ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

ELIMINACIÓN DEL RUIDO EN IMÁGENES A TRAVÉS DE UN PROBLEMA DE OPTIMIZACIÓN VARIACIONAL BINIVEL CON PARÁMETRO POLINOMIAL

PROYECTO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE MATEMÁTICO

DIEGO FRANCISCO GARZÓN QUEZADA

diego.garzon.q@gmail.com

DIRECTOR: JUAN CARLOS DE LOS REYES BUENO, PH.D.

juan.delosreyes@epn.edu.ec

Quito, Enero 2015

DECLARACIÓN

Yo DIEGO FRANCISCO GARZÓN QUEZADA, declaro bajo juramento que el trabajo aquí escrito 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 Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

DIEGO FRANCISCO GARZÓN QUEZADA

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por DIEGO FRANCISCO GARZÓN QUEZADA, bajo mi supervisión

Juan Carlos De los Reyes Bueno, Ph.D.
DIRECTOR

AGRADECIMIENTOS

En primer lugar agradezco a mi director de tesis Juan Carlos por haber hecho posible la conclusión exitosa de este trabajo.
Quisiera agradecer con infinito cariño a mi padre por haberme brindado sobre todo su sensatez y sabiduría, y a mi madre por haberme dado su fé y su coraje. Ambos me han ayudado a llegar hasta aquí y, más aún, me han llevado a comprender hasta donde puedo llegar en el futuro. Este trabajo es nuestro.
Finalmente, guardaré eterna gratitud a mis amigos que supieron acompañarme de la mejor manera posible en este largo e intenso camino hacia la madurez.

DEDICATORIA

A mis padres que son el camino que me une a la divinidad, y al resto de personas, que son el otro camino.

Índice de Contenido

Lista de Figuras

Lista de Tablas

Resumen

La presencia completamente aleatoria de ruido o distorsión en una imagen es un resultado inevitable en la adquisición de imágenes digitales, por lo que es necesario estudiar un método de disminución de este ruido. En este proyecto partimos del modelo de Variación Total que fue introducido para el control y la reconstrucción de imagenes en un articulo de 1992 por Rudin, Osher y Fatemi ya que este modelo se ha vuelto, al igual que otros modelos variacionales, clásico para los problemas de análisis de imágenes. Basados en una metodología de variación total, planteamos el problema de determinación del peso de un ruido como un problema de estimación óptima de parámetros. Consideramos en este trabajo tres tipos de peso: un parámetro real, un polinomio cuadrático de coeficientes positivos y el caso de una función en general. Debido a que el problema subyacente consiste en uno de optimización no diferenciable en dimensión infinita, la caracterización y cálculo de la solución al problema global (binivel) resultan complejos. Estudiamos el problema de manera analítica, obteniendo resultados de existencia, aproximación y caracterización a través de condiciones necesarias de primer orden, y numérica, considerando métodos cuasi-Newton en combinación con métodos de Newton generalizados.

Palabras clave: Procesamiento de Imágenes, Control Optimo, Distribución de Ruido, Optimización de EDPs con restricciones, Regularización de Huber

Abstract

The presence of noise or random distortions is an inevitable result when acquiring digital images, so it is necessary to study a method for reducing this noise.

In this project we take, as a starting point, the Total Variation Model that was introduced for the control and reconstruction of images in a 1992 article by Rudin, Osher and Fatemi, because this model has become, like other variational models, a classic for understanding image analysis problems.

Based on the Total Variation methodology, we propose the problem of finding the noise weight as a problem of optimal parameter estimation. We consider in this paper three types of weight: a real parameter, a quadratic polynomial with nonnegative coefficients and a general function. Because the underlying problem is not differentiable in the infinite dimensional case, characterization and computation of the solution to the global problem (dual) are complex.

We study the problem analytically, obtaining existence, approach and characterization results through first order necessary conditions, and numerically considering quasi-Newton methods in combination with generalized Newton methods.

Keywords: Image processing, Optimal Control, noise distribution, PDE-constrained optimization, Huber regularization

CAPÍTULO 1

El modelo de control de imágenes

El modelo de Variación Total fue introducido para el control y la reconstrucción de imagenes en un articulo de 1992 por Rudin, Osher y Fatemi [1]. Este modelo se ha vuelto, al igual que otros modelos variacionales, clásico para los problemas de análisis de imágenes. En este capítulo estudiaremos las características que llevan a que la regularización en espacios de Variación Acotada sea usada para modelar este problema. Además, para obtener continuidad en el operador solución y, por lo tanto, convergencia de los parámetros óptimos regularizados, procederemos a estudiar el problema en espacios de Hilbert.

1.1 Motivación

Una imagen digital es una representación de una imagen en una región discreta, especificada por los valores que la describen. Estos valores pueden ser descritos de manera vectorial o matricial; una imagen de descripción vectorial almacena características representadas por formas geométricas, a diferencia de una imagen matricial que almacena características representadas como un "mapa de píxeles".

Definimos entonces, una imagen como una función bidimensional $f(x, y)$ donde x e y son las coordenadas espaciales, y el valor de la función f en cualquier par de coordenadas (x, y) es el valor del píxel en dicho punto.

A partir de una imagen obtenida, denominamos ruido a cualquier píxel representado computacionalmente que no corresponda con la imagen original. La presencia, completamente aleatoria del ruido o distorsión en imágenes es un fenómeno inevitable introducido por el proceso de formación de la imagen o en el momento de archivarla, decodificarla, transmitirla, etcétera. Esta distorsión complica el tratamiento de dichas imágenes, pues incluso una cantidad "muy pequeña"de ruido puede ser perjudicial cuando altos niveles de exactitud son requeridos (por ejemplo en el análisis por píxeles de una imagen).

Se pueden identificar distintos tipos de ruido en una imagen (ver figura 1.1) dada basados en la distribución que este siga. Dependiendo del tipo de imagen, y la aplicación que esta tenga, este ruido sigue determinada distribución, así tenemos:

- Ruido Gaussiano (Tomografías de Resonancia Magnética, MRI)
- Ruido de Poisson (Tomografías de Emisión de Positrones, PET o medidas con radar)

- Ruido de Impulso (Errores de transmición o pixeles alterados en sensores de cámaras)

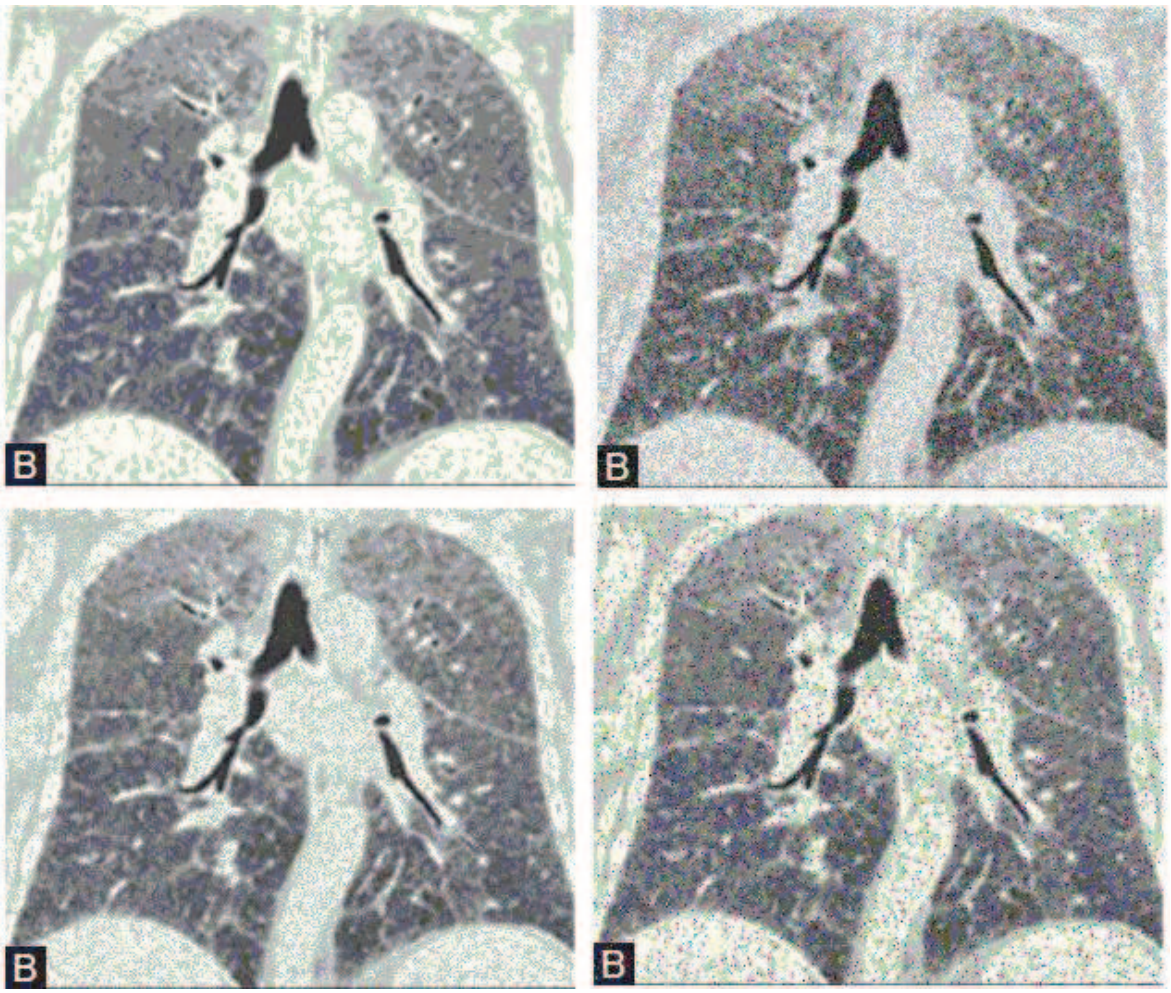

Figura 1.1: *Primera fila: Imagen Original (izquierda) e imagen con ruido Gaussiano de varianza 0.02 (derecha). Segunda fila: imagenes con ruido de Poisson (izquierda) y con ruido de impulso (derecha) Imagen original tomada de [2]*

Es común el uso de un método basado en una función de tipo gaussiano para remover el ruido de una imagen y suavizarla. Sin embargo, métodos de este tipo causan también un efecto de blur (aspecto borroso) a la imagen lo que hace que esta pierda informaciones importantes sobre los bordes de los objetos [3]

Dada la aleatoriedad de la presencia de dicho ruido, es natural pensar en que un modelo adecuado para representar dicho fenómeno deberá tener raíces estadísticas. Es así, que una estimación de mínimos cuadrados podría funcionar en muchas imágenes de distinto tipo; este proceso es dependiente de la norma en L_2. Sin embargo, debido a la naturaleza geométrica del modelamiento de imágenes, se concluye [4] que una norma apropiada para imágenes es la norma de variación total TV y no la norma en L_2.

Las normas TV son, en esencia, normas L_1 de derivadas, por lo que se pueden utilizar procesos de estimación validos en L_1 en problemas de imágenes.

Un estudio del espacio de funciones de variación acotada BV es necesario, pues gracias a sus propiedades es posible la estimación de discontinuidades en las soluciones del problema [4].

Basados en los trabajos [1] y [4] tomaremos como punto de partida para la eliminación del ruido de una imagen, al problema de minimización de la variación total de la solución estimada, donde las restricciones reflejan la distribución estadística.

El modelo desarrollado en este trabajo se encuentra motivado por el siguiente problema en restauración de imágenes:

Dada una imagen con ruido $f : \Omega \to \mathbb{R}$, donde Ω es un subconjunto abierto y acotado de $\mathbb{R}^2$. Si llamamos $u : \Omega \to \mathbb{R}$ a la imagen sin ruido que deseamos obtener, entonces se cumple que:

$$f(x, y) = u(x, y) + n(x, y) \tag{1.1}$$

donde n es el ruido. Lo que buscamos es reconstruir u a partir de f. Una primera opción es usar la ecuación del calor (por ser una ecuación que modela problemas de difusión), sin embargo esto provoca difusión isotrópica de la imagen (es decir, en el intento por remover el ruido de la imagen se disminuyen porciones importantes de esta, como bordes por ejemplo) por lo que en algunos modelos [5] se ha intentado encontrar un método de difusión que se comporte de manera isotrópica en regiones planas y detenerla cerca de los bordes propios de la imagen alterando, de alguna forma, la ecuación de calor.

A partir de (1.1), en [1] se propone el siguiente modelo de minimización en el que las restricciones, así como el espacio funcional escogido, tienen el fin de detener el efecto isotrópico de la ecuación de calor original:

$$\text{mín} \int_\Omega \nabla u \, dx \tag{1.2a}$$

sujeto a

$$\int_\Omega u \, dx = \int_\Omega f \, dx, \tag{1.2b}$$

$$\int_\Omega |f - u|^2 \, dx = \sigma^2 \tag{1.2c}$$

donde σ^2 es la varianza global estimada del ruido y u y f son definidas como antes.

(1.2b) es una restricción lineal, que significa que el ruido $n(x, y)$ es de medida nula; mientras que (1.2c) es una restricción no lineal que aporta la información a-priori de que la desviación estándar del ruido debe ser igual a σ.

El problema (1.2) puede ser reformulado como un problema extendido de minimización, llamado Modelo de Variación Total para el Procesamiento de Imágenes:

$$u_\lambda = \text{arg} \min_u \{|u|_{BV_\Omega} + \lambda \phi(u)\} \tag{1.3}$$

en el que se impone una condición de Neumann en el borde $\partial\Omega$ para asegurar la condición

de la media: $\int_\Omega u = \int_\Omega f$. El multiplicador λ es un parámetro positivo, que puede ser una constante real o una función perteneciente a un espacio funcional X tal que $X \hookrightarrow L^2(\Omega)$ que determina el nivel de reducción de ruido en la imagen (ver figura 1.2) y $\phi : L^p(\Omega) \to \mathbb{R}$, con $p \geq 1$, es una distancia llamada 'término de fidelidad' que depende de la distribución estadística del ruido, así:

Cuando el ruido sigue una distribución normal, el ruido de f es gaussiano y ϕ es la norma en L^2 de la cantidad $u - f$. Cuando el ruido sigue una distribución de Poisson, $\phi(u) = \int_\Omega \lambda(u - f \ \log \ u) \ dx$ (distancia de Kullback - Leibler). Y, en presencia de ruido de impulso, ϕ es la norma en $\mathbb{L}^1$ de la cantidad $u - f$.

Figura 1.2: *Primera fila, de derecha a izquierda: Imagen Original, imagen con ruido Gaussiano, e imagen con $\lambda = 1 \times 10^6$. Segunda fila, de derecha a izquierda imagenes con $\lambda = 5000$, $\lambda = 1000$ y $\lambda = 100$ respectivamente. Observar que a un menor valor de λ disminuye la presencia del ruido. Sin embargo, con valores demasiado bajos se obtienen resultados distorcionados por lo que es necesaria la búsqueda de un valor óptimo.*

En general, (1.3) pertenece a la clase de problemas de minimización del tipo:

$$u_\lambda = \arg\min_u \{J(u) + H(u, f)\} \tag{1.4}$$

donde J es un funcional de regularización convexo no negativo, y el funcional de ajuste H es convexo y no negativo respecto a u para f fijo.

En la siguiente sección, veremos algunas propiedades útiles del espacio de variación acotada (BV). Después, se plantea la formulación del problema en este espacio, para finalmente transferir el problema a un espacio de Hilbert y así garantizar continuidad en el operador solución.

1.2 Algunas propiedades de los espacios en Variación Acotada (BV)

Muchos problemas en física, mecánica o procesamiento de imágenes son modelados en el Espacio de Variación Acotada (BV) ya que este permite discontinuidades en la solución. En fases de transición, o en la segmentación de una imagen, en teoría de plasticidad y en el estudio de fisuras, las soluciones de los problemas presentan discontinudiades a lo largo de variedades de codimensión uno y la solución de este tipo de problemas no se encuentra en los espacios de Sobolev clásicos, por lo que esta teoría debe ser complementada con los resultados del espacio BV.

1.2.1 Funciones de Variación Acotada de una variable

Definición 1.1. *La Variación Total de una función de valores reales, o complejos, f es definida en el intervalo $[a,b] \subset \mathbb{R}$ como la cantidad:*

$$V_a^b = \sup_{P \in \boldsymbol{P}} \sum_{i=0}^{np-1} |f(x_{i+1}) - f(x_i)|,$$

donde el supremo está tomado sobre el conjunto de particiones del intervalo $[a,b]$

Cuando f es diferenciable y su derivada es Riemann Integrable, su variación total es:

$$V_a^b = \int_a^b |f'(x)| dx$$

Definición 1.2. *Una función de valores reales, se dice de Variación Acotada (BV) en un intervalo $[a,b] \subset \mathbb{R}$ si su variación total es finita. Es decir:*

$$f \in BV([a,b]) \Longleftrightarrow V_a^b(f) < \infty$$

1.2.2 Funciones de Variación Acotada de varias variables

En las siguientes definiciones, tomadas de [6], Ω es un subconjunto abierto de $\mathbb{R}^N$ y el espacio $M(\Omega, \mathbb{R}^N)$ denota al espacio de todas las medidas de Borel con valores en $\mathbb{R}^N$.

Definición 1.3. *Diremos que una función $u : \Omega \to \mathbb{R}^N$ es una función de Variación Acotada, es decir $u \in BV$ si y solo si pertenece a $L^1(\Omega)$ y su gradiente Du, en el sentido de las distribuciones, está en $M(\Omega, \mathbb{R}^N)$ es decir es una medida de Borel con valores en $\mathbb{R}^N$*

Definición 1.4. *Dada una función u perteneciente a $L^1(\Omega)$ la variación total de u en Ω se define como:*

$$V(u,\Omega) := \sup \left\{ \int_\Omega u(x) \nabla.\mathbf{g} \; dx : g \in C_c^1(\Omega, \mathbb{R}^N), ||g||_\infty \leq 1 \right\}$$

donde $\nabla.\mathbf{g}$ representa a la divergencia de la función g y esta función pertenece al espacio C_c^1 de funciones continuamente diferenciables de soporte compacto contenido en Ω; notar que

para esta definición no se requiere que $\Omega \in \mathbb{R}^n$ sea un intervalo acotado. Las funciones g están equipadas con la norma: $||g||_\infty = \left(\sum_{i=1}^{N} \sup_{x\in\Omega} |g(x)|^2\right)^{1/2}$

Observación 1.1 (Forma de la Variación Total de una función diferenciable en varias variables)**.** *Dada una función diferenciable f definida en un abierto acotado $\Omega \subseteq \mathbb{R}^n$, la Variación Total de f puede ser representada por la siguiente expresión:*

$$V(f,\Omega) = \int_\Omega |Df(x)|dx$$

donde D es el vector gradiente de la función y $|\cdot|$ es la norma vectorial L^2

El siguiente teorema, nos proveé de una caracterización del espacio BV:

Teorema 1.1. *Las siguientes proposiciones son equivalentes.*

1. $u \in BV(\Omega)$
2. $u \in L^1(\Omega)$ *y* $\frac{\partial u}{\partial x_i} \in M(\Omega)$ *para todo* $i = 1, 2, ..., N$
3. $u \in L^1(\Omega)$ *y* $V(u,\Omega) < \infty$
4. $u \in L^1(\Omega)$ *y* $V(u,\Omega) < \infty$ *donde* $V(u,\Omega) := \sup\{\langle Du, g\rangle \ : \ g \in C_c^1(\Omega,\mathbb{R}^N) \ , \ ||g||_\infty \leq 1\}$ *con:*

$$\langle Du, g\rangle := \sum_{i=1}^{N} \int_\Omega g_i \frac{\partial u}{\partial x_i} \, dx$$

De las últimas dos afirmaciones, podemos concluir que la variación total $V(u,\Omega)$ se puede definir, de manera equivalente por:

$$|u|_{BV} = \sup\{\langle Du, g\rangle \ : \ g \in C_c(\Omega,\mathbb{R}^N) \ , \ ||g||_\infty \leq 1\} \tag{1.5}$$

Definición 1.5 (Norma del espacio BV)**.** *El Espacio de Variación Acotada está dotado de la norma:*

$$||u||_{BV} = ||u||_{L^1} + |u|_{BV}$$

Con la norma definida de esta forma, el espacio de Funciones de Variación Acotada es completo. La completitud de el espacio BV se sigue de la completitud del espacio L^1 y la propiedad de semicontinuidad expresada a continuación [6]

Teorema 1.2 (Semicontinuidad)**.** *Si $\{u_n\}$ es una sucesión en $BV(\Omega)$ con $\sup_n |u_n|_{BV} < \infty$ que converge fuertemente $u_n \to u$ en $L_1(\Omega)$, entonces $|u|_{BV} \leq \liminf_{n\to\infty} |u_n|_{BV}$ y $u \in BV(\Omega)$*

Observación 1.2. *Las funciones pertenecientes a este espacio tienen únicamente discontinuidades de salto. Es decir que los límites laterales en la dirección negativa*

$$L^- = \lim_{x\to x_0^-} f(x)$$

y en la dirección positiva,

$$L^+ = \lim_{x \to x_0^+} f(x)$$

siempre existen y son finitos, sin embargo pueden no ser iguales (en los puntos de discontinuidad)

Observación 1.3. *El funcional $V(\cdot, \Omega) : BV(\Omega) \to \mathbb{R}^+$ de la definición (1.4) es semicontinuo inferiormente en $BV(\Omega)$*

Observación 1.4. *$BV(\Omega)$ no es separable.*

Observación 1.5. *El espacio de Sobolev $W^{1,1}(\Omega)$ es un subespacio propio de $BV(\Omega)$. Es decir, se pueden encontrar funciones de Variación Acotada que no pertenecen a $W^{1,1}$, por ejemplo: cualquier función escalonada con un salto no trivial en dimensión 1.*

1.3 El Modelo de Control de Imágenes, problema de minimización a ser resuelto.

1.3.1 Introducción

Sea $f \in L^p(\Omega)$, con $p = 1$ o 2 y $\Omega \subset \mathbb{R}^2$, una imagen con ruido. Para remover el ruido consideraremos la regularización de Variación Total propuesta en [7] y que es considerada de manera frecuente en este tipo de problemas (ver [8], [9], [10] y las citas en [7]). Eso nos permite reconstruir una versión "limpia"de f, a la que notaremos como u, a través de la minimización del funcional (1.3) . Más precisamente:

$$\mathcal{J}(u) = |Du|(\Omega) + \int_\Omega \lambda \phi(u)\, dx \tag{1.6}$$

Donde

$$|Du|(\Omega) = \sup_{g \in C_0^\infty(\Omega;\mathbb{R}^2), ||g||_\infty \leq 1} \int_\Omega u \nabla . \mathbf{g}\, dx$$

es la seminorma definida en (1.5) (aquí, $C_0^\infty(\Omega)$ denota al espacio de las funciones continuamente diferenciables con soporte compacto [11]) y, como habíamos dicho antes: λ puede ser un parámetro real de valor positivo o una función perteneciente a un espacio $X \hookrightarrow L^2$. λ depende de la presencia del ruido, es decir que este parámetro refleja la intensidad del ruido en una imagen dada; y ϕ es la distancia llamada 'término de fidelidad' que depende de la distribución estadística del ruido.

Nuestro objetivo, es plantear un problema de optimización de EDPs para determinar el valor de λ (tanto en el caso real como en el caso funcional) y, de esta manera, entender la distribución del ruido presente en la medición de f. Aún más, trabajaremos con una extensión del modelo (1.6) para una versión más general, que permita representar en los datos múltiples tipos de distribución de ruido. Para ello seguiremos el método propuesto en [7]: extendemos (1.6) y ahora consideramos:

$$\min_{u \in BV} \left\{ |Du|(\Omega) + \sum_{i=1}^{d} \int_{\Omega} \lambda_i \phi_i(u, f) dx \right\} \tag{1.7}$$

como nuestro problema de minimización. Aquí, $\phi_i, i = 1, ..., d$ son funciones convexas y diferenciables respecto a u y los λ_i son parámetros reales no negativos, o funciones dependientes de las coordenadas x, y.

Las funciones ϕ_i modelan diferentes elecciones de fidelidad en los datos, mientras que los parámetros λ_i sirven como pesos de las funciones ϕ_i. En muchas imágenes conviven distintas distribuciones de ruido por lo que en [7] se propone esta generalización en base a la cual seguiremos nuestro estudio.[1]

Trataremos a (1.7) como restricción de un problema de optimización cuya meta sea la minimización del peso de la variable λ. Más adelante reemplazaremos dicho problema por una condición de optimalidad necesaria (en forma de desigualdad variacional). Necesitaremos después de una regularización de tipo Huber para poder así plantear nuestro problema en un espacio de Hilbert.

1.3.2 Problema de Optmización el espacio de Variación Acotada BV

Usando (1.7) como restricción del problema de optimización, ahora nos concentramos en el siguiente modelo:

$$\min_{\lambda_i \geq 0, i=1,...,d} g(\widetilde{u}) + \beta \sum_{i=1}^{d} ||\lambda_i||_X^2 \tag{1.8a}$$

sujeto a la restricción:

$$\widetilde{u} = \text{argmin}_{u \in BV \cap \mathcal{A}} \left\{ J(u) = |Du|(\Omega) + \sum_{i=1}^{d} \int_{\Omega} \lambda_i \phi_i(u, f) dx \right\} \tag{1.8b}$$

donde $X = \mathbb{R}$ o $X \hookrightarrow L^2(\Omega)$ cuando los λ_i son parámetros funcionales, $g : \mathbb{L}^2(\Omega) \longmapsto \mathbb{R}$ es un funcional de tipo C^1 a ser minimizado y $\beta > 0$.

El conjunto de funciones admisibles $\mathcal{A}$ se escoge de acuerdo a las distancias ϕ_i; así, $BV \cap \mathcal{A}$ restringe el conjunto de funciones BV en Ω a aquellas para las cuales $\phi_i : L^p(\Omega) \to \mathbb{R}$, con $p \geq 1$, está definida $\forall i = 1, ...d$ esto es:

1. ϕ_i debe estar bien definida,
2. ϕ_i debe ser diferenciable
3. ϕ_i debe ser convexa en u y debe estar acotada inferiormente
4. Además, se debe cumplir con la condición de coercividad:

[1] En este trabajo no se profundiza en ejemplos de imágenes donde conviven distintos tipos de ruido, sin embargo lo que se pretende es mostrar un modelo general que pueda servir de referencia para quien deseé estudiar estos casos. Los ejemplos que se estudian más adelante, siguen este modelo pero en particular para una distribución estadística de ruido (es decir, un parámetro λ y una función ϕ)

$$\int_\Omega \phi_i(u,f)dx \geq C_1||u||^p_{L^p} - C_2 \qquad \forall u \in L^p(\Omega) \cap \mathcal{A} \tag{1.9}$$

para constantes no negativas C_1 y C_2 y al menos un valor de $p = 1$ o $p = 2$ y para todo $i = 1, ..., d$

Ejemplos de ϕ_i's que cumplen con estas condiciones se pueden encontrar en [7].

En este trabajo, estudiaremos un ejemplo del Modelo de Ruido Gaussiano. En este modelo, la función de distancia ϕ es igual al cuadrado de la norma euclídea, es decir: $\phi = |u - f|^2$. La integral $\int_\Omega \phi(u)dx = ||u - f||^2_{L^2(\Omega)}$ cumple con la condición de coercividad para $p = 2$ y el conjunto admisible $\mathcal{A} = L^2(\Omega)$.

La función $g(\widetilde{u})$ es el funcional a ser minimizado (por lo general, la diferencia entre una imagen aproximada a la original y u). El término $\beta \sum_{i=1}^d ||\lambda_i||^2_X$ permite que el peso con el se aplica la reducción de ruido en la imagen no sea muy violento (evitando así, que la imagen quede borrosa). Mientras que la función de distancia ϕ nos indica que el ruido es de medida nula.

Para la solución numérica de $(1.8\mathrm{b})$ usaremos métodos iterativos para los que el gradiente de la Variación Total del modelo debe estar definido y ser único. Es decir, el minimizador de $(1.8\mathrm{b})$ debe estar caracterizado de manera única por la solución de la ecuación de Euler- Lagrange correspondiente. Como la Variación Total no es diferenciable, su derivada es en realidad caracterizada por un conjunto de subgradientes (el subdiferencial); por eso, es necesario usar una versión regularizada, a través de la función de Huber, de la variación total.

Una imagen con ruido usualmente contiene grandes diferencias entre pixeles cercanos, para eliminar estas diferencias se propuso un modelo en Variación Acotada ya que este elimina el efecto isotrópico del problema original previniendo el efecto borroso que se puede producir.

Al escoger una regularización a la Variación Acotada, es necesario escoger una función que haga exactamente lo mismo: eliminar el efecto isotrópico evitando que la imagen se vuelva borrosa. Se ha sugerido usar la transformada de Fourier, la transformada a coseno discreto, y una transformada wavelet (ver referencias en [12]); con estos métodos una parte del dominio es aislada de la imagen logrando así que el componente con ruido se elimine de manera selectiva y la imagen sin ruido se reconstruye tomando la transformada inversa, sin embargo estos métodos dan resultados alterados. De igual manera, se han implementado métodos estadísticos para la disminución del ruido dando buenos resultados (ver referencias en [12]) por lo que en [12] se estudia una extensión no local a estos métodos. A través de la regularización no local de la variación total se logra la preservación de bordes, sin embargo esta deja residuos de ruido en la imagen. Para eliminar completamente este ruido, se utiliza una regularización en H^1; pero, esta última deja una imagen resultante borrosa en comparación a la regularización no local. En el mismo trabajo, se propone una forma de fusionar ambas regularizaciones de manera selectiva: la regularización no local de Huber, esto se logra con un valor de cota en base al cual se realiza esta distinción. El valor de la regularización de Huber aplica una penalidad lineal (la regularización no local de la varia-

ción total) para valores más grandes que la cota propuesta, y una penalidad cuadrática (la regularización H^1) para valores más pequeños.

En [13] se define por:

Definición 1.6 (Función de Huber).

$$\begin{cases} L_\delta(a) = (1/2)a^2 & \textit{cuando } |a| \leq \delta \\ L_\delta(a) = \delta(|a| - \delta/2), & \textit{caso contrario.} \end{cases}$$

Esta función es cuadrática para valores pequeños de a y lineal para valores más grandes, con valores iguales y pendientes en las secciones para las que $|a| = \delta$.

La función de Huber es siempre continua (o semicontinua inferiormente)

Ahora, nosotros usaremos una regularización de tipo Huber para la variación total, como se procede en [7]. Sea $\gamma >> 1$:

$$|\nabla u|_\gamma = \begin{cases} |\nabla u| - \frac{1}{2\gamma} & \text{si } |\nabla u| \geq \frac{1}{\gamma} \\ |\nabla u|^2 \frac{\gamma}{2} & \text{si } |\nabla u| < \frac{1}{\gamma} \end{cases} \tag{1.10}$$

Luego, tenemos la siguiente versión regularizada de (1.8)

$$\min_{\lambda_i \geq 0, i=1,\ldots,d} g(\widetilde{u}) + \beta \sum_{i=1}^{d} ||\lambda_i||_X^2 \tag{1.11a}$$

sujeta a:

$$\widetilde{u} = \text{argmin}_{u \in W^{1,1} \cap \mathcal{A}} \left\{ \mathcal{J}^\gamma(u) = \int_\Omega |\nabla u|_\gamma \, dx + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i(u, f) dx \right\} \tag{1.11b}$$

donde el espacio X, g , las distancias ϕ_i y $\beta > 0$ se definen como antes. Asumimos que el conjunto de funciones admisibles $\mathcal{A}$ es un subconjunto convexo y cerrado de $W^{1,1}(\Omega)$, escogido de acuerdo a los ϕ_i (ver Observación (1.5))

En [7] se demuestra la existencia de una solución óptima para (1.11b) por el método de relajación, para ello se extiende la definición $\mathcal{J}^\gamma$ al espacio $BV(\Omega)$ como:

$$\mathcal{J}^\gamma_{ext}(u) = \begin{cases} \mathcal{J}^\gamma & u \in W^{1,1} \cap \mathcal{A} \\ +\infty & u \in BV(\Omega) \setminus (W^{1,1} \cap \mathcal{A}) \end{cases} \tag{1.12}$$

y se prueba la existencia de un minimizador para la envoltura semicontinua de $\mathcal{J}^\gamma_{ext}$, a través de su versión relajada, definida como:

$$\mathcal{J}^\gamma_{relax}(u) = \int_\Omega |\nabla u|_\gamma dx + C \int_\Omega |D_\mathcal{S} u| + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i(u, f) dx \tag{1.13}$$

Notar que

$$\mathcal{J}^\gamma_{relax}(u) \leq \mathcal{J}^\gamma_{ext}(u), \quad \text{para} \quad u \in BV(\Omega)$$

y que $\mathcal{J}^\gamma_{relax}(u) = \mathcal{J}^\gamma_{ext}(u)$ para $u \in W^{1,1}(\Omega) \cap \mathcal{A}$.

