha)f_2(y) - \kappa^2\alpha(1 - \alpha)\|x - y\|^2/2 \quad (3.111) \\ &< \alpha f_2(x) + (1 - \alpha)f_2(y). \end{aligned}$$

(iii) : It follows from (3.110) that T is injective. Therefore, by (ii)(b), there is at most one solution. Regarding existence, Proposition 3.3.22(i) asserts that it suffices to show that $f_1 + f_2$ is coercive. Since $f \in \Gamma_0(\mathcal{K})$, it is minorized by a continuous affine functional [76, Theorem 2.2.6(iii)], say $\langle \cdot \mid u \rangle + \eta/2$ where $u \in \mathcal{K} \setminus \{0\}$ and $\eta \in \mathbb{R}$. Hence, we derive from (3.110) that

$$\begin{aligned} (\forall x \in \mathcal{H}) \quad &2(f_1(x) + f_2(x)) \\ &\geq 2\langle Lx \mid u \rangle + \eta + \|Tx - z\|^2 \\ &= 2\langle x \mid L^*u \rangle + \eta + \|Tx\|^2 - 2\langle x \mid T^*z \rangle + \|z\|^2 \\ &= \|x + L^*u - T^*z\|^2 + (\|Tx\|^2 - \|x\|^2) - \|L^*u - T^*z\|^2 + \|z\|^2 + \eta \quad (3.112) \\ &\geq (\|x\| - \|L^*u - T^*z\|)^2 + (\kappa^2 - 1)\|x\|^2 - \|L^*u - T^*z\|^2 + \|z\|^2 + \eta \\ &\geq (\kappa\|x\| - \|L^*u - T^*z\|/\kappa)^2 - \|L^*u - T^*z\|^2/\kappa^2 + \|z\|^2 + \eta, \end{aligned}$$

and we obtain $\lim_{\|x\| \rightarrow +\infty} f_1(x) + f_2(x) = +\infty$.

(iv) : This follows from Proposition 3.3.22(iii) and Lemma 3.3.9. \square

injective Next, we turn our attention to the stability of the solutions to Problem 3.3.39 with respect to perturbations of the observed data z .

Proposition 3.3.42 *Suppose that T satisfies (3.110). Let \tilde{z} be a point in \mathcal{G} , and let x and \tilde{x} be the unique solutions to Problem 3.3.39 associated with z and \tilde{z} , respectively. Then*

$$\|x - \tilde{x}\| \leq \|z - \tilde{z}\|/\kappa. \quad (3.113)$$

Proof. The existence and uniqueness of x and \tilde{x} follow from Proposition 3.3.41(iii). Next, we derive from Proposition 3.3.41(iv)(c) that

$$\begin{cases} \langle T\tilde{x} - Tx \mid z - Tx \rangle + f(Lx) \leq f(L\tilde{x}) \\ \langle Tx - T\tilde{x} \mid \tilde{z} - T\tilde{x} \rangle + f(L\tilde{x}) \leq f(Lx). \end{cases} \quad (3.114)$$

Adding these two inequalities, we obtain $\|T(x - \tilde{x})\|^2 \leq \langle T(x - \tilde{x}) \mid z - \tilde{z} \rangle$ and, by the Cauchy-Schwarz inequality, $\|T(x - \tilde{x})\| \leq \|z - \tilde{z}\|$. Using (3.110), we conclude that $\kappa\|x - \tilde{x}\| \leq \|z - \tilde{z}\|$. \square

In the context of Problem 3.3.39, the forward-backward splitting algorithm (3.73) assumes the following form, which can be described as an inexact, relaxed proximal Landweber method, as it alternates between an inexact Landweber step $x_n \mapsto x_n + \gamma_n(T^*(z - Tx_n) - b_n)$ and a relaxed inexact proximal step.

Theorem 3.3.43 (Proximal Landweber method) *Suppose that $G \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/\|T\|^2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left((L^* \circ \text{prox}_{\gamma_n f} \circ L)(x_n + \gamma_n(T^*(z - Tx_n) - b_n)) + a_n - x_n \right). \quad (3.115)$$

Then :

- (i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.
- (ii) $\sum_{n \in \mathbb{N}} \|T^*T(x_n - x)\|^2 < +\infty$.
- (iii) $\sum_{n \in \mathbb{N}} \|(L^* \circ \text{prox}_{\gamma_n f} \circ L)(x_n + \gamma_n T^*(z - Tx_n)) - x_n\|^2 < +\infty$.
- (iv) $(x_n)_{n \in \mathbb{N}}$ converges strongly to x if and only if $\underline{\lim} d_G(x_n) = 0$. In particular, strong convergence occurs in each of the following cases :
 - (a) $\text{int } G \neq \emptyset$.
 - (b) f satisfies Condition 3.3.23 on $L(G)$.
 - (c) T is bounded below.
 - (d) $\text{Id} - T^*T$ is demicompact.

Proof. Let f_1 , f_2 , and β be as in Proposition 3.3.40. Then, in view of Lemma 3.3.9, (3.73) reduces to (3.115) in the present setting. Thus, items (i)–(iii), as well as the main claim in item (iv) and item (iv)(a) are consequences of their counterparts in Theorem 3.3.25.

(iv)(b) : In view of Theorem 3.3.25(iv)(b), it suffices to show that $f \circ L$ satisfies Condition 3.3.23 on G . To this end, take sequences $(y_n)_{n \in \mathbb{N}}$ and $(v_n)_{n \in \mathbb{N}}$ in \mathcal{H} , and

points $y \in G$ and $v \in \partial(f \circ L)(y) = L^*(\partial f(Ly))$ such that $y_n \rightharpoonup y$, $v_n \rightarrow v$, and $(\forall n \in \mathbb{N}) v_n \in \partial(f \circ L)(y_n) = L^*(\partial f(Ly_n))$ (see Lemma 3.3.3(i)). Since L is linear and bounded, it is weakly and strongly continuous. Therefore, we have $Ly_n \rightharpoonup Ly \in L(G)$ and $Lv_n \rightarrow Lv \in \partial f(Ly)$. On the other hand, $(\forall n \in \mathbb{N}) Lv_n \in \partial f(Ly_n)$. Hence, since f satisfies Condition 3.3.23 on $L(G)$, there exists a subsequence $(y_{k_n})_{n \in \mathbb{N}}$ such that $Ly_{k_n} \rightarrow Ly$. It follows from assumption (iii) in Problem 3.3.39 that $y_{k_n} \rightarrow y$.

(iv)(c) : It follows from (3.111) that f_2 is strongly convex. Hence the claim follows from Proposition 3.3.27(vii) and Theorem 3.3.25(iv)(c). \square

(iv)(d) : In this case $\text{Id} - \nabla f_2$ is demicompact. Hence the claim follows from Proposition 3.3.27(ii) and Theorem 3.3.25(iv)(c). \square

3.3.5.3 Constrained least-squares problems

The least-squares problem associated with (3.105) is

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad \frac{1}{2} \|Tx - z\|^2. \quad (3.116)$$

A natural way to regularize this problem is to force the solutions to lie in a given closed convex set modeling *a priori* constraints [35, 48, 63]. This leads to the following formulation.

Problem 3.3.44 Let

- (i) $T: \mathcal{H} \rightarrow \mathcal{G}$ be a nonzero bounded linear operator ;
- (ii) C be a nonempty closed convex subset of \mathcal{H} .

The objective is to

$$\underset{x \in C}{\text{minimize}} \quad \frac{1}{2} \|Tx - z\|^2. \quad (3.117)$$

The set of solutions to this problem is denoted by G .

Proposition 3.3.45 *Problem 3.3.44 is a special case of Problem 3.3.39 with $\mathcal{K} = \mathcal{H}$, $L = \text{Id}$, and $f = \iota_C$.*

Proof. The proof is a direct consequence of (3.19). \square

Proposition 3.3.46

- (i) *Problem 3.3.44 possesses at least one solution if one of the following conditions is satisfied :*
 - (a) *C is bounded.*
 - (b) *$T(C)$ is closed.*

(ii) *Problem 3.3.44 possesses at most one solution if one of the following conditions is satisfied :*

(a) *Problem (3.116) has no solution in C , and C is strictly convex, i.e.,*

$$(\forall(x, y) \in C^2) \quad (x + y)/2 \in \text{int } C. \quad (3.118)$$

(b) *T is injective.*

(iii) *Problem 3.3.44 possesses exactly one solution if T is bounded below.*

(iv) *Let $x \in \mathcal{H}$ and $\gamma \in]0, +\infty[$. Then the following statements are equivalent :*

(a) *x solves Problem 3.3.44.*

(b) *$x = P_C(x + \gamma T^*(z - Tx))$.*

(c) *$x \in C$ and $(\forall y \in C) \quad \langle Ty - Tx \mid z - Tx \rangle \leq 0$.*

Proof. (i)(a) : This follows from Proposition 3.3.45 and Proposition 3.3.41(i) since ι_C is coercive.

(i)(b) : Since T is linear and C is convex, $T(C)$ is convex. Hence the assumptions imply that $T(C)$ is a nonempty closed convex subset of \mathcal{G} . As a result, z admits a projection p onto $T(C)$ and, therefore, there exists a point $x \in C$ such that $p = Tx$ and x solves (3.117).

(ii)(a) : By Fermat's rule (3.24), if (3.116) has no solution in C , then we have $(\forall x \in C) \quad 0 \notin \partial \|Tx - z\|^2/2$ and the result therefore follows from [47, Theorem 1.3].

Finally, items (ii)(b), (iii), and (iv) follow from Proposition 3.3.45 and their counterparts in Proposition 3.3.41, with the help of (3.31) in (iv)(b) and of (3.19) in (iv)(c). \square

Corollary 3.3.47 *Suppose that $G \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/\|T\|^2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(P_C(x_n + \gamma_n(T^*(z - Tx_n) - b_n)) + a_n - x_n \right). \quad (3.119)$$

Then :

(i) *$(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.*

(ii) *$(x_n)_{n \in \mathbb{N}}$ converges strongly to x if and only if $\underline{\lim} d_G(x_n) = 0$.*

Proof. Specialize Theorem 3.3.43(i)&(iv) to the setting described in Proposition 3.3.45 and use (3.19). \square

Remark 3.3.48 As in Theorem 3.3.43(iv), we obtain strong convergence in particular when $\text{int } G \neq \emptyset$, when T is bounded below, or when $\text{Id} - T^*T$ is demicompact. Another example is when C is boundedly compact, since in this case ι_C satisfies condition (i) in Proposition 3.3.27 and we can therefore conclude with Theorem 3.3.43(iv)(b).

Remark 3.3.49 (Projected Landweber iteration) Corollary 3.3.47 improves upon the results of [35, Section 3.1], which considered the special case when $\lambda \equiv 1$, $\gamma_n \equiv \gamma \in]0, 2/\|T\|^2[$, $a_n \equiv 0$, and $b_n \equiv 0$. In this particular scenario, (3.119) reduces to the classical projected Landweber iteration

$$x_{n+1} = P_C(x_n + \gamma T^*(z - Tx_n)), \quad \text{where } 0 < \gamma < 2/\|T\|^2, \quad (3.120)$$

item (i) can be found in [35, Theorem 3.2(v)], and item (ii) implies [35, Theorem 3.2(vi)] and, in turn, [35, Theorem 3.3].

Remark 3.3.50 (Disproving a conjecture) In [35, Section 3.1], it was conjectured that, for any C , \mathcal{G} , T , and z in Problem 3.3.44 such that $G \neq \emptyset$, any sequence $(x_n)_{n \in \mathbb{N}}$ generated by the projected Landweber iteration (3.120) converges strongly to a point in G . This conjecture is not true, as we now show. Take $\mathcal{G} = \mathbb{R}$, $z = 0$, and $T: x \mapsto \langle x | u \rangle$, where $u \in \mathcal{H} \setminus \{0\}$. Furthermore set $H = \ker T$ and $\gamma = 1/\|T\|^2$. Then (3.120) can be rewritten as

$$x_{n+1} = P_C \left(x_n - \frac{1}{\|T\|^2} T^* T x_n \right) = P_C \left(x_n - \frac{\langle x_n | u \rangle}{\|u\|^2} u \right) = (P_C \circ P_H) x_n. \quad (3.121)$$

However, it was shown in [41] that, for a particular choice of x_0 , u , and of a closed convex cone C , the sequence $(x_n)_{n \in \mathbb{N}}$ produced by this alternating projection iteration converges weakly but not strongly to a point in G .

3.3.5.4 Sparse regularization problems

In nonlinear approximation theory, statistics, and signal processing, a powerful idea is to decompose a function into an orthonormal basis and to transform the coefficients of the decomposition to construct sparse approximations or estimators, e.g., [18, 20, 30, 31, 33, 34, 49]. In the context of infinite-dimensional inverse problems, a variational formulation of this concept is the following (the specialization to the finite dimensional setting is straightforward).

Problem 3.3.51 Let

- (i) $T: \mathcal{H} \rightarrow \mathcal{G}$ be a nonzero bounded linear operator;
- (ii) $(e_k)_{k \in \mathbb{N}}$ be an orthonormal basis of \mathcal{H} ;
- (iii) $(\phi_k)_{k \in \mathbb{N}}$ be functions in $\Gamma_0(\mathbb{R})$ such that $(\forall k \in \mathbb{N}) \phi_k \geq 0$ and $\phi_k(0) = 0$.

The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad \frac{1}{2} \|Tx - z\|^2 + \sum_{k \in \mathbb{N}} \phi_k(\langle x | e_k \rangle). \quad (3.122)$$

The set of solutions to this problem is denoted by G .

Proposition 3.3.52 *Problem 3.3.51 is a special case of Problem 3.3.39 with $\mathcal{K} = \ell^2(\mathbb{N})$, $L: x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{N}}$, and $f: (\xi_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \phi_k(\xi_k)$.*

Proof. See proof of Example 3.3.20. \square

Proposition 3.3.53

- (i) *Problem 3.3.51 possesses at least one solution if there exists a function $c: [0, +\infty[\rightarrow [0, +\infty[$ such that $c(0) = 0$, $\lim_{t \rightarrow +\infty} c(t) = +\infty$, and*

$$(\forall (\xi_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})) \quad \sum_{k \in \mathbb{N}} \phi_k(\xi_k) \geq c\left(\sum_{k \in \mathbb{N}} |\xi_k|^2\right). \quad (3.123)$$

- (ii) *Problem 3.3.51 possesses at most one solution if one of the following conditions is satisfied :*

- (a) *The functions $(\phi_k)_{k \in \mathbb{N}}$ are strictly convex.*
(b) *T is injective.*

- (iii) *Problem 3.3.51 possesses exactly one solution if T is bounded below.*

- (iv) *Let $x \in \mathcal{H}$ and $\gamma \in]0, +\infty[$. Then the following statements are equivalent :*

- (a) *x solves Problem 3.3.51.*
(b) $(\forall k \in \mathbb{N}) \quad \langle x | e_k \rangle = \text{prox}_{\gamma \phi_k}(x + \gamma T^*(z - Tx) | e_k)$.
(c) $(\forall k \in \mathbb{N})(\forall \eta \in \mathbb{R}) \quad (\eta - \langle x | e_k \rangle) \langle z - Tx | Te_k \rangle + \phi_k(\langle x | e_k \rangle) \leq \phi_k(\eta)$.

Proof. In view of Proposition 3.3.52, we can invoke Proposition 3.3.41. Let f and L be as in Proposition 3.3.52.

(i) : By Proposition 3.3.41(i), it is enough to show that f is coercive. Let $\mathbf{x} = (\xi_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})$. Then it follows from (3.123) that $f(\mathbf{x}) = \sum_{k \in \mathbb{N}} \phi_k(\xi_k) \geq c(\sum_{k \in \mathbb{N}} |\xi_k|^2) = c(\|\mathbf{x}\|^2)$. Therefore, $\|\mathbf{x}\| \rightarrow +\infty \Rightarrow f(\mathbf{x}) \rightarrow +\infty$.

(ii)(a) : In view of Proposition 3.3.41(ii)(a), it is enough to show that f is strictly convex. Let $\mathbf{x} = (\xi_k)_{k \in \mathbb{N}}$ and $\mathbf{y} = (\eta_k)_{k \in \mathbb{N}}$ be two distinct points in $\text{dom } f$ (if $\text{dom } f$ is a singleton, the conclusion is clear) and let $\alpha \in]0, 1[$. Then there exists an index $l \in \mathbb{N}$ such that $\xi_l \neq \eta_l$, $\phi_l(\xi_l) < +\infty$, and $\phi_l(\eta_l) < +\infty$. Moreover, by strict convexity of

$\phi_l, \phi_l(\alpha\xi_l + (1 - \alpha)\eta_l) < \alpha\phi_l(\xi_l) + (1 - \alpha)\phi_l(\eta_l)$. Consequently, since the functions $(\phi_k)_{k \in \mathbb{N}}$ are convex,

$$\begin{aligned} f(\alpha x + (1 - \alpha)y) &= \sum_{k \in \mathbb{N}} \phi_k(\alpha\xi_k + (1 - \alpha)\eta_k) \\ &< \sum_{k \in \mathbb{N}} \alpha\phi_k(\xi_k) + (1 - \alpha)\phi_k(\eta_k) \\ &= \alpha f(x) + (1 - \alpha)f(y), \end{aligned} \tag{3.124}$$

which proves the strict convexity of f .

Finally, items (ii)(b), (iii), and (iv) follow from their counterpart in Proposition 3.3.41, with the help of Example 3.3.20 in (iv). \square

We now turn our attention to the numerical solution of Problem 3.3.51.

Corollary 3.3.54 *Suppose that $G \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/\|T\|^2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1[$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(b_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Moreover, for every $n \in \mathbb{N}$, let $(\alpha_{n,k})_{k \in \mathbb{N}}$ be a sequence in $\ell^2(\mathbb{N})$ and suppose that $\sum_{n \in \mathbb{N}} \sqrt{\sum_{k \in \mathbb{N}} |\alpha_{n,k}|^2} < +\infty$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(\sum_{k \in \mathbb{N}} (\alpha_{n,k} + \text{prox}_{\gamma_n \phi_k} \langle x_n + \gamma_n(T^*(z - Tx_n) - b_n) \mid e_k \rangle) e_k - x_n \right). \tag{3.125}$$

Then :

- (i) $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in G$.
- (ii) $\sum_{n \in \mathbb{N}} \|T^*T(x_n - x)\|^2 < +\infty$.
- (iii) $\sum_{n \in \mathbb{N}} \left\| \text{prox}_{\gamma_n f_1} (x_n + \gamma_n T^*(z - Tx_n)) - x_n \right\|^2 < +\infty$, where $f_1: y \mapsto \sum_{k \in \mathbb{N}} \phi_k(\langle y \mid e_k \rangle)$.
- (iv) $(x_n)_{n \in \mathbb{N}}$ converges strongly to x if and only if $\underline{\lim} d_G(x_n) = 0$.

Proof. It follows from Example 3.3.20 that (3.125) is a special case of (3.115) with $(\forall n \in \mathbb{N}) a_n = \sum_{k \in \mathbb{N}} \alpha_{n,k} e_k$. In view of Proposition 3.3.52, the corollary is therefore an application of Theorem 3.3.43. \square

Specific strong convergence conditions are given in Theorem 3.3.43(iv). Let us now provide two illustrations of the above results.

Example 3.3.55 Suppose that T is bounded below. Then (without further assumptions on the sequence $(\phi_k)_{k \in \mathbb{N}}$), Problem 3.3.51 has a unique solution x (Proposition 3.3.53(iii)) and we obtain the strong convergence of any sequence generated by (3.125) to x (see Theorem 3.3.43(iv)(c)). Moreover, as the data z vary, the solutions are stable in the sense of (3.113).

Problem 3.3.56 We revisit a problem investigated in [30] with different tools (see also [10, 31, 36, 64, 65] for related frameworks and special cases). Let

- (i) $T: \mathcal{H} \rightarrow \mathcal{G}$ be a nonzero bounded linear operator;
- (ii) $(e_k)_{k \in \mathbb{N}}$ be an orthonormal basis of \mathcal{H} ;
- (iii) $p \in [1, 2]$ and $(\omega_k)_{k \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $\omega = \inf_{k \in \mathbb{N}} \omega_k > 0$.

The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad \frac{1}{2} \|Tx - z\|^2 + \sum_{k \in \mathbb{N}} \omega_k |\langle x | e_k \rangle|^p. \quad (3.126)$$

Clearly, Problem 3.3.56 is a special case of Problem 3.3.51 with $(\forall k \in \mathbb{N}) \phi_k: \xi \mapsto \omega_k |\xi|^p$. Moreover, since $p \in [1, 2]$, we have $(\forall (\xi_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})) (\sum_{k \in \mathbb{N}} |\xi_k|^p)^{1/p} \geq (\sum_{k \in \mathbb{N}} |\xi_k|^2)^{1/2}$. Accordingly,

$$\begin{aligned} (\forall (\xi_k)_{k \in \mathbb{N}} \in \ell^2(\mathbb{N})) \quad \sum_{k \in \mathbb{N}} \phi_k(\xi_k) &= \sum_{k \in \mathbb{N}} \omega_k |\xi_k|^p \\ &\geq \omega \sum_{k \in \mathbb{N}} |\xi_k|^p \\ &\geq \omega \left(\sum_{k \in \mathbb{N}} |\xi_k|^2 \right)^{p/2}. \end{aligned} \quad (3.127)$$

Therefore (3.123) holds with $c: t \mapsto \omega t^{p/2}$. Hence, as a result of Proposition 3.3.53(i), Problem 3.3.56 admits at least one solution. Moreover, we deduce from Proposition 3.3.53(ii)(a) that it admits exactly one solution if $1 < p \leq 2$. Now, let the sequences $(\gamma_n)_{n \in \mathbb{N}}$, $(\lambda_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$, and $(\alpha_{n,k})_{n \in \mathbb{N}}$ be as in Corollary 3.3.54 and define, for every $(k, n) \in \mathbb{N}^2$,

$$\pi_{n,k} = \text{prox}_{\omega_{n,k} |\cdot|^p} \langle x_n + \gamma_n (T^*(z - Tx_n) - b_n) | e_k \rangle, \quad \text{where} \quad \omega_{n,k} = \gamma_n \omega_k. \quad (3.128)$$

Then we can rewrite (3.125) as

$$x_{n+1} = x_n + \lambda_n \left(\sum_{k \in \mathbb{N}} (\alpha_{n,k} + \pi_{n,k}) e_k - x_n \right). \quad (3.129)$$

We deduce at once from Corollary 3.3.54(i) that

$$(x_n)_{n \in \mathbb{N}} \text{ converges weakly to a solution } x \text{ to Problem 3.3.56.} \quad (3.130)$$

In [30], (3.130) was obtained with the additional assumptions $\lambda_n \equiv 1$, $\|T\| < 1$, $\gamma_n \equiv 1$, $\alpha_{n,k} \equiv 0$, and $b_n \equiv 0$ (see [30, Proposition 3.11]). Furthermore, it was shown that, in this particular case, strong convergence is achieved [30, Theorem 3.1]. Let us now extend this result.

