

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

ALGORITMO FORWARD-BACKWARD MULTIMALLA CON
APLICACIÓN AL PROBLEMA DE SUPRESIÓN DE RUIDO EN UNA
IMAGEN.

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL
TÍTULO DE MAGÍSTER EN OPTIMIZACIÓN MATEMÁTICA

EDISON FELIPE GUERRA URGILES
gflip_05j@hotmail.es

Director: Tuomo Valkonen, p.H.D.
tuomo.valkonen@epn.edu.ec

Codirector: Sergio González Andrade, p.H.D.
sergio.gonzalez@epn.edu.ec

Quito, enero 2023

Declaración de autoría.

Yo, Edison Felipe Guerra Urgilés, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Escuela Politécnica Nacional puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por la Ley de la Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Edison Felipe Guerra Urgilés

Certificación

Certifico que el presente trabajo fue realizado y desarrollado por Edison Felipe Guerra Urgilés, bajo mi supervisión.

Tuomo Valkonen, p.H.D.
DIRECTOR

Sergio González Andrade, p.H.D.
CODIRECTOR

DEDICATORIA

*A mi Señor y Salvador Jesús de Nazaret,
y a mi hija Daniela Aseneth.*

Agradecimientos

Primeramente, agradezco a mi Dios por haberme rescatado hace más de 10 años y darme la oportunidad de vivir estos momentos en el ámbito científico. Luego, agradezco a mi amada esposa Anita, que me ha apoyado incondicionalmente. También a mi mamá que con sus consejos alentó para no desmayar. Finalmente, quiero agradecer a mi director Tuomo Valkonen y a mi codirector Sergio González, quienes me supieron guiar en todo momento.

Índice general

1. Introducción	6
2. Preliminares	11
2.1. Análisis Convexo	12
2.2. Método Forward-Backward Splitting	18
2.3. Métodos Multimalla	19
2.3.1. Operadores de Restricción y Prolongación	22
2.4. Procesamiento de Imágenes	23
3. Algoritmo Forward-Backward Multimalla	26
3.1. Condición de Coherencia	27
3.1.1. Primer Intento	28
3.1.2. Problema en la malla gruesa	29
3.2. Búsqueda lineal	30
3.3. Análisis Importante	34
3.4. Algoritmo Forward-Backward Multigrid	38
3.5. Problema de Supresión de Ruido en una Imagen	42
3.5.1. Caso General	44
3.5.2. Caso Especial	45
4. Proyecciones	50
4.1. Operadores de Restricción y Prolongación	51
4.2. Operador proximal de la función no suave en la malla gruesa	54
5. Experimentación Numérica	79
5.1. Problema en la malla fina	79
5.2. Problema en la malla gruesa	79
5.2.1. Algoritmo FBMG	81
5.3. FB vs. FBMG	82
6. Conclusiones y Recomendaciones	86

Capítulo 1

Introducción

A lo largo de la historia, los problemas de optimización han generado gran interés en la comunidad científica, dando lugar al desarrollo de teorías y algoritmos que determinan la solución, o aproximación de la misma, de esta rama de la matemática. Los problemas de optimización cubren diversos tipos de problemas de control óptimo modelados mediante ecuaciones diferenciales parciales. Este campo ha tomado un gran interés en el ámbito científico gracias a la vasta variedad de aplicaciones en áreas de flujo de fluidos, crecimiento de cristales o fenómenos de calor, proporcionando diversas técnicas y algoritmos que permitan hallar la solución de dichos problemas [9]. Varios de los algoritmos creados para aproximar la solución para este tipo de problemas de control óptimo aluden a técnicas primal-dual, métodos de Newton o quasi-Newton [7, 18, 2].

Por su naturaleza, los problemas de control óptimo (también conocidos como problemas de optimización de ecuaciones diferenciales parciales con restricciones) discretizados, tienden a ser problemas que contemplan un gran número de variables. De modo que, desarrollar un algoritmo multimalla que mejore la eficiencia computacional a la hora de aproximar la solución es relevante en este ámbito.

Asimismo, el procesamiento de imágenes es un campo que involucra a la optimización. Esta rama de investigación ha sido objeto de estudio, con mayor intensidad en las últimas décadas, por la diversidad de aplicaciones como: imágenes satelitales, imágenes de rayos X, tomografías computarizadas, etc. [6]. Muchos problemas de procesamiento de imágenes se pueden categorizar como: denoising (eliminación de ruido), deblurring (desenfocar), edge detection (detección de bordes), optical flow computation (cálculo de flujo óptico), restoration – inpainting (restauración – repintura), etc. [3].

Debemos mencionar, además, que dentro de la ciencia de imágenes existen problemas conocidos como problemas inversos. Dichos problemas tienen diversas aplicaciones en Magnetic Resonance Imaging (MRI), Computerized Tomography (CT), Positron Emission Tomography (PET), Electrical Impedance Tomography (EIT) y similares.

Informalmente, los problemas directos e inversos se pueden describir con dos acciones: causa y efecto. En otras palabras, los problemas directos encuentran el efecto de una causa, por el contrario, el problema inverso recibe un efecto y quiere recuperar la causa [16].

Esta formulación de los problemas inversos involucra diversas herramientas matemáticas como: la transformada de Radon, EDP, regularizaciones de Tikhonov, etc. [12]. Mediante las herramientas matemáticas descritas dentro de los problemas inversos, surgen modelos de problemas de optimización como

$$\min_{x \in X} \frac{1}{2} \|A(x) - b\|^2 + R(x), \quad (1.1)$$

donde A es un operador “forward” y R es la regularización que se aplica a los datos del problema. Los problemas de optimización en que estamos interesados en estudiar poseen una estructura similar al problema (1.1).

En la ciencia de imágenes, diversos métodos se han creado por la necesidad de encontrar numéricamente la solución de este bagaje de problemas en el procesamiento de imágenes. Sin embargo, existen varios métodos distinguidos: forward-backward splitting [15], primal-dual proximal splitting [20], alternating direction method of multipliers (ADMM) [11, 22]. Por tanto, desarrollar un algoritmo multimalla forward-backward implica progresar en la ciencia de imágenes, ya que los métodos multimalla apelan a la eficiencia computacional.

Dentro de este diverso conjunto de aplicaciones asociadas a problemas de optimización surgen aquellos que por su extensa cantidad de variables son llamados problemas de optimización a gran escala. Muchos métodos de optimización se han desarrollado para resolver los problemas a gran escala, manteniendo un nivel tolerable en el costo computacional y de almacenamiento [18]. Dentro de los métodos a gran escala, nacen los métodos multimalla.

Los métodos multimalla fueron creados originalmente para resolver ecuaciones diferenciales ordinarias con condiciones de borde de Dirichlet. Posteriormente, se desarrollaron para resolver ecuaciones diferenciales parciales con diversas condiciones de frontera [4]. Sin embargo, Stephen G. Nash descubrió, creativamente, como usar los métodos multimalla para resolver problemas de optimización a gran escala [17], asumiendo que la función objetivo es diferenciable. La idea básica de los métodos multimalla para problemas de optimización es usar problemas de menor dimensión (problemas de optimización en la “malla gruesa”) que proporcionen información útil para aproximar la solución en el problema original (problema de optimización en la “malla fina”) [17]. Los operadores de transferencia, llamados operadores de restricción y prolongación, dan forma a la idea detrás de los métodos multimalla, es decir, los operadores de trans-

ferencia pasan la información del problema original al problema de menor dimensión y viceversa [4].

Pese a la excelente idea de usar los métodos multimalla para resolver problemas de optimización a gran escala con función objetivo diferenciable, el método propuesto por Stephen G. Nash no es posible aplicar a varios problemas del tipo

$$\min_{x \in X} G(x) + F(x), \quad (\text{P})$$

donde F es diferenciable, pero G no es diferenciable. Hace varios años, Panos Parpas describe un método multimalla para aproximar la solución del problema de optimización a gran escala (P), mediante un proceso de suavizar la función G , [19], cuando la función G es convexa, no diferenciable.

Existen varios métodos que aproximan la solución del problema (P) sin necesidad de recurrir a un proceso de suavizar la función G . Uno de estos métodos es conocido como *Forward-Backward Splitting Method* [15]. Pero cuando el problema (P) es un problema de optimización a gran escala, el algoritmo derivado del método *Forward-Backward Splitting Method*, demanda un esfuerzo computacional proporcional al número de variables del problema. Entonces emerge una pregunta natural, ¿es posible diseñar un algoritmo multimalla procedente del método *Forward-Backward Splitting Method* de tal forma que aproximen la solución del problema de optimización a gran escala (P)?

Con el objetivo de responder a la pregunta realizada, nos disponemos en el presente trabajo de investigación a desarrollar un algoritmo multimalla derivado del método *Forward-Backward Splitting Method* para resolver el problema del tipo (P). Dicho algoritmo será implementado para resolver el problema de supresión de ruido en una imagen, por su importancia en el ámbito de ciencia de imágenes.

El problema de supresión de ruido en una imagen surge del hecho que el ruido en una imagen es inevitable, debido al proceso de formación de una imagen, grabación de imágenes o transmisión de imágenes [21]. Con la innovación digital se supondría que el ruido en una imagen captada por una cámara digital evita la adherencia de ruido en la imagen, sin embargo, ya que la cámara digital realiza un conteo de fotones para medir el brillo en la imagen, y este proceso es aleatorio, el conteo de fotones también es un proceso aleatorio, dando como resultado cierto nivel de ruido en la medición de la imagen [3]. Diferentes modelos se pueden tomar para expresar el ruido en una imagen [3], pero un modelo adecuado para representar el ruido en una imagen, se realiza mediante las normas de variación total, es decir, normas L_1 del gradiente [21]. Usar este tipo de normas en este problema implica usar métodos no suaves de aproximación de la solución, en otras palabras, métodos como *forward-backward splitting method*, *primal-dual proximal splitting*, *alternating direction method of multipliers (ADMM)* (los cuales

ya hemos mencionado previamente) entre otros.

El modelo que usaremos nosotros en el presente trabajo de investigación para expresar y/o remover el ruido de una imagen es

$$\min_{x \in \mathbb{R}^{n_1 n_2}} \frac{1}{2} \|x - b\|^2 + \alpha \|\nabla x\|_{2,1}.$$

Trataremos con mucha más atención este problema en la siguiente sección, donde indicaremos su formulación dual y la forma de adaptar al algoritmo derivado del método forward-backward para aproximar la imagen sin ruido como solución del modelo.

Distribuimos el contenido de nuestra investigación de la siguiente manera. En el capítulo 2 describimos la notación que vamos a usar; también mencionamos resultados importantes del análisis convexo y optimización no suave (herramientas fundamentales para tratar problemas no diferenciables) junto a una breve introducción de los métodos multimalla para que el lector se familiarice con el entorno de investigación [7, 4]. Finalmente, tratamos con más detalle el problema que modela la eliminación de ruido en una imagen, previamente mencionado [21, 3].

En el capítulo 3 nos enfocaremos en diseñar el algoritmo *Forward-Backward Multigrid* para aproximar la solución del problema del tipo (P). Para ello, usaremos la estructura del método *Forward-Backward Multigrid* para diseñar la cierta condición que relacione el problema original con el problema de menor dimensión, de modo que garanticemos que la información obtenida hasta cierta iteración no se pierda al pasar al problema en la malla gruesa. Luego, usaremos la solución del problema de menor dimensión para definir una dirección de descenso con respecto a la derivada direccional de la función objetivo del problema del tipo (P), de tal forma que obtengamos una mejor aproximación de la última iteración del problema original, mediante una búsqueda lineal. Finalmente, presentaremos varios resultados de validación del algoritmo *Forward-Backward Multigrid* propuesto.

El capítulo 4 tiene como objetivo describir la construcción de ciertos operadores (operadores de restricción y prolongación) que nos permitirán pasar del problema original al problema de menor dimensión, específicamente para el problema de supresión de ruido en una imagen. Además, explicamos el cálculo del operador proximal de la función no diferenciable en el problema de menor dimensión, en dos versiones diferentes, con la intención de facilitar la implementación del problema de supresión de ruido, en el lenguaje Matlab, y posterior comparación numérica.

El capítulo 5 está dedicado enteramente a comparar numéricamente el algoritmo *Forward-Backward Multigrid* diseñado en los capítulos previos con su homólogo, describiendo las instancias usadas para la experimentación numérica y los parámetros usados.

Finalmente, en el capítulo 6 damos a conocer las conclusiones y recomendaciones que surgieron durante el estudio del presente proyecto de investigación.

Capítulo 2

Preliminares

En este capítulo describiremos la notación y los espacios en los cuales vamos a trabajar durante todo el desarrollo de la investigación.

Durante el presente trabajo de investigación usaremos frecuentemente el término “problema en la malla fina” para referirnos al problema de optimización del tipo (P). De modo que, asumiremos que el espacio $X \subseteq \mathbb{R}^h$ es un espacio de Hilbert. Usaremos la letra $x \in X$ para la variable del problema de optimización del tipo (P).

El problema de menor dimensión, por el contrario, será mencionado como “problema en la malla gruesa”, tiene un papel relevante dentro del algoritmo multimalla. De igual forma, el espacio de Hilbert en el cual vamos a construir el problema en la malla gruesa es $X_H \subseteq \mathbb{R}^H$, y la variable del problema de menor dimensión la identificaremos con el símbolo $\zeta \in X_H$. Rápidamente, damos una breve estructura del problema en la malla gruesa:

$$\min_{\zeta \in X_H} G_H^k(\zeta) + F_H(\zeta) + \langle w_H^k, \zeta \rangle, \quad (2.1)$$

donde $G_H^k : X_H \rightarrow \mathbb{R}$ y $w_H^k \in X_H$ se construyen en la k -ésima iteración que aproxima la solución del problema en la malla fina (P). Evidentemente, al referirnos al problema en la malla gruesa, implícitamente afirmamos que $H < h \in \mathbb{N}$.

Diremos que una función F es G-diferenciable si es Gâteaux diferenciable y notaremos la derivada de Gâteaux de la función, por ejemplo F , como ∇F . Además, diremos que una función tiene gradiente L -Lipschitz si la derivada de Gâteaux de una función es Lipschitz continua con constante $L > 0$. Por lo contrario, denotaremos la derivada direccional de una función F en un punto x en la dirección d como

$$F'(x, d).$$

También por facilidad abreviaremos el método forward-backward splitting method

por FB, cuando hagamos referencia al uso en la resolución del problema (2.25).

2.1. Análisis Convexo

Las siguientes definiciones usaremos con mucha regularidad en el análisis teórico del algoritmo derivado del FB multimalla. La identidad de tres puntos es una herramienta fundamental para analizar la convergencia de los métodos FB, Douglas-Rachford, PDPS, etc. [7, Capítulo 9,11]). Además, es necesario aclarar que la desigualdad de tres puntos suave, si bien en el presente trabajo de investigación es presentada como una definición, es consecuencia de que la derivada de Gâteaux de una función sea Lipschitz continua en un espacio de Hilbert [7, Capítulo 7].

Definición 2.1.1 (Identidad de tres puntos). *Sea X un espacio de Hilbert. La identidad de tres puntos es*

$$\langle x - z, x - \hat{x} \rangle = \frac{1}{2} \|x - z\|^2 - \frac{1}{2} \|z - \hat{x}\|^2 + \frac{1}{2} \|x - \hat{x}\|^2, \quad \forall z, x, \hat{x} \in X. \quad (2.2)$$

Definición 2.1.2 (Desigualdad de tres puntos suave). *Sean X un espacio de Hilbert, $F : X \rightarrow \mathbb{R}$ Gâteaux diferenciable con gradiente L -Lipschitz. Entonces la desigualdad de tres puntos suave es:*

$$\langle \nabla F(z), x - \hat{x} \rangle \geq F(x) - F(\hat{x}) - \frac{L}{2} \|x - z\|^2, \quad \forall z, x, \hat{x} \in X. \quad (2.3)$$

La siguiente definición puede estudiarse con más detalle en [7, Capítulo 6].

Definición 2.1.3 (Operador Monótono). *Sea X un espacio de Hilbert. Un mapeo conjunto-valuado $A : X \rightrightarrows X^*$ se llama monótono si $\text{graph}(A) \neq \emptyset$ (para excluir los casos triviales) y*

$$\langle y_1 - y_2, x_1 - x_2 \rangle \geq 0, \quad \forall (x_1, y_1), (x_2, y_2) \in \text{graph}(A).$$

Durante el presente trabajo de investigación usaremos un operador fundamental para desarrollar el método FB, conocido como operador proximal.

Definición 2.1.4. *Sean X un espacio de Hilbert, una función $J : X \rightarrow \mathbb{R}$ propia, convexa y semicontinua inferior, se define el operador proximal de J como*

$$\text{prox}_F : X \rightarrow X, \quad \text{prox}_J(x) = \underset{z \in Z}{\text{argmin}} \frac{1}{2} \|z - x\|_X^2 + J(z).$$

Una identidad equivalente para el operador proximal también es

$$\text{prox}_F = (I + \partial F)^{-1}. \quad (2.4)$$

Para más detalles sobre los operadores proximales y su definición, referimos al lector a [7].

A continuación hablaremos brevemente de la construcción del método FB junto a resultados relevantes que nos ayudarán en el análisis de convergencia del algoritmo multimalla derivado del método FB. Para ello, es necesario enunciar en el presente trabajo de investigación resultados sumamente relevantes como lemas, dejando al lector la inquietud de estudiar las demostraciones de los siguientes resultados en [7].

El primer lema que vamos a presentar, nos da una caracterización de las condiciones de optimalidad para un problema de optimización convexo en la forma

$$\min_{x \in X} J(x). \quad (2.5)$$

Lema 2.1.1 (Principio de Fermat). *Sea X un espacio de Hilbert. También, sean $J : X \rightarrow \mathbb{R}$ y $x^* \in \text{dom } J$. Entonces los siguientes enunciados son equivalentes:*

- i) $0 \in \partial J(x^*)$,*
- ii) $J(x^*) = \min_{x \in X} J(x)$.*

Seguidamente, exponemos dos reglas fundamentales del sub diferencial de una o varias funciones convexas.

Lema 2.1.2. *Sean X un espacio de Hilbert y $J : X \rightarrow \mathbb{R}$ Gâteaux diferenciable en x . Entonces, $\partial J(x) = \{\nabla J(x)\}$.*

Lema 2.1.3 (Regla de la Suma). *Sean X un espacio de Hilbert, $F, G : X \rightarrow \mathbb{R}$ convexas y semicontinuas inferiormente y $x \in \text{dom } F \cap \text{dom } G$. Entonces*

$$\partial F(x) + \partial G(x) \subset \partial(F + G)(x).$$

con la igualdad si existe $x_0 \in \text{int}(\text{dom } F) \cap \text{dom } G$.

El siguiente resultado nos da una caracterización entre el sub diferencial de una función convexa con el operador proximal de dicha función. Este resultado es útil para desarrollar varios métodos proximales separados (splitting proximal methods).

Lema 2.1.4 (Sub diferencial - Operador Proximal). *Sean X un espacio de Hilbert, $J : X \rightarrow \mathbb{R}$ propia, convexa y semicontinua inferiormente. También, sea $x, \hat{x} \in X$.*

Entonces para todo $\eta > 0$,

$$\hat{x} \in \partial J(x) \quad \Leftrightarrow \quad x = \text{prox}_{\eta J}(x + \eta \hat{x}).$$

Un resultado importante que nos ayudará en el capítulo 4 cuando analicemos el operador proximal de la función no suave en el problema en la malla gruesa, es el siguiente:

Lema 2.1.5. *Sean X un espacio de Hilbert, $J_1, J_2 : X \rightarrow \mathbb{R}$ propias, convexas y semicontinuas inferiormente. Si $J(x, y) = J_1(x) + J_2(y)$ y $\gamma > 0$, entonces*

$$\text{prox}_{\gamma J}(x, y) = \begin{pmatrix} \text{prox}_{\gamma J_1}(x) \\ \text{prox}_{\gamma J_2}(y) \end{pmatrix}.$$

Una de las herramientas fundamentales dentro de la optimización convexa es la *dualidad*. Por tanto, es necesario introducir la definición de la conjugada de Fenchel (o conjugada convexa)¹. Sea X un espacio de Hilbert y $J : X \rightarrow \mathbb{R}$ una función propia, convexa y semicontinua inferiormente. Se define la conjugada de Fenchel de la función J por $J^* : X \rightarrow \mathbb{R}$ tal que

$$J^*(\hat{x}) = \sup_{x \in X} \{\langle \hat{x}, x \rangle - J(x)\}. \quad (2.6)$$

La conjugada de Fenchel posee diversas propiedades que permiten relacionar la función con su conjugada de Fenchel [7].

Lema 2.1.6 (Fenchel-Moreau-Rockafellar). *Sean X un espacio de Hilbert y $J : X \rightarrow \mathbb{R}$ propia. Entonces*

- i) $J^{**} \leq J$
- ii) $J^{**}(x) = \sup \{a(x) : a \text{ continua afín tal que } a(\tilde{x}) \leq J(\tilde{x}), \forall \tilde{x} \in X\}$
- iii) $J^{**} = J$ sí y solo sí J es convexa y semicontinua inferiormente.

Además, podemos obtener un sub gradiente de la conjugada de Fenchel a partir de un sub gradiente de su función en un punto. Resumimos este resultado en el siguiente lema.

Lema 2.1.7. *Sean X un espacio de Hilbert y $J : X \rightarrow \mathbb{R}$ propia, convexa y semicontinua inferiormente. Entonces, los siguientes enunciados son equivalentes para todo $x, \hat{x} \in X$:*

¹Referimos al lector a [7, Cap. 5] donde se describe detalladamente la definición de la conjugada de Fenchel.

$$i) \langle \hat{x}, x \rangle_X = J(x) + J^*(\hat{x});$$

$$ii) \hat{x} \in \partial J(x);$$

$$iii) x \in \partial J^*(\hat{x}).$$

Otra de las propiedades que tiene la conjugada de Fenchel particularmente para una función de traslación, se describe a continuación.

Lema 2.1.8. Sean X un espacio de Hilbert, $J : X \rightarrow \mathbb{R}$ propia, convexa y semicontinua inferiormente, y $x_0 \in X$. Entonces la conjugada de Fenchel de las funciones $J_1(x) = J(x - x_0)$ y $J_2(x) = J(x) + \langle x_0, x \rangle$ tiene la siguiente relación $J_1^* = J_2$ y $J_2^* = J_1$, respectivamente.

El siguiente resultado permite definir la formulación dual de un problema de optimización convexo del tipo

$$\min_{x \in X} G(Kx) + F(x), \quad (2.7)$$

donde K es un operador lineal y continuo. Dicho resultado describe las condiciones suficientes de optimalidad para un problema de optimización del tipo (2.7). Dichas condiciones de optimalidad pueden intercambiarse usando las propiedades de la conjugada de Fenchel presentadas previamente [7]. En particular, este resultado nos ayudará en el problema de supresión de ruido en una imagen, de forma que podamos expresar la solución del problema en términos sencillos de una de sus funciones.

Lema 2.1.9 (Fenchel-Rockafellar). Sean X y Y dos espacios de Hilbert, además, $F : X \rightarrow \mathbb{R}$ y $G : Y \rightarrow \mathbb{R}$ propias, convexas y semicontinuas inferiormente, y $K \in \mathcal{L}(X; Y)$. Supongamos, aún más,

$$i) \text{ el problema primal admite una solución } x^* \in X;$$

$$ii) \text{ existe } \hat{x} \in \text{dom}(G \circ K) \cap \text{dom} F, \text{ con } K\hat{x} \in \text{int}(\text{dom} G).$$

Entonces, el problema dual admite una solución $y^* \in Y$ y

$$\min_{x \in X} F(x) + G(Kx) = \max_{y \in Y} -G^*(y) - F^*(-K^*y).$$

Asimismo, x^* y y^* son soluciones de problema primal y dual, respectivamente, si y solo si

$$\begin{cases} y^* \in \partial G(Kx^*) \\ -K^*y^* \in \partial F(x^*) \end{cases}$$

Uno de los teoremas clásicos dentro de las ramas de optimización convexa y análisis matemático es el teorema de la proyección sobre un convexo cerrado. Para mayor detalles acerca del teorema de la proyección, referimos al lector a [14].

Lema 2.1.10 (Teorema de la Proyección). *Sean X un espacio de Hilbert, $E \subseteq X$ un conjunto no vacío, convexo y cerrado. Para todo $x \in X$, existe un único, $z \in E$ tal que*

$$\|x - z\| = \min_{w \in E} \|x - w\|.$$

Equivalentemente, a decir que para todo $x \in X$, existe un único, $z \in E$ tal que

$$\langle x - z, w - z \rangle \leq 0, \quad \forall w \in E.$$

Finalmente, es importante mencionar ciertos resultados acerca de la derivada direccional de una función convexa, ya que nos ayudará en la definición de cierta condición que relacione los problemas en la malla fina y en la malla gruesa (conocida como condición de coherencia), además de encontrar una dirección de descenso.

El primer resultado que debemos mencionar nos da algunas propiedades básicas y la definición de la derivada direccional para una función convexa. Este resultado fue tomando de [7].

Lema 2.1.11. *Sea $F : X \rightarrow \bar{\mathbb{R}}$ convexa. Sea, además, $x \in \text{dom } F$ y $d \in X$ fijo. Entonces*

(i) *la función*

$$\phi : (0, \infty) \rightarrow \bar{\mathbb{R}}, \quad t \mapsto \frac{F(x + td) - F(x)}{t}$$

es creciente;

(ii) *existe un límite $F'(x; d) = \lim_{t \searrow 0} \phi(t) \in [-\infty, \infty]$, que satisface*

$$F'(x; d) \leq F(x + d) - F(x);$$

(iii) *si $x \in \text{int}(\text{dom } F)$, el límite $F'(x; d)$ es finito.*

El segundo resultado que vamos a presentar caracteriza de la derivada direccional mediante el sub diferencial de una función.