En las demostraciones propuestas para este problema en [7], son necesarias las propiedades asumidas para las funciones ϕ_i, es decir: continuidad de las funciones, la convexidad (que implica la convexidad del conjunto $\mathcal{A}$ y la propiedad de coercividad (1.9))

Después de probar la existencia de una solución óptima para la versión relajada del funcional (1.13) [7] prueba que $\mathcal{J}^{\gamma}_{relax}$ Γ-converge (Ver A.1.4 en Apéndices) a la Variación Total $\mathcal{J}$ cuando $\gamma \to \infty$. Luego, el minimizador de $\mathcal{J}^{\gamma}_{relax}$ converge al minimizador de $\mathcal{J}$ cuando γ tiende hacia el infinito.

Estos resultados referentes al problema de minimización en el espacio BV no son suficientes para concluir convergencia de los pesos óptimos regularizados en [7], pues para esto necesitamos continuidad en la aplicación solución $\lambda \to u(\lambda)$. Existen muchos trabajos (ver referencias en [7]) que presentan resultados en esta dirección, sin embargo no son lo suficientemente fuertes. De hecho, este es un tema en investigación.

1.4 Problema de Optmización en el espacio de Hilbert H^1_0

A pesar de las propiedades favorables para la reconstrucción de imágenes, el problema (1.7) posee varias dificultades. A nivel analítico, el principal problema reside en que (1.7) está definido en un espacio de Banach que no es reflexivo, con lo cual su dual es difícil de caracterizar. Por otro lado, a nivel numérico, el sistema de optimalidad asociado con (1.7) consiste en una EDP no lineal que no se puede implementar numéricamente de manera directa.

Para obtener continuidad del operador solución, y por ende convergencia de los parámetros óptimos regularizados λ_i seguiremos el lineamiento propuesto en [7].

Para ello reemplazamos el problema (1.7) por la versión elíptico-regularizada:

$$\min_u \left(\frac{\varepsilon}{2} ||Du||^2_{L^2} + |Du|(\Omega) + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i(u)\, dx \right) \tag{1.14}$$

donde $0 < \varepsilon << 1$ es un parámetro de difusión artificial. A continuación, usaremos una caracterización de (1.14) a través de una desigualdad variacional de segundo tipo, para llegar a un sistema de Optimalidad que nos permita implementar numéricamente el problema.

1.4.1 Resultado de existencia para una desigualdad variacional del segundo tipo

Sea X un Espacio de Hilbert con producto escalar $(\ \cdot\ )_X$ y norma inducida $||\cdot||$. Sean además, $a : X \times X \to \mathbb{R}$ una forma bilineal coerciva y continua, y $j : X \to (-\infty, +\infty]$ una función propia semicontinua inferiormente. Para un elemento $f \in X$ dado, se considera el problema de encontrar $u \in X$ tal que:

$$a(u, v-u) + j(v) - j(u) \geq (f, v-u)_X, \qquad \forall v \in X \tag{1.15}$$

Este problema se conoce como desigualdad variacional elíptica de segundo tipo, un estudio detallado de problemas de control óptimo gobernados por este tipo de desigual-

dades se puede encontrar en [14]. En [15] se demuestra la existencia de soluciones para este tipo de desigualdades gracias al Teorema de Hartman-Stampacchia y la definición de 'operadores submonótonos'. Ahora, mostraremos una caracterización de dicha solución.

Definimos

$$\mathcal{J}(u) = \frac{1}{2}a(u,u) + j(v) - (f,v)_X \tag{1.16}$$

Teorema 1.3. *Un elemento $u \in X$ es solución de (1.15) si y solo si, es un minimizador de $\mathcal{J}$, es decir:*

$$\mathcal{J}(v) \geq \mathcal{J}(u), \qquad \forall v \in X$$

Demostración. para $v \in X$, si u es solución de (1.15), se tiene:

$$\begin{aligned}\mathcal{J}(v) - \mathcal{J}(u) &= a(u, v-u) + j(v) - j(u) - (f, v-u)_X + \frac{1}{2}a(u-v, u-v)\\ &\geq \frac{1}{2}a(u-v, u-v)\\ &\geq C||u-v||_X^2 \geq 0\end{aligned}$$

donde C es una constante.

Por otro lado, supongamos que u es un mínimo de $\mathcal{J}$. Usando la definición (1.16):

$$\frac{1}{2}a(v,v) + j(v) - (f,v)_X \geq \frac{1}{2}a(u,u) + j(u) - (f,u)_X, \qquad \forall v \in X$$

en particular, para $v = u + t(v-u)$ donde $t > 0$:

$$\frac{1}{2}a(u+t(v-u), u+t(v-u)) + j(u+t(v-u)) - (f, u+t(v-u))_X \geq \frac{1}{2}a(u,u) + j(u) - (f,u)_X$$

de donde:

$$t \;\; a(u, v-u) + \frac{t^2}{2}a(u-v, u-v) + t(j(v) - j(u)) \geq t(f, v-u)_X$$

Cuando $t \to 0$:

$$a(u, v-u) + j(v) - j(u) \geq (f, v-u)_X \qquad \forall v \in X$$

□

De este teorema, se sigue:

Teorema 1.4. *Sea $f \in X$, entonces existe un único elemento $u \in X$ tal que (1.15) se cumple.*

Demostración. Supongamos que existen dos elementos u_1 y u_2 en X que satisfacen (1.15). Por el teorema anterior, esto es equivalente a admitir que:

$$\mathcal{J}(v) \geq \mathcal{J}(u_1) \quad \forall v \in X$$

y

$$\mathcal{J}(v) \geq \mathcal{J}(u_2) \quad \forall v \in X$$

se cumplen a la vez, lo cual es contradictorio ya que, en particular u_2 pertenece a X.

□

1.4.2 Caracterización del problema con una desigualdad variacional del segundo tipo y existencia de solución

Si notamos

$$\mathcal{J}_H = \frac{\varepsilon}{2}||Du||^2_{L^2} + |Du|(\Omega) + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i(u)\, dx$$

Tenemos que una condición de optimalidad necesaria está dada por la siguiente desigualdad:

$$\mathcal{J}'_H(u)\,(v-u) \geq 0 \quad \forall v \in H^1_0 \tag{1.17}$$

Encontraremos la derivada direccional del funcional $\mathcal{J}_H$.

Sean $\varphi \in H^1_0$ una dirección y $h > 0$,

$$\begin{aligned}
\mathcal{J}_H(u+h\varphi) - \mathcal{J}_H(u) &= \frac{\varepsilon}{2}\int_\Omega (|D(u+h\varphi)|^2 - |Du|^2)dx + |Du+h\varphi|(\Omega) - |Du|(\Omega) \\
&+ \sum_{i=1}^{d} \int_\Omega \lambda_i[\phi_i(u+h,f) - \phi_i(u,f)]\, dx \\
&= \frac{\varepsilon}{2}\int_\Omega (2h|Du||D\varphi| + h^2|D\varphi|^2)dx + \int_\Omega |Du+h\varphi|dx - \int_\Omega |Du|dx \\
&+ \sum_{i=1}^{d} \int_\Omega \lambda_i[\phi_i(u+h,f) - \phi_i(u,f)]\, dx
\end{aligned}$$

dividiendo por h:

$$\begin{aligned}
\frac{1}{h}(\mathcal{J}_H(u+h\varphi) - \mathcal{J}_H(u)) &= \varepsilon \int_\Omega \left(|Du||D\varphi| + \frac{h}{2}|D\varphi|^2\right) dx + \int_\Omega \frac{1}{h}\left(|Du+h\varphi| - |Du|\right) dx \\
&+ \sum_{i=1}^{d} \int_\Omega \lambda_i \frac{1}{h}[\phi_i(u+h,f) - \phi_i(u,f)]\, dx
\end{aligned}$$

cuando $h \to 0$ tenemos:

$$\mathcal{J}'_H(u).\varphi = \varepsilon \int_\Omega |Du||D\varphi|dx + \int_\Omega |D\varphi|dx + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi'_i(u,f)\varphi\, dx \tag{1.18}$$

Con este resultado, podemos calcular (1.17):

$$\varepsilon(Du, D(v-u))_{L^2} + \int_\Omega |Dv|dx - \int_\Omega |Du|dx + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i'(u,f)(v-u)\, dx \geq 0 \quad \forall v \in H_0^1 \quad (1.19)$$

Tras sumar el término coercivo en (1.14), hacemos que el espacio solución esté en H_0^1 (dado que ahora es necesario que la primera derivada de u sea L^2 integrable).

De manera alternativa, si usamos el término $\varepsilon||u||_{H^1(\Omega)}$ en (1.14) el espacio solución será $H^1(\Omega)$, lo que permite escoger nuevas condiciones de borde (ver [16]).

Ahora, nuestro problema se transforma en determinar el conjunto de parámetros λ_i con $i = 1, 2, ..., d$ óptimos que resuelven:

$$\min_{\lambda_i \geq 0, i=1,\ldots,d} g(u) + \beta \sum_{i=1}^{d} ||\lambda_i||_X^2 \quad (1.20\text{a})$$

sujeta a:

$$\varepsilon(Du, D(v-u))_{L^2} + \int_\Omega |Dv|dx - \int_\Omega |Du|dx + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i'(u,f)(v-u)\, dx \geq 0 \quad \forall v \in H_0^1 \quad (1.20\text{b})$$

donde $X = \mathbb{R}$ o es un espacio funcional de Hilbert.

El problema (1.20) corresponde a un problema de optimización generado por desigualdades variacionales del segundo tipo. El teorema 1.3 nos proporciona una herramienta de resolución para este tipo de desigualdad.

A continuación se analizará el problema (1.20) desde la perspectiva propuesta en [14] y [7]: después de probar existencia de una solución óptima, veremos una regularización del problema que nos permita probar continuidad de la aplicación solución y convergencia de los parámetros óptimos, lo que nos llevará a encontrar el sistema de optimalidad, para ambos casos, de parámetros reales y funcionales, que caracterice la solución óptima de (1.20).

La existencia de una solución óptima para el problema (1.20) se puede ver en [7, Teorema 3.1], donde se prueba la existencia de un par óptimo para (1.20). Adicionalmente se explora el caso $\beta = 0$: si $X = \mathbb{R}$ el teorema de existencia funciona, mientras que para $X = L^2(\Omega)$ se necesitan restricciones de caja (es decir $0 \leq \lambda_i(x) \leq b$ c.t.p en Ω) para garantizar la existencia de una solución óptima.

Sin embargo, como se explica en detalle en [14], es necesaria una regularización de conjuntos activos e inactivos, para poder caracterizar el problema a través de un sistema de optimalidad adecuado y se pueda seguir los pasos propuestos en [17] para obtener una aproximación numérica del tipo Newton Semismooth.

1.4.3 Problema de dualidad de una desigualdad variacional de 2do tipo

Definición 1.7 (Desigualdad Variacional de Segundo Tipo [14]). *Sean Y, U espacios funcionales de Hilbert definidos en un dominio acotado $\Omega \subset \mathbb{R}^n$, $a(\cdot,\cdot)$ una forma bilineal continua*

y coerciva, y $j : Y \to \mathbb{R}$ *un funcional convexo no diferenciable de la forma:*

$$j(v) := g \int_S |\mathcal{K}v| ds,$$

con $S \subset \bar{\Omega}, g > 0$, *y* $\mathcal{K} : Y \to L^2(S)$ *un operador lineal acotado. Se define una desigualdad variacional de segundo tipo como el problema:*

Dado $f \in Y'$, *encontrar* $y \in Y$ *tal que:*

$$a(y, v - y) + j(v) - j(y) \geq \langle f, v - y \rangle_{Y',Y} \qquad \forall v \in Y \tag{1.21}$$

Se puede observar que la desigualdad $(1.20\mathrm{b})$ entra en esta categoría, por lo tanto el siguiente análisis, visto en detalle en [14], bien puede aplicarse en lo que refiere a nuestro problema.

La desigualdad (1.21) se puede reformular como:

$$\min_{y \in Y} \mathcal{F}(y) + g \int_S |\mathcal{K}y| ds \tag{P}$$

donde $\mathcal{F}(y) = \frac{1}{2} a(y, y) - \langle f, y \rangle_{Y',Y}$.

Por otro lado, el dual de Fenchel (ver apéndice A.1) del problema (P) está dado por:

$$\sup_{q \in L^2(S)} \{-\mathcal{F}^\star(-\mathcal{K}^\star q) - \mathcal{G}^\star q\} \tag{D}$$

que se puede reescribir como:

$$\begin{cases} \text{máx}_{q \in \mathcal{K}_g} -\frac{1}{2} a(y, y) \\ \text{s.a:} \quad a(y, w) + (q, \mathcal{K}w) = \langle f, w \rangle_{Y',Y} \quad \forall w \in Y \end{cases} \tag{1.22}$$

Como los funcionales involucrados son propios y convexos, y Y es reflexivo, se tiene que [18, p.60]:

$$\min_{y \in Y} \left\{ \mathcal{F}(y) + \alpha \int_S |\mathcal{K}y| ds \right\} = \sup_{q \in L^2(S)} \{-\mathcal{F}^\star(-\mathcal{K}^\star q - \mathcal{G}^\star q)\} \tag{1.23}$$

es decir, no hay diferencia de dualidad entre (P) y (1.22). Más aún, las soluciones y de (P) y q de (1.22) están ligadas por el sistema

$$\begin{cases} \mathcal{K}^\star q = \mathcal{F}'(y) \\ q \in \partial \mathcal{G}(\mathcal{K}y) \end{cases} \tag{1.24}$$

donde $\partial \mathcal{G}$ es el operador subdiferencial (ver A.4 en A.1).

El sistema (1.24) se puede expresar como:

$$a(y, v) + (q, \mathcal{K}v) = \langle f, v \rangle_{Y',Y} \quad \forall v \in Y \tag{1.25a}$$

$$\langle q(x), \mathcal{K}y(x) \rangle_{Y',Y} = g|\mathcal{K}y(x)| \quad \text{c.t.p } S \tag{1.25b}$$

$$|q(x)| \leq g \quad \text{c.t.p } S \tag{1.25c}$$

Se puede observar que q no está determinada en los sectores donde $\mathcal{K}y(x) = 0$. Más aún (ver referencias en [14]), que si $\ker \mathcal{K}^\star \neq 0$, entonces no existe un multiplicador $q \in L^2(S)$ que satisfaga el sistema (1.25). Estos hechos hacen que sea necesario determinar alguna solución para que el problema (1.25) esté bien definido.

En [14] se propone una regularización de Tykhonov para obtener una versión bien definida de (1.25). Dicha regularización se ha aplicado numéricamente en muchos trabajos para resolver desigualdad variacionales del segundo tipo con métodos del tipo newthon semismooth (ver referencias en [14]) y más aún: su eficiencia numérica es comparada con otros algoritmos puestos a prueba en [17].

1.4.4 Existencia y convergencia de soluciones del problema regularizado

Con las consideraciones realizadas en el apartado anterior, definimos ahora la siguiente familia de problemas regularizados:

$$\varepsilon(Du_\gamma, Dv)_{L^2} + (h_\gamma(Du_\gamma), Dv)_{L^2} + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i'(u_\gamma, f) v \, dx = 0 \quad \forall v \in H_0^1 \tag{1.26}$$

donde $h_\gamma(Du_\gamma)$ corresponde a una regularización de conjuntos activo e inactivo del subdiferencial de $|Du_\gamma|$, es decir h_γ coincide con todos los elementos del subdiferencial en una vecindad pequeña. Hay que notar que en lugar de $Du(x)$ estamos escribiendo Du por facilidad.

La elección natural es la función:

$$h_\gamma(z) = \frac{\gamma z}{\text{máx}(1; \gamma|z|)} \tag{1.27}$$

que corresponde a la derivada de la Función de Huber definida en (1.10).

Una regularización alternativa es propuesta en [7], dada por la función C^1:

$$h_\gamma(z) = \begin{cases} \frac{z}{|z|} & \text{si} \quad \gamma|z| \geq 1 + \frac{1}{2\gamma} \\ \frac{z}{|z|}\left(1 - \frac{\gamma}{2}\left(1 - \gamma|z| + \frac{1}{2\gamma}\right)^2\right) & \text{si} \quad 1 - \frac{1}{2\gamma} \leq \gamma|z| \leq 1 + \frac{1}{2\gamma} \\ \gamma z & \text{si} \quad \gamma|z| \leq 1 - \frac{1}{2\gamma} \end{cases} \tag{1.28}$$

En adelante, se notará:

$$\mathcal{A}^\gamma := \left\{ x \in \Omega \ : \ \gamma|z| - 1 \geq \frac{1}{2\gamma} \right\} \quad \text{al conjunto activo}$$

$$\mathcal{S}^\gamma := \left\{ x \in \Omega \ : \ |\gamma|z| - 1| \leq \frac{1}{2\gamma} \right\} \quad \text{al conjunto semi-activo}$$

y

$$\mathcal{I}^\gamma := \left\{ x \in \Omega \ : \ 1 - \gamma|z| > \frac{1}{2\gamma} \right\} = \mathcal{S} \setminus (\mathcal{A}^\gamma \cup \mathcal{S}^\gamma) \quad \text{al conjunto inactivo}$$

Esta regularización se usa cuando es necesaria la diferenciabilidad de la función regularizadora. Ambas funciones, (1.27) y (1.28), regularizan u de manera local y poseen una estructura de conjuntos activo e inactivo; es decir: la norma de la función alcanza la cota 1 fuera de la vecindad: $\left\{ x : |x| \leq \frac{1}{\gamma} \right\}$

Observación 1.6. *En [14] se verifica que* (1.26) *tiene solución única. Más aún, se demuestra que la sucesión de soluciones regularizadoras* $\{u_\gamma\}$ *converge fuertemente en* $H_0^1(\Omega)$ *a la solución de* (1.19)

Tomando en cuenta la versión regularizada (1.26) de la desigualdad variacional, nuestro problema se convierte en:

$$\min_{\lambda_i \geq 0,\ i=1,..,d} g(u) + \beta \sum_{i=1}^{d} ||\lambda_i||_X^2 \tag{1.29a}$$

sujeto a:

$$\varepsilon(Du_\gamma, Dv)_{L^2} + (h_\gamma(Du_\gamma), Dv)_{L^2} + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i'(u_\gamma, f) v \, dx = 0 \quad \forall v \in H_0^1 \tag{1.29b}$$

donde $0 < \varepsilon << 1$.

En este contexto, en [7] se prueba la existencia de una solución óptima para el problema regularizado (1.29) a través del siguiente teorema:

Teorema 1.5. *Existe una solución óptima para cada problema regularizado* (1.29)*. Más aún, la sucesión* $\{\lambda_\gamma\}$ *de parámetros óptimos regularizado está acotada en* X^d *y cada subsucesión que converge débilmente, converge hacia una solución óptima de* (1.20)

1.4.5 Sistema de Optimalidad

Definimos el operador $\mathcal{G}_\gamma : X^d \mapsto H_0^1(\Omega)$ como el operador solución que asigna a cada parámetro λ_γ la solución correspondiente a la desigualdad variacional regularizada (1.26) con la función h_γ definida en (1.28). En [7, Proposición 3.4] se prueba que este operador es Gâteaux diferenciable y, más aún, que la solución de dicha derivada en una dirección dada es única.

Gracias a esta propiedad es posible obtener un sistema de optimalidad para la caracterización de las soluciones óptimas del problema regularizado [7, Teorema 3.5]

Teorema 1.6 (Sistema de Optimalidad del Problema Regularizado)**.** *Sean* $(\lambda_\gamma, u_\gamma)$ *las soluciones óptimas del problema regularizado* (1.29)*. Existen multiplicadores de Lagrange* $(p_\gamma, \lambda) \in H_0^1(\mathbb{R}) \times X^d$ *tales que se cumple el sistema de optimalidad:*

$$\varepsilon(Du_\gamma, Dv)_{L^2} + (q, Dv)_{L^2} + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i'(u_\gamma, f)v\ dx = 0 \quad , \quad \forall v \in H_0^1(\Omega) \tag{1.30a}$$

$$q - h_\gamma(Du_\gamma) = 0 \quad ctp\ en\Omega \tag{1.30b}$$

$$\varepsilon(Dv, Dp_\gamma)_{L^2} + (\xi, Dv)_{(H_0^1(\Omega))'} + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i''(u_\gamma, f)pv\ dx = -(g'(u_\gamma), v) \quad , \quad \forall v \in H_0^1(\Omega) \tag{1.30c}$$

$$\xi - h_\gamma'(Du_\gamma)Dp = 0 \tag{1.30d}$$

$$2\beta(\lambda_\gamma, \lambda - \lambda_\gamma)_{X^d} + \sum_{i=1}^{d} \int_\Omega \phi_i'(u_\gamma, f)p(\lambda_i - \lambda_{\gamma_i})\ dx \geq 0 \quad , \ \forall \lambda \geq 0 \tag{1.30e}$$

Observación 1.7. *A continuación procederemos a calcular dicho sistema de optimalidad basándonos en la forma clásica de deducir un sistema de optimalidad a partir de la definición de Lagrangeano asociado al problema. Sin embargo, dicho proceso no es una demostración formal de la existencia de los multiplicadores $p \in H_0^1(\Omega)$ y $\lambda \in X^d$ ni del sistema de optimalidad como tal, si no más bien una herramienta para derivar el sistema de optimalidad asociado al problema. Para una demostración formal de su existencia y validez ver [7, Teorema 3.6] y [14, Teorema 6.2] , en donde a partir de la demostración de la existencia, y deducción, del sistema de optimalidad del problema relajado* (1.29)*, se concluye dicho teorema.*

Demostración. Definimos el Lagrangeano:

$$L(u_\gamma, \lambda_\gamma, p_\gamma) = g(u_\gamma) + \beta|\lambda_\gamma|^2 + \varepsilon(Du_\gamma, Dp_\gamma) + (h_\gamma(Du_\gamma), Dp_\gamma) + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\phi_i'(u_\gamma, f)\ p_\gamma\ dx \tag{1.31}$$

donde $p_\gamma \in H_0^1(\Omega)$.

Ahora, sea $v \in H_0^1(\Omega)$, las condiciones necesarias de primer orden son:

- $$\frac{\partial L}{\partial p_\gamma}(u_\gamma, \lambda_\gamma, p_\gamma)\ v = 0 \quad \forall v \in H_0^1 \tag{1.32}$$
- $$\frac{\partial L}{\partial u_\gamma}(u_\gamma, \lambda_\gamma, p_\gamma)v = 0 \quad \forall v \in H_0^1 \tag{1.33}$$
- $$\frac{\partial L}{\partial \lambda_\gamma}(u_\gamma, \lambda_\gamma, p)(\lambda - \lambda_\gamma) \geq 0 \quad \forall \lambda \geq 0 \tag{1.34}$$

Primero encontraremos el valor de la derivada en (1.32). Para esto, sea $h > 0$ y sea $v \in H_0^1(\Omega)$:

$$\begin{aligned}
L(u_\gamma, \lambda_\gamma, p_\gamma + hv) - L(u_\gamma, \lambda_\gamma, p_\gamma) &= \varepsilon(Du_\gamma, D(p_\gamma + hv)) + (h_\gamma(Du_\gamma), D(p_\gamma + vh)) \\
&+ \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i'(u_\gamma, f)(p_\gamma + vh)\ dx \\
&- \varepsilon(Du_\gamma, Dp_\gamma) - (h_\gamma(Du_\gamma), Dp_\gamma) - \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i'(u_\gamma, f) p_\gamma\ dx \\
&= \varepsilon(Du_\gamma, h\ Dv) + (h_\gamma(Du_\gamma), h\ Dv) + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i'(u_\gamma, f) vh\ dx \\
&= h\ \varepsilon(Du_\gamma, Dv) + h\ (h_\gamma(Du_\gamma), Dv) + h \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i'(u_\gamma, f) v\ dx
\end{aligned}$$

dividiendo por h y cuando $h \to 0$ se tiene que:

$$\frac{\partial L}{\partial p_\gamma}(u_\gamma, \lambda_\gamma, p_\gamma)v = \varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\ \phi_i'(u_\gamma, f) v\ dx = 0$$

si es que, adicionalmente, decidimos notar: $q = h_\gamma(Du_\gamma)$ tenemos como resultado el sistema (1.30a) y (1.30b).

Ahora bien. Veremos que la derivada (1.33) es equivalente a las condiciones de optimalidad (1.30c) y (1.30d). Para ello: sean $h > 0$ y $v \in H_0^1(\Omega)$,

$$\begin{aligned}
L(u_\gamma + hv, \lambda_\gamma, p_\gamma) - L(u_\gamma, \lambda_\gamma, p_\gamma) &= g(u_\gamma + hv) + \varepsilon(D(u_\gamma + hv), Dp_\gamma) + (h_\gamma(Du_\gamma + h\ Dv), Dp_\gamma) \\
&+ \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} \phi_i'(u_\gamma + hv, f)\ p_\gamma\ dx - g(u_\gamma) - \varepsilon(Du_\gamma, Dp_\gamma) \\
&- (h_\gamma(Du_\gamma), Dp_\gamma) - \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} \phi_i'(u_\gamma, f)\ p_\gamma\ dx \\
&= [g(u_\gamma + hv) - g(u_\gamma)] + \varepsilon[(Du_\gamma + h\ Dv, Dp_\gamma) - (Du_\gamma, Dp_\gamma)] \\
&+ (h_\gamma(D(u_\gamma + hv)) - h_\gamma(Du_\gamma), Dp_\gamma) \\
&+ \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} [\phi_i'(u_\gamma + h\ v, f) - \phi_i'(u_\gamma, f)]\ p_\gamma\ dx \\
&= [g(u_\gamma + hv) - g(u_\gamma)] + \varepsilon h(Dv, Dp_\gamma) + (h_\gamma(D(u_\gamma + hv)) \\
&- h_\gamma(Du_\gamma), Dp_\gamma) + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} [\phi_i'(u_\gamma + h\ v, f) - \phi_i'(u_\gamma, f)]\ p_\gamma\ dx
\end{aligned}$$

Dividiendo esta expresión por h:

$$\frac{1}{h}[L(u_\gamma + hv, \lambda_\gamma, p_\gamma) - L(u_\gamma, \lambda_\gamma, p_\gamma)] = \frac{1}{h}[g(u_\gamma + hv) - g(u_\gamma)] + \varepsilon(Dv, Dp_\gamma)$$
$$+ \frac{1}{h}(h_\gamma(D(u_\gamma + hv)) - h_\gamma(Du_\gamma), Dp_\gamma)$$
$$+ \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} \ \frac{1}{h}[\phi_i'(u_\gamma + h\ v, f) - \phi_i'(u_\gamma, f)]\ dx$$

y cuando $h \to 0$:

$$\frac{\partial L}{\partial u_\gamma}(u_\gamma, \lambda_\gamma, p_\gamma)v = [g'(u_\gamma)v] + \varepsilon(Dv, Dp_\gamma) + (\langle h_\gamma'(Du_\gamma), Dv\rangle_{(H_0^1(\Omega))'}, Dp_\gamma)$$
$$+ \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} \phi_i''(u_\gamma, f)v\ p_\gamma\ dx = 0$$

de donde, para todo $v \in H_0^1(\Omega)$:

$$\varepsilon(Dv, Dp_\gamma) + (h_\gamma'(Du_\gamma)Dv, Dp_\gamma) + \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i} \phi_i''(u_\gamma, f)v\ p_\gamma\ dx = -[g'(u_\gamma)v] \tag{1.35}$$

Si notamos $\xi = h_\gamma'(Du_\gamma)Dp$ tenemos que (1.35) es equivalente al sistema (1.30c) y (1.30d). Finalmente, calcularemos (1.34) para obtener la condición de optimalidad (1.30e).
Sean $h > 0$ y sea $v \in X^d$

$$L(u_\gamma, \lambda_\gamma + hv, p_\gamma) - L(u_\gamma, \lambda_\gamma, p_\gamma) = \beta|\lambda_\gamma + hv|^2 + \sum_{i=1}^{d} \int_\Omega (\lambda_{\gamma_i} + hv_i)\phi_i'(u_\gamma, f)p_\gamma\ dx$$
$$- \beta|\lambda_\gamma|^2 - \sum_{i=1}^{d} \int_\Omega \lambda_{\gamma_i}\phi_i'(u_\gamma, f)p_\gamma\ \ dx$$
$$= \beta(|\lambda_\gamma + hv|^2 - |\lambda_\gamma|^2) + \sum_{i=1}^{d} \int_\Omega hv_i\ \phi_i'(u_\gamma, f)p_\gamma\ dx$$
$$= \beta[2h\ (v, \lambda_\gamma)_{X^d} - h^2|v|^2] + \sum_{i=1}^{d} \int_\Omega hv_i\ \phi_i'(u_\gamma, f)p_\gamma\ dx$$

dividiendo por $h > 0$:

$$\frac{1}{h}[L(u_\gamma, \lambda_\gamma + hv, p_\gamma) - L(u_\gamma, \lambda_\gamma, p_\gamma)] = \beta[2\ (v, \lambda_\gamma)_{X^d} - h|v|^2] + \sum_{i=1}^{d} \int_\Omega v_i\ \phi_i'(u_\gamma, f)p_\gamma\ dx$$

y cuando $h \to 0$:

$$\frac{\partial L}{\partial \lambda_\gamma}(u_\gamma, \lambda_\gamma, p_\gamma)v = 2\beta(v, \lambda_\gamma)_{X^d} + \sum_{i=1}^{d} \int_\Omega v_i \ \phi_i'(u_\gamma, f)p_\gamma \ dx$$

Ahora, para todo $\lambda \geq 0$ si tomamos $v = \lambda - \lambda_\gamma$, entonces se cumple que:

$$\begin{aligned}\frac{\partial L}{\partial \lambda_\gamma}(u_\gamma, \lambda_\gamma, p_\gamma)(\lambda - \lambda_\gamma) &= 2\beta(\lambda - \lambda_\gamma, \lambda_\gamma)_{X^d} + \sum_{i=1}^{d} \int_\Omega (\lambda_i - \lambda_{\gamma_i}) \ \phi_i'(u_\gamma, f)p_\gamma \ dx \\ &\geq 0 \qquad \forall \lambda \geq 0\end{aligned}$$

Es decir, se cumple la condición de optimalidad (1.30e) □

Observación 1.8. *Como consecuencia de [7, Teorema 3.5] se tiene la siguiente caracterización del gradiente del funcional de costo reducido:*

$$\nabla \mathcal{F}(\lambda) = 2\beta\lambda + \int_\Omega \phi'(u, f)p \ dx \tag{1.36}$$

Esta caracterización será de vital importancia en la resolución numérica de nuestro problema

Observación 1.9. *Para el caso $X^d = \mathbb{R}^d$, la condición de optimalidad* (1.30e) *puede reescribirse como el sistema de complementariedad:*

$$\mu_i = 2\beta\lambda_{\gamma_i} + \int_\Omega p_\gamma \phi_i'(u - \gamma, f) \ dx \qquad i = 1, .., d \tag{1.37a}$$

$$\mu_i \geq 0, \quad \lambda_{\gamma_i} \geq 0, \quad \mu_i \lambda_{\gamma_i} = 0 \qquad i = 1, .., d \tag{1.37b}$$

Demostración. Tomando $\lambda = 0$ en (1.30e) tenemos que para cada $i = 1, ..., d$:

$$-2\beta\lambda_{\gamma_i}^2 - \int_\Omega \lambda_{\gamma_i} \phi_i'(u_{\gamma_i}, f)p_{\gamma_i} \ dx \geq 0$$

de la misma manera, si tomamos $\lambda = 2\lambda_{\gamma_i}$ en (1.30e):

$$2\beta\lambda_{\gamma_i}^2 + \int_\Omega \lambda_{\gamma_i} \phi_i'(u_\gamma, f)p_\gamma \ dx \geq 0$$

luego,

$$\begin{aligned}2\beta\lambda_{\gamma_i}^2 + \int_\Omega \lambda_{\gamma_i} \phi_i'(u_\gamma, f)p \ dx &= \lambda_{\gamma_i} \left(2\beta\lambda_{\gamma_i} + \int_\Omega \phi_i'(u_\gamma, f)p \ dx \right) \\ &= 0, \qquad \forall i = 1, .., d\end{aligned}$$

si notamos, para cada $i = 1, .., d$:

$$\mu_i = 2\beta\lambda_{\gamma_i} + \int_\Omega p\phi_i'(u_\gamma, f) \ dx \qquad i = 1, .., d$$

tenemos el resultado.