Corollary 3.3.57 *Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/\|T\|^2$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(b_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{H} such that $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Moreover, for every $n \in \mathbb{N}$, let $(\alpha_{n,k})_{k \in \mathbb{N}}$ be a sequence in $\ell^2(\mathbb{N})$ and suppose that $\sum_{n \in \mathbb{N}} \sqrt{\sum_{k \in \mathbb{N}} |\alpha_{n,k}|^2} < +\infty$. Fix $x_0 \in \mathcal{H}$ and let $(x_n)_{n \in \mathbb{N}}$ be a sequence generated by (3.128)–(3.129). Then $(x_n)_{n \in \mathbb{N}}$ converges strongly to a solution x to Problem 3.3.56.*

Proof. As seen in (3.130), $x_n \rightharpoonup x$, where x solves Problem 3.3.56. Now set $f_1: y \mapsto \sum_{k \in \mathbb{N}} \omega_k |\langle y | e_k \rangle|^p$, $f_2: y \mapsto \|Ty - z\|^2/2$, and, for every $n \in \mathbb{N}$, set $h_n = x + \gamma_n T^*(z - Tx) = x - \gamma_n \nabla f_2(x)$ and $v_n = x_n - x$. Then we must show that $v_n \rightarrow 0$. Proposition 3.3.22(iii) yields $(\forall n \in \mathbb{N}) \operatorname{prox}_{\gamma_n f_1} h_n = x$. Hence, it follows from Lemma 3.3.5 that

$$\begin{aligned} & (\forall n \in \mathbb{N}) \quad \|\operatorname{prox}_{\gamma_n f_1}(v_n + h_n) - v_n - \operatorname{prox}_{\gamma_n f_1} h_n\| \\ &= \|\operatorname{prox}_{\gamma_n f_1}(x_n + \gamma_n T^*(z - Tx)) - x_n\| \\ &\leq \|\operatorname{prox}_{\gamma_n f_1}(x_n + \gamma_n T^*(z - Tx)) - \operatorname{prox}_{\gamma_n f_1}(x_n + \gamma_n T^*(z - Tx_n))\| \\ &\quad + \|\operatorname{prox}_{\gamma_n f_1}(x_n + \gamma_n T^*(z - Tx_n)) - x_n\| \\ &\leq \gamma_n \|T^*T(x_n - x)\| + \|\operatorname{prox}_{\gamma_n f_1}(x_n + \gamma_n T^*(z - Tx_n)) - x_n\|. \end{aligned} \quad (3.131)$$

Therefore, the boundedness of $(\gamma_n)_{n \in \mathbb{N}}$ and Corollary 3.3.54(ii)&(iii) yield

$$\|v_n - \operatorname{prox}_{\gamma_n f_1}(v_n + h_n) + \operatorname{prox}_{\gamma_n f_1} h_n\| \rightarrow 0. \quad (3.132)$$

On the other hand, (3.130) states that

$$v_n \rightarrow 0. \quad (3.133)$$

The remainder of the proof is patterned after that of [30, Lemma 3.18]. There, it was shown that, if $\|T\| < 1$ and $\gamma_n \equiv 1$ (hence $h_n \equiv x + T^*(z - Tx)$), then (3.132) and (3.133) imply that $v_n \rightarrow 0$. We shall show that this conclusion remains true in our more general setting. Define, for every n and k in \mathbb{N} , $\eta_{n,k} = \langle h_n | e_k \rangle$, $\nu_{n,k} = \langle v_n | e_k \rangle$, $\xi_k = \langle x | e_k \rangle$, $\chi_k = \langle q | e_k \rangle$, and $\rho_k = \langle r | e_k \rangle$, where $q = T^*(z - Tx)$ and $r = 2q/\|T\|^2$. Since, for every $n \in \mathbb{N}$, $h_n = x + \gamma_n q$ and $\gamma_n < 2/\|T\|^2$, we have

$$(\forall k \in \mathbb{N})(\forall n \in \mathbb{N}) \quad |\eta_{n,k}|^2/2 \leq |\xi_k|^2 + \gamma_n^2 |\chi_k|^2 \leq |\xi_k|^2 + |\rho_k|^2. \quad (3.134)$$

Now let $\delta = \inf_{(n,k) \in \mathbb{N}^2} \omega_{n,k}$. Note that (3.128) and our assumptions yield

$$\delta > 0. \quad (3.135)$$

As in [30, Lemma 3.18], we treat the cases $1 < p \leq 2$ and $p = 1$ separately.

First, suppose that $1 < p \leq 2$. We derive from (3.133) that $\sup_{n \in \mathbb{N}} \|v_n\| \leq B$ for some $B \in]0, +\infty[$. Now define $\mathbb{K} = \{k \in \mathbb{N} \mid (\exists n \in \mathbb{N}) |\eta_{n,k}| \geq B\}$. Then we derive from (3.134) that

$$(\forall k \in \mathbb{K})(\exists n \in \mathbb{N}) \quad |\xi_k|^2 + |\rho_k|^2 \geq |\eta_{n,k}|^2/2 \geq B^2/2. \quad (3.136)$$

Consequently, since x and r lie in \mathcal{H} , we have

$$+\infty > \sum_{k \in \mathbb{N}} (|\xi_k|^2 + |\rho_k|^2) \geq \sum_{k \in \mathbb{K}} (|\xi_k|^2 + |\rho_k|^2) \geq (\text{card } \mathbb{K})B^2/2, \quad (3.137)$$

and \mathbb{K} is therefore a finite set. Consequently, it results from (3.133) that $\sum_{k \in \mathbb{K}} |\nu_{n,k}|^2 \rightarrow 0$. To show that $\|v_n\|^2 \rightarrow 0$, it remains to show that $\sum_{k \in \mathbb{N} \setminus \mathbb{K}} |\nu_{n,k}|^2 \rightarrow 0$. The definition of \mathbb{K} yields

$$(\forall k \in \mathbb{N} \setminus \mathbb{K})(\forall n \in \mathbb{N}) \quad |\eta_{n,k}| < B. \quad (3.138)$$

Using (3.135), (3.138) and proceeding as in [30, Lemma 3.18], we obtain a constant $\mu \in]0, +\infty[$ depending only on p , δ , and B such that

$$\begin{aligned} \sum_{k \in \mathbb{N} \setminus \mathbb{K}} |\nu_{n,k}|^2 &\leq \mu \sum_{k \in \mathbb{N} \setminus \mathbb{K}} |\nu_{n,k} - \text{prox}_{\omega_{n,k}|\cdot|^p}(\nu_{n,k} + \eta_{n,k}) + \text{prox}_{\omega_{n,k}|\cdot|^p} \eta_{n,k}|^2 \\ &\leq \mu \|v_n - \text{prox}_{\gamma_n f_1}(v_n + h_n) + \text{prox}_{\gamma_n f_1} h_n\|^2. \end{aligned} \quad (3.139)$$

Hence it follows from (3.132) that $\sum_{k \in \mathbb{N} \setminus \mathbb{K}} |\nu_{n,k}|^2 \rightarrow 0$, as desired.

Finally, suppose that $p = 1$. Since x and r lie in \mathcal{H} , (3.134) yields

$$\sum_{k \in \mathbb{N}} \sup_{n \in \mathbb{N}} |\eta_{n,k}|^2 / 2 \leq \sum_{k \in \mathbb{N}} (|\xi_k|^2 + |\rho_k|^2) < +\infty. \quad (3.140)$$

Hence, by (3.135), there exists an integer $K \in \mathbb{N}$ such that $(\forall n \in \mathbb{N}) \sum_{k > K} |\eta_{n,k}|^2 \leq (\delta/2)^2$, and it follows from (3.133) that $\sum_{k=0}^K |\nu_{n,k}|^2 \rightarrow 0$. It now remains to show that $\sum_{k > K} |\nu_{n,k}|^2 \rightarrow 0$. Invoking the same arguments as in [30, Lemma 3.18], this follows from (3.52) and (3.132). \square

Remark 3.3.58 Let us make a few of comments about Corollary 3.3.57.

- (i) In [30, Remark 3.14], the assumption $p \leq 2$ in Problem 3.3.56 was made to ensure that the sequence $(x_n)_{n \in \mathbb{N}}$ is bounded, whereas here it is made to ensure that Problem 3.3.56 has solutions. These two conditions are perfectly consistent. Indeed, the algorithm of [30] iterates $x_{n+1} = R x_n = R^{n+1} x_0$, where $R = \text{prox}_{f_1} \circ (\text{Id} + T^*(z - T))$ is nonexpansive (actually averaged nonexpansive [25]) and its fixed point set $\text{Fix } R$ is the set of solutions to Problem 3.3.56 (Proposition 3.3.22(iii) or Proposition 3.3.53(iv)). Hence, $(R^n x_0)_{n \in \mathbb{N}}$ is bounded if and only if $\text{Fix } R \neq \emptyset$; actually, $\text{Fix } R = \emptyset \Rightarrow \|R^n x_0\| \rightarrow +\infty$ [11, Corollary 9(b)].
- (ii) Let $f: (\xi_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \omega_k |\xi_k|^p$. Then, since $\inf_{k \in \mathbb{N}} \omega_k > 0$, f is strongly convex on $\ell^p(\mathbb{N})$ for $p = 2$, and strong convergence can be deduced directly from Theorem 3.3.43(iv)(b) and Proposition 3.3.27(vii). However, for $1 < p < 2$, we cannot conclude via Theorem 3.3.43(iv)(b) and Proposition 3.3.27(vi) since, even for $(\omega_k)_{k \in \mathbb{N}}$ constant, f is known to be uniformly convex on bounded sets only in $\ell^p(\mathbb{N})$ [75].

- (iii) For $p = 1$, it follows from Corollary 3.3.57 and Example 3.3.21 that (3.128)–(3.129) is a strongly convergent iterative soft-thresholding method. This result extends the theoretical foundations of the multiresolution schemes proposed in [10, 31, 36, 64, 65].

3.3.6 Denoising problems

3.3.6.1 Problem statement and basic properties

In denoising problems, $\mathcal{G} = \mathcal{H}$ and $T = \text{Id}$ in (3.105), which leads to the data formation equation

$$z = \bar{x} + w. \quad (3.141)$$

In other words, z is a noisy observation of \bar{x} . We derive from Problem 3.3.39 the following formulation.

Problem 3.3.59 Let

- (i) \mathcal{K} be a real Hilbert space;
- (ii) $L: \mathcal{H} \rightarrow \mathcal{K}$ be a bijective bounded linear operator such that $L^{-1} = L^*$;
- (iii) $f \in \Gamma_0(\mathcal{K})$.

The objective is to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f(Lx) + \frac{1}{2} \|x - z\|^2. \quad (3.142)$$

Proposition 3.3.60 *Problem 3.3.59 possesses exactly one solution z^\oplus , which is characterized by one of the following equivalent conditions :*

- (i) $z^\oplus = \text{prox}_{f \circ L} z = (L^* \circ \text{prox}_f \circ L)z$.
- (ii) $(\forall y \in \mathcal{H}) \quad \langle y - z^\oplus | z - z^\oplus \rangle + f(Lz^\oplus) \leq f(Ly)$.

Proof. Since $T = \text{Id}$, the existence and uniqueness of a solution follow from Proposition 3.3.41(iii). The characterizations are obtained by applying Proposition 3.3.41(iv) with $T = \text{Id}$ and $\gamma = 1$. \square

Example 3.3.61 (Wavelet soft-thresholding) Suppose that, in Problem 3.3.59, $(e_k)_{k \in \mathbb{N}}$ is an orthonormal wavelet basis of \mathcal{H} , $\mathcal{K} = \ell^2(\mathbb{N})$, $L: x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{N}}$, $(\omega_k)_{k \in \mathbb{N}}$ is a sequence in $]0, +\infty[$, and $f: (\xi_k)_{k \in \mathbb{N}} \mapsto \sum_{k \in \mathbb{N}} \omega_k |\xi_k|$. Then, it follows from Example 3.3.21 that z^\oplus is the wavelet soft-thresholded transformation of z [33, 34].

Proposition 3.3.60 states that Problem 3.3.59 admits a unique solution $z^\oplus = \text{prox}_{f \circ L} z$. According to Moreau's decomposition principle (3.42), the signal z can be decomposed as

$$z = z^\oplus + z^\ominus, \quad (3.143)$$

where $z^\ominus = \text{prox}_{(f \circ L)^*} z = \text{prox}_{f^* \circ L} z$ is by definition the solution to the dual problem

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad f^*(Lx) + \frac{1}{2} \|x - z\|^2. \quad (3.144)$$

Moreover, $f(Lz^\oplus) + f^*(Lz^\ominus) = \langle z^\oplus | z^\ominus \rangle$. Schematically, the action of the dual filters $\text{prox}_{f \circ L}$ and $\text{prox}_{f^* \circ L}$ can be represented as in Figure 1.

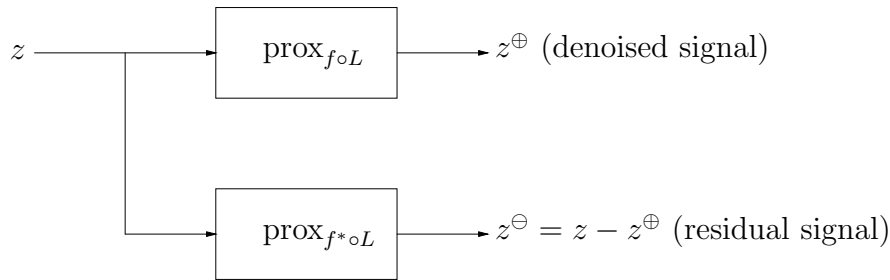


Fig. 1 : Proximal signal denoising.

Moreau's decomposition principle tells us that the component of the signal that is filtered out, namely z^\ominus , is actually dually related to the denoised component z^\oplus since it is obtained by applying the same type of proximal operation to z , except that the function f is now replaced by its conjugate f^* . In practice, deeper insights into the properties of the denoising procedure can be gained from the availability of the two components z^\oplus and z^\ominus in the decomposition of z . This is particularly important in standard linear hierarchical signal analysis [49], as well as in certain nonlinear extensions thereof [24, 28].

As a classical illustration of this denoising decomposition, consider the case when $\mathcal{H} = L^2(\mathbb{R})$, $\mathcal{K} = \mathcal{H} \times \mathcal{H}$, and the Fourier transform \widehat{x} of the original signal in (3.141) lies mostly in some low frequency band B , whereas the Fourier transform \widehat{w} of the noise lies mostly in a higher frequency range. Then it is natural to obtain the denoised signal z^\oplus by low-pass filtering z [58]. Now let $L: \mathcal{H} \rightarrow \mathcal{K}$ be the Fourier transform operator, let V be the closed vector subspace of \mathcal{K} of signals with support B , and set $f = \iota_V$. Then, as is well-known [23, 66], the above low-pass filtering operation can be written as $z^\oplus = P_V(Lz)$ which, in light of (3.31), we can rewrite as $z^\oplus = \text{prox}_{f \circ L} z$. Since $f^* = \iota_V^* = \iota_{V^\perp}$, the signal that is filtered out is indeed the high-pass component $z^\ominus = P_{V^\perp}(Lz) = \text{prox}_{f^* \circ L} z$.

As a second example, take $\mathcal{H} = L^2(\Omega)$, where Ω is a bounded open domain in \mathbb{R}^2 , and let f be the total variation, i.e.,

$$f: x \mapsto \begin{cases} \gamma \int_{\Omega} |\nabla x(\omega)| d\omega, & \text{if } x \in H^1(\Omega); \\ +\infty, & \text{otherwise,} \end{cases} \quad (3.145)$$

where $\gamma \in]0, +\infty[$. Then the denoising problem (3.142) was initially proposed in [62]. In this case, the proximal decomposition (3.143) appears implicitly in [51], where z^{\oplus} was described as a bounded variation component of the image z carrying most of its structure, while z^{\ominus} was described as a texture/noise component. One will find in [51] a detailed and insightful analysis of this decomposition model.

We conclude this section with a stability result. Let $\tilde{z} \in \mathcal{H}$ be a perturbation of the data z and, as above, let \tilde{z}^{\oplus} and z^{\oplus} be the associated solutions produced by (3.142). Then it follows at once from Proposition 3.3.42 with $T = \text{Id}$ that $\|z^{\oplus} - \tilde{z}^{\oplus}\| \leq \|z - \tilde{z}\|$. Here is a sharpening of this result.

Proposition 3.3.62 *Let $\tilde{z} \in \mathcal{H}$. Then $\|z^{\oplus} - \tilde{z}^{\oplus}\|^2 + \|z^{\ominus} - \tilde{z}^{\ominus}\|^2 \leq \|z - \tilde{z}\|^2$, where we have used the same notation as in (3.143).*

Proof. Using Lemma 3.3.5 and (3.143), we obtain

$$\begin{aligned} \|z^{\oplus} - \tilde{z}^{\oplus}\|^2 &= \|\text{prox}_{f \circ L} z - \text{prox}_{f \circ L} \tilde{z}\|^2 \\ &\leq \|z - \tilde{z}\|^2 - \|(\text{Id} - \text{prox}_{f \circ L})z - (\text{Id} - \text{prox}_{f \circ L})\tilde{z}\|^2 \\ &= \|z - \tilde{z}\|^2 - \|(z - z^{\oplus}) - (\tilde{z} - \tilde{z}^{\oplus})\|^2 \\ &= \|z - \tilde{z}\|^2 - \|z^{\ominus} - \tilde{z}^{\ominus}\|^2. \end{aligned} \quad (3.146)$$

□

3.3.6.2 A split denoising problem

As seen in Proposition 3.3.60, the solution to Problem 3.3.59 is $z^{\oplus} = L^*(\text{prox}_f(Lx))$. In Section 3.3.2.6 we have provided examples of proximity operators that could be computed in closed form. In some problems, however, it may be more difficult to evaluate prox_f directly and some iterative procedure may be required. We address this question in the case when f can be split into the sum of two functions, say

$$f = \varphi + \psi, \quad (3.147)$$

where

- (i) $\varphi \in \Gamma_0(\mathcal{K})$ and the operator prox_{φ} is relatively easy to implement;

- (ii) $\psi \in \Gamma_0(\mathcal{K})$ is differentiable and $\nabla\psi$ is α -Lipschitz continuous on \mathcal{K} for some $\alpha \in]0, +\infty[$.

Problem 3.3.59 then becomes

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad \varphi(Lx) + \psi(Lx) + \frac{1}{2}\|x - z\|^2. \quad (3.148)$$

We now provide a strongly convergent algorithm to construct the solution z^\oplus to (3.148).

Theorem 3.3.63 *Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/(\alpha + 1)$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1[$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{K} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{K}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(\text{prox}_{\gamma_n \varphi} \left((1 - \gamma_n)x_n + \gamma_n(Lz - \nabla\psi(x_n) - b_n) \right) + a_n - x_n \right). \quad (3.149)$$

Then $(x_n)_{n \in \mathbb{N}}$ converges strongly to $x = \text{prox}_f(Lz)$ and $z^\oplus = L^*x$.

Proof. In view of assumption (ii) in Problem 3.3.59, (3.148) is equivalent to

$$\underset{x \in \mathcal{K}}{\text{minimize}} \quad \varphi(x) + \psi(x) + \frac{1}{2}\|x - Lz\|^2. \quad (3.150)$$

Now set $f_1 = \varphi$ and $f_2 = \psi + \|\cdot - Lz\|^2/2$. Then ∇f_2 is Lipschitz continuous with constant $1/\beta = \alpha + 1$ and (3.150) is a special case of Problem 3.3.1 transposed in \mathcal{K} . Moreover, (3.149) is a special case of (3.73). We also observe that, since $\|\cdot - Lz\|^2/2$ is strongly convex, f_2 is likewise. It therefore follows from Proposition 3.3.27(vii) that f_2 satisfies Condition 3.3.23. Hence, we derive from Theorem 3.3.25(iv)(c) that $(x_n)_{n \in \mathbb{N}}$ converges strongly to the solution $x = Lz^\oplus$ to (3.150). \square

Remark 3.3.64 The continuity of L^* yields $L^*x_n \rightarrow z^\oplus$ in Theorem 3.3.63.

A noteworthy special case of (3.148) is when $\varphi = \iota_C$, for some nonempty closed convex set $C \subset \mathcal{K}$. In this case, we seek the optimal solution to the denoising problem relative to $\psi \circ L$ over the feasibility set $L^{-1}(C)$, i.e.,

$$\underset{Lx \in C}{\text{minimize}} \quad \psi(Lx) + \frac{1}{2}\|x - z\|^2. \quad (3.151)$$

This formulation makes it possible to incorporate more *a priori* information in terms of constraints on Lz^\oplus . As a direct corollary to Theorem 3.3.63 we obtain the following corollary.

Corollary 3.3.65 *Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/(\alpha + 1)$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{K} such that $\sum_{n \in \mathbb{N}} \|a_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$. Fix $x_0 \in \mathcal{K}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(P_C \left((1 - \gamma_n)x_n + \gamma_n(Lz - \nabla\psi(x_n) - b_n) \right) + a_n - x_n \right). \quad (3.152)$$

*Then $(x_n)_{n \in \mathbb{N}}$ converges strongly to $x = \text{prox}_f(Lz)$ and $z^\oplus = L^*x$.*

3.4 Bibliographie

- [1] F. ACKER AND M. A. PRESTEL, *Convergence d'un schéma de minimisation alternée*, Ann. Fac. Sci. Toulouse V. Sér. Math., 2 (1980), pp. 1–9.
- [2] H. C. ANDREWS AND B. R. HUNT, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
- [3] G. AUBERT AND L. VESE, *A variational method in image recovery*, SIAM J. Numer. Anal., 34 (1997), pp. 1948–1979.
- [4] J.-P. AUBIN AND H. FRANKOWSKA, *Set-Valued Analysis*, Birkhäuser, Boston, MA, 1990.
- [5] J.-F. AUJOL, G. AUBERT, L. BLANC-FÉRAUD, AND A. CHAMBOLLE, *Image decomposition into a bounded variation component and an oscillating component*, J. Math. Imaging Vision, 22 (2005), pp. 71–88.
- [6] J.-F. AUJOL AND A. CHAMBOLLE, *Dual norms and image decomposition models*, Int. J. Comput. Vision, 63 (2005), pp. 85–104.
- [7] J.-F. AUJOL, G. GILBOA, T. CHAN, AND S. OSHER, *Structure-texture image decomposition - Modeling, algorithms, and parameter selection*, Int. J. Comput. Vision, 67 (2006), pp. 111–136.
- [8] J.-B. BAILLON AND G. HADDAD, *Quelques propriétés des opérateurs angle-bornés et n -cycliquement monotones*, Israel J. Math., 26 (1977), pp. 137–150.
- [9] H. H. BAUSCHKE, P. L. COMBETTES, AND S. REICH, *The asymptotic behavior of the composition of two resolvents*, Nonlinear Anal., 60 (2005), pp. 283–301.
- [10] J. BECT, L. BLANC-FÉRAUD, G. AUBERT, AND A. CHAMBOLLE, *A ℓ^1 unified variational framework for image restoration*, in Proc. Eighth Europ. Conf. Comput. Vision, Prague, 2004, T. Pajdla and J. Matas, eds., Lecture Notes in Comput. Sci. 3024, Springer-Verlag, New York, 2004, pp. 1–13.
- [11] J. BORWEIN, S. REICH, AND I. SHAFRIR, *Krasnoselski-Mann iterations in normed spaces*, Canad. Math. Bull., 35 (1992), pp. 21–28.

- [12] D. BUTNARIU, A. N. IUSEM, AND C. ZĂLINESCU, *On uniform convexity, total convexity and convergence of the proximal point and outer Bregman projection algorithms in Banach spaces*, J. Convex Anal., 10 (2003), pp. 35–61.
- [13] C. L. BYRNE, *Iterative oblique projection onto convex sets and the split feasibility problem*, Inverse Problems, 18 (2002), pp. 441–453.
- [14] J. A. CADZOW AND Y. SUN, *Sequences with positive semidefinite Fourier transforms*, IEEE Trans. Acoust. Speech Signal Process., 34 (1986), pp. 1502–1510.
- [15] Y. CENSOR AND T. ELFVING, *A multiprojection algorithm using Bregman projections in a product space*, Numer. Algorithms, 8 (1994), pp. 221–239.
- [16] Y. CENSOR AND S. A. ZENIOS, *Parallel Optimization : Theory, Algorithms and Applications*, Oxford University Press, New York, 1997.
- [17] A. CHAMBOLLE, *An algorithm for total variation minimization and applications*, J. Math. Imaging Vision, 20 (2004), pp. 89–97.
- [18] A. CHAMBOLLE, R. A. DEVORE, N. Y. LEE, AND B. J. LUCIER, *Nonlinear wavelet image processing : Variational problems, compression, and noise removal through wavelet shrinkage*, IEEE Trans. Image Process., 7 (1998), pp. 319–335.
- [19] A. CHAMBOLLE AND P. L. LIONS, *Image recovery via total variation minimization and related problems*, Numer. Math., 76 (1997), pp. 167–188.
- [20] A. COHEN, *Numerical Analysis of Wavelet Methods*, Elsevier, New York, 2003.
- [21] P. L. COMBETTES, *The foundations of set theoretic estimation*, Proc. IEEE, 81 (1993), pp. 182–208.
- [22] P. L. COMBETTES, *Inconsistent signal feasibility problems : Least-squares solutions in a product space*, IEEE Trans. Signal Process., 42 (1994), pp. 2955–2966.
- [23] P. L. COMBETTES, *The convex feasibility problem in image recovery*, in Advances in Imaging and Electron Physics, P. Hawkes, ed., Vol. 95, Academic Press, New York, 1996, pp. 155–270.
- [24] P. L. COMBETTES, *Convexité et signal*, in Actes du Congrès de Mathématiques Appliquées et Industrielles SMAI'01, Pompadour, France, May 28–June 1, 2001, pp. 6–16.
- [25] P. L. COMBETTES, *Solving monotone inclusions via compositions of nonexpansive averaged operators*, Optimization, 53 (2004), pp. 475–504.
- [26] P. L. COMBETTES AND P. BONDON, *Hard-constrained inconsistent signal feasibility problems*, IEEE Trans. Signal Process., 47 (1999), pp. 2460–2468.
- [27] P. L. COMBETTES AND S. A. HIRSTOAGA, *Equilibrium programming in Hilbert spaces*, J. Nonlinear Convex Anal., 6 (2005), pp. 117–136.
- [28] P. L. COMBETTES AND J.-C. PESQUET, *Convex multiresolution analysis*, IEEE Trans. Pattern Anal. Machine Intell., 20 (1998), pp. 1308–1318.
- [29] P. L. COMBETTES AND J.-C. PESQUET, *Wavelet-constrained image restoration*, Int. J. Wavelets Multiresolut. Inf. Process., 2 (2004), pp. 371–389.