Lema 2.1.12. Sea F una función propia, convexa y semicontinua inferior. Entonces la derivada direccional de F en el punto x y en la dirección d se escribe como

$$F'(x, d) = \sup_{f^* \in \partial F(x)} \langle f^*, d \rangle \quad (2.8)$$

La demostración de este resultado es una consecuencia de la definición de subgradiente y el Lema 4.5 en [7]. Invitamos al lector a profundizar en este resultado.

El tercer resultado que vamos a presentar nos va a ayudar a calcular la derivada direccional de una función cuando el subdiferencial está definido mediante el producto cartesiano de conjuntos.

Lema 2.1.13. Sean $X \subset \mathbb{R}^n$, $F : X \rightarrow \mathbb{R}$ una función propia, convexa y semicontinua inferior y $x \in X$. Además, supongamos que

$$F(x) = \sum_{i=1}^n f_i(x_i),$$

donde $f_i : \mathbb{R} \rightarrow \mathbb{R}$ es convexa para todo $i = 1, \dots, n$. Ya que

$$\partial F(x) = \{a^* \in \mathbb{R}^n : a_i^* \in \partial f_i(x_i), \quad 1 \leq i \leq n\}.$$

para una dirección $d \in X$ se tiene que

$$F'(x, d) = \sum_{i=1}^n \sup_{a_i^* \in \partial f_i(x_i)} a_i^* d_i \quad (2.9)$$

Demostración. Sea $x, d \in X$. De la caracterización de la derivada direccional en el Lema (2.1.12) se tiene

$$\begin{aligned} F'(x, d) &= \sup_{a^* \in \partial F(x)} \langle a^*, d \rangle \\ &= \sup_{a^* \in \partial F(x)} \sum_{i=1}^n a_i^* d_i \\ &= \sup_{a_1^* \in \partial f_1(x_1)} \cdots \sup_{a_n^* \in \partial f_n(x_n)} \sum_{i=1}^n a_i^* d_i \\ &= \sum_{i=1}^n \sup_{a_i^* \in \partial f_i(x_i)} a_i^* d_i. \end{aligned}$$

□

Este resultado es muy importante para calcular la derivada direccional en funciones separables como la norma 1 o la indicatriz de un producto de conjuntos.

Por último, el siguiente resultado usa dos propiedades inherentes de la derivada direccional: si la función es convexa, entonces la derivada direccional es convexa en las direcciones y la homogeneidad con escalares no negativos respecto a las direcciones. En otras palabras,

$$F'(x; \lambda d_1 + (1 - \lambda) d_2) \leq \lambda F'(x; d_1) + (1 - \lambda) F'(x; d_2), \quad \forall d_1, d_2 \in X, \quad \lambda \in [0, 1],$$

$$F'(x; td) = tF'(x; d), \quad \forall t \geq 0.$$

Lema 2.1.14. *Sea $F : X \rightarrow \mathbb{R}$ convexa. Entonces para $x \in \text{int}(\text{dom } F)$ y las direcciones $d_1, \dots, d_n \in X$ tenemos*

$$F' \left(x; \sum_{i=1}^n d_i \right) \leq \sum_{i=1}^n F'(x; d_i). \quad (2.10)$$

Este resultado es muy sencillo de mostrar usando la desigualdad de Jensen y la homogeneidad de la derivada direccional.

2.2. Método Forward-Backward Splitting

Aquí veremos brevemente como se deriva el método FB, mediante los lemas previamente citados. El detalle de la construcción del método se puede consultar en [15], [7].

Consideremos el problema del tipo (P)

$$\min_{x \in X} G(x) + F(x).$$

Uno de los métodos para aproximar la solución de un problema de tipo (P), donde G es no diferenciable, es el método forward-backward. El cual consiste en aplicar la regla de la suma al principio de Fermat, Lemas 2.1.3 y 2.1.1 respectivamente, para luego usar el Lema 2.1.4 en la función G . Dicho de otra forma, el método forward-backward se puede expresar de forma implícita como

$$0 \in \partial G(x^{k+1}) + \nabla F(x^k) + \tau_k^{-1}(x^{k+1} - x^k). \quad (2.11)$$

Consecuentemente, las iteraciones del método forward-backward, cuando F es Gâteaux diferenciable, vienen dadas por

$$x^{k+1} = \text{prox}_{\tau_k G}(x^k - \tau_k \nabla F(x^k)). \quad (2.12)$$

A partir de la ecuación implícita (2.11), podemos inferir varias propiedades del

método forward-backward.

Proposición 2.2.1. *Resolvemos el problema (P) con la ecuación explícita (2.11) del método forward-backward. Si $\tau_k < 2L^{-1}$, las iteraciones generadas por método forward-backward cumplen*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2, \quad (2.13)$$

$$F(x^{k+1}) + G(x^{k+1}) \leq F(x^k) + G(x^k), \quad (2.14)$$

donde $x^* \in [\partial F + G]^{-1}(0)$.

La demostración de los resultados se pueden consultar en [7, Capítulos 9, 11]. Una sucesión que satisface la desigualdad (2.13) se dice que la sucesión es Féjer monótona.

Las dos propiedades enunciadas en la Proposición 2.2.1 nos servirán para demostrar la convergencia del algoritmo multimalla derivado del FB.

Presentamos el algoritmo FB, [7], para aproximar la solución del problema de tipo (P) con el fin de comparar el método FB con el su homólogo derivado de los métodos multimalla posteriormente:

Algorithm 1 Forward-Backward Splitting (FB)

- 1: Inicialización: $x^0 \in X$, $\tau > 0$ tal que $\tau < 2L^{-1}$, $k = 0$.
 - 2: **while** criterio de parada **do**
 - 3: Calcule $\nabla F(x^k)$.
 - 4: Calcule $x^{k+1} = \text{prox}_{\tau G}(x^k + \tau \nabla F(x^k))$
 - 5: $k \leftarrow k + 1$
 - 6: **end while**
-

2.3. Métodos Multimalla

Ciertamente, un problema físico gobernado por una ecuación diferencial corresponde a un problema abstracto en un dominio espacial. Al discretizar el problema abstracto en un conjunto de puntos en una malla, el problema discreto resultante se convierte en un sistema de ecuaciones equivalentes al número de puntos en la malla, [4].

Los métodos multimalla fueron creados para mejorar la aproximación computacional en la solución de problemas físicos gobernados por ecuaciones diferenciales.

Con la intención de explicar brevemente el funcionamiento de los métodos multimalla en un problema físico gobernado por una ecuación diferencial, consideremos el modelo con condición inicial y final en \mathbb{R}

$$-u''(t) + \sigma u(t) = f(t), \quad 0 < t < 1, \quad \sigma \geq 0, \quad (2.15)$$

$$u(0) = u(1) = 0. \quad (2.16)$$

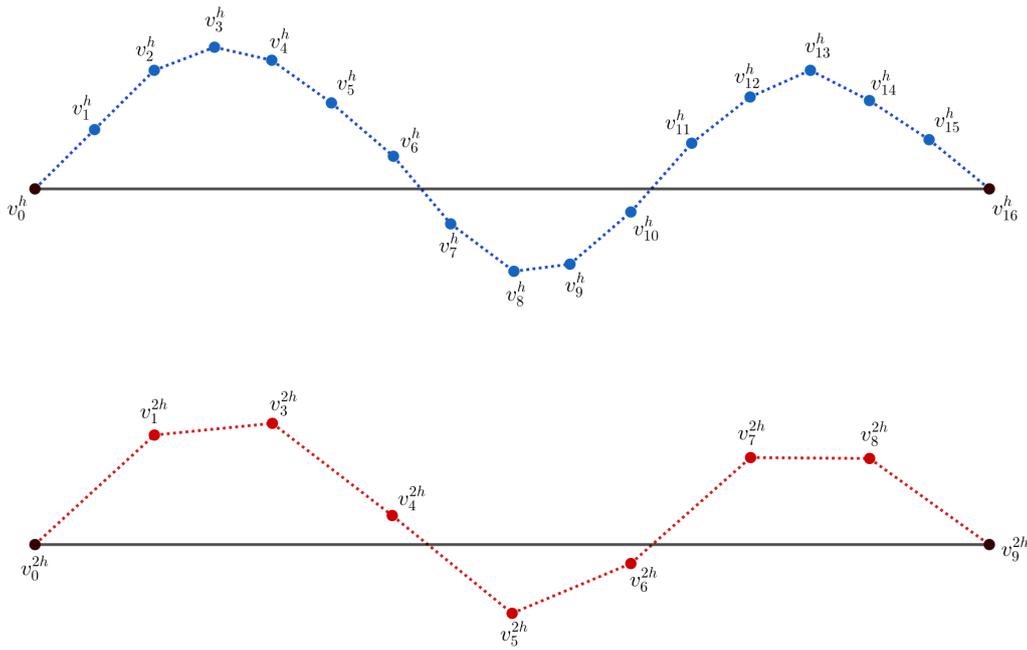


Figura 2.1: Transición de una malla de 16 nodos a una malla de 9 nodos, mediante el operador de restricción I_h^{2h} .

1. Empieza con el sistema $A\mathbf{v} = \mathbf{f}$ en una malla fina.
2. Mejora la solución del sistema lineal $A\mathbf{v} = \mathbf{f}$ usando un método iterativo.
3. Calcula el residuo de la solución aproximada del paso anterior.
4. Estima la restricción del residuo en la siguiente malla (gruesa), mediante el operador de restricción.
5. Resuelve recursivamente el sistema lineal $A^{2h}\mathbf{v}^{2h} = I_h^{2h}\mathbf{r}$, en la malla gruesa con dato inicial cero, donde $I_h^{2h}\mathbf{r}$ es la restricción del residuo.
6. Interpola la solución del sistema en la malla gruesa a la malla fina, mediante el operador de prolongación.
7. Actualiza la solución aproximada restando la corrección calculada en la malla gruesa del paso anterior, de la solución aproximada en la malla fina.
8. Nuevamente, mejora la solución usando un método iterativo.

Uno de los métodos iterativos más usados en los métodos multimalla es el método de Jacobi ponderado, con peso $\omega = 2/3$. Se ha estudiado que el método de Jacobi ponderado es un excelente método para reducir el error de alta frecuencia. Nuestra intención es dar una breve introducción de los métodos multimalla, de modo que referimos al

lector a la bibliografía básica de los métodos multimalla [4, 10] para profundizar en este campo.

La idea detrás de los métodos multimalla unidimensional se puede extender a problema físicos gobernados por ecuaciones diferenciales parciales como

$$\begin{cases} \mathcal{A}(u(t)) = f(t) \\ + \text{condiciones de frontera} \end{cases} ,$$

donde \mathcal{A} es un operador diferencial de segundo orden, corresponde a un problema físico en un dominio Ω .

Si bien, los métodos multimalla surgieron de la necesidad de reducir el costo computacional en problemas físicos gobernados por ecuaciones diferenciales, su aplicación se desarrolló también en otro tipo problemas que involucran sistemas lineales. Es decir, los métodos multimalla poseen un componente geométrico en el cual se definió una malla, sin embargo, es posible resolver con los métodos multimalla un sistema lineal carente de un componente geométrico en su mallado. A esta técnica se le conoce como métodos multimalla algebraicos, ya que la malla se define de forma algebraica y no geométrica [4].

Asimismo, los métodos multimalla progresaron hasta el área de optimización, involucrando los tres componentes fundamentales de los métodos multimalla con un ligero cambio de concepto en el operador solución, en otras palabras, usa métodos de optimización -como: el descenso del gradiente, el gradiente conjugado, Newton, quasi-Newton, etc. [18]- en lugar de los métodos iterativos para el operador de solución. Además, incorpora nuevos conceptos que mantienen la coherencia entre las distintas mallas [17].

Hemos mencionado los tres componentes fundamentales de los métodos multimalla, pero dos de ellos merecen una atención especial. El paso de información a las distintas mallas se realizan mediante los operadores de transferencia: operador de restricción y operador de prolongación, por lo que es relevante estudiar brevemente su estructura.

2.3.1. Operadores de Restricción y Prolongación

Los operadores restricción y prolongación [4], notados por I_h^H y I_H^h respectivamente, son operadores de transferencia en las distintas mallas, donde el operador restricción nos va a permitir pasar del problema en la malla fina al problema en la malla gruesa. Por el contrario, el operador prolongación nos permitirá pasar del problema en la malla gruesa al problema en la malla fina.

La forma más sencilla de definir los operadores de restricción y prolongación, que es efectiva para los métodos multimalla, es usar la inyección canónica o ponderación completa y una interpolación lineal, respectivamente [4]. Sin embargo, independiente-

Asimismo, notaremos con la función indicadora (o indicatriz de un conjunto) de un conjunto $A \subset \mathbb{R}^n$ como

$$\delta_A(x) = \begin{cases} 0, & x \in A, \\ \infty, & x \notin A. \end{cases}$$

Finalmente, ya que estamos tratando el problema de supresión de ruido en un espacio de dimensión finita, ∇ representa el gradiente de discretizado.

El problema de supresión de ruido en una imagen $b \in \mathbb{R}^{n_1 n_2}$ se describe como

$$\min_{x \in \mathbb{R}^{n_1 n_2}} \frac{1}{2} \|x - b\|^2 + \alpha \|\nabla x\|_{2,1}, \quad (2.20)$$

donde la norma isotrópica de campos de vectores $\|\cdot\|_{2,1}$ está definida por

$$\|A\|_{2,1} = \sum_{i=1}^{n_1 n_2} \|A_i\|_2.$$

Aquí, A_i es la i -ésima columna de la matriz $A \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$ representando el gradiente discreto de la imagen. La imagen $x \in \mathbb{R}^{n_1 n_2}$, de tamaño $n_1 \times n_2$, está expresado como un vector de píxeles.

Uno de los inconvenientes que presenta el problema (2.20) es que no es posible aplicar directamente el algoritmo del método forward-backward, ya que el operador proximal de $\alpha \|\nabla x\|_{2,1}$ no tiene una forma sencilla. Sin embargo, podemos resolver el problema (2.20) en su formulación dual y usar el Lema 2.1.9 para recuperar la solución.

La formulación dual de un problema de optimización convexa del tipo (P) se deriva de la conjugada de Fenchel o la conjugada convexa, de la siguiente manera. Observe que el problema de supresión de ruido en una imagen (2.20) se puede expresar como el problema de minimización del tipo (2.7) tomando las funciones

$$G(\cdot) = \|\cdot\|_{2,1}, \quad F(\cdot) = \frac{1}{2} \|\cdot - b\|^2, \quad K = \nabla.$$

De esta manera, y usando la conjugada de Fenchel, escribimos la formulación dual del problema (2.20) como sigue a continuación

$$\max_{y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}} - \left[\alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0,\alpha)}(y_i) + \frac{1}{2} \|\nabla^* y\|^2 - \langle \nabla^* y, b \rangle \right], \quad (2.21)$$

donde $y_i \in \mathbb{R}^2$, para todo $i = 1, \dots, n_1 n_2$.

Resolver el problema dual (2.21) es equivalente a resolver el problema de minimi-

zación

$$\min_{y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}} \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0, \alpha)}(y_i) + \frac{1}{2} \|\nabla^* y\|^2 - \langle \nabla^* y, b \rangle, \quad (2.22)$$

o equivalentemente

$$\min_{y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}} \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0, \alpha)}(y_i) + \frac{1}{2} \|\nabla^* y - b\|^2. \quad (2.23)$$

Puesto que haremos uso constante de las funciones del problema (2.22) para diseñar el algoritmo multimalla derivado del método forward-backward, notaremos a dichas funciones como

$$G(y) = \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0, \alpha)}(y_i) \quad \text{y} \quad F(y) = \frac{1}{2} \|\nabla^* y - b\|^2. \quad (2.24)$$

De esta forma podemos escribir el problema de supresión de ruido en su formulación dual como

$$\min_{y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}} G(y) + F(y). \quad (2.25)$$

Capítulo 3

Algoritmo Forward-Backward Multimalla

En esta sección estudiaremos las propiedades del método forward-backward para diseñar un algoritmo multimalla que aproxime la solución del problema de tipo (P). Recordemos que el problema de tipo (P) es

$$\min_{x \in X} G(x) + F(x). \quad (3.1)$$

El algoritmo forward-backward multimalla que queremos proponer consiste en realizar un bloque de iteraciones del algoritmo derivado del método forward-backward para el problema (3.1), luego construir el problema en la malla gruesa (2.1) de manera que podamos obtener una dirección de descenso después de haber iterado con el algoritmo derivado del método forward-backward, de igual manera. Una vez que hemos encontrado la dirección de descenso, realizamos la búsqueda lineal con la regla de Armijo y la derivada direccional de la función objetivo del problema (3.1). El proceso se repite cierta cantidad de veces o hasta alcanzar un criterio de parada.

En el contexto de este capítulo usaremos x para representar la variable del problema en la malla fina (3.1), junto k para identificar las iteraciones del algoritmo derivado del método forward-backward. Asimismo, notaremos con ζ a la variable del problema en la malla gruesa (2.1), a la vez que usamos j para notar las iteraciones del algoritmo derivado del método forward-backward. Debemos mencionar que, cuando escribamos $\zeta^{k,l}$, queremos enfatizar que el problema en la malla gruesa se construyó en la k -ésima iteración del problema en la malla fina.

La siguiente definición va a ser usada con frecuencia en los posteriores capítulos, con el objetivo de identificar la iteración previa al paso de información del problema en la malla fina al problema en la malla gruesa.

Definición 3.0.1 (Iteración de Transición). *Decimos que una iteración es una*

iteración de transición si es la última iteración del método FB que se realizó para resolver el problema (3.1), antes de pasar a resolver el problema en la malla gruesa.

3.1. Condición de Coherencia

En esta sección construiremos el problema de menor dimensión, el cual usaremos para describir el algoritmo multimalla derivado del método forward-backward. Además, veremos con mucho más detalle como construir una condición de coherencia para el algoritmo multimalla derivado del método FB.

Para entender de que forma tenemos que construir la condición de coherencia, analicemos primeramente el caso diferenciable, es decir,

$$\min_{x \in X} \Psi(x),$$

donde Ψ es diferenciable. Por supuesto, este análisis ya lo hizo S. Nash en [17], de modo que nosotros simplemente estudiaremos su definición y el método de construcción de la condición de coherencia.

S. Nash propuso construir el problema en la malla gruesa con una inclinación lineal, es decir, si $\Psi : X \rightarrow \mathbb{R}$ es la función dada por el usuario, es posible construir Ψ_H , como por ejemplo: $\Psi_H : X_H \rightarrow \mathbb{R}$ puede tener la misma ley de asignación de Ψ , sin embargo, está definido en X_H , [17, 19].

$$\min_{\zeta \in X_H} \tilde{\Psi}_H(\zeta) = \Psi_H(\zeta) + \langle v_H^k, \zeta \rangle, \quad (3.2)$$

con la intención de lograr emparejar en cierto modo la restricción del gradiente de la función Ψ en la iteración de transición con el gradiente del problema en la malla gruesa (3.2) en dato inicial. Por tanto, si k es la iteración de transición y $\zeta^{k,0} = I_h^H x^k$ es el dato inicial para el problema en la malla gruesa (3.2), entonces la construcción de v_H toma la forma

$$v_H^k = I_h^H \nabla \Psi(x^k) - \nabla \Psi_H(\zeta^{k,0}).$$

Consecuentemente,

$$\nabla \tilde{\Psi}_H(\zeta^{k,0}) = I_h^H \nabla \Psi(x^k). \quad (3.3)$$

A esta identidad se la conoce como **condición de coherencia** (también llamada condición de coherencia de primer orden en el caso diferencial [19]). De esta forma, la información obtenida hasta la iteración de transición se incluye en el problema de la

malla gruesa. Para ello, tenemos que desarrollar una versión no suave de (3.3).

3.1.1. Primer Intento

En nuestro caso, usaremos una idea similar a la de S. Nash para definir el problema en la malla gruesa y precisar un tipo de condición de coherencia. Es importante mencionar que la inclinación o penalización que vamos a construir va cambiando cada vez que se use el problema en la malla gruesa para aproximar la solución de (3.44), [17, 19].

La penalización del problema en la malla gruesa que nos permitirá definir una condición de coherencia la podemos dividir en dos partes: la parte diferenciable y la parte no diferenciable. Sea k la iteración de transición y el dato inicial del problema en la malla gruesa $\zeta^{k,0} = I_h^H x^k$; en la parte diferenciable podemos usar la misma idea de Nash, [17], esto es,

$$w_H^k = I_h^H \nabla F(x^k) - \nabla F_H(\zeta^{k,0}). \quad (3.4)$$

Para la parte no diferenciable quisiéramos realizar algo similar a (3.4). Una primera idea es escoger un sub gradiente tanto de G como de G_H de forma que podamos escribir

$$v_H^k = I_h^H g - g_H,$$

donde $g \in \partial G(x^k)$ y $g_H \in \partial G_H(\zeta^{k,0})$. Así, diseñamos el problema en la malla gruesa como

$$\min_{\zeta \in X_H} J_H(\zeta) = G_H(\zeta) + F_H(\zeta) + \langle v_H^k + w_H^k, \zeta \rangle. \quad (3.5)$$

Si diseñamos el problema en la malla gruesa de esta manera, podemos definir una condición de coherencia particular para los sub gradientes g y g_H , es decir, podemos emparejar la restricción del sub gradiente de J , de (3.44), en la iteración de transición con el sub gradiente de J_H , de (3.5), en el dato inicial $\zeta^{k,0}$.

Una dificultad que encontraremos más tarde cuando intentemos incorporar la información obtenida del problema en la malla gruesa mediante una búsqueda lineal es que no vamos a poder asegurar que obtengamos un parámetro $\beta > 0$ que satisfaga la condición de suficiente descenso de Armijo. Esto se debe a que el sub gradiente no es un “buen” elemento para realizar la búsqueda lineal (ver Figura 3.1).

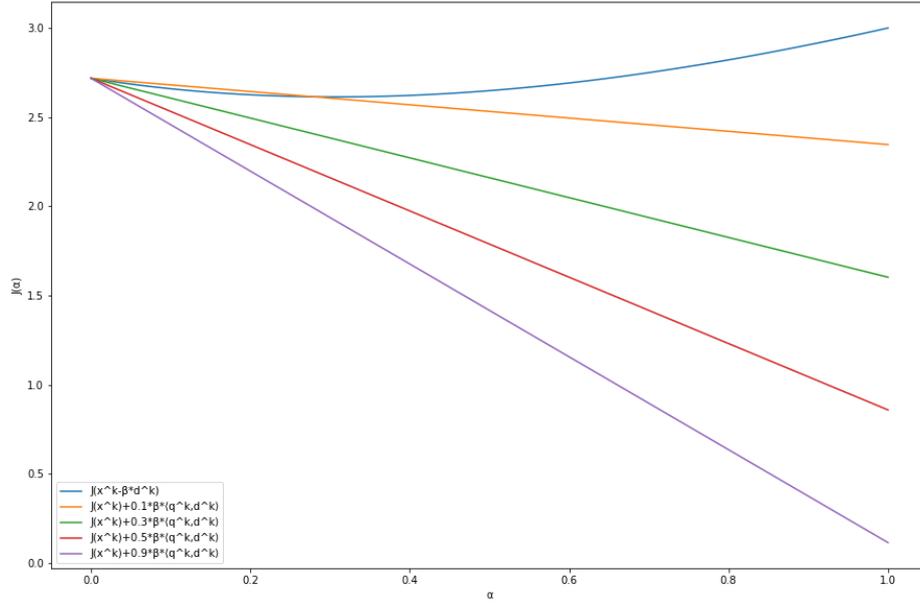


Figura 3.1: Posible caso en la búsqueda lineal con la regla de Armijo.

3.1.2. Problema en la malla gruesa

Naturalmente, la condición de coherencia está relacionada fuertemente con el problema en la malla gruesa. De modo que, tenemos que construir el problema en la malla gruesa de tal forma que evite el caso de la sección anterior mediante la condición de coherencia. Tomamos el dato inicial $\zeta^{k,0} = I_h^H x^k$ y w_H^k como (3.4).

Suposición 3.1.1 (Condición de Coherencia). *Sea k la iteración de transición. Para el problema de optimización del tipo (P),*

$$\min_{x \in X} J(x) = G(x) + F(x), \quad (3.6)$$

donde G y F son funciones convexas, propias y semicontinuas inferiormente. El problema en la malla gruesa asociado

$$\min_{\zeta \in X_H} J_H(\zeta) = G_H^k(\zeta) + F_H(\zeta) + \langle w_H^k, \zeta \rangle, \quad (3.7)$$

donde G_H^k y F_H son convexas, propias y semicontinuas inferiormente. Además, w_H^k viene dado por (3.4). Para el dato inicial se cumple $\zeta^{k,0} = I_h^H x^k$

$$I_h^H \partial G(x^k) \subseteq \partial G_H^k(\zeta^{k,0}). \quad (3.8)$$

Observe que a diferencia del caso diferenciable, la condición de coherencia que proponemos es relativa a las funciones no suaves G y G_H^k .