□

El resultado de este último teorema es de vital importancia para la resolución númerica de nuestro problema.

Estos resultados, junto con el Teorema 1.5 concluyen con la demostración [7, Teorema 3.6] de la existencia de los multiplicadores de Lagrange en el sistema de Optimalidad que caracteriza a las solución del problema original (1.20).

Teorema 1.7 (Sistema de Optimalidad). *Sea $\bar{\lambda} \in X^d$ una solución óptima de (1.20) y un punto de acumulación débil de la sucesión $\{\lambda_\gamma\}$ de soluciones regularizadas de (1.29). Luego, existen multiplicadores $p \in H_0^1(\Omega)$ y $\xi \in (H_0^1(\Omega))'$ tales que se satisface el siguiente sistema de optimalidad:*

$$\varepsilon(D\bar{u}, Dv)_{L^2} + (q, Dv)_{L^2} + \sum_{i=1}^{d} \int_\Omega \bar{\lambda}_i \ \phi_i'(\bar{u}, f) v \ dx = 0 \quad , \quad \forall v \in H_0^1(\Omega) \tag{1.38a}$$

$$\langle q, \nabla \bar{u} \rangle_{\mathbb{R}^2} = |\nabla \bar{u}| \qquad ctp \ \ de \ \ \Omega \tag{1.38b}$$

$$\varepsilon(Dv, Dp)_{L^2} + (\xi, Dv)_{(\nabla H_0^1(\Omega))'} + \sum_{i=1}^{d} \int_\Omega \bar{\lambda}_i \ \phi_i''(\bar{u}, f) p v \ dx = -(g'(\bar{u}), v) \ \ , \ \ \forall v \in H_0^1(\Omega) \tag{1.38c}$$

$$\langle \xi, Dp \rangle_{(\nabla H_0^1(\Omega))'} \geq 0 \tag{1.38d}$$

$$\langle \xi, D\bar{u} \rangle_{(\nabla H_0^1(\Omega))'} = 0 \tag{1.38e}$$

$$2\beta(\bar{\lambda}, \lambda - \bar{\lambda})_{X^d} + \sum_{i=1}^{d} \int_\Omega \phi_i'(\bar{u}, f) p (\lambda_i - \bar{\lambda}) \ dx \geq 0 \quad , \ \forall \lambda \geq 0 \tag{1.38f}$$

1.4.6 Sistema de Optimalidad del Problema Regularizado para el Caso Real

Gracias a la observación 1.8 tenemos que el problema regularizado (1.29) cuando $X^d = \mathbb{R}^d$ puede caracterizarse por el siguiente sistema de optimalidad:

$$\varepsilon(Du_\gamma, Dv)_{L^2} + (q, Dv)_{L^2} + \sum_{i=1}^{d} \int_\Omega \bar{\lambda}_i \ \phi_i'(u_\gamma, f) v \ dx = 0 \quad , \quad \forall v \in H_0^1(\Omega) \tag{1.39a}$$

$$q - h_\gamma(Du_\gamma) = 0 \quad \text{ctp en} \Omega \tag{1.39b}$$

$$\varepsilon(Dv, Dp_\gamma)_{L^2} + (\xi, Dv)_{(H_0^1(\Omega))'} + \sum_{i=1}^{d} \int_\Omega \bar{\lambda}_i \, \phi_i''(u_\gamma, f) pv \, dx = -(g'(u_\gamma), v) \ , \ \forall v \in H_0^1(\Omega) \tag{1.39c}$$

$$\xi - h_\gamma'(Du_\gamma) Dp_\gamma = 0 \tag{1.39d}$$

$$\mu_i = 2\beta\bar{\lambda}_i + \int_\Omega p\phi_i'(u_\gamma, f) \, dx \qquad i = 1, .., d \tag{1.39e}$$

$$\mu_i \geq 0, \ \ \bar{\lambda}_i \geq 0, \ \ \mu_i\bar{\lambda}_i = 0 \qquad i = 1, .., d \tag{1.39f}$$

Y gracias a este, el sistema de Optimalidad para el problema (1.20) en el caso real es el expuesto en el Teorema 1.7

1.4.7 Sistema de Optimalidad para el caso de un polinomio cuadrático con coeficientes positivos

La contribución genuina de este trabajo es la resolución del problema (1.20) para el caso polinomial.

Para ello, supondremos que el parámetro λ es un polinomio cuadrático de la forma

$$\lambda(x, y) = Ax^2 + By^2 + Cxy + Dx + Ey + F \tag{1.40}$$

donde A, B, C, D, E, F son números reales. Además, imponemos como condición sobre los coeficientes que estos sean no negativos, es decir: $A \geq 0, \ B \geq 0, \ C \geq 0, \ D \geq 0, \ E \geq 0, \ F \geq 0$

Nuestra meta será encontrar los parámetros óptimos $\bar{A}, \bar{B}, \bar{C}, \bar{D}, \bar{E}, \bar{F}$ tales que $\bar{\lambda}(x, y) = \bar{A}x^2 + \bar{B}y^2 + \bar{C}xy + \bar{D}x + \bar{E}y + \bar{F}$ sea óptimo.

En particular, de (1.20), este problema es:

$$\min_{\lambda = Ax^2 + By^2 + Cxy + Dx + Ey + F} g(u) + \beta|\lambda^2|_{L^2} \tag{1.41a}$$

$$\varepsilon(Du, D(v-u))_{L^2(\Omega)} + \int_\Omega \lambda(x, y)\phi'(u, f) \, (v-u) \, dx + \int_\Omega |Dv| \, dx - \int_\Omega |Du| \, dx \geq 0 \quad \forall v \in H_0^1 \tag{1.41b}$$

$$\bar{A} \geq 0, \ \bar{B} \geq 0, \ \bar{C} \geq 0, \ \bar{D} \geq 0, \ \bar{E} \geq 0, \ \bar{F} \geq 0 \tag{1.41c}$$

Y el problema regularizado tiene la forma de (1.29)

Teorema 1.8 (Sistema de Optimalidad del Problema Regularizado). *Sean $A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma$ tales que $\lambda_\gamma(x, y) = A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma$ es solución del problema regularizado (1.29), entonces existen multiplicadores $p \in H_0^1(\Omega)$ y $A, B, C, D, E, F \in \mathbb{R}^+$, donde $\mathbb{R}^+ := [0, +\infty)$, tales que se satisface el siguiente sistema:*

$$\varepsilon(Du_\gamma, Dv) + (q, Dv) + \int_\Omega \bar{\lambda}\ \phi'(u_\gamma, f) v\ dx = 0 \quad , \quad \forall v \in H_0^1 \tag{1.42a}$$

$$q - h_\gamma(Du_\gamma) = 0 \tag{1.42b}$$

$$\varepsilon(Dv, Dp) + Dp(\xi, Dv)_{(H_0^1(\Omega))'} + \int_\Omega \bar{\lambda}\ \phi''(u_\gamma, f) v\ p\ dx = -(g'(u_\gamma), v) \quad , \quad \forall v \in H_0^1 \tag{1.42c}$$

$$\xi - h'_\gamma(Du_\gamma) Dp = 0 \tag{1.42d}$$

$$2\beta((A-A_\gamma)\ x^2\ ,\ A_\gamma x^2+B_\gamma y^2+C_\gamma xy+D_\gamma x+E_\gamma y+F_\gamma)+\int_\Omega (A-A_\gamma)\ x^2\ \phi'(u_\gamma, f)\ p\ dx \geq 0, \forall A \in \mathbb{R}^+ \tag{1.43a}$$

$$2\beta((B-B_\gamma)\ y^2\ ,\ A_\gamma x^2+B_\gamma y^2+C_\gamma xy+D_\gamma x+E_\gamma y+F_\gamma)+\int_\Omega (B-B_\gamma)\ y^2\ \phi'(u_\gamma, f))\ p\ dx \geq 0, \forall B \in \mathbb{R}^+ \tag{1.43b}$$

$$2\beta((C-C_\gamma)\ xy\ ,\ A_\gamma x^2+B_\gamma y^2+C_\gamma xy+D_\gamma x+E_\gamma y+F_\gamma)+\int_\Omega (C-C_\gamma)\ xy\ \phi'(u_\gamma, f)\ p\ dx \geq 0, \forall C \in \mathbb{R}^+ \tag{1.43c}$$

$$2\beta((D-D_\gamma)\ x\ ,\ A_\gamma x^2+B_\gamma y^2+C_\gamma xy+D_\gamma x+E_\gamma y+F_\gamma)+\int_\Omega (D-D_\gamma)\ x\ \phi'(u_\gamma, f)\ p\ dx \geq 0, \forall D \in \mathbb{R}^+ \tag{1.43d}$$

$$2\beta((E-E_\gamma)\ y\ ,\ A_\gamma x^2+B_\gamma y^2+C_\gamma xy+D_\gamma x+E_\gamma y+F_\gamma)+\int_\Omega (E-E_\gamma)\ y\ \phi'(u_\gamma, f)\ p\ dx \geq 0, \forall E \in \mathbb{R}^+ \tag{1.43e}$$

$$2\beta((F-F_\gamma)\ ,\ A_\gamma x^2+B_\gamma y^2+C_\gamma xy+D_\gamma x+E_\gamma y+F_\gamma)+\int_\Omega (F-F_\gamma)\ \phi'(u_\gamma, f)\ p\ dx \geq 0, \forall F \in \mathbb{R}^+ \tag{1.43f}$$

Definimos el Lagrangeano:

$$L(u_\gamma, \lambda, p) = g(u_\gamma) + \beta|\lambda_\gamma|^2 + \varepsilon(Du_\gamma, Dp) + (h_\gamma(Du_\gamma), Dp) + \int_\Omega \lambda\phi'(u_\gamma, f)\ p\ dx \tag{1.44}$$

donde $p \in H_0^1(\Omega)$.

Si λ tiene la forma (1.40), el lagrangeano se puede escribir como:

$$\begin{aligned} L(u_\gamma, A_\gamma, B_\gamma, .., F_\gamma, p) = {} & g(u_\gamma) + \beta||A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma||_{L^2}^2 \\ & + \varepsilon(Du_\gamma, Dp) + (h_\gamma(Du_\gamma), Dp) \\ & + \int_\Omega (A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)\ \ \phi'(u_\gamma, f)\ p\ dx \end{aligned}$$

Ahora, sea $v \in H_0^1(\Omega)$, las condiciones necesarias de primer orden son:

$$\frac{\partial L}{\partial u}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma, \bar{p})v = 0 \quad \forall v \in H_0^1(\Omega) \tag{1.45a}$$

$$\frac{\partial L}{\partial p}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma, \bar{p})\ v = 0 \quad \forall v \in H_0^1(\Omega) \tag{1.45b}$$

$$\frac{\partial L}{\partial A}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})(A - A_\gamma) \geq 0 \quad \forall A \in \mathbb{R}^+ \tag{1.45c}$$

$$\frac{\partial L}{\partial B}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})(B - B_\gamma) \geq 0 \quad \forall B \in \mathbb{R}^+ \tag{1.45d}$$

$$\frac{\partial L}{\partial C}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})(C - C_\gamma) \geq 0 \quad \forall C \in \mathbb{R}^+ \tag{1.45e}$$

$$\frac{\partial L}{\partial D}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})(D - D_\gamma) \geq 0 \quad \forall D \in \mathbb{R}^+ \tag{1.45f}$$

$$\frac{\partial L}{\partial E}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})(E - E_\gamma) \geq 0 \quad \forall E \in \mathbb{R}^+ \tag{1.45g}$$

$$\frac{\partial L}{\partial F}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})(F - F_\gamma) \geq 0 \quad \forall F \in \mathbb{R}^+ \tag{1.45h}$$

Calculando la derivada $(1.45\mathrm{a})$, obtenemos las ecuaciones $(1.42\mathrm{a})$ y $(1.42\mathrm{b})$ del sistema de optimalidad (de la misma manera que en el Teorema 1.6). De igual forma: calculando la derivada $(1.45\mathrm{b})$, obtenemos las ecuación adjunta: $(1.42\mathrm{c})$ y $(1.42\mathrm{d})$ del sistema de optimalidad.

Sea $v \in \mathbb{R}$ una dirección, calcularemos la derivada $(1.45\mathrm{c})$

Para ello, hacemos primero:

$$\begin{aligned}
L(u_\gamma, A_\gamma + h, B_\gamma, .., F_\gamma, p) - L(u_\gamma, A_\gamma, .., F_\gamma, p) = \\
&\beta||(A_\gamma + h)x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma||^2_{L^2} \\
&+ \int_\Omega [(A_\gamma + h)x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma]\ \phi'(u, f)\ p\ dx \\
&- \beta||A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma||_{L^2} \\
&- \int_\Omega (A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)\ \phi'(u, f)\ p\ dx \\
&= \beta(hv\ x^2, h\ v\ x^2 + 2[A_\gamma x^2 B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma])_{L^2} \\
&- \int_\Omega [hv\ x^2\ \phi'(u, f)\ p]\ dx
\end{aligned}$$

Dividiendo por h:

$$\begin{aligned}
\frac{1}{h}[L(u_\gamma, A_\gamma + h, B_\gamma, .., F_\gamma, p) - L(u_\gamma, A_\gamma, .., F_\gamma, p)] = \\
&\beta(v\ x^2, h\ v\ x^2 + 2[A_\gamma x^2 B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma])_{L^2} \\
&- \int_\Omega [v\ x^2\ \phi'(u, f)\ p]\ dx
\end{aligned}$$

y, cuando $h \to 0$

$$\frac{\partial L}{\partial A}(u_\gamma, A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma \bar{p})\ v = 2\beta(vx^2, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega [vx^2\ \phi'(u, f)\ p]\ dx$$

Realizando un procedimiento similar para calcular las derivadas de $(1.45\mathrm{c})$ a $(1.45\mathrm{h})$,

obtenemos los siguientes resultados para direcciones $v \in \mathbb{R}$:

$$\frac{\partial L}{\partial A} v = 2\beta(v\ x^2, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega v\ x^2 \phi'(u)\ p\ dx \tag{1.46a}$$
$$\frac{\partial L}{\partial B} v = 2\beta(v\ y^2, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega v\ y^2 \phi'(u)\ p\ dx \tag{1.46b}$$
$$\frac{\partial L}{\partial C} v = 2\beta(v\ xy, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega v\ xy \phi'(u)\ p\ dx \tag{1.46c}$$
$$\frac{\partial L}{\partial D} v = 2\beta(v\ x, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega v\ x \phi'(u)\ p\ dx \tag{1.46d}$$
$$\frac{\partial L}{\partial E} v = 2\beta(v\ y, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega v\ y \phi'(u)\ p\ dx \tag{1.46e}$$
$$\frac{\partial L}{\partial F} v = 2\beta(v, A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega v\ \phi'(u)\ p\ dx \tag{1.46f}$$

Ahora, si es que en (1.46) tomamos $v = A - A_\gamma$, $v = B - B_\gamma$, $v = C - C_\gamma$, $v = D - D_\gamma$, $v = E - E_\gamma$, $v = F - F_\gamma$ según corresponda, obtenemos el sistema (1.43).

Observación 1.10. *Por un razonamiento similar a la observación 1.9, el sistema* (1.43) *puede reescribirse como el sistema de complementaridad:*

$$\mu_1 = 2\beta(x^2\ ,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega x^2\ \phi'(u)\ p\ dx \tag{1.47a}$$

$$\mu_2 = 2\beta(y^2\ ,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega y^2\ \phi'(u)\ p\ dx \tag{1.47b}$$

$$\mu_3 = 2\beta(xy\ ,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega xy\ \phi'(u)\ p\ dx \tag{1.47c}$$

$$\mu_4 = 2\beta(x\ ,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega x\ \phi'(u)\ p\ dx \tag{1.47d}$$

$$\mu_5 = 2\beta(y\ ,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega y\ \phi'(u)\ p\ dx \tag{1.47e}$$

$$\mu_6 = 2\beta(1,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} + \int_\Omega \phi'(u)\ p\ dx \tag{1.47f}$$

$$\mu_i \geq 0\ \ \forall i = 1,2,3,4,5,6\ \ ,\ \ A_\gamma \geq 0, B_\gamma \geq 0, C_\gamma \geq 0, D_\gamma \geq 0, E_\gamma \geq 0, F_\gamma \geq 0 \tag{1.47g}$$
$$A_\gamma \mu_1 = 0,\ B_\gamma \mu_2 = 0,\ C_\gamma \mu_3 = 0,\ D_\gamma \mu_4 = 0,\ E_\gamma \mu_5 = 0,\ F_\gamma \mu_6 = 0$$

Ahora, si en (1.47a) *desarrollamos el producto en* L^2*:*

$$(x^2\ ,\ A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)_{L^2} =$$
$$= \int_\Omega x^2 (A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma)\ dx$$

$$
\begin{aligned}
&= \int_0^1 \int_0^1 (A_\gamma x^4 + B_\gamma x^2 y^2 + C_\gamma x^3 y + D_\gamma x^3 + E_\gamma x^2 y + F_\gamma x^2)\ dy\ dx \\
&= A_\gamma \int_0^1 \int_0^1 x^4\ dy\ dx + B_\gamma \int_0^1 \int_0^1 x^2 y^2\ dy\ dx \\
&+ C_\gamma \int_0^1 \int_0^1 x^3 y\ dy\ dx + D_\gamma \int_0^1 \int_0^1 x^3\ dy\ dx \\
&+ E_\gamma \int_0^1 \int_0^1 x^2 y\ dy\ dx + F_\gamma \int_0^1 \int_0^1 x^2\ dy\ dx \\
&= \frac{1}{5} A_\gamma + \frac{1}{9} B_\gamma + \frac{1}{8} C_\gamma + \frac{1}{4} D_\gamma + \frac{1}{6} E_\gamma + \frac{1}{3} F_\gamma
\end{aligned}
$$

De donde tenemos que:

$$
\mu_1 = 2\beta \left(\frac{1}{5} A_\gamma + \frac{1}{9} B_\gamma + \frac{1}{8} C_\gamma + \frac{1}{4} D_\gamma + \frac{1}{6} E_\gamma + \frac{1}{3} F_\gamma \right) + \int_\Omega x^2\ \phi'(u)\ p\ dx
$$

Realizando un proceso similar en (1.47b) - (1.47f) *tenemos que el sistema de optimalidad* (1.42) - (1.43) *puede ser reescrito como:*

$$
\varepsilon(Du_\gamma, Dv) + (q, Dv) + \int_\Omega \bar{\lambda}\ \phi'(u_\gamma, f) v\ dx = 0 \quad , \quad \forall v \in H_0^1 \tag{1.48a}
$$

$$
q - h_\gamma(Du_\gamma) = 0 \tag{1.48b}
$$

$$
\varepsilon(Dv, Dp) + Dp(\xi, Dv)_{(H_0^1(\Omega))'} + \int_\Omega \bar{\lambda}\ \phi''(u_\gamma, f) v\ p\ dx = -(g'(u_\gamma), v) \quad , \quad \forall v \in H_0^1 \tag{1.48c}
$$

$$
\xi - h'_\gamma(Du_\gamma) Dp = 0 \tag{1.48d}
$$

$$
\mu_1 = 2\beta \left(\frac{1}{5} A_\gamma + \frac{1}{9} B_\gamma + \frac{1}{8} C_\gamma + \frac{1}{4} D_\gamma + \frac{1}{6} E_\gamma + \frac{1}{3} F_\gamma \right) + \int_\Omega x^2\ \phi'(u)\ p\ dx \tag{1.49a}
$$

$$
\mu_2 = 2\beta \left(\frac{1}{9} A_\gamma + \frac{1}{5} B_\gamma + \frac{1}{8} C_\gamma + \frac{1}{6} D_\gamma + \frac{1}{4} E_\gamma + \frac{1}{3} F_\gamma \right) + \int_\Omega y^2\ \phi'(u)\ p\ dx \tag{1.49b}
$$

$$
\mu_3 = 2\beta \left(\frac{1}{8} A_\gamma + \frac{1}{8} B_\gamma + \frac{1}{9} C_\gamma + \frac{1}{6} D_\gamma + \frac{1}{4} E_\gamma + \frac{1}{3} F_\gamma \right) + \int_\Omega xy\ \phi'(u)\ p\ dx \tag{1.49c}
$$

$$
\mu_4 = 2\beta \left(\frac{1}{4} A_\gamma + \frac{1}{6} B_\gamma + \frac{1}{6} C_\gamma + \frac{1}{3} D_\gamma + \frac{1}{4} E_\gamma + \frac{1}{2} F_\gamma \right) + \int_\Omega x\ \phi'(u)\ p\ dx \tag{1.49d}
$$

$$
\mu_5 = 2\beta \left(\frac{1}{6} A_\gamma + \frac{1}{4} B_\gamma + \frac{1}{6} C_\gamma + \frac{1}{4} D_\gamma + \frac{1}{3} E_\gamma + \frac{1}{2} F_\gamma \right) + \int_\Omega y\ \phi'(u)\ p\ dx \tag{1.49e}
$$

$$
\mu_6 = 2\beta \left(\frac{1}{3} A_\gamma + \frac{1}{3} B_\gamma + \frac{1}{4} C_\gamma + \frac{1}{2} D_\gamma + \frac{1}{2} E_\gamma + F_\gamma \right) + \int_\Omega \phi'(u)\ p\ dx \tag{1.49f}
$$

$$
\mu_i \geq 0 \quad \forall i = 1, 2, 3, 4, 5, 6 \quad , \quad A_\gamma \geq 0, B_\gamma \geq 0, C_\gamma \geq 0, D_\gamma \geq 0, E_\gamma \geq 0, F_\gamma \geq 0 \tag{1.49g}
$$

En el siguiente capítulo analizaremos la aproximación numérica para el sistema de optimalidad para un λ real (1.39) así como también para el sistema de optimalidad cuando λ es un polinomio cuadrático.

1.5 Sistema de Optimalidad para el modelo en el que λ es una Función en General

En la sección anterior estudiamos el caso en el que λ es un número real y el caso en el que es un polinomio de segundo grado con parámetros no negativos. Ahora, modificaremos estos modelos para obtener el sistema de optimalidad para una función $Q(\lambda)$.

Sea λ un vector en $\mathbb{R}^n$ de la forma $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)$ a ser optimizado. Definimos $Q(\lambda)$ como combinación de estos parámetros como una función de L^2, que es diferenciable dos veces y es sobreyectiva. Así, el problema (1.20) para este caso consiste en encontrar el vector de valores óptimos $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{R}^n$ y la función $u(\lambda) \in H_0^1(\Omega)$ tales que sean solución del problema de optimización:

$$\min_{(u,\lambda)\in H_0^1(\Omega)\times\mathbb{R}^n} g(u) + \beta|\lambda|^2_{\mathbb{R}^n} \tag{1.50a}$$

sujeta a:

$$\varepsilon(Du, D(v-u))_{L^2(\Omega)} + \int_\Omega Q(\lambda)\phi'(u,f)\,(v-u)\,dx + \int_\Omega |Dv|\,dx - \int_\Omega |Du|\,dx \geq 0 \quad \forall v \in H_0^1 \tag{1.50b}$$

$$Q(\lambda) \geq 0 \quad \text{c.t.p } \Omega \tag{1.50c}$$

De manera análoga a los casos anteriores, planteamos el problema regularizado (1.29):

$$\min_{(u_\gamma,\lambda_\gamma)\in H_0^1(\Omega)\times\mathbb{R}^n} g(u_\gamma) + \beta|\lambda_\gamma|^2_{\mathbb{R}^n} \tag{1.51a}$$

sujeta a:

$$\varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \int_\Omega Q(\lambda_\gamma)\phi'(u_\gamma)v\,dx = 0 \quad \forall v \in H_0^1(\Omega) \tag{1.51b}$$

$$Q(\lambda_\gamma) \geq 0 \quad \text{c.t.p } \Omega \tag{1.51c}$$

En este caso, el funcional de costo $\mathcal{J}(u,\lambda) = g(u) + \beta|\lambda|^2$ esta definido en el espacio de Banach de dimensión infinita $X := H_0^1(\Omega)\times\mathbb{R}^n$ por lo que es necesario verificar la condición de regularidad para garantizar la existencia de un multiplicador de Lagrange [19] [20].

Sea $Y := H^{-1}(\Omega) \times L^2(\Omega)$, agrupando las condiciones (1.51b) y (1.51c) definimos la aplicación $G : X \to Y$ como:

$$G(u_\gamma, \lambda_\gamma) = \begin{pmatrix} \varepsilon(Du_\gamma, D\cdot) + (h_\gamma(Du_\gamma), D\cdot) + \int_\Omega Q(\lambda_\gamma)\phi'(u_\gamma)\cdot\ dx \\ Q(\lambda_\gamma) \end{pmatrix} \tag{1.52}$$

Primero, debemos ver que la región factible $K = \{0\} \times \left\{v \in L^2(\Omega) : \ v \geq 0 \ \text{ ctp } \Omega\right\}$ es un cono convexo cerrado en Y. Para ello, basta tomar dos elementos: $w = (w_1, w_2) \in K$ y $v = (v_1, v_2) \in K$; esto es: $v_1 = w_1 = 0$ y $v_2 \geq 0$, $w_2 \geq 0$.

Sean $\alpha > 0$ y $\beta > 0$:

$$\begin{aligned}\alpha v + \beta w &= \alpha(v_1, v_2) + \beta(w_1, w_2)\\ &= (\alpha v_1, \alpha v_2) + (\beta w_1, \beta w_2)\\ &= (\alpha v_1 + \beta w_1, \alpha v_2 + \beta w_2)\\ \Rightarrow &\alpha v_1 + \beta w_1 = 0,\\ &\alpha v_2 + \beta w_2 \geq 0\end{aligned}$$

Luego, $\alpha v + \beta w \in K$ para todo par de elementos v, w de K y todo par de escalares α, β, de donde el cono K es convexo. Ahora bien, es inmediato ver que 0 es cerrado; por otra parte el conjunto de funciones $\{v \in L^2(\Omega) : \quad v \geq 0 \quad \text{ctp } \Omega\}$ es también cerrado, ya que toda sucesión de funciones mayores que cero converge a una función mayor a cero en casi tipo punto.

Por lo que podemos concluir que K es un cono convexo cerrado de Y.

Definición 1.8 (Punto Regular). *Si llamamos $G'(u_\gamma, \lambda_\gamma)$ a la derivada de Fréchet de G (1.52), un punto factible (u^*, λ^*) del problema de optimización regularizado (1.51) se dice regular si:*

$$0 \in int(G(u^*, \lambda^*) + G'(u^*, \lambda^*)X - K)$$

o, de manera equivalente [19]:

$$G'(u^*, \lambda^*)X - K(G(u^*, \lambda^*)) = Y$$

Ahora bien, basta demostrar que la solución óptima local (u^*, λ^*) de (1.51) es un punto regular para obtener la existencia de los multiplicadores de Lagrange que nos permitirán obtener un sistema de optimalidad para nuestro problema [19].

Teorema 1.9. *Sea (u^*, λ^*) una solución óptima local de (1.51), entonces (u^*, λ^*) es también un punto regular.*

Demostración. Para obtener el resultado, es suficiente verificar que $G'(u_\gamma, \lambda_\gamma) : X \to Y$ es una función sobreyectiva [20]. Es decir, debemos probar que para cada $(y_1, y_2) \in G^{-1}(\Omega) \times L^2(\Omega) = Y$ existe un par $(w_1, w_2) \in H_0^1(\Omega) \times \mathbb{R}^n = X$ tal que $G'(u_\gamma, \lambda_\gamma)(w_1, w_2) = (y_1, y_2)$. La derivada de Fréchet de $G'(u_\gamma, \lambda_\gamma)$ en la dirección (w_1, w_2) está dada por:

$G'(u_\gamma, \lambda_\gamma)(w_1, w_2) =$

$$\begin{pmatrix} \varepsilon(Dw_1, D\cdot) + (h'_\gamma(Du_\gamma)Dw_1, D\cdot) + \int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)w_1 \cdot \ dx + \int_\Omega Q'(\lambda_\gamma)w_2\phi'(u_\gamma) \cdot \ dx \\ Q'(\lambda_\gamma)w_2 \end{pmatrix} \tag{1.53}$$

Sea $y_2 \in L^2(\Omega)$, si fijamos w_2 como la solución de $Q'(\lambda_\gamma)w_2 = y_2$, tendríamos que probar que dado $y_1 \in H^{-1}(\Omega)$ existe $w_1 \in H_0^1(\Omega)$ tal que, para todo $v \in H_0^1(\Omega)$:

$$\langle y_1, v\rangle - \int_\Omega Q'(\lambda_\gamma)\, w_2\, \phi'(u_\gamma)\, v\, dx = \varepsilon(Dw_1, Dv) + (h'_\gamma(Du_\gamma)Dw_1, Dv) + \int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)w_1\, v\, dx$$

La existencia de $w_1 \in H_0^1(\Omega)$ está garantizada por el teorema de Lax-Milgram, pues si definimos la forma bilineal:

$$a(w_1, v) = \varepsilon(Dw_1, Dv) + (h'_\gamma(Du_\gamma)Dw_1, Dv) + \int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)w_1\ v\ dx$$

esta es continua y coerciva en $H_0^1(\Omega)$, como se probará a continuación.

Para comprobar la continuidad, vemos primero que la integral:

$$\int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)w_1\ v\ dx$$

está bien definida. En efecto: dado que, $\phi : L^p(\Omega) \to \mathbb{R}$ con $p \geq 1$, su segunda derivada es $\phi''(u) : L^p(\Omega) \to L^{p'}(\Omega)$ con $\frac{1}{p} + \frac{1}{p'} = 1$.

En particular, si tomamos $p = \frac{3}{2}$: $\phi''(u) : L^{\frac{3}{2}}(\Omega) \to L^3(\Omega)$, de donde para $w_1 \in H_0^1(\Omega) \hookrightarrow L^{\frac{3}{2}}(\Omega)$, $\phi''(u)[w_1] \in L^3(\Omega)$

Sea $v \in H_0^1(\Omega)$, en particular $v \in H_0^1(\Omega) \hookrightarrow L^6(\Omega)$ y $\phi''(u)[w_1]v \in L^2(\Omega)$. Además, como $Q(\lambda_\gamma) \in L^2(\Omega)$, tenemos que el producto $Q(\lambda_\gamma)\phi''(u_\gamma)w_1\ v$ pertenece a $L^1(\Omega)$ (pues aplicando la observación 2 del teorema 4.6 en [21] tenemos que $\frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1$). Con lo que la integral

$$\int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)w_1\ v\ dx$$

está bien definida.