- [30] I. DAUBECHIES, M. DEFRISE, AND C. DE MOL, *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Comm. Pure Appl. Math., 57 (2004), pp. 1413–1457.
- [31] C. DE MOL AND M. DEFRISE, *A note on wavelet-based inversion algorithms*, Contemp. Math., 313 (2002), pp. 85–96.
- [32] D. DOBSON AND O. SCHERZER, *Analysis of regularized total variation penalty methods for denoising*, Inverse Problems, 12 (1996), pp. 601–617.
- [33] D. DONOHO AND I. JOHNSTONE, *Ideal spatial adaptation via wavelet shrinkage*, Biometrika, 81 (1994), pp. 425–455.
- [34] D. L. DONOHO, I. M. JOHNSTONE, G. KERKYACHARIAN, AND D. PICARD, *Wavelet shrinkage : Asymptopia ?*, J. R. Statist. Soc. B., 57 (1995), pp. 301–369.
- [35] B. EICKE, *Iteration methods for convexly constrained ill-posed problems in Hilbert space*, Numer. Funct. Anal. Optim., 13 (1992), pp. 413–429.
- [36] M. A. T. FIGUEIREDO AND R. D. NOWAK, *An EM algorithm for wavelet-based image restoration*, IEEE Trans. Image Process., 12 (2003), pp. 906–916.
- [37] R. W. GERCHBERG, *Super-resolution through error energy reduction*, Optica Acta, 21 (1974), pp. 709–720.
- [38] M. GOLDBURG AND R. J. MARKS II, *Signal synthesis in the presence of an inconsistent set of constraints*, IEEE Trans. Circuits and Systems, 32 (1985), pp. 647–663.
- [39] G. T. HERMAN, *Image Reconstruction from Projections, the Fundamentals of Computerized Tomography*, Academic Press, New York, 1980.
- [40] U. HERMANN AND D. NOLL, *Adaptive image reconstruction using information measures*, SIAM J. Control Optim., 38 (2000), pp. 1223–1240.
- [41] H. S. HUNDAL, *An alternating projection that does not converge in norm*, Non-linear Anal., 57 (2004), pp. 35–61.
- [42] B. R. HUNT, *The inverse problem of radiography*, Math. Biosciences, 8 (1970), pp. 161–179.
- [43] B. R. HUNT, *The application of constrained least-squares estimation to image restoration by digital computer*, IEEE Trans. Computers, 22 (1973), pp. 805–812.
- [44] N. HURT, *Phase Retrieval and Zero Crossings : Mathematical Methods in Image Reconstruction*, Kluwer, Boston, 1989.
- [45] S. L. KEELING, *Total variation based convex filters for medical imaging*, Appl. Math. Comput., 139 (2003), pp. 101–119.
- [46] A. LANNES, S. ROQUES, AND M. J. CASANOVE, *Stabilized reconstruction in signal and image processing. Part 1 : Partial deconvolution and spectral extrapolation with limited field*, J. Modern Opt., 34 (1987), pp. 161–226.
- [47] E. S. LEVITIN AND B. T. POLYAK, *Constrained minimization methods*, USSR Comput. Math. Math. Phys., 6 (1966), pp. 1–50.

- [48] A. J. LEVY, *A fast quadratic programming algorithm for positive signal restoration*, IEEE Trans. Acoust. Speech Signal Process., 31 (1983), pp. 1337–1341.
- [49] S. G. MALLAT, *A Wavelet Tour of Signal Processing*, 2nd ed, Academic Press, New York, 1999.
- [50] P. MARÉCHAL, D. TOGANE, AND A. CELLER, *A new reconstruction methodology for computerized tomography : FRECT (Fourier regularized computed tomography)*, IEEE Trans. Nucl. Sc., 47 (2000), pp. 1595-1601.
- [51] Y. MEYER, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS, Providence, RI, 2001.
- [52] J.-J. MOREAU, *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires*, C. R. Acad. Sci. Paris Sér. A Math., 255 (1962), pp. 238–240.
- [53] J.-J. MOREAU, *Fonctions convexes duales et points proximaux dans un espace hilbertien*, C. R. Acad. Sci. Paris Sér. A Math., 255 (1962), pp. 2897–2899.
- [54] J.-J. MOREAU, *Propriétés des applications 'prox,'* C. R. Acad. Sci. Paris Sér. A Math., 256 (1963), pp. 1069–1071.
- [55] J.-J. MOREAU, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France, 93 (1965), pp. 273–299.
- [56] D. NOLL, *Reconstruction with noisy data : An approach via eigenvalue optimization*, SIAM J. Optim., 8 (1998), pp. 82–104.
- [57] S. OSHER, A. SOLÉ, AND L. VESE, *Image decomposition and restoration using total variation minimization and the H^{-1} norm*, Multiscale Model. Simul., 1 (2003), pp. 349–370.
- [58] A. PAPOULIS, *The Fourier Integral and Its Applications*, McGraw-Hill, New York, 1962.
- [59] A. PAPOULIS, *A new algorithm in spectral analysis and band-limited extrapolation*, IEEE Trans. Circuits and Systems, 22 (1975), pp. 735–742.
- [60] J.-C. PESQUET AND P. L. COMBETTES, *Wavelet synthesis by alternating projections*, IEEE Trans. Signal Process., 44 (1996), pp. 728–732.
- [61] R. T. ROCKAFELLAR, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., 14 (1976), pp. 877–898.
- [62] L. I. RUDIN, S. OSHER, AND E. FATEMI, *Nonlinear total variation based noise removal algorithms*, Physica D, 60 (1992), pp. 259–268.
- [63] A. SABHARWAL AND L. C. POTTER, *Convexly constrained linear inverse problems : Iterative least-squares and regularization*, IEEE Trans. Signal Process., 46 (1998), pp. 2345–2352.
- [64] J.-L. STARCK, D. L. DONOHO, AND E. J. CANDÈS, *Astronomical image representation by the curvelet transform*, Astronom. and Astrophys., 398 (2003), pp. 785–800.

- [65] J.-L. STARCK, M. K. NGUYEN, AND F. MURTAGH, *Wavelets and curvelets for image deconvolution : A combined approach*, Signal Process., 83 (2003), pp. 2279–2283.
- [66] H. STARK, ED., *Image Recovery : Theory and Application*, Academic Press, San Diego, CA, 1987.
- [67] H. STARK AND Y. YANG, *Vector Space Projections : A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics*, Wiley, New York, 1998.
- [68] A. M. THOMPSON AND J. KAY, *On some Bayesian choices of regularization parameter in image restoration*, Inverse Problems, 9 (1993), pp. 749–761.
- [69] D. M. TITTERINGTON, *General structure of regularization procedures in image reconstruction*, Astronom. and Astrophys., 144 (1985), pp. 381–387.
- [70] S. TWOMEY, *The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements*, J. Franklin Inst., 279 (1965), pp. 95–109.
- [71] L. A. VESE AND S. J. OSHER, *Modeling textures with total variation minimization and oscillating patterns in image processing*, J. Sci. Comput., 19 (2003), pp. 553–572.
- [72] L. A. VESE AND S. J. OSHER, *Image denoising and decomposition with total variation minimization and oscillatory functions*, J. Math. Imaging Vision, 20 (2004), pp. 7–18.
- [73] D. C. YOULA, *Generalized image restoration by the method of alternating orthogonal projections*, IEEE Trans. Circuits and Systems, 25 (1978), pp. 694–702.
- [74] D. C. YOULA AND H. WEBB, *Image restoration by the method of convex projections : Part 1—Theory*, IEEE Trans. Medical Imaging, 1 (1982), pp. 81–94.
- [75] C. ZĂLINESCU, *On uniformly convex functions*, J. Math. Anal. Appl., 95 (1983), pp. 344–374.
- [76] C. ZĂLINESCU, *Convex Analysis in General Vector Spaces*, World Scientific, River Edge, NJ, 2002.
- [77] E. ZEIDLER, *Nonlinear Functional Analysis and Its Applications I : Fixed-Point Theorems*, Springer-Verlag, New York, 1993.

Chapitre 4

Formulation Variationnelle de Problèmes Inverses sur des Trames

Ce chapitre expose les travaux de recherche qui ont donné lieu à la publication rapportée à la Section 4.4. Après une introduction, les résultats essentiels de cette dernière sont présentés, pour finir par l'article publié en anglais.

4.1 Description générale

Ce chapitre vient élargir les hypothèses du Problème 3.3.51, problème inverse linéaire de régularisation parcimonieuse traité à la Section 3.3.5.4 du Chapitre 3 et rappelé ci-après :

Problème 4.1.1 Soient

- (i) \mathcal{H} et \mathcal{G} des espaces de Hilbert réels ;
- (ii) $T : \mathcal{H} \rightarrow \mathcal{G}$ un opérateur linéaire borné non nul ;
- (iii) $(e_k)_{k \in \mathbb{N}}$ une base hilbertienne de \mathcal{H} ;
- (iv) $(\phi_k)_{k \in \mathbb{N}}$ des fonctions de $\Gamma_0(\mathbb{R})$ telles que $(\forall k \in \mathbb{N}) \phi_k \geq \phi_k(0) = 0$.

L'objectif est de

$$\underset{x \in \mathcal{H}}{\text{minimiser}} \frac{1}{2} \|Tx - z\|^2 + \sum_{k \in \mathbb{N}} \phi_k(\langle x | e_k \rangle). \quad (4.1)$$

L'existence, l'unicité et la caractérisation des solutions sont étudiées dans la Section 3.3.5.4 du Chapitre 3. Un théorème de convergence de l'algorithme explicite-implicite y est également fourni.

Parmi les problèmes inverses étudiés sur un espace de Hilbert réel \mathcal{H} , certains sont soumis à des contraintes provenant d'informations a priori sur la solution recherchée \bar{x} . Pour tenir compte au mieux de ces données, et de plus, des propriétés physiques de ce signal \bar{x} , il est souvent plus judicieux de représenter \bar{x} par une décomposition selon une famille de vecteurs $(e_k)_{k \in \mathbb{K} \subset \mathbb{N}}$ de la forme

$$\bar{x} = \sum_{k \in \mathbb{K}} \bar{\xi}_k e_k, \quad (4.2)$$

en supposant détenir des informations a priori sur les coefficients $(\bar{\xi}_k)_{k \in \mathbb{K}}$. Le plus courant est de choisir pour $(e_k)_{k \in \mathbb{K}}$ une base hilbertienne de \mathcal{H} , par exemple, en utilisant les décompositions de Fourier ou en base d'ondelettes ou en base de bandelettes [8, 29, 30].

Dans ce chapitre, nous employons une représentation qui généralise les bases hilbertiennes, *la trame*, assez récemment exploitée [6, 7, 16, 21, 37]. Cette notion est définie et illustrée dans la Section 4.2.

De tels problèmes inverses se formulent, de manière synthétique, suivant le problème de minimisation :

Problème 4.1.2 Soient $(\phi_k)_{k \in \mathbb{K}}$ des fonctions de $\Gamma_0(\mathbb{R})$ telles que $\mathbb{K} = \{1, \dots, K\}$ avec $K \in \mathbb{N}$, ou $\mathbb{K} = \mathbb{N}$ et il existe un sous-ensemble \mathbb{L} de \mathbb{K} tel que

- (i) $\mathbb{K} \setminus \mathbb{L}$ est fini ;
- (ii) $(\forall k \in \mathbb{L}) \phi_k \geq 0$;
- (iii) Il existe une suite $(\zeta_k)_{k \in \mathbb{L}}$ réelle telle que :

$$\sum_{k \in \mathbb{L}} |\zeta_k|^2 < +\infty, \quad \sum_{k \in \mathbb{L}} |\text{prox}_{\phi_k} \zeta_k|^2 < +\infty, \quad \text{et} \quad \sum_{k \in \mathbb{L}} \phi_k(\zeta_k) < +\infty. \quad (4.3)$$

De plus, soit $\Psi \in \Gamma_0(\mathcal{H})$ différentiable sur \mathcal{H} de gradient τ -lipschitz, pour un certain $\tau \in]0, +\infty[$. L'objectif est de

$$\underset{(\xi_k)_{k \in \mathbb{K}} \in \ell^2(\mathbb{K})}{\text{minimiser}} \sum_{k \in \mathbb{K}} \phi_k(\xi_k) + \Psi \left(\sum_{k \in \mathbb{K}} \xi_k e_k \right). \quad (4.4)$$

La fonctionnelle à minimiser dans (4.4) s'écrit comme une somme de deux termes séparables : le premier ne dépend que des ϕ_k et pénalise chaque coefficient de la trame, tandis que le second ne dépend que de Ψ et pénalise la décomposition $x = \sum_{k \in \mathbb{K}} \xi_k e_k$ pour modéliser ainsi les contraintes directement imposées sur la solution cherchée \bar{x} .

En particulier, les formulations bayésiennes par approche statistique et les problèmes inverses sous contraintes parcimonieuses entrent dans ce cadre, comme l'explique la Section 4.4.3.3.

En approfondissant l'étude de l'outil théorique de base qu'est l'opérateur proximal (Section 4.4.2 et Section 4.4.4), nous résolvons le Problème 4.1.2 et proposons notamment un algorithme déduit de [15] (Section 4.4.3.2 et Section 4.4.5). Des simulations numériques de restauration d'images basée sur des trames viennent conclure cet article (Section 4.4.6).

4.2 Les trames

L'objectif originel des *trames* (*frames* en anglais) est d'extraire, par une représentation redondante, l'information a priori pertinente d'un signal ou d'un processus d'observations, présente mais cachée par des représentations complexes.

Commençons par définir cette notion :

Définition 4.2.1 Une famille $(e_k)_{k \in \mathbb{K}}$ de \mathcal{H} est une trame s'il existe deux constantes μ et ν de $]0, +\infty[$ telle que

$$(\forall x \in \mathcal{H}) \quad \mu \|x\|^2 \leq \sum_{k \in \mathbb{K}} |\langle x | e_k \rangle|^2 \leq \nu \|x\|^2. \quad (4.5)$$

Si $\mu = \nu$, la trame $(e_k)_{k \in \mathbb{K}}$ est dite ajustée ou étroite.

À la définition de la trame est étroitement liée celle d'opérateur de trame :

Définition 4.2.2 Soit $(e_k)_{k \in \mathbb{K}}$ une trame de \mathcal{H} . L'opérateur de trame associé est l'opérateur linéaire injectif borné

$$F: \mathcal{H} \rightarrow \ell^2(\mathbb{K}): x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{K}}. \quad (4.6)$$

L'adjoint de F est l'opérateur linéaire surjectif borné

$$F^*: \ell^2(\mathbb{K}) \rightarrow \mathcal{H}: (\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \xi_k e_k. \quad (4.7)$$

Les trames ont été introduites, en 1946, par Gabor [25, 26] pour apporter une représentation tenant compte des informations locales temporelles, contrairement aux transformées de Fourier. En effet, beaucoup de signaux évoluent avec le temps comme la parole, la musique. L'idée est de décomposer toute fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ selon la famille de fonctions élémentaires réelles $(g_{m,n})_{(m,n) \in \mathbb{Z}^2}$ telles que pour $(a, b) \in \mathbb{R}^2$ fixé et pour tout $(m, n) \in \mathbb{Z}^2$,

$$(\forall x \in \mathbb{R}) \quad g_{m,n}(x) = g(x - na)e^{2\pi imbx}, \quad (4.8)$$

de la manière suivante : il existe une famille de réels $(c_{m,n})_{(m,n) \in \mathbb{Z}^2}$ telle que

$$f = \sum_{(n,m) \in \mathbb{Z}^2} c_{m,n} g_{m,n}. \quad (4.9)$$

La théorie des trames a été ensuite approfondie par Duffin et Schaeffer [22] et raffinée dans les années 1980 par divers travaux tels que [18]. L'intérêt de la propriété de redondance des trames est alors mise en valeur : introduire la redondance permet de mieux prendre en compte le fait qu'un signal varie peu localement. Par exemple, un signal sonore (parole ou musique) ne change pas du tout au tout entre deux courts instants. De même, les pixels voisins d'une image sont pratiquement identiques. Plus encore, en autorisant des redondances, les trames révèlent d'autant mieux les aspects géométriques de l'image. En effet, elles tiennent compte des contours alors que les ondelettes s'intéressent uniquement aux directions horizontales, verticales et diagonales de l'image.

La théorie des trames a, de plus, une grande portée dans divers champs d'applications [25, 37] : la parole, la spectrographie, la mécanique et l'électricité pour représenter et identifier des systèmes linéaires, l'imagerie et la vision biologique avec le cortex visuel. Dans ce chapitre, nous nous limitons à des applications en déconvolution d'images.

Dressons maintenant une liste d'exemples de trames :

Exemple 4.2.3

- (i) D'un point de vue historique, comme nous venons de le mentionner, les trames de Gabor ont joué un grand rôle dans les problèmes inverses [16, 37].
- (ii) Les trames ajustées ($\mu = \nu = m$) : union de m bases hilbertiennes. Dans le cas particulier où $\mathcal{H} = L^2(\mathbb{R}^2)$, la décomposition d'ondelettes en arbre dual est l'union de deux bases d'ondelettes orthonormales [7].
- (iii) Les trames dont la dénomination termine par « let », comme les curvelets [6] ou les contourlets [21], ont l'avantage de fournir des décompositions directionnelles pour mieux préserver les contours et certaines propriétés géométriques de l'image.
- (iv) La base de Riesz de \mathcal{H} est aussi une trame : ceci correspond au cas particulier où F est bijective. Il existe alors une unique base biorthogonale $(\check{e}_k)_{k \in \mathbb{K}}$ telle que, pour tout $x \in \mathcal{H}$ et tout $(\xi_k)_{k \in \mathbb{K}} \in \ell^2(\mathbb{K})$,

$$x = F^*(\xi_k)_{k \in \mathbb{K}} \Leftrightarrow (\forall k \in \mathbb{K}) \quad \xi_k = \langle x \mid \check{e}_k \rangle. \quad (4.10)$$

Si $F^{-1} = F^*$, $(e_k)_{k \in \mathbb{K}}$ est une base hilbertienne et $(\check{e}_k)_{k \in \mathbb{K}} = (e_k)_{k \in \mathbb{K}}$. Des exemples de bases de Riesz dans $L^2(\mathbb{R}^2)$ sont fournis dans [9].

4.3 Les principaux résultats

4.3.1 Quelques cas d'applications du Problème 4.1.2

Présentons les deux cadres essentiels d'applications en traitement du signal, que nous allons explorer : la représentation parcimonieuse et l'approche bayésienne.

4.3.1.1 Représentation parcimonieuse

Une représentation parcimonieuse de la solution cherchée \bar{x} est obtenue quand la majorité des coefficients $(\bar{\xi}_k)_{k \in \mathbb{K}}$ de la décomposition (4.2) sont nuls. Pour tout $k \in \mathbb{K}$, tout revient à choisir un intervalle $\Omega_k \subset \mathbb{R}$ et à remplacer par 0 le k -ième coefficient de la décomposition du signal observé s'il appartient à Ω_k .

Un tel problème se modélise par le Problème 4.1.2 en choisissant :

- $\phi_k = \psi_k + \sigma_{\Omega_k}$, où ψ_k , différentiable en 0, vérifie $0 = \psi_k(0) \leq \psi_k \in \Gamma_0(\mathbb{R})$ et où σ_{Ω_k} est la fonction d'appui de Ω_k [14] ;
- reste à choisir Ψ qui révèle les contraintes imposées à \bar{x} . Rappelons, pour cela, le contexte du problème traité : il s'agit de restituer $\bar{x} \in \mathcal{H}$ à partir de q observations

$$z_i = T_i \bar{x} + v_i, \quad 1 \leq i \leq q, \quad (4.11)$$

où $T_i \in \mathcal{B}(\mathcal{H}, \mathcal{G}_i)$, \mathcal{G}_i est un espace de Hilbert réel, $z_i \in \mathcal{G}_i$, et $v_i \in \mathcal{G}_i$ est la réalisation du bruit. Nous supposons, de plus, connaître des contraintes imposées à \bar{x} et représentées par des ensembles convexes fermés $(S_i)_{1 \leq i \leq m}$ de \mathcal{H} [10, 35]. De toutes ces données, nous déduisons que la fonction Ψ utilisée dans le Problème 4.1.2 est :

$$\Psi : x \mapsto \frac{1}{2} \sum_{i=1}^q \alpha_i \|T_i x - z_i\|^2 + \frac{1}{2} \sum_{i=1}^m \vartheta_i d_{S_i}^2(x), \quad (4.12)$$

où $(\alpha_i)_{1 \leq i \leq q}$ et $(\vartheta_i)_{1 \leq i \leq m}$ sont des réels strictement positifs. Alors, $\Psi \in \Gamma_0(\mathcal{H})$ est bien différentiable sur \mathcal{H} de gradient τ -lipschitz [14, Section 5.1] avec

$$\tau = \left\| \sum_{i=1}^q \alpha_i T_i^* T_i \right\| + \sum_{i=1}^m \vartheta_i. \quad (4.13)$$

4.3.1.2 Approche bayésienne

Un problème inverse classique est de restituer $\bar{x} \in \mathcal{H}$ à partir d'une observation

$$z = T \bar{x} + v, \quad (4.14)$$

dans un espace de Hilbert réel \mathcal{G} , où $T \in \mathcal{B}(\mathcal{H}, \mathcal{G})$ et $v \in \mathcal{G}$ représente la réalisation du bruit. Si $\bar{x} = (\bar{\xi}_k)_{k \in \mathbb{K}}$ est la suite des coefficients de \bar{x} dans $(e_k)_{k \in \mathbb{K}}$, (4.14) peut s'écrire

$$z = TF^*\bar{x} + v. \quad (4.15)$$

Précisons le cadre dans lequel nous travaillons :

Hypothèse 4.3.1

- (i) $\mathcal{H} = \mathbb{R}^N$, $\mathcal{G} = \mathbb{R}^M$, et $\mathbb{K} = \{1, \dots, K\}$, où $K \geq N$.
- (ii) Les vecteurs \bar{x} , z et v sont respectivement les réalisations des variables aléatoires réelles \bar{X} , Z , et V définies sur le même espace de probabilisé.
- (iii) \bar{X} et V sont mutuellement indépendantes et ont pour densité de probabilité respectivement les fonctions $f_{\bar{X}}$ et f_V .
- (iv) Les composantes de \bar{X} sont indépendantes, de densités log-concave semi-continues supérieurement.
- (v) La fonction $\ln f_V$ est concave et différentiable de gradient lipschitz.