Lema 3.1.1. *Sea (3.1) el problema de optimización tal que las funciones G y F del problema de optimización (3.1) son convexas, propias y semicontinuas inferiormente. Suponemos que la condición de coherencia dada en la Suposición 3.1.1 se cumple. Además, supongamos que $\text{int}(\text{dom } F) \cap \text{dom } G \neq \emptyset$. Entonces*

$$I_h^H \partial J(x^k) \subset \partial J_H(\zeta^{k,0}). \quad (3.9)$$

Demostración. Es sencillo verificar que la condición de coherencia (Suposición 3.1.1) nos garantiza (3.9). Efectivamente, gracias a que G_H^k y F_H son convexas y semicontinuas inferiormente, de la Suposición 3.1.1), podemos usar la regla de la suma sobre los subdiferenciales, es decir, del Lema 2.1.3 obtenemos

$$\partial G_H^k(\zeta^{k,0}) + \nabla F_H(\zeta^{k,0}) + w_H^k \subset \partial J_H(\zeta^{k,0}),$$

pero en virtud de la definición de la penalización w_H^k , en (3.4), tenemos que

$$\partial G_H^k(\zeta^{k,0}) + I_h^H \nabla F(x^k) = \partial G_H^k(\zeta^{k,0}) + \nabla F_H(\zeta^{k,0}) + w_H^k \subset \partial J_H(\zeta^{k,0}). \quad (3.10)$$

Luego, de (3.8) obtenemos

$$I_h^H \partial G(x^k) + I_h^H \nabla F(x^k) \subset \partial J_H(\zeta^{k,0}).$$

Finalmente, ya que $\text{int}(\text{dom } F) \cap \text{dom } G \neq \emptyset$, nuevamente, de la regla de la suma para los subdiferenciales y la linealidad del operador de restricción I_h^H , obtenemos

$$I_h^H \partial J(x^k) = I_h^H \partial G(x^k) + I_h^H \nabla F(x^k) \subset \partial J_H(\zeta^{k,0}).$$

□

Observe que no tenemos una igualdad en los subdiferenciales de J_H y J , pero la propiedad (3.8), que satisfacen los subdiferenciales de G_H^k y G , nos basta para garantizar que la derivada direccional de J en x^k en la cierta dirección sea finita, y además sea una dirección de descenso.

3.2. Búsqueda lineal

En esta sección describiremos como encontrar una dirección de descenso para J en la iteración de transición k , resolviendo el problema en la malla gruesa (3.7) con el algoritmo derivado del método forward-backward (Algoritmo 1).

Antes de enunciar el resultado que nos indica cuál es la dirección de descenso para

J en la iteración de transición k , es importante mencionar que si $\zeta^{k,0} = I_h^H x^k$ es un mínimo del problema en la malla gruesa (3.7), entonces aproximar el problema en la malla gruesa no aporta ninguna información relevante para la malla fina. Por tanto, en esta sección vamos a considerar el caso cuando $\zeta^{k,0} = I_h^H x^k$ no es un mínimo para el problema en la malla gruesa (3.7).

Teorema 3.2.1. *Sea k la iteración de transición. Supongamos $\zeta^{k,0} = I_h^H x^k$ no es un mínimo local del problema en la malla gruesa (3.7), es decir, $0 \notin \partial J_H(\zeta^{k,0})$. Usando el Algoritmo 1 para aproximar la solución del problema (3.7), donde w_H^k viene dado por (3.4), con tamaño de paso $\tau \leq \min\{L^{-1}, L_H^{-1}\}$ y $m > 1$ iteraciones como criterio de parada. La dirección*

$$d = I_H^h (\zeta^{k,m} - \zeta^{k,0}) \quad (3.11)$$

satisface

$$\sup_{g \in \partial G(x^k)} \langle g + \nabla F(x^k), d \rangle \leq -\frac{1}{2\tau} \sum_{j=0}^{m-1} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 < 0. \quad (3.12)$$

Demostración. Debido a que ya se han realizado $m > 1$ iteraciones en el Algoritmo 1 para el problema en la malla gruesa (3.7), se cumple la ecuación implícita

$$0 \in \partial G_H^k(\zeta^{k,j+1}) + \nabla F_H(\zeta^{k,j}) + w_H^k + \tau^{-1}(\zeta^{k,j+1} - \zeta^{k,j}), \quad (3.13)$$

para todo $j = 0, \dots, m-1$.

De la desigualdad del sub diferencial de G_H^k y la desigualdad de tres puntos suave (2.3), para F_H , se tiene que

$$\begin{aligned} \langle \partial G_H^k(\zeta^{k,j+1}) + \nabla F_H(\zeta^{k,j}), \zeta^{k,j+1} - \zeta^{k,j} \rangle &\geq G_H^k(\zeta^{k,j+1}) - G_H^k(\zeta^{k,j}) \\ &+ F_H(\zeta^{k,j+1}) - F_H(\zeta^{k,j}) - \frac{L_H}{2} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2. \end{aligned}$$

Testeando (3.13) en $\langle \cdot, \zeta^{k,j+1} - \zeta^{k,j} \rangle$ obtenemos

$$\begin{aligned} \frac{1}{\tau} \langle \zeta^{k,j+1} - \zeta^{k,j}, \zeta^{k,j+1} - \zeta^{k,j} \rangle + \langle w_H^k, \zeta^{k,j+1} - \zeta^{k,j} \rangle &\leq G_H^k(\zeta^{k,j}) - G_H^k(\zeta^{k,j+1}) \\ &+ F_H(\zeta^{k,j}) - F_H(\zeta^{k,j+1}) + \frac{L_H}{2} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2. \end{aligned}$$

Operando en el lado izquierdo de la última desigualdad tenemos

$$\begin{aligned} \frac{1}{2\tau} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 + \frac{1}{2\tau} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 + \langle w_H^k, \zeta^{k,j+1} - \zeta^{k,0} \rangle &\leq G_H^k(\zeta^{k,j}) - \\ &G_H^k(\zeta^{k,j+1}) + F_H(\zeta^{k,j}) - F_H(\zeta^{k,j+1}) + \frac{L_H}{2} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2. \end{aligned}$$

Ya que $\tau \leq \min \{L^{-1}, L_H^{-1}\}$, se sigue que $\tau^{-1} - L_H \geq 0$. Por tanto, logramos

$$\begin{aligned} \langle w_H^k, \zeta^{k,j+1} - \zeta^{k,j} \rangle + \frac{1}{2\tau} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 &\leq G_H^k(\zeta^{k,j}) - G_H^k(\zeta^{k,j+1}) \\ &+ F_H(\zeta^{k,j}) - F_H(\zeta^{k,j+1}). \end{aligned} \quad (3.14)$$

Sumando (3.14) sobre $j = 0, \dots, m-1$, se infiere

$$\begin{aligned} \langle w_H^k, \zeta^{k,m} - \zeta^{k,0} \rangle + \frac{1}{2\tau} \sum_{j=0}^{m-1} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 &\leq G_H^k(\zeta^{k,0}) - G_H^k(\zeta^{k,m}) \\ &+ F_H(\zeta^{k,0}) - F_H(\zeta^{k,m}). \end{aligned} \quad (3.15)$$

Expandamos w_H^k . De la definición de w_H^k en (3.4) tenemos

$$\begin{aligned} \langle w_H^k, \zeta^{k,m} - \zeta^{k,0} \rangle &= \langle I_h^H \nabla F(x^k) - \nabla F_H(\zeta^{k,0}), \zeta^{k,m} - \zeta^{k,0} \rangle \\ &= \langle I_h^H \nabla F(x^k), \zeta^{k,m} - \zeta^{k,0} \rangle + \langle \nabla F_H(\zeta^{k,0}), \zeta^{k,0} - \zeta^{k,m} \rangle, \end{aligned}$$

De la convexidad de F_H , obtenemos

$$\langle w_H^k, \zeta^{k,m} - \zeta^{k,0} \rangle \geq \langle I_h^H \nabla F(x^k), \zeta^{k,m} - \zeta^{k,0} \rangle + F_H(\zeta^{k,0}) - F_H(\zeta^{k,m}). \quad (3.16)$$

Reemplazando (3.16) en (3.15), y simplificando, obtenemos

$$\langle I_h^H \nabla F(x^k), \zeta^{k,m} - \zeta^{k,0} \rangle + \frac{1}{2\tau} \sum_{j=0}^{m-1} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 \leq G_H^k(\zeta^{k,0}) - G_H^k(\zeta^{k,m}).$$

Seguidamente usamos la Proposición 4.1.1 y la definición de la dirección d , dada por (3.11), logramos

$$\langle \nabla F(x^k), d \rangle + \frac{1}{2\tau} \sum_{j=0}^{m-1} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 \leq G_H^k(\zeta^{k,0}) - G_H^k(\zeta^{k,m}). \quad (3.17)$$

Ahora, debido a que construimos el problema en la malla gruesa (3.46) de tal forma

que se cumpla la condición de coherencia (3.8) podemos escribir,

$$\langle I_h^H g, \zeta - \zeta^{k,0} \rangle \leq G_H^k(\zeta) - G_H^k(\zeta^{k,0}), \quad \forall \zeta \in X_H, \quad \forall g \in \partial G(x^k).$$

En particular tomando $\zeta = \zeta^{k,m}$ se sigue

$$\langle I_h^H g, \zeta^{k,m} - \zeta^{k,0} \rangle \leq G_H^k(\zeta^{k,m}) - G_H^k(\zeta^{k,0}), \quad \forall g \in \partial G(x^k). \quad (3.18)$$

Sumando las desigualdades (3.17) y (3.18), y nuevamente usando la Proposición 4.1.1 junto a la definición de la dirección d , (3.11), conseguimos la siguiente desigualdad

$$\langle g, d \rangle + \langle \nabla F(x^k), d \rangle + \frac{1}{2\tau} \sum_{j=0}^{m-1} \|\zeta^{k,j+1} - \zeta^{k,j}\|^2 \leq 0, \quad \forall g \in \partial G(x^k).$$

Tomando el supremo sobre $g \in \partial G(x^k)$, llegamos al resultado (3.12). \square

Corolario 3.2.2. *Bajo los mismos supuestos del Teorema 3.2.1 la dirección d , dado por (3.11), es una dirección de descenso para la función J del problema (3.6), es decir,*

$$J'(x^k; d) < 0.$$

Demostración. El resultado se obtiene fácilmente de la caracterización de la derivada direccional de una función convexa dada por el Lema 2.1.12 en el resultado del Teorema 3.2.1. \square

El siguiente resultado nos garantiza que podemos realizar una búsqueda lineal para la función J del problema (3.44) mediante la regla de Armijo.

Teorema 3.2.3. *Sea k la iteración de transición. Supongamos $\zeta^{k,0} = I_h^H x^k$ no es un mínimo local del problema en la malla gruesa (3.7). Usando el Algoritmo 1 para aproximar la solución del problema (3.7), donde w_H^k viene dado por (3.4), con tamaño de paso $\tau \leq \min\{L^{-1}, L_H^{-1}\}$ y $m > 1$ iteraciones como criterio de parada. Para todo $0 < \delta < 1$, existe $\beta > 0$ tal que*

$$J(x^k + \beta d) < J(x^k) + \delta \beta J'(x^k; d) \quad (3.19)$$

donde d viene dado por la expresión (3.11).

Demostración. Debido al Corolario 3.2.2 sabemos que d , dada por la expresión (3.11), es una dirección de descenso para J , es decir,

$$J'(x^k; d) < 0. \quad (3.20)$$

Recordemos que la definición de la derivada direccional¹ de una función en el punto x^k en una dirección d es

$$J'(x^k, d) = \lim_{\beta \searrow 0} \frac{J(x^k + \beta d) - J(x^k)}{\beta}.$$

Equivalentemente podemos escribir: para todo $\varepsilon > 0$, existe $\beta_\varepsilon > 0$ tal que

$$\left| \frac{J(x^k + \beta d) - J(x^k)}{\beta} - J'(x^k, d) \right| < \varepsilon,$$

cuando $\beta \in (0, \beta_\varepsilon)$. Esto implica que

$$J(x^k + \beta d) - J(x^k) < \beta\varepsilon + \beta J'(x^k, d),$$

cuando $\beta \in (0, \beta_\varepsilon)$. Puesto que la derivada direccional satisface (3.20), podemos tomar $\delta \in (0, 1)$ y $\varepsilon = (\delta - 1) J'(x^k, d) > 0$ de modo que obtengamos

$$J(x^k + \beta d) - J(x^k) < \delta\beta J'(x^k, d).$$

Reordenando la última desigualdad obtenemos la regla de Armijo (3.19). □

3.3. Análisis Importante

Si bien pudimos realizar la búsqueda lineal, mediante la regla de Armijo (3.19), para validar el Algoritmo Forward-Backward Multigrid necesitamos que el elemento resultante de la búsqueda lineal, notado como $x^{k+1} = x^k + \beta d$, satisfaga una de las siguientes propiedades:

(I) Monotonía de Fejér respecto a $[\partial J]^{-1}(0)$, es decir, si $x^* \in [\partial J]^{-1}(0)$

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2.$$

(II) Quasi-monotonía de Fejér respecto a $[\partial J]^{-1}(0)$, es decir, si $x^* \in [\partial J]^{-1}(0)$

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 + \gamma.$$

(III) Quasi-monotonía de la función $J^* = J + \frac{1}{2r} \|\cdot - x^*\|^2$, respecto a x^k , tal que

¹Ver el Lema 2.1.11.

$x^* \in [\partial J]^{-1}(0)$. En otras palabras,

$$J(x^{k+1}) + \frac{1}{2\tau} \|x^{k+1} - x^*\|^2 \leq J(x^k) + \frac{1}{2\tau} \|x^k - x^*\|^2 + \theta_k,$$

donde $\theta \geq 0$ sea sumable.

Puesto que la monotonía de Fejér es una de las propiedades del método forward-backward [7], deseáramos que x^{k+1} satisfaga la propiedad (I). Hay varios estudios acerca de la monotonía de Fejér junto a la búsqueda lineal aplicada al algoritmo del gradiente proyectado [8, 13, 1]. Sin embargo, al usar una dirección de descenso distinta a menos el gradiente de la función F no podemos garantizar ninguna de las propiedades (I) y (II). Lo que nos deja la propiedad (III) como última opción.

Afortunadamente, si podemos mostrar la propiedad (III). Sin embargo, necesitamos reemplazar la función J por $J^k = J + \frac{\theta}{2\tau} \|\cdot - x^k\|^2$ en la regla de Armijo (3.19), donde $\theta > 0$.

Teorema 3.3.1. *Bajo los mismos supuestos del Teorema 3.2.3, existe $\beta > 0$, tal que*

$$J^k(x^k + \beta d) < J^k(x^k) + \delta\beta(J^k)'(x^k; d), \quad (3.21)$$

equivalente a

$$J(x^k + \beta d) + \frac{\theta}{2\tau} \|\beta d\|^2 < J(x^k) + \delta\beta J'(x^k; d). \quad (3.22)$$

Demostración. Por un lado, Basta demostrar que la derivada direccional de J^k y J coinciden en x^k en la dirección p . Evidentemente, se tiene

$$(J^k)'(x^k; d) = J'(x^k; d)$$

ya que $J^k(x^k) = J(x^k)$. El resultado se obtiene aplicando el mismo razonamiento del Teorema 3.2.3. Por otro lado, usando la definición de la función J^k

$$J^k(x) = J(x) + \frac{\theta}{2\tau} \|x - x^k\|^2. \quad (3.23)$$

y operando en la regla de Armijo (3.21) llegamos a la expresión (3.22). \square

Algo que tenemos que identificar en este punto es que la función G_H^k puede tomar varias formas, siempre y cuando satisfaga la condición de coherencia (3.8). Por ejemplo, al tener la función G podemos construir $G_H : X_H \rightarrow \mathbb{R}$ de tal forma que podamos

escribir

$$G_H^k(\zeta) := G_H(\zeta) + E(\zeta) \quad (3.24)$$

donde $E : X_H \rightarrow \mathbb{R}$ es una función convexa, propia, semicontinua inferiormente, tal que satisface la condición de coherencia (3.8).

Observación 3.3.1. *Al poder escoger diferentes tipos de expresiones para G_H , la dirección de descenso, d dado por (3.11), cambia para cada problema en la malla gruesa. La ventaja de esto se aprecia numéricamente, ya que podemos escoger G_H de tal forma que sea sencillo de implementar y nos de buenos resultados en cuanto a la dirección de descenso.*

Independientemente como construyamos la función G_H , podemos obtener un resultado general de la propiedad (III) para el problema de supresión de ruido en una imagen.

Teorema 3.3.2. *Sean $G, F : X \rightarrow \mathbb{R}$ convexas, propias y semicontinuas inferiormente, tal que F es G -diferenciable con gradiente L -Lipschitz. Además, suponemos que las funciones $G_H^k, F_H : X_H \rightarrow \mathbb{R}$ también son convexas, propias y semicontinuas inferiormente, tal que F_H es G -diferenciable con gradiente L_H -Lipschitz. Sean $x^* \in [\partial J]^{-1}(0)$, donde $J := G + F$, la dirección $d \in X$ dada por la expresión (3.11) y la regla de Armijo (3.22). Para $\theta_k > 1$ tal que*

$$\sum_{k \in \mathbb{N}} \frac{1}{\theta_k - 1} < +\infty,$$

cumple

$$J(x^{k+1}) + \frac{1}{2\tau} \|x^{k+1} - x^*\|^2 - \frac{1}{2\tau(\theta_k - 1)} \|x^k - x^*\|^2 < J(x^k) + \frac{1}{2\tau} \|x^k - x^*\|^2 \quad (3.25)$$

donde $x^{k+1} = x^k + \beta d$ para $\beta > 0$.

Demostración. Tenemos para todo $\lambda \in (0, 1)$

$$\frac{\theta}{2\tau} \|x^{k+1} - x^k\|^2 = \frac{\theta\lambda}{2\tau} \|x^{k+1} - x^k\|^2 + \frac{\theta(1-\lambda)}{2\tau} \|x^{k+1} - x^k\|^2. \quad (3.26)$$

Reordenando la identidad de tres puntos (2.2) para x^k, x^{k+1}, x^* logramos

$$\frac{1}{2} \|x^{k+1} - x^k\|^2 = \langle x^k - x^{k+1}, x^k - x^* \rangle + \frac{1}{2} \|x^{k+1} - x^*\|^2 - \frac{1}{2} \|x^k - x^*\|^2. \quad (3.27)$$

Reemplazando (3.27) en (3.26) llegamos a

$$\begin{aligned} \frac{\theta}{2\tau} \|x^{k+1} - x^k\|^2 &= \frac{\theta\lambda}{2\tau} \|x^{k+1} - x^k\|^2 + \frac{\theta(1-\lambda)}{\tau} \langle x^k - x^{k+1}, x^k - x^* \rangle \\ &\quad + \frac{\theta(1-\lambda)}{2\tau} \|x^{k+1} - x^*\|^2 - \frac{\theta(1-\lambda)}{2\tau} \|x^k - x^*\|^2. \end{aligned} \quad (3.28)$$

Usamos la desigualdad de Young en la forma

$$|\langle a, b \rangle| \leq \frac{1}{2\mu} \|a\|^2 + \frac{\mu}{2} \|b\|^2,$$

para $\mu > 0$, en (3.28) obtenemos

$$\begin{aligned} \frac{\theta}{2\tau} \|x^{k+1} - x^k\|^2 &\geq \frac{\theta\lambda}{2\tau} \|x^{k+1} - x^k\|^2 - \frac{\theta(1-\lambda)}{2\tau\mu} \|x^k - x^*\|^2 - \frac{\theta\mu(1-\lambda)}{2\tau} \|x^{k+1} - x^k\| \\ &\quad + \frac{\theta(1-\lambda)}{2\tau} \|x^{k+1} - x^*\|^2 - \frac{\theta(1-\lambda)}{2\tau} \|x^k - x^*\|^2. \end{aligned}$$

Tomando $\mu = \frac{\lambda}{1-\lambda}$ y simplificando, logramos en la última desigualdad

$$\frac{\theta}{2\tau} \|x^{k+1} - x^k\|^2 \geq \frac{\theta(1-\lambda)}{2\tau} \|x^{k+1} - x^*\|^2 - \frac{\theta(1-\lambda)}{2\lambda\tau} \|x^k - x^*\|^2. \quad (3.29)$$

Luego, fijando $\theta = \theta_k > 1$ y $\lambda = 1 - \frac{1}{\theta_k}$ en (3.29), logramos

$$\frac{\theta_k}{2\tau} \|x^{k+1} - x^k\|^2 \geq \frac{1}{2\tau} \|x^{k+1} - x^*\|^2 - \frac{1}{2\tau} \left(1 + \frac{1}{\theta_k - 1}\right) \|x^k - x^*\|^2. \quad (3.30)$$

Reemplazando (3.30) en la regla de Armijo (3.22), reordenando obtenemos (3.25). \square

Definimos el algoritmo de búsqueda lineal con la regla de Armijo.

Algorithm 2 Búsqueda Lineal para el FBMG (caso general)

- 1: Escoja $0 < \beta$ y fije $\delta, \rho \in (0, 1)$, $\theta > 1$
 - 2: **while** $J(x^k + \beta d) + \frac{\theta}{2\tau} \|\beta d\|^2 > J(x^k) + \delta\beta J'(x^k; d)$ **do**
 - 3: $\beta \leftarrow \beta\rho$
 - 4: **end while**
 - 5: **return** $y^{k+1} = y^k + \beta d$
-

3.4. Algoritmo Forward-Backward Multigrid

Damos a conocer el Algoritmo Forward-Backward Multigrid. Con la intención de aligerar la carga en la notación del algoritmo, consideremos los siguientes índices: $n_i = \kappa i$ y $m_i = \kappa(i + 1)$.

Algorithm 3 Forward-Backward Multigrid con dos niveles (FBMG)

- 1: Inicialización: x^0 , $\tau \leq \min \{L^{-1}, L_H^{-1}\}$, $\mathbb{N} \ni N, \kappa, m > 2$.
- 2: **for** $i = 0, \dots, N - 1$ **do**
- 3: **for** $k = n_i, \dots, m_i - 1$ **do**
- 4: $x^{k+1} = \text{prox}_G(x^k - \tau \nabla F(x^k))$
- 5: **end for**
- 6: Fije

$$\begin{aligned} \zeta^{m_i,0} &= I_h^H x^{m_i}, \\ w_H^{m_i} &= I_h^H \nabla F(x^{m_i}) - \nabla F_H(\zeta^{m_i,0}), \end{aligned}$$

- 7: **for** $j = 0, \dots, m - 1$ **do**
- 8: $\zeta^{m_i,j+1} = \text{prox}_{G_H^k}(\zeta^{m_i,j} - \tau(\nabla F_H(\zeta^{m_i,j}) + w_H^{m_i}))$,
- 9: **end for**
- 10: Fije $d = I_H^h(\zeta^{m_i,m} - \zeta^{m_i,0})$.
- 11: Emplear el Algoritmo 2 para obtener

$$x^{m_i+1} = x^{m_i} + \beta d,$$

para algún $\beta > 0$.

- 12: **end for**
-

Observación 3.4.1. Dadas las funciones G, F del problema (3.1), es posible construir las funciones F_H y G_H^k , donde G_H^k está dado por la expresión (3.24), con G_H dado por el usuario².

La flexibilidad del Algoritmo 3 permite empezar con una corrección en la malla gruesa, simplemente cambiando de posición los pasos 3-5 luego de los pasos 6-11.

Teorema 3.4.1. Sean $G, F : X \rightarrow \mathbb{R}$ convexas, propias y semicontinuas inferiormente, tal que F es G -diferenciable con gradiente L -Lipschitz. Además, suponemos que las funciones $G_H^k, F_H : X_H \rightarrow \mathbb{R}$ también son convexas, propias y semicontinuas inferiormente, tal que F_H es G -diferenciable con gradiente L_H -Lipschitz. Supongamos que $[\partial J]^{-1}(0) \neq \emptyset$, donde $J := G + F$, junto al problema en la malla gruesa (3.7) con G_H^k y F_H dadas por el usuario. Suponemos, también, que usamos el Algoritmo 2

²Ver la Sección 3.3

para la búsqueda lineal en la línea 11 del Algoritmo 3. Si

$$\sup_{i \in \mathbb{N}} \|x^{i(\kappa+1)} - x^*\|^2 < \infty, \quad \sum_{i \in \mathbb{N}} \frac{1}{\theta_{i\kappa} - 1} < \infty,$$

donde κ es la iteración de transición en el Algoritmo 3 y $\theta_{i\kappa}$ dado por el Teorema 3.3.2; además $x^* \in [\partial J]^{-1}(0)$ y $x^0 \in X$ es el dato inicial, el Algoritmo 3 satisface

$$J(\tilde{x}^N) - J(x^*) \leq \frac{1}{2\tau N\kappa} \left(\|x^0 - x^*\|^2 + \sup_{i \in \mathbb{N}} \|x^{i(\kappa+1)} - x^*\|^2 \sum_{i=0}^{N-1} \frac{1}{\theta_{i\kappa} - 1} \right), \quad (3.31)$$

$$J(x^{N(\kappa+1)}) - J(x^*) \leq \frac{1}{2\tau N\kappa} \left(\|x^0 - x^*\|^2 + \sup_{i \in \mathbb{N}} \|x^{i(\kappa+1)} - x^*\|^2 \sum_{i=0}^{N-1} \frac{1}{\theta_{i\kappa} - 1} \right), \quad (3.32)$$

donde

$$\tilde{x}^N = \frac{1}{N\kappa} \sum_{i=0}^{N-1} \sum_{\substack{k=i(\kappa+1) \\ k \neq \kappa-1}}^{i(\kappa+1)+\kappa} x^{k+1} \quad (3.33)$$

Demostración. En primer lugar, consideremos la i -ésima iteración del Algoritmo 3. Puesto que usamos el algoritmo derivado del método forward-backward (Algoritmo 1) en los pasos 3-4, se cumple la ecuación implícita,

$$0 \in \tau [\partial G(x^{k+1}) + \nabla F(x^k)] + (x^{k+1} - x^k), \quad (3.34)$$

para todo $k = i(\kappa + 1), \dots, i(\kappa + 1) + \kappa - 1$.