Además, notemos que $h'_\gamma(Du_\gamma)Dw_1 \in L^\infty(\Omega)$; en efecto, derivando (1.28) tenemos que:

$$h'_\gamma(Du)Dv = \begin{cases} \frac{Dv}{|Du|} - \frac{Du\langle Dv, Du\rangle}{|Du|^3} & \text{en } A^\gamma \\ \frac{Dv}{|Du|}\left(1 - \frac{\gamma}{2}B^2\right) + \frac{Du\langle Dv, Du\rangle}{|Du|^2}\left(-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right) & \text{en } S^\gamma \\ \gamma Dv & \text{en } I^\gamma \end{cases} \tag{1.54}$$

donde $B = \left(1 - \gamma|Du| + \frac{1}{2\gamma}\right)$. Se conserva la notación

$$\mathcal{A}^\gamma := \left\{x \in \Omega\ :\ \gamma|Du| - 1 \geq \frac{1}{2\gamma}\right\} \quad \text{es el conjunto activo}$$

$$\mathcal{S}^\gamma := \left\{x \in \Omega\ :\ |\gamma|Du| - 1| \leq \frac{1}{2\gamma}\right\} \quad \text{es el conjunto semi-activo}$$

y

$$\mathcal{I}^\gamma := \left\{x \in \Omega\ :\ 1 - \gamma|Du| > \frac{1}{2\gamma}\right\} = \mathcal{S} \setminus (\mathcal{A}^\gamma \cup \mathcal{S}^\gamma) \quad \text{es el conjunto inactivo}$$

<u>En A^γ</u> tenemos que: $\frac{1}{|Du|} \leq \frac{2\gamma^2}{1+2\gamma}$, de donde:

$$|h'_\gamma(Du)Dv| = \left|\frac{Dv}{|Du|} - \frac{Du\,\langle Dv, Du\rangle}{|Du|^3}\right|$$

$$
\begin{aligned}
&\leq \left|\frac{Dv}{|Du|}\right| + \left|\frac{Du\,\langle Dv, Du\rangle}{|Du|^3}\right| \\
&\leq \frac{|Dv|}{|Du|} + \frac{|Du|\,|\langle Dv, Du\rangle|}{|Du|^3} \\
&\leq \frac{|Dv|}{|Du|} + \frac{|Du|^2|Dv|}{|Du|^3} \\
&= 2\frac{|Dv|}{|Du|} \leq \frac{4|Dv|\gamma^2}{1+2\gamma}
\end{aligned}
$$

<u>En S^γ</u> tenemos que $\frac{1}{|Du|} \leq \frac{2\gamma^2}{2\gamma-1}$, de donde se obtiene que $B = \left(1 - \gamma|Du| + \frac{1}{2\gamma}\right)$ cumple con: $0 \leq B \leq \frac{1}{\gamma}$. Además, como $\gamma \geq 1$,

$$
0 \leq B \leq \frac{1}{\gamma} \leq 1
$$

luego: $0 \leq B^2 \leq \frac{1}{\gamma^2} \leq 1$

Por otro lado,

$$
\begin{aligned}
\left|-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right| &= \left|-\frac{1}{|Du|}\left(1 - \frac{\gamma}{2}\right) + \gamma^2 B\right| \\
&\leq \left|-\frac{1}{|Du|}\left(1 - \frac{\gamma}{2}\right)\right| + \left|\gamma^2 B\right| \\
&\leq \frac{1}{|Du|}\left|1 - \frac{\gamma}{2}\right| + \left|\gamma^2 B\right| \\
&\leq \frac{2\gamma^2}{2\gamma - 1}\left(1 + \frac{\gamma}{2}\right) + \gamma^2 := c_1
\end{aligned}
$$

Ahora, podemos calcular:

$$
\begin{aligned}
|h'_\gamma(Du)Dv| &= \left|\frac{Dv}{|Du|}\left(1 - \frac{\gamma}{2}B^2\right) + \frac{Du\,\langle Dv, Du\rangle}{|Du|^2}\left(-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right)\right| \\
&\leq \left|\frac{Dv}{|Du|}\left(1 - \frac{\gamma}{2}B^2\right)\right| + \left|\frac{Du\,\langle Dv, Du\rangle}{|Du|^2}\left(-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right)\right| \\
&\leq \frac{|Dv|}{|Du|}\left|1 - \frac{\gamma}{2}B^2\right| + \frac{|Du|^2|Dv|}{|Du|^2}\left|-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right| \\
&\leq \frac{2\gamma^2}{2\gamma - 1}\left(1 + \frac{\gamma}{2}\right)|Dv| + c_1|Dv| := \widetilde{c_1}|Dv|
\end{aligned}
$$

Finalmente, <u>en I^γ</u>:

$$
|h'_\gamma(Du)Dv| = |\gamma|Dv|| = \gamma|Dv|
$$

Por lo que se puede concluir que

$$
|h'_\gamma(Du)Dv| \leq C|Dv|,
$$

donde la constante $C \geq 0$ es $C := \text{máx}\left\{\frac{4\gamma^2}{1+2\gamma},\ \widetilde{c_1},\ \gamma\right\}$, y $h'_\gamma(Du) \in L^\infty(\Omega)$.

Con estos resultados, probamos la continuidad de la forma bilineal $a(w_1, v)$:

$$\begin{aligned}
a(w_1, v) &= \varepsilon(Dw_1, Dv)_{L^2(\Omega)} + (h'_\gamma(Du_\gamma)Dw_1, Dv)_{L^2(\Omega)} + (Q(\lambda_\gamma), \phi''(u_\gamma)w_1\ v)_{L^2(\Omega)} \\
&\leq \varepsilon|(Dw_1, Dv)_{L^2(\Omega)}| + |(h'_\gamma(Du_\gamma)Dw_1, Dv)_{L^2(\Omega)}| + |(Q(\lambda_\gamma), \phi''(u_\gamma)w_1\ v)_{L^2(\Omega)}| \\
&\leq \varepsilon||Dw_1||_{L^2(\Omega)}||Dv||_{L^2(\Omega)} + ||h'_\gamma(Du_\gamma)Dw_1||_{L^2(\Omega)}||Dv||_{L^2(\Omega)} \\
&+ ||Q(\lambda_\gamma)||_{L^2(\Omega)}||\phi''(u_\gamma)w_1\ v||_{L^2(\Omega)}
\end{aligned}$$

Sabemos que existe una constante positiva K_1 (puede depender de u_γ) tal que:

$$||\phi''(u_\gamma)w_1||_{L^3(\Omega)} \leq K_1||w_1||_{H_0^1(\Omega)}$$

Además, si llamamos $K_2 := ||Q(\lambda_\gamma)||_{L^2(\Omega)}$ tenemos que:

$$\begin{aligned}
a(w_1, v) &\leq \varepsilon||w_1||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)} + ||h'_\gamma(Du_\gamma)||_{L^\infty(\Omega)}||Dw_1||_{L^2(\Omega)}||Dv||_{L^2(\Omega)} \\
&+ K_2||\phi''(u_\gamma)w_1||_{L^3(\Omega)}||v||_{H_0^1(\Omega)} \\
&\leq \varepsilon||w_1||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)} + C||w_1||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)} \\
&+ K_1K_2||w_1||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)} \\
&\leq (\varepsilon + C + K_1K_2)||w_1||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)} \\
&= \widetilde{K}||w_1||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)}
\end{aligned}$$

Con $\widetilde{K} := \varepsilon + C + K_1K_2 > 0$ queda demostrada la continuidad de la forma bilineal.

Para continuar con las hipótesis del Teorema de Lax Milgram, es necesario probar la coercividad de:

$$a(v, v) = \varepsilon(Dv, Dv)_{L^2(\Omega)} + (h'_\gamma(Du_\gamma)Dv, Dv)_{L^2(\Omega)} + (Q(\lambda_\gamma), \phi''(u_\gamma)\ v^2)_{L^2(\Omega)}$$

para $v \in H_0^1(\Omega)$.

Para esto, hay que notar primero que, gracias a que $\phi(u)$ es convexa, $\phi''(u_\gamma)\ v^2 \geq 0$, y como $Q(\lambda_\gamma)$ es no negativa, entonces $(Q(\lambda_\gamma), \phi''(u_\gamma)\ v^2)_{L^2(\Omega)} \geq 0$.

A continuación veremos que $(h'_\gamma(Du_\gamma)Dv, Dv)_{L^2(\Omega)} \geq 0$ a partir de un análisis de (1.54).

En $\mathcal{A}^\gamma$ se tiene, como consecuencia de la desigualdad de Cauchy- Schwartz:

$$\begin{aligned}
(h'_\gamma(Du_\gamma)Dv, Dv)_{L^2(\Omega)} &= \frac{\langle Dv, Dv\rangle}{|Du|} - \frac{\langle Dv, Du\rangle^2}{|Du|^3} \\
&= \frac{|Dv|^2}{|Du|} - \frac{\langle Dv, Du\rangle^2}{|Du|^3} \\
&= \frac{|Dv|^2|Du|^2 - \langle Dv, Du\rangle^2}{|Du|^3} \geq 0
\end{aligned}$$

Usando el mismo razonamiento, y las propiedades del conjunto S^γ se tiene que, en este intervalo:

$$\begin{aligned}(h'_\gamma(Du_\gamma)Dv, Dv)_{L^2(\Omega)} &= \frac{\langle Dv, Dv\rangle}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) + \frac{\langle Dv, Du\rangle^2}{|Du|^2}\left(-\frac{1}{|Du|}+\frac{1}{|Du|}\frac{\gamma}{2}B^2+\gamma^2 B\right)\\ &= \frac{|Dv|^2}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) + \frac{\langle Dv, Du\rangle^2}{|Du|^2}\left(-\frac{1}{|Du|}+\frac{1}{|Du|}\frac{\gamma}{2}B^2+\gamma^2 B\right)\\ &\geq 0\end{aligned}$$

y de manera inmediata, para el conjunto inactivo I^γ:

$$\begin{aligned}(h'_\gamma(Du_\gamma)Dv, Dv)_{L^2(\Omega)} &= \gamma\,\langle Dv, Dv\rangle\\ &\geq 0\end{aligned}$$

Con lo que podemos probar la coercividad de la forma bilineal:

$$\begin{aligned}a(v,v) &= \varepsilon(Dv, Dv) + (h'_\gamma(Du_\gamma)Dv, Dv) + (Q(\lambda_\gamma)\phi''(u_\gamma)v\cdot\ v)\\ &= \varepsilon||Dv||^2 + (h'_\gamma(Du_\gamma)Dv, Dv) + (Q(\lambda_\gamma)\phi''(u_\gamma)v\cdot\ v)\\ &\geq \varepsilon||Dv||^2_{L^2(\Omega)}\\ &= c||v||^2_{H^1_0(\Omega)}\end{aligned}$$

con c>0

□

1.5.1 Sistema de Optimalidad

Definamos el Cono Polar de un subconjunto $A \subset X$ como:

$$A^+ = \left\{z \in X' : \langle z, a\rangle_{X',X} \geq 0\ ,\ \forall a \in A\right\}$$

Ahora, el cono polar de K es:

$$K^+ = \left\{\begin{pmatrix} z_1\\ z_2\end{pmatrix} \in H^1_0(\Omega)\times L^2(\Omega) : \begin{matrix}\langle z_1, \varphi_1\rangle_{H^1_0(\Omega),H^{-1}(\Omega)} \geq 0 & \forall\varphi_1 = 0\\ (z_2, \varphi_2)_{L^2(\Omega)} \geq 0 & \forall\varphi_2 \geq 0 \text{ c.t.p } \Omega\end{matrix}\right\}$$

que se puede reducir al conjunto: $K = \left\{(z_1, z_2) : z_1 \in H^1_0(\Omega)\ \text{ y }\ z_2 \in L^2(\Omega) \text{ tal que } z_2 \geq 0 \text{ c.t.p } \Omega\right\}$

Gracias al resultado obtenido en el teorema 1.9, podemos formular el siguiente teorema:

Teorema 1.10 (Sistema de Optimalidad asociado al problema regularizado en el que el parámetro λ es una función). *Sea $(u_\gamma, \lambda_\gamma)$ una solución óptima local del problema de optimización regularizado* (1.51)*, entonces existen multiplicadores de Lagrange $p \in H^1_0(\Omega)$ y $\mu \in L^2(\Omega)$, tal que $\mu \geq 0$ c.t.p Ω, tales que se tiene el siguiente sistema de optimalidad:*

$$\varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \int_\Omega Q(\lambda_\gamma)\phi'(u_\gamma)v \; dx = 0 \quad \forall v \in H_0^1 \tag{1.55a}$$

$$\varepsilon(Dv, Dp) + (h'_\gamma(Du_\gamma)Dv, Dp) + \int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)v \; p \; dx = (g'(u_\gamma), v) \quad , \quad \forall v \in H_0^1 \tag{1.55b}$$

$$\int_\Omega Q'(\lambda_\gamma)\phi'(u_\gamma)p \; dx + \int_\Omega Q'(\lambda_\gamma)\mu \; dx = 2\beta\lambda_\gamma \tag{1.55c}$$

$$\int_\Omega Q(\lambda_\gamma)\mu \; dx = 0 \tag{1.55d}$$

$$Q(\lambda_\gamma) \geq 0 \qquad c.t.p \; \Omega \tag{1.55e}$$

Demostración. Notemos el funcional de costo de nuestro problema regularizado (1.51)

$$\mathcal{J}(u_\gamma, \lambda_\gamma) = g(u_\gamma) + \beta|\lambda_\gamma|^2$$

De acuerdo a [19] [20] los multiplicadores de lagrange $p \in H_0^1(\Omega)$ y $\mu \in L^2(\Omega)$, tal que $\mu \geq 0$ c.t.p Ω, cumplen con el sistema:

$$\mathcal{J}'(u_\gamma, \lambda_\gamma)(v, h) - \langle (p, \mu), G'(u_\gamma, \lambda_\gamma)(v, h)\rangle_{H_0^1(\Omega)\times L^2(\Omega), H^{-1}(\Omega)\times L^2(\Omega)} = 0 \tag{1.56a}$$

$$\langle (p, \mu), G(u_\gamma, \lambda_\gamma)\rangle_{H_0^1(\Omega)\times L^2(\Omega), H^{-1}(\Omega)\times L^2(\Omega)} = 0 \tag{1.56b}$$

para todo par $(v, h) \in H_0^1(\Omega) \times \mathbb{R}^n$. $G(u_\gamma, \lambda_\gamma)$ se definió en (1.52) y su derivada $G'(u_\gamma, \lambda_\gamma)$ en (1.53).

Derivando $\mathcal{J}$ en la dirección $(v, h) \in H_0^1(\Omega) \times \mathbb{R}^n$ obtenemos:

$$\mathcal{J}'(u_\gamma, \lambda_\gamma)(v, h) = (g'(u_\gamma), v) + 2\beta\lambda_\gamma h$$

Por otro lado, notemos que el producto en dualidad $(H_0^1(\Omega) \times L^2(\Omega), H^{-1}(\Omega) \times L^2(\Omega))$ no es más que la suma de los productos en dualidad de cada componente. Sean U y W dos espacios de Hilbert, si notamos la suma directa de dichos espacios como $V = U \times W$, el producto en dualidad se define como:

$$\langle (a_1, b_1), (a_2, b_2)\rangle_{V'\times V} = \langle a_1, a_2\rangle_{U',U} + \langle b_1, b_2\rangle_{W',W}$$

Por lo que, el producto (1.56b) nos queda:

$$\begin{aligned}\langle (p, \mu), G(u_\gamma, \lambda_\gamma)\rangle_{H_0^1(\Omega)\times L^2(\Omega), H^{-1}(\Omega)\times L^2(\Omega)} = \\ \varepsilon(Du_\gamma, Dp) + (h_\gamma(Du_\gamma), Dp) + \int_\Omega Q(\lambda_\gamma)\phi'(u_\gamma)p \; dx \\ + \langle Q(\lambda_\gamma), \mu\rangle_{L^2(\Omega)} = 0\end{aligned} \tag{1.57}$$

de donde sale la condición de complementariedad (1.55d) (aplicando la ecuación de estado).

De igual manera, desarrollando el producto en dualidad, (1.56a) nos queda:

$$\begin{aligned}
\mathcal{J}'(u_\gamma,\lambda_\gamma)(v,h) &- \langle (p,\mu), G'(u_\gamma,\lambda_\gamma)(v,h)\rangle_{H_0^1(\Omega)\times L^2(\Omega),H^{-1}(\Omega)\times L^2(\Omega)} \\
&= \mathcal{J}'(u_\gamma,\lambda_\gamma)(v,h) \\
&- \langle (p,\mu), (G_1'(u_\gamma,\lambda_\gamma)(v,h), G_2'(u_\gamma,\lambda_\gamma)(v,h)\rangle_{H_0^1(\Omega)\times L^2(\Omega),H^{-1}(\Omega)\times L^2(\Omega)} \\
&= \mathcal{J}'(u_\gamma,\lambda_\gamma)(v,h) - \langle p, G_1'(u_\gamma,\lambda_\gamma)(v,h)\rangle_{H_0^1(\Omega),H^{-1}(\Omega)} \\
&- \langle \mu, G_2'(u_\gamma,\lambda_\gamma)(v,h)\rangle_{L^2(\Omega)} \\
&= (g'(u_\gamma),v) + 2\beta\lambda_\gamma h - \varepsilon(Dv,Dp) - (h_\gamma'(Du_\gamma)Dv,Dp) - \int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)v\ p\ dx \\
&- \int_\Omega Q'(\lambda_\gamma)h\phi'(u_\gamma)p\ dx - \int_\Omega Q'(\lambda_\gamma)h\mu\ dx = 0 \qquad (1.58)
\end{aligned}$$

$\forall v \in H_0^1(\Omega), \quad \forall h \in \mathbb{R}^n$

Si en (1.58) $v = 0$:

$$2\beta\lambda_\gamma h - \int_\Omega Q'(\lambda_\gamma)h\phi'(u_\gamma)p\ dx - \int_\Omega Q'(\lambda_\gamma)h\mu\ dx = 0$$

que se puede expresar también como:

$$2\beta\lambda_\gamma - \int_\Omega Q'(\lambda_\gamma)\phi'(u_\gamma)p\ dx - \int_\Omega Q'(\lambda_\gamma)\mu\ dx = 0$$

y así se obtiene (1.55c).

Ahora, si en (1.58) $h = 0$:

$$(g'(u_\gamma),v) - \varepsilon(Dv,Dp) - (h_\gamma'(Du_\gamma)Dv,Dp) - \int_\Omega Q(\lambda_\gamma)\phi''(u_\gamma)v\ p\ dx = 0,$$

para todo $v \in H_0^1(\Omega)$ de donde se obtiene la ecuación adjunta (1.55b) de nuestro sistema de optimalidad. □

A continuación, veremos dos ejemplos específicos de este problema y escribiremos sus sistemas de optimalidad asociados.

1.5.2 Problema de Optimización donde el parámetro λ es un polinomio cuadrático con valores reales

Con el método visto en esta sección, veremos el sistema de optimalidad para un modelo en el que el ruido λ tiene la forma de un polinomio cuadrátido $Ax^2 + By^2 + Cxy + Dx + Ey + F$. La diferencia con el modelo visto anteriormente en (1.41) es que ahora no tendremos las restricciones de positividad sobre los coeficientes A, B, C, D, E y F, sino que la restricción de no negatividad será para toda la función $Q(\lambda) = Ax^2 + By^2 + Cxy + Dx + Ey + F \geq 0$ dejando la posibilidad de obtener respuestas en las que algunos parámetros tengan signo negativo sin que esto afecte a la positividad del polinomio cuadrático.

Así, si llamamos $\lambda = (A, B, C, D, E, F) \in \mathbb{R}^6$, planteamos el problema (1.50) para este caso específico:

$$\min_{(u,\lambda)\in H_0^1(\Omega)\times\mathbb{R}^6} g(u) + \beta|\lambda|^2 \tag{1.59a}$$

sujeta a:

$$\begin{aligned}\varepsilon(Du, D(v-u))_{L^2(\Omega)} + \int_\Omega [Ax^2 + By^2 + Cxy + Dx + Ey + F]\phi'(u,f)\,(v-u)\,dx \\ + \int_\Omega |Dv|\,dx - \int_\Omega |Du|\,dx \geq 0 \quad \forall v \in H_0^1\end{aligned} \tag{1.59b}$$

$$Ax^2 + By^2 + Cxy + Dx + Ey + F \geq 0 \quad \text{c.t.p } \Omega \tag{1.59c}$$

y su versión regularizada, a partir de (1.51):

$$\min_{(u_\gamma,\lambda_\gamma)\in H_0^1(\Omega)\times\mathbb{R}^6} g(u_\gamma) + \beta|\lambda_\gamma|^2 \tag{1.60a}$$

sujeta a:

$$\varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \int_\Omega [A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma]\phi'(u_\gamma)v\,dx = 0 \quad \forall v \in H_0^1(\Omega) \tag{1.60b}$$

$$A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma \geq 0 \quad \text{c.t.p } \Omega \tag{1.60c}$$

donde $\lambda_\gamma = (A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma)$.

Luego, obtenemos el sistema de optimalidad (visto en el teorema 1.10 de forma general) asociado a este problema.

Teorema 1.11 (Sistema de Optimalidad asociado al problema regularizado en el que el parámetro λ es un polinomio cuadrático). *Sea $(u_\gamma, \lambda_\gamma)$ una solución óptima local del problema de optimización regularizado (1.60), donde $\lambda_\gamma = (A_\gamma, B_\gamma, C_\gamma, D_\gamma, E_\gamma, F_\gamma)$, entonces existen multiplicadores de Lagrange $p \in H_0^1(\Omega)$ y $\mu \in L^2(\Omega)$, tal que $\mu \geq 0$ c.t.p Ω, tales que se tiene el siguiente sistema de optimalidad:*

$$\varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \int_\Omega [A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma]\phi'(u_\gamma)v\,dx = 0 \quad \forall v \in H_0^1 \tag{1.61a}$$

$$\begin{aligned}\varepsilon(Dv, Dp) + (h'_\gamma(Du_\gamma)Dv, Dp) \\ + \int_\Omega [A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma]\ \phi''(u_\gamma)v\ p\ dx \\ = (g'(u_\gamma), v) \quad , \forall v \in H_0^1\end{aligned} \tag{1.61b}$$

$$\begin{pmatrix} \int_\Omega x^2[\phi'(u_\gamma)p+\mu]\ dx \\ \int_\Omega y^2[\phi'(u_\gamma)p+\mu]\ dx \\ \int_\Omega xy\ [\phi'(u_\gamma)p+\mu]\ dx \\ \int_\Omega x\ [\phi'(u_\gamma)p+\mu]\ dx \\ \int_\Omega y\ [\phi'(u_\gamma)p+\mu]\ dx \\ \int_\Omega [\phi'(u_\gamma)p+\mu]\ dx \end{pmatrix} = 2\beta \begin{pmatrix} A_\gamma \\ B_\gamma \\ C_\gamma \\ D_\gamma \\ E_\gamma \\ F_\gamma \end{pmatrix} \tag{1.61c}$$

$$\int_\Omega [A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma] \cdot \mu\ dx = 0 \tag{1.61d}$$

$$[A_\gamma x^2 + B_\gamma y^2 + C_\gamma xy + D_\gamma x + E_\gamma y + F_\gamma] \geq 0 \qquad c.t.p\ \Omega \tag{1.61e}$$

1.5.3 Problema de Optimización donde el parámetro λ es una combinación de funciones trigonométricas

Ahora, aplicaremos el método visto en esta sección para encontrar el sistema de optimalidad asociado a un problema en el que el ruido λ está modelado como una función trigonométrica de la forma $\lambda_1 \sin(\lambda_2\ x) + \lambda_3 \cos(\lambda_4\ xy)$. Aquí, la restricción de no negatividad se aplica sobre la función: $Q(\lambda) = \lambda_1 \sin(\lambda_2\ x) + \lambda_3 \cos(\lambda_4\ xy) \geq 0$ mientras que los coeficientes $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ pueden tomar el valor de cualquier número real. Así, si llamamos $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{R}^4$, podemos plantear el problema de optimización general (1.50) para este caso específico:

$$\min_{(u,\lambda)\in H_0^1(\Omega)\times\mathbb{R}^4} g(u) + \beta|\lambda|^2 \tag{1.62a}$$

sujeta a:

$$\begin{aligned} \varepsilon(Du, D(v-u))_{L^2(\Omega)} + \int_\Omega [\lambda_1 \sin(\lambda_2\ x) + \lambda_3 \cos(\lambda_4\ xy)]\phi'(u,f)\ (v-u)\ dx \\ + \int_\Omega |Dv|\ dx - \int_\Omega |Du|\ dx \geq 0 \quad \forall v \in H_0^1 \end{aligned} \tag{1.62b}$$

$$\lambda_1 \sin(\lambda_2\ x) + \lambda_3 \cos(\lambda_4\ xy) \geq 0 \quad \text{c.t.p}\ \ \Omega \tag{1.62c}$$

y su versión regularizada, a partir de (1.51):

$$\min_{(u_\gamma,\lambda_\gamma)\in H_0^1(\Omega)\times\mathbb{R}^4} g(u_\gamma) + \beta|\lambda_\gamma|^2 \tag{1.63a}$$

sujeta a:

$$\varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \int_\Omega [\lambda_{1\gamma} \sin(\lambda_{2\gamma}\ x) + \lambda_{3\gamma} \cos(\lambda_{4\gamma}\ xy)]\phi'(u_\gamma)v\ dx = 0 \quad \forall v \in H_0^1(\Omega) \tag{1.63b}$$

$$\lambda_{1\gamma} \sin(\lambda_{2\gamma}\ x) + \lambda_{3\gamma} \cos(\lambda_{4\gamma}\ xy) \geq 0 \quad \text{c.t.p}\ \ \Omega \tag{1.63c}$$

donde $\lambda_\gamma = (\lambda_{1\gamma}, \lambda_{2\gamma}, \lambda_{3\gamma}, \lambda_{4\gamma})$

De donde obtenemos el siguiente sistema de optimalidad que fue visto de manera general en el teorema 1.10

Teorema 1.12 (Sistema de Optimalidad asociado al problema regularizado en el que el parámetro λ es una combinación de funciones trigonométricas). *Sea* $(u_\gamma, \lambda_\gamma)$ *una solución óptima local del problema de optimización regularizado* (1.63), *donde* $\lambda_\gamma = (\lambda_{1\gamma}, \lambda_{2\gamma}, \lambda_{3\gamma}, \lambda_{4\gamma}) \in \mathbb{R}^4$, *entonces existen multiplicadores de Lagrange* $p \in H_0^1(\Omega)$ *y* $\mu \in L^2(\Omega)$, *tal que* $\mu \geq 0$ *c.t.p* Ω, *tales que se tiene el siguiente sistema de optimalidad:*

$$\varepsilon(Du_\gamma, Dv) + (h_\gamma(Du_\gamma), Dv) + \int_\Omega [\lambda_{1\gamma} \sin(\lambda_{2\gamma}\ x) + \lambda_{3\gamma} \cos(\lambda_{4\gamma}\ xy)]\phi'(u_\gamma)v\ dx = 0 \qquad \forall v \in H_0^1(\Omega) \quad (1.64a)$$

$$\varepsilon(Dv, Dp) + (h'_\gamma(Du_\gamma)Dv, Dp) + \int_\Omega [\lambda_{1\gamma} \sin(\lambda_{2\gamma}\ x) + \lambda_{3\gamma} \cos(\lambda_{4\gamma}\ xy)]\ \phi''(u_\gamma)v\ p\ dx = (g'(u_\gamma), v) \quad , \forall v \in H_0^1 \quad (1.64b)$$

$$\begin{pmatrix} \int_\Omega \sin(\lambda_{2\gamma}\ x)\ [\phi'(u_\gamma)p + \mu]\ dx \\ \int_\Omega \lambda_{1\gamma} x \cos(\lambda_{2\gamma} x)\ [\phi'(u_\gamma)p + \mu]\ dx \\ \int_\Omega \cos(\lambda_{4\gamma}\ xy)\ [\phi'(u_\gamma)p + \mu]\ dx \\ -\int_\Omega \lambda_{3\gamma} xy \sin(\lambda_{4\gamma}\ xy)\ [\phi'(u_\gamma)p + \mu]\ dx \end{pmatrix} = 2\beta \begin{pmatrix} \lambda_{1\gamma} \\ \lambda_{2\gamma} \\ \lambda_{3\gamma} \\ \lambda_{4\gamma} \end{pmatrix} \quad (1.64c)$$

$$\int_\Omega [\lambda_{1\gamma} \sin(\lambda_{2\gamma}\ x) + \lambda_{3\gamma} \cos(\lambda_{4\gamma}\ xy)]\mu\ dx = 0 \quad (1.64d)$$

$$[\lambda_{1\gamma} \sin(\lambda_{2\gamma}\ x) + \lambda_{3\gamma} \cos(\lambda_{4\gamma}\ xy)] \geq 0 \qquad c.t.p\ \Omega \quad (1.64e)$$

CAPÍTULO 2

Implementación Numérica

De aquí en adelante estudiaremos la solución numérica del problema de optimización (1.20).

Para la optimización del valor de los parámetros óptimos se usó un método cuasi-Newton (La fórmula Davidon–Fletcher–Powell o DFP) junto con un algoritmo de búsqueda lineal que sigue la regla de Armijo.

La ecuación de estado se resolvió a través de un algoritmo Newton Semismooth (NSS), juntando una discretización en elementos finitos para las imágenes con un mallado homogéneo de $h > 0$ y una discretización en diferencias finitas para el gradiente. Se usó como inicialización del algoritmo NSS la imagen computada en la iteración previa. Por otro lado, el problema adjunto se resolvió como una ecuación lineal, discretizada con los mismos métodos que la ecuación de estado.

En resumen, se procedió en 3 partes:

1. Uso del algoritmo NSS para resolver la ecuación de estado: $(1.39\mathrm{a})$ - $(1.39\mathrm{b})$ en el caso de un parámetro real y $(1.42\mathrm{a})$ - $(1.42\mathrm{b})$ en el caso polinomial.

2. Resolución lineal del problema adjunto: $(1.39\mathrm{c})$ - $(1.39\mathrm{d})$ en el caso de un parámetro real y $(1.42\mathrm{c})$ - $(1.42\mathrm{d})$ en el caso polinomial.

3. Optimización de los parámetros óptimos, a través de un método cuasi Newton (DFP) y un algoritmo de búsqueda lineal: $(1.39\mathrm{e})$ - $(1.39\mathrm{f})$ en el caso de un parámetro real y (1.43) en el caso polinomial.

De manera alternativa se podría haber considerado un método de tipo Newton para la solución de los sistemas de optimalidad (1.39) y (1.42) y de esa manera realizar el cálculo en una sola parte. Sin embargo la solución se hubiera convertido en un proceso muy costoso (al tener que calcular las ecuaciones de estado y adjunta de una vez) sin especificar técnicas de precondicionamiento específicas para el problema. Este es un tema en investigación.

2.1 Implementación del método de elementos finitos

Como se dijo en la introducción de este capítulo: para la solución de la ecuación de estado $(1.39\mathrm{a})$-$(1.39\mathrm{b})$ y la ecuación adjunta $(1.39\mathrm{c})$ - $(1.39\mathrm{d})$ en el caso real, y las ecuación de

estado y adjunta (1.42) del caso polinomial, procederemos a utilizar una discretización en elementos finitos de dichas ecuaciones para lo que es necesario determinar con exactitud las matrices de masa y rigidez que son claves para lograr nuestro objetivo. Para encontrar dichas matrices, se procedió primero a estudiar un problema no lineal en elementos finitos más simple y de esta manera se llegó a encontrar las matrices de malla y rigidez que más adelante nos sirvieron para resolver las ecuaciones de estado y adjunta de nuestros problemas.

Sea $\Omega = [0,1] \times [0,1]$ un dominio poligonal de $\mathbb{R}^2$ de frontera Γ. Consideremos la función $u = u(x)$ para $x = (x,y) \in \Omega$ solución del problema no lineal modelo:

$$-\triangle u + u^3 = f \qquad \text{en} \quad \Omega \tag{2.1a}$$

$$u = 0 \qquad \text{en} \quad \Gamma \tag{2.1b}$$

cuya formulación variacional está dada por:

$$\int_\Omega \nabla v \nabla u \, dx + \int_\Omega u^3 v \, dx = \int_\Omega f v \, dx \quad \forall v \in H_0^1 \tag{2.2}$$

Si notamos $a(u,v) := \int_\Omega \nabla v \nabla u \, dx$ y $(\cdot,\cdot)$ al producto en L^2, podemos reescribir (2.2) de la forma $F(u) = 0$:

$$F(u) = a(u,v) + (u^3, v) - (f,v) = 0 \tag{2.3}$$

Luego, si aplicamos el método de Newton para encontrar las raíces de (2.3) llegamos al siguiente esquema:

$$F'(u)\delta_u = -F(u) \tag{2.4}$$

donde $\delta_u \in H_0^1(\Omega)$ es la dirección de la derivada.

Si F es como en (2.3),

$$F'(u)\delta_u = a(\delta_u, v) + (3u^2\delta_u, v) \quad \forall v \in H_0^1$$

con lo que, el problema a resolver (2.4) es:

$$a(\delta_u, v) + (3u^2\delta_u, v) = -a(u,v) + (u^3, v) - (f,v) \tag{2.5}$$

Sea $V_h \subset H_0^1(\Omega)$ un subespacio de dimensión finita. El método de elementos finitos (FEM) es una técnica para construir una base del subespacio V_h a través de interpolación polinomial. Consideremos la partición $\mathcal{I}_h$ del intervalo $[0,1]$ en n subintervalos $I_j^1 = [x_j, x_{j+1}]$ (y de manera análoga $I_j^2 = [y_j, y_{j+1}]$) de tamaño $h = x_{j+1} - x_j$, $j = 0, ..., n-1$ con:

$$0 = x_0 < x_1 < ... < x_{n-1} < x_n = 1$$

Consideremos la familia de polinomios a trozos:

$$X_h^k := \{v \in C([0,1]) : v|_{I_j} \in \mathbb{P}_k(I_j), \forall I_j \in \mathcal{I}_h\}$$

Sea adicionalmente

$$V_h = X_k^{k,0} = \{v \in X_h^k : v_h(x) = 0 \quad \text{en} \ \ \Gamma\}$$

cuya dimensión es $n^2 - 1$

En este caso particular, si en nuestro dominio $\Omega = [0,1] \times [0,1]$ utilizamos los pasos de discretización $h = k = \frac{1}{n+1}$ obtenemos la malla:

$$\bar{\Omega}_{hk} = \{(x_i, y_j) : x_i = ih, \ y_j = jh, \ i = 0, ..n+1, \ j = 0, ..n+1\}$$

(ver la figura 2.1)

Consideramos la aproximación de Galerkin [23] [24] de (2.5): encontrar $u_h(x)$ tal que:

$$a(\delta_{uh}, v_h) + (3u_h^2 \delta_{uh}, v_h) = -a(u_h, v_h) + (u_h^3, v_h) - (f_h, v_h) \quad \forall v_h \in V_h \tag{2.6}$$

donde $u_h = 0$ en Γ

Considerando una base $\{\varphi_j\}_{j=1,...,n^2}$ la solución se puede expresar como:

$$u_h(t) = \sum_{j=1}^{N} u_j(t)\varphi_j \tag{2.7}$$

donde u_j son los valores de la función u evaluada en el punto $x_j = (x_k, y_l)$

Para la construcción de una base $(\varphi_1, \varphi_2, ..., \varphi_{n^2})$ consideramos las funciones φ_i tales que:

$$\varphi_i(p_j) = \delta_{ij} \tag{2.8}$$

donde los $p_j = (x_k, y_l)$ son los nodos de la malla (ver la figura 2.1).