L'approche bayésienne consiste à estimer \bar{x} à partir de z par application de la règle du maximum a posteriori [3, 4, 36] et se ramène au Problème 4.1.2 en choisissant pour les $(\phi_k)_{k \in \mathbb{K}}$ les fonctions potentiels des densités de probabilité de \bar{X} et

$$(\forall x \in \mathcal{H}) \quad \Psi(x) = -\ln f_V(z - Tx). \quad (4.16)$$

De l'Hypothèse 4.3.1(v), nous déduisons que $\Psi \in \Gamma_0(\mathcal{H})$ est bien différentiable sur \mathcal{H} de gradient lipschitz.

Des remarques traitent le cas où V est un vecteur gaussien de moyenne nulle (Remarque 4.4.16 et Remarque 4.4.18).

4.3.2 Construction de solutions au Problème 4.1.2

Pour déterminer les solutions du Problème 4.1.2, l'outil théorique essentiel est, là encore, l'opérateur proximal. Dans cette étude, nous utilisons des résultats du Chapitre 3 et dégageons la Proposition 4.4.10, formule de décomposition proximale d'une fonction sur une base hilbertienne d'un espace de Hilbert réel \mathcal{X} . Des calculs explicites d'opérateurs proximaux de fonctions potentiels de distribution de probabilité sont également établis à des fins numériques à la Section 4.4.4. Toute cette théorie est détaillée dans le Chapitre 2.

Nous obtenons, par ce biais, tout d'abord des conditions d'existence de solutions au Problème 4.1.2 :

Proposition 4.3.2 *Supposons que l'une des propriétés suivantes est vérifiée.*

- (i) La fonction $(\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \phi_k(\xi_k) + \Psi(F^*(\xi_k)_{k \in \mathbb{K}})$ est coercive.
- (ii) $\inf_{k \in \mathbb{K}} \inf \phi_k(\mathbb{R}) > -\infty$, Ψ est coercive et $(e_k)_{k \in \mathbb{K}}$ est une base de Riesz.
- (iii) $\inf \Psi(\mathcal{H}) > -\infty$ et l'une des propriétés suivantes est vérifiée.
 - (a) La fonction $(\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \phi_k(\xi_k)$ est coercive.
 - (b) Il existe $\omega \in]0, +\infty[$ et $p \in [1, 2]$ tels que $(\forall k \in \mathbb{K}) \phi_k \geq \omega |\cdot|^p$.
 - (c) \mathbb{K} est fini et les fonctions $(\phi_k)_{k \in \mathbb{K}}$ sont coercives.

Alors le Problème 4.1.2 admet une solution.

Nous proposons, ensuite, une résolution du Problème 4.1.2 par un algorithme de type explicite-implicite :

Algorithme 4.3.3 Fixer $\mathbf{x}_0 \in \ell^2(\mathbb{K})$ et construire une suite $(\mathbf{x}_n)_{n \in \mathbb{N}} = ((\xi_{n,k})_{k \in \mathbb{K}})_{n \in \mathbb{N}}$ en posant pour tout $n \in \mathbb{N}$,

$$(\forall k \in \mathbb{K}) \quad \xi_{n+1,k} = \xi_{n,k} + \lambda_n \left(\text{prox}_{\gamma_n \phi_k} \left(\xi_{n,k} - \gamma_n (\eta_{n,k} + \beta_{n,k}) \right) + \alpha_{n,k} - \xi_{n,k} \right), \quad (4.17)$$

où $\lambda_n \in]0, 1]$, $\gamma_n \in]0, +\infty[$, $\{\alpha_{n,k}\}_{k \in \mathbb{K}} \subset \mathbb{R}$, $(\eta_{n,k})_{k \in \mathbb{K}} = F(\nabla \Psi(F^* \mathbf{x}_n))$ et $(\beta_{n,k})_{k \in \mathbb{K}} = Fb_n$, où $b_n \in \mathcal{H}$.

La convergence de cet algorithme est fournie sous certaines hypothèses supplémentaires par rapport à celles imposées dans le Problème 4.1.2 :

Hypothèse 4.3.4

- (i) Le Problème 4.1.2 admet une solution.
- (ii) $\inf_{n \in \mathbb{N}} \lambda_n > 0$.
- (iii) $\inf_{n \in \mathbb{N}} \gamma_n > 0$ et $\sup_{n \in \mathbb{N}} \gamma_n < 2/\beta$, où $\beta \in]0, +\infty[$ est tel que $F \circ \nabla \Psi \circ F^*$ est β -lipschitz.
- (iv) $\sum_{n \in \mathbb{N}} \sqrt{\sum_{k \in \mathbb{K}} |\alpha_{n,k}|^2} < +\infty$ et $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$.

Théorème 4.3.5 Soit $(\mathbf{x}_n)_{n \in \mathbb{N}}$ une suite générée par l'Algorithme 4.3.3 sous l'Hypothèse 4.3.4. Alors $(\mathbf{x}_n)_{n \in \mathbb{N}}$ converge faiblement vers une solution du Problème 4.1.2.

Le cas particulier où $(e_k)_{k \in \mathbb{K}}$ est une base de Riesz est étudié au Corollaire 4.4.39. Des exemples de problèmes conséquents en traitement du signal sont aussi déduits du Problème 4.1.2, dans la Remarque 4.4.40, lorsque $\mathbb{K} = \mathbb{N}$ et $(e_k)_{k \in \mathbb{K}}$ est une base hilbertienne de \mathcal{H} .

4.3.3 Simulations numériques

Pour chaque application exposée à la Section 4.3.1, nous présentons un exemple de restauration d'images.

L'espace de Hilbert est $\mathcal{H} = \mathbb{R}^{512 \times 512}$. Pour l'approche bayésienne, nous choisissons comme fonctions potentiels, celles de l'Exemple 4.4.24, du maximum d'entropie : $\phi_k = \omega_k |\cdot| + \tau_k |\cdot|^2 + \kappa_k |\cdot|^{p_k}$, où $p_k \in \{4/3, 3/2, 3, 4\}$ et $\{\omega_k, \tau_k, \kappa_k\} \subset]0, +\infty[$. Pour ces valeurs de p_k , l'opérateur proximal de ϕ_k est explicite. Les autres paramètres $\omega_k, \tau_k, \kappa_k$ sont choisis pour chaque sous-bande d'ondelettes par une approche par maximum de vraisemblance.

Pour mesurer l'efficacité de notre méthode, nous mesurons, en décibels, l'erreur relative entre les images z et \bar{x} , définie par $20 \log_{10}(\|\bar{x}\|/\|z - \bar{x}\|)$. Le critère visuel est aussi concluant.

4.4 Article en anglais

A VARIATIONAL FORMULATION FOR FRAME-BASED INVERSE PROBLEMS¹

Abstract : A convex variational framework is proposed for solving inverse problems in Hilbert spaces with a priori information on the representation of the target solution in a frame. The objective function to be minimized consists of a separable term penalizing each frame coefficient individually and of a smooth term modeling the data formation model as well as other constraints. Sparsity-constrained and Bayesian formulations are examined as special cases. A splitting algorithm is presented to solve this problem and its convergence is established in infinite-dimensional spaces under mild conditions on the penalization functions, which need not be differentiable. Numerical simulations demonstrate applications to frame-based image restoration.

4.4.1 Introduction

In inverse problems, certain physical properties of the target solution \bar{x} are most suitably expressed in terms of the coefficients $(\bar{\xi}_k)_{k \in \mathbb{K} \subset \mathbb{N}}$ of its representation $\bar{x} = \sum_{k \in \mathbb{K}} \bar{\xi}_k e_k$ with respect to a family of vectors $(e_k)_{k \in \mathbb{K}}$ in a Hilbert space $(\mathcal{H}, \|\cdot\|)$. Traditionally, such linear representations have been mostly centered on orthonormal bases as, for instance, in Fourier, wavelet, or bandlet decompositions [8, 29, 30]. Recently, attention has shifted towards more general, overcomplete representations

¹C. Chaux, P. L. Combettes, J.-C. Pesquet and V. R. Wajs, A variational formulation for frame-based inverse problems, *Inverse Problems*, vol. 23, pp. 1495–1518, 2007.

known as *frames*; see [6, 7, 16, 21, 37] for specific examples. Recall that a family of vectors $(e_k)_{k \in \mathbb{K}}$ in \mathcal{H} constitutes a frame if there exist two constants μ and ν in $]0, +\infty[$ such that

$$(\forall x \in \mathcal{H}) \quad \mu \|x\|^2 \leq \sum_{k \in \mathbb{K}} |\langle x | e_k \rangle|^2 \leq \nu \|x\|^2. \quad (4.18)$$

The associated frame operator is the injective bounded linear operator

$$F: \mathcal{H} \rightarrow \ell^2(\mathbb{K}): x \mapsto (\langle x | e_k \rangle)_{k \in \mathbb{K}}, \quad (4.19)$$

the adjoint of which is the surjective bounded linear operator

$$F^*: \ell^2(\mathbb{K}) \rightarrow \mathcal{H}: (\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \xi_k e_k. \quad (4.20)$$

When $\mu = \nu$ in (4.18), $(e_k)_{k \in \mathbb{K}}$ is said to be a *tight frame*. A simple example of a tight frame is the union of m orthonormal bases, in which case $\mu = \nu = m$. For instance, in $\mathcal{H} = L^2(\mathbb{R}^2)$, a real dual-tree wavelet decomposition is the union of two orthonormal wavelet bases [7, 34]. Curvelets [6] constitute another example of a tight frame of $L^2(\mathbb{R}^2)$. Historically, Gabor frames [16, 37] have played an important role in many inverse problems. Another common example of a frame is a Riesz basis, which corresponds to the case when $(e_k)_{k \in \mathbb{K}}$ is linearly independent or, equivalently, when F is bijective. In such instances, there exists a unique Riesz basis $(\check{e}_k)_{k \in \mathbb{K}}$ such that $(e_k)_{k \in \mathbb{K}}$ and $(\check{e}_k)_{k \in \mathbb{K}}$ are biorthogonal. Furthermore, for every $x \in \mathcal{H}$ and $(\xi_k)_{k \in \mathbb{K}} \in \ell^2(\mathbb{K})$,

$$x = F^*(\xi_k)_{k \in \mathbb{K}} \quad \Leftrightarrow \quad (\forall k \in \mathbb{K}) \quad \xi_k = \langle x | \check{e}_k \rangle. \quad (4.21)$$

When $F^{-1} = F^*$, $(e_k)_{k \in \mathbb{K}}$ is an orthonormal basis and $(\check{e}_k)_{k \in \mathbb{K}} = (e_k)_{k \in \mathbb{K}}$. Examples of Riesz bases of $L^2(\mathbb{R}^2)$ include biorthogonal bases of compactly supported dyadic wavelets having certain symmetry properties [9]. Further constructions as well as a detailed account of frame theory in Hilbert spaces can be found in [27].

The goal of the present paper is to propose a flexible convex variational framework for solving inverse problems in which a priori information (e.g., sparsity, distribution, statistical properties) is available about the representation of the target solution in a frame. Our analysis and our numerical algorithm will rely heavily on proximity operators. Section 4.4.2 is devoted to these operators. Our main variational formulation is presented and analyzed in Section 4.4.3. It consists (see Problem 4.4.11) of minimizing the sum of a separable, possibly nondifferentiable function penalizing each coefficient of the frame decomposition individually, and of a smooth function which combines other information on the problem and the data formation model. Connections with sparsity-constrained and Bayesian formulations are also established. In connection with the latter, we derive in Section 4.4.4 closed-form expressions for the proximity operators associated with a variety of univariate

log-concave distributions. A proximal algorithm for solving Problem 4.4.11 is presented in Section 4.4.5 and its convergence is established in infinite-dimensional spaces under mild assumptions on the penalization functions. An attractive feature of this algorithm is that it is fully split in that, at each iteration, all the functions appearing in the problem are activated individually. Finally, applications to image recovery are demonstrated in Section 4.4.6.

4.4.2 Basic tool : proximity operator

4.4.2.1 Notation

Throughout, \mathcal{X} is a separable real Hilbert space with scalar product $\langle \cdot | \cdot \rangle$, norm $\|\cdot\|$, and distance d . $\Gamma_0(\mathcal{X})$ is the class of lower semicontinuous convex functions from \mathcal{X} to $] -\infty, +\infty]$ which are not identically equal to $+\infty$. The indicator function of a subset S of \mathcal{X} is

$$\iota_S: \mathbf{x} \mapsto \begin{cases} 0, & \text{if } \mathbf{x} \in S; \\ +\infty, & \text{if } \mathbf{x} \notin S, \end{cases} \quad (4.22)$$

its support function is $\sigma_S: \mathcal{X} \rightarrow [-\infty, +\infty]: \mathbf{u} \mapsto \sup_{\mathbf{x} \in S} \langle \mathbf{x} | \mathbf{u} \rangle$, and its distance function is $d_S: \mathcal{X} \rightarrow [0, +\infty]: \mathbf{x} \mapsto \inf \|S - \mathbf{x}\|$. If S is nonempty, closed, and convex then, for every $\mathbf{x} \in \mathcal{X}$, there exists a unique point $P_S \mathbf{x}$ in S , called the projection of \mathbf{x} onto S , such that $\|\mathbf{x} - P_S \mathbf{x}\| = d_S(\mathbf{x})$ (further background on convex analysis will be found in [40]).

4.4.2.2 Background

Let $\varphi \in \Gamma_0(\mathcal{X})$. The subdifferential of φ at $\mathbf{x} \in \mathcal{X}$ is the set

$$\partial\varphi(\mathbf{x}) = \{\mathbf{u} \in \mathcal{X} \mid (\forall \mathbf{y} \in \mathcal{X}) \langle \mathbf{y} - \mathbf{x} | \mathbf{u} \rangle + \varphi(\mathbf{x}) \leq \varphi(\mathbf{y})\}. \quad (4.23)$$

If φ is Gâteaux differentiable at \mathbf{x} with gradient $\nabla\varphi(\mathbf{x})$, then $\partial\varphi(\mathbf{x}) = \{\nabla\varphi(\mathbf{x})\}$. The conjugate of φ is the function $\varphi^* \in \Gamma_0(\mathcal{X})$ defined by

$$(\forall \mathbf{u} \in \mathcal{X}) \quad \varphi^*(\mathbf{u}) = \sup_{\mathbf{x} \in \mathcal{X}} \langle \mathbf{x} | \mathbf{u} \rangle - \varphi(\mathbf{x}). \quad (4.24)$$

The continuous convex function

$$\gamma\varphi: \mathcal{X} \rightarrow \mathbb{R}: \mathbf{x} \mapsto \inf_{\mathbf{y} \in \mathcal{X}} \varphi(\mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{y}\|^2 \quad (4.25)$$

is the Moreau envelope of index $\gamma \in]0, +\infty[$ of φ .

Definition 4.4.1 [32] Let $\varphi \in \Gamma_0(\mathcal{X})$. Then, for every $\mathbf{x} \in \mathcal{X}$, the function $\mathbf{y} \mapsto \varphi(\mathbf{y}) + \|\mathbf{x} - \mathbf{y}\|^2/2$ achieves its infimum at a unique point denoted by $\text{prox}_\varphi \mathbf{x}$. The operator $\text{prox}_\varphi: \mathcal{X} \rightarrow \mathcal{X}$ thus defined is the *proximity operator* of φ . Moreover, for every \mathbf{x} and \mathbf{p} in \mathcal{X}

$$\mathbf{p} = \text{prox}_\varphi \mathbf{x} \Leftrightarrow \mathbf{x} - \mathbf{p} \in \partial\varphi(\mathbf{p}) \quad (4.26)$$

$$\Leftrightarrow (\forall \mathbf{y} \in \mathcal{X}) \langle \mathbf{y} - \mathbf{p} \mid \mathbf{x} - \mathbf{p} \rangle + \varphi(\mathbf{p}) \leq \varphi(\mathbf{y}). \quad (4.27)$$

Example 4.4.2 Let $\gamma \in]0, +\infty[$, let S be a nonempty convex subset of \mathcal{X} , and set $\varphi = \iota_S$. Then it follows at once from (4.22), (4.25), and Definition 4.4.1 that $\gamma\varphi = d_S^2/(2\gamma)$ and $\text{prox}_{\gamma\varphi} = P_S$.

Here are basic properties of the proximity operator.

Lemma 4.4.3 [15, Section 2] *Let $\varphi \in \Gamma_0(\mathcal{X})$. Then the following hold.*

- (i) $(\forall \mathbf{x} \in \mathcal{X}) \mathbf{x} \in \text{Argmin } \varphi \Leftrightarrow 0 \in \partial\varphi(\mathbf{x}) \Leftrightarrow \text{prox}_\varphi \mathbf{x} = \mathbf{x}$.
- (ii) $(\forall \mathbf{x} \in \mathcal{X})(\forall \mathbf{y} \in \mathcal{X}) \|\text{prox}_\varphi \mathbf{x} - \text{prox}_\varphi \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$.
- (iii) $(\forall \mathbf{x} \in \mathcal{X})(\forall \gamma \in]0, +\infty[) \mathbf{x} = \text{prox}_{\gamma\varphi} \mathbf{x} + \gamma \text{prox}_{\varphi^*/\gamma}(\mathbf{x}/\gamma)$.

In Lemma 4.4.3, (i) states that the minimizers of φ are characterized as the zeros of the subdifferential of φ (Fermat's rule) or, equivalently, as the fixed points of prox_φ ; (ii) states that prox_φ is nonexpansive, which turns out to be an essential property in the convergence of iterative methods [13]; finally, (iii) is Moreau's decomposition principle [31], which provides a powerful nonlinear decomposition rule parametrized by φ and extends in particular the standard orthogonal decomposition rule [15, Remark 2.11].

Lemma 4.4.4 *Let $\varphi \in \Gamma_0(\mathcal{X})$, let $\gamma \in]0, +\infty[$, and set $\psi = \gamma\varphi$. Then the following hold.*

- (i) ψ is Fréchet-differentiable on \mathcal{X} .
- (ii) $\nabla\psi = (\text{Id} - \text{prox}_{\gamma\varphi})/\gamma = \text{prox}_{\varphi^*/\gamma}(\cdot/\gamma)$.
- (iii) $\nabla\psi$ is $(1/\gamma)$ -Lipschitz continuous.

Proof. (i) and (ii) : A routine extension of [32, Proposition 7.d], where $\gamma = 1$. (iii) : Since $\varphi^*/\gamma \in \Gamma_0(\mathcal{X})$ and (ii) asserts that $\nabla\psi = \text{prox}_{\varphi^*/\gamma}(\cdot/\gamma)$, this is a direct consequence of Lemma 4.4.3(ii). \square

Next, we record some proximal calculus rules that will allow us to derive new proximity operators from existing ones.

Lemma 4.4.5 [15, Lemma 2.6] *Let $\varphi \in \Gamma_0(\mathcal{X})$ and let $\mathbf{x} \in \mathcal{X}$. Then the following hold.*

- (i) Let $\psi = \varphi + \alpha \|\cdot\|^2/2 + \langle \cdot | \mathbf{u} \rangle + \beta$, where $\mathbf{u} \in \mathcal{X}$, $\alpha \in [0, +\infty[$, and $\beta \in \mathbb{R}$. Then $\text{prox}_\psi \mathbf{x} = \text{prox}_{\varphi/(\alpha+1)}((\mathbf{x} - \mathbf{u})/(\alpha + 1))$.
- (ii) Let $\psi = \varphi(\cdot - \mathbf{z})$, where $\mathbf{z} \in \mathcal{X}$. Then $\text{prox}_\psi \mathbf{x} = \mathbf{z} + \text{prox}_\varphi(\mathbf{x} - \mathbf{z})$.
- (iii) Let $\psi = \varphi(\cdot/\rho)$, where $\rho \in \mathbb{R} \setminus \{0\}$. Then $\text{prox}_\psi \mathbf{x} = \rho \text{prox}_{\varphi/\rho^2}(\mathbf{x}/\rho)$.
- (iv) Let $\psi: \mathbf{y} \mapsto \varphi(-\mathbf{y})$. Then $\text{prox}_\psi \mathbf{x} = -\text{prox}_\varphi(-\mathbf{x})$.

We conclude this section with some properties of proximity operators on the real line.

Lemma 4.4.6 [14, Section 2] *Let $\phi \in \Gamma_0(\mathbb{R})$. Then the following hold.*

- (i) $\text{prox}_\phi: \mathbb{R} \rightarrow \mathbb{R}$ is increasing.
- (ii) Suppose that ϕ admits 0 as a minimizer. Then

$$(\forall \xi \in \mathbb{R}) \quad \begin{cases} 0 \leq \text{prox}_\phi \xi \leq \xi, & \text{if } \xi > 0; \\ \text{prox}_\phi \xi = 0, & \text{if } \xi = 0; \\ \xi \leq \text{prox}_\phi \xi \leq 0, & \text{if } \xi < 0. \end{cases} \quad (4.28)$$

This is true in particular when ϕ is even.

- (iii) Suppose that $\phi = \psi + \sigma_\Omega$, where $\psi \in \Gamma_0(\mathbb{R})$ is differentiable at 0 with $\psi'(0) = 0$, and where $\Omega \subset \mathbb{R}$ is a nonempty closed interval. Then $\text{prox}_\phi = \text{prox}_\psi \circ \text{soft}_\Omega$, where

$$\text{soft}_\Omega = \text{prox}_{\sigma_\Omega}: \mathbb{R} \rightarrow \mathbb{R}: \xi \mapsto \begin{cases} \xi - \underline{\omega}, & \text{if } \xi < \underline{\omega}; \\ 0, & \text{if } \xi \in \Omega; \\ \xi - \bar{\omega}, & \text{if } \xi > \bar{\omega}, \end{cases} \quad \text{with} \quad \begin{cases} \underline{\omega} = \inf \Omega, \\ \bar{\omega} = \sup \Omega, \end{cases} \quad (4.29)$$

is the soft thresholder on Ω . In particular, if $\Omega = [-\omega, \omega]$ for some $\omega \in]0, +\infty[$, we obtain

$$\text{soft}_{[-\omega, \omega]} = \text{prox}_{\omega|\cdot|}: \mathbb{R} \rightarrow \mathbb{R}: \xi \mapsto \text{sign}(\xi) \max\{|\xi| - \omega, 0\}. \quad (4.30)$$

The soft-thresholding operation described in Lemma 4.4.6(iii) is illustrated in Fig. 4.1.

4.4.2.3 Forward-backward splitting

In this section, we consider the following abstract variational framework, that will cover our main problem (Problem 4.4.11).

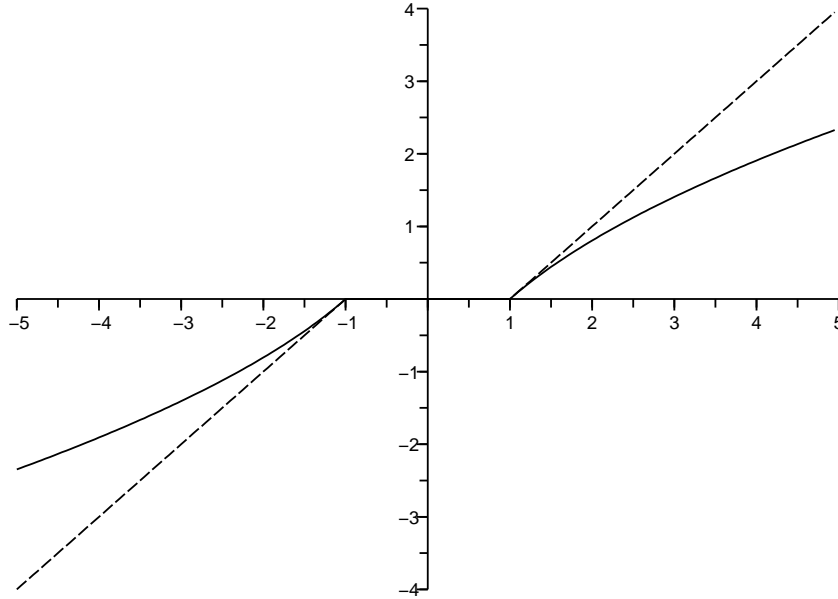


FIG. 4.1 – $\Omega = [-1, 1]$. Graphs of soft_Ω (dashed line) and prox_ϕ (solid line) in Lemma 4.4.6(iii) with $\psi = 0.1|\cdot|^3$.