Al igual que en la demostración del Teorema (3.2.1) usaremos la desigualdad del sub diferencial de G y la desigualdad de tres puntos suave (2.3) para la función F , obteniendo

$$\langle \partial G(x^{k+1}) + \nabla F(x^k), x^{k+1} - x^* \rangle \geq J(x^{k+1}) - J(x^*) - \frac{L}{2} \|x^{k+1} - x^k\|^2 \quad (3.35)$$

Testeamos (3.34) en $\langle \cdot, x^{k+1} - x^* \rangle$ y usamos (3.35) junto a la identidad de tres puntos (2.2), logrando

$$\left(\frac{1}{2\tau} - \frac{L}{2} \right) \|x^{k+1} - x^k\|^2 + \frac{1}{2\tau} \|x^{k+1} - x^*\|^2 \leq J(x^*) - J(x^{k+1}) + \frac{1}{2\tau} \|x^k - x^*\|^2$$

Ya que $\tau \leq \min \{L^{-1}, L_H^{-1}\}$, se tiene

$$J(x^{k+1}) - J(x^*) + \frac{1}{2\tau} \|x^{k+1} - x^*\|^2 \leq \frac{1}{2\tau} \|x^k - x^*\|^2, \quad (3.36)$$

para todo $k = i(\kappa + 1), \dots, i(\kappa + 1) + \kappa - 1$. Además, del Teorema 3.3.2 sabemos que, reordenando

$$\begin{aligned} J(x^{i(\kappa+1)+\kappa+1}) - J(x^{i(\kappa+1)+\kappa}) + \frac{1}{2\tau} \|x^{i(\kappa+1)+\kappa+1} - x^*\|^2 \\ < \frac{1}{2\tau} \|x^{i(\kappa+1)+\kappa} - x^*\|^2 + \frac{1}{2\tau(\theta_{i\kappa} - 1)} \|x^{i(\kappa+1)+\kappa} - x^*\|^2. \end{aligned} \quad (3.37)$$

Sumando sobre $k = i(\kappa + 1), \dots, i(\kappa + 1) + \kappa - 1$ en (3.36) junto a (3.37), se sigue

$$\begin{aligned} \sum_{\substack{k=i(\kappa+1) \\ k \neq \kappa-1}}^{i(\kappa+1)+\kappa} [J(x^{k+1}) - J(x^*)] + \frac{1}{2\tau} \|x^{i(\kappa+1)+\kappa+1} - x^*\|^2 \\ \leq \frac{1}{2\tau} \|x^{i(\kappa+1)} - x^*\|^2 + \frac{1}{2\tau(\theta_{i\kappa} - 1)} \|x^{i(\kappa+1)+\kappa} - x^*\|^2. \end{aligned} \quad (3.38)$$

Sin embargo, al usar la monotonía de Fejér dada por el algoritmo derivado del método forward-backward³, esto es,

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2$$

para todo $k = i(\kappa + 1), \dots, i(\kappa + 1) + \kappa - 1$, obtenemos de (3.38)

$$\begin{aligned} \sum_{\substack{k=i(\kappa+1) \\ k \neq \kappa-1}}^{i(\kappa+1)+\kappa} [J(x^{k+1}) - J(x^*)] + \frac{1}{2\tau} \|x^{i(\kappa+1)+\kappa+1} - x^*\|^2 \\ \leq \frac{1}{2\tau} \|x^{i(\kappa+1)} - x^*\|^2 + \frac{1}{2\tau(\theta_{i\kappa} - 1)} \|x^{i(\kappa+1)} - x^*\|^2. \end{aligned} \quad (3.39)$$

Nuevamente, sumando sobre $i = 0, \dots, N - 1$ en (3.39) obtenemos

$$\begin{aligned} \sum_{i=0}^{N-1} \sum_{\substack{k=i(\kappa+1) \\ k \neq \kappa-1}}^{i(\kappa+1)+\kappa} [J(x^{k+1}) - J(x^*)] + \sum_{i=0}^{N-1} \frac{1}{2\tau} \|x^{i(\kappa+1)+\kappa+1} - x^*\|^2 \\ \leq \sum_{i=0}^{N-1} \frac{1}{2\tau} \|x^{i(\kappa+1)} - x^*\|^2 + \sum_{i=0}^{N-1} \frac{1}{2\tau(\theta_{i\kappa} - 1)} \|x^{i(\kappa+1)} - x^*\|^2, \end{aligned} \quad (3.40)$$

pero, $i(\kappa + 1) + \kappa + 1 = (i + 1)(\kappa + 1)$. Por tanto, obtenemos una suma telescópica de

³Ver Lema 2.2.1.

la siguiente forma

$$\frac{1}{2\tau} \sum_{i=0}^{N-1} \left[\|x^{i(\kappa+1)+\kappa+1} - x^*\|^2 - \|x^{i(\kappa+1)} - x^*\|^2 \right] = \frac{1}{2\tau} \|x^{N(\kappa+1)} - x^*\|^2 - \frac{1}{2\tau} \|x^0 - x^*\|^2,$$

donde $(N-1)(\kappa+1) + \kappa + 1 = N(\kappa+1)$. Analizando de esta manera la desigualdad (3.40) obtenemos

$$\begin{aligned} & \sum_{i=0}^{N-1} \sum_{\substack{k=i(\kappa+1) \\ k \neq \kappa-1}}^{i(\kappa+1)+\kappa} [J(x^{k+1}) - J(x^*)] + \frac{1}{2\tau} \|x^{N(\kappa+1)} - x^*\|^2 \\ & \leq \frac{1}{2\tau} \|x^0 - x^*\|^2 + \sum_{i=0}^{N-1} \frac{1}{2\tau(\theta_{i\kappa} - 1)} \|x^{i(\kappa+1)} - x^*\|^2, \\ & \leq \frac{1}{2\tau} \|x^0 - x^*\|^2 + \frac{1}{2\tau} \sup_{i \in \mathbb{N}} \|x^{i(\kappa+1)} - x^*\|^2 \sum_{i=0}^{N-1} \frac{1}{\theta_{i\kappa} - 1}, \end{aligned} \quad (3.41)$$

De donde, usando la desigualdad de Jensen en la función J de la última desigualdad (3.41) a la vez que la definición de \tilde{x}^N en (3.33) y simplificando llegamos a (3.31).

Finalmente, para obtener (3.32) basta con mostrar que

$$J(x^{k+1}) \leq J(x^k), \quad \forall i(\kappa+1) \leq k \leq i(\kappa+1) + \kappa, \quad \forall 0 \leq i \leq N-1. \quad (3.42)$$

Por un lado, sabemos que el algoritmo derivado del método forward-backward genera una sucesión monótona⁴ respecto a la función J , es decir, la monotonía (3.42) se cumple para todo $i(\kappa+1) \leq k \leq i(\kappa+1) + \kappa - 1$ y para todo $0 \leq i \leq N-1$. Sin embargo, ya que realizamos la búsqueda lineal (3.22) es evidente que

$$J(x^{i(\kappa+1)+\kappa+1}) < J(x^{i(\kappa+1)+\kappa-1}),$$

para todo $0 \leq i \leq N-1$. Por tanto, (3.42) se cumple para todas las iteraciones generadas por el Algoritmo 3. Regresando a la desigualdad (3.41) y aplicando la monotonía de la función J respecto a las iteraciones generadas, obtenemos

$$\begin{aligned} & N\kappa (J(x^{N(\kappa+1)}) - J(x^*)) + \frac{1}{2\tau} \|x^{N(\kappa+1)} - x^*\|^2 \\ & \leq \frac{1}{2\tau} \|x^0 - x^*\|^2 + \frac{1}{2\tau} \sup_{i \in \mathbb{N}} \|x^{i(\kappa+1)} - x^*\|^2 \sum_{i=0}^{N-1} \frac{1}{\theta_{i\kappa} - 1}. \end{aligned}$$

En conclusión, luego de simplificar en la última desigualdad logramos (3.32). \square

⁴Ver Proposición 2.2.1.

Observación 3.4.2. *El paso 11 del Algoritmo 3 permite en cierto modo “pegar” la iteración de transición con la iteración inicial del siguiente bloque iteraciones del método forward-backward (líneas 3-5 del Algoritmo 3).*

3.5. Problema de Supresión de Ruido en una Imagen

Como mencionamos en el capítulo 2, gracias a las condiciones de optimalidad dadas por el Teorema de Fenchel-Rockafellar (Lema 2.1.9) podemos recuperar la solución del problema (3.1). Es decir, si $y^* \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$ es la solución del problema dual (3.1), entonces la solución $x^* \in \mathbb{R}^{n_1 n_2}$ del problema de supresión de ruido en una imagen en su formulación primal (2.20) viene dada por la expresión

$$x^* = -\nabla^* y^* + b. \quad (3.43)$$

Recordemos que el problema de minimización que queremos resolver, asociado al problema de supresión de ruido, tiene la forma

$$\min_{y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}} J(y) = G(y) + F(y), \quad (3.44)$$

donde G y F vienen dadas por (2.24).

En primer lugar, describamos la forma de las funciones G_H y F_H . Si $y \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$, entonces

$$G_H(\zeta) = \alpha \sum_{l=1}^{N_1 N_2} \delta_{B(0, \alpha)}(\zeta_l), \quad F_H(\zeta) = \frac{1}{2} \|\nabla^* \zeta - I_h^H b\|^2. \quad (3.45)$$

Proponemos en el presente proyecto de investigación el problema en la malla gruesa para el problema de supresión de ruido en una imagen (3.44)

$$\min_{\zeta \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}} J_H(\zeta) = G_H^k(\zeta) + F_H(\zeta) + \langle w_H^k, \zeta \rangle, \quad (3.46)$$

donde w_H^k viene dado por (3.4) con $\zeta^{k,0} = I_h^H y^k$, igualmente

$$G_H^k(\zeta) = G_H(\zeta) + \delta_\Omega(\zeta) \quad (3.47)$$

tal que

$$\Omega = \prod_{l=1}^{N_1 N_2} \Omega_l, \quad \Omega_l = \{[\zeta^{k,0}]_l + \zeta_l : \langle \zeta_l, \rho \rangle \leq 0, \quad \forall \rho \in \Gamma_l\}, \quad (3.48)$$

$$\Gamma_l = [I_h^H \partial G(y^k)]_l = \{[I_h^H g]_l : g \in \partial G(y^k)\}. \quad (3.49)$$

Proposición 3.5.1. *Debido a la forma en que construimos la función G_H^k , garantizamos que se cumple la condición de coherencia (3.8), con $\zeta^{k,0} = I_h^H y^k$.*

Demostración. Por un lado, sean $l \in \{1, \dots, N_1 N_2\}$ y $\rho \in \Gamma_l$, donde Γ_l es como (3.49), cualesquiera. Queremos mostrar que $\rho \in \partial \delta_{\Omega_l}(\zeta_l^{k,0})$ lo cual es equivalente a mostrar que

$$\langle \rho, z - \zeta_l^{k,0} \rangle \leq 0,$$

para todo $z \in \mathbb{R}^2$ tal que

$$\langle z - \zeta_l^{k,0}, q \rangle \leq 0, \quad q \in \Gamma_l.$$

Sin embargo, la equivalencia nos da directamente que

$$\Gamma_l \subseteq \partial \delta_{\Omega_l}(\zeta_l^{k,0}).$$

Por consiguiente hemos mostrado que

$$\Gamma_l \subseteq \partial \delta_{\Omega_l}(\zeta_l^{k,0}), \quad \forall l = 1, \dots, N_1 N_2.$$

Por otro lado, por la forma de G_H , dado por (3.45), sabemos que

$$\{0\} \subset [\partial G_H(\zeta^{k,0})]_l, \quad \forall l = 1, \dots, N_1 N_2.$$

De modo que

$$\Gamma_l = \{0\} + \Gamma_l \subseteq [\partial G_H(\zeta^{k,0})]_l + \partial \delta_{\Omega_l}(\zeta_l^{k,0}), \quad \forall l = 1, \dots, N_1 N_2.$$

Seguidamente usamos el Lema 2.1.5 para poder concluir que

$$I_h^H \partial G(y^k) \subset \prod_{l=1}^{N_1 N_2} \Gamma_l \subseteq \prod_{l=1}^{N_1 N_2} [\partial G_H(\zeta^{k,0})]_l + \partial \delta_{\Omega_l}(\zeta_l^{k,0}) \subseteq \partial G_H(\zeta^{k,0}) + \partial \delta_{\Omega}(\zeta^{k,0}).$$

Luego, usando la regla de la suma para los sub diferenciales obtenemos

$$\begin{aligned} I_h^H \partial G(y^k) &\subset \partial G_H(\zeta^{k,0}) + \partial \delta_\Omega(\zeta^{k,0}) \\ &\subseteq \partial(G_H + \delta_\Omega)(\zeta^{k,0}) \\ &= \partial G_H^k(\zeta^{k,0}). \end{aligned}$$

□

La Proposición 3.5.1 nos sugiere que podemos tomar distintas formas para la función G_H^k que dependen de G_H , ya que en el Teorema 3.4.1 la acotación depende del elemento encontrado por la búsqueda lineal en el Algoritmo 2.

3.5.1. Caso General

Observe que la formulación dual del problema de supresión de ruido en una imagen (3.44) verifica las condiciones del Teorema 3.4.1 fácilmente. Resumimos este hecho en el siguiente corolario.

Corolario 3.5.1. *El problema de supresión de ruido en una imagen en su formulación (3.44), junto al problema en la malla gruesa dado por la expresión (3.46), donde F_H y G_H^k vienen dados por (3.45) y (3.47), respectivamente, tal que G_H está dado por el usuario⁵. Además, asumimos que los operadores de restricción y prolongación vienen dado por la Definición 4.1.1. Si $\theta_{i\kappa} = 1 + a^i$, donde $a > 1$ y $y^* \in [G + F]^{-1}(0)$*

$$J(\tilde{y}^N) - J(y^*) \leq \frac{1}{2\tau N\kappa} \left(\|y^0 - y^*\|^2 + \frac{4a(n_1 n_2 \alpha)^2}{a-1} \right), \quad (3.50)$$

$$J(y^{N(\kappa+1)}) - J(y^*) \leq \frac{1}{2\tau N\kappa} \left(\|y^0 - y^*\|^2 + \frac{4a(n_1 n_2 \alpha)^2}{a-1} \right), \quad (3.51)$$

donde

$$\tilde{y}^N = \frac{1}{N\kappa} \sum_{i=0}^{N-1} \sum_{\substack{k=i(\kappa+1) \\ k \neq \kappa-1}}^{i(\kappa+1)+\kappa} y^{k+1}. \quad (3.52)$$

Demostración. Es sencillo verificar que las funciones G y F , dadas por (2.24), del problema de supresión de ruido en una imagen (3.44) son propias, convexas y semicontinuas inferiormente. Asimismo, las funciones G_H , F_H y G_H^k , dadas por (3.45) y (3.47) respectivamente, son propias, convexas y semicontinuas inferiormente.

⁵Puede considerarse $G_H = 0$ o como en la expresión (3.45)

De igual manera, es elemental verificar que F y F_H son G -diferenciables, tales que su gradiente sea Lipschitz con constante, de acuerdo a [5], $L = L_H = \sqrt{8}$.

Luego, podemos calcular

$$\sup_{i \in \mathbb{N}} \|y^{i(\kappa+1)} - y^*\|^2,$$

donde κ es la iteración de transición. Efectivamente, observe que la sucesión $\{y^{i(\kappa+1)}\}_{i \in \mathbb{N}}$ consiste del dato inicial y^0 y las iteraciones dadas por la búsqueda lineal. Gracias al Teorema 3.3.2, existe $\beta_{i(\kappa+1)} > 0$ tal que $y_j^{i(\kappa+1)} \in B(0, \alpha)$, para todo $j = 1, \dots, n_1 n_2$. Por tanto,

$$\begin{aligned} \|y^{i(\kappa+1)} - y^*\|^2 &\leq 2 \left(\|y^{i(\kappa+1)}\|^2 + \|y^*\|^2 \right), \\ &= 2 \left(\left(\sum_{j=1}^{n_1 n_2} \|y_j^{i(\kappa+1)}\|_2 \right)^2 + \left(\sum_{j=1}^{n_1 n_2} \|y_j^*\| \right)^2 \right), \\ &\leq 2 \left((n_1 n_2 \alpha)^2 + (n_1 n_2 \alpha)^2 \right), \\ &= 4 (n_1 n_2 \alpha)^2. \end{aligned}$$

También tenemos

$$\sum_{i=0}^{N-1} \frac{1}{\theta_{i\kappa} - 1} = \sum_{i=0}^{N-1} \frac{1}{1 + a^i - 1} = \sum_{i=0}^{N-1} \left(\frac{1}{a} \right)^i \leq \sum_{i \in \mathbb{N}} \left(\frac{1}{a} \right)^i = \frac{a}{a-1}.$$

De donde, usando el Teorema 3.4.1 obtenemos (3.50) y (3.51). \square

3.5.2. Caso Especial

En esta sección proponemos una expresión de G_H que nos ayudará a demostrar que la Propiedad (III). Pero tenemos que cambiar levemente el dominio de la función G_H , dada por la expresión en (3.45), y también tenemos que normalizar la dirección de descenso (3.11).

Cambiamos el dominio de G_H por

$$\prod_{l=1}^{N_1 N_2} B(0, \alpha_l), \quad \text{con} \quad \alpha_l = \left\| \zeta_l^{k,0} \right\|_2, \quad (3.53)$$

con $\zeta^{k,0} = I_h^H y^k$. Por tanto, la función G_H que vamos a usar en la función G_H^k (3.47)

del problema en la malla gruesa (3.46) queda

$$G_H(\zeta) = \sum_{l=1}^{N_1 N_2} \alpha_l \delta_{B(0, \alpha_l)}(\zeta). \quad (3.54)$$

Normalizar la dirección d no tiene ninguna complicación ni altera los resultados del Corolario 3.2.2, el Teorema 3.2.3 y el Teorema 3.3.1. Resumimos este resultado en el siguiente teorema.

Teorema 3.5.2. *Bajo los mismos supuestos del Corolario 3.2.2, el Teorema 3.2.3 y el Teorema 3.3.1. Fijando $\theta = 1$, para la dirección de descenso normalizada*

$$p = \frac{d}{\|d\|}, \quad (3.55)$$

existe $\beta > 0$ tal que

$$J'(y^k; p) < 0, \quad (3.56)$$

$$J(y^k + \beta p) < J(y^k) + \delta \beta J'(y^k; p), \quad (3.57)$$

$$J(y^k + \beta p) + \frac{1}{2\tau} \|\beta p\|^2 < J(y^k) + J'(y^k; p). \quad (3.58)$$

donde $0 < \delta < 1$.

Demostración. Gracias a la homogeneidad de la derivada direccional con respecto a escalares no negativos, tenemos

$$J'(y^k; p) = \frac{1}{\|d\|} J'(y^k; d) < 0$$

La demostración de la regla de Armijo (3.57) es similar a la demostración del Teorema (3.2.3), usando (3.56) en lugar de (3.20). Finalmente, ya que

$$J^{k'}(y^k; p) = J'(y^k; p) < 0$$

se tiene (3.58), siguiendo los mismos pasos del Teorema 3.2.3. \square

A partir de estos dos cambios vamos a demostrar que la propiedad (III) se cumple.

Teorema 3.5.3. *Sean $y^* \in [\partial J]^{-1}(0)$, la dirección $p \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_1}$ dada por la expresión (3.55) y la regla de Armijo (3.58). Si $\zeta^{k,0} = I_h^H y^k$, para $\theta_k > 0$ tal que*

$$\sum_{k \in \mathbb{N}} \theta_k < +\infty,$$

la función J del problema de supresión de ruido en su formulación dual (3.44) cumple

$$J(y^{k+1}) + \frac{1}{2\tau} \|y^{k+1} - y^*\|^2 < J(y^k) + \frac{1}{2\tau} \|y^k - y^*\|^2 + \theta_k \|y^*\| \quad (3.59)$$

donde $y^{k+1} = y^k + \beta p$ para $\beta \in (0, \min\{\beta_\varepsilon, \theta_k \tau\})$.

Demostración. En primer lugar, gracias a la forma de la función G_H en la expresión (3.54) podemos deducir la siguiente desigualdad

$$\langle p, y^k \rangle \leq 0. \quad (3.60)$$

Efectivamente, de la propiedad de los operadores, restricción e interpolación en la Proposición 4.1.1 a la par de la definición de las direcciones p en (3.55) y d en (3.11), tenemos

$$\langle p, y^k \rangle = \frac{1}{\|d\|} \langle d, y^k \rangle = \frac{1}{\|d\|} \langle I_H^h(\zeta^{k,m} - \zeta^{k,0}), y^k \rangle = \frac{v}{\|d\|} \langle \zeta^{k,m} - \zeta^{k,0}, I_h^H y^k \rangle.$$

Además, si reemplazamos la definición del dato inicial para el problema en la malla gruesa, $\zeta^{k,0} = I_h^H y^k$, a la par con la desigualdad de Cauchy–Schwarz, obtenemos

$$\langle p, y^k \rangle \leq \frac{v}{\|d\|} \left(\|\zeta^{k,m}\| \|\zeta^{k,0}\| - \|\zeta^{k,0}\|^2 \right).$$

El cálculo de la norma de $\zeta^{k,m}$ y $\zeta^{k,0}$ nos da una expresión interesante de la última desigualdad. Esto es,

$$\|\zeta^{k,m}\| = \sum_{l=1}^{N_1 N_2} \|\zeta_l^{k,m}\| \leq \sum_{l=1}^{N_1 N_2} \alpha_l, \quad \|\zeta^{k,0}\| = \sum_{l=1}^{N_1 N_2} \|\zeta_l^{k,0}\| = \sum_{l=1}^{N_1 N_2} \alpha_l,$$

donde hemos usado a la expresión de la función G_H en (3.54). Por tanto, obtenemos

$$\langle p, y^k \rangle \leq \frac{v}{\|d\|} \left(\left(\sum_{l=1}^{N_1 N_2} \alpha_l \right) \left(\sum_{l=1}^{N_1 N_2} \alpha_l \right) - \left(\sum_{l=1}^{N_1 N_2} \alpha_l \right)^2 \right) = 0.$$

Lo cual muestra que se cumple (3.60).

En segundo lugar, de la identidad de tres puntos (2.2) para y^k, y^{k+1}, y^* a la par de

la definición de $y^{k+1} = y^k + \beta p$, con $0 < \beta < \min\{\beta_\varepsilon, \theta_k \tau\}$, tenemos

$$\begin{aligned} \frac{1}{2\tau} \|y^{k+1} - y^k\|^2 &= \frac{1}{\tau} \langle y^k - y^{k+1}, y^k - y^* \rangle + \frac{1}{2\tau} \|y^{k+1} - y^*\|^2 - \frac{1}{2\tau} \|y^k - y^*\|^2 \\ &= \frac{1}{\tau} \langle -\beta p, y^k - y^* \rangle + \frac{1}{2\tau} \|y^{k+1} - y^*\|^2 - \frac{1}{2\tau} \|y^k - y^*\|^2 \\ &\geq \frac{1}{\tau} \langle \beta p, y^* \rangle + \frac{1}{2\tau} \|y^{k+1} - y^*\|^2 - \frac{1}{2\tau} \|y^k - y^*\|^2, \end{aligned} \quad (3.61)$$

donde en la última línea usamos la desigualdad (3.60).

Finalmente, usando la desigualdad (3.61) en la regla de Armijo (3.58), obtenemos

$$J(y^{k+1}) + \frac{1}{2\tau} \|y^{k+1} - y^*\|^2 < J(y^k) + \frac{1}{2\tau} \|y^k - y^*\|^2 - \frac{\beta}{\tau} \langle p, y^* \rangle + \delta \beta J'(y^k; p).$$

Luego de emplear la desigualdad de Cauchy-Schwarz y el hecho que p es una dirección de descenso (3.56), logramos

$$\begin{aligned} J(y^{k+1}) + \frac{1}{2\tau} \|y^{k+1} - y^*\|^2 &< J(y^k) + \frac{1}{2\tau} \|y^k - y^*\|^2 + \frac{\beta}{\tau} \|y^*\|, \\ &< J(y^k) + \frac{1}{2\tau} \|y^k - y^*\|^2 + \theta_k \|y^*\|, \end{aligned}$$

donde en la última desigualdad usamos el hecho que $0 < \beta < \min\{\beta_\varepsilon, \theta_k \tau\}$. \square

Definimos el algoritmo de búsqueda lineal con la regla de Armijo, en el contexto del problema de supresión de ruido en una imagen.

Algorithm 4 Búsqueda Lineal para el FBMG (caso especial)

- 1: Escoja $0 < \beta < \theta_k \tau$, $\delta, \rho \in (0, 1)$.
 - 2: **while** $J(y^k + \beta p) + \frac{\beta^2}{2\tau} > J(y^k) + \delta \beta J'(y^k; p)$ **do**
 - 3: $\beta \leftarrow \beta \rho$
 - 4: **end while**
 - 5: **return** $y^{k+1} = y^k + \beta p$
-

Corolario 3.5.4. *Supongamos que $[\partial J]^{-1}(0) \neq \emptyset$, donde J es la función objetivo del problema de supresión de ruido en una imagen en su formulación dual (3.44), junto al problema en la malla gruesa (3.46) con G_H dada la expresión (3.54). Suponemos, también, que usamos el Algoritmo 4 para la búsqueda lineal en la línea 11 del Algoritmo 3. Si la sucesión $\{\theta_{i_\kappa}\}_{i \in \mathbb{N}}$ es sumable, es decir, existe $\hat{\theta} \in \mathbb{R}_+$ tal que*

$$\sum_{i \in \mathbb{N}} \theta_{i_\kappa} = \hat{\theta},$$

donde κ es la iteración de transición en el Algoritmo 3 y θ_{i_κ} dado por el Teorema 3.5.3;

además $y^* \in [\partial J]^{-1}(0)$ y $y^0 \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$ es el dato inicial, el Algoritmo 3 satisface

$$J(\tilde{y}^N) - J(y^*) \leq \frac{1}{2\tau N\kappa} \left(\|y^0 - y^*\|^2 + \hat{\theta} \|y^*\| \right), \quad (3.62)$$

$$J(y^{N(\kappa+1)}) - J(y^*) \leq \frac{1}{2\tau N\kappa} \left(\|y^0 - y^*\|^2 + \hat{\theta} \|y^*\| \right), \quad (3.63)$$

donde \tilde{y}^N viene dado por (3.33).