Así el soporte de (φ_i) consiste en todos los triangulos que incluyen al vértice p_i, y las funciones test tienen la forma de pirámides hexagonales, cuya cúspide tiene una altura de 1 en el punto del mallado que sea el centro de la función test (ver la figura 2.2).

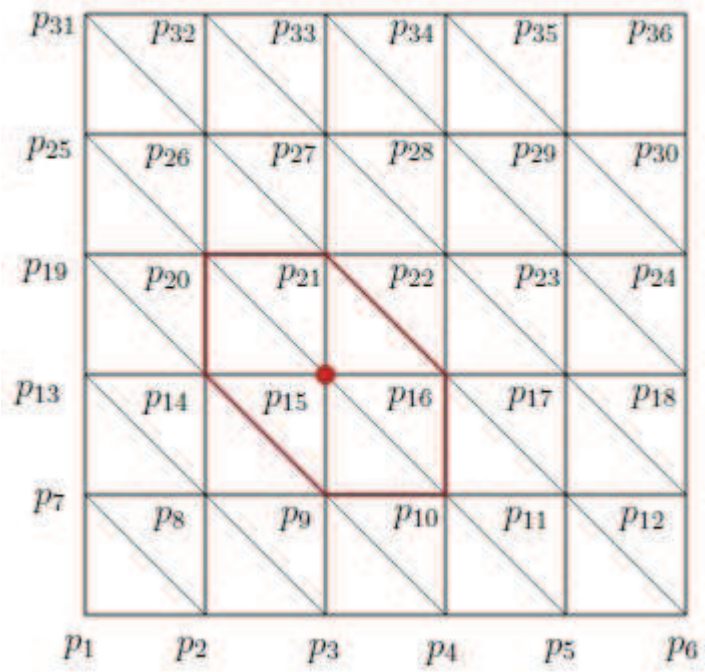

Figura 2.1: *Malla para una discretización del dominio Ω con 4 puntos de interpolación ($h = 1/5$) , con un total de 36 puntos. El mallado consiste en 50 triángulos que, unidos, conforman el dominio. Se puede ver (en rojo) la base de la función test φ_{15} que comprende todos los triangulos que contienen al punto $p_{15} = (x_3, y_3)$*

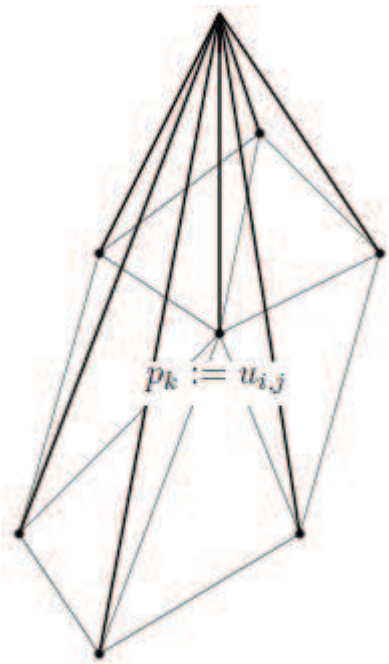

Figura 2.2: *Función test φ_k cuyo punto central es p_k*

Si miramos en detalle la base de una de las funciones φ_i (figura 2.3), podemos identificar 6 regiones, que notaremos tal y como se puede apreciar en la figura: A, B, C, D, E, F:

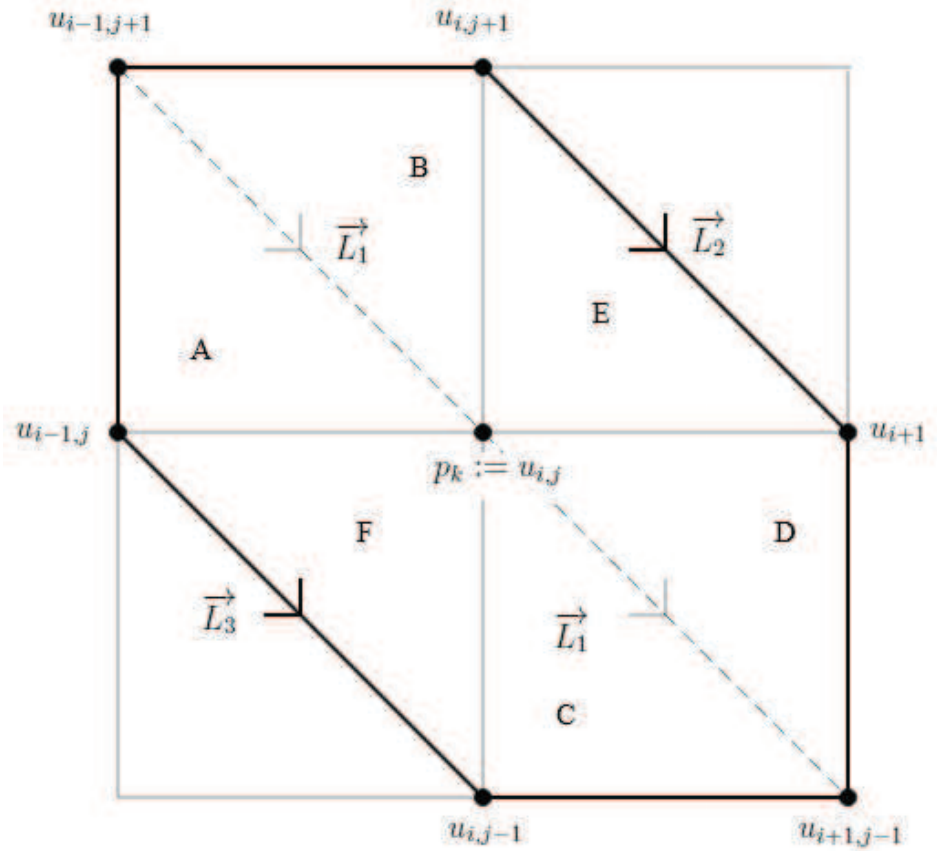

Figura 2.3: *base de la función test φ_k. Se notarán las regiones: A, B, C, D, E, F para poder mostrar de manera explícita la ecuación de la función test*

$$A \text{ es } \begin{cases} x_{i-1} \leq x \leq x_i \\ y_j \leq y \leq L_1 \end{cases}, \quad B \text{ es } \begin{cases} x_{i-1} \leq x \leq x_i \\ L_1 \leq y \leq y_{j+1} \end{cases}$$

$$C \text{ es } \begin{cases} x_i \leq x \leq x_{i+1} \\ y_{j-1} \leq y \leq L_1 \end{cases}, \quad D \text{ es } \begin{cases} x_i \leq x \leq x_{i+1} \\ L_1 \leq y \leq y_j \end{cases}$$

$$E \text{ es } \begin{cases} x_i \leq x \leq x_{i+1} \\ y_j \leq y \leq L_2 \end{cases}, \quad F \text{ es } \begin{cases} x_{i-1} \leq x \leq x_i \\ L_3 \leq y \leq y_j \end{cases}$$

donde:

$$L_1 : y = y_j - (x - x_i)$$
$$L_2 : y = y_j - (x - x_{i+1})$$
$$L_3 : y = y_j - (x - x_{i-1})$$

Usando, la ecuación general de un plano que pasa por tres puntos, concluimos que nuestras funciones test $\{\varphi_l\}$ cunplen con la siguiente ecuación:

Sea φ_l la función test cuyo centro en la base es el punto $u(x_i, y_j)$,

$$\varphi_l(x,y) = \begin{cases} \frac{x-x_{i-1}}{h} & \text{en} \quad A \\ \frac{y_{j+1}-y}{h} & \text{en} \quad B \\ \frac{y-y_{j-1}}{h} & \text{en} \quad C \\ \frac{x_{i+1}-x}{h} & \text{en} \quad D \\ -\frac{x}{h} - \frac{y}{h} - \frac{y_j x_i - x_{i+1} y_{j+1}}{h^2} & \text{en} \quad E \\ -\frac{y_j x_i - x_{i-1} y_{j-1}}{h^2} + \frac{x}{h} + \frac{y}{h} & \text{en} \quad F \\ 0 & \text{caso contrario} \end{cases} \tag{2.9}$$

y sus derivadas direccionales:

$$D_x\varphi_l(x,y) = \begin{cases} 1/h & \text{en} \quad A \\ 0 & \text{en} \quad B \\ 0 & \text{en} \quad C \\ -1/h & \text{en} \quad D \\ -1/h & \text{en} \quad E \\ 1/h & \text{en} \quad F \\ 0 & \text{caso contrario} \end{cases} , \quad D_y\varphi_l(x,y) = \begin{cases} 0 & \text{en} \quad A \\ -1/h & \text{en} \quad B \\ 1/h & \text{en} \quad C \\ 0 & \text{en} \quad D \\ -1/h & \text{en} \quad E \\ 1/h & \text{en} \quad F \\ 0 & \text{caso contrario} \end{cases}$$

Ahora bien. De (2.6) y (2.7) tenemos que el esquema (2.5) queda:

$$a(\delta_u h, v_h) + (3u_h^2 \delta_u h, v_h) = -a(u_h, v_h) + (u_h^3, v_h) - (f_h, v_h) \ \ \forall v_h \in V_h \tag{2.10}$$

Toda función $g \in V_h$ se puede representar por:

$$g(x) = \sum_{i=1}^{N} \alpha_i \varphi_i$$

donde $\alpha_i = g(p_i)$.

Entonces, podemos aproximar: $\delta_u \approx \delta_{uh} = \sum_{i=1}^{(N} \delta_{uh_i}\varphi_i$ en (2.10). De aquí en adelante, para evitar confusión en la nomenclatura, usaremos la notación u en lugar de u_h y δ_u en lugar de δ_{uh}, entendiendo que estas funciones no son iguales entre sí, pero sí se aproximan una con otra. Si además hacemos: $v = \varphi_j$, nos queda:

$$\begin{aligned} a\left(\sum_{i=1}^{N} \delta_{u_i}\varphi_i, \varphi_j\right) + \left(\sum_{i=1}^{N}\left[\sum_{k=1}^{N} u_k^2\varphi_k\right]\delta_{u_i}\varphi_i, \varphi_j\right)_{L^2(\Omega)} = \\ - a\left(\sum_{i=1}^{N} u_i\varphi_i, \varphi_j\right) + \left(\sum_{i=1}^{N} u_i^3\varphi_i, \varphi_j\right)_{L^2(\Omega)} - \left(\sum_{i=1}^{N} f_i\varphi_i, \varphi_j\right)_{L^2(\Omega)} \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^{N} \delta_{u_i}(D\varphi_i, D\varphi_j)_{L^2(\Omega)} + 3\sum_{i=1}^{N} \delta_{u_i}\left[\sum_{k=1}^{N} u_k^2\varphi_k\right](\varphi_i, \varphi_j)_{L^2(\Omega)} = \\ - \sum_{i=1}^{N} u_i(D\varphi_i, D\varphi_j)_{L^2(\Omega)} + \sum_{i=1}^{N} u_i^3(\varphi_i, \varphi_j)_{L^2(\Omega)} - \sum_{i=1}^{N} f_i(\varphi_i, \varphi_j)_{L^2(\Omega)} \end{aligned}$$

Si notamos como $\overrightarrow{\delta_u}$ al vector de los (δ_{u_i})
$\overrightarrow{u}$ al vector de los (u_i)
$\overrightarrow{f}$ al vector de los (f_i)
s a la cantidad $s = \sum_{k=1}^{(n+1)^2} u_k^2 \varphi_k$
$\overrightarrow{u^3}$ al vector de los (u_i^3)

M_m a la matriz que contiene a los productos $(\varphi_i, \varphi_j)_{L^2(\Omega)}$, a la que nos referiremos de aquí en adelante como *Matriz de Masa* (figura 2.4)

M_R a la matriz que contiene a los productos $(D\varphi_i, D\varphi_j)_{L^2(\Omega)}$, a la que nos referiremos de aquí en adelante como *Matriz de Rigidez* (figura 2.5)

Ambas matrices fueron construidas a partir de (2.9), tomando los productos en L^2 de las funciones φ_i y almacenándolos en matrices de dimensión $M \times N$, donde N es el número de puntos del mallado. En las figuras 2.4 y figura 2.5 se puede ver el patrón de dispersión de las matrices de masa y rigidez, respectivamente, y en sus descripciones un detalle de sus valores.

Entonces podemos reescribir lo anterior como:

$$M_R \overrightarrow{\delta_u} + 3sM_m \overrightarrow{\delta_u} = -M_R \overrightarrow{u} + M_m \overrightarrow{u^3} - M_m \overrightarrow{f}$$

de donde obtenemos que el paso del esquema de Newton (2.4) es:

$$\overrightarrow{\delta_u} = [M_R \overrightarrow{u} + M_m (\overrightarrow{u^3} - \overrightarrow{f})] * [M_R + 3\ s\ M_m]^{-1} \tag{2.11}$$

Observación 2.1. *Notar que si n es el número de puntos de interpolación en ambas componentes, se obtiene un mallado con $N = (n+2)^2$ puntos p_k por lo que las Matrices de Malla y Rigidez tendrán una dimensión de $N \times N$*

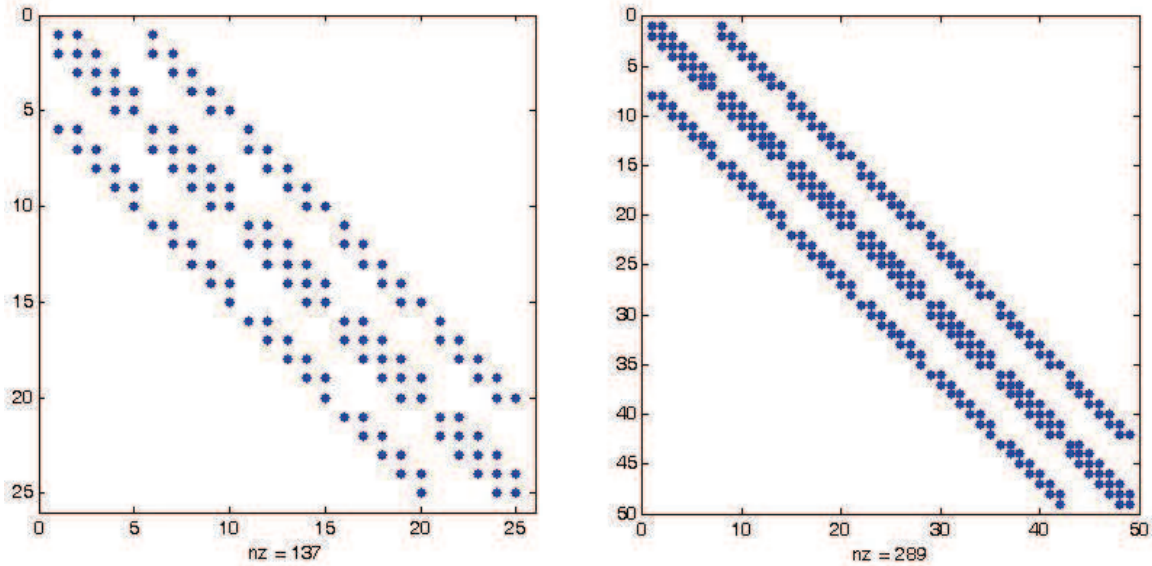

Figura 2.4: *Patrón de dispersión de la matriz de masa para un mallado de* $h = 1/4$. *A la izquierda está representada la matriz de masa con triangulación regular, es decir cuando no se toman los nodos del filo del mallado; los valores distintos de cero son todos iguales a* $h^2/12$ *y en la diagonal principal* $h^2/2$. *A la derecha, se puede apreciar la misma matriz de masa pero, como se sugiere en [24], se incluyen los nodos en el filo del mallado, dando como resultado una matriz similar a la de la izquierda, pero con fracciones de los valores en los productos que involucran los nodos periféricos*

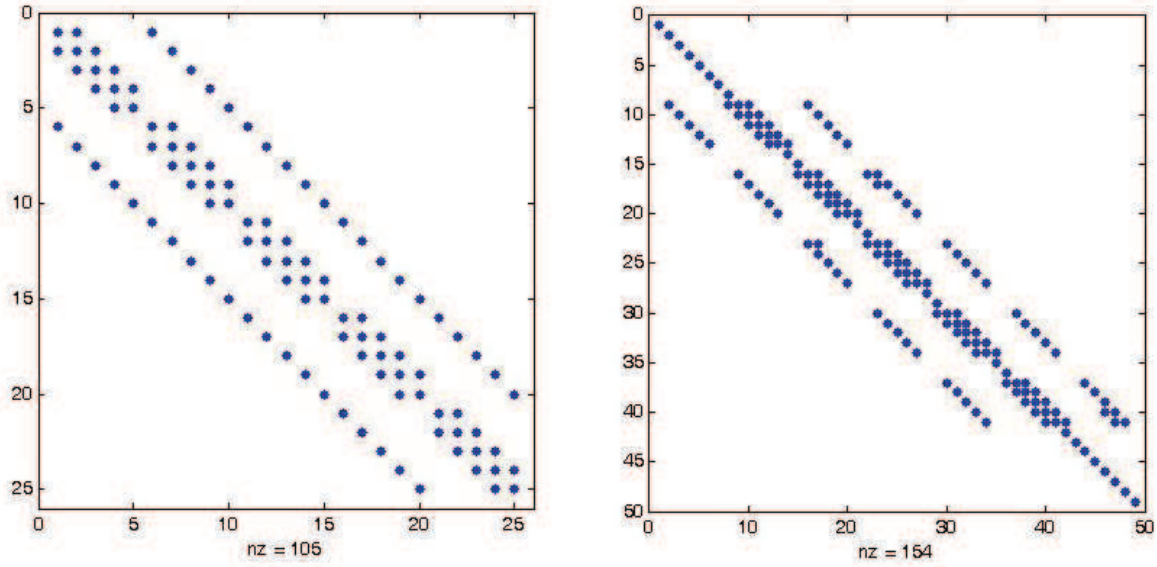

Figura 2.5: *Patrón de dispersión de la matriz de rigidez para un mallado de* $h = 1/4$, *en la que los valores distintos de cero son todos iguales a* -1 *y en la diagonal principal* 4 *(esta matriz no depende del valor de* h*). Como en el gráfico 2.4: A la izquierda está representada la matriz de rigidez con triangulación regular y a la derecha se puede apreciar la misma matriz pero incluidos los nodos en el filo del mallado*

Con el resultado (2.11), somos capaces de proponer el algoritmo 1 para la resolución del problema no lineal modelo (2.1), a través del esquema de Newton planteado en (2.4), usando la aproximación por elementos finitos propuesta en [24].

Algoritmo 1 Algoritmo de resolución del problema modelo (2.1)

1: Inicializar con $u_0 = 0$ y fijar $k = 0$
2: **mientras** $|\delta_u| > tol$ **hacer**
3: Resolver el esquema (2.4): Calcular $\overrightarrow{\delta_u}$ con (2.11)
4: Actualizar $u_{k+1} = u_k + \overrightarrow{\delta_u}$ y fijar $k = k + 1$
5: **fin mientras**

Si usamos este algoritmo para calcular la solución del problema modelo (2.1) con $f = x^3y^3(x-1)^3(y-1)^3 - 2y(y-1) - 2x(x-1)$ obtenemos la función que se puede apreciar en la figura 2.6 que se aproxima a la solución exacta del problema $u = xy(x-1)(y-1)$ con un error de $0{,}02$.

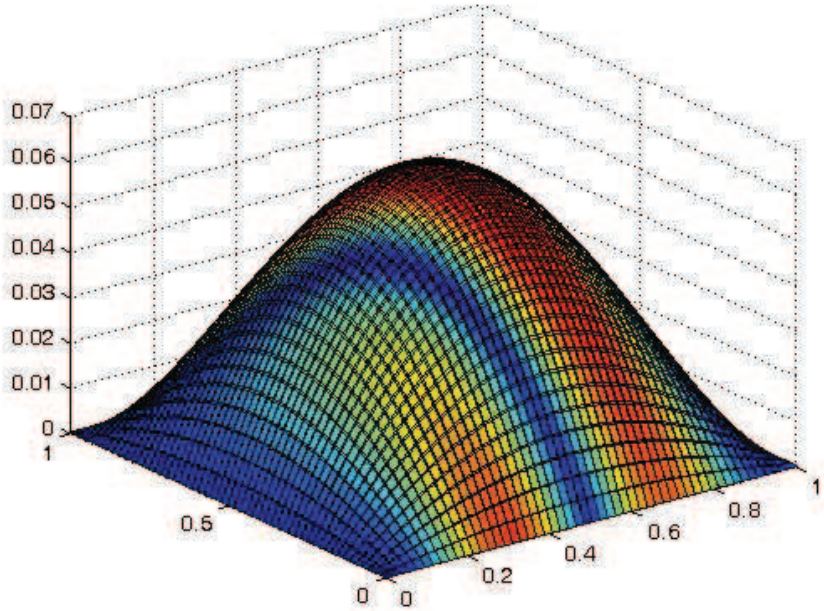

Figura 2.6: *Solución del problema modelo* (2.1) *con* $f = x^3y^3(x-1)^3(y-1)^3 - 2y(y-1) - 2x(x-1)$ *y* $h = 1/50$, *obtenida al implementar el algoritmo 1 graficada junto con la solución exacta del problema* $u = xy(x-1)(y-1)$

En la siguiente sección implementaremos un algoritmo NSS para encontrar la solución de la ecuación de estado (1.39a)- (1.39b) que es de tipo no lineal. Para ello utilizamos la discretización en elementos finitos analizada en esta seccón.

2.2 Algoritmo NSS para la ecuación de estado

Sea Y un espacio vectorial de Hilbert tal que $u \in Y$, si definimos el operador $F : Y \times Y \to Y' \times Y'$:

$$F(u,q) = \begin{pmatrix} \varepsilon(Du,Dv) + (q,Dv) + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i'(u,f)v\, dx \\ q - h_\gamma(Du) \end{pmatrix} \tag{2.12}$$

podemos escribir la ecuación de estado (1.39a) - (1.39b) como:

$$F(u,q) = 0 \tag{2.13}$$

Aplicaremos un método de tipo Newton para la resolución de (2.13):

$$F'(u,q)(\delta_u,\delta_q) = -F(u,q) \tag{2.14}$$

donde, δ_u y δ_q son las direcciones en las cuales se deriva el Jacobiano F' del operador F:

$$F'(u,q)(\delta_u,\delta_q) = \begin{pmatrix} \varepsilon(D\delta_u,Dv) + (\delta_q,Dv) + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i''(u)[\delta_u]v \\ \delta_q - h_\gamma'(Du)D\delta_u \end{pmatrix} \tag{2.15}$$

con $h_\gamma(Du)$ como en (1.28) y $h_\gamma'(Du)$ como en (1.54)

Siguiendo con el esquema de Newton (2.14), igualamos las primeras componentes de (2.12) y (2.15):

$$\varepsilon(D\delta_u,Dv) + (\delta_q,Dv) + \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i''(u)[\delta_u]v = -\varepsilon(Du,Dv) - (q,Dv) - \sum_{i=1}^{d} \int_\Omega \lambda_i \phi_i'(u,f)v\, dx \tag{2.16}$$

Ahora introduciremos un problema de ejemplo, en el que la función ϕ esté dada de manera explícita, para obtener una aproximación de (2.16) con elementos finitos:

Sea $f \in L^2$ una imagen con ruido dada. Consideraremos ahora el problema de eliminación de un único ruido de tipo gaussiano en una imagen dada con un parámetro real. Así, con $\phi = ||u-f||_{L^2}^2$ y $g(u) = \frac{1}{2}||u-u_0||_{L^2}^2$, donde $u_0 \in L^2$ es la imagen original o una aproximación a ella. A partir de (1.20) consideraremos el siguiente problema de control óptimo:

$$\min_\lambda \frac{1}{2}||u-u_o||_{L^2}^2 + \beta\lambda^2 \tag{2.17a}$$

sujeto a la restricción:

$$\varepsilon(Du, D(v-u))_{L^2} + \int_\Omega |Dv|dx - \int_\Omega |Du|dx + \lambda \int_\Omega (u-f)(v-u)dx \geq 0 \quad \forall v \in H_0^1 \tag{2.17b}$$

Ahora podemos reescribir (2.16) exclusivamente para este problema:

$$\varepsilon(D\delta_u,Dv) + (\delta_q,Dv) + \lambda(\delta_u,v) = -\varepsilon(Du,Dv) - (q,Dv) - \lambda(u-f,v) \tag{2.18}$$

Usaremos la aproximación en elementos finitos analizada en la sección anterior para resolver (2.18) numéricamente. Para ello tomamos: $v = \varphi_j$, $\delta u = \sum_{i=1}^{N} \delta u_i \varphi_i$, $\delta q = \sum_{k=1}^{M} \delta q_k \phi_k$, $q = \sum_{k=1}^{M} q_k \phi_k$, $u = \sum_{i=1}^{N} u_i \phi_i$

Y nos queda:

$$\varepsilon \left(\sum_{i=1}^{N} \delta u_i D\varphi_i, D\varphi_j \right) + \left(\sum_{k=1}^{M} \delta q_k \phi_k, D\varphi_j \right) + \lambda \left(\sum_{i=1}^{N} \delta u_i \varphi_i, \varphi_j \right)$$
$$= -\varepsilon \left(\sum_{i=1}^{N} u_i D\varphi_i, D\varphi_j \right) - \left(\sum_{k=1}^{M} q_k \phi_k, D\varphi_j \right) - \lambda \left(\sum_{i=1}^{N} (u_i - f_i)\varphi_i, \varphi_j \right)$$

Hay que notar que hemos utilizado dos discretizaciones distintas: la primera con la base conformada con las N funciones test φ para las funciones u , y la segunda con la base conformada por las funciones test ϕ para las funciones q que tienen una dimensión M ya que están definidos para cada triángulo del mallado (ver la figura 2.1).

Usando las propiedades del producto escalar $(.,.)$ tenemos:

$$\varepsilon \sum_{i=1}^{N} \delta u_i \left(D\varphi_i, D\varphi_j \right) + \sum_{k=1}^{M} \delta q_k \left(\phi_k, D\varphi_j \right) + \lambda \sum_{i=1}^{N} \delta u_i \left(\varphi_i, \varphi_j \right) =$$
$$- \varepsilon \sum_{i=1}^{N} u_i \left(D\varphi_i, D\varphi_j \right) - \sum_{k=1}^{M} q_k \left(\phi_k, D\varphi_j \right) - \lambda \sum_{i=1}^{N} u_i \left(\varphi_i, \varphi_j \right) + \lambda \sum_{i=1}^{N} f_i \left(\varphi_i, \varphi_j \right)$$

Si notamos como $\delta u = (\delta u_1, \delta u_2, ..., \delta u_N)$

La Matriz de Rigidez: $M_R = ((D\varphi_i, D\varphi_j))_{i,j}$ (figura 2.5)

$\delta q = (\delta q_1, \delta q_2, ..., \delta q_M)$

$A_2 = ((D_X \varphi_i, \phi_j) \quad (D_Y \varphi_i, \phi_j)) \in \mathbb{R}^{N,2M}$

La Matriz de Masa: $M_m = ((\varphi_i, \varphi_j))_{i,j}$ (figura 2.4)

$u = (u_1, u_2, ..., u_N)$

$f = (f_1, f_2, ..., f_N)$

$q = (q_1, q_2, ..., q_M)$,

obtenemos que (2.18) es equivalente a resolver:

$$\varepsilon M_R \delta u + A_2 \delta q + \lambda M_m \delta u = -\varepsilon M_R u - A_2 q - \lambda M_m u + \lambda M_m f$$

y así, tenemos:

$$(\varepsilon M_R + \lambda M_m)\delta u + A_2 \delta q = (-\varepsilon M_R - \lambda M_m)u - A_2 q + \lambda M_m f \tag{2.19}$$

Observación 2.2.

- *Un mallado con n puntos de interpolación está conformado por $(n+2)^2$ puntos p_k (ver la figura 2.1), por lo que la solución u_h del esquema numérico tendrá $N = (n+2)^2$ valores.*

- *Las Matrices de Masa y Rigidez tienen las propiedades vistas en la observación 2.1*

- *$M := 2t$ donde t es el número de triángulos en el mallado.*
- *Las funciones ϕ_k tienen el valor de 1 dentro del triángulo k-ésimo y 0 fuera de este, mientras que las derivadas direccionales $D_x\varphi$ y $D_y\varphi$ están dadas por (2.9), por lo que la matriz A_2 tiene la dispersión que se indica en la figura 2.9*

Ahora, para continuar con el esquema de Newton igualamos las segundas componentes de (2.12) y (2.15) , es decir:

$$\delta q - h'_\gamma(Du)D\delta u = -q + h_\gamma(Du) \tag{2.20}$$

Para la derivada $h'_\gamma(Du)D\delta u$ del lado izquierdo recurrimos a (1.54) (que fue obtenida a partir de la expresión (1.28) de la regularización de conjuntos activo e inactivo del subdiferencial) mientras que en el lado derecho de la igualdad usamos la expresión (1.27) de la regularización que es equivalente a (1.28).

Así (2.20) nos queda:

$$\delta_q - h'_\gamma(Du)D\delta_u = -q + \frac{\gamma Du}{\text{máx}(1, \gamma|Du|)}$$

o lo que es lo mismo:

$$\text{máx}(1, \gamma|Du|)\delta_q - \text{máx}(1, \gamma|Du|)\ \ h'_\gamma(Du)D\delta_u = -q\ \ \text{máx}(1, \gamma|Du|) + \gamma Du \tag{2.21}$$

Para resolver la ecuación (2.21) fue necesaria la discretización del grandiente D a través de un esquema en diferencias finitas. En el siguiente apartado se introduce la matriz G, que aproxima al gradiente D, hallada a partir de dicha discretización que fue vital para continuar con la resolución de la ecuación.

2.2.1 Discretización en diferencias finitas del Gradiente

En ambas dimensiones, se utilizó una discretización en diferencias finitas progresivas con un mallado de $h = \frac{1}{n+1}$ tomando en cuenta la aproximación de la derivada en cada triángulo (2.9).