Problem 4.4.7 Let f_1 and f_2 be functions in $\Gamma_0(\mathcal{X})$ such that f_2 is differentiable on \mathcal{X} with a β -Lipschitz continuous gradient for some $\beta \in]0, +\infty[$. The objective is to

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} \quad f_1(\mathbf{x}) + f_2(\mathbf{x}). \quad (4.31)$$

A key consequence of Fermat’s rule (Lemma 4.4.3(i)) and (4.26) is the following characterization of the solutions to Problem 4.4.7 which, in itself, attests the central role played by proximity operators.

Proposition 4.4.8 [15, Proposition 3.1(iii)] *Let $\mathbf{x} \in \mathcal{X}$ and let $\gamma \in]0, +\infty[$. Then \mathbf{x} is a solution to Problem 4.4.7 if and only if $\mathbf{x} = \text{prox}_{\gamma f_1}(\mathbf{x} - \gamma \nabla f_2(\mathbf{x}))$.*

Let $\gamma \in \mathbb{R}$ and set $T = \text{prox}_{\gamma f_1} \circ (\text{Id} - \gamma \nabla f_2)$. Proposition 4.4.8 asserts that a point $\mathbf{x} \in \mathcal{X}$ solves Problem 4.4.7 if and only if $\mathbf{x} = T\mathbf{x}$. This fixed point characterization suggests solving Problem 4.4.7 via the successive approximation method $\mathbf{x}_{n+1} = T\mathbf{x}_n$, for a suitable value of the “step size” parameter γ . The next result describes an algorithm in this vein, which is based on the forward-backward splitting

method for monotone operators [13]. It allows for inexact evaluations of the operators prox_{f_1} and ∇f_2 via the incorporation of the error sequences $(\mathbf{a}_n)_{n \in \mathbb{N}}$ and $(\mathbf{b}_n)_{n \in \mathbb{N}}$, respectively, as well as for iteration-dependent relaxation parameters $(\lambda_n)_{n \in \mathbb{N}}$ and step sizes $(\gamma_n)_{n \in \mathbb{N}}$.

Theorem 4.4.9 [15, Theorem 3.4(i)] *Suppose that $\text{Argmin}(f_1 + f_2) \neq \emptyset$. Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, +\infty[$ such that $0 < \inf_{n \in \mathbb{N}} \gamma_n \leq \sup_{n \in \mathbb{N}} \gamma_n < 2/\beta$, let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that $\inf_{n \in \mathbb{N}} \lambda_n > 0$, and let $(\mathbf{a}_n)_{n \in \mathbb{N}}$ and $(\mathbf{b}_n)_{n \in \mathbb{N}}$ be sequences in \mathcal{X} such that $\sum_{n \in \mathbb{N}} \|\mathbf{a}_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|\mathbf{b}_n\| < +\infty$. Fix $\mathbf{x}_0 \in \mathcal{X}$ and, for every $n \in \mathbb{N}$, set*

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \lambda_n \left(\text{prox}_{\gamma_n f_1} \left(\mathbf{x}_n - \gamma_n (\nabla f_2(\mathbf{x}_n) + \mathbf{b}_n) \right) + \mathbf{a}_n - \mathbf{x}_n \right). \quad (4.32)$$

Then $(\mathbf{x}_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 4.4.7.

4.4.2.4 Decomposition formula

The following decomposition property, which extends [15, Example 2.19], will be instrumental in our analysis.

Proposition 4.4.10 *Set $\Upsilon : \mathcal{X} \rightarrow]-\infty, +\infty] : \mathbf{x} \mapsto \sum_{i \in \mathbb{I}} \psi_i(\langle \mathbf{x} | \mathbf{o}_i \rangle)$, where :*

- (i) $\emptyset \neq \mathbb{I} \subset \mathbb{N}$;
- (ii) $(\mathbf{o}_i)_{i \in \mathbb{I}}$ is an orthonormal basis of \mathcal{X} ;
- (iii) $(\psi_i)_{i \in \mathbb{I}}$ are functions in $\Gamma_0(\mathbb{R})$;
- (iv) Either \mathbb{I} is finite, or there exists a subset \mathbb{J} of \mathbb{I} such that :
 - (a) $\mathbb{I} \setminus \mathbb{J}$ is finite;
 - (b) $(\forall i \in \mathbb{J}) \psi_i \geq 0$;
 - (c) there exists a sequence $(\zeta_i)_{i \in \mathbb{J}}$ in \mathbb{R} such that $\sum_{i \in \mathbb{J}} |\zeta_i|^2 < +\infty$, $\sum_{i \in \mathbb{J}} |\text{prox}_{\psi_i} \zeta_i|^2 < +\infty$, and $\sum_{i \in \mathbb{J}} \psi_i(\zeta_i) < +\infty$.

Then $\Upsilon \in \Gamma_0(\mathcal{X})$ and $(\forall \mathbf{x} \in \mathcal{X}) \text{prox}_{\Upsilon} \mathbf{x} = \sum_{i \in \mathbb{I}} (\text{prox}_{\psi_i} \langle \mathbf{x} | \mathbf{o}_i \rangle) \mathbf{o}_i$.

Proof. We treat only the case when \mathbb{I} is infinite as the case when \mathbb{I} is finite will follow trivially the arguments presented below. Fix, for every $i \in \mathbb{I} \setminus \mathbb{J}$, $\zeta_i \in \mathbb{R}$ such that $\psi_i(\zeta_i) < +\infty$ and set $\mathbf{z} = \sum_{i \in \mathbb{I}} \zeta_i \mathbf{o}_i$. Then (iv) implies that $\sum_{i \in \mathbb{I}} \zeta_i^2 < +\infty$ and, in view of (ii), that $\mathbf{z} \in \mathcal{X}$. Moreover, $\Upsilon(\mathbf{z}) = \sum_{i \in \mathbb{J}} \psi_i(\zeta_i) + \sum_{i \in \mathbb{I} \setminus \mathbb{J}} \psi_i(\zeta_i) < +\infty$.

Let us show that $\Upsilon \in \Gamma_0(\mathcal{X})$. As just seen, $\Upsilon(\mathbf{z}) < +\infty$ and, therefore, $\Upsilon \neq +\infty$. Next, we observe that, by virtue of (iii), the functions $(\psi_i(\langle \cdot | \mathbf{o}_i \rangle))_{i \in \mathbb{I}}$ are lower semicontinuous and convex. As a result, $\sum_{i \in \mathbb{I} \setminus \mathbb{J}} \psi_i(\langle \cdot | \mathbf{o}_i \rangle)$ is lower semicontinuous and convex, as a finite sum of such functions. Thus, to show that $\Upsilon \in \Gamma_0(\mathcal{X})$, it

remains to show that $\Upsilon_{\mathbb{J}} = \sum_{i \in \mathbb{J}} \psi_i(\langle \cdot | \mathbf{o}_i \rangle)$ is lower semicontinuous and convex. It follows from (iv)(b) that

$$\Upsilon_{\mathbb{J}} = \sup_{\substack{\mathbb{J}' \subset \mathbb{J} \\ \mathbb{J}' \text{ finite}}} \sum_{i \in \mathbb{J}'} \psi_i(\langle \cdot | \mathbf{o}_i \rangle). \quad (4.33)$$

However, as above, each finite sum $\sum_{i \in \mathbb{J}'} \psi_i(\langle \cdot | \mathbf{o}_i \rangle)$ is lower semicontinuous and convex. Therefore, $\Upsilon_{\mathbb{J}}$ is likewise as the supremum of a family of lower semicontinuous convex functions.

Now fix $\mathbf{x} \in \mathcal{X}$ and set

$$(\forall i \in \mathbb{I}) \quad \xi_i = \langle \mathbf{x} | \mathbf{o}_i \rangle \quad \text{and} \quad \pi_i = \text{prox}_{\psi_i} \xi_i. \quad (4.34)$$

It follows from (iv)(a) and (iv)(c) that

$$\sum_{i \in \mathbb{I}} |\text{prox}_{\psi_i} \zeta_i|^2 = \sum_{i \in \mathbb{J}} |\text{prox}_{\psi_i} \zeta_i|^2 + \sum_{i \in \mathbb{I} \setminus \mathbb{J}} |\text{prox}_{\psi_i} \zeta_i|^2 < +\infty. \quad (4.35)$$

Hence, we derive from Lemma 4.4.3(ii) and (ii) that

$$\begin{aligned} \frac{1}{2} \sum_{i \in \mathbb{I}} |\pi_i|^2 &\leq \sum_{i \in \mathbb{I}} |\pi_i - \text{prox}_{\psi_i} \zeta_i|^2 + \sum_{i \in \mathbb{I}} |\text{prox}_{\psi_i} \zeta_i|^2 \\ &= \sum_{i \in \mathbb{I}} |\text{prox}_{\psi_i} \xi_i - \text{prox}_{\psi_i} \zeta_i|^2 + \sum_{i \in \mathbb{I}} |\text{prox}_{\psi_i} \zeta_i|^2 \\ &\leq \sum_{i \in \mathbb{I}} |\xi_i - \zeta_i|^2 + \sum_{i \in \mathbb{I}} |\text{prox}_{\psi_i} \zeta_i|^2 \\ &= \|\mathbf{x} - \mathbf{z}\|^2 + \sum_{i \in \mathbb{I}} |\text{prox}_{\psi_i} \zeta_i|^2 \\ &< +\infty. \end{aligned} \quad (4.36)$$

Let us set $\mathbf{p} = \sum_{i \in \mathbb{I}} \pi_i \mathbf{o}_i$. Then it follows from (4.36) and (ii) that $\mathbf{p} \in \mathcal{X}$. On the other hand, we derive from (4.34) and (4.27) that

$$(\forall i \in \mathbb{I})(\forall \eta \in \mathbb{R}) \quad (\eta - \pi_i)(\xi_i - \pi_i) + \psi_i(\pi_i) \leq \psi_i(\eta). \quad (4.37)$$

Hence, by Parseval and (ii),

$$\begin{aligned} (\forall \mathbf{y} \in \mathcal{X}) \quad \langle \mathbf{y} - \mathbf{p} | \mathbf{x} - \mathbf{p} \rangle + \Upsilon(\mathbf{p}) &= \sum_{i \in \mathbb{I}} \langle \mathbf{y} - \mathbf{p} | \mathbf{o}_i \rangle \langle \mathbf{x} - \mathbf{p} | \mathbf{o}_i \rangle + \sum_{i \in \mathbb{I}} \psi_i(\pi_i) \\ &= \sum_{i \in \mathbb{I}} (\langle \mathbf{y} | \mathbf{o}_i \rangle - \pi_i)(\xi_i - \pi_i) + \psi_i(\pi_i) \\ &\leq \sum_{i \in \mathbb{I}} \psi_i(\langle \mathbf{y} | \mathbf{o}_i \rangle) \\ &= \Upsilon(\mathbf{y}). \end{aligned} \quad (4.38)$$

Invoking (4.27) once again, we conclude that $\mathbf{p} = \text{prox}_{\Upsilon} \mathbf{x}$. \square

4.4.3 Problem formulation

4.4.3.1 Assumptions and problem statement

Throughout, \mathcal{H} is a separable real Hilbert space with scalar product $\langle \cdot | \cdot \rangle$, norm $\|\cdot\|$, and distance d . The index set \mathbb{K} is either $\{1, \dots, K\}$ ($K \in \mathbb{N}$) or \mathbb{N} , according as \mathcal{H} is finite or infinite dimensional. Moreover, $(e_k)_{k \in \mathbb{K}}$ is a frame in \mathcal{H} with constants μ and ν (see (4.18)) and frame operator F (see (4.19)). Finally, the sequence of frame coefficients of a generic point $x \in \mathcal{H}$ will be denoted by \mathbf{x} , i.e., $\mathbf{x} = (\xi_k)_{k \in \mathbb{K}}$, where $x = \sum_{k \in \mathbb{K}} \xi_k e_k$.

Let $\bar{x} \in \mathcal{H}$ be the target solution of the underlying inverse problem. Our basic premise is that a priori information is available about the coefficients $(\bar{\xi}_k)_{k \in \mathbb{K}}$ of the decomposition

$$\bar{x} = \sum_{k \in \mathbb{K}} \bar{\xi}_k e_k \quad (4.39)$$

of \bar{x} in $(e_k)_{k \in \mathbb{K}}$. To recover \bar{x} , it is therefore natural to formulate a variational problem in the space $\ell^2(\mathbb{K})$ of frame coefficients, where a priori information on $(\bar{\xi}_k)_{k \in \mathbb{K}}$ can be easily incorporated. More precisely, a solution will assume the form $\tilde{x} = \sum_{k \in \mathbb{K}} \tilde{\xi}_k e_k$, where $(\tilde{\xi}_k)_{k \in \mathbb{K}}$ is a solution to the following problem.

Problem 4.4.11 Let $(\phi_k)_{k \in \mathbb{K}}$ be functions in $\Gamma_0(\mathbb{R})$ such that either $\mathbb{K} = \{1, \dots, K\}$ with $K \in \mathbb{N}$, or $\mathbb{K} = \mathbb{N}$ and there exists a subset \mathbb{L} of \mathbb{K} such that

- (i) $\mathbb{K} \setminus \mathbb{L}$ is finite;
- (ii) $(\forall k \in \mathbb{L}) \phi_k \geq 0$;
- (iii) there exists a sequence $(\zeta_k)_{k \in \mathbb{L}}$ in \mathbb{R} such that $\sum_{k \in \mathbb{L}} |\zeta_k|^2 < +\infty$, $\sum_{k \in \mathbb{L}} |\text{prox}_{\phi_k} \zeta_k|^2 < +\infty$, and $\sum_{k \in \mathbb{L}} \phi_k(\zeta_k) < +\infty$.

In addition, let $\Psi \in \Gamma_0(\mathcal{H})$ be differentiable on \mathcal{H} with a τ -Lipschitz continuous gradient for some $\tau \in]0, +\infty[$. The objective is to

$$\underset{(\xi_k)_{k \in \mathbb{K}} \in \ell^2(\mathbb{K})}{\text{minimize}} \sum_{k \in \mathbb{K}} \phi_k(\xi_k) + \Psi \left(\sum_{k \in \mathbb{K}} \xi_k e_k \right). \quad (4.40)$$

Remark 4.4.12

- (i) The functions $(\phi_k)_{k \in \mathbb{K}}$ in Problem 4.4.11 need not be differentiable. As will be seen in Section 4.4.3.3.1, this feature is essential in sparsity-constrained problems.
- (ii) Suppose that $\mathbb{K} = \mathbb{N}$. Then Conditions (ii) and (iii) in Problem 4.4.11 hold when, for every $k \in \mathbb{L}$, ϕ_k admits a minimizer ζ_k such that $\phi_k(\zeta_k) = 0$ and $\sum_{k \in \mathbb{L}} |\zeta_k|^2 < +\infty$. Indeed, we then have $\sum_{k \in \mathbb{L}} |\text{prox}_{\phi_k} \zeta_k|^2 = \sum_{k \in \mathbb{L}} |\zeta_k|^2 < +\infty$ (this follows from Lemma 4.4.3(i)) and $\sum_{k \in \mathbb{L}} \phi_k(\zeta_k) = 0$. In particular, Conditions (ii) and (iii) in Problem 4.4.11 hold when $(\forall k \in \mathbb{L}) \phi_k \geq \phi_k(0) = 0$, which amounts to setting $\zeta_k \equiv 0$.

4.4.3.2 Existence and characterization of solutions

We first address the issue of the existence of solutions to Problem 4.4.11. Recall that a function $\varphi: \mathcal{H} \rightarrow]-\infty, +\infty]$ is said to be coercive if $\lim_{\|x\| \rightarrow +\infty} \varphi(x) = +\infty$.

Proposition 4.4.13 *Suppose that one of the following holds.*

- (i) *The function $(\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \phi_k(\xi_k) + \Psi(F^*(\xi_k)_{k \in \mathbb{K}})$ is coercive.*
- (ii) *$\inf_{k \in \mathbb{K}} \inf \phi_k(\mathbb{R}) > -\infty$, Ψ is coercive, and $(e_k)_{k \in \mathbb{K}}$ is a Riesz basis.*
- (iii) *$\inf \Psi(\mathcal{H}) > -\infty$ and one of the following properties is satisfied.*
 - (a) *The function $(\xi_k)_{k \in \mathbb{K}} \mapsto \sum_{k \in \mathbb{K}} \phi_k(\xi_k)$ is coercive.*
 - (b) *There exists $\omega \in]0, +\infty[$ and $p \in [1, 2]$ such that $(\forall k \in \mathbb{K}) \phi_k \geq \omega |\cdot|^p$.*
 - (c) *\mathbb{K} is finite and the functions $(\phi_k)_{k \in \mathbb{K}}$ are coercive.*

Then Problem 4.4.11 admits a solution.

Proof. We denote by $\mathbf{x} = (\xi_k)_{k \in \mathbb{K}}$ a generic element in $\ell^2(\mathbb{K})$ and by $\|\mathbf{x}\| = \sqrt{\sum_{k \in \mathbb{K}} |\xi_k|^2}$ its norm. Set

$$f_1: \mathbf{x} \mapsto \sum_{k \in \mathbb{K}} \phi_k(\xi_k) \quad \text{and} \quad f_2 = \Psi \circ F^*. \quad (4.41)$$

First, suppose that (i) holds. Then it follows from the assumptions on $(\phi_k)_{k \in \mathbb{K}}$ in Problem 4.4.11 and Proposition 4.4.10 that $f_1 \in \Gamma_0(\ell^2(\mathbb{K}))$. On the other hand, since Ψ is a finite function in $\Gamma_0(\mathcal{H})$ and $F^*: \ell^2(\mathbb{K}) \rightarrow \mathcal{H}$ is linear and bounded, f_2 is a finite function in $\Gamma_0(\ell^2(\mathbb{K}))$. Altogether, $f_1 + f_2 \in \Gamma_0(\ell^2(\mathbb{K}))$ and the claim follows from [76, Theorem 2.5.1(ii)].

Next, suppose that (ii) holds. In view of (i), since f_1 is bounded below, it is enough to show that f_2 is coercive. Since $(e_k)_{k \in \mathbb{K}}$ is a Riesz basis, we have [30]

$$(\forall \mathbf{x} \in \ell^2(\mathbb{K})) \quad \|F^*\mathbf{x}\| \geq \sqrt{\mu} \|\mathbf{x}\|. \quad (4.42)$$

In turn, the coercivity of Ψ implies that $\lim_{\|\mathbf{x}\| \rightarrow +\infty} \|\Psi(F^*\mathbf{x})\| = +\infty$.

Now, suppose that (iii) holds. In case (iii)(a), since Ψ is bounded below, f_2 is likewise. In turn, the coercivity of f_1 implies that of $f_1 + f_2$, hence the result by (i). Now suppose that (iii)(b) is satisfied and let $\mathbf{x} \in \ell^2(\mathbb{K})$. Then

$$f_1(\mathbf{x}) = \sum_{k \in \mathbb{K}} \phi_k(\xi_k) \geq \omega \sum_{k \in \mathbb{K}} |\xi_k|^p \geq \omega \|\mathbf{x}\|^p. \quad (4.43)$$

Therefore f_1 is coercive and the claim follows from (iii)(a). Finally, suppose that (iii)(c) is satisfied. In view of (iii)(a), it is enough to show that f_1 is coercive. To this end, fix $\rho \in]0, +\infty[$ and recall that $\mathbb{K} = \{1, \dots, K\}$. Let us set $\lambda = \min_{k \in \mathbb{K}} \inf \phi_k(\mathbb{R})$.

Since the functions $(\phi_k)_{k \in \mathbb{K}}$ are coercive and in $\Gamma_0(\mathbb{R})$, it follows from [40, Theorem 2.5.1(ii)] that $\lambda \in \mathbb{R}$. Coercivity also implies that we can find $\delta \in]0, +\infty[$ such that

$$(\forall \xi \in \mathbb{R}) \quad |\xi| \geq \delta/\sqrt{K} \quad \Rightarrow \quad \min_{k \in \mathbb{K}} \phi_k(\xi) \geq \rho + (1 - K)\lambda. \quad (4.44)$$

Now take $\mathbf{x} \in \ell^2(\mathbb{K})$ such that $\|\mathbf{x}\| \geq \delta$ and fix $\ell \in \mathbb{K}$ such that $|\xi_\ell| = \max_{k \in \mathbb{K}} |\xi_k|$. Then $|\xi_\ell| \geq \delta/\sqrt{K}$ and therefore (4.44) yields

$$f_1(\mathbf{x}) = \sum_{k \in \mathbb{K}} \phi_k(\xi_k) \geq \rho + (1 - K)\lambda + \sum_{k \in \mathbb{K} \setminus \{\ell\}} \phi_k(\xi_k) \geq \rho + (1 - K)\lambda + (K - 1)\lambda = \rho, \quad (4.45)$$

which shows that f_1 is coercive. \square

Next, we turn our attention to the characterization of the solutions to Problem 4.4.11.

Proposition 4.4.14 *Let $(\xi_k)_{k \in \mathbb{K}} \in \ell^2(\mathbb{K})$, let $(\eta_k)_{k \in \mathbb{K}} = (F \circ \nabla \Psi \circ F^*)(\xi_k)_{k \in \mathbb{K}}$, and let $\gamma \in]0, +\infty[$. Then $(\xi_k)_{k \in \mathbb{K}}$ solves Problem 4.4.11 if and only if $(\forall k \in \mathbb{K}) \xi_k = \text{prox}_{\gamma \phi_k}(\xi_k - \gamma \eta_k)$.*

Proof. Set $\mathcal{X} = \ell^2(\mathbb{K})$ and let $(\mathbf{o}_k)_{k \in \mathbb{K}}$ denote the canonical orthonormal basis of $\ell^2(\mathbb{K})$. Then (4.40) can be written as

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{minimize}} \quad \sum_{k \in \mathbb{K}} \phi_k(\langle \mathbf{x} \mid \mathbf{o}_k \rangle) + \Psi(F^* \mathbf{x}). \quad (4.46)$$

Now set

$$f_1 = \sum_{k \in \mathbb{K}} \phi_k(\langle \cdot \mid \mathbf{o}_k \rangle) \quad \text{and} \quad f_2 = \Psi \circ F^*. \quad (4.47)$$

Then, in the light of the assumptions on $(\phi_k)_{k \in \mathbb{K}}$ in Problem 4.4.11, Proposition 4.4.10 yields $f_1 \in \Gamma_0(\mathcal{X})$. On the other hand, since Ψ is a finite function in $\Gamma_0(\mathcal{H})$ and $F^*: \mathcal{X} \rightarrow \mathcal{H}$ is linear and bounded, we have $f_2 \in \Gamma_0(\mathcal{X})$. In addition, since $\nabla \Psi$ is Lipschitz continuous, so is $\nabla f_2 = F \circ \nabla \Psi \circ F^*$. Altogether, (4.46) conforms to the format of Problem 4.4.7. Now set $\mathbf{x} = (\xi_k)_{k \in \mathbb{K}}$. Then it follows from Proposition 4.4.8, (4.47), and Proposition 4.4.10 that

$$\begin{aligned} (\xi_k)_{k \in \mathbb{K}} \text{ solves Problem 4.4.11} &\Leftrightarrow \mathbf{x} = \text{prox}_{\gamma f_1}(\mathbf{x} - \gamma \nabla f_2(\mathbf{x})) \\ &\Leftrightarrow \mathbf{x} = \text{prox}_{\gamma f_1}(\mathbf{x} - \gamma(F \circ \nabla \Psi \circ F^*)(\mathbf{x})) \\ &\Leftrightarrow (\xi_k)_{k \in \mathbb{K}} = \text{prox}_{\gamma f_1}(\xi_k - \gamma \eta_k)_{k \in \mathbb{K}} \\ &\Leftrightarrow (\xi_k)_{k \in \mathbb{K}} = \left(\text{prox}_{\gamma \phi_k}(\xi_k - \gamma \eta_k) \right)_{k \in \mathbb{K}}, \end{aligned} \quad (4.48)$$

which provides the desired characterization. \square

4.4.3.3 Specific frameworks

In Problem 4.4.11, the functions $(\phi_k)_{k \in \mathbb{K}}$ penalize the frame coefficients $(\xi_k)_{k \in \mathbb{K}}$, while the function Ψ penalizes $x = F^*(\xi_k)_{k \in \mathbb{K}} = \sum_{k \in \mathbb{K}} \xi_k e_k$, thereby modeling direct constraints on \bar{x} . This flexible framework makes it possible to model a wide range of inverse problems. Two important instances are presented below.