Demostración. La demostración es semejante a la demostración del Teorema 3.4.1, salvo que reemplazamos la desigualdad (3.38) por la desigualdad (3.59), dada por el Teorema 3.5.3. \square

Capítulo 4

Proyecciones

En este capítulo describiremos como calcular las proyecciones del problema en la malla gruesa (3.46) para el problema de supresión de ruido en una imagen en su formulación dual (3.44), considerando dos casos especiales de G_H . Además, escribimos la construcción de los operadores de restricción y prolongación que usaremos en la implementación del algoritmo 3.

Recordemos el problema de minimización asociado al problema de supresión de ruido en una imagen en su formulación dual:

$$\min_{y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}} \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0, \alpha)}(y_i) + \frac{1}{2} \|\nabla_h^* y - b\|^2, \quad (4.1)$$

donde

$$G(y) = \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0, \alpha)}(y_i) \quad y \quad F(y) = \frac{1}{2} \|\nabla_h^* y - b\|^2. \quad (4.2)$$

Hemos tomado ∇_h^* como el operador divergencia discretizado presentado en [5].

Durante todo el capítulo vamos a tomar la iteración de transición k y el punto inicial en el problema de la malla gruesa como

$$\zeta^{k,0} = I_h^H y^k.$$

Además, puesto que haremos uso constante de los puntos en la frontera de la bola $B(0, \alpha)$, notaremos a la frontera como

$$\text{bd}B(0, \alpha).$$

Es conveniente escribir también el problema de minimización en la malla gruesa, es

decir,

$$\min_{\zeta \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}} J_H(\zeta) = G_H^k(\zeta) + F_H(\zeta) + \langle w_H^k, \zeta \rangle \quad (4.3)$$

tal que, para cierta función $G_H : \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2} \rightarrow \mathbb{R}$ que precisaremos más adelante,

$$G_H^k(\zeta) = G_H(\zeta) + \delta_\Omega(\zeta), \quad F_H(\zeta) = \frac{1}{2} \|\nabla_H^* \zeta - I_h^H b\|^2, \quad (4.4)$$

$$w_H^k = I_h^H \nabla F(y^k) - \nabla F_H(\zeta^{k,0}) \quad (4.5)$$

y además,

$$\Omega = \prod_{l=1}^{N_1 N_2} \Omega_l, \quad \Omega_l = \left\{ \zeta_l^{k,0} + \zeta_l : \langle \zeta_l, \rho \rangle \leq 0, \quad \forall \rho \in \Gamma_l \right\}, \quad (4.6)$$

$$\Gamma_l = [I_h^H \partial G(y^k)]_l = \{ [I_h^H g]_l : g \in \partial G(y^k) \}. \quad (4.7)$$

Asimismo, ∇_H^* fue considerado como el operador divergencia discretizado en el espacio $\mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$, tal como se presenta en [5].

4.1. Operadores de Restricción y Prolongación

Para construir el problema de dimensión menor, necesitamos precisar los operadores de restricción y prolongación que vamos a usar junto a ciertas propiedades.

Definición 4.1.1. *Consideremos una partición disjunta homogénea de índices en el conjunto $\{1, 2, \dots, n_1 n_2\}$, es decir, sean $A_l \subset \{1, 2, \dots, n_1 n_2\}$, existe $v \in \mathbb{N}$ tal que*

$$|A_l| = v, \quad \bigcup_{l=1}^{N_1 N_2} A_l = \{1, \dots, n_1 n_2\}, \quad A_l \cap A_p = \emptyset, \quad \forall p, l \in \{1, 2, \dots, N_1 N_2\}.$$

Definimos el operador restricción $I_h^H : \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2} \rightarrow \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$ como

$$[I_h^H y]_l = \frac{1}{v} \sum_{j \in A_l} y_j, \quad \forall l \in \{1, \dots, N_1 N_2\}. \quad (4.8)$$

Asimismo, definimos el operador prolongación $I_H^h : \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2} \rightarrow \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$ como

$$[I_H^h \zeta]_j = \zeta_l, \quad \forall j \in A_l, \quad \forall l \in \{1, \dots, N_1 N_2\}. \quad (4.9)$$

A continuación presentamos ciertas propiedades de los operadores de restricción y prolongación, dados por (4.8) y (4.9) respectivamente.

Proposición 4.1.1. *Los operadores de restricción y prolongación, descritos según la Definición 4.1.1, satisfacen*

$$(I_h^H)^* = \frac{1}{v} I_H^h, \quad (I_H^h)^* = v I_h^H,$$

donde v es la cardinalidad de la partición homogénea.

Demostración. Sean $y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$, $\zeta \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$, cualesquiera. Luego,

$$\begin{aligned} \langle I_h^H y, \zeta \rangle &= \sum_{l=1}^{N_1 N_2} \langle [I_h^H y]_l, \zeta_l \rangle \\ &= \sum_{l=1}^{N_1 N_2} \left(\frac{1}{|A_l|} \sum_{j \in A_l} y_{j,1} \zeta_{l,1} + \frac{1}{|A_l|} \sum_{j \in A_l} y_{j,2} \zeta_{l,2} \right) \\ &= \frac{1}{v} \sum_{j \in A_l} \sum_{l=1}^{N_1 N_2} \left(y_{j,1} [I_H^h \zeta]_{j,1} + y_{j,2} [I_H^h \zeta]_{j,2} \right) \\ &= \frac{1}{v} \sum_{i=1}^{n_1 n_2} \langle y_i, [I_H^h \zeta]_i \rangle \\ &= \frac{1}{v} \langle y, I_H^h \zeta \rangle. \end{aligned}$$

Lo cual muestra que la adjunto del operador restricción satisface $(I_h^H)^* = \frac{1}{v} I_H^h$. Usando las propiedades de los operadores adjuntos en espacios de Hilbert, se obtiene $(I_H^h)^* = v I_h^H$, [14]. \square

Proposición 4.1.2. *Consideremos los operadores restricción y prolongación dados por la Definición 4.1.1. Además, sean $\alpha > 0$, $y \in \prod_{i=1}^{n_1 n_2} B(0, \alpha) \subset \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$ y $\zeta \in \prod_{l=1}^{N_1 N_2} B(0, \alpha) \subset \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$. Entonces*

$$I_h^H y \in \prod_{l=1}^{N_1 N_2} B(0, \alpha), \quad I_H^h \zeta \in \prod_{i=1}^{n_1 n_2} B(0, \alpha).$$

Demostración. Es muy sencillo verificar que si $\zeta \in \prod_{l=1}^{N_1 N_2} B(0, \alpha)$, entonces $I_H^h \zeta \in \prod_{i=1}^{n_1 n_2} B(0, \alpha)$, gracias a la definición del operador prolongación en (4.9). Por otro lado, para el operador de restricción basta verificar que

$$[I_h^H y]_l \in B(0, \alpha),$$

para todo $l \in \{1, \dots, N_1 N_2\}$. Sea $l \in \{1, \dots, N_1 N_2\}$, cualquiera. Recordemos que tenemos

la propiedad

$$\left(\sum_{i=1}^n a_i \right)^2 \leq n \sum_{i=1}^n a_i^2. \quad (4.10)$$

Así,

$$\begin{aligned} \|[I_h^H y]_l\|_2^2 &= \left(\frac{1}{|A_l|} \sum_{j \in A_l} y_{j,1} \right)^2 + \left(\frac{1}{|A_l|} \sum_{j \in A_l} y_{j,2} \right)^2 \\ &\leq \frac{1}{|A_l|^2} |A_l| \sum_{j \in A_l} y_{j,1}^2 + \frac{1}{|A_l|^2} |A_l| \sum_{j \in A_l} y_{j,2}^2 \\ &= \frac{1}{|A_l|} \left(\sum_{j \in A_l} (y_{j,1}^2 + y_{j,2}^2) \right) \\ &\leq \frac{1}{|A_l|} \sum_{j \in A_l} \alpha^2 \\ &= \alpha^2 \end{aligned}$$

Por tanto,

$$[I_h^H x]_l \in B(0, \alpha).$$

Como el índice fue arbitrario, se concluye que $I_h^H y \in \prod_{l=1}^{N_1 N_2} B(0, \alpha)$. \square

La demostración de la Proposición 4.1.2 nos permite identificar una propiedad topológica para $[I_h^H y]_l$ con respecto a $B(0, \alpha)$. Es decir, si dentro de la partición usada para el operador restricción I_h^H existen dos puntos en la bola $B(0, \alpha) \subset \mathbb{R}^2$ distintos, entonces la restricción en dicho índice de la partición es un punto interior de la bola $B(0, \alpha) \subset \mathbb{R}^2$. Escribimos este resultado en el siguiente corolario.

Corolario 4.1.1. *Dados los operadores de restricción y prolongación por la Definición 4.1.1. Además, sean $y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$ y $\alpha > 0$ tales que*

$$\begin{aligned} y_j \in B(0, \alpha), \quad \forall j \in A_l, \quad \forall l \in \{1, \dots, N_1 N_2\}, \\ \exists p, i \in A_l, \quad y_p \neq y_i. \end{aligned}$$

Entonces

$$[I_h^H y]_l \in \text{int} B(0, \alpha), \quad \forall l \in \{1, \dots, N_1 N_2\}.$$

Demostración. Sean $l \in \{1, \dots, N_1 N_2\}$ y A_l un conjunto de la partición de $\{1, \dots, n_1 n_2\}$. La demostración se basa en la propiedad (4.10); ya que si $a_i > 0$, la igualdad se

alcanza únicamente si todos los elementos son iguales. Por tanto, usando monotonía estrictamente creciente de la raíz cuadrada, obtenemos

$$\| [I_h^H y]_l \|_2 < \alpha.$$

A partir de esto, es un ejercicio elemental mostrar que $[I_h^H y]_l \in \text{int}B(0, \alpha)$. \square

4.2. Operador Proximal de G_H^k

En esta sección introduciremos dos definiciones para la función G_H en (4.4) de forma que podamos calcular el operador proximal de G_H^k . Recordemos que la función del problema en la malla gruesa no suave es

$$G_H^k(\zeta) = G_H(\zeta) + \delta_\Omega(\zeta), \quad (4.11)$$

donde $\Omega = \prod \Omega_l$ viene dado por (4.6), además, es importante recordar que Ω_l en (4.6) viene dado por

$$\Omega_l = \left\{ \zeta_l^{k,0} + \zeta_l : \langle \zeta_l, \rho \rangle \leq 0, \quad \forall \rho \in \Gamma_l \right\},$$

o también

$$\Omega_l = \left\{ \zeta_l^{k,0} \right\} + \Gamma_l^\circ,$$

donde Γ_l° es el cono polar de Γ , definido por

$$\Gamma_l^\circ = \left\{ \rho \in \mathbb{R}^2 : \langle \rho, q \rangle \leq 0, \quad \forall q \in \Gamma_l \right\}.$$

Es conocido que el operador proximal de una indicatriz sobre un conjunto convexo generaliza el concepto de la proyección sobre conjuntos convexos [7]. De modo que, nuestra intención es definir G_H como una indicatriz sobre un producto de bolas de centro cero y radio $\alpha_l > 0$.

Observe que gracias a la forma de Ω , la indicatriz de Ω es separable, es decir,

$$\delta_\Omega(\zeta) = \sum_{l=1}^{N_1 N_2} \delta_{\Omega_l}(\zeta_l),$$

de modo que para expresar el operador proximal de G_H^k nos basta analizar la estructura de Ω_l para todo $l = 1, \dots, N_1 N_2$, gracias al Lema 2.1.5.

Para el siguiente resultado necesitamos definir el conjunto de índices

$$D_l = \{j : y_j \in \text{bd}B(0, \alpha), \quad j \in A_l\}. \quad (4.12)$$

Además, debemos enfatizar que haremos constante uso de los conos convexos, por tanto, es imperativo escribir su definición:

$$\text{cone}(\{x_1, \dots, x_n\}) = \left\{ \sum_{i=1}^n \beta_i x_i : \beta_i \geq 0, \quad 1 \leq i \leq n \right\}. \quad (4.13)$$

Es preciso recordar la condición de coherencia dada por la Proposición 3.5.1, es decir,

$$I_h^H \partial G(y^k) \subset \partial G_H^k(\zeta^{k,0}).$$

También debemos mencionar el sub diferencial de G_H , dada como la indicatriz del producto de bolas de centro cero y radio α_l , es decir, si G_H está definido por

$$G_H(\zeta) = \delta_{\prod_{l=1}^{N_1 N_2} B(0, \alpha_l)}(\zeta) = \sum_{l=1}^{N_1 N_2} \delta_{B_2(0, \alpha_l)}(\zeta_l) \quad (4.14)$$

el sub diferencial asociado es:

$$[\partial G_H(\zeta)]_l = \partial \delta_{B(0, \alpha)}(\zeta_l) = \begin{cases} \{0\}, & \text{si } \zeta_l \in \text{int}(B(0, \alpha)), \\ \{\beta \zeta_l : \beta \geq 0\}, & \text{si } \zeta_l \in \text{bd}(B(0, \alpha)), \\ \emptyset & \text{caso contrario,} \end{cases}$$

para todo $l \in \{1, \dots, N_1 N_2\}$.

En cierta parte de la demostración del siguiente Teorema, vamos a usar una versión del Lema de Farkas. Por lo cual, mencionaremos como un lema la versión del Lema de Farkas que vamos a usar:

Lema 4.2.1. Sean $A \in \mathbb{R}^{m \times d}$, $b \in \mathbb{R}^m$. Exactamente, una de las dos proposiciones es verdadera:

- $\exists x \in \mathbb{R}^d$ tal que $Ax \leq b$.
- $\exists z \in \mathbb{R}^m$ tal que $z \geq 0$, $A^T z = 0$ y $z^T b < 0$.

Referimos al lector a [23] para consultar la demostración de esta versión del Lema de Farkas y profundizar en su conocimiento.

Teorema 4.2.1. Sea $G : X \longrightarrow \mathbb{R}$ dada por la expresión (4.2) y los conjuntos

$$\Omega = \prod_{l=1}^{N_1 N_2} \Omega_l, \quad \Omega_l = \left\{ \zeta_l^{k,0} \right\} + \Gamma_l^\circ, \quad (4.15)$$

$$\Gamma_l = [I_h^H \partial G(y^k)]_l = \left\{ [I_h^H g]_l : g \in \partial G(y^k) \right\}. \quad (4.16)$$

donde Γ_l° es el cono polar de Γ_l . Si usamos el operador de restricción I_h^H definido por (4.8) y el dato inicial $\zeta^{k,0} = I_h^H y^k$, entonces para todo $l \in \{1, \dots, N_1 N_2\}$ tal que $A_l \subset \{1, \dots, n_1 n_2\}$ es un conjunto de la partición de índices $\{1, \dots, n_1 n_2\}$, los conjuntos Ω_l se pueden escribir como

$$\Omega_l = \left\{ \zeta_l^{k,0} \right\} + \Gamma_l^\circ, \quad (4.17)$$

donde Γ_l° se puede escribir como

$$\Gamma_l^\circ = \begin{cases} \mathbb{R}^2, & \text{si } D_l = \emptyset, \\ \Gamma_{j,l}^\circ, & \text{si } D_l = \{j\}, \\ \Gamma_{j,l}^\circ \cap \Gamma_{n,l}^\circ, & \text{si } \exists j, n \in D_l, \Gamma_l = \text{cone}(\{y_j, y_n\}), \\ \{\beta y_j^\perp : \beta \geq 0\}, & \text{si } \exists j, n, r \in D_l, y_j = -y_n \in \Gamma_l = \text{cone}(\{y_j, y_r, y_n\}), \\ & \text{tal que } y_r \text{ no es paralelo a } y_j \text{ o a } y_n, \\ \{0\}, & \text{si } \Gamma_l = \mathbb{R}^2. \end{cases} \quad (4.18)$$

donde y_j^\perp es el vector ortogonal a y_j que no está en Γ_l , y además $\Gamma_{j,l}^\circ \subset \mathbb{R}^2$ está definido por

$$\Gamma_{j,l}^\circ = \left\{ \rho \in \mathbb{R}^2 : \langle \rho, y_j \rangle \leq 0 \right\}. \quad (4.19)$$

Demostración. Para demostrar este resultado vamos a necesitar calcular el sub diferencial de

$$G(y) = \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0,\alpha)}(y_i),$$

puesto que influye directamente en la definición del conjunto Γ_l en (4.16). Pero calcular el sub diferencial de G es sencillo gracias al Lema 2.1.5. En otras palabras, el sub

diferencial de G dado por (4.2) es

$$[\partial G(y)]_i = \partial \delta_{B(0,\alpha)}(y_i) = \begin{cases} \{0\}, & \text{si } y_i \in \text{int}(B(0,\alpha)), \\ \{\beta y_i : \beta \geq 0\}, & \text{si } y_i \in \text{bd}(B(0,\alpha)), \\ \emptyset & \text{caso contrario,} \end{cases} \quad (4.20)$$

para todo $i \in \{1, \dots, n_1 n_2\}$.

Además, si $\zeta \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$ necesitamos definir el conjunto de puntos con la propiedad

$$\langle \zeta_l, \eta \rangle \leq 0, \quad \forall \eta \in \Gamma_l,$$

donde Γ_l viene dado por (4.15). Observe que el conjunto con la propiedad mencionada corresponde al cono polar de Γ_l , es decir,

$$\Gamma_l^\circ = \{\zeta_l \in \mathbb{R}^2 : \langle \zeta_l, \eta \rangle \leq 0, \quad \forall \eta \in \Gamma_l\}, \quad (4.21)$$

para todo $l \in \{1, \dots, N_1 N_2\}$.

Sea $l \in \{1, \dots, N_1 N_2\}$ tal que $A_l \subset \{1, \dots, n_1 n_2\}$ es un conjunto de la partición de índices $\{1, \dots, n_1 n_2\}$. Dividimos la demostración por casos. El primer caso que trataremos es cuando $|D_l| = 0$. Observe que en este caso todos los puntos de la malla fina en el conjunto de índices A_l son puntos interiores de la bola $B(0, \alpha)$. Entonces Γ_l es cero. Efectivamente, si

$$y_j \in \text{int}(B_2(0, \alpha)), \quad \forall j \in A_l,$$

entonces de (4.20) concluimos que

$$[\partial G(y)]_j = \{0\}, \quad \forall j \in A_l.$$

Por tanto, de la definición del operador de restricción (4.8) deducimos que

$$\Gamma_l = \{[I_h^H g]_l : g \in \partial G(y)\} = \{0\}.$$

De modo que $\Gamma_l^\circ = \mathbb{R}^2$. Consecuentemente $\Omega_l = \mathbb{R}^2$.

El segundo caso que vamos a tratar es cuando $|D_l| = 1$. Asimismo, podemos notar que en este caso todos los puntos cuyo índice esté en el conjunto A_l , excepto uno, son puntos interiores de la bola $B(0, \alpha)$ (ver Figura 4.1). En otras palabras, existe un único

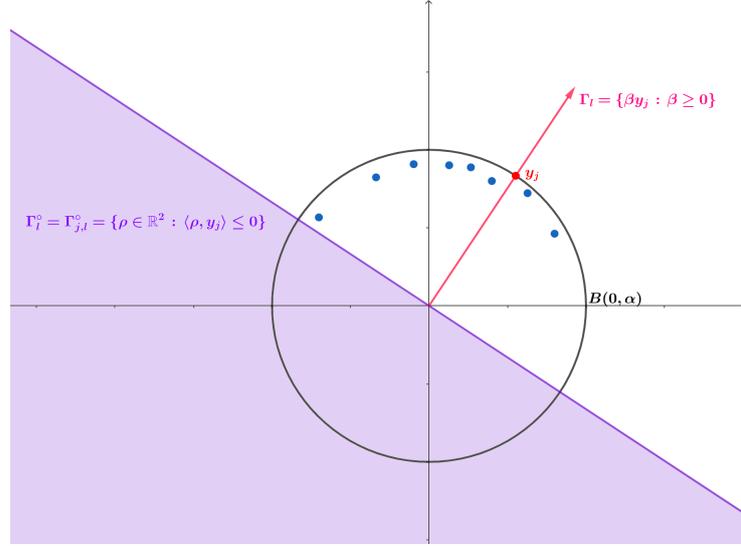


Figura 4.1: Todos los puntos con índices en el conjunto A_l son puntos interiores, excepto uno, es decir, $D_l = \{j\}$.

índice $j \in A_l$ tal que

$$y_k \in \text{int}B(0, \alpha), \quad \forall k \in A_l \setminus \{j\}, \quad y_j \in \text{bd}B(0, \alpha).$$

Sin dudas, de (4.20) aseguramos que

$$\begin{aligned} [\partial G(y)]_k &= \{0\}, \quad \forall k \in A_l \setminus \{j\}, \\ [\partial G(y)]_j &= \{\beta y_j : \beta \geq 0\}. \end{aligned}$$

Nuevamente, de la definición del operador de restricción en (4.8) concluimos que

$$\Gamma_l = \{\beta y_j : \beta \geq 0\}.$$

Por tanto, Γ_l° es el cono convexo definido por

$$\Gamma_l^\circ = \{\rho \in \mathbb{R}^2 : \langle \rho, y_j \rangle \leq 0\} = \Gamma_{j,l}^\circ.$$

Finalmente, el tercer y último caso es cuando $|D_l| > 1$. El hecho que $|D_l| > 1$ implica que existen más de un punto en la frontera de la bola $B(0, \alpha)$, cuyos índices están en el conjunto A_l . Dicho de otra manera,

$$y_k \in \text{int}B(0, \alpha), \quad \forall k \in A_l \setminus D_l, \quad y_p \in \text{bd}B(0, \alpha), \quad \forall p \in D_l.$$

De (4.20) sabemos que

$$\begin{aligned} [\partial G(y)]_k &= \{0\}, \quad \forall k \in A_l \setminus D_l, \\ [\partial G(y)]_p &= \{\beta y_p : \beta \geq 0\}, \quad \forall p \in D_l. \end{aligned}$$

Entonces el conjunto Γ_l es

$$\Gamma_l = \left\{ \sum_{p \in D_l} \beta_p y_p : \beta_p \geq 0, \quad \forall p \in D_l \right\}. \quad (4.22)$$

Dado que Γ_l podría contener elementos que se pueden construir a partir de otros elementos de Γ_l es importante eliminar dichos vectores para especificar Γ_l° . Por un lado, sí $|D_l| = 2$, es evidente que

$$\Gamma_l^\circ = \Gamma_{j,l}^\circ \cap \Gamma_{n,l}^\circ.$$

Por otro lado, si $|D_l| > 2$ debemos mencionar es que gracias a que $\Gamma_l \subseteq \mathbb{R}^2$, el cono Γ_l va a tener al menos dos vectores directores. De modo que tenemos que realizar un proceso de eliminación de aquellos vectores dependientes en Γ_l .

Para eliminar aquellos vectores dependientes de Γ_l , primeramente eliminamos todos los vectores en la frontera que estén duplicados. Seguidamente, analizamos el sistema de ecuaciones

$$\beta_j y_j + \beta_n y_n = y_r, \quad (4.23)$$

donde $j, n, r \in D_l$ y $y_j, y_n, y_r \in \text{bd}B(0, \alpha)$. Puesto que hemos eliminado los vectores duplicados, podemos garantizar que $y_j \neq y_n \neq y_r$ en el sistema de ecuaciones (4.23).

Dentro de este proceso de eliminación de los vectores redundantes o dependientes en Γ_l , dividimos el análisis en seis casos:

- i) Supongamos que la solución del sistema (4.23) es $\beta_j, \beta_n \geq 0$. Entonces $y_r \in \text{cone}(\{y_j, y_n\})$, lo cual indica que y_r es irrelevante en el conjunto Γ_l , ya que y_r se puede construir con y_j y y_n . Por tanto, eliminamos el índice r de D_l , dando como resultado

$$\Gamma_l = \left\{ \sum_{p \in D_l \setminus \{r\}} \beta_p y_p : \beta_p \geq 0, \quad \forall p \in D_l \setminus \{r\} \right\}. \quad (4.24)$$

- ii) Si la solución del sistema (4.23) es $\beta_j > 0$ y $\beta_n < 0$. Es posible reordenar el sistema de ecuaciones (4.23) de tal forma que y_j esté en el cono $\text{cone}(\{y_n, y_r\})$.

Dicho de otra manera, despejando y_j del sistema (4.23) conseguimos

$$y_j = \frac{1}{\beta_j} y_r - \frac{\beta_n}{\beta_j} y_n.$$

Lo cual muestra, claramente, que $y_j \in \text{cone}(\{y_n, y_r\})$. Asimismo, esto indica que y_j es irrelevante para Γ_l . De modo que eliminamos el índice j de D_l , dando como resultado

$$\Gamma_l = \left\{ \sum_{p \in D_l \setminus \{j\}} \beta_p y_p : \beta_p \geq 0, \quad \forall p \in D_l \setminus \{j\} \right\}. \quad (4.25)$$

- III) Ahora, si la solución del sistema (4.23) es $\beta_j < 0$ y $\beta_n > 0$. Semejante al caso anterior, podemos reordenar el sistema de ecuaciones (4.23) para verificar que $y_n \in \text{cone}(\{y_j, y_r\})$. En definitiva, debemos eliminar el índice n de D_l , consecuentemente

$$\Gamma_l = \left\{ \sum_{p \in D_l \setminus \{n\}} \beta_p y_p : \beta_p \geq 0, \quad \forall p \in D_l \setminus \{n\} \right\}. \quad (4.26)$$

- IV) Cuando la solución del sistema (4.23) es $\beta_j = 0$ y $\beta_n < 0$, podemos calcular el valor de β_n . Efectivamente, el sistema de ecuaciones (4.23) queda $\beta_n y_n = y_r$, además, ya que $y_n, y_r \in \text{bd}B(0, \alpha)$, tenemos

$$\alpha = \|y_r\| = \|\beta_n y_n\| = |\beta_n| \|y_n\| = |\beta_n| \alpha,$$

de donde concluimos que $\beta_n = -1$. Esto nos indica que el ángulo entre y_r y y_n es 180° . Puesto que todavía tenemos que realizar más comparaciones con los elementos restantes de Γ_l , debemos eliminar uno de los índices de los vectores opuestos con la intención de evitar incoherencias en las siguientes comparaciones. Por ejemplo, si no eliminamos el índice n de S_l es probable que el sistema de ecuaciones (4.23) no tenga solución en la siguiente comparación, ya que es probable que no existan $\beta_p, \beta_n \in \mathbb{R}$, tal que

$$y_{\hat{p}} = \beta_p y_p + \beta_n y_n = (\beta_p + \beta_n) y_p.$$

Así, eliminamos el índice n de D_l y guardamos los índices de los vectores opuestos

en otro conjunto notado por S_l . En otras palabras, obtenemos los conjuntos

$$\Gamma_l = \left\{ \sum_{p \in D_l \setminus \{n\}} \beta_p y_p : \beta_p \geq 0, \quad \forall p \in D_l \setminus \{n\} \right\}, \quad S_l = \{n, r\}. \quad (4.27)$$

El conjunto S_l nos ayudará a escribir explícitamente el conjunto $\bigcap_{p \in D_l} \Xi_l^p$ al final de todas las comparaciones.