La matriz D del gradiente en dierencias finitas está conformada en realidad por dos matrices: la discretización de la derivada en x, D_x y la discretización de la derivada en y, D_y (figura 2.9)

Así, en el punto (x_i, y_j), usando la notación que se puede ver en la figura 2.8, las derivadas se aproximan por:

$$D_x u_{i,j} \approx \frac{u_{i+1,j} - u_{i,j}}{h} = \frac{p_i - p_{i-(n+2)}}{h} \tag{2.22a}$$

$$D_y u_{i,j} \approx \frac{u_{i,j+1} - u_{i,j}}{h} = \frac{p_{i+1} - p_i}{h} \tag{2.22b}$$

En las figura 2.7 se puede ver, de manera gráfica, la forma en la que se construyeron las ecuaciones (2.22)

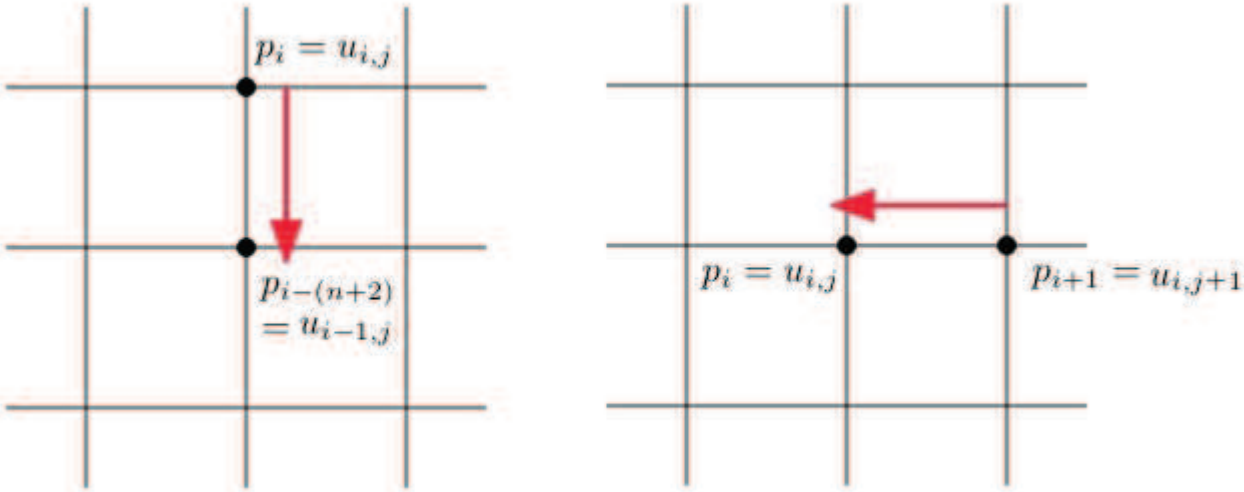

Figura 2.7: *Aproximación en diferencias finitas progresivas de la derivada en la dirección x (izquierda) y en la dirección y (derecha)*

De esta forma, se construye la matriz aproximación del gradiente D a la que notaremos $G = [G_x; G_y]$ donde G_x y G_y son las aproximaciones de D_x y D_y respectivamente. En la práctica, las matrices G_x y G_y tienen $(n+2)^2$ columnas (número de puntos p_i del mallado) y $2 \times (n+1)^2$ filas (número de triángulos), ver la figura 2.8. Además se puede observar la disperción de estas matrices en la figura 2.9

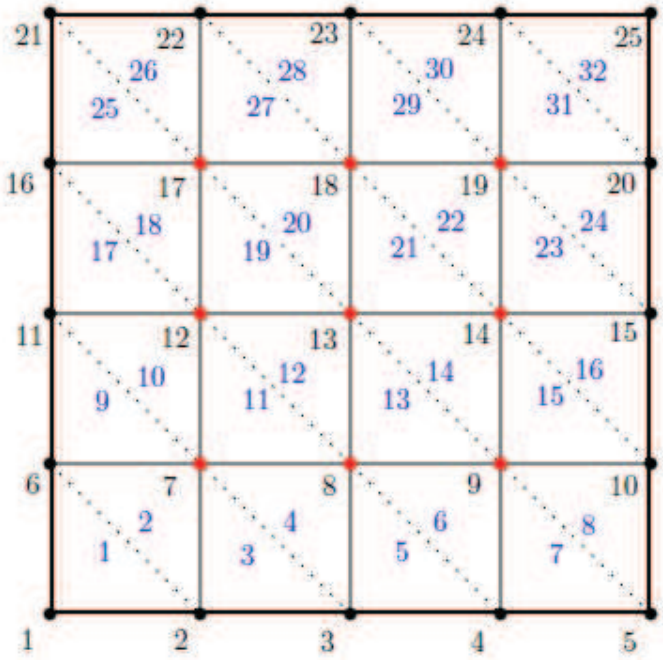

Figura 2.8: *Notación usada para numerar los puntos del mallado (en color negro) y los triángulos en los que se ha dividido el dominio (en letras azules)*

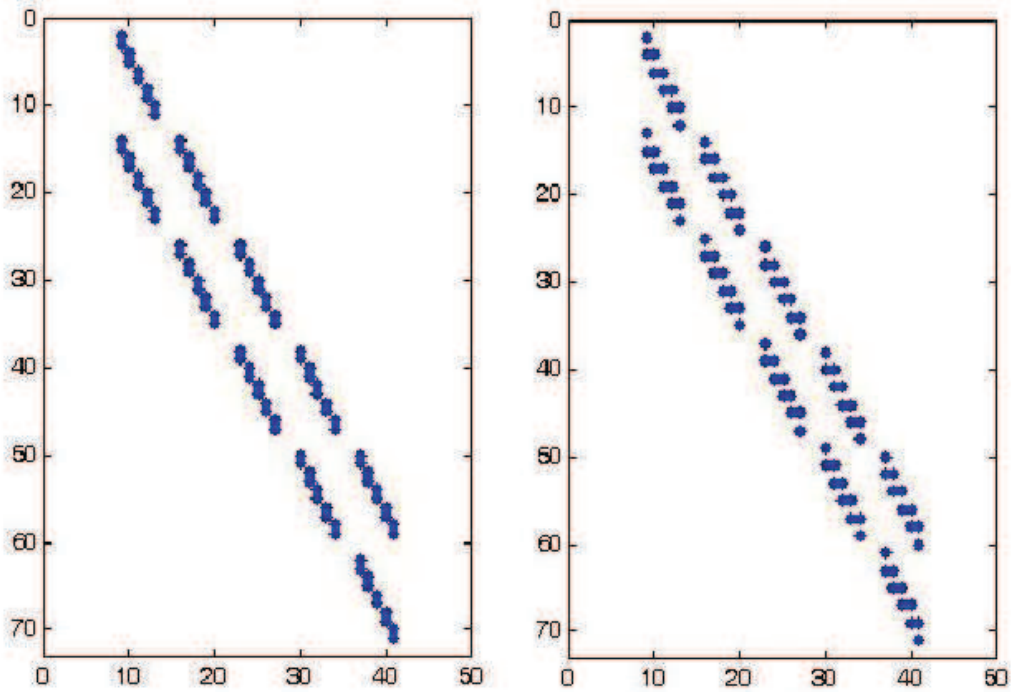

Figura 2.9: *Gráfico de disperción de las matrices G_x (a la izquierda) y G_y (a la derecha) de las aproximaciones por diferencias finitas de las derivadas parciales D_x y D_y respectivamente.*

2.2.2 Aproximación por diferencias finitas en los conjuntos activo, activo débil e inactivo

En el conjunto activo A^γ $\text{máx}(1, \gamma|Du|) = \gamma|Du|$, por lo que (2.21) se convierte en:

$$\gamma|Du|\delta_q - \gamma|Du| \;\; h'_\gamma|_{A^\gamma} = -q\gamma|Du| + \gamma Du$$

reemplazando (1.54):

$$\gamma|Du|\delta_q - \gamma|Du| \left(\frac{D\delta_u}{|Du|} - \frac{Du\langle Du, D\delta u\rangle}{|Du|^3} \right) = -q\gamma|Du| + \gamma Du$$

Como $Du\langle Du, D\delta u\rangle = Du^T D\delta u Du$ la fórmula anterior es equivalente a:

$$\gamma|Du|\delta_q - \gamma|Du| \left(\frac{D\delta_u}{|Du|} - \frac{Du^T D\delta u Du}{|Du|^3} \right) = -q\gamma|Du| + \gamma Du \tag{2.23}$$

Podemos observar, por como se define q en (1.39b), que $q = \frac{Du}{|Du|}$ en c.t.p. de A^γ. Por lo que podemos reemplazar el término $\frac{Du^T D\delta u Du}{|Du|^3}$ por $\frac{Du^T D\delta u}{|Du|^2} q$ y, proyectando en el conjunto factible, tenemos que:

$$\frac{Du^T D\delta u Du}{|Du|^3} = \frac{Du^T}{|Du|^2} \frac{q}{\text{máx}(1, |q|)} D\delta u$$

con lo que (2.23) nos queda:

$$\gamma|Du|\delta_q - \gamma\left(1 - \frac{Du^T}{|Du|}\frac{q}{\text{máx}(1,|q|)}\right)D\delta u = -q\gamma|Du| + \gamma Du$$

y usando la discretización del gradiente G:

$$\gamma|G\cdot u|\delta_q - \gamma\left(1 - \frac{(G\cdot u)^T}{|G\cdot u|}\frac{q}{\text{máx}(1,|q|)}\right)G\cdot\delta u = -q\gamma|G\cdot u| + \gamma G\cdot u \tag{2.24}$$

En el conjunto activo débil S^γ si notamos $M := \text{máx}(1,|Du|)$ podemos expresar (2.21) como:

$$M\delta_q - M\left[\frac{D\delta u}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) + \frac{Du\langle Du, D\delta u\rangle}{|Du|^2}\left(-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right)\right] = -qM + \gamma Du \tag{2.25}$$

donde $B = \left(1 - \gamma|Du| + \frac{1}{2\gamma}\right)$

Como $Du\langle Du, D\delta u\rangle = Du^T D\delta u Du$ podemos reescribir (2.25):

$$M\delta_q - M\left[\frac{D\delta u}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) + \frac{Du^T D\delta u Du}{|Du|^2}\left(-\frac{1}{|Du|} + \frac{1}{|Du|}\frac{\gamma}{2}B^2 + \gamma^2 B\right)\right] = -qM + \gamma Du$$

$$M\delta_q - M\left\{\frac{D\delta u}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) + \frac{Du^T D\delta u Du}{|Du|^2}\left[-\frac{1}{|Du|}\left(1-\frac{\gamma}{2}B^2\right)B^2 + \gamma^2 B\right]\right\} = -qM + \gamma Du$$

$$M\delta_q - M\left[\frac{D\delta u}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) - \frac{Du^T D\delta u Du}{|Du|^2}\left(1-\frac{\gamma}{2}B^2\right) + \frac{Du^T D\delta u Du}{|Du|^2}\gamma^2 B\right] = -qM + \gamma Du \tag{2.26}$$

De manera similar a A^γ podemos realizar el reemplazo:

$$\frac{Du^T D\delta u Du}{|Du|^3}\left(1-\frac{\gamma}{2}B^2\right) = \frac{Du^T D\delta_u}{|Du|^2}\cdot\frac{q}{\text{máx}(1,|q|)}$$

ya que en S^γ:

$$\begin{aligned}\left(1-\frac{\gamma}{2}B^2\right)\frac{Du}{|Du|} &= \left[1-\frac{\gamma}{2}\left(1-\gamma|Du| + \frac{1}{2\gamma}\right)^2\right]\frac{Du}{|Du|}\\ &= \frac{q}{\text{máx}(1,|q|)}\end{aligned}$$

y (2.26) puede reescribirse como:

$$M\delta_q - M\left[\frac{D\delta u}{|Du|}\left(1-\frac{\gamma}{2}B^2\right) - \frac{Du^T D\delta u}{|Du|^2}\frac{q}{\text{máx}(1,|q|)} + \frac{Du^T D\delta u Du}{|Du|^2}\gamma^2 B\right] = -qM + \gamma Du$$

Usando la discretización del gradiente G:

$$M\delta_q - M\left[\frac{1}{|G\cdot u|}\left(1-\frac{\gamma}{2}B^2\right) - \frac{(G\cdot u)^T}{|G\cdot u|^2}\frac{q}{\text{máx}(1,|q|)} + \frac{(G\cdot u)^T G\cdot u}{|G\cdot u|^2}\gamma^2 B\right] G\cdot\delta u = -qM + \gamma G\cdot u \tag{2.27}$$

En el conjunto Inactivo I^γ $\text{máx}(1,|Du|) = 1$ por lo que (2.21) se convierte, en este conjunto, en:

$$\delta_q - \gamma D\delta_u = -q + \gamma Du$$

y usando la discretización en diferencias finitas del gradiente:

$$\delta_q - \gamma G\cdot\delta_u = -q + \gamma G\cdot u \tag{2.28}$$

2.2.3 Esquema de Newton

Juntando los resultados (2.19), (2.24), (2.27) y (2.28) definimos las matrices:

$$\Psi := \begin{pmatrix} \varepsilon M_R + \lambda M_m & A_2 \\ -R_1\cdot G & M \end{pmatrix} \tag{2.29}$$

donde:

$$R_1 = \begin{cases} \gamma\left(1 - \frac{(G\cdot u)^T}{|G\cdot u|}\frac{q}{\text{máx}(1,|q|)}\right) & \text{en} \quad A^\gamma \\ M\left[\frac{1}{|G\cdot u|}\left(1-\frac{\gamma}{2}B^2\right) - \frac{(G\cdot u)^T}{|G\cdot u|}\left(\frac{1}{|G\cdot u|}\frac{q}{\text{máx}(1,|q|)} + \gamma^2 BG\cdot u\right)\right] & \text{en} \quad S^\gamma \\ \gamma & \text{en} \quad I^\gamma \end{cases} \tag{2.30}$$

y:

$$\Lambda := \begin{pmatrix} -\varepsilon M_r - \lambda M_m & -A_2 & \lambda M_m \\ \gamma G & -M & 0 \end{pmatrix} \tag{2.31}$$

y la ecuación de incrementos que resolvemos está dada por:

$$\Psi\begin{pmatrix}\delta_u \\ \delta_q\end{pmatrix} = \Lambda\begin{pmatrix} u \\ q \\ f\end{pmatrix} \tag{2.32}$$

Un esquema para la solución de cada sistema regularizado de la ecuación de estado está dado en el Algoritmo 2

Ejemplo: Para ilustrar las propiedades principales del algoritmo Newton SemiSmooth propuesto consideramos el problema de control óptimo del fluido de Bingham [14] [25].

Para la aproximación numérica se realizó con el esquema en diferencias finitas visto en la sección anterior. La solución para los parámetros $f \equiv 10, \gamma = 100, \beta = 1^{-12}, h = 1/40$ se obtiene después de 18 iteraciones. En la figura 2.10 se puede observar el control óptimo

Algoritmo 2 Algoritmo de resolución del esquema NSS

1: Inicializar $(u_0, p_0) \in Y \times Y$ y fijar $k = 0$
2: definir: $\mathcal{A}^\gamma := \left\{x \in \Omega \ : \ \gamma|z| - 1 \geq \frac{1}{2\gamma}\right\}$, $\mathcal{S}^\gamma := \left\{x \in \Omega \ : \ |\gamma|z| - 1| \leq \frac{1}{2\gamma}\right\}$ y $\mathcal{I}^\gamma = \mathcal{S} \setminus (\mathcal{A}^\gamma \cup \mathcal{S}^\gamma)$
3: **mientras** $|\delta_u| > tol$ **hacer**
4: Resolver a ecuación de incremento:

$$F'(u,q)(\delta_u, \delta_q) = -F(u,q)$$

dada por (2.32)
5: Actualizar $u_{k+1} = u_k + \overrightarrow{\delta_u}$, $p_{k+1} = p_k + \overrightarrow{\delta_p}$ y fijar $k = k + 1$
6: **fin mientras**

resultante, que tiene un valor constante en el conjunto inactivo.

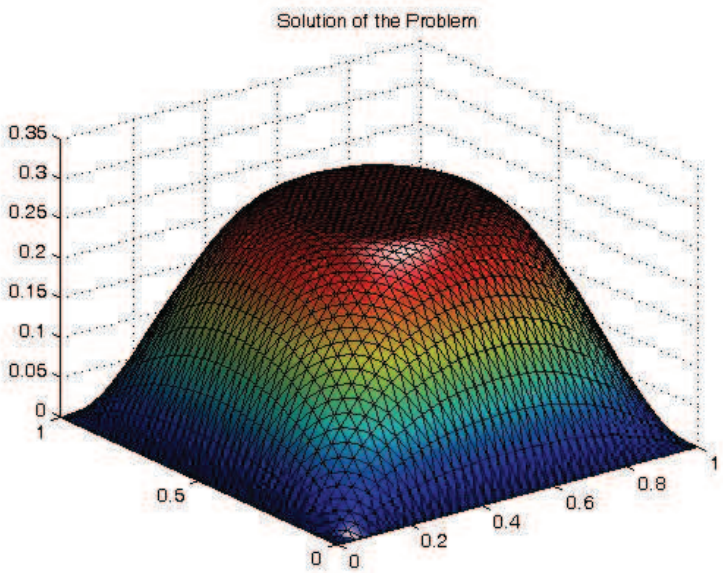

Figura 2.10: *Control óptimo de un fluido de Bingham con* $f \equiv 10, \gamma = 100, \beta = 1^{-12}, h = 1/40$

2.2.4 Solución de la ecuación de Estado para el Problema con Parámetro Polinomial de coeficientes positivos(1.41)

En el capítulo anterior, planteamos el problema de encontrar un control óptimo polinomial para la disminución de ruido en imágenes. De manera específica, en (1.41) se planteó dicho problema de minimización y más adelante, a partir del Teorema 1.8, se llegó al sistema de optimalidad (1.42)- (1.43) del que buscaremos una solución usando técnicas similares a las vistas anteriormente para la resolución de (2.17)

Ahora introduciremos un problema de ejemplo en el que deseamos hallar el parámetro λ^* polinomial óptimo en el que la función ϕ está dada de manera explícita:

Sea $f \in L^2$ una imagen con ruido dada. Consideraremos ahora el problema de eliminación de un único ruido de tipo gaussiano en una imagen dada con un parámetro polinomial de la forma: $\lambda(x,y) = Ax^2 + By^2 + Cxy + Dx + Ey + F$, donde A, B, C, D, E, F son números reales no negativos. Así, con $\phi = ||u - f||^2_{L^2}$ y $g(u) = \frac{1}{2}||u - u_0||^2_{L^2}$, donde $u_0 \in L^2$ es la imagen original o una aproximación a ella, a partir de (1.20) consideraremos el siguiente problema de control óptimo:

$$\min_{\lambda = Ax^2+By^2+Cxy+Dx+Ey+F} \frac{1}{2}||u - u_o||^2_{L^2} + \beta|\lambda|^2 \tag{2.33a}$$

sujeto a las restricciones:

$$\varepsilon(Du, Dv) + (h_\gamma(Du), Dv) + \int_\Omega \lambda(x,y)\phi'(u,f)\, v\, dx = 0 \quad \forall v \in H^1_0 \tag{2.33b}$$

$$A \geq 0,\ B \geq 0,\ C \geq 0,\ D \geq 0,\ E \geq 0,\ F \geq 0 \tag{2.33c}$$

Al igual que en el caso real aplicamos el método Newton Semismooth (2.14) para resolver la ecuación de estado (1.42a)- (1.42b).

En este caso no podemos aislar el parámetro $\lambda(x,y)$ de las integrales, como se hizo para obtener (2.18) ya que λ depenede de las variables de integración (x,y). Tomando esto en cuenta e igualando las primeras componentes del sistema (2.14) tenemos:

$$\varepsilon(D\delta_u, Dv) + (\delta_q, Dv) + (\lambda\delta_u, v) = -\varepsilon(Du, Dv) - (q, Dv) - (\lambda[u - f], v) \tag{2.34}$$

Al igual que en el caso real, usamos la aproximación en elementos finitos analizada en la sección 2.1 para resolver (2.34) numéricamente. Para ello tomamos: $v = \varphi_j$, $\delta u = \sum_{i=1}^N \delta u_i \varphi_i$, $\delta q = \sum_{k=1}^M \delta q_k \phi_k$, $q = \sum_{k=1}^M q_k \phi_k$, $u = \sum_{i=1}^N u_i \phi_i$. Si llamamos λ_i al valor de la función $\lambda(x,y)$ evaluada en el punto $p_i = (x_k, y_l)$ tenemos la aproximación de (2.34) por elementos finitos:

$$\begin{aligned}\varepsilon \sum_{i=1}^N \delta u_i \left(D\varphi_i, D\varphi_j\right) + \sum_{k=1}^M \delta q_k \left(\phi_k, D\varphi_j\right) + \sum_{i=1}^N \lambda_i \delta u_i \left(\varphi_i, \varphi_j\right) = \\ - \varepsilon \sum_{i=1}^N u_i \left(D\varphi_i, D\varphi_j\right) - \sum_{k=1}^M q_k \left(\phi_k, D\varphi_j\right) - \sum_{i=1}^N \lambda_i u_i \left(\varphi_i, \varphi_j\right) + \sum_{i=1}^N \lambda_i f_i \left(\varphi_i, \varphi_j\right)\end{aligned}$$

Y con la misma notación que se usó en el caso real para plantear (2.19), obtenemos que (2.34) es equivalente a:

$$(\varepsilon M_R + \lambda M_m)\delta u + A_2 \delta q = (-\varepsilon M_R - \lambda M_m)u - A_2 q + \lambda M_m f \tag{2.35}$$

donde $\lambda = \{\lambda\}_{i,j}$ es la matriz que almacena los valores de la función $\lambda(x,y)$ al evaluarla en los puntos (x_i, y_j).

Como la condición $(1.42\mathrm{b})$ del sistema de optimalidad no depende de λ, podemos usar el análisis en diferencias finitas realizado en la subsección 2.2.2 para obtener la ecuación de incrementos de nuestro algoritmo NSS. Así, definimos las matrices:

$$\Psi := \begin{pmatrix} \varepsilon M_R + \lambda M_m & A_2 \\ -R_1 \cdot G & M \end{pmatrix}$$

donde R_1 se define como en (2.30), y:

$$\Lambda := \begin{pmatrix} -\varepsilon M_r - \lambda M_m & -A_2 & \lambda M_m \\ \gamma G & -M & 0 \end{pmatrix}$$

Donde $\lambda = \{\lambda\}_{i,j}$ tiene forma de una matriz, y la ecuación de incrementos que resolvemos está dada por:

$$\Psi \begin{pmatrix} \delta_u \\ \delta_q \end{pmatrix} = \Lambda \begin{pmatrix} u \\ q \\ f \end{pmatrix} \tag{2.36}$$

Un algoritmo para la solución de cada sistema regularizado de la ecuación de estado del problema (2.33) está dado por el algoritmo 2 visto en la sección 2.2.2.

2.3 Solución del Problema Adjunto

2.3.1 Solución del Problema Adjunto para el caso de parámetros reales (2.17)

Ahora, encontraremos la solución p del problema adjunto $(1.39\mathrm{c})$ - $(1.39\mathrm{d})$ como una ecuación lineal.

Para ello recurriremos nuevamente al problema modelo (2.17). En este problema, el sistema adjunto puede escribirse como:

$$\begin{cases} \varepsilon(Dp, Dv) + (q, Dv) + \lambda(v, p) = -(u - u_o, v) \\ q = h'_\gamma(Du)Dp \end{cases} \tag{2.37}$$

donde $u_0 \in L^2(\Omega)$ es la imagen original, o una aproximación a esta basada en un estudio previo.

Ahora, discretizamos usando el método de elementos finitos analizado en la sección 2.1

Sea $v = \varphi_j$, haciendo un análisis similar al de la aproximación por elementos finitos de la ecuación de estado:

$$\varepsilon \left(\sum_{i=1}^{N} p_i D\varphi_i, D\varphi_j \right) + \left(\sum_{i=1}^{M} q_i \phi_i, D\varphi_j \right) + \lambda \left(\sum_{i=1}^{N} p_i \varphi_i, \varphi_j \right) = - \left(\sum_{i=1}^{N} (u_i - u_{o_i}) \varphi_i, \varphi_j \right)$$

lo que equivale a:

$$\varepsilon \sum_{i=1}^{N} p_i \left(D\varphi_i, D\varphi_j\right) + \sum_{i=1}^{M} q_i \left(\phi_i, D\varphi_j\right) + \lambda \sum_{i=1}^{N} p_i \left(\varphi_i, \varphi_j\right) = -\sum_{i=1}^{N} (u_i - f_i) \left(\varphi_i, \varphi_j\right) \tag{2.38}$$

Si notamos como p al vector que almacena todos los valores de p_i y adoptamos la notación usada en la sección anterior, podemos plantear (2.38) como:

$$\varepsilon M_R.p + A_2 q + \lambda M_m.p = -M_m(u - u_o) \tag{2.39}$$

En este caso usamos la información obtenida al aplicar el algoritmo 2 y obtenemos u.

Por otro lado, la ecuación $(1.39\mathrm{d})$ para nuestro ejemplo es:

$$h'_\gamma(Du)Dp - q = 0 \tag{2.40}$$

Para resolver esta parte de nuestro problema adjunto, utlizamos la discretización en diferencias finitas (2.30) de $h'_\gamma(Du)$. Así, juntando (2.39) y (2.40) tenemos que, para hallar p es necesario resolver:

$$\begin{pmatrix} \varepsilon M_R + \lambda M_m & A_2 \\ R_1 & -I \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} -M_m(u - u_o) \\ 0 \end{pmatrix} \tag{2.41}$$

donde R_1 está dada en (2.30)

2.3.2 Solución del Problema Adjunto para el caso de parámetro polinomial (2.33)

El problema adjunto $(1.42\mathrm{c})$ - $(1.42\mathrm{d})$ del ejemplo (2.33) para el caso polinomial se puede reescribir como:

$$\begin{cases} \varepsilon(Dp, Dv) + (q, Dv) + (\lambda v, p) = -(u - u_o, v) \\ q = h'_\gamma(Du)Dp \end{cases} \tag{2.42}$$

donde u_o es la imagen original o una aproximación basada en un estudio previo.

Usando una discretización en elementos finitos para la primera ecuación $(1.42\mathrm{c})$ tenemos que esta es equivalente a:

$$\varepsilon M_R.p + A_2 q + \lambda M_m.p = -M_m(u - u_o) \tag{2.43}$$

donde λ es la matriz que almacena los valores $\lambda(x, y)$ de $(x, y) \in \Omega$.

Para la segunda ecuación, $(1.42\mathrm{d})$, utilizamos la misma discretización en diferencias finitas que usamos en (2.30). Así, para hallar la solución p del problema adjunto resolvemos el sistema:

$$\begin{pmatrix} \varepsilon M_R + \lambda M_m & A_2 \\ R_1 & -I \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} -M_m(u - u_o) \\ 0 \end{pmatrix} \tag{2.44}$$

2.4 Método cuasi-Newton para la optimización de EDPs con restricciones

Consideramos el problema general de control óptimo:

$$\begin{cases} \text{mín}\, J(y,u) \\ \text{sujeto a:} \\ e(y,u) = 0 \end{cases} \tag{2.45}$$

donde Y, U, W son espacios de Hilbert, $J : Y \times U \to \mathbb{R}$ y $e : Y \times U \to W$

Si asumimos que existe una solución única $y(u)$ a $e(y,u) = 0$ podemos reescribir el problema de control óptimo en su forma reducida:

$$\underset{u \in U}{\text{mín}}\, f(u) := J(y(u), u)$$

si f es continuamente diferenciable, en el sentido de Fréchet, podemos usar un método de descenso para obtener un punto estacionario del problema de control óptimo (2.45).

La idea de los métodos de descenso consiste en encontrar, en una iteración dada u_k, una dirección de descenso d_k, es decir:

$$f(u_k + d_k) < f(u_k)$$

como f decrece más rápido en la dirección $\nabla f(u_n)$ una primera elección para la dirección es $d_k = -\nabla f(u_k)$. A esta elección se le llama Método de descenso.

Como en el método de Newton, usamos ahora una aproximación de segundo orden para encontrar el mínimo de la función $f(u)$; la expansión de Taylor alrededor de u_k es:

$$f(u_k + d_k) \approx f(u_k) + \nabla f(u_k)^T d_k + \frac{1}{2} d_k^T B d_k$$

donde B es una aproximación de la matriz Hessiana. El gradiente de esta aproximación es:

$$\nabla f(u_k + d_k) \approx \nabla f(u_k) + B d_k$$

Igualando este gradiente a cero (que es el objeivo de la optimización) obtenemos el paso de Newton

$$d_k = -B^{-1} \nabla f(x_k) \tag{2.46}$$

Este tipo de métodos toman el nombre de Métodos Cuasi-Newton, y la variedad de métodos de este tipo se da por la elección de la matriz B; sin embargo en la mayoría de los casos se cumple que la solución es simétrica $B^T = B$ y siempre se trata de llegar a que la actualización B_{k+1} sea tan cercana como sea posible a B_k en alguna norma.

El método cuasi-Newton que usamos sigue la fórmula DFP (Davidon - Fletcher - Powell):

$$H_{k+1} = H_k - \frac{H_k.y_k.y_k^t.H_k}{y_k^t.H_k.y_k} + \frac{s_k.s_k^t}{y_k^t.s_k} \tag{2.47}$$

donde $s_k = u_{k+1} - u_k$ y $y_k = \nabla f(u_{k+1}) - \nabla f(u_k)$.

Hay que tomar en cuenta que si u es un parámetro real, los métodos cuasi-Newton no difieren entre sí.

Ya que se ha determinado una dirección de descenso d_K apropiada es importante saber qué tanto nos moveremos en dicha dirección. Para eso introduciremos una constante α_k y nuestra iteración se verá como:

$$x_{k+1} = x_k + \alpha_k d_k \tag{2.48}$$

con d_k como en (2.46)

A pesar de que cualquier dirección de la forma: $d_k = -\alpha_k H_k \nabla f(u_k)$ con $\alpha > 0$ es de descenso [26]. Lo más sensato seria escoger un α tal que sea la solución del problema

$$\alpha_k = \text{arg} \min_{\alpha > 0} f(u_k + \alpha p_k) \tag{2.49}$$

Sin embargo, en la práctica es muy costoso resolver (2.49), por lo que en su lugar se toma un algoritmo de búsqueda lineal para encontrar el parámetro α_k.

Regla de Armijo

Una estrategia de búsqueda lineal popular es la Regla de Armijo que consiste en que dada una dirección de descenso d_k de f en u_k, se escoge:$\alpha \in \{1, \frac{1}{2}, \frac{1}{4}, ...\}$ tal que se satisface la regla,

$$f(u_k + \alpha_k d_k) - f(x_k) \leq \gamma \alpha_k \nabla f(u_k)^T d_k \tag{2.50}$$

donde $\gamma \in (0, 1)$ es una constante dada (por lo general $\gamma = 10^{-4}$)

Si nos remitimos a nuestro problema original de control de imágenes (1.29), y definimos $G_\gamma : X^d \to H_0^1(\Omega)$ el operador solución que asigna a cada parámetro λ la solución de (1.29b). Tenemos que el funcional de costo del problema está dado por:

$$\mathcal{F} = g(G_\gamma(\lambda)) + \beta ||\lambda||_X^2 \tag{2.51}$$

dado que $g \in C^1$, se tiene que $\mathcal{F}$ es continuamente F-diferenciable, por lo tanto podemos usar un método cuasi Newton para encontrar el control óptimo.

A continuación analizaremos la estructura de las iteraciones del método cuasi-Newton con una fórmula DFP, siguiendo el esquema (2.48), para los problemas (2.17) para el caso real y (2.33) para el caso de un parámetro polinomial.

2.4.1 Método cuasi-Newton para la optimización del problema de Control de imágenes con parámetros reales

De (2.48) tenemos que en cada iteración λ está dado por:

$$\lambda_{k+1} = \lambda_k + \alpha_k d_k \tag{2.52}$$

donde:

$$d_k = -H\nabla\mathcal{F}(\lambda_k) \tag{2.53}$$

H sigue la fórmula DFP (2.47) que simplificada nos da:

$$H_{k+1} = \frac{s_k}{y_k} = \frac{\lambda_{k+1} - \lambda_k}{\nabla\mathcal{F}(\lambda_{k+1}) - \nabla\mathcal{F}(\lambda_k)}$$

Por otro lado, en la observación 1.8 se tiene la caracterización del gradiente $\nabla\mathcal{F}$ dada por (1.36); así, en cada paso este gradiente se actualiza a:

$$\nabla\mathcal{F}(\lambda_k) = 2\beta\lambda_k + 2\int_\Omega (u_k - f)p_k dx \tag{2.54}$$

u_k es la solución de la ecuación de estado (2.14), obtenida gracias al algoritmo 2, para el paso k-ésimo y p_k es la solución del problema adjunto (2.37) en la k-ésima iteración.

El uso de un algoritmo de búsqueda lineal disminuye el costo computacional de la resolución numérica del problema, por lo que se sigue la regla de Armijo (2.50) para determinar λ_k en (2.52):

se escoge $\alpha_k \in \left\{1, \frac{1}{2}, \frac{1}{4}, ...\right\}$ tal que se satisface:

$$\frac{1}{2}||u_{k+1} - u_0||^2_{L^2(\Omega)} + \beta\lambda^2_{k+1} - \frac{1}{2}||u_k - u_o||^2_{L^2(\Omega)} - \beta\lambda^2_k \leq 10^{-4}\alpha_k\nabla\mathcal{F}(\lambda_k)^T d_k \tag{2.55}$$

Algoritmo 3 Algoritmo de resolución del método cuasi-Newton, con la regla de Armijo, para encontrar el parámetro óptimo real.