4.4.3.3.1 Inverse problems with sparsity constraints A common objective in selecting the frame $(e_k)_{k \in \mathbb{K}}$ is to obtain a sparse representation of the target solution \bar{x} in the sense that most of the coefficients $(\bar{\xi}_k)_{k \in \mathbb{K}}$ in (4.39) are zero. By choosing $\phi_k = \omega_k |\cdot|$ with $\omega_k > 0$ in Problem 4.4.11, one aims at setting to zero the k th coefficient if it falls into the interval $[-\omega_k, \omega_k]$, hence promoting sparsity (see [17, 24, 38] for special cases). Note that in this case, it follows from Proposition 4.4.14 and (4.30) that a solution $(\xi_k)_{k \in \mathbb{K}}$ to Problem 4.4.11 is characterized by the soft thresholding identities (see also Fig. 4.1)

$$(\forall k \in \mathbb{K}) \quad \xi_k = \text{prox}_{\omega_k |\cdot|}(\xi_k - \eta_k) = \text{soft}_{[-\omega_k, \omega_k]}(\xi_k - \eta_k), \quad (4.49)$$

where $(\eta_k)_{k \in \mathbb{K}} = (F \circ \nabla \Psi \circ F^*)(\xi_k)_{k \in \mathbb{K}}$. More generally, to aim at zeroing a coefficient falling into a closed interval $\Omega_k \subset \mathbb{R}$, one can use the function $\phi_k = \psi_k + \sigma_{\Omega_k}$, where ψ_k satisfies $0 = \psi_k(0) \leq \psi_k \in \Gamma_0(\mathbb{R})$ and is differentiable at 0 [14]. This construct actually characterizes all thresholders on Ω_k that have properties suitable to their use in iterative methods [14, Theorem 3.3]. A decomposition rule for computing the resulting thresholders is supplied in Lemma 4.4.6(iii).

Let us now discuss possible choices for the smooth function Ψ . Suppose that the problem under consideration is to recover $\bar{x} \in \mathcal{H}$ from q observations

$$z_i = T_i \bar{x} + v_i, \quad 1 \leq i \leq q, \quad (4.50)$$

where T_i is a bounded linear operator from \mathcal{H} to a real Hilbert space \mathcal{G}_i , $z_i \in \mathcal{G}_i$, and $v_i \in \mathcal{G}_i$ is the realization of a noise process. A standard data fidelity criterion in such instances is the function $x \mapsto \sum_{i=1}^q \alpha_i \|T_i x - z_i\|^2$, where $(\alpha_i)_{1 \leq i \leq q}$ are strictly positive reals, see e.g., [12, 23]. In addition, assume that a priori information is available that constrains \bar{x} to lie in some closed convex subsets $(S_i)_{1 \leq i \leq m}$ of \mathcal{H} (see [10, 35] and the references therein for examples). These constraints can be aggregated via the cost function $x \mapsto \sum_{i=1}^m \vartheta_i d_{S_i}^2(x)$, where $(\vartheta_i)_{1 \leq i \leq m}$ are strictly positive reals [5, 22]. These two objectives can be combined by using the function

$$\Psi: x \mapsto \frac{1}{2} \sum_{i=1}^q \alpha_i \|T_i x - z_i\|^2 + \frac{1}{2} \sum_{i=1}^m \vartheta_i d_{S_i}^2(x) \quad (4.51)$$

in Problem 4.4.11. This function is indeed differentiable and its gradient

$$\nabla \Psi: x \mapsto \sum_{i=1}^q \alpha_i T_i^*(T_i x - z_i) + \sum_{i=1}^m \vartheta_i (x - P_{S_i} x) \quad (4.52)$$

has Lipschitz constant [14, Section 5.1]

$$\tau = \left\| \sum_{i=1}^q \alpha_i T_i^* T_i \right\| + \sum_{i=1}^m \vartheta_i. \quad (4.53)$$

In instances when $\left\| \sum_{i=1}^q \alpha_i T_i^* T_i \right\|$ cannot be evaluated directly, it can be majorized by $\sum_{i=1}^q \alpha_i \|T_i\|^2$. It should be noted that, more generally, Lemma 4.4.4(iii) implies that Ψ remains Lipschitz continuous if the term $\sum_{i=1}^m \vartheta_i d_{S_i}^2(x)$ in (4.51) is replaced by a sum of Moreau envelopes (see [25, Section 6.3] and [15, Section 4.1] for related frameworks).

4.4.3.3.2 Bayesian statistical framework A standard linear inverse problem is to recover $\bar{x} \in \mathcal{H}$ from an observation

$$z = T\bar{x} + v, \quad (4.54)$$

in a real Hilbert space \mathcal{G} , where $T: \mathcal{H} \rightarrow \mathcal{G}$ is a bounded linear operator and where $v \in \mathcal{G}$ stands for an additive noise perturbation. If $\bar{x} = (\bar{\xi}_k)_{k \in \mathbb{K}}$ denotes the coefficients of \bar{x} in $(e_k)_{k \in \mathbb{K}}$, (4.54) can be written as

$$z = TF^*\bar{x} + v. \quad (4.55)$$

For the sake of simplicity, the following assumptions regarding (4.55) are made in this section (with the usual convention $\ln 0 = -\infty$).

Assumption 4.4.15

- (i) $\mathcal{H} = \mathbb{R}^N$, $\mathcal{G} = \mathbb{R}^M$, and $\mathbb{K} = \{1, \dots, K\}$, where $K \geq N$.
- (ii) The vectors \bar{x} , z , and v are, respectively, realizations of real-valued random vectors \bar{X} , Z , and V defined on the same probability space.
- (iii) The random vectors \bar{X} and V are mutually independent and have probability density functions $f_{\bar{X}}$ and f_V , respectively.
- (iv) The components of \bar{X} are independent with upper-semicontinuous log-concave densities.
- (v) The function $\ln f_V$ is concave and differentiable with a Lipschitz continuous gradient.

Under Assumption 4.4.15, a common Bayesian approach for estimating \bar{x} from z consists in applying a maximum a posteriori (MAP) rule [3, 4, 36], which amounts to maximizing the posterior probability density $f_{\bar{X}|Z=z}$. Thus, \tilde{x} is a MAP estimate of \bar{x} if

$$(\forall \mathbf{x} \in \mathbb{R}^K) \quad f_{\bar{X}|Z=z}(\tilde{x}) \geq f_{\bar{X}|Z=z}(\mathbf{x}). \quad (4.56)$$

Using Bayes' formula, this amounts to solving

$$\underset{x \in \mathbb{R}^K}{\text{minimize}} \quad -\ln f_{\bar{X}}(x) - \ln f_{Z|\bar{X}=x}(z). \quad (4.57)$$

In view of (4.55), this is also equivalent to solving

$$\underset{x \in \mathbb{R}^K}{\text{minimize}} \quad -\ln f_{\bar{X}}(x) - \ln f_V(z - TF^*x). \quad (4.58)$$

Under Assumption 4.4.15, this convex optimization problem is a special case of Problem 4.4.11. Indeed, Assumption 4.4.15(iv) allows us to write, without loss of generality, the prior density as

$$(\forall (\xi_k)_{k \in \mathbb{K}} \in \mathbb{R}^K) \quad f_{\bar{X}}((\xi_k)_{k \in \mathbb{K}}) \propto \prod_{k=1}^K \exp(-\phi_k(\xi_k)), \quad (4.59)$$

where $(\phi_k)_{k \in \mathbb{K}}$ are the so-called potential functions of the marginal probability density functions of \bar{X} . It also follows from Assumption 4.4.15(iv) that the functions $(\phi_k)_{k \in \mathbb{K}}$ are in $\Gamma_0(\mathbb{R})$. Now set

$$(\forall x \in \mathcal{H}) \quad \Psi(x) = -\ln f_V(z - Tx). \quad (4.60)$$

Then Assumption 4.4.15(v) asserts that $\Psi \in \Gamma_0(\mathcal{H})$ is differentiable with a Lipschitz continuous gradient. Altogether, (4.58) reduces to Problem 4.4.11.

Remark 4.4.16 In the simple case when V is a zero-mean Gaussian vector with an invertible covariance matrix Λ , the function Ψ reduces (up to an additive constant) to the residual energy function $x \mapsto \langle \Lambda^{-1}(z - Tx) \mid z - Tx \rangle / 2$. When \bar{X} is further assumed to be Gaussian, the solution to Problem 4.4.11 is a linear function of z . Recall that the MAP estimate coincides with the minimum mean-square error estimate under such Gaussian models for both V and \bar{X} [39, Section 2.4].

Remark 4.4.17 An alternative Bayesian strategy would be to determine a MAP estimate of \bar{x} . This would lead to

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad -\ln f_{\bar{X}}(x) - \ln f_V(z - Tx), \quad (4.61)$$

where $f_{\bar{X}}$ can be deduced from (4.59) through the change of variable $\bar{X} = F^*\bar{X}$. In the case of an orthonormal basis decomposition, it is easy to check that (4.61) is equivalent to problem (4.58). By contrast, when F corresponds to an overcomplete frame, the expression of $f_{\bar{X}}$ becomes involved and (4.61) is usually much less tractable than Problem 4.4.11. As will be seen in Section 4.4.5, the latter can be solved via a simple splitting algorithm.

Remark 4.4.18 Let us decompose the observation vector as $z = [z_1^\top, \dots, z_q^\top]^\top$ and the matrix representing T as $[T_1^\top, \dots, T_q^\top]^\top$ where, for every $i \in \{1, \dots, q\}$, $z_i \in \mathbb{R}^{M_i}$ and $T_i \in \mathbb{R}^{M_i \times N}$ with $\sum_{i=1}^q M_i = M$. Furthermore, assume that V is a zero-mean Gaussian vector with diagonal covariance matrix

$$\Lambda = \begin{bmatrix} \alpha_1^{-1} I_{M_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_q^{-1} I_{M_q} \end{bmatrix}, \quad (4.62)$$

where $(\alpha_i)_{1 \leq i \leq q}$ are strictly positive reals and I_{M_i} , $1 \leq i \leq q$, is the identity matrix of size $M_i \times M_i$. Then Ψ reduces to the first term in (4.51) where $\mathcal{G}_i = \mathbb{R}^{M_i}$ and the MAP estimation problem under Assumption 4.4.15 becomes a special case of the problem addressed in Section 4.4.3.3.1 with $m = 0$.

4.4.4 Proximity operators associated with log-concave densities

As discussed in Section 4.4.3.3.2, the functions $(\phi_k)_{k \in \mathbb{K}}$ in (4.40) act as the potential functions of log-concave univariate probability densities modeling the frame coefficients individually in Bayesian formulations. On the other hand, the proximity operators of such functions will, via Proposition 4.4.10, play a central role in Section 4.4.5. Hereafter, we derive closed-form expressions for these proximity operators in the case of some classical log-concave univariate probability densities [20, Chapters VII&IX].

Let us start with a few observations.

Remark 4.4.19 Let $\phi \in \Gamma_0(\mathbb{R})$.

- (i) It follows from Definition 4.4.1 that $(\forall \xi \in \mathbb{R}) \phi(\text{prox}_\phi \xi) < +\infty$.
- (ii) If ϕ is even, then it follows from Lemma 4.4.5(iv) that prox_ϕ is odd. Therefore, in such instances, it will be enough to determine $\text{prox}_\phi \xi$ for $\xi \geq 0$ and to extend the result to $\xi < 0$ by antisymmetry.
- (iii) Let $\xi \in \mathbb{R}$. If ϕ is differentiable at $\text{prox}_\phi \xi$, then (4.26) yields

$$(\forall \pi \in \mathbb{R}) \quad \pi = \text{prox}_\phi \xi \quad \Leftrightarrow \quad \pi + \phi'(\pi) = \xi. \quad (4.63)$$

We now examine some concrete examples.

Example 4.4.20 (Laplace distribution) Let $\omega \in]0, +\infty[$ and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \omega|\xi|. \quad (4.64)$$

Then, for every $\xi \in \mathbb{R}$, $\text{prox}_\phi \xi = \text{soft}_{[-\omega, \omega]} \xi = \text{sign}(\xi) \max\{|\xi| - \omega, 0\}$.

Proof. Apply Lemma 4.4.6(iii) with $\psi = 0$ and $\Omega = [-\omega, \omega]$. \square

Example 4.4.21 (Gaussian distribution) Let $\tau \in]0, +\infty[$ and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \tau|\xi|^2. \quad (4.65)$$

Then, for every $\xi \in \mathbb{R}$, $\text{prox}_\phi \xi = \xi/(2\tau + 1)$.

Proof. Apply Lemma 4.4.5(i) with $\mathcal{X} = \mathbb{R}$, $\varphi = 0$, $\alpha = 2\tau$, and $\mathbf{u} = 0$. \square

Example 4.4.22 (generalized Gaussian distribution) Let $p \in]1, +\infty[$, $\kappa \in]0, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \kappa|\xi|^p. \quad (4.66)$$

Then, for every $\xi \in \mathbb{R}$, $\text{prox}_\phi \xi = \text{sign}(\xi)\varrho$ where ϱ is the unique solution in $[0, +\infty[$ to

$$\varrho + p\kappa\varrho^{p-1} = |\xi|. \quad (4.67)$$

In particular, the following hold :

- (i) $\text{prox}_\phi \xi = \xi + \frac{4\kappa}{3 \cdot 2^{1/3}} \left((\chi - \xi)^{1/3} - (\chi + \xi)^{1/3} \right)$, where $\chi = \sqrt{\xi^2 + 256\kappa^3/729}$, if $p = 4/3$;
- (ii) $\text{prox}_\phi \xi = \xi + 9\kappa^2 \text{sign}(\xi) (1 - \sqrt{1 + 16|\xi|/(9\kappa^2)})/8$, if $p = 3/2$;
- (iii) $\text{prox}_\phi \xi = \text{sign}(\xi) (\sqrt{1 + 12\kappa|\xi|} - 1)/(6\kappa)$, if $p = 3$;
- (iv) $\text{prox}_\phi \xi = \left(\frac{\chi + \xi}{8\kappa} \right)^{1/3} - \left(\frac{\chi - \xi}{8\kappa} \right)^{1/3}$, where $\chi = \sqrt{\xi^2 + 1/(27\kappa)}$, if $p = 4$.

Proof. Let $\xi \in \mathbb{R}$ and set $\pi = \text{prox}_\phi \xi$. As seen in Remark 4.4.19(ii), because ϕ is even, it is enough to assume that $\xi \geq 0$. Since ϕ is differentiable, it follows from (4.28) and (4.63) that π is the unique solution in $[0, +\infty[$ to

$$\pi + p\kappa\pi^{p-1} = \xi, \quad (4.68)$$

which provides (4.67). For $p = 3$, π is the solution in $[0, +\infty[$ to the equation $\pi + 3\kappa\pi^2 - \xi = 0$, i.e., $\pi = (\sqrt{1 + 12\kappa\xi} - 1)/(6\kappa)$ and we obtain (iii) by antisymmetry. In turn, since $(2|\cdot|^{3/2}/3)^* = |\cdot|^{3/3}$, Lemma 4.4.3(iii) with $\gamma = 3\kappa/2$ yields $\pi = \text{prox}_{\gamma(2|\cdot|^{3/2}/3)} \xi = \xi - \gamma \text{prox}_{(3\gamma)^{-1}|\cdot|^{3/2}}(\xi/\gamma) = \xi + 9\kappa^2 \text{sign}(\xi) (1 - \sqrt{1 + 16|\xi|/(9\kappa^2)})/8$, which proves (ii). Now, let $p = 4$. Then (4.68) asserts that π is the unique solution in $[0, +\infty[$ to the third degree equation $4\kappa\pi^3 + \pi - \xi = 0$, namely $\pi = (\sqrt{\alpha^2 + \beta^3} - \alpha)^{1/3} - (\sqrt{\alpha^2 + \beta^3} + \alpha)^{1/3}$, where $\alpha = -\xi/(8\kappa)$ and $\beta = 1/(12\kappa)$. Since this expression is an odd function of ξ , we obtain (iv). Finally, we deduce (i) from (iv) by observing that, since $(3|\cdot|^{4/3}/4)^* = |\cdot|^{4/4}$, Lemma 4.4.3(iii) with $\gamma = 4\kappa/3$ yields $\pi = \text{prox}_{\gamma(3|\cdot|^{4/3}/4)} \xi = \xi - \gamma \text{prox}_{(4\gamma)^{-1}|\cdot|^{4/3}}(\xi/\gamma)$, hence the result after simple algebra. \square

Example 4.4.23 (Huber distribution) Let $\omega \in]0, +\infty[$, $\tau \in]0, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \begin{cases} \tau\xi^2, & \text{if } |\xi| \leq \omega/\sqrt{2\tau}; \\ \omega\sqrt{2\tau}|\xi| - \omega^2/2, & \text{otherwise.} \end{cases} \quad (4.69)$$

Then, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \begin{cases} \frac{\xi}{2\tau + 1}, & \text{if } |\xi| \leq \omega(2\tau + 1)/\sqrt{2\tau}; \\ \xi - \omega\sqrt{2\tau} \text{sign}(\xi), & \text{if } |\xi| > \omega(2\tau + 1)/\sqrt{2\tau}. \end{cases} \quad (4.70)$$

Proof. Let $\xi \in \mathbb{R}$ and set $\pi = \text{prox}_\phi \xi$. Since ϕ is even, we assume that $\xi \geq 0$ (see Remark 4.4.19(ii)). In addition, since ϕ is differentiable, it follows from (4.28) and (4.63) that π is the unique solution in $[0, \xi]$ to $\pi + \phi'(\pi) = \xi$. First, suppose that $\pi = \omega/\sqrt{2\tau}$. Then $\phi'(\pi) = \omega\sqrt{2\tau}$ and, therefore, $\xi = \pi + \phi'(\pi) = \omega(2\tau + 1)/\sqrt{2\tau}$. Now, suppose that $\xi \leq \omega(2\tau + 1)/\sqrt{2\tau}$. Then it follows from Lemma 4.4.6(i) that $\pi \leq \text{prox}_\phi(\omega(2\tau + 1)/\sqrt{2\tau}) = \omega/\sqrt{2\tau}$. In turn, (4.69) yields $\phi'(\pi) = 2\tau\pi$ and the identity $\xi = \pi + \phi'(\pi)$ yields $\pi = \xi/(2\tau + 1)$. Finally, if $\xi > \omega(2\tau + 1)/\sqrt{2\tau}$, then Lemma 4.4.6(i) yields $\pi \geq \text{prox}_\phi(\omega(2\tau + 1)/\sqrt{2\tau}) = \omega/\sqrt{2\tau}$ and, in turn, $\phi'(\pi) = \omega\sqrt{2\tau}$, which allows us to conclude that $\pi = \xi - \omega\sqrt{2\tau}$. \square

Example 4.4.24 (maximum entropy distribution) This density is obtained by maximizing the entropy subject to the knowledge of the first, second, and p -th order absolute moments, where $2 \neq p \in]1, +\infty[$ [28]. Let $\omega \in]0, +\infty[$, $\tau \in [0, +\infty[$, $\kappa \in]0, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \omega|\xi| + \tau|\xi|^2 + \kappa|\xi|^p. \quad (4.71)$$

Then, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \text{sign}(\xi) \text{prox}_{\kappa|\cdot|^p/(2\tau+1)} \left(\frac{1}{2\tau + 1} \max\{|\xi| - \omega, 0\} \right) \quad (4.72)$$

where the expression of $\text{prox}_{\kappa|\cdot|^p/(2\tau+1)}$ is supplied by Example 4.4.22.

Proof. The function ϕ is a quadratic perturbation of the function $\varphi = \omega|\cdot| + \kappa|\cdot|^p$. Applying Lemma 4.4.6(iii) with $\psi = \kappa|\cdot|^p$ and $\Omega = [-\omega, \omega]$, we get ($\forall \xi \in \mathbb{R}$) $\text{prox}_\varphi \xi = \text{prox}_{\kappa|\cdot|^p}(\text{soft}_{[-\omega, \omega]} \xi) = \text{sign}(\xi) \text{prox}_{\kappa|\cdot|^p}(\max\{|\xi| - \omega, 0\})$. Hence, the result follows from Lemma 4.4.5(i) where $\mathcal{X} = \mathbb{R}$, $\alpha = 2\tau$, and $\mathbf{u} = 0$. \square

Example 4.4.25 (smoothed Laplace distribution) Let $\omega \in]0, +\infty[$ and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \omega|\xi| - \ln(1 + \omega|\xi|). \quad (4.73)$$

This potential function is sometimes used as a differentiable approximation to (4.64), e.g., [33]. We have, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \text{sign}(\xi) \frac{\omega|\xi| - \omega^2 - 1 + \sqrt{|\omega|\xi| - \omega^2 - 1|^2 + 4\omega|\xi|}}{2\omega}. \quad (4.74)$$

Proof. According to Remark 4.4.19(ii), since ϕ is even, we can focus on the case when $\xi \geq 0$. As ϕ achieves its infimum at 0, Lemma 4.4.6(ii) yields $\pi = \text{prox}_\phi \xi \geq 0$. We deduce from (4.63) that π is the unique solution in $[0, +\infty[$ to the equation

$$\omega\pi^2 + (\omega^2 + 1 - \omega\xi)\pi - \xi = 0, \quad (4.75)$$

which leads to (4.74). \square

Example 4.4.26 (exponential distribution) Let $\omega \in]0, +\infty[$ and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} \omega\xi, & \text{if } \xi \geq 0; \\ +\infty, & \text{if } \xi < 0. \end{cases} \quad (4.76)$$

Then, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \begin{cases} \xi - \omega & \text{if } \xi \geq \omega; \\ 0 & \text{if } \xi < \omega. \end{cases} \quad (4.77)$$

Proof. Set $\varphi = \iota_{[0, +\infty[}$. Then Example 4.4.2 yields $\text{prox}_\varphi = P_{[0, +\infty[}$. In turn, since ϕ is a linear perturbation of φ , the claim results from Lemma 4.4.5(i), where $\mathcal{X} = \mathbb{R}$, $\alpha = 0$, and $\mathbf{u} = \omega$. \square

Example 4.4.27 (gamma distribution) Let $\omega \in]0, +\infty[$, $\kappa \in]0, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} -\kappa \ln(\xi) + \omega\xi, & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \leq 0. \end{cases} \quad (4.78)$$

Then, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \frac{\xi - \omega + \sqrt{|\xi - \omega|^2 + 4\kappa}}{2}. \quad (4.79)$$

Proof. Set

$$\varphi: \mathbb{R} \rightarrow]-\infty, +\infty] : \xi \mapsto \begin{cases} -\kappa \ln(\xi), & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \leq 0. \end{cases} \quad (4.80)$$

We easily get from Remark 4.4.19(i)&(iii) that

$$(\forall \xi \in \mathbb{R}) \quad \text{prox}_\varphi \xi = \frac{\xi + \sqrt{\xi^2 + 4\kappa}}{2}. \quad (4.81)$$

In turn, since ϕ is a linear perturbation of φ , the claim results from Lemma 4.4.5(i), where $\mathcal{X} = \mathbb{R}$, $\alpha = 0$, and $\mathbf{u} = \omega$. \square

Example 4.4.28 (chi distribution) Let $\kappa \in]0, +\infty[$ and let

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \begin{cases} -\kappa \ln(\xi) + \xi^2/2, & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \leq 0. \end{cases} \quad (4.82)$$

Then, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \frac{\xi + \sqrt{\xi^2 + 8\kappa}}{4}. \quad (4.83)$$

Proof. Since ϕ is a quadratic perturbation of the function φ defined in (4.80), the claim results from Lemma 4.4.5(i), where $\mathcal{X} = \mathbb{R}$, $\alpha = 1$, and $\mathbf{u} = 0$. \square

Example 4.4.29 (uniform distribution) Let $\omega \in]0, +\infty[$ and set $\phi = \iota_{[-\omega, \omega]}$. Then it follows at once from Example 4.4.2 that, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = P_{[-\omega, \omega]} \xi = \begin{cases} -\omega, & \text{if } \xi < -\omega; \\ \xi, & \text{if } |\xi| \leq \omega; \\ \omega, & \text{if } \xi > \omega. \end{cases} \quad (4.84)$$

Example 4.4.30 (triangular distribution) Let $\underline{\omega} \in]-\infty, 0[$, let $\bar{\omega} \in]0, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \begin{cases} -\ln(\xi - \underline{\omega}) + \ln(-\underline{\omega}), & \text{if } \xi \in]\underline{\omega}, 0]; \\ -\ln(\bar{\omega} - \xi) + \ln(\bar{\omega}), & \text{if } \xi \in]0, \bar{\omega}[; \\ +\infty, & \text{otherwise.} \end{cases} \quad (4.85)$$

Then, for every $\xi \in \mathbb{R}$,

$$\text{prox}_\phi \xi = \begin{cases} \frac{\xi + \underline{\omega} + \sqrt{|\xi - \underline{\omega}|^2 + 4}}{2}, & \text{if } \xi < 1/\underline{\omega}; \\ \frac{\xi + \bar{\omega} - \sqrt{|\xi - \bar{\omega}|^2 + 4}}{2}, & \text{if } \xi > 1/\bar{\omega}; \\ 0 & \text{otherwise.} \end{cases} \quad (4.86)$$

Proof. Let $\xi \in \mathbb{R}$ and set $\pi = \text{prox}_\phi \xi$. Let us first note that $\partial\phi(0) = [1/\underline{\omega}, 1/\overline{\omega}]$. Therefore, (4.26) yields

$$\pi = 0 \quad \Leftrightarrow \quad \xi \in [1/\underline{\omega}, 1/\overline{\omega}]. \quad (4.87)$$

Now consider the case when $\xi > 1/\overline{\omega}$. Since ϕ admits 0 as a minimizer, it follows from Lemma 4.4.6(ii) and (4.87) that $\pi \in]0, \xi]$. Hence, we derive from (4.63) that π is the only solution in $]0, \xi]$ to $\pi + 1/(\overline{\omega} - \pi) = \xi$, i.e., $\pi = (\xi + \overline{\omega} - \sqrt{|\xi - \overline{\omega}|^2 + 4})/2$. Likewise, if $\xi < 1/\underline{\omega}$, it follows from Lemma 4.4.6(ii), (4.87), and (4.63) that π is the only solution in $[\xi, 0[$ to $\pi - 1/(\pi - \underline{\omega}) = \xi$, which yields $\pi = (\xi + \underline{\omega} + \sqrt{|\xi - \underline{\omega}|^2 + 4})/2$. \square

The next example is an extension of Example 4.4.28.