- v) Asimismo, cuando la solución del sistema (4.23) es $\beta_j < 0$ y $\beta_n = 0$, podemos hacer un análisis idéntico al caso anterior para concluir que tenemos que eliminar el índice j de D_l y guardar los índices j, r en el conjunto S_l .
- vi) Por el contrario, si la solución del sistema de ecuaciones (4.23) es $\beta_j, \beta_n < 0$, es posible demostrar que $\Gamma_l = \mathbb{R}^2$. Sea $x \in \mathbb{R}^2$, cualquiera. Queremos mostrar que

$$x = \alpha_1 y_j + \alpha_2 y_n + \theta y_r,$$

pero, usando la dependencia lineal de y_r dado por la solución del sistema de ecuaciones (4.23), podemos escribir

$$x = (\alpha_1 + \theta \beta_j) y_j + (\alpha_2 + \theta \beta_n) y_n.$$

A partir de esto, nos basta mostrar que existen $\alpha_1, \alpha_2, \theta \in \mathbb{R}$ tal que

$$\phi_1 = \alpha_1 + \theta \beta_j \geq 0, \quad \phi_2 = \alpha_2 + \theta \beta_n \geq 0.$$

Consideremos la matriz A y el vector b , como

$$A = \begin{bmatrix} 1 & 0 & \beta_j \\ 0 & 1 & \beta_n \end{bmatrix}, \quad b = \begin{bmatrix} -\phi_1 \\ -\phi_2 \end{bmatrix}.$$

Observe que no existe un vector $a \in \mathbb{R}^2$ tal que satisfaga

$$a \geq 0, \quad A^T a = 0, \quad a^T b < 0,$$

ya que el kernel de A^T es cero. Por tanto, el Lema de Farkas, citado en el Lema 4.2.1, nos garantiza la existencia de $\hat{a} = [\alpha_1, \alpha_2, \theta]$ tal que

$$A\hat{a} \leq b,$$

es decir,

$$\begin{aligned}\alpha_1 + \theta\beta_j &= \phi_1 \geq 0, \\ \alpha_2 + \theta\beta_n &= \phi_2 \geq 0.\end{aligned}$$

Esto muestra que $x \in \Gamma_l$. Así, aseguramos que $\Gamma_l = \mathbb{R}^2$. Consecuentemente,

$$\Gamma_l^\circ = \{0\}.$$

Como ya hemos calculado Γ_l° , no es necesario hacer más comparaciones.

Una vez que hemos descrito todos los casos posibles en el proceso de eliminación de vectores dependientes de Γ_l podemos escribir la forma de Γ_l° . Luego de realizar $|D_l| - 2$ comparaciones, existen solo tres formas de Γ_l° :

1. Supongamos que durante las $|D_l| - 2$ comparaciones únicamente se cumplieron uno de los casos [I\)-III\)](#) en cada comparación. Entonces el conjunto de índices D_l tiene únicamente dos índices, digamos que dichos índices son j, n , ver [Figura 4.2](#). Consecuentemente, es obvio que

$$\Gamma_l^\circ = \Gamma_{j,l}^\circ \cap \Gamma_{n,l}^\circ.$$

2. Ahora supongamos que durante las $|D_l| - 2$ comparaciones uno de los casos [IV\)](#) o [V\)](#) se cumplieron. Entonces el conjunto S_l contiene los índices de los vectores opuestos. Además, luego de las $|D_l| - 2$ comparaciones, el conjunto D_l también tiene dos índices. Digamos que los conjuntos S_l y D_l se escriben como

$$S_l = \{\hat{j}, r\}, \quad D_l = \{j, n\}.$$

- 2.1. Supongamos que uno de los índices en S_l está en D_l , por ejemplo $\hat{j} = j$. En tal caso, los tres vectores con índices j, n, r forman un semiespacio que define Γ_l , ver [Figura 4.4](#). En otras palabras, aseguramos que

$$\Gamma_l = \text{cone}(\{y_j, y_n\}) \cup \text{cone}(\{y_n, y_r\}).$$

Observe que Γ_l es el semiespacio que contiene y_n . Una vez que hemos definido la estructura del conjunto Γ_l como un semiespacio, es obvio que el cono polar de Γ_l , definido en [\(4.21\)](#) se cumple únicamente para βy_j^\perp , donde $\beta \geq 0$ y y_j^\perp es perpendicular a y_j tal que $\langle y_j^\perp, g \rangle \leq 0$, para todo $g \in \Gamma_l$. Por tanto,

podemos escribir

$$\Gamma_l^\circ = \{\beta y_j^\perp : \beta \geq 0\}.$$

2.2. En el caso de que ningún índice en el conjunto S_l esté en D_l significa que uno de los vectores opuestos (con índices en S_l) se encuentra en el cono formado por los vectores cuyos índices están en D_l . Supongamos que $y_j \in \text{cone}(\{y_j, y_n\})$, lo que es equivalente a decir que existen $\beta_j, \beta_n \geq 0$ tal que

$$\beta_j y_j + \beta_n y_n = y_j. \quad (4.28)$$

Afirmamos que $\beta_j, \beta_n > 0$, ya que si no $\hat{j} \in D_l$. En efecto, ya que tomamos tres índices consecutivos para resolver el sistema de ecuaciones (4.23) y suponemos que β_j o β_n es cero, implica que

$$y_j = y_j \quad \text{o} \quad y_n = y_j$$

de donde por el caso **1)** deberíamos eliminar el índice j de D_l y conservar el índice $\hat{j} \in D_l$, lo cual es falso. De esta manera, aseguramos que la solución del sistema de ecuaciones (4.28) tiene como solución escalares positivos. Luego, debido a que en S_l están los índices de los vectores opuestos, sabemos que $y_j = -y_r$. Reemplazado en la identidad (4.28) obtenemos

$$\beta_j y_j + \beta_n y_n = -y_r,$$

de donde, si multiplicamos por menos uno a la identidad, deducimos que

$$y_r \notin \text{cone}(\{y_j, y_n\}).$$

Es decir, existen dos escalares negativos, $-\beta_j$ y $-\beta_n$, tales que

$$-\beta_j y_j - \beta_n y_n = y_r.$$

Observe que hemos llegado al caso **vi)**, por esta razón podemos garantizar que

$$\Gamma_l^\circ = \{0\}.$$

□

La demostración del anterior teorema nos permite escribir un algoritmo que devuel-

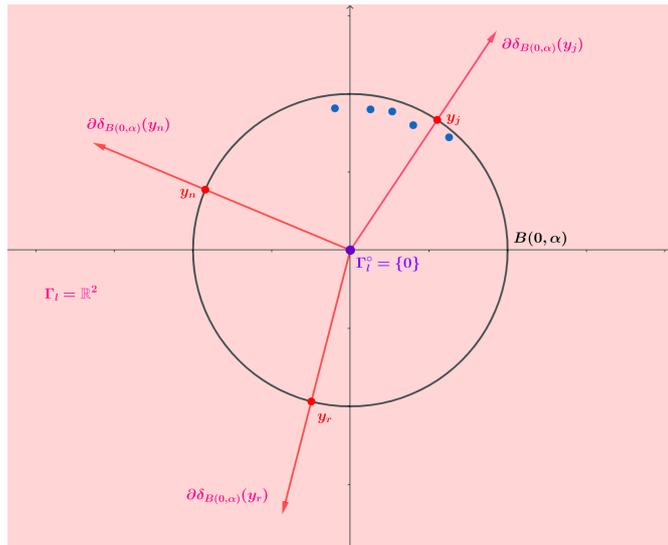


Figura 4.2: Conjunto $\Gamma_l = \mathbb{R}^2$ cuando se cumple el caso VI) del Teorema 4.2.1.

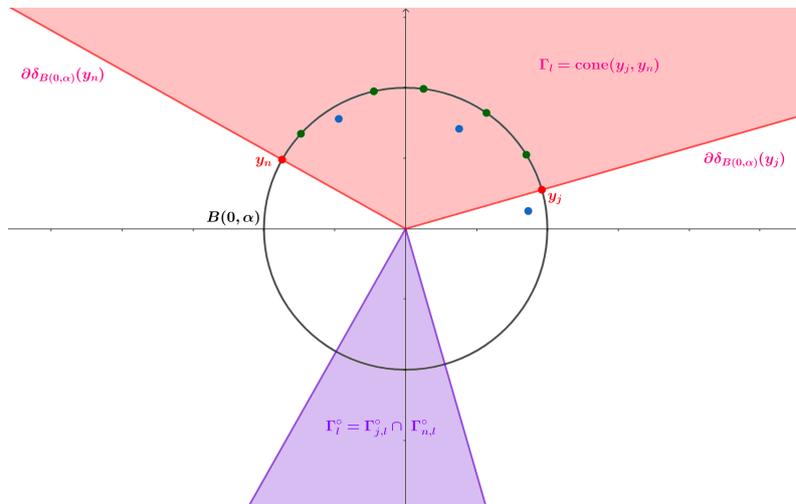


Figura 4.3: Conjunto Γ_l luego de $|D_l| - 2$ comparaciones para eliminar los vectores dependientes.

ve la estructura de Ω_l en cualquier forma del conjunto D_l .

Algorithm 5 Vector director cono convexo Ω_l (VDCC).

Require: $D_l \subset A_l$, $\zeta_l^{k,0}$, eliminación de duplicados en D_l . ▷ Pre-procesamiento.

1: Inicialización: $S_l = \emptyset$.
 2: **if** $|D_l| = 0$ **then**
 3: **return** $\Omega_l = \mathbb{R}^2$.
 4: **else if** $|D_l| = 1$ **then**
 5: **return** $\Omega_l = \left\{ \zeta_l^{k,0} \right\} + \Xi_l^j$ ▷ Suponemos que $D_l = \{j\}$.
 6: **else if** $|D_l| = 2$ **then**
 7: **return** $\Omega_l = \left\{ \zeta_l^{k,0} \right\} + \Xi_l^j \cap \Xi_l^n$ ▷ Suponemos que $D_l = \{j, n\}$.
 8: **else**
 9: **while** $|D_l| > 2$ **do**
 10: Escoja tres índices consecutivos $j, n, r \in D_l$ y resuelva el sistema de ecuaciones en \mathbb{R}^2

$$\beta_j y_j + \beta_n y_n = y_r.$$

11: **if** $\beta_j \geq 0$ **and** $\beta_n \geq 0$ **then**
 12: $D_l = D_l \setminus \{r\}$
 13: **else if** $\beta_j > 0$ **and** $\beta_n < 0$ **then**
 14: $D_l = D_l \setminus \{j\}$
 15: **else if** $\beta_j < 0$ **and** $\beta_n > 0$ **then**
 16: $D_l = D_l \setminus \{n\}$
 17: **else if** $\beta_j = 0$ **and** $\beta_n < 0$ **then**
 18: $S_l = \{n, r\}$
 19: $D_l = D_l \setminus \{n\}$
 20: **else if** $\beta_j < 0$ **and** $\beta_n = 0$ **then**
 21: $S_l = \{j, r\}$
 22: $D_l = D_l \setminus \{j\}$
 23: **else if** $\beta_j < 0$ **and** $\beta_n < 0$ **then**
 24: **return** $\Omega_l = \left\{ \zeta_l^{k,0} \right\}$
 25: **end if**
 26: **end while** ▷ Al final del lazo **while** obtenemos el conjunto D_l con dos índices.
 27: Sean $\hat{j}, r \in S_l$ ▷ Suponemos que $D_l = \{j, n\}$ hasta el final del pseudocódigo.
 28: **if** $S_l = \emptyset$ **then**
 29: **return** $\Omega_l = \left\{ \zeta_l^{k,0} \right\} + \Xi_l^j \cap \Xi_l^n$
 30: **else if** $\hat{j} \in D_l$ **o** $r \in D_l$ **then**
 31: Escribimos $\bigcap_{p \in D_l} \Xi_l^p = \left\{ \beta y_j^\perp : \langle y_j^\perp, y_j \rangle = 0, \beta \geq 0, \langle y_j^\perp, y_n \rangle \leq 0 \right\}$ o la misma expresión pero con el índice n .
 32: **return** $\Omega_l = \left\{ \zeta_l^{k,0} \right\} + \bigcap_{p \in D_l} \Xi_l^p$.
 33: **else if** $j, n \notin E_l$ **then**
 34: **return** $\Omega_l = \left\{ \zeta_l^{k,0} \right\}$
 35: **end if**
 36: **end if**

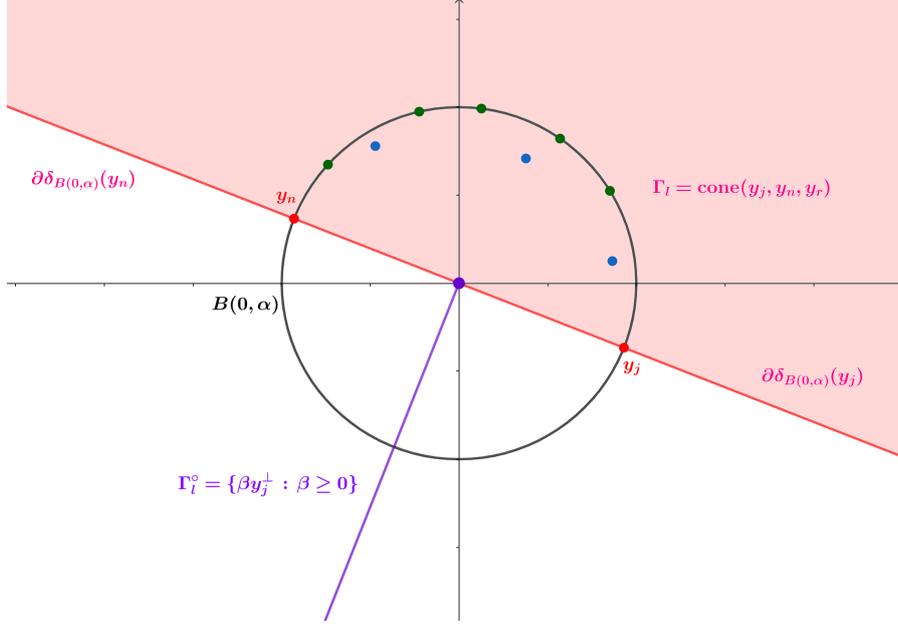


Figura 4.4: Conjunto Γ_l cuando hay dos vectores opuestos y todos los vectores en la frontera restantes estan en un mismo sub-espacio de \mathbb{R}^2 .

Recordemos que queremos expresar el operador proximal de G_H^k ,

$$G_H^k(\zeta) = G_H(\zeta) + \delta_\Omega(\zeta)$$

tomando G_H como la indicatriz del producto de bolas de centro cero y radio α_l . Es decir, la funci3n G_H^k tiene la forma

$$G_H^k(\zeta) = G_H(\zeta) + \delta_\Omega(\zeta) = \sum_{l=1}^{N_1 N_2} \delta_{B_2(0, \alpha_l)}(\zeta_l) + \sum_{l=1}^{N_1 N_2} \delta_{\Omega_l}(\zeta_l), \quad (4.29)$$

donde $\alpha_l > 0$ vamos a especificar a continuaci3n. Gracias al Corolario 4.1.1 sabemos que el dato inicial $\zeta^{k,0} \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$ es un punto interior para la bola $B(0, \alpha)$; esto es importante para definir la estructura de G_H . Ademas, sı $\Gamma_l^\circ = \Gamma_{j,l}^\circ \cap \Gamma_{n,l}^\circ$, para algun $l \in \{1, \dots, N_1 N_2\}$ y $j, n \in D_j$; ademas, si el radio de la l -esima bola es menor a la norma de la l -esima coordenada, podrıamos tener el caso que se muestra en la siguiente figura:

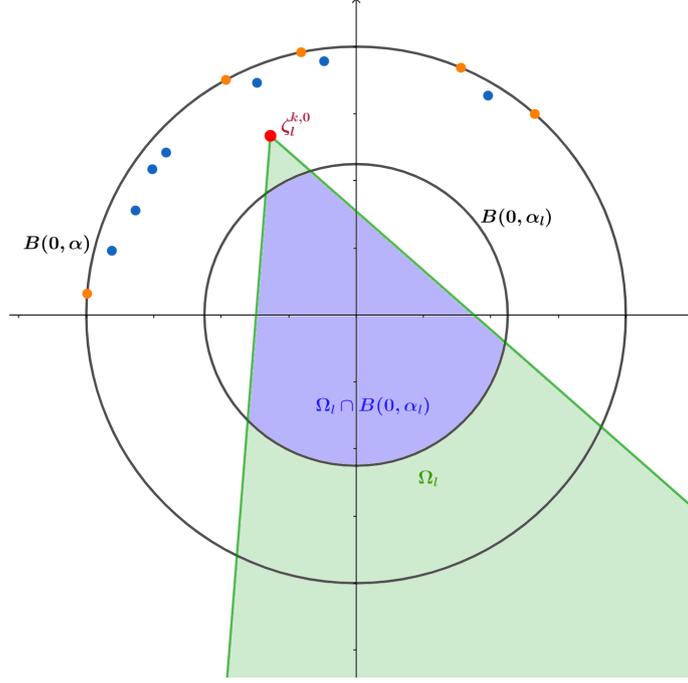


Figura 4.5: Cuatro puntos de intersección entre $\text{bd}(B(0, \alpha_l))$ y Ω_l .

Para evitar este engorroso caso, consideramos que la función G_H viene definida como

$$G_H(\zeta) = \sum_{l=1}^{N_1 N_2} \delta_{B_2(0, \alpha_l)}(\zeta), \quad (4.30)$$

donde $\alpha_l > \left\| \zeta_l^{k,0} \right\|_2$, para todo $l \in \{1, \dots, N_1 N_2\}$.

Expresaremos el operador proximal de G_H^k en sus distintas estructuras con el siguiente lema.

Lema 4.2.2. *Sea $\zeta^{k,0} = I_h^H y^k$ el dato inicial donde se construye el problema en la malla gruesa, tal que y^k es la aproximación de la solución del problema en la malla fina (4.1) luego de k iteraciones con el algoritmo derivado del método forward-backward. Además, sean G_H^k la función del problema en la malla gruesa, dada por*

$$G_H^k(\zeta) = \sum_{l=1}^{N_1 N_2} \delta_{B_2(0, \alpha_l)}(\zeta) + \sum_{l=1}^{N_1 N_2} \delta_{\Omega_l}(\zeta),$$

tal que $\alpha \geq \alpha_l > \left\| \zeta_l^{k,0} \right\|_2$, para todo $l \in \{1, \dots, N_1 N_2\}$; y D_l el conjunto de índices dado por (4.12). También usamos el escalar

$$\lambda_j = \frac{\max \left\{ 0, \left\langle \zeta_l - \zeta_l^{k,0}, y_j \right\rangle \right\}}{\|y_j\|^2}. \quad (4.31)$$

para definir el operador proximal de G_H^k en sus distintos casos. Entonces, para todo $l \in \{1, \dots, N_1 N_2\}$, el operador proximal, de la función G_H^k , asociado, se puede escribir como

$$\left[\text{prox}_{G_H^k}(\zeta) \right]_l = \begin{cases} v_1, & \text{si } D_l = \emptyset, \\ v_2, & \text{si } D_l = \{j\}, \\ v_3, & \text{si } \exists j, n \in D_l, \Gamma_l = \text{cone}(\{y_j, y_n\}), \\ v_4, & \text{si } \exists j, n, r \in D_l, y_j = -y_n \in \Gamma_l = \text{cone}(\{y_j, y_r, y_n\}), \\ & \text{tal que } y_r \text{ no es paralelo a } y_j \text{ o a } y_n, \\ v_5, & \text{si } \Gamma_l = \mathbb{R}^2. \end{cases}$$

tal que v_1, v_2, v_3, v_4, v_5 vienen dados por las expresiones (4.32a)-(4.32e), como se presentan a continuación:

$$v_1 = \frac{\alpha_l \zeta_l}{\max\{\alpha_l, \|\zeta_l\|_2\}}, \quad (4.32a)$$

$$v_2 = \begin{cases} \hat{\zeta}, & \text{si } \Phi_l^1, \\ \bar{\zeta}, & \text{si } \Phi_l^2, \\ \zeta_l - \lambda_j y_j, & \text{si } \Phi_l^3, \\ \frac{\alpha_l \zeta_l}{\max\{\alpha_l, \|\zeta_l\|_2\}}, & \text{si } \Phi_l^4, \end{cases} \quad (4.32b)$$

donde λ_j viene dado por (4.31) y

$$\begin{aligned} \Phi_l^1 &= \text{cone}(\{y_j, \hat{\zeta}\}) + \{\hat{\zeta}\} \\ \Phi_l^2 &= \text{cone}(\{y_j, \bar{\zeta}\}) + \{\bar{\zeta}\} \\ \Phi_l^3 &= \left\{ \lambda \hat{\zeta} + (1 - \lambda) \bar{\zeta} + \beta y_j : \lambda \in (0, 1), \beta \geq 0 \right\}, \\ \Phi_l^4 &= (\Phi_l^1 \cup \Phi_l^2 \cup \Phi_l^3)^C. \end{aligned}$$

Además,

$$v_3 = \begin{cases} \zeta_l^{k,0}, & \text{si } \Delta_l^1, \\ \hat{\zeta}, & \text{si } \Delta_l^2, \\ \bar{\zeta}, & \text{si } \Delta_l^3, \\ \zeta_l - \lambda_j y_j, & \text{si } \Delta_l^4, \\ \zeta_l - \lambda_n y_n, & \text{si } \Delta_l^5, \\ \frac{\alpha_l \zeta_l}{\max\{\alpha_l, \|\zeta_l\|_2\}}, & \text{si } \Delta_l^6, \end{cases} \quad (4.32c)$$

tales que λ_j, λ_n vienen dados por (4.31) y

$$\begin{aligned}\Delta_l^1 &= \Gamma_l + \{\zeta_l^{k,0}\}, \\ \Delta_l^2 &= \text{cone}\left(\{y_j, \hat{\zeta}\}\right) + \{\hat{\zeta}\} \\ \Delta_l^3 &= \text{cone}\left(\{y_n, \bar{\zeta}\}\right) + \{\bar{\zeta}\} \\ \Delta_l^4 &= \left\{\lambda\hat{\zeta} + (1-\lambda)\zeta_l^{k,0} + \beta y_j : \lambda \in (0,1), \beta \geq 0\right\}, \\ \Delta_l^5 &= \left\{\lambda\bar{\zeta} + (1-\lambda)\zeta_l^{k,0} + \beta y_n : \lambda \in (0,1), \beta \geq 0\right\}, \\ \Delta_l^6 &= (\Delta_l^1 \cup \Delta_l^2 \cup \Delta_l^3 \cup \Delta_l^4 \cup \Delta_l^5)^C.\end{aligned}$$

También escribimos

$$v_4 = \begin{cases} \zeta_l^{k,0}, & \text{si } \Pi_l^1, \\ \hat{\zeta}, & \text{si } \Pi_l^2, \\ \zeta_l^{k,0} + y_j \frac{\max\{0, \langle \zeta_l - \zeta_l^{k,0}, y_j^\perp \rangle\}}{\|y_j^\perp\|^2}, & \text{si } \Pi_l^3, \end{cases} \quad (4.32d)$$

donde

$$\begin{aligned}\Pi_l^1 &= \Gamma_l + \{\zeta_l^{k,0}\}, \\ \Pi_l^2 &= -\Gamma_l + \{\hat{\zeta}\} \\ \Pi_l^3 &= (\Pi_l^1 \cup \Pi_l^2)^C.\end{aligned}$$

y $-\Gamma_l$ es el conjunto Γ_l con escalares no positivos. Finalmente,

$$v_5 = \zeta_l^{k,0}, \quad (4.32e)$$

Demostración. Sean $\zeta \in \mathbb{R}^2 \times \mathbb{R}^{N_1 N_2}$ y $l \in \{1, \dots, N_1 N_2\}$, cualesquiera. De igual manera que en la sección anterior, nos basta analizar el operador proximal por coordenadas gracias al Lema 2.1.5.