1: fijar el contador $k = 0$, e inicializar con: $\lambda_0 = 0$, $H_0 = 1$ y $\alpha_0 = 1$
2: **mientras** $\alpha||\nabla\mathcal{F}(\lambda_k)|| < tol$ **hacer**
3: calcular $\nabla\mathcal{F}_k$ con la fórmula (2.54) con los valores de la ecuación de estado (algoritmo 2) $u_k = u(\lambda_k)$ y de la ecuación adjunta (2.41) $p_k = p(\lambda_k)$
4: calcular el tamaño del paso k-ésimo $d_k = -H_k\nabla\mathcal{F}_k$ con H_k dado por (2.47).
5: para aplicar la regla de Armijo (2.55), fijar el contador inicial $j = 0$ y hacer una aproximación inicial $\alpha_0 = 1$
6: **mientras** α_j no cumpla con la Regla de Armijo (2.55) **hacer**
7: $\alpha_{j+1} = \left(\frac{1}{2}\right)^j$
Salida: α_j
8: **fin mientras**
9: realizar la proyección $\lambda_{k+1} = \text{máx}(\lambda_k + \alpha_k d_k, 0)$
10: **fin mientras**

Observación 2.3 (Criterio de Parada). *Se ha escogido como criterio de parada al gradiente del funcional de costo multiplicado por el valor de α en cada iteración, es decir se espera que $\alpha||\nabla\mathcal{F}(\lambda_k)||$ sea menor a una tolerancia deseada.*

Observación 2.4. *Para el cálculo de las integrales se sigue la regla trapezoidal*

$$\int_a^b \int_c^d f(x,y)\ dy\ dx \approx T2D(f,h)$$

donde

$$\begin{aligned} T2D(f,h) = \frac{1}{4}h^2[&f(a,c)+f(b,c)+f(a,d)+f(b,d)+2\sum_{i=1}^{n-1} f(x_i,c) \\ &+2\sum_{i=1}^{n-1} f(x_i,d)+2\sum_{i=1}^{n-1} f(a,y_i)+2\sum_{i=1}^{n-1} f(b,y_i) \\ &+4\sum_{i=1}^{n-1}\left(\sum_{j=1}^{n-1} f(x_i,y_j)\right)] \end{aligned}$$

2.4.2 Método cuasi-Newton para la optimización del problema de Control de imágenes con parámetros polinomiales

Para obtener el polinomio $\lambda = Ax^2 + By^2 + Cxy + Dx + Ey + F$ óptimo, definimos el vector $V_k = (A_k, B_k, C_k, D_k, E_k, F_k)^T$ formado por los valores de los coeficientes del polinomio λ_k en la iteración k-ésima.

De tal manera que, cada iteración del método cuasi-Newton estará dada por:

$$V_{k+1} = V_k + \alpha_k d_k \tag{2.56}$$

donde $d_k = -H\nabla\mathcal{F}(V_k)$

H sigue la fórmula DFP (2.47) y mediante un proceso análogo al de la observación 1.8 se tiene la caracterización del gradiente obtenida a partir del sistema (1.49):

$$\nabla\mathcal{F}(V_k) = \begin{pmatrix} 2\beta\left(\frac{1}{5}A_k+\frac{1}{9}B_k+\frac{1}{8}C_k+\frac{1}{4}D_k+\frac{1}{6}E_k+\frac{1}{3}F_k\right)+\int_\Omega x^2\ (u_k-f)\ p_k\ dx \\ 2\beta\left(\frac{1}{9}A_k+\frac{1}{5}B_k+\frac{1}{8}C_k+\frac{1}{6}D_k+\frac{1}{4}E_k+\frac{1}{3}F_k\right)+\int_\Omega y^2\ (u_k-f)\ p_k\ dx \\ 2\beta\left(\frac{1}{8}A_k+\frac{1}{8}B_k+\frac{1}{9}C_k+\frac{1}{6}D_k+\frac{1}{4}E_k+\frac{1}{3}F_k\right)+\int_\Omega xy\ (u_k-f)\ p_k\ dx \\ 2\beta\left(\frac{1}{4}A_k+\frac{1}{6}B_k+\frac{1}{6}C_k+\frac{1}{3}D_k+\frac{1}{4}E_k+\frac{1}{2}F_k\right)+\int_\Omega x\ (u_k-f)\ p_k\ dx \\ 2\beta\left(\frac{1}{6}A_k+\frac{1}{4}B_k+\frac{1}{6}C_k+\frac{1}{4}D_k+\frac{1}{3}E_k+\frac{1}{2}F_k\right)+\int_\Omega y\ (u_k-f)\ p_k\ dx \\ 2\beta\left(\frac{1}{3}A_k+\frac{1}{3}B_k+\frac{1}{4}C_k+\frac{1}{2}D_k+\frac{1}{2}E_k+F_k\right)+\int_\Omega (u_k-f)\ p_k\ dx \end{pmatrix} \tag{2.57}$$

donde u_k es la solución de la ecuación de estado (2.34), obtenida gracias al algoritmo 2, para el paso k-ésimo y p_k es la solución del problema adjunto (2.44) en la k-ésima iteración.

Para determinar α_k se sigue el criterio de Armijo (2.50):

$$\mathcal{F}(V_k + \alpha_k d_k) - \mathcal{F}(V_k) \leq 10^{-4}\alpha_k \nabla\mathcal{F}(V_k)^T$$

$$\frac{1}{2}||u_{k+1}-u_0||^2_{L^2(\Omega)}+\beta\lambda(V_{k+1})^2-\frac{1}{2}||u_k-u_o||^2_{L^2(\Omega)}-\beta\lambda(V_k)^2\leq 10^{-4}\alpha_k\nabla\mathcal{F}(V_k)^T d_k \quad (2.58)$$

Finalmente, se realiza una proyección para así asegurar que: $\bar{A}\geq 0,\ \bar{B}\geq 0,\ \bar{C}\geq 0,\ \bar{D}\geq 0,\ \bar{E}\geq 0,\ \bar{F}\geq 0$

Con lo que el proceso de resolución del método cuasi-Newton con la regla de Armijo para encontrar el parámetro óptimo polinomial está dado por el algoritmo 4

Algoritmo 4 Algoritmo de resolución del método cuasi-Newton, con la regla de Armijo, para encontrar el parámetro óptimo polinomial de coeficientes no negativos

1: fijar el contador $k=0$, e inicializar con: $V_0=(0,0,0,0,0,0)^T$, $H_0=I_{6\times 6}$ y $\alpha_0=1$
2: **mientras** $\alpha||\nabla\mathcal{F}(V_k)||<tol$ **hacer**
3: calcular $\nabla\mathcal{F}_k$ con la fórmula (2.57) con los valores de la ecuación de estado (algoritmo 2) $u_k=u(\lambda_k)$ y de la ecuación adjunta (2.41) $p_k=p(\lambda_k)$
4: calcular el tamaño del paso k-ésimo $d_k=-H_k\nabla\mathcal{F}_k$ con H_k dado por (2.47).
5: para aplicar la regla de Armijo (2.58), fijar el contador inicial $j=0$ y hacer una aproximación inicial $\alpha_0=1$
6: **mientras** α_j no cumpla con la Regla de Armijo (2.58) **hacer**
7: $\alpha_{j+1}=\left(\frac{1}{2}\right)^j$
Salida: α_j
8: **fin mientras**
9: realizar la proyección $V_{k+1}=\text{máx}(V_k-\alpha_k d_k,0)$
10: **fin mientras**

CAPÍTULO 3

Resultados

3.1 La solución óptima tiene dependencia de la varianza en la distribución del ruido

Como un primer ejemplo, nos enfocamos en el problema de minimización del ruido en una imagen con parámetro real (2.17) equivalente al sistema de optimalidad (1.39). El problema consistió en encontrar la elección óptima para el parámetro de regularización de variación total si se conoce la imagen original (en la práctica, se utiliza un conjunto de entrenamiento para obtener una imagen aproximada a la original con una tolerancia dada [7])

En las figuras siguientes, presentamos los resultados de 3 experimentos relativos a este programa, en los que se muestra una misma imagen siguiendo una distribución con 3 valores distintos de varianza. Observamos que, incrementando los valores de la varianza del ruido en los datos, los parámetros óptimos varían significativamente. Esto se debe a que, al aumentar la distorsión en una imagen hay menos información inicial que se puede obtener, es aquí cuando una regularización en variación total juega un rol importante.

Después analizamos el ejemplo del caso polinomial (2.33) , en el que se puede observar el mismo comportamiento que en el caso real con respecto a la dependencia de los resultados de la varianza de la distribución del ruido en la imagen inicial.

En la figura 3.1 presentamos los resultados del primer experimento, en el que el parámetro óptimo de regularización calculado es $\lambda^* = 1612{,}8$ para el caso escalar. Este resultado fue obtenido después de 23 iteraciones. Cuando el parámetro óptimo de regularización tiene la forma de un polinomio cuadrático, el resultado calculado es $\lambda^* = 406x^2 + 362y^2 + 297,6xy + 618{,}5x + 600y + 1232{,}7$. Este resultado fue obtenido después de 23 iteraciones.

La imagen con ruido y ambas imagenes con el ruido minimizado se muestran en la figura 3.1. Estos resultados fueron obtenidos para los valores: $\gamma = 100$, $\varepsilon = 1e - 12$, $\beta = 1e - 10$ y $h = 1/100$:

En la figura 3.2 presentamos los resultados del segundo experimento, en el que el parámetro óptimo de regularización calculado es, $\lambda^* = 2055{,}7$ para el caso escalar. Este resultado fue obtenido después de 24 iteraciones. Cuando el parámetro óptimo de regularización tiene la forma de un polinomio cuadrático, el resultado calculado es $\lambda^* =$

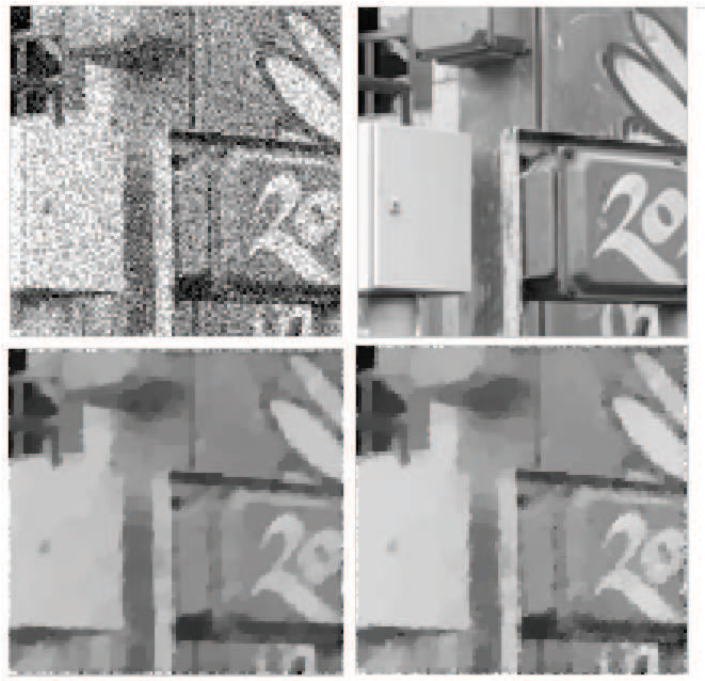

Figura 3.1: *Primera fila: Imagen con ruido (izquierda) e imagen original (derecha). Segunda fila: Imágenes con ruido minimizado de parámetros óptimos* $\lambda^* = 1612{,}8$ *y* $\lambda^* = 406x^2 + 362y^2 + 297{,}6xy + 618{,}5x + 600y + 1232{,}7$ *respectivamente*

$579{,}1x^2 + 519{,}1y^2 + 430{,}4xy + 824{,}5x + 818{,}9y + 1518{,}6$. Este resultado fue obtenido después de 27 iteraciones.

La imagen con ruido y ambas imagenes con el ruido minimizado se muestran en la figura 3.2. Estos resultados fueron obtenidos para los valores: $\gamma = 100$, $\varepsilon = 1e - 12$, $\beta = 1e - 10$ y $h = 1/100$.

En la figura 3.3 presentamos los resultados del tercer experimento: en el que el parámetro óptimo de regularización calculado es $\lambda^* = 1383{,}1$ para el caso escalar. Este resultado fue obtenido después de 19 iteraciones. Cuando el parámetro óptimo de regularización tiene la forma de un polinomio cuadrático, el resultado calculado es $\lambda^* = 317{,}5762x^2 + 282{,}3121y^2 + 232{,}8364xy + 482{,}8629x + 467{,}9869y + 960{,}1977$. Este resultado fue obtenido después de 22 iteraciones.

La imagen con ruido y ambas imagenes con el ruido minimizado se muestran en la figura 3.3. Estos resultados fueron obtenidos para los valores: $\gamma = 100$, $\varepsilon = 1e - 12$, $\beta = 1e - 10$ y $h = 1/100$.

En el Anexo B se encuentran los codigos en Matlab que nos llevaron a obtener estos resultados.

En el grafico 3.4 se puede ver la evolución del funcional de costo para el segundo experimento, para los casos real y polinomial. En ambos casos se llega al mismo valor, sin embargo para el caso polinomial esto toma más iteraciones.

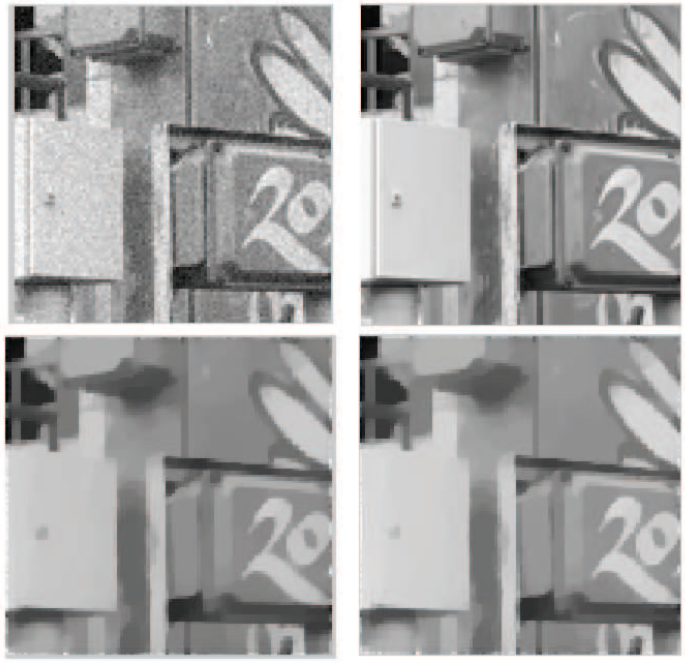

Figura 3.2: *Primera fila: Imagen con ruido (izquierda) e imagen original (derecha). Segunda fila: Imágenes con ruido minimizado de parámetros óptimos* $\lambda^* = 2055{,}7$ *y* $\lambda^* = 579{,}1x^2 + 519{,}1y^2 + 430{,}4xy + 824{,}5x + 818{,}9y + 1518{,}6$ *respectivamente*

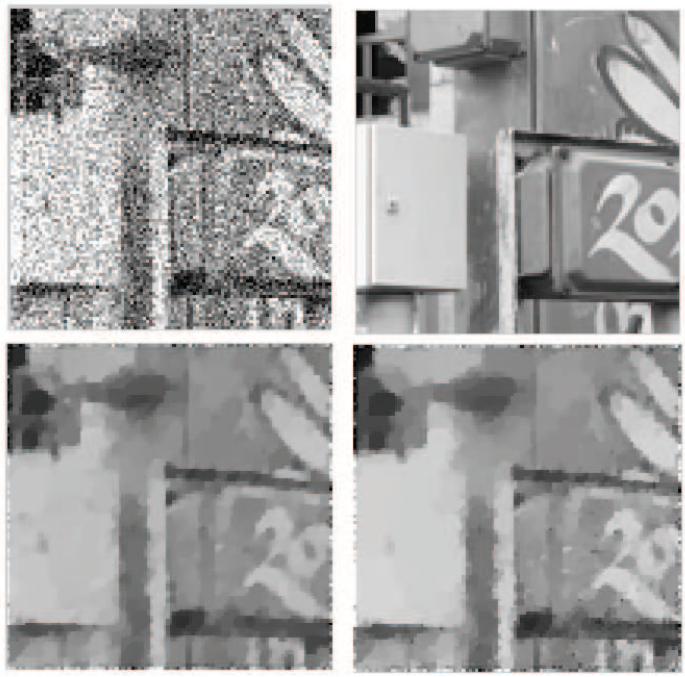

Figura 3.3: *Primera fila: Imagen con ruido (izquierda) e imagen original (derecha). Segunda fila: Imágenes con ruido minimizado de parámetros óptimos* $\lambda^* = 1383{,}1$ *y* $\lambda^* = 317{,}5762x^2 + 282{,}3121y^2 + 232{,}8364xy + 482{,}8629x + 467{,}9869y + 960{,}1977$ *respectivamente*

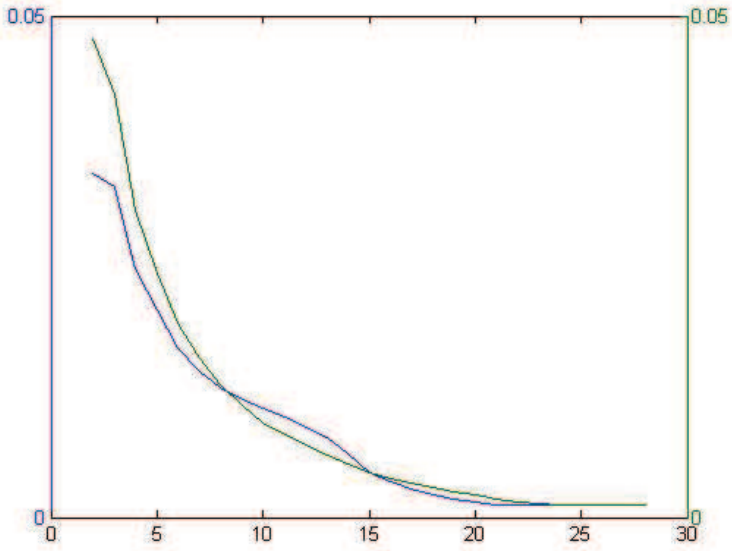

Figura 3.4: *Gráfico de la evolución del funcional de costo vs el número de iteración para el segundo experimento. En verde para el caso polinomial y en azul para el caso real.*

3.2 Dependencia del mallado

Con respecto a la dependencia del mallado en la tabla 3.1 constan el número iteraciones del método cuasi Newton para diferentes tamaños de malla para el caso real así como el valor óptimo λ^* obtenido. Se puede observar un comportamiento dependiente del número de puntos del mallado, pues para cada uno de ellos se obtiene un control óptimo de distinto valor. Sin embargo, el control varía relativamente poco lo que implica un comportamiento robusto de los valores.

# puntos en el mallado	60	65	70	75
# iteraciones para el caso real	22	21	25	23
# λ^*	1062.9	1134.2	1003	1437

Tabla 3.1: *relación entre el mallado para el método cuasi Newton para el primer experimento en el caso real*

En la tabla 3.2 constan el número iteraciones del método cuasi Newton para diferentes tamaños de malla para el caso polinomial, así como los valores $A^*, B^*, C^*, D^*, E^*, F^*$ del polinomio óptimo obtenido. Se puede observar un comportamiento dependiente del número de puntos del mallado, pues para cada uno de ellos se obtiene un control óptimo de distinto valor. Sin embargo, al igual que en el caso real, el control varía relativamente poco lo que implica un comportamiento robusto de los valores.

# puntos en el mallado	60	65	70	75
# iteraciones para el caso real	23	23	24	24
# A^*	260.16	285.09	309.96	321.79
# B^*	221.32	242.69	264.97	276.53
# C^*	187.89	205.84	223.06	231.76
# D^*	392.08	429.29	469.27	488.87
# E^*	369.09	403.35	440.44	459.39
# F^*	774.86	842.59	922.38	958.03

Tabla 3.2: *relación entre el mallado para el método cuasi Newton para el primer experimento en el caso polinomial*

3.3 Comparación del ruido como un parámetro real y polinomial

En los siguientes experimentos se pretende minimizar la presencia del ruido de la figura 3.5 usando ambos ejemplos, el real (2.17) y el polinomial (2.33)

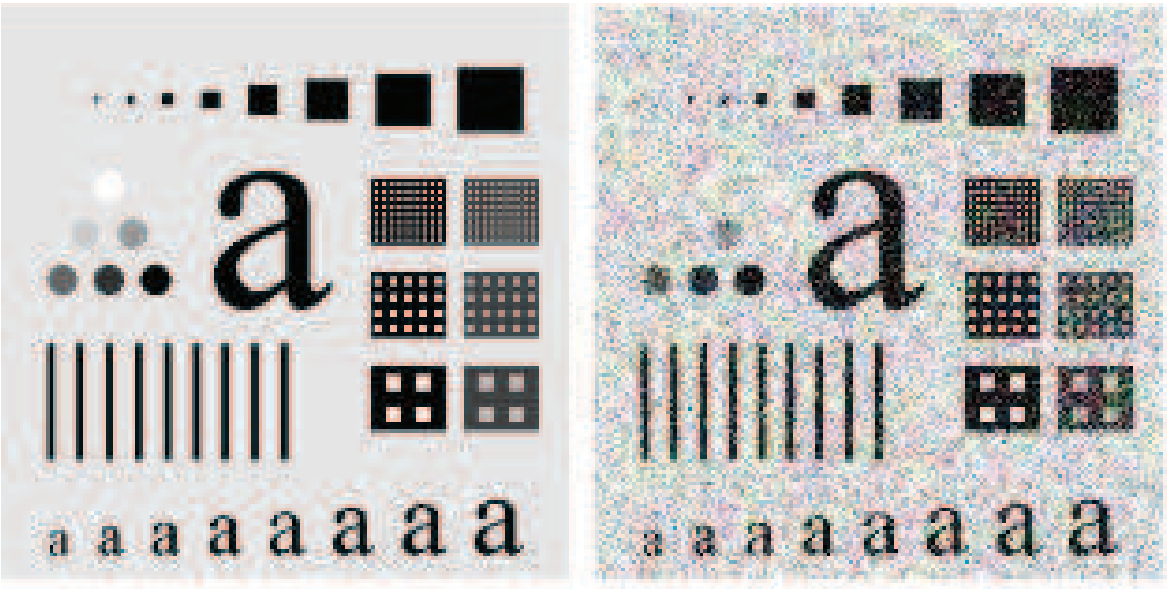

Figura 3.5: *Imagen original (izquierda) y con un ruido de varianza igual a* $0{,}05$ *(derecha)*

En el primer experimento se resolvió el problema de minimización del ruido en la imagen 3.5 con parámetro real (2.17) equivalente al sistema de optimalidad (1.39). El problema consistió en encontrar la elección óptima para el parámetro de regularización de variación total si se conoce la imagen original. Después de 22 iteraciones se obtuvo el parámetro óptimo de $\lambda^* = 1915{,}8$, y el funcional de costo calculado fue de $0{,}012$.

Después analizamos el ejemplo del caso polinomial (2.33) para el que, después de 20 iteraciones, el polinomio cuadrático óptimo es: $\lambda^* = 394{,}2x^2 + 324{,}4y^2 + 278{,}5xy + 580{,}8x + 550{,}2y + 1162$ con un funcional de costo de $0{,}012$

En la figura 3.6 presentamos los resultados de ambos experimentos. Estos resultados

fueron obtenidos para los valores: $\gamma = 100$, $\varepsilon = 1e-12$, $\beta = 1e-10$ y $h = 1/150$.

Figura 3.6: *Imágenes con el parámetro λ óptimo para el ejemplo mostrado en la figura 3.5. A la izquierda, tenemos el resultado de la minimización del ruido en la imagen con un parámetro real, mientras que a la derecha tenemos la misma imagen resuelta para un parámetro λ polinomial cuadratico. En el primer caso, la solución óptima de $\lambda^* = 1915{,}8$ nos da como resultado un funcional de costo igual a $0{,}012$, mientras que el parámetro polinomial cuadrático computado en el segundo experimento $\lambda^* = 394{,}2x^2 + 324{,}4y^2 + 278{,}5xy + 580{,}8x + 550{,}2y + 1162$ nos lleva a una imagen con ruido disminuido con un funcional de costo igual a $0{,}012$*

En ambos casos, el funcional de costo alcanza $0{,}012$, sin embargo para el caso polinomial esto sucedió en menos iteraciones que en el caso real, estas fueron de 19 y 22 respectivamente.

En los siguientes ejemplos, veremos dos casos: en uno de ellos esta tendencia (de lograr el mismo valor del funcional de costo pero en menos iteraciones) se cumple, pero en el otro no.

3.4 Valores del Parámetro de Regularización γ

En la primera sección, se dijo que una condición para que el operador relajado (1.13) converja hacia $\mathcal{J}$ es que el parámetro de la regularización de Huber γ en (1.10) se acerque a infinito. En la práctica, valores mayores a 3 nos dan convergencia del método hacia los parámetros de regularización óptimos.

En los siguientes experimentos se pretende encontrar el parámetro λ óptimo para los caso real (2.17) y polinomial (2.33) con distintos valores de γ para las imágenes 3.7 y 3.8

En la figura 3.7 se puede observar la imagen original y la versión con ruido gaussiano con varianza igual a $0{,}05$ de esta, mientras que en la figura 3.8 se presenta otra imagen con una presencia de ruido mucho más agresiva: de media igual a $0{,}02$ y varianza igual a $0{,}1$.

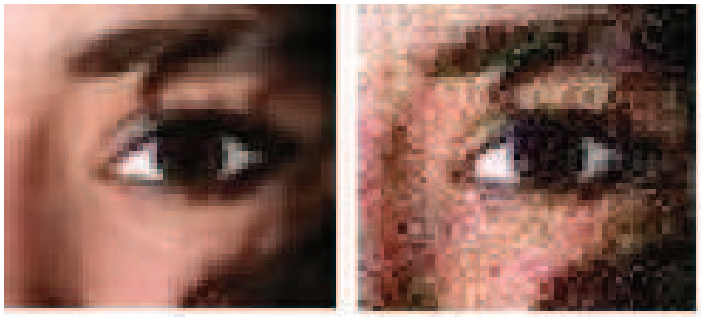

Figura 3.7: *Imagen original (izquierda) y con un ruido de varianza igual a* 0,05 *(derecha)*

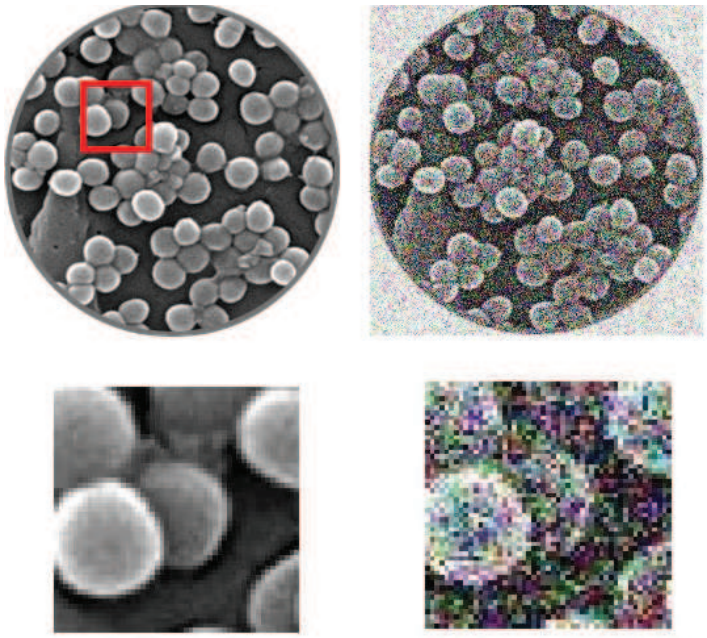

Figura 3.8: *Imagen original (superior izquierda) y con un ruido de media igual a* 0,02 *y varianza igual a* 0,1 *(superior, derecha). En los siguientes experimentos hemos tomado la porción aquí mostrada (imagenes en la parte inferior)*

Esta diferencia se ve presente en los resultados, donde el funcional de costo de la segunda imagen es significativamente mayor al obtenido para la primera.

En la tabla 3.3 se pueden ver los valores del experimento para el problema real en ambas imágenes. En la figura 3.9 se muestran las imágenes con ruido disminuido.

γ	Figura 3.7			Figura 3.8		
	λ	F. de Costo	# de It.	λ	F. de Costo	# de It.
0	No converge			No converge		
0.1	No converge			No converge		
1	No converge			No converge		
3	1857.2	0.001	26	344.94	0.014	23
5	1845.7	0.001	23	470.18	0.013	19
100	2025.8	0.001	22	370.71	0.013	18
1000	1687.2	0.001	21	414.35	0.013	19
10^8	1691.3	0.001	21	416.16	0.013	19

Tabla 3.3: *Resultados del parámetro óptimo real λ para distintos valores de γ en las figuras 3.7 y 3.8 con $\beta = 10^{-12}$, $\varepsilon = 10^{-10}$ y un mallado de 50 puntos*

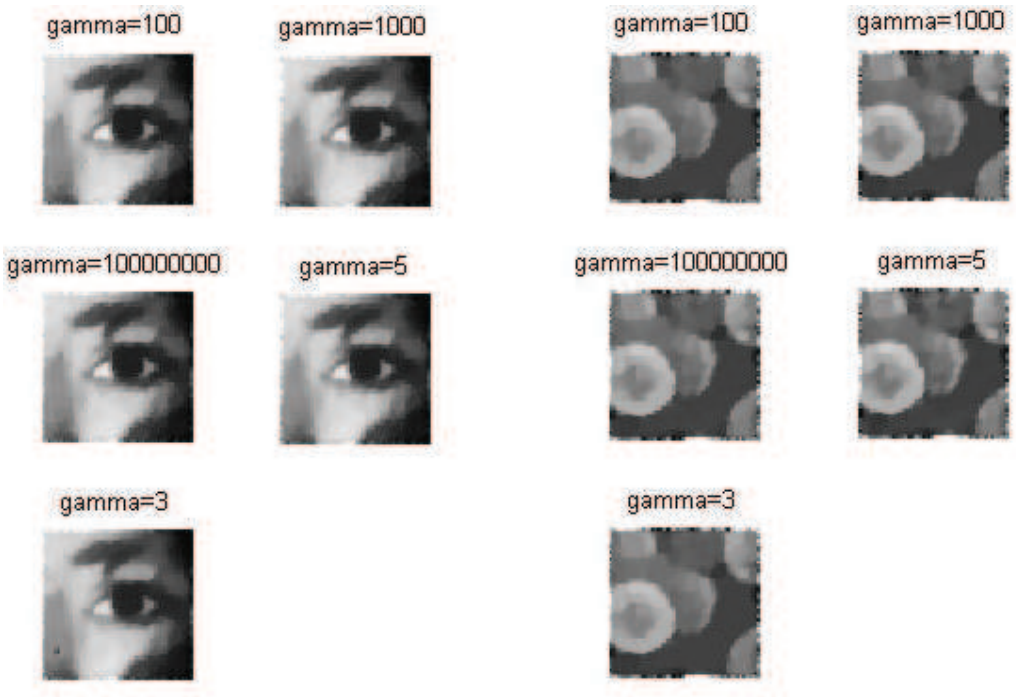

Figura 3.9: *Resultados del parámetro óptimo real λ para distintos valores de γ en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.3 con $\beta = 10^{-12}$, $\varepsilon = 10^{-10}$ y un mallado de 50 puntos*

A continuación, se realizó el mismo experimento para encontrar el parámetro λ óptimo pero esta vez para el caso polinomial (2.33) con distintos valores de γ para las imágenes 3.7 y 3.8

En la tabla 3.4 se pueden ver los valores del experimento para el problema polinomial

en ambas imágenes. En la figura 3.10 se muestran los resultados de dichos experimentos.