Example 4.4.31 (Weibull distribution) Let $\omega \in]0, +\infty[$, $\kappa \in]0, +\infty[$, and $p \in]1, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \begin{cases} -\kappa \ln(\xi) + \omega \xi^p, & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \leq 0. \end{cases} \quad (4.88)$$

Then, for every $\xi \in \mathbb{R}$, $\pi = \text{prox}_\phi \xi$ is the unique strictly positive solution to

$$p\omega\pi^p + \pi^2 - \xi\pi = \kappa. \quad (4.89)$$

Proof. Since ϕ is differentiable on $]0, +\infty[$, it follows from Remark 4.4.19(i)&(iii) that π is the unique solution in $]0, +\infty[$ to $\pi + \phi'(\pi) = \xi$ or, equivalently, to (4.89). \square

A similar proof can be used in the following two examples.

Example 4.4.32 (generalized inverse Gaussian) Let $\omega \in]0, +\infty[$, $\kappa \in]0, +\infty[$, and $\rho \in]0, +\infty[$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \begin{cases} -\kappa \ln(\xi) + \omega \xi + \rho/\xi, & \text{if } \xi > 0; \\ +\infty, & \text{if } \xi \leq 0. \end{cases} \quad (4.90)$$

Then, for every $\xi \in \mathbb{R}$, $\pi = \text{prox}_\phi \xi$ is the unique strictly positive solution to

$$\pi^3 + (\omega - \xi)\pi^2 - \kappa\pi = \rho. \quad (4.91)$$

Example 4.4.33 (Pearson type I) Let $\underline{\kappa}$ and $\overline{\kappa}$ be in $]0, +\infty[$, let $\underline{\omega}$ and $\overline{\omega}$ be reals such that $\underline{\omega} < \overline{\omega}$, and set

$$\phi: \mathbb{R} \rightarrow]-\infty, +\infty]: \xi \mapsto \begin{cases} -\underline{\kappa} \ln(\xi - \underline{\omega}) - \overline{\kappa} \ln(\overline{\omega} - \xi), & \text{if } \xi \in]\underline{\omega}, \overline{\omega}[; \\ +\infty, & \text{otherwise.} \end{cases} \quad (4.92)$$

Then, for every $\xi \in \mathbb{R}$, $\pi = \text{prox}_\phi \xi$ is the unique solution in $]\underline{\omega}, \overline{\omega}[$ to

$$\pi^3 - (\underline{\omega} + \overline{\omega} + \xi)\pi^2 + (\underline{\omega}\overline{\omega} - \underline{\kappa} - \overline{\kappa} + (\underline{\omega} + \overline{\omega})\xi)\pi = \underline{\omega}\overline{\omega}\xi - \underline{\omega}\overline{\kappa} - \overline{\omega}\underline{\kappa}. \quad (4.93)$$

Remark 4.4.34

- (i) The chi-square distribution with $n > 2$ degrees of freedom is a special case of the gamma distribution (Example 4.4.27) with $(\omega, \kappa) = (1/2, n/2 - 1)$.
- (ii) The normalized Rayleigh distribution is a special case of the chi distribution (Example 4.4.28) with $\kappa = 1$.
- (iii) The beta distribution and the Wigner distribution are special cases of the Pearson type I distribution (Example 4.4.33) with $(\underline{\omega}, \bar{\omega}) = (0, 1)$, and $-\underline{\omega} = \bar{\omega}$ and $\underline{\kappa} = \bar{\kappa} = 1/2$, respectively.
- (iv) The proximity operator associated with translated and/or scaled versions of the above densities can be obtained via Lemma 4.4.5(ii)&(iii).
- (v) For log-concave densities for which the proximity operator of the potential function is difficult to express in closed form (e.g., Kumaraswamy or logarithmic distributions), one can turn to simple procedures to solve (4.26) or (4.63) numerically.

4.4.5 Algorithm

We propose the following algorithm to solve Problem 4.4.11.

Algorithm 4.4.35 Fix $\mathbf{x}_0 \in \ell^2(\mathbb{K})$ and construct a sequence $(\mathbf{x}_n)_{n \in \mathbb{N}} = ((\xi_{n,k})_{k \in \mathbb{K}})_{n \in \mathbb{N}}$ by setting, for every $n \in \mathbb{N}$,

$$(\forall k \in \mathbb{K}) \quad \xi_{n+1,k} = \xi_{n,k} + \lambda_n \left(\text{prox}_{\gamma_n \phi_k} \left(\xi_{n,k} - \gamma_n (\eta_{n,k} + \beta_{n,k}) \right) + \alpha_{n,k} - \xi_{n,k} \right), \quad (4.94)$$

where $\lambda_n \in]0, 1]$, $\gamma_n \in]0, +\infty[$, $\{\alpha_{n,k}\}_{k \in \mathbb{K}} \subset \mathbb{R}$, $(\eta_{n,k})_{k \in \mathbb{K}} = F(\nabla \Psi(F^* \mathbf{x}_n))$, and $(\beta_{n,k})_{k \in \mathbb{K}} = Fb_n$, where $b_n \in \mathcal{H}$.

The chief advantage of this algorithm is to be fully split in the sense that the functions $(\phi_k)_{k \in \mathbb{K}}$ and Ψ appearing in (4.40) are used separately. First, the current iterate \mathbf{x}_n is transformed into a point in $F^* \mathbf{x}_n$ in \mathcal{H} , and the gradient of Ψ is evaluated at this point to within some tolerance b_n . Next, we obtain the sequence $(\eta_{n,k})_{k \in \mathbb{K}} = F(\nabla \Psi(F^* \mathbf{x}_n))$ to within some tolerance $(\beta_{n,k})_{k \in \mathbb{K}} = Fb_n$. Then one chooses $\gamma_n > 0$, and, for every $k \in \mathbb{K}$, applies the operator $\text{prox}_{\gamma_n \phi_k}$ to $\xi_{n,k} - \gamma_n (\eta_{n,k} + \beta_{n,k})$. An error $\alpha_{n,k}$ is tolerated in this computation. Finally, the k th component $\xi_{n+1,k}$ of \mathbf{x}_{n+1} is obtained by applying a relaxation of parameter λ_n to this inexact proximal step. Let us note that the computation of the proximal steps can be performed in parallel.

To study the asymptotic behavior of the sequences generated by Algorithm 4.4.35, we require the following set of assumptions.

Assumption 4.4.36 In addition to the standing assumptions of Problem 4.4.11, the following hold.

- (i) Problem 4.4.11 admits a solution.
- (ii) $\inf_{n \in \mathbb{N}} \lambda_n > 0$.
- (iii) $\inf_{n \in \mathbb{N}} \gamma_n > 0$ and $\sup_{n \in \mathbb{N}} \gamma_n < 2/\beta$, where β is a Lipschitz constant of $F \circ \nabla \Psi \circ F^*$.
- (iv) $\sum_{n \in \mathbb{N}} \sqrt{\sum_{k \in \mathbb{K}} |\alpha_{n,k}|^2} < +\infty$ and $\sum_{n \in \mathbb{N}} \|b_n\| < +\infty$.

Remark 4.4.37 As regards Assumption 4.4.36(i), sufficient conditions can be found in Proposition 4.4.13. Let us now turn to the parameter β in Assumption 4.4.36(iii), which determines the range of the step sizes $(\gamma_n)_{n \in \mathbb{N}}$. It follows from the assumptions of Problem 4.4.11 and (4.18) that, for every \mathbf{x} and \mathbf{y} in $\ell^2(\mathbb{K})$,

$$\begin{aligned}
\|F(\nabla \Psi(F^* \mathbf{x})) - F(\nabla \Psi(F^* \mathbf{y}))\| &\leq \|F\| \|\nabla \Psi(F^* \mathbf{x}) - \nabla \Psi(F^* \mathbf{y})\| \\
&\leq \tau \|F\| \|F^* \mathbf{x} - F^* \mathbf{y}\| \\
&\leq \tau \|F\|^2 \|\mathbf{x} - \mathbf{y}\| \\
&\leq \tau \nu \|\mathbf{x} - \mathbf{y}\|.
\end{aligned} \tag{4.95}$$

Thus, the value $\beta = \tau \nu$ can be used in general. In some cases, however, a sharper bound can be obtained, which results in a wider range for the step sizes $(\gamma_n)_{n \in \mathbb{N}}$. For example, in the problem considered in Section 4.4.3.3.1, if the norm of $R = \sum_{i=1}^q \alpha_i F T_i^* T_i F^*$ can be evaluated, it follows from (4.51) and the nonexpansivity of the operators $(\text{Id} - P_{S_i})_{1 \leq i \leq m}$ that one can take

$$\beta = \|R\| + \nu \sum_{i=1}^m \vartheta_i. \tag{4.96}$$

Theorem 4.4.38 *Let $(\mathbf{x}_n)_{n \in \mathbb{N}}$ be an arbitrary sequence generated by Algorithm 4.4.35 under Assumption 4.4.36. Then $(\mathbf{x}_n)_{n \in \mathbb{N}}$ converges weakly to a solution to Problem 4.4.11.*

Proof. Set $\mathcal{X} = \ell^2(\mathbb{K})$, $f_1 = \sum_{k \in \mathbb{K}} \phi_k(\langle \cdot | \mathbf{o}_k \rangle)$, and $f_2 = \Psi \circ F^*$, where $(\mathbf{o}_k)_{k \in \mathbb{K}}$ denotes the canonical orthonormal basis of $\ell^2(\mathbb{K})$. Then $\nabla f_2 = F \circ \nabla \Psi \circ F^*$ is β -Lipschitz continuous (see Assumption 4.4.36(iii)) and, as seen in the proof of Proposition 4.4.14, (4.40) conforms to the format of Problem 4.4.7. Furthermore, it follows from Proposition 4.4.10 that we can rewrite (4.94) as

$$\begin{aligned}
\mathbf{x}_{n+1} &= \mathbf{x}_n + \lambda_n \left(\sum_{k \in \mathbb{K}} (\text{prox}_{\gamma_n \phi_k} \langle \mathbf{x}_n - \gamma_n F(\nabla \Psi(F^* \mathbf{x}_n)) + b_n | \mathbf{o}_k \rangle + \alpha_{n,k}) \mathbf{o}_k - \mathbf{x}_n \right) \\
&= \mathbf{x}_n + \lambda_n \left(\text{prox}_{\gamma_n f_1} (\mathbf{x}_n - \gamma_n (\nabla f_2(\mathbf{x}_n) + \mathbf{b}_n)) + \mathbf{a}_n - \mathbf{x}_n \right),
\end{aligned} \tag{4.97}$$

where $\mathbf{a}_n = (\alpha_{n,k})_{k \in \mathbb{K}}$ and $\mathbf{b}_n = F b_n$. Since Assumption 4.4.36(iv) and (4.18) imply that $\sum_{n \in \mathbb{N}} \|\mathbf{a}_n\| < +\infty$ and $\sum_{n \in \mathbb{N}} \|\mathbf{b}_n\| \leq \sqrt{\nu} \sum_{n \in \mathbb{N}} \|b_n\| < +\infty$, the claim therefore follows from Theorem 4.4.9. \square

Let $(x_n)_{n \in \mathbb{N}}$ be a sequence generated by Algorithm 4.4.35 under Assumption 4.4.36 and set $(\forall n \in \mathbb{N}) x_n = F^*x_n$. On the one hand, Theorem 4.4.38 asserts that $(x_n)_{n \in \mathbb{N}}$ converges weakly to a solution x to Problem 4.4.11. On the other hand, since F^* is linear and bounded, it is weakly continuous and, therefore, $(x_n)_{n \in \mathbb{N}}$ converges weakly to F^*x . However, it is not possible to express (4.94) as an iteration in terms of the sequence $(x_n)_{n \in \mathbb{N}}$ in \mathcal{H} in general. The following corollary addresses the case when F is surjective, which does lead to an algorithm in \mathcal{H} .

Corollary 4.4.39 *Suppose that $(e_k)_{k \in \mathbb{K}}$ is a Riesz basis with companion biorthogonal basis $(\check{e}_k)_{k \in \mathbb{K}}$. Fix $x_0 \in \mathcal{H}$ and, for every $n \in \mathbb{N}$, set*

$$x_{n+1} = x_n + \lambda_n \left(\sum_{k \in \mathbb{K}} \left(\text{prox}_{\gamma_n \phi_k} \left(\langle x_n | \check{e}_k \rangle - \gamma_n \langle \nabla \Psi(x_n) + b_n | e_k \rangle \right) + \alpha_{n,k} \right) e_k - x_n \right), \quad (4.98)$$

where $\lambda_n \in]0, 1]$, $\gamma_n \in]0, +\infty[$, $\{\alpha_{n,k}\}_{k \in \mathbb{K}} \subset \mathbb{R}$, and $b_n \in \mathcal{H}$. Suppose that Assumption 4.4.36 is in force. Then $(x_n)_{n \in \mathbb{N}}$ converges weakly to a point $x \in \mathcal{H}$ and $(\langle x | \check{e}_k \rangle)_{k \in \mathbb{K}}$ is a solution to Problem 4.4.11.

Proof. Set $(\forall n \in \mathbb{N})(\forall k \in \mathbb{K}) \xi_{n,k} = \langle x_n | \check{e}_k \rangle$, $\eta_{n,k} = \langle \nabla \Psi(x_n) | e_k \rangle$, and $\beta_{n,k} = \langle b_n | e_k \rangle$. Then, for every $n \in \mathbb{N}$, it follows from (4.21) that $x_n = F^*(\xi_{n,k})_{k \in \mathbb{K}}$ and, in turn, that

$$(\eta_{n,k})_{k \in \mathbb{K}} = F(\nabla \Psi(F^*x_n)), \quad \text{where } x_n = (\xi_{n,k})_{k \in \mathbb{K}}. \quad (4.99)$$

Furthermore, for every $n \in \mathbb{N}$, it follows from (4.98) and the biorthogonality of $(e_k)_{k \in \mathbb{K}}$ and $(\check{e}_k)_{k \in \mathbb{K}}$ that

$$\begin{aligned} (\forall k \in \mathbb{K}) \quad \xi_{n+1,k} &= \langle x_{n+1} | \check{e}_k \rangle \\ &= \langle x_n | \check{e}_k \rangle + \lambda_n \left(\text{prox}_{\gamma_n \phi_k} \left(\langle x_n | \check{e}_k \rangle \right. \right. \\ &\quad \left. \left. - \gamma_n \langle \nabla \Psi(x_n) + b_n | e_k \rangle \right) + \alpha_{n,k} - \langle x_n | \check{e}_k \rangle \right) \\ &= \xi_{n,k} + \lambda_n \left(\text{prox}_{\gamma_n \phi_k} \left(\xi_{n,k} - \gamma_n (\eta_{n,k} + \beta_{n,k}) \right) + \alpha_{n,k} - \xi_{n,k} \right). \end{aligned} \quad (4.100)$$

Since Theorem 4.4.38 states that $(x_n)_{n \in \mathbb{N}}$ converges weakly to a solution x to Problem 4.4.11, $(x_n)_{n \in \mathbb{N}} = (F^*x_n)_{n \in \mathbb{N}}$ converges weakly to $x = F^*x$. Consequently, (4.21) asserts that we can write $x = (\langle x | \check{e}_k \rangle)_{k \in \mathbb{K}}$. \square

Remark 4.4.40 Suppose that $\mathbb{K} = \mathbb{N}$ and that $(e_k)_{k \in \mathbb{K}}$ is an orthonormal basis of \mathcal{H} . Then (4.98) reduces to

$$x_{n+1} = x_n + \lambda_n \left(\sum_{k \in \mathbb{K}} \left(\text{prox}_{\gamma_n \phi_k} \left(\langle x_n - \gamma_n (\nabla \Psi(x_n) + b_n) | e_k \rangle \right) + \alpha_{n,k} \right) e_k - x_n \right).$$

$$(4.101)$$

In this particular setting, some results related to Corollary 4.4.39 are the following.

- (i) Suppose that $\Psi: x \mapsto \|Tx - z\|^2/2$, where T is a nonzero bounded linear operator from \mathcal{H} to a real Hilbert space \mathcal{G} and $z \in \mathcal{G}$. Suppose that, in addition, $(\forall k \in \mathbb{K}) \phi_k \geq \phi_k(0) = 0$. Then the convergence of (4.101) is discussed in [15, Corollary 5.16].
- (ii) Suppose that $(\Omega_k)_{k \in \mathbb{K}}$ are closed intervals of \mathbb{R} such that $0 \in \text{int} \bigcap_{k \in \mathbb{K}} \Omega_k$ and that

$$(\forall k \in \mathbb{K}) \quad \phi_k = \psi_k + \sigma_{\Omega_k}, \quad (4.102)$$

where $\psi_k \in \Gamma_0(\mathbb{R})$ is differentiable at 0 and $\psi_k \geq \psi_k(0) = 0$. Then (4.101) is the thresholding algorithm proposed and analyzed in [14], namely

$$x_{n+1} = x_n + \lambda_n \left(\sum_{k \in \mathbb{K}} \left(\text{prox}_{\gamma_n \psi_k} \left(\text{soft}_{\gamma_n \Omega_k} \langle x_n - \gamma_n (\nabla \Psi(x_n) + b_n) | e_k \rangle \right) + \alpha_{n,k} \right) e_k - x_n \right), \quad (4.103)$$

where $\text{soft}_{\gamma_n \Omega_k}$ is defined in (4.29).

- (iii) Suppose that the assumptions of both (i) and (ii) hold and that, in addition, we set $\lambda_n \equiv 1$, $\|T\| < 1$, $\gamma_n \equiv 1$, $\alpha_{n,k} \equiv 0$, $b_n \equiv 0$, and $(\forall k \in \mathbb{K}) \psi_k = 0$ and $\Omega_k = [-\omega_k, \omega_k]$. Then (4.103) becomes

$$x_{n+1} = \sum_{k \in \mathbb{K}} \left(\text{soft}_{\Omega_k} \langle x_n + T^*(z - Tx_n) | e_k \rangle \right) e_k. \quad (4.104)$$

Algorithm 4.4.35 can be regarded as a descendant of this original method, which is investigated in [17] and [19].

4.4.6 Numerical results

The proposed framework is applicable to a wide array of variational formulations for inverse problems over frames. We provide a couple of examples to illustrate its applicability in wavelet-based image restoration in the Euclidean space $\mathcal{H} = \mathbb{R}^{512 \times 512}$. The choice of the potential functions $(\phi_k)_{k \in \mathbb{K}}$ in Problem 4.4.11 is guided by the observation that regular images typically possess sparse wavelet representations and that the resulting wavelet coefficients often have even probability density functions [30]. Among the candidate potential functions investigated in Section 4.4.4, those of Example 4.4.24 appear to be the most appropriate for modeling wavelet coefficients

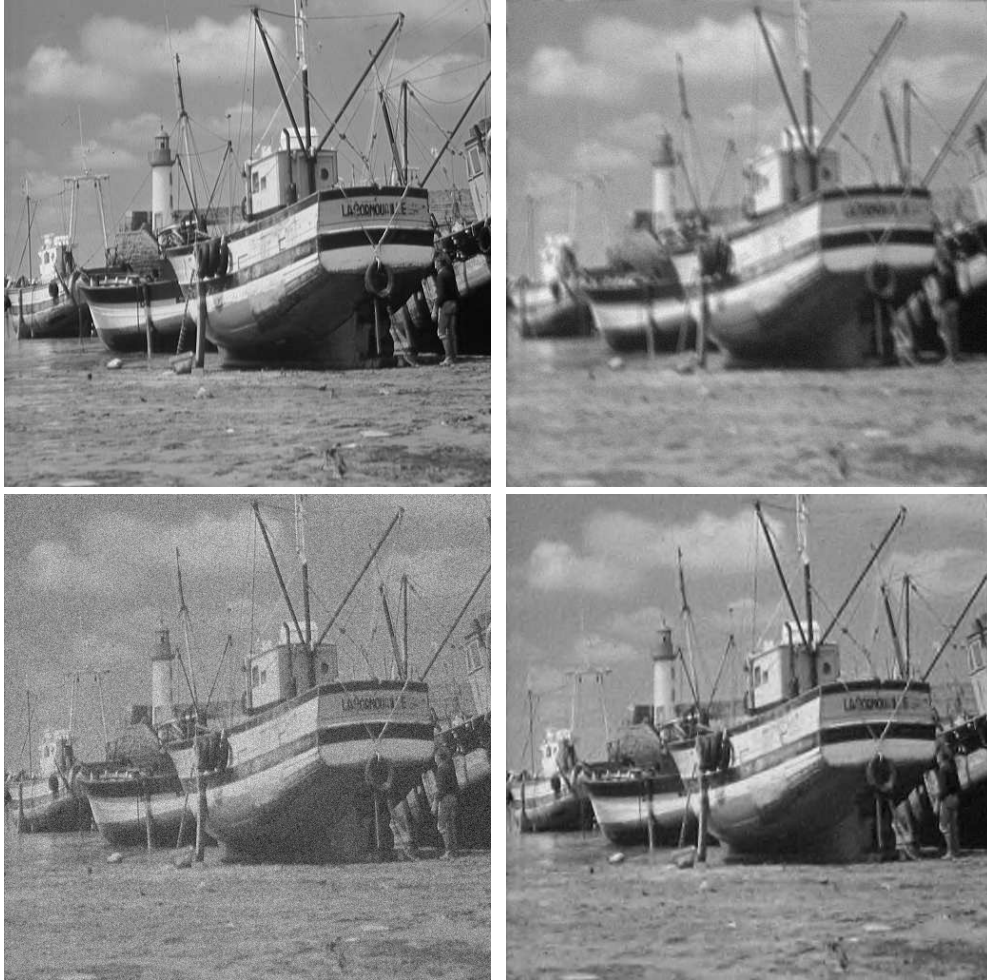


FIG. 4.2 – Example 1 – Original image (top left), first observation (top right), second observation (bottom left), and image restored with 200 iterations of Algorithm 4.4.35 (bottom right).

on two counts. First, they provide flexible models of even potentials. Second, as shown in Lemma 4.4.6(iii), their proximity operators are thresholders and they therefore promote sparsity. More precisely, we employ potential functions of the form $\phi_k = \omega_k |\cdot| + \tau_k |\cdot|^2 + \kappa_k |\cdot|^{p_k}$, where $p_k \in \{4/3, 3/2, 3, 4\}$ and $\{\omega_k, \tau_k, \kappa_k\} \subset]0, +\infty[$. Note that prox_{ϕ_k} can be obtained explicitly via (4.72) and Examples 4.4.22(i)-(iv). In addition, it follows from Proposition 4.4.13(iii)(b) that, with such potential functions, Problem 4.4.11 does admit a solution. The values of the parameters ω_k , τ_k , κ_k , and p_k are chosen for each wavelet subband via a maximum likelihood approach. The first example uses a biorthogonal wavelet basis and the second one uses an M -band dual-tree wavelet frame. Let us emphasize that such decompositions cannot be dealt with using the methods developed in [14], which are limited to orthonormal basis representations. Algorithm 4.4.35 is implemented with $\lambda_n \equiv 1$ and large step sizes (i.e., γ_n close to $2/\beta$) since such values have been observed to provide a good speed of convergence in our experiments.