En primer lugar, sí $D_l = \emptyset$, por el Teorema 4.2.1 sabemos que $\Omega_l = \mathbb{R}^2$. Entonces

$$\Omega_l \cap B(0, \alpha_l) = B(0, \alpha_l),$$

así, el operador proximal de G_H^k viene dado por la expresión

$$\left[\text{prox}_{G_H^k}(\zeta)\right]_l = \frac{\alpha_l \zeta_l}{\max\{\alpha_l, \|\zeta_l\|_2\}}.$$

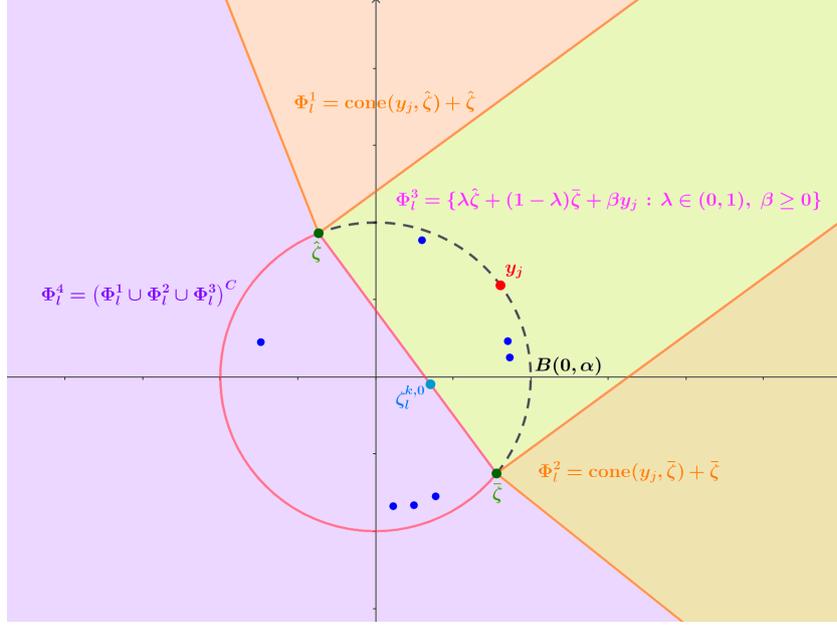


Figura 4.6: División de conjuntos para el cálculo del operador proximal de G_H^k , dado por (4.29), cuando $D_l = \{j\}$.

En segundo lugar, si $D_l = \{j\}$ del Teorema 4.2.1 sabemos que el conjunto Ω_l tiene la forma

$$\Omega_l = \zeta_l^{k,0} + \Gamma_{j,l}^\circ,$$

por tanto, $\Omega_l \cap B(0, \alpha_l)$ es una porción de la bola, $B(0, \alpha_l)$ como se muestra en la Figura 4.6. Llamamos $\hat{\zeta}, \bar{\zeta}$ a los puntos de intersección entre la frontera de la bola $B(0, \alpha_l)$ y Ω_l . Vamos a mostrar que \mathbb{R}^2 se puede particionar en cuatro conjuntos

$$\begin{aligned} \Phi_l^1 &= \text{cone}(\{y_j, \hat{\zeta}\}) + \{\hat{\zeta}\} \\ \Phi_l^2 &= \text{cone}(\{y_j, \bar{\zeta}\}) + \{\bar{\zeta}\} \\ \Phi_l^3 &= \left\{ \lambda \hat{\zeta} + (1 - \lambda) \bar{\zeta} + \beta y_j : \lambda \in (0, 1), \beta \geq 0 \right\}, \\ \Phi_l^4 &= (\Phi_l^1 \cup \Phi_l^2 \cup \Phi_l^3)^C. \end{aligned}$$

Sin embargo, en este caso es más que obvio verificar que

$$\mathbb{R}^2 = \Phi_l^1 \cup \Phi_l^2 \cup \Phi_l^3 \cup \Phi_l^4,$$

gracias a la definición de los conjuntos Φ_l^j , con $j = 1, 2, 3, 4$. Nos basta mostrar que los conjuntos son disjuntos entre sí. Es claro que Φ_l^4 es disjunto con Φ_l^1, Φ_l^2 y Φ_l^3 por su propia definición. Luego, nos basta mostrar que Φ_l^1, Φ_l^2 y Φ_l^3 son disjuntos entre sí.

Para ello es necesario especificar la forma de $\hat{\zeta}$ y $\bar{\zeta}$. Observe que un punto de

intersección entre las fronteras de Ω_l y $B(0, \alpha_l)$ debe satisfacer

$$\langle \hat{\zeta} - \zeta_l^{k,0}, y_j \rangle = 0, \quad \|\hat{\zeta}\| = \alpha_l.$$

Por tanto, un buen candidato a punto de intersección tiene que tener la forma

$$p = ay_j^\perp + \zeta_l^{k,0},$$

donde y_j^\perp es el vector ortogonal y de igual norma a y_j . Claramente, se tiene que $p \in \text{bd}\Omega_l$; nos resta encontrar la forma de a tal que satisfaga $p \in \text{bd}B(0, \alpha_l)$. Hallemos los valores de a :

$$\alpha_l^2 = \|p\|^2 = \|ay_j^\perp + \zeta_l^{k,0}\|^2 = a^2 \|y_j^\perp\|^2 + \|\zeta_l^{k,0}\|^2 + 2a \langle y_j^\perp, \zeta_l^{k,0} \rangle,$$

de donde, resolviendo la anterior ecuación cuadrática para a obtenemos que

$$a_1 = \frac{-\langle y_j^\perp, \zeta_l^{k,0} \rangle + \sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)}}{\|y_j^\perp\|^2}, \quad (4.33)$$

$$a_2 = \frac{-\langle y_j^\perp, \zeta_l^{k,0} \rangle - \sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)}}{\|y_j^\perp\|^2}. \quad (4.34)$$

Gracias a que $\alpha_l > \|\zeta_l^{k,0}\|$ concluimos que

$$\kappa = \langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right) > 0. \quad (4.35)$$

Además, podemos escribir $\hat{\zeta} = a_1 y_j^\perp + \zeta_l^{k,0}$ y $\bar{\zeta} = a_2 y_j^\perp + \zeta_l^{k,0}$. Ahora, mostremos que $\Phi_l^1 \cap \Phi_l^2 = \emptyset$. En efecto, por reducción al absurdo supongamos que existe $z \in \Phi_l^1 \cap \Phi_l^2$, lo cual es equivalente a decir que existen $\beta_1, \beta_2, \beta_3, \beta_4 \geq 0$ tal que

$$\beta_1 y_j + (\beta_2 + 1) \hat{\zeta} = z = \beta_3 y_j + (\beta_4 + 1) \bar{\zeta}, \quad (4.36)$$

de donde, si reordenamos (4.36) obtenemos

$$(\beta_1 - \beta_3) y_j = (\beta_4 + 1) \bar{\zeta} - (\beta_2 + 1) \hat{\zeta}.$$

Luego, por un lado, usando la ortogonalidad de y_j y y_j^\perp logramos

$$0 = \langle y_j^\perp, (\beta_1 - \beta_3) y_j \rangle = \langle y_j^\perp, (\beta_4 + 1) \bar{\zeta} - (\beta_2 + 1) \hat{\zeta} \rangle,$$

de donde

$$(\beta_2 + 1) \langle y_j^\perp, \hat{\zeta} \rangle = (\beta_4 + 1) \langle y_j^\perp, \bar{\zeta} \rangle. \quad (4.37)$$

Por otro lado, si expandimos el producto interno entre y_j^\perp y $\hat{\zeta}$ y usamos la definición de a_1 en (4.33), llegamos a escribir

$$\begin{aligned} \langle y_j^\perp, \hat{\zeta} \rangle &= \langle y_j^\perp, a_1 y_j^\perp + \zeta_l^{k,0} \rangle, \\ &= a_1 \|y_j^\perp\|^2 + \langle y_j^\perp, \zeta_l^{k,0} \rangle, \\ &= \frac{-\langle y_j^\perp, \zeta_l^{k,0} \rangle + \sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)}}{\|y_j^\perp\|^2} \|y_j^\perp\|^2 + \langle y_j^\perp, \zeta_l^{k,0} \rangle, \\ &= -\langle y_j^\perp, \zeta_l^{k,0} \rangle + \sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)} + \langle y_j^\perp, \zeta_l^{k,0} \rangle, \\ &= \sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)}. \end{aligned}$$

Asimismo para el producto interno entre y_j^\perp y $\bar{\zeta}$ junto a la definición de a_2 en (4.34) escribimos

$$\langle y_j^\perp, \bar{\zeta} \rangle = -\sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)}.$$

Reemplazando en (4.37) y agrupando tenemos

$$(\beta_2 + \beta_4 + 2) \sqrt{\langle y_j^\perp, \zeta_l^{k,0} \rangle^2 - \|y_j^\perp\|^2 \left(\|\zeta_l^{k,0}\|^2 - \alpha_l \right)} = 0.$$

Puesto que el argumento de la raíz cuadrada es positivo, según (4.35), la única posibilidad para que se cumpla la identidad anterior es que

$$\beta_2 + \beta_4 + 2 = 0$$

lo cual es falso, ya que β_2 y β_4 son no negativos. Esto muestra que $\Phi_l^1 \cap \Phi_l^2 = \emptyset$.

El mismo razonamiento para demostrar que $\Phi_l^1 \cap \Phi_l^2 = \emptyset$ se usa para mostrar

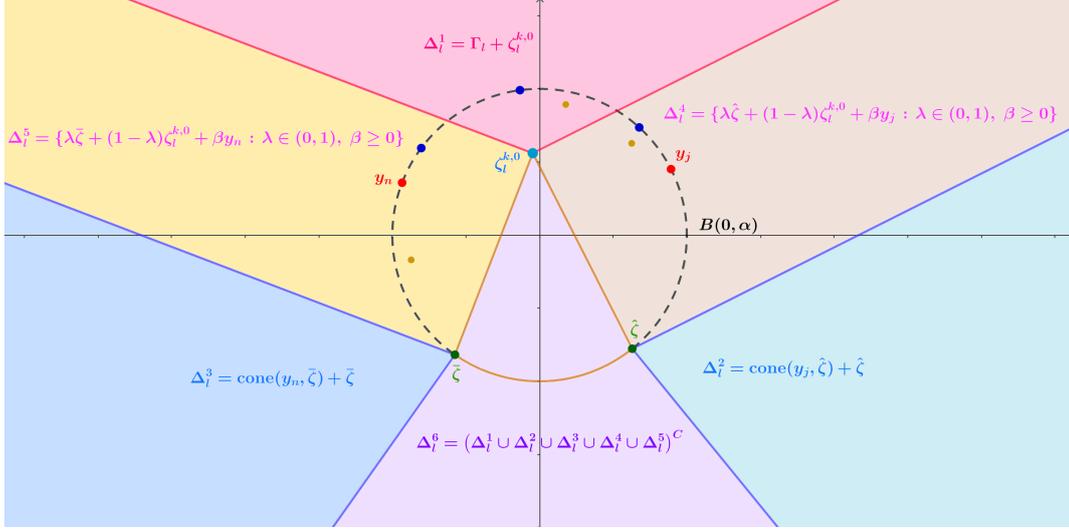


Figura 4.7: División de conjuntos para el cálculo del operador proximal de G_H^k , dado por (4.29), cuando $\Gamma_l = \text{cone}(\{y_j, y_n\})$.

$\Phi_l^1 \cap \Phi_l^3 = \emptyset$ y $\Phi_l^2 \cap \Phi_l^3 = \emptyset$, teniendo en cuenta la simetría de la combinación convexa en el conjunto Φ_l^3 . Esto muestra que los conjuntos Φ_l^j , con $j = 1, 2, 3, 4$, forman una partición de \mathbb{R}^2 , de modo que, podemos escribir el operador proximal de G_H^k como

$$\left[\text{prox}_{G_H^k}(\zeta) \right]_l = \begin{cases} \hat{\zeta}, & \text{si } \Phi_l^1, \\ \bar{\zeta}, & \text{si } \Phi_l^2, \\ \zeta_l - \lambda_j y_j, & \text{si } \Phi_l^3, \\ \frac{\alpha_l \zeta_l}{\max\{\alpha_l, \|\zeta_l\|_2\}}, & \text{si } \Phi_l^4. \end{cases}$$

En tercer lugar, si existen $j, n \in D_l$ tal que $\Gamma_l = \text{cone}(\{y_j, y_n\})$, entonces por el Teorema 4.2.1 sabemos que Ω_l tiene la forma

$$\Omega_l = \zeta_l^{k,0} + \Gamma_{j,l}^\circ \cap \Gamma_{n,l}^\circ.$$

Similar al caso anterior, el conjunto $\Omega_l \cap B(0, \alpha_l)$ es una porción de la bola $B(0, \alpha_l)$ como se muestra en la Figura 4.7.

Nuevamente, vamos a mostrar que \mathbb{R}^2 se puede particionar en los conjuntos

$$\begin{aligned}\Delta_l^1 &= \Gamma_l + \{\zeta_l^{k,0}\}, \\ \Delta_l^2 &= \text{cone}\left(\{y_j, \hat{\zeta}\}\right) + \{\hat{\zeta}\} \\ \Delta_l^3 &= \text{cone}\left(\{y_n, \bar{\zeta}\}\right) + \{\bar{\zeta}\} \\ \Delta_l^4 &= \left\{\lambda\hat{\zeta} + (1-\lambda)\zeta_l^{k,0} + \beta y_j : \lambda \in (0,1), \beta \geq 0\right\}, \\ \Delta_l^5 &= \left\{\lambda\bar{\zeta} + (1-\lambda)\zeta_l^{k,0} + \beta y_n : \lambda \in (0,1), \beta \geq 0\right\}, \\ \Delta_l^6 &= (\Delta_l^1 \cup \Delta_l^2 \cup \Delta_l^3 \cup \Delta_l^4 \cup \Delta_l^5)^C.\end{aligned}$$

Ciertamente, de la definición de Δ_l^6 concluimos que

$$\mathbb{R}^2 = \bigcup_{j=1}^6 \Delta_l^j.$$

Una vez más, nos basta mostrar que los conjuntos Δ_l^j , con $j = 1, 2, 3, 4, 5$, son disjuntos entre sí (excluimos el conjunto Δ_l^6 ya que es disjunto con todos los demás conjuntos por definición). Igualmente, al caso anterior, necesitamos precisar los puntos de intersección, lo cuales llamamos $\hat{\zeta}$ y $\bar{\zeta}$, asociados a los semiespacios tangentes a y_j y y_n , respectivamente. Ya conocemos que los puntos de intersección tienen la forma

$$\hat{\zeta} = ay_j^\perp + \zeta_l^{k,0}, \quad \bar{\zeta} = dy_n^\perp + \zeta_l^{k,0}$$

tales que

$$\langle y_j^\perp, y_n \rangle < 0, \quad \langle y_n^\perp, y_j \rangle < 0.$$

Necesitamos verificar que los puntos de intersección estén en Ω_l . Para ello, necesitamos que los vectores ortogonales y_j^\perp y y_n^\perp satisfagan, además,

$$\langle y_j^\perp, \zeta_l^{k,0} \rangle \leq 0, \quad \langle y_n^\perp, \zeta_l^{k,0} \rangle \leq 0.$$

Tomando de esta forma los vectores ortogonales y_j^\perp y y_n^\perp es sencillo verificar que los

puntos de intersección $\hat{\zeta}$ y $\bar{\zeta}$ están en Ω_l , donde los valores de a y d vienen dados por

$$a = \frac{-\langle y_j^\perp, \zeta_l^{k,0} \rangle + \sqrt{\kappa}}{\|y_j^\perp\|^2}, \quad (4.38)$$

$$d = \frac{-\langle y_n^\perp, \zeta_l^{k,0} \rangle + \sqrt{\kappa}}{\|y_n^\perp\|^2}. \quad (4.39)$$

A partir de esto, con el mismo razonamiento del caso anterior, vamos a demostrar que $\Delta_l^1 \cap \Delta_l^2 = \emptyset$. Por contradicción, supongamos que existe $z \in \Delta_l^1 \cap \Delta_l^2$, esto quiere decir que existen $\beta_1, \beta_2, \beta_3, \beta_4 \geq 0$ tal que

$$\beta_1 y_j + \beta_2 y_n + \zeta_l^{k,0} = z = \beta_3 y_j + (\beta_4 + 1) \hat{\zeta}.$$

Por un lado, si usamos la definición de $\hat{\zeta}$ y reordenando, obtenemos

$$(\beta_1 - \beta_3) y_j = (\beta_4 + 1) (a y_j^\perp + \zeta_l^{k,0}) - \beta_2 y_n - \zeta_l^{k,0} = (\beta_4 + 1) a y_j^\perp + \beta_4 \zeta_l^{k,0} - \beta_2 y_n.$$

Por otro lado, usando la ortogonalidad de y_j^\perp y y_j y reemplazando la última desigualdad junto a la definición de a en (4.38), obtenemos

$$\begin{aligned} 0 &= \langle y_j^\perp, (\beta_1 - \beta_3) y_j \rangle, \\ &= \langle y_j^\perp, (\beta_4 + 1) a y_j^\perp + \beta_4 \zeta_l^{k,0} - \beta_2 y_n \rangle, \\ &= (\beta_4 + 1) a \|y_j^\perp\|^2 + \beta_4 \langle y_j^\perp, \zeta_l^{k,0} \rangle - \beta_2 \langle y_j^\perp, y_n \rangle \\ &= (\beta_4 + 1) \left(-\langle y_j^\perp, \zeta_l^{k,0} \rangle + \sqrt{\kappa} \right) + \beta_4 \langle y_j^\perp, \zeta_l^{k,0} \rangle - \beta_2 \langle y_j^\perp, y_n \rangle \\ &= -\langle y_j^\perp, \zeta_l^{k,0} \rangle + (\beta_4 + 1) \sqrt{\kappa} - \beta_2 \langle y_j^\perp, y_n \rangle \end{aligned}$$

Sin embargo, por la forma de escoger el vector ortogonal y_j^\perp y el hecho que $\beta_4 \geq 0$ y $\kappa > 0$ podemos asegurar que

$$0 = -\langle y_j^\perp, \zeta_l^{k,0} \rangle + (\beta_4 + 1) \sqrt{\kappa} - \beta_2 \langle y_j^\perp, y_n \rangle > 0,$$

lo cual es una contradicción. Por tanto, $\Delta_l^1 \cap \Delta_l^2 = \emptyset$. Análogo al caso anterior (usando el mismo esquema de demostración junto a la elección de los vectores ortogonales y_j^\perp y y_n^\perp) y siguiendo un análisis similar, se puede mostrar que Δ_l^j , con $j = 1, 2, 3, 4, 5$, son disjuntos entre sí. Consecuentemente, los conjuntos Δ_l^j , con $j = 1, 2, 3, 4, 5, 6$, forman

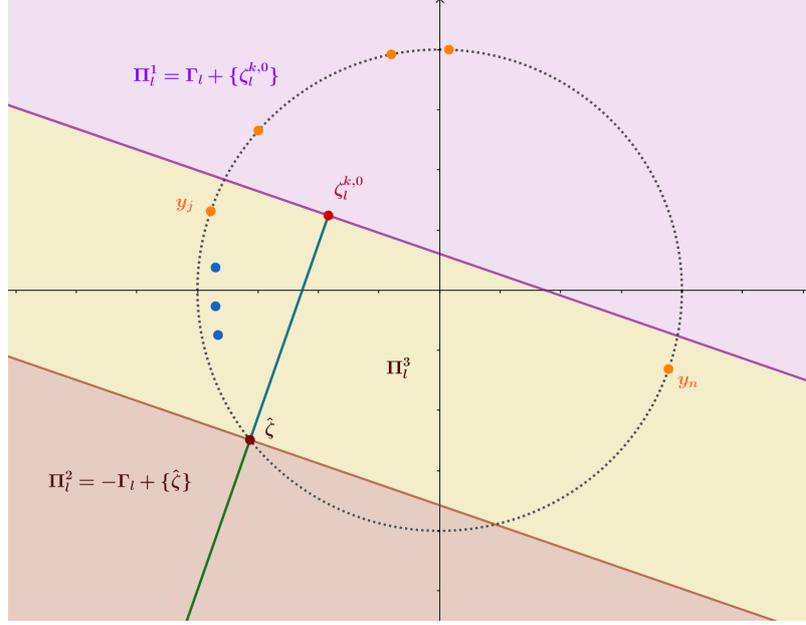


Figura 4.8: División de conjuntos para el cálculo del operador proximal de G_H^k , dado por (4.29), cuando $|D_l| > 1$ y Γ_l tiene vectores opuestos que forman un semi-espacio en \mathbb{R}^2 .

una partición de \mathbb{R}^2 ; por consiguiente, escribimos el operador proximal de G_H^k como

$$\left[\text{prox}_{G_H^k}(\zeta) \right]_l = \begin{cases} \zeta_l^{k,0}, & \text{si } \Delta_l^1, \\ \hat{\zeta}, & \text{si } \Delta_l^2, \\ \bar{\zeta}, & \text{si } \Delta_l^3, \\ \zeta_l - \lambda_j y_j, & \text{si } \Delta_l^4, \\ \zeta_l - \lambda_n y_n, & \text{si } \Delta_l^5, \\ \frac{\alpha_l \zeta_l}{\max\{\alpha_l, \|\zeta_l\|_2\}}, & \text{si } \Delta_l^6. \end{cases}$$

En cuarto lugar, si existen $j, n, r \in D_l$ tal que $\Gamma_l = \text{cone}(\{y_j, y_r, y_n\})$, entonces por el Teorema 4.2.1 sabemos que Ω_l se escribe como

$$\Omega_l = \zeta_l^{k,0} + \{\beta y_j^\perp : \beta \geq 0\},$$

donde y_j^\perp es el vector ortogonal a y_j que no está en Γ_l . De esta manera, podemos concluir que $\Omega_l \cap B(0, \alpha_l)$ es una porción de Ω_l tal como se muestra en la Figura 4.8. Reiteramos que es posible particionar \mathbb{R}^2 en tres conjuntos

$$\begin{aligned}\Pi_l^1 &= \Gamma_l + \{\zeta_l^{k,0}\}, \\ \Pi_l^2 &= -\Gamma_l + \{\hat{\zeta}\} \\ \Pi_l^3 &= (\Pi_l^1 \cup \Pi_l^2)^C,\end{aligned}$$

donde $-\Gamma$ es el conjunto Γ_l con escalares no positivos, es decir,

$$-\Gamma_l = \left\{ \sum_{j \in A_l} b_j y_j : \beta_j \leq 0, \quad \forall j \in A_l \right\}.$$

En este caso, es mucho más sencillo mostrar que \mathbb{R}^2 se puede particionar en $\Pi_l^1, \Pi_l^2, \Pi_l^3$, ya que únicamente tenemos que mostrar $\Pi_l^1 \cap \Pi_l^2 = \emptyset$.

Gracias al Teorema 4.2.1 sabemos que Γ_l es un semiespacio formado que contienen vectores opuestos, es decir,

$$\Gamma_l = \text{cone}(\{y_j, y_r, y_n\}).$$

A partir de esta definición es sencillo verificar que $\Pi_l^1 \cap \Pi_l^2 = \emptyset$ usando el mismo razonamiento del caso previo a este, y considerando que y_j^\perp satisface

$$\langle y_j^\perp, \zeta_l^{k,0} \rangle \leq 0, \quad \text{y} \quad \langle y_j^\perp, \rho \rangle \leq 0, \quad \forall \rho \in \Gamma_l.$$

De modo que \mathbb{R}^2 se puede particionar en los conjuntos Π_l^1, Π_l^2 y Π_l^3 ; consecuentemente podemos escribir el operador proximal de G_H^k como

$$\left[\text{prox}_{G_H^k}(\zeta) \right]_l = \begin{cases} \zeta_l^{k,0}, & \text{si } \Pi_l^1, \\ \hat{\zeta}, & \text{si } \Pi_l^2, \\ \zeta_l^{k,0} + y_j \frac{\max\{0, \langle \zeta_l - \zeta_l^{k,0}, y_j^\perp \rangle\}}{\|y_j^\perp\|^2}, & \text{si } \Pi_l^3. \end{cases}$$

En último lugar, sí $\Gamma_l = \mathbb{R}^2$, del Teorema 4.2.1 sabemos que el conjunto Ω_l se escribe como $\Omega_l = \zeta_l^{k,0}$. Por tanto,

$$\Omega_l \cap B(0, \alpha_l) = \zeta_l^{k,0},$$

de donde, el operador proximal de G_H^k es

$$\left[\text{prox}_{G_H^k}(\zeta) \right]_l = \zeta_l^{k,0}.$$

□

Observación 4.2.1. *Cuando $\alpha_l = \left\| \zeta_l^{k,0} \right\|$ también podemos escribir un operador proximal asociado a G_H^k , sin embargo, el análisis es mucho más extenso y Ω_l puede tomar más de cinco formas.*

Capítulo 5

Experimentación Numérica

En esta sección describiremos las instancias que usamos para comparar el algoritmo derivado del método forward-backward con su homólogo multimalla, implementado en lenguaje de programación Matlab.

5.1. Problema en la malla fina

El problema de supresión de ruido en la malla fina tienen varios parámetros que tenemos que especificar antes de presentar los resultados numéricos. Puesto que vamos a usar el algoritmo derivado del método forward-backward para aproximar la solución del problema de supresión de ruido en su formulación dual (4.1), describiremos los parámetros que usaremos durante todos los experimentos numéricos.

Para los experimentos numéricos, consideramos tres imágenes con ruido de tamaños: 300×300 , 512×768 y 768×1024 . Cada imagen corresponde al parámetro b en el problema de supresión de ruido (4.1). Además, fijamos el valor de $\tau = (4\sqrt{2})^{-1}$, el cual es el tamaño de paso que usa el Algoritmo 1, tomando en cuenta que la constante de Lipschitz continuidad del gradiente de F es $\sqrt{8}$, de acuerdo a [5].

Recordemos que la función G del problema de supresión de ruido en una imagen es

$$G(y) = \alpha \sum_{i=1}^{n_1 n_2} \delta_{B(0, \alpha)}(y_i),$$

de modo que fijamos $\alpha = 0,15$ para todas las instancias.