4.4.6.1 Example 1

We provide a multiview restoration example in a biorthogonal wavelet basis. The original image \bar{x} is the standard test image displayed in Fig. 4.2 (top left). Two observations (see Fig. 4.2 top right and bottom left) conforming to the model (4.50) are available. In our experiment, $\mathcal{G}_1 = \mathcal{G}_2 = \mathcal{H}$ and v_1 and v_2 are realizations of two independent zero-mean Gaussian white noise processes. Moreover, the operator T_1 models a motion blur in the diagonal direction and satisfies $\|T_1\| = 1$, whereas $T_2 = \text{Id}/2$. The blurred image-to-noise ratio is higher for the first observation (22.79 dB versus 15.18 dB) and so is the relative error (18.53 dB versus 5.891 dB) (the decibel value of the relative error between an image z and \bar{x} is $20 \log_{10} (\|\bar{x}\|/\|z - \bar{x}\|)$). The function Ψ in Problem 4.4.11 is given by (4.51), where $\alpha_1 = 4.00 \times 10^{-2}$ and $\alpha_2 = 6.94 \times 10^{-3}$ are the inverses of the variances of the noise corrupting each observation. In addition, we set $m = 1$, $\vartheta_1 = 10^{-2}$, and $S_1 = [0, 255]^{512 \times 512}$ to enforce the known range of the pixel values. A discrete biorthogonal spline 9-7 decomposition [2] is used over 3 resolution levels. Algorithm 4.4.35 is used to solve Problem 4.4.11. By numerically evaluating $\|R\|$ in (4.96), we obtain $\beta = 0.230$ and the step sizes are chosen to be $\gamma_n \equiv 1.99/\beta = 8.66$. The resulting restored image, shown in Fig. 4.2 (bottom right), yields a relative error of 23.84 dB.

4.4.6.2 Example 2

The original SPOT5 satellite image \bar{x} is shown in Fig. 4.3 (top) and the degraded image z in $\mathcal{G} = \mathcal{H}$ is shown in Fig. 4.3 (center). The degradation model is given by (4.54), where T is a 7×7 uniform blur with $\|T\| = 1$, and where v is a realization of a zero-mean Gaussian white noise process. The blurred image-to-noise ratio is 28.08 dB and the relative error is 12.49 dB.

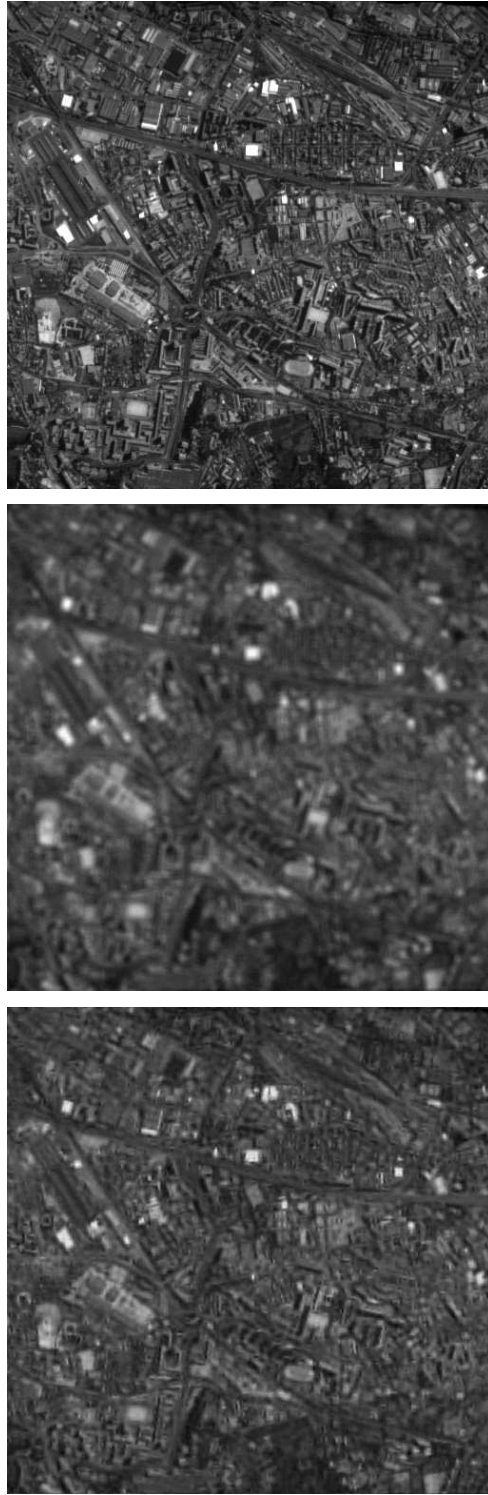


FIG. 4.3 – Example 2 – Original image (top); degraded image (center); image restored in a dual-tree wavelet frame with 100 iterations of Algorithm 4.4.35 (bottom).

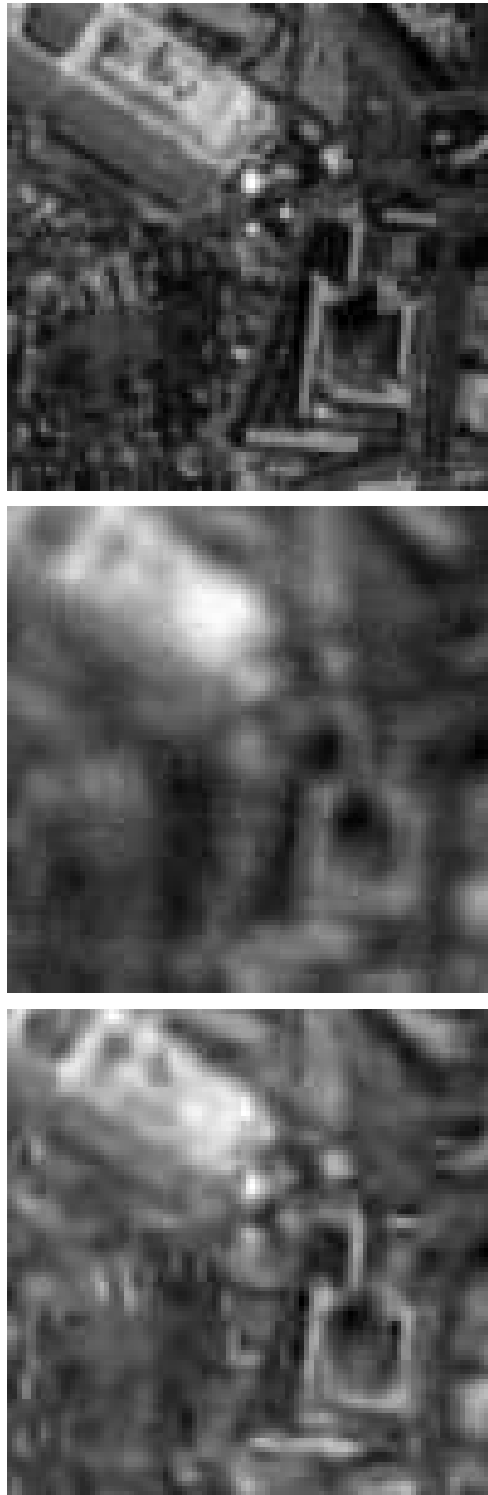


FIG. 4.4 – Example 2 – Zoom on a 100×100 portion of the SPOT5 satellite image. Original image (top); degraded image (center); image restored in a dual-tree wavelet frame with 100 iterations of Algorithm 4.4.35 (bottom).

In this example, we perform a restoration in a discrete two-dimensional version of an M -band dual-tree wavelet frame [7]. This decomposition has a redundancy factor of 2 (i.e., with the notation of Section 4.4.3.3.2, $K/N = 2$). In our experiments, decompositions over 2 resolution levels are performed with $M = 4$ using the filter bank proposed in [1]. The function Ψ in Problem 4.4.11 is given by (4.60), where f_V is the probability density function of the Gaussian noise. A solution is obtained via Algorithm 4.4.35. For the representation under consideration, we derive from (4.96) that $\beta = 2$ and we set $\gamma_n \equiv 0.995$. The restored image, shown in Fig. 4.3 (bottom), yields a relative error of 15.68 dB, i.e., a significant improvement of over 3 dB in terms of signal-to-noise ratio. A more precise inspection of the magnified areas displayed in Fig. 4.4 shows that the proposed method makes it possible to recover sharp edges while removing noise in uniform areas. The good behavior in terms of edge recovery may be attributed to the choice of the M -band dual-tree wavelet decomposition which is known to provide a good representation of directional features such as edges [7].

4.5 Bibliographie

- [1] O. Alkin and H. Caglar, Design of efficient M -band coders with linear-phase and perfect-reconstruction properties, *IEEE Trans. Signal Process.*, vol. 43, pp. 1579–1590, 1995.
- [2] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, Image coding using wavelet transform, *IEEE Trans. Image Process.*, vol. 1, pp. 205–220, 1992.
- [3] J. M. Bioucas-Dias, Bayesian wavelet-based image deconvolution : a GEM algorithm exploiting a class of heavy-tailed priors, *IEEE Trans. Image Process.*, vol. 15, pp. 937–951, 2006.
- [4] C. Bouman and K. Sauer, A generalized Gaussian image model for edge-preserving MAP estimation, *IEEE Trans. Image Process.*, vol. 2, pp. 296–310, 1993.
- [5] C. L. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, *Inverse Problems*, vol. 20, pp. 103–120, 2004.
- [6] E. J. Candès and D. L. Donoho, Recovering edges in ill-posed inverse problems : Optimality of curvelet frames, *Ann. Statist.*, vol. 30, pp. 784–842, 2002.
- [7] C. Chaux, L. Duval, and J.-C. Pesquet, Image analysis using a dual-tree M -band wavelet transform, *IEEE Trans. Image Process.*, vol. 15, pp. 2397–2412, 2006.
- [8] A. Cohen, *Numerical Analysis of Wavelet Methods*. Elsevier, New York, 2003.
- [9] A. Cohen, I. Daubechies, and J.-C. Feauveau, Biorthogonal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

- [10] P. L. Combettes, The foundations of set theoretic estimation, *Proc. IEEE*, vol. 81, pp. 182–208, 1993.
- [11] P. L. Combettes, Inconsistent signal feasibility problems : Least-squares solutions in a product space, *IEEE Trans. Signal Process.*, vol. 42, pp. 2955–2966, 1994.
- [12] P. L. Combettes, A block-iterative surrogate constraint splitting method for quadratic signal recovery, *IEEE Trans. Signal Process.*, vol. 51, pp. 1771–1782, 2003.
- [13] P. L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization*, vol. 53, pp. 475–504, 2004.
- [14] P. L. Combettes and J.-C. Pesquet, Proximal thresholding algorithm for minimization over orthonormal bases, *SIAM J. Optim.*, to appear.
- [15] P. L. Combettes and V. R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.*, vol. 4, pp. 1168–1200, 2005.
- [16] I. Daubechies, *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992.
- [17] I. Daubechies, M. Defrise, and C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Comm. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.
- [18] I. Daubechies, A. Grossmann, and Y. Meyer, Painless nonorthogonal expansions, *J. Math. Phys.*, vol. 27, pp. 1271–1283, 1986.
- [19] C. de Mol and M. Defrise, A note on wavelet-based inversion algorithms, *Contemp. Math.*, vol. 313, pp. 85–96, 2002.
- [20] L. Devroye, *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.
- [21] M. N. Do and M. Vetterli, The contourlet transform : An efficient directional multiresolution image representation, *IEEE Trans. Image Process.*, vol. 14, pp. 2091–2106, 2005.
- [22] R. J. Duffin and A. C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Amer. Math. Soc.*, vol. 72, pp. 341–366, 1952.
- [23] M. Elad and A. Feuer, Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images, *IEEE Trans. Image Process.*, vol. 6, pp. 1646–1658, 1997.
- [24] O. D. Escoda, L. Granai, and P. Vandergheynst, On the use of a priori information for sparse signal approximations, *IEEE Trans. Signal Process.*, vol. 54, pp. 3468–3482, 2006.
- [25] H. G. Feichtinger and T. Strohmer, *Gabor Analysis and Algorithms : Theory and Applications*, Birkhäuser, Boston, 1998.
- [26] D. Gabor, Theory of communication. *J. IEE (London)*, vol. 93, pp. 429–457, 1946.

- [27] D. Han and D. R. Larson, *Frames, Bases, and Group Representations*. Mem. Amer. Math. Soc., vol. 147, 2000.
- [28] J. N. Kapur and H. K. Ksevan, *Entropy Optimization Principles with Applications*. Academic Press, Boston, 1992.
- [29] E. Le Pennec and S. G. Mallat, Sparse geometric image representations with bandelets, *IEEE Trans. Image Process.*, vol. 14, pp. 423–438, 2005.
- [30] S. G. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, New York, 1999.
- [31] J.-J. Moreau, Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris Sér. A Math.*, vol. 255, pp. 2897–2899, 1962.
- [32] J.-J. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France*, vol. 93, pp. 273–299, 1965.
- [33] M. Nikolova and M. K. Ng, Analysis of half-quadratic minimization methods for signal and image recovery, *SIAM J. Sci. Comput.*, vol. 27, pp. 937–966, 2005.
- [34] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, The dual-tree complex wavelet transform, *IEEE Signal Process. Mag.*, vol. 22, pp. 123–151, 2005.
- [35] H. Stark, ed., *Image Recovery : Theory and Application*. Academic Press, San Diego, CA, 1987.
- [36] A. M. Thompson and J. Kay, On some Bayesian choices of regularization parameter in image restoration, *Inverse Problems*, vol. 9, pp. 749–761, 1993.
- [37] R. Tolimieri and M. An, *Time-Frequency Representations*. Birkhäuser, Boston, MA, 1998.
- [38] J. A. Tropp, Just relax : Convex programming methods for identifying sparse signals in noise, *IEEE Trans. Inform. Theory*, vol. 52, pp. 1030–1051, 2006.
- [39] H. L. Van Trees, *Detection, Estimation, and Modulation Theory – Part I*. Wiley, New York, 1968.
- [40] C. Zălinescu, *Convex Analysis in General Vector Spaces*. World Scientific, River Edge, NJ, 2002.

Chapitre 5

Bilan et Perspectives

5.1 Bilan

Cette thèse a abordé des problèmes non linéaires de traitement du signal et de l'image par le biais de l'analyse convexe. Nous avons mis en valeur le rôle central que joue l'opérateur proximal dans l'analyse et la résolution de tels problèmes.

Le Chapitre 2 est une synthèse des propriétés de cet opérateur, pour la plupart, utilisées dans toute la thèse. De plus, nous avons montré comment l'opérateur proximal est un outil d'analyse et de décomposition des images, tant d'un point de vue théorique que par des simulations numériques.

Le Chapitre 3 a présenté un problème d'une large portée applicative : nous avons exhibé une grande classe de problèmes variationnels inverses du traitement du signal, à première vue complètement dissociés, qui peuvent s'interpréter comme un problème de minimisation de somme de deux fonctions soumises à certaines propriétés de régularité. Pour le résoudre, des résultats de convergence de l'algorithme de type explicite-implicite, basé sur l'opérateur proximal, ont été établis.

Enfin, dans le Chapitre 4, nous nous sommes focalisés sur un des exemples traités dans le Chapitre 3 pour l'étendre et l'appliquer aux problèmes inverses linéaires traités par trames. Des simulations numériques convaincantes viennent illustrer ce chapitre.

5.2 Perspectives

Cette thèse donne lieu à des perspectives sous divers angles : analyse convexe, étude algorithmique, traitement du signal et autres applications des mathématiques.

- Le point de vue théorique : vu le rôle moteur de l'opérateur proximal, une première piste d'étude est de déterminer d'avantage de propriétés et de calculs explicites d'opérateurs proximaux. Par exemple, sous quelles conditions sur $\varphi \in \Gamma_0(\mathcal{H})$ et $\psi \in \Gamma_0(\mathcal{H})$ peut-on obtenir une formule générale pour $\text{prox}_\varphi \circ \text{prox}_\psi$? Un autre exemple : afin d'unifier et d'étendre des études déjà menées, il serait intéressant, dans le cas où l'espace de Hilbert est $H^1(\Omega)$, Ω étant un domaine ouvert de \mathbb{R}^m , d'obtenir des propriétés et/ou caractérisations de l'opérateur proximal (à défaut d'une expression générale) de la fonctionnelle intégrale f , évoquée au Chapitre 3, de la forme

$$f: x \mapsto \gamma \int_{\Omega} \varphi(\omega, x(\omega), \nabla x(\omega)) d\omega, \quad (5.1)$$

où $\gamma \in]0, +\infty[$. Dans la Section 3.3.5 du Chapitre 3, l'importance d'une telle étude est mise en valeur dans divers domaines du traitement du signal. En particulier, si φ est fonction de x uniquement, il est possible, sous certaines conditions sur φ , d'obtenir une expression explicite [11]. Une telle fonctionnelle, dont les propriétés ont été abordées dans [12], englobe, de plus, les entropies qui ont déjà fait l'objet d'études variées [3, 4, 8, 10, 15] mais dont l'opérateur proximal demeure inconnu. Une autre fonction dont l'opérateur proximal serait utile est l'extension de l'enveloppe de Moreau introduite dans [1] (notons que l'enveloppe de Pasch-Hausdorff en est un cas particulier), et plus généralement, une convolution infimale quelconque.

Pour élargir les perspectives théoriques, rappelons que la propriété clé qui fait de l'opérateur proximal un outil puissant d'analyse des signaux est de posséder une propriété de décomposition selon un opérateur maximal monotone (le sous-différentiel d'une fonction de $\Gamma_0(\mathcal{H})$, où \mathcal{H} est un espace de Hilbert réel). Une première voie s'ouvre à nous : décomposer des signaux à l'aide d'une décomposition selon un opérateur maximal monotone, différent du sous-différentiel d'une fonction de $\Gamma_0(\mathcal{H})$, par exemple, un opérateur antisymétrique. Une autre possibilité est de travailler avec une variante de la résolvante du sous-différentiel d'une fonction $\varphi \in \Gamma_0(\mathcal{H})$, définie en remplaçant l'opérateur Id dans $(\text{Id} + \partial\varphi)^{-1}$ par un opérateur $H \in \mathcal{S}_{++}$ et d'étudier les décompositions de l'opérateur $(H + \partial\varphi)^{-1}$. Cette méthode est employée dans [5] pour obtenir des vitesses de convergence de l'algorithme explicite-implicite et nous conduit au second point. Les propriétés de l'opérateur $(H + A)^{-1}$, où $A: \mathcal{H} \rightarrow 2^{\mathcal{H}}$ est monotone, sont étudiées dans [7].

- Le point de vue algorithmique : En rapport avec le point de vue théorique, la première idée serait de discrétiser la fonctionnelle f évoquée ci-dessus à des fins numériques, basées sur l'algorithme explicite-implicite. Une seconde idée serait de généraliser ce dernier algorithme dans le cas de métriques variables pour unifier et synthétiser les résultats de [2, 9], ou dans le cas de [5], et étudier les cas de convergence.

- Le point de vue applicatif : les applications de cette thèse appartiennent essentiellement au domaine de l'imagerie. Il reste à faire, d'une part, une étude plus approfondie des propriétés de chaque composante de la décomposition proximale de l'image analysée, ainsi dégagées au Chapitre 2 à partir du choix des fonctionnelles. D'autre part, il serait intéressant d'appliquer la théorie élaborée dans cette thèse en utilisant d'autres techniques spécifiques au traitement de l'image. Les applications pourraient, de plus, s'élargir à d'autres signaux que les images, et même, à d'autres domaines que le traitement du signal, en rapport avec les problèmes d'équilibre en économie [6] ou la morphologie mathématique [13, 14].

Paris, avril 2007.

5.3 Bibliographie

- [1] H. ATTOUCH AND R. J.-B. WETS, *Epigraphical analysis*, Analyse Non Linéaire, (H. Attouch, J.-P. Aubin, F. Clarke, and I. Ekeland Eds.) Gauthier-Villars, Paris (1989), pp. 73–100.
- [2] J. F. BONNANS, J. CH. GILBERT, C. LEMARÉCHAL, AND C. A. SAGASTIZÁBAL, *A family of variable metric proximal methods*, Mathematical Programming, 68 (1995), pp. 15–47.
- [3] J. M. BORWEIN, A. S. LEWIS, AND D. NOLL, *Maximum entropy reconstruction using derivative information, I : Fisher information and convex duality*, Mathematics of Operations Research, 21 (1996), pp. 442–468.
- [4] J. M. BORWEIN, A. S. LEWIS, M. N. LIMBER, AND D. NOLL, *Maximum entropy reconstruction using derivative information, Part 2 : computational results*, Numerische Mathematik, 69 (1995), pp. 243–256.
- [5] G. H.-G. CHEN AND R. T. ROCKAFELLAR, *Convergence rates in forward-backward splitting*, SIAM Journal on Optimization, 7 (1997), pp. 421–444.
- [6] P. L. COMBETTES AND S. A. HIRSTOAGA, *Equilibrium programming in Hilbert spaces*, Journal of Nonlinear and Convex Analysis, 6 (2005), pp. 117–136.
- [7] Y.-P. FANG AND N.-J. HUANG, *H-monotone operator and resolvent operator technique for variational inclusions*, Applied Mathematics and Computation, 145 (2003), pp. 795–803.
- [8] U. HERMANN AND D. NOLL, *Adaptive image reconstruction using information measures*, SIAM Journal on Control and Optimization, 38 (2000), pp. 1223–1240.

- [9] C. LEMARÉCHAL AND C. SAGASTIZÁBAL, *Variable metric bundle methods : from conceptual to implementable forms*, Mathematical Programming, 76 (1997), pp. 393–410.
- [10] D. NOLL, *Reconstruction with noisy data : An approach via eigenvalue optimization*, SIAM Journal on Control and Optimization, 8 (1998), pp. 82–104.
- [11] R. T. ROCKAFELLAR, *Integrals which are convex functionals*, Pacific Journal of Mathematics, 24 (1968), pp. 525–539.
- [12] R. T. ROCKAFELLAR, *Integrals which are convex functionals II*, Pacific Journal of Mathematics, 39 (1971), pp. 439–469.
- [13] J. SERRA, *Image Analysis and Mathematical Morphology I*, Academic Press, London, 1982.
- [14] J. SERRA, *Image Analysis and Mathematical Morphology II : Theoretical Advances*, Academic Press, London, 1988.
- [15] L. A. VESE AND S. J. OSHER, *Numerical methods for p -harmonic flows and applications to image processing*, SIAM Journal on Numerical Analysis, 40 (2002), pp. 2085–2104.