5.2. Problema en la malla gruesa

Debemos también precisar los parámetros del problema en la malla gruesa (4.3). En primer lugar, los operadores de transferencia, conocidos en el presente trabajo de

investigación como operadores de restricción I_h^H y prolongación I_H^h , fueron construidos con dos matrices, ya que almacenar la matriz de restricción o prolongación estándar es demasiado costoso, computacionalmente. Para usar los operadores de transferencia, numéricamente, conviene usar los datos $y \in \mathbb{R}^2 \times \mathbb{R}^{n_1 n_2}$ en la forma $y \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^2$. De este modo, el operador de restricción se puede escribir:

$$I_h^H y_1 = \frac{N_1 N_2}{n_1 n_2} M_l \cdot y_1 \cdot M_d, \quad I_h^H y_2 = \frac{N_1 N_2}{n_1 n_2} M_l \cdot y_2 \cdot M_d,$$

donde¹ y_1 es la primera matriz de tamaño $n_1 \times n_2$ de y . De igual manera, y_2 es la segunda matriz de tamaño $n_1 \times n_2$ de y . Además, $M_l = \text{Id}_{N_1} \otimes R_1$ y $M_d = \text{Id}_{N_2} \otimes R_2$; aquí \otimes representa el producto de Kronecker, Id_{N_1} es la matriz identidad de tamaño $N_1 \times N_1$ y

$$R_1 = \underbrace{[1, 1, \dots, 1]}_{n_1/N_1 \text{ veces}}, \quad R_2 = \left. \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\} n_2/N_2 \text{ veces}$$

Asimismo, el operador de prolongación se puede calcular usando la Proposición 4.1.1, de modo que si el dato en la malla gruesa $\zeta \in \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \times \mathbb{R}^2$ entonces

$$I_H^h \zeta_1 = M_l^T \cdot \zeta_1 \cdot M_d^T, \quad I_H^h \zeta_2 = M_l^T \cdot \zeta_2 \cdot M_d^T$$

De igual manera, ζ_1 y ζ_2 son las matrices de tamaño $N_1 \times N_2$ de ζ . De esta manera, construimos los datos $b_H = I_h^H b$ y $\zeta^{k,0} = I_h^H y^k$, donde este último es el punto inicial para el problema en la malla gruesa.

Puesto que queremos aproximar el problema en la malla gruesa (4.3) mediante el algoritmo derivado del método forward-backward, fijamos los parámetros del algoritmo 1 con $\tau = (4\sqrt{2})^{-1}$. Además, establecimos la dimensión del problema en la malla gruesa como: 12×12 , 64×64 y 64×128 , respectivamente a las imágenes de tamaño 300×300 , 512×768 y 768×1024 .

De igual manera, fijamos $\alpha_l = 0,15$ para la función G_H del problema en la malla gruesa (4.3), definida por

$$G_H(\zeta) = \sum_{l=1}^{N_1 N_2} \delta_{B_2(0, \alpha_l)}(\zeta_l).$$

¹Usamos la notación del lenguaje de matlab para mejor comprensión de la implementación de los operadores de transferencia.

5.2.1. Algoritmo FBMG

Una vez definidos los parámetros de los problemas en la malla fina (4.1) y en la malla gruesa (4.3), indicaremos los parámetros que usamos en el algoritmo forward-backward multimalla definido en el Algoritmo 3.

Realizamos una pequeña modificación en el algoritmo, escribiendo una sentencia **while** en lugar de la sentencia **for** en la segunda línea, con la intención de escribir un criterio de parada del algoritmo. Como criterio de parada usamos la brecha dual del problema de supresión de ruido en una imagen en la malla fina, es decir,

$$e^k = \frac{1}{2} \|x^k - b\|^2 + \alpha \|\nabla x^k\|_{2,1} + J(y^k),$$

donde J es la función objetivo del problema de minimización equivalente al problema de supresión de ruido en su formulación dual (3.44) y x^k se obtiene de usar la expresión (3.43) con y^k , es decir,

$$x^k = -\nabla^* y^k + b.$$

De esta manera, tomamos en todas las instancias el criterio de parada: $e > \text{tol}$, donde $\text{tol} = e^0 * 1 \times 10^{-5}$.

Delimitamos las iteraciones del algoritmo forward-backward en la malla fina por 4, 7, 10, 13, es decir, tomamos la iteración de transición $\kappa = 3, 6, 9, 12$, y realizamos una iteración luego de la búsqueda lineal. En otras palabras, consideramos $n_i \bmod \kappa = 0$ y $m_i \bmod \kappa = 0$, donde $n_i = \kappa i$ y $m_i = \kappa(i + 1)$. Asimismo, establecimos las iteraciones en la malla gruesa con $m = 50$ para todas las instancias.

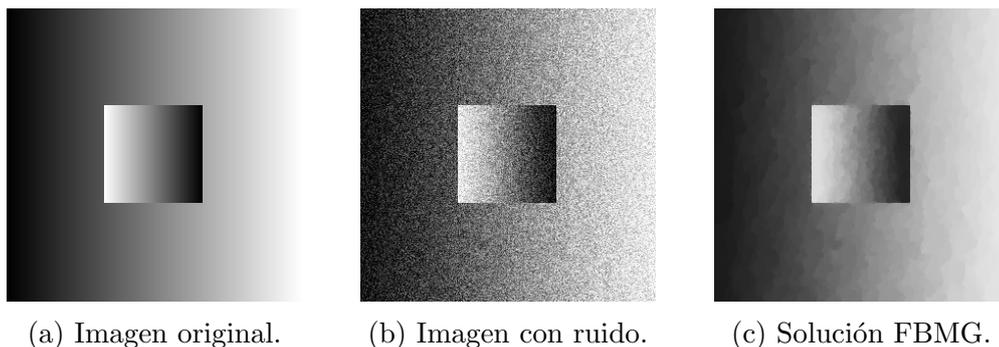
Luego, para la búsqueda lineal, fijamos $\theta_i = 1 + (1,1)^i$, donde i es la iteración del FBMG. Observe que la elección de θ_i satisface

$$\sum_{i \in \mathbb{N}} \frac{1}{\theta_i - 1} = \sum_{i \in \mathbb{N}} \frac{1}{(1,1)^i} = \sum_{i \in \mathbb{N}} \left(\frac{10}{11}\right)^i < \infty.$$

Otro aspecto importante dentro de la búsqueda lineal es la derivada direccional de G , pero sabemos que la derivada direccional de una indicatriz tiene que ser cero, siempre y cuando exista. Puesto que ya sabemos de antemano que la derivada direccional de G es finita, gracias al Lema 2.1.13, aseguramos que

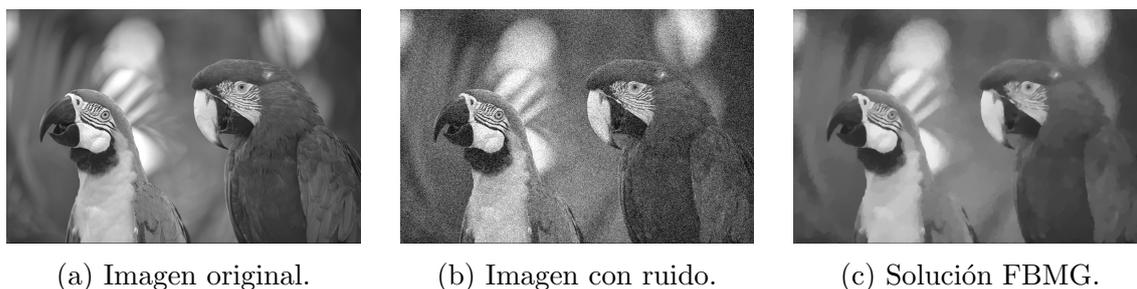
$$G(y^k; d) = \sup_{g \in \partial G(y^k)} \langle g, d \rangle = 0,$$

donde $0 = g \in \partial G(y^k)$ y d es la dirección dada por el Teorema 3.2.1. Finalmente, en la búsqueda lineal tomamos $\delta = 1 \times 10^{-3}$.



(a) Imagen original. (b) Imagen con ruido. (c) Solución FBMG.

Figura 5.1: Primera instancia con una imagen de tamaño 300×300 .



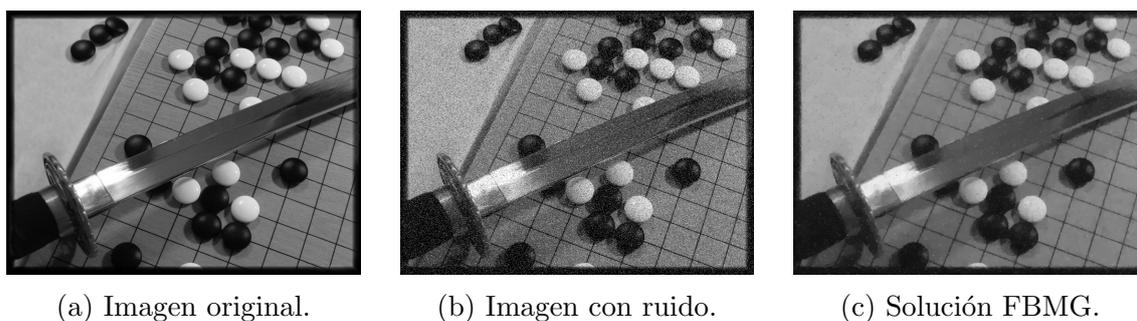
(a) Imagen original. (b) Imagen con ruido. (c) Solución FBMG.

Figura 5.2: Segunda instancia con una imagen de tamaño 512×768 .

5.3. FB vs. FBMG

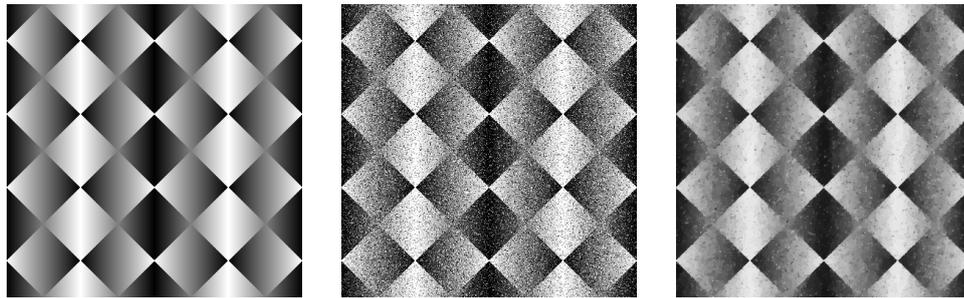
Dentro de la simulación comparamos el tiempo del CPU que se tarda en alcanzar el criterio de parada establecido. Del mismo modo, comparamos el número de iteraciones totales del algoritmo forward-backward dentro de cada iteración del algoritmo FBMG con la brecha dual alcanzada.

Comparamos el algoritmo forward-backward con su homólogo con seis instancias. En las figuras 5.1-5.6 se encuentran las seis instancias que usamos, las cuales contienen la imagen original, la imagen con ruido (al 50% en todas las imágenes) y la solución del algoritmo forward-backward multimalla. Además, en las figuras 5.7 y 5.8 comparamos el tiempo que le toma al algoritmo FBMG alcanzar una solución aproximada para las seis instancias. Como se puede observar, el algoritmo FBMG aumenta su velocidad de



(a) Imagen original. (b) Imagen con ruido. (c) Solución FBMG.

Figura 5.3: Tercera instancia con una imagen de tamaño 768×1024 .



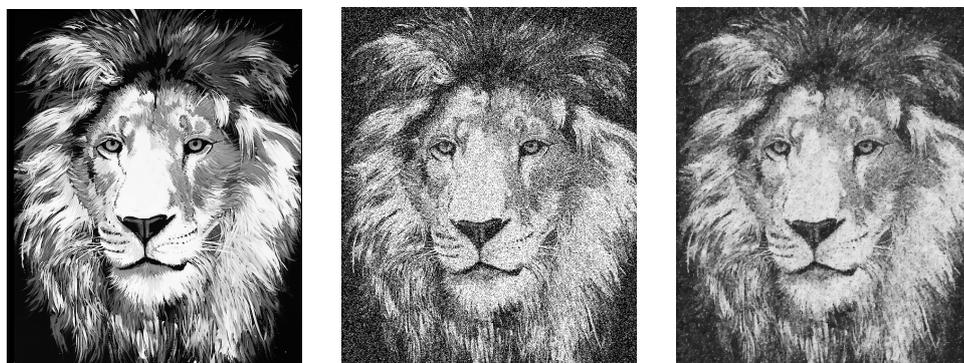
(a) Imagen original. (b) Imagen con ruido. (c) Solución FBMG.

Figura 5.4: Cuarta instancia con una imagen de tamaño 400×400 .



(a) Imagen original. (b) Imagen con ruido. (c) Solución FBMG.

Figura 5.5: Quinta instancia con una imagen de tamaño 910×602 .



(a) Imagen original. (b) Imagen con ruido. (c) Solución FBMG.

Figura 5.6: Sexta instancia con una imagen de tamaño 555×684 .

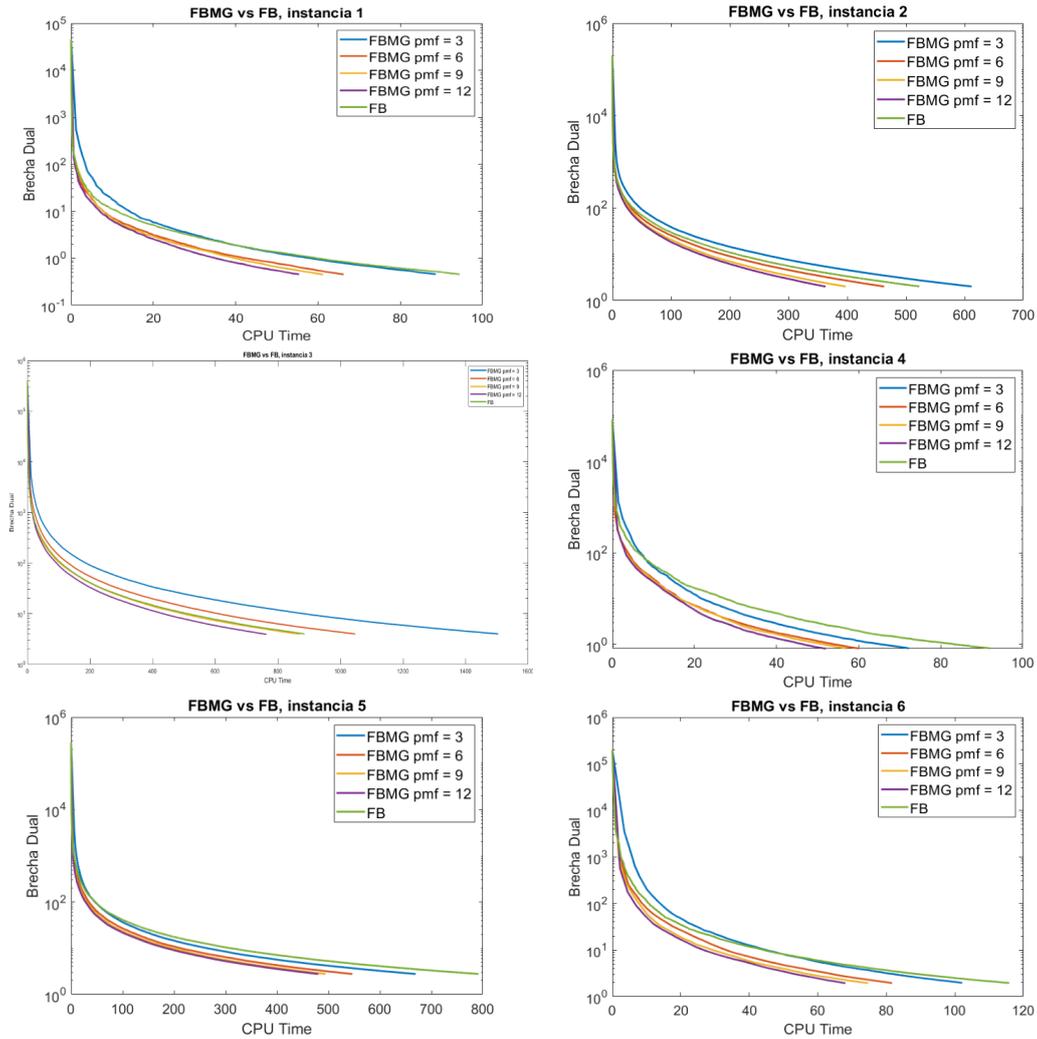


Figura 5.7: Comparación del tiempo de ejecución y la brecha dual del problema de supresión de ruido en las tres instancias propuestas.

ejecución mientras aumentan las iteraciones en la malla fina. Este fenómeno se debe al cálculo numérico de la brecha dual, ya que el lenguaje de matlab no es muy eficiente en la asignación de memoria.

La figura 5.8 nos da una idea de como es el funcionamiento del algoritmo FBMG, comparando el total de las iteraciones del algoritmo forward-backward tanto en la malla fina como en la malla gruesa.

De acuerdo a las instancias usadas para comparar el algoritmo forward-backward con su homólogo multimalla, podemos concluir que en imágenes sencillas que no contengan varias direcciones en los gradientes de la imagen el algoritmo multimalla actúa ligeramente mejor. Sin embargo, cuando la imagen presenta varias direcciones en el gradiente, el algoritmo multimalla necesita mejorar el paso de información a la malla gruesa, es decir, deberíamos usar mallas adaptativas que analicen los gradientes de la imagen (específicamente los operadores de transferencia).

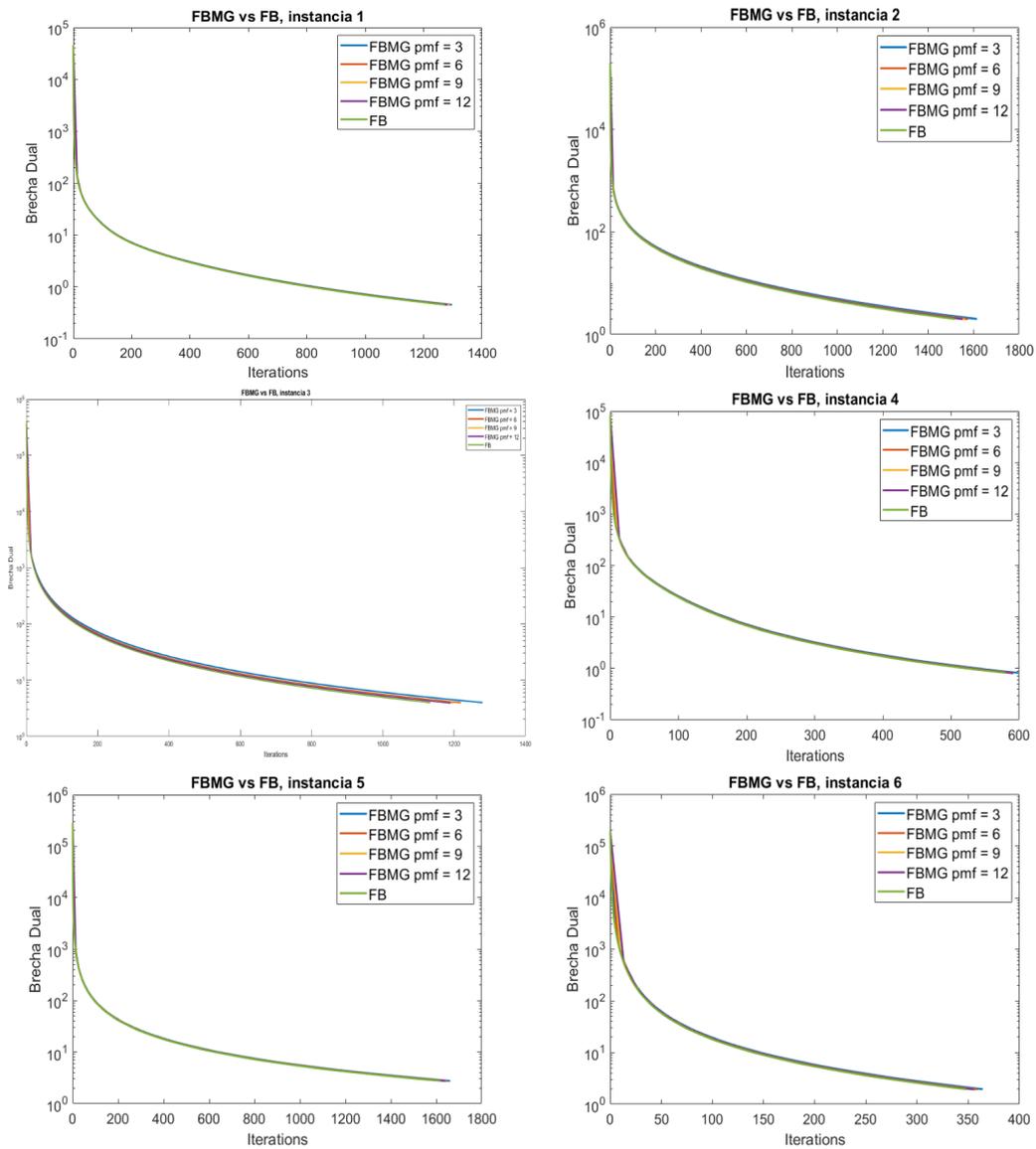


Figura 5.8: Comparación del número de iteraciones y la brecha dual del problema de supresión de ruido en las tres instancias propuestas.

Capítulo 6

Conclusiones y Recomendaciones

Diseñar un algoritmo multimalla derivado del método forward-backward fue un reto complejo que nos abrió el camino a estudiar más a fondo las propiedades del método forward-backward, la condición de coherencia para problemas no diferenciables y el diseño e implementación de un problema en menor dimensión asociado al problema de supresión de ruido en una imagen. Tras el diseño del algoritmo forward-backward multimalla pudimos demostrar la convergencia del mismo y dar ciertos resultados específicos al problema de supresión de ruido en una imagen.

En el presente trabajo de investigación hemos diseñado, creativamente, un tipo de problema de menor dimensión que nos permite encontrar una dirección de descenso (respecto a la derivada direccional de la función objetivo del problema de minimización) para el problema original, aportando una condición no suave que relaciona ambos problemas mediante la contención de la restricción del sub diferencial de la parte no suave del problema original en el sub diferencial de la función no suave del problema de menor dimensión.

La implementación numérica del algoritmo forward-backward multimalla fue todo un reto en la programación, puesto que implementar el operador proximal de G_H^k conllevó a examinar la mejor expresión de las proyecciones. Todavía es posible mejorar la implementación del algoritmo FBMG en lenguajes como Python, Julia, etc.

Los resultados obtenidos nos indican que podemos seguir mejorando el algoritmo forward-backward multimalla, tanto en la implementación como en el diseño del algoritmo. Hemos encontrado el camino inicial para acoplar una búsqueda lineal a un método no diferenciable.

En el capítulo 4 vimos la forma explícita del operador proximal de G_H^k , dado por (4.29), considerando $\alpha_l > \left\| \zeta_l^{k,0} \right\|$. Sin embargo, consideramos que analizar el caso cuando $\alpha_l = \left\| \zeta_l^{k,0} \right\|$ requiere mucho más cuidado, por las posibles intersecciones entre Ω_l y $B(0, \alpha_l)$. Esto nos deja la oportunidad de continuar innovando el algoritmo forward-backward multimalla, con aplicación en el problema de supresión de ruido en una

imagen, tanto en la parte teórica como en la parte práctica.

Un nuevo avance en el algoritmo forward-backward multimalla para el problema de supresión de ruido en una imagen, consistiría en implementar operadores de transferencia adaptivos que analicen las direcciones de los gradientes en la imagen. Implementar estos nuevos operadores no cambian la teoría de convergencia del algoritmo, sin embargo, mejoraría notablemente el rendimiento del algoritmo forward-backward multimalla

El método multimalla que hemos propuesto en el presente trabajo de investigación promueve a la comunidad de la Escuela Politécnica Nacional el interés por estudiar, investigar y diseñar nuevos métodos multimalla no suaves (o no diferenciables) con aplicaciones en el procesamiento de imágenes.

Bibliografía

- [1] H. H. Bauschke and P. L. Combettes. A weak-to-strong convergence principle for fejér-monotone methods in hilbert spaces. *Mathematics of operations research*, 26(2):248–264, 2001.
- [2] A. Beck. *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. SIAM, 2014.
- [3] K. Bredies, D. Lorenz, et al. *Mathematical Image Processing*. Springer, 2018.
- [4] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A multigrid tutorial*. SIAM, 2000.
- [5] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, 20(1-2):89–97, 2004.
- [6] T. F. Chan and J. Shen. *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. SIAM, 2005.
- [7] C. Clason and T. Valkonen. Introduction to nonsmooth analysis and optimization. *arXiv preprint arXiv:2001.00216*, 2020.
- [8] P. L. Combettes. Quasi-fejérian analysis of some optimization algorithms. In *Studies in Computational Mathematics*, volume 8, pages 115–152. Elsevier, 2001.
- [9] J. C. De los Reyes. *Numerical PDE-constrained optimization*. Springer, 2015.
- [10] J. W. Demmel. *Applied numerical linear algebra*. SIAM, 1997.
- [11] D. Gabay. Chapter ix applications of the method of multipliers to variational inequalities. In *Studies in mathematics and its applications*, volume 15, pages 299–331. Elsevier, 1983.
- [12] M. Hanke. *A taste of inverse problems: basic theory and examples*. SIAM, 2017.
- [13] A. N. Iusem. On the convergence properties of the projected gradient method for convex optimization. *Computational & Applied Mathematics*, 22:37–52, 2003.

- [14] E. Kreyszig. *Introductory functional analysis with applications*, volume 1. wiley New York, 1978.
- [15] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
- [16] J. L. Mueller and S. Siltanen. *Linear and nonlinear inverse problems with practical applications*. SIAM, 2012.
- [17] S. G. Nash. A multigrid approach to discretized optimization problems. *Optimization Methods and Software*, 14(1-2):99–116, 2000.
- [18] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.
- [19] P. Parpas. A multilevel proximal gradient algorithm for a class of composite optimization problems. *SIAM Journal on Scientific Computing*, 39(5):S681–S701, 2017.
- [20] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An algorithm for minimizing the mumford-shah functional. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1133–1140. IEEE, 2009.
- [21] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.
- [22] T. Valkonen. Preconditioned admm with nonlinear operator constraint. In *System Modeling and Optimization: 27th IFIP TC 7 Conference, CSMO 2015, Sophia Antipolis, France, June 29-July 3, 2015, Revised Selected Papers*, volume 494, page 117. Springer, 2017.
- [23] G. M. Ziegler. *Lectures on polytopes*, volume 152. Springer Science & Business Media, 2012.