γ	**Figura 3.7**		
	λ	**F. de Costo**	**# de It**
0	No Converge		
0.1	No Converge		
1	$15{,}85x^2 + 10{,}12y^2 + 10{,}22xy + 22{,}62x + 18{,}19y + 42{,}35$	0.704	15
3	$200{,}57x^2 + 125{,}27y^2 + 126{,}55xy + 294{,}03x + 235{,}30y + 560{,}62$	0.002	19
5	$270{,}02x^2 + 167{,}52y^2 + 169{,}74xy + 396{,}09x + 315{,}25y + 753{,}98$	0.002	19
100	$290{,}06x^2 + 182{,}46y^2 + 184{,}24xy + 425{,}35x + 341{,}78y + 807{,}30$	0.002	20
1000	$291{,}64x^2 + 183{,}50y^2 + 185{,}28xy + 427{,}68x + 343{,}09y + 811{,}70$	0.001	20
10^8	$291{,}82x^2 + 183{,}62y^2 + 185{,}40xy + 427{,}95x + 343{,}91y + 812{,}19$	0.001	20
γ	**Figura 3.8**		
	λ	**F. de Costo**	**# de It.**
0	No converge		
0.1	No converge		
1	$4{,}94x^2 + 3{,}51y^2 + 3{,}34xy + 6{,}7259x + 5{,}4471y + 11{,}83$	0.094	8
3	$84{,}06x^2 + 72{,}02y^2 + 58{,}95xy + 132{,}72x + 123{,}91y + 278{,}24$	0.013	26
5	$86{,}03x^2 + 72{,}81y^2 + 59{,}93xy + 136{,}59x + 126{,}57y + 288{,}21$	0.013	22
100	$92{,}94x^2 + 79{,}26y^2 + 65{,}23xy + 147{,}71x + 137{,}72y + 311{,}45$	0.013	24
1000	$93{,}40x^2 + 79{,}73y^2 + 65{,}60xy + 148{,}44x + 138{,}51y + 312{,}98$	0.013	24
10^8	$93{,}41x^2 + 79{,}75y^2 + 65{,}62xy + 148{,}47x + 138{,}54y + 313{,}03$	0.013	24

Tabla 3.4: *Resultados del parámetro óptimo polinomial λ para distintos valores de γ en las figuras 3.7 y 3.8 con $\beta = 10^{-12}$, $\varepsilon = 10^{-10}$ y un mallado de 50 puntos*

Tanto en la tabla 3.3 como en la tabla 3.4 se puede ver que el programa converge a partir de $\gamma = 3$, y que el funcional de costo disminuye a medida que γ aumenta. En este sentido, los resultados experimentales concuerdan con lo dicho en la primera sección respecto al parámetro de regularización de Huber (1.10): a medida que γ tiende hacia infinito, la función de Huber se acerca más al valor del subdiferencial. Además, se ratifica el hecho de que el funcional de costo alcanza los mismos valores para el caso real y el polinomial.

3.5 Valores del Parámetro de Difusión Artificial ε

En la primera sección, se introdujo un parámetro de difusión artificial ε para obtener (1.14) y así reducir el problema de Reducción del Ruido en una Imagen al espacio funcional H_0^1. Veremos que, a través de la experimentación, el parámetro ε no afecta en la búsqueda de soluciones óptimas, siempre y cuando este sea mucho menor a 1.

En los siguientes experimentos se pretende encontrar el parámetro λ óptimo para los caso real (2.17) y polinomial (2.33) con distintos valores de ε para las imágenes 3.7 y 3.8

En la tabla 3.5 se pueden ver los valores del experimento para el problema real en ambas imágenes. En la figura 3.11 se muestran las imágenes con ruido disminuido.

A continuación, se realizó el mismo experimento para encontrar el parámetro λ óptimo pero esta vez para el caso polinomial (2.33) con distintos valores de ε para las imágenes

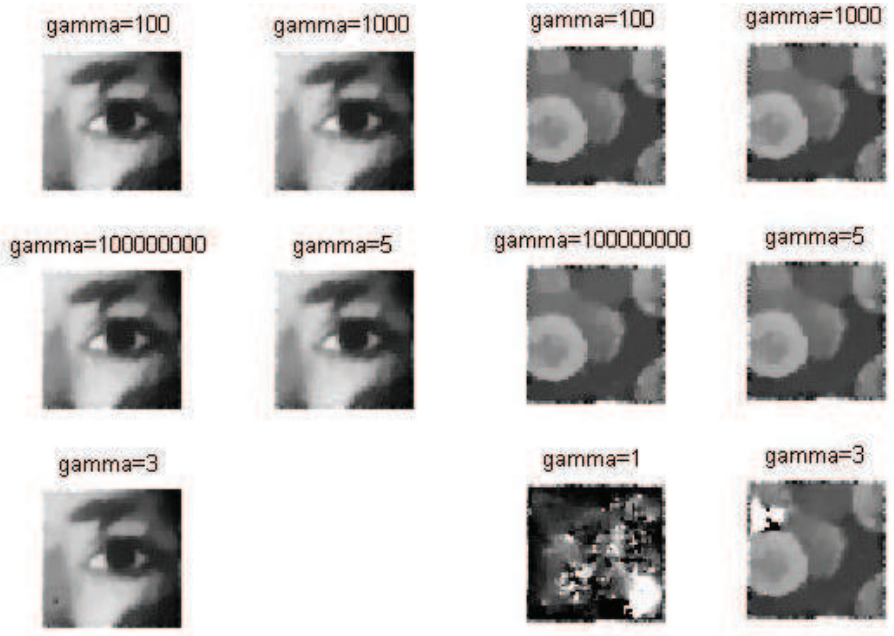

Figura 3.10: *Resultados del parámetro óptimo polinomial λ para distintos valores de γ en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.4 con $\beta = 10^{-12}$, $\varepsilon = 10^{-10}$ y un mallado de 50 puntos*

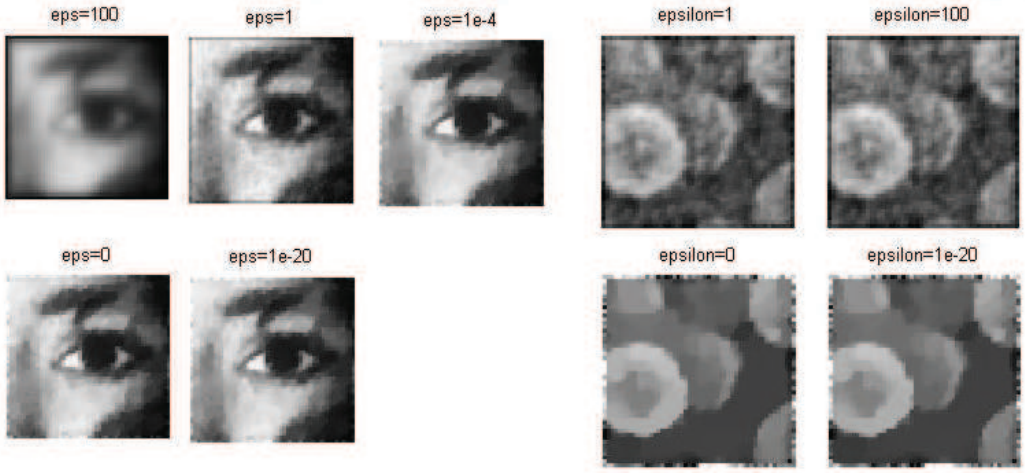

Figura 3.11: *Resultados del parámetro óptimo real λ para distintos valores de ε en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.5 con $\beta = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

ε	**Figura 3.7**			**Figura 3.8**		
	λ	**F. de Costo**	**# de it.**	λ	**F. de Costo**	**# de it.**
0	2025.8	0.001	22	370.71	0.013	18
10^{-20}	2025.8	0.001	22	370.71	0.013	18
10^{-4}	1816.3	0.001	26	1.0612	0.113	26
1	1685.8	0.002	21	4912.1	0.014	18
100	8036.4	0.031	15	4234.4	0.013	19

Tabla 3.5: *Resultados del parámetro óptimo real λ para distintos valores de ε en las figuras 3.7 y 3.8 con $\beta = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

3.7 y 3.8

En la tabla 3.6 se pueden ver los valores del experimento para el problema polinomial en ambas imágenes. En la figura 3.12 se muestran los resultados de dichos experimentos.

ε	**Figura 3.7**		
	λ	**F. de Costo**	**# de it**
0	$290{,}06x^2 + 182{,}46y^2 + 184{,}24xy + 425{,}35x + 341{,}78y + 807{,}30$	0.002	20
10^{-20}	$290{,}06x^2 + 182{,}46y^2 + 184{,}24xy + 425{,}35x + 341{,}78y + 807{,}30$	0.002	20
10^{-4}	$298{,}53x^2 + 175{,}96y^2 + 176{,}99xy + 441{,}73x + 329{,}03y + 859{,}44$	0.002	20
1	$942{,}6x^2 + 613y^2 + 609{,}5xy + 1370x + 1137{,}6y + 2601{,}6$	0.015	15
100	$2340{,}9x^2 + 1524{,}8y^2 + 1515{,}9xy + 3402{,}8x + 2829{,}6y + 6458{,}9$	0.099	7
ε	**Figura 3.8**		
	λ	**F. de Costo**	**# de it**
0	$92{,}94x^2 + 79{,}26y^2 + 65{,}23xy + 147{,}71x + 137{,}72y + 311{,}45$	0.0130	24
10^{-20}	$92{,}94x^2 + 79{,}26y^2 + 65{,}23xy + 147{,}71x + 137{,}72y + 311{,}45$	0.013	24
10^{-4}	$90{,}57x^2 + 72{,}84y^2 + 60{,}12xy + 144{,}28x + 126{,}90y + 307{,}04$	0.013	26
1	$1231{,}9x^2 + 1097{,}3y^2 + 893xy + 1935{,}3x + 1893{,}3y + 4088{,}7$	0.015	17
100	$14163x^2 + 12652y^2 + 10288xy + 22256x + 21820y + 47024$	0.023	12

Tabla 3.6: *Resultados del parámetro óptimo polinomial λ para distintos valores de ε en las figuras 3.7 y 3.8 con $\beta = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

Tanto en la tabla 3.5 como en la tabla 3.6 se puede ver que el programa converge para todos los valores propuestos, sin embargo que se empiezan a obtener valores óptimos de λ, con funcional de costo reducido, para cantidades de ε por debajo de 10^{-4}, y que el funcional de costo disminuye a medida que ε se acerca a 0. En este sentido, los resultados experimentales concuerdan con lo dicho en la primera sección respecto al parámetro de difusión artificial ε en (1.14): a medida que ε tiende a 0, se obtiene una respuesta más cercana al problema original en el Espacio de Variación Acotada.

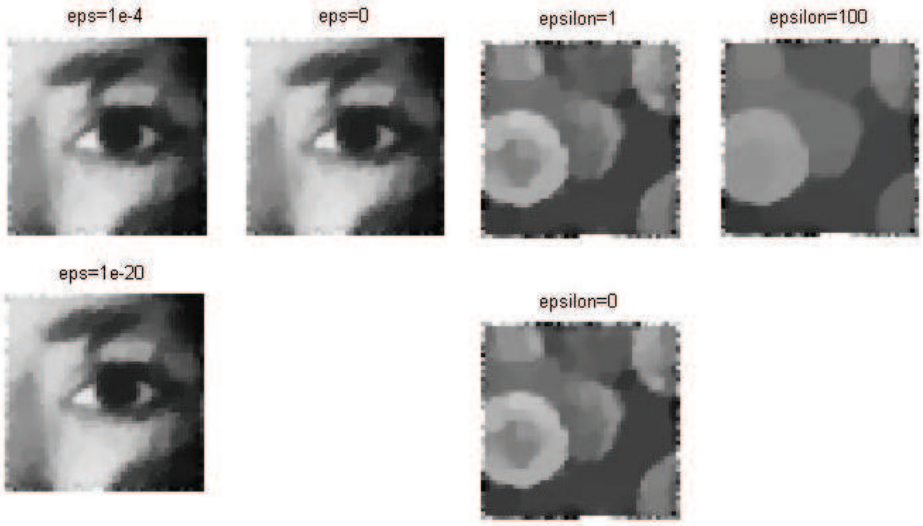

Figura 3.12: *Resultados del parámetro óptimo polinomial λ para distintos valores de ε en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.6 con $\beta = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

3.6 Valores del Peso β

En la primera sección definimos el funcional a ser minimizado como $(1.8a)$. Hay dos objetivos al tomar este como nuestro funcional a ser minimizado: el primero, es reducir lo mejor posible la diferencia entre la imagen que se obtendrá al final y la imagen original a través de la función g; el segundo objetivo es obtener un λ no muy grande (pues mientras más grande es λ mayor es la presencia del ruido); sin embargo, para que el segundo objetivo no entre en clonflicto con la función g, es necesario asignarle un peso β al valor de λ en el funcional de costo. Como veremos en esta sección, mientras menor valor tenga este peso será mejor para nuestro modelo.

En los siguientes experimentos se pretende encontrar el parámetro λ óptimo para los caso real (2.17) y polinomial (2.33) con distintos valores de β para las imágenes 3.7 y 3.8

En la tabla 3.7 se pueden ver los valores del experimento para el problema real en ambas imágenes. En la figura 3.13 se muestran las imágenes con ruido disminuido.

β	**Figura 3.7**			**Figura 3.8**		
	λ	**F. de Costo**	**# de it.**	λ	**F. de Costo**	**# de it.**
0	2032.8	0.001	22	370.71	0.013	18
10^{-20}	2032.8	0.001	22	370.71	0.013	18
10^{-7}	148.79	0.009	23	148.77	0.018	15
10^{-5}	No Converge			No Converge		
1	0.1879	0.191	5	No Converge		

Tabla 3.7: *Resultados del parámetro óptimo real λ para distintos valores de β en las figuras 3.7 y 3.8 con $\varepsilon = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

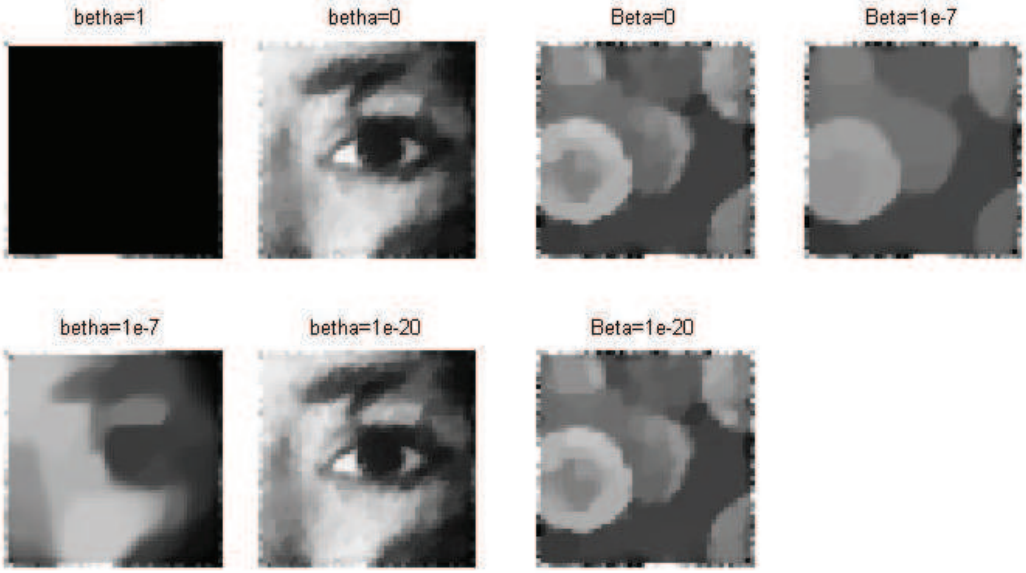

Figura 3.13: *Resultados del parámetro óptimo real λ para distintos valores de β en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.7 con $\varepsilon = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

A continuación, se realizó el mismo experimento para encontrar el parámetro λ óptimo pero esta vez para el caso polinomial (2.33) con distintos valores de β para las imágenes 3.7 y 3.8

En la tabla 3.8 se pueden ver los valores del experimento para el problema polinomial en ambas imágenes. En la figura 3.14 se muestran los resultados de dichos experimentos.

β	**Figura 3.7**		
	λ	**F. de Costo**	**# de it**
0	$290{,}43x^2 + 184{,}58y^2 + 186{,}38xy + 430{,}30x + 345{,}75y + 816{,}70$	0.001	19
10^{-20}	$290{,}43x^2 + 184{,}58y^2 + 186{,}38xy + 430{,}30x + 345{,}75y + 816{,}70$	0.001	20
10^{-7}	$41{,}58x^2 + 26{,}48y^2 + 26{,}73xy + 60{,}58x + 48{,}58y + 114{,}21$	0.016	13
10^{-5}	$11{,}06x^2 + 7{,}33y^2 + 7{,}39xy + 15{,}77x + 12{,}88y + 29{,}06$	0.015	15
1	No Converge		
β	**Figura 3.8**		
	λ	**F. de Costo**	**# de it**
0	$92{,}94x^2 + 79{,}26y^2 + 65{,}23xy + 147{,}71x + 137{,}72y + 311{,}45$	0.013	24
10^{-20}	$92{,}94x^2 + 79{,}26y^2 + 65{,}23xy + 147{,}71x + 137{,}72y + 311{,}45$	0.013	24
10^{-7}	$931{,}35x^2 + 26{,}84y^2 + 22{,}21xy + 49{,}42x + 46{,}13y + 103{,}45$	0.022	13
10^{-5}	No Converge		
1	$0{,}19x^2 + 0{,}07y^2 + 0{,}22xy + 0{,}08x - 0{,}05y + 0{,}12$	0.245	5

Tabla 3.8: *Resultados del parámetro óptimo polinomial λ para distintos valores de β en las figuras 3.7 y 3.8 con $\varepsilon = 10^{-12}$, γ y un mallado de 50 puntos*

Tanto en la tabla 3.7 como en la tabla 3.8 se puede ver que el programa no converge para valores más grandes que $\beta = 10^{-7}$, y que el funcional de costo disminuye a medida

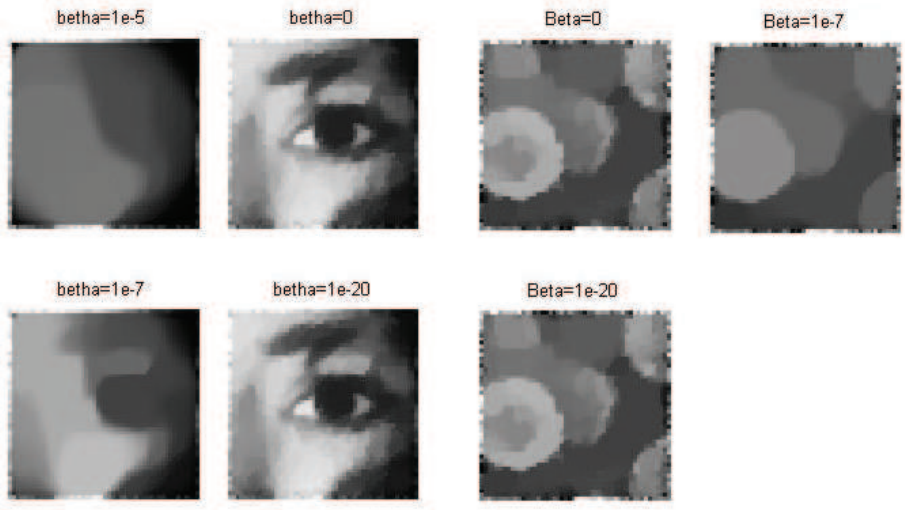

Figura 3.14: *Resultados del parámetro óptimo polinomial λ para distintos valores de β en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.8 con $\varepsilon = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

que β se acerca a 0. Se puede observar, además, que incluso para $\beta = 0$ se obtiene el valor óptimo de λ con funconal de costo reducido. En este sentido, los resultados experimentales concuerdan con lo dicho en la primera sección respecto al peso β para $(1.8\mathrm{a})$.

3.7 Dependencia del Valor Inicial λ_0

Finalmente, es importante comparar el desempeño de los algoritmos 3 para el caso real y 4 para el caso polinomial con distintos valores iniciales λ_0

En los siguientes experimentos se pretende encontrar el parámetro λ óptimo para los caso real (2.17) y polinomial (2.33) con distintos valores iniciales λ_0 para las imágenes 3.7 y 3.8

En la tabla 3.9 se pueden ver los valores del experimento para el problema real en ambas imágenes. En la figura 3.15 se muestran las imágenes con ruido disminuido.

A continuación, se realizó el mismo experimento para encontrar el parámetro λ óptimo pero esta vez para el caso polinomial (2.33) con distintos valores iniciales λ_0 para las imágenes 3.7 y 3.8

En la tabla 3.10 se pueden ver los valores del experimento para el problema polinomial de la experimentación en ambas imágenes. En la figura 3.16 se muestran los resultados de dichos experimentos.

Tanto en la tabla 3.9 como en la tabla 3.10 se puede ver que el programa no converge para valores iniciales de λ muy pequeños (menores que 10^{-6}), y que el funcional de costo disminuye a medida que el λ_0 inicial se acerca a la respuesta obtenida (en nuestro caso, valores cercanos a 100, sin que este hecho afecte en gran medida al funcional de costo obtenido (pues el valor del funcional de costo para todos los casos varía ligeramente para distintos valores de λ_0 a partir de 1).

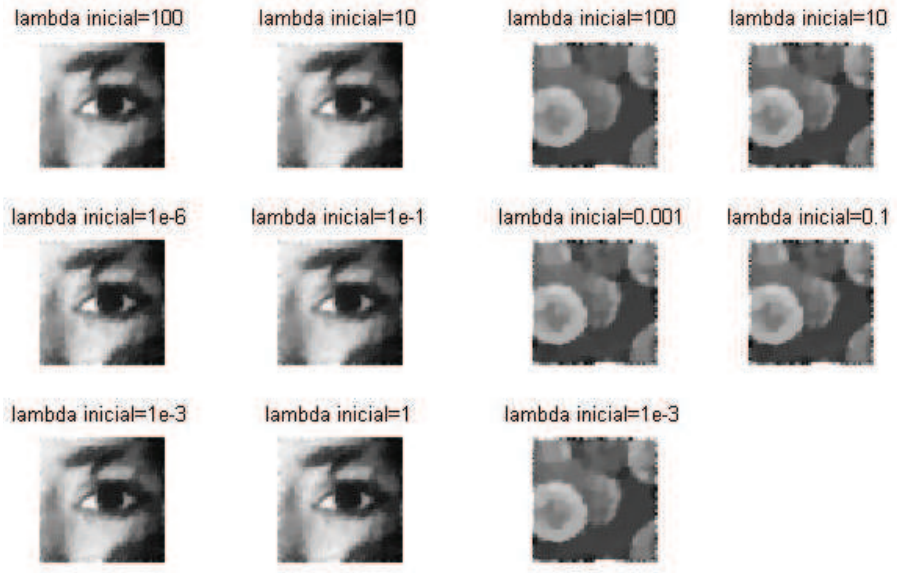

Figura 3.15: *Resultados del parámetro óptimo real λ para distintos valores iniciales de λ en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.9 con $\beta = 10^{-12}$, $\varepsilon = 10^{-10}$, $\gamma = 100$ y un mallado de 50 puntos*

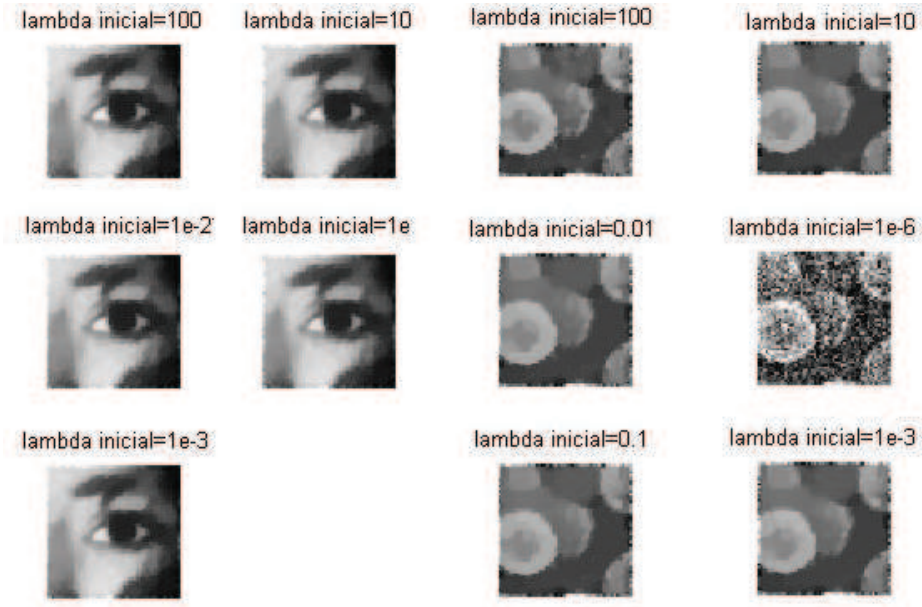

Figura 3.16: *Resultados del parámetro óptimo polinomial λ para distintos valores iniciales de λ en las figuras 3.7 (izquierda) y 3.8 (derecha) que se pueden ver en detalle en la tabla 3.10 con $\beta = 10^{-20}$, $\varepsilon = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos. Los valores de λ que se pueden ver en cada imagen, corresponden al valor de una constante que se multiplica por un vector de unos para obtener así el vector λ con el que se inician las iteraciones*

λ_0	**Figura 3.7**			**Figura 3.8**		
	λ	**F. de Costo**	**# de it**	λ	**F. de Costo**	**#de it.**
0	No Converge			No Converge		
10^{-6}	2643.6	0.001	5	No Converge		
10^{-3}	1836.1	0.001	24	426.54	0.013	10
0.01	677.44	0.001	26	461.19	0.013	16
0.1	1711	0.001	22	456.707	0.013	19
1	2032.8	0.001	22	438.70	0.013	18
10	1715.6	0.001	24	438.92	0.013	18
100	1792.7	0.001	14	366.58	0.013	7

Tabla 3.9: *Resultados del parámetro óptimo real λ para distintos valores inciales de λ en las figuras 3.7 y 3.8 con $\beta = 10^{-20}$, $\varepsilon = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

λ_0	**Figura 3.7**		
	λ	**F. de Costo**	**# de it.**
0	No Converge		
10^{-6}	No Converge		
10^{-3}	$231{,}46x^2 + 151{,}37y^2 + 132{,}51xy + 338{,}45x + 268{,}51y + 733{,}84$	0.003	10
0.01	$271{,}69x^2 + 170{,}54y^2 + 172{,}74xy + 396{,}52x + 318{,}95y + 748{,}32$	0.002	20
0.1	$262{,}28x^2 + 166{,}05y^2 + 167{,}07xy + 385{,}05x + 310{,}80y + 731{,}48$	0.002	19
10	$273{,}35x^2 + 157{,}73y^2 + 166{,}57xy + 389{,}81x + 297{,}16y + 735{,}03$	0.002	17
100	$291{,}10x^2 + 202{,}36y^2 + 186{,}88xy + 409{,}76x + 312{,}82y + 936{,}21$	0.001	8
λ_0	**Figura 3.8**		
	λ	**F. de Costo**	**# de it.**
0	No Converge		
10^{-6}	$20396x^2 + 11522y^2 + 10227xy + 32578x + 21638y + 68727$	0.027	5
10^{-3}	$93{,}40x^2 + 51{,}19y^2 + 44{,}72xy + 151{,}71x + 104{,}11y + 337{,}17$	0.013	16
0.01	$91{,}20x^2 + 79{,}65y^2 + 65{,}06xy + 145{,}43x + 138{,}44y + 307{,}70$	0.013	26
0.1	$91{,}99x^2 + 78{,}70y^2 + 64{,}39xy + 146{,}69x + 136{,}95y + 310{,}24$	0.013	26
10	$97{,}50x^2 + 81{,}22y^2 + 68{,}65xy + 153{,}03x + 139{,}64y + 321{,}94$	0.013	21
100	$164{,}38x^2 + 82{,}14y^2 + 95{,}64xy + 201{,}32x + 116{,}78y + 418{,}31$	0.013	8

Tabla 3.10: *Resultados del parámetro óptimo polinomial λ para distintos valores iniciales de λ (los valores que se pueden ver en la primera columna de la tabla corresponden a la constate que se multiplica por un vector de unos) en las figuras 3.7 y 3.8 con $\beta = 10^{-20}$, $\varepsilon = 10^{-12}$, $\gamma = 100$ y un mallado de 50 puntos*

CAPÍTULO 4

Conclusiones y Recomendaciones

En este trabajo se ha propuesto una metodología para determinar un arreglo óptimo para un problema genérico (1.8) de control de imágenes en el espacio de Variación Acotada en el que se permiten distintos tipos de ruido presentes en un conjunto de datos. Para transformar nuestro problema a una versión que permita una resolución numérica más cómoda, empleamos una regularización de tipo Huber para el término en el que se veía involucrada la Variación Total (1.11). Más aún, llevamos este problema al espacio de Hilbert H_0^1 sumando una regularización elíptica al modelo original (1.14) que, de acuerdo a la experimentación, no afecta a la solución del problema original. Basados en trabajos previos, que se pueden encontrar como referencias bibliográficas, se garantiza la existencia, consistencia en la aproximación y diferenciabilidad del operador solución. Está última característica nos permitió obtener un sistema de optimalidad para tres casos: el primero, en el que el parámetro óptimo es un número real, el segundo para el que el parámetro óptimo λ^* es un polinomio cuadrático con coeficientes positivos, y el tercero para una función en L^2 en general.

En la segunda parte de este trabajo, se encuentra una amplia discusión acerca de la resolución numérica de los sistemas de optimalidad para el problema de parámetro real (1.39) y del polinomio con coeficientes positivos (1.42) encontrados en la primera parte.

Los parámetros óptimos para ambos casos son calculados usando un método quasi-Newton, en conjunto con un método de tipo Newton semismooth.

Se verifica de manera empírica, en base al cálculo del funcional de costo, que dicha solución óptima en efecto minimiza el ruido presente en una imagen para ambos casos: el caso polinomial y real, para los cuales se alcanza un funcional de costo minimizado. A pesar de que para algunos casos el modelo polinomial mostró convergencia en menos iteraciones, esto no sucedió en todos los casos por lo que no se puede llegar a ninguna conclusión al respecto más que se alcanza el mismo valor de funcional de costo para ambos casos.

Los modelos aquí expuestos presentan una desventaja: para el cálculo del parámetro óptimo de minimización del ruido es necesario conocer la imagen original. Sin embargo, con un algoritmo de aprendizaje se puede obtener una aproximación a dicha imagen original en base a un conjunto grande de entrenamiento.

Otro aspecto interesante es la forma del parámetro λ estudiado en este trabajo: en la primera sección se planteó la posibilidad de ampliar el modelo para funciones en general (sean polinomios, ecuaciones trigonométricas, etc), sin embargo esta vez escogimos, a ma-

nera de ejemplo, trabajar con polinomios cuadráticos de coeficientes positivos como paso inicial para comprender el comportamiento del algoritmo de mejoramiento de imágenes, sin embargo el método aquí expuesto se puede usar como punto de partida para el estudio de distintos casos en los que el parámetro λ sea modelado por distintos tipos de funciones.

Referencias

[1] E. Fatemi L. Rudin, S. Osher. *Nonlinear total variation based noise removal algorithms*, volume 60, pages: 259-268. Physica D. Journal, 1992.

[2] S. Kumar B. Satija. *Indian Journal of Radiology and Imaging*, volume 23 Issue 4, pp 285-398. Publimed, 2014.

[3] Rodrigo Farias Rezino. *Contagem de Pessoas em Tempo Real Utilizando Técnicas de Computação Gráfica*. Universidade Norte do Paraná, Centro de Ciencias Exatas e Tecnológicas, Engenharia da Computação, 1993.

[4] L. Rudin. *Images, numerical analysis of singularities and shock filters*. Caltech, C.S. Department, California Institute of Technology, 1987. Cap. 1, pp 24-25.

[5] P. Perona and J. Malik. *Scale-space and edge detection using anisotropic diffusion*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12, 1990, pages: 629–639.

[6] G. Michaille H. Attouch, G. Butazzo. *Variational Analysis in Sobolev and BV Spaces*. Siam, 2006, Cap. 10, pp 371-406.

[7] C.B Schönlieb JC De Los Reyes. *Image Denoising: Learning Noise Distribution via PDE-Constrained Optimization*. to appear in Inverse Problems and Imaging, DAMTP-Technical Report, 2014.

[8] M. Novaga V. Caselles, A. Chambolle. *Regularity for solutions of the total variation denoising problem*, volume 27, pp 232–252. Revista Matemática Iberoamericanal.

[9] V. Caselles. *Total Variation Based Image Denoising and Restoration*. European Mathematican Society, 2006.

[10] C. Sbert J. Duran, B. Coll. *Chambolle's Projection Algorithm for Total Variation Denoising*. Imagen Processing On Line, 2013.

[11] William Faris. *Real Analysis: Part II*. Notas de Curso, 2004, cap: 1.

[12] S. I. Yoo S. Son, D. Kang. *Non-local Huber Regularization for Image Denoising: A Hybrid Approach of Two Non-local Regularizations*, volume ICPRAM, pp 554-559. SciTePress, 2013.

[13] Peter J Huber. *Robust Estimation of a Location Parameter*. Annals of Statistics 53, 1964, pp 73-102.

[14] J.C. De los Reyes. *Optimal control of a class of variational inequalities of the second kind*, volume 49, pp 1629-1658. SIAM Journal on Control and Optimization, 2011.

[15] N.C. Wong J.C. Yao B.T. Kien, M. Wong. *Solution Existence of Variational Inequalities with Pseudomonotone Operators in the Sense of Brézis*, volume 140, pp 249-263. Journal of Optimization Theory and Applications, 2009.

[16] M. Hintermüller and K. Kunisch. *Total bounded variation regularization as a bilaterally constrained optimization problem*, volume 64. SIAM Journal on Applied Mathematics, 2004.

[17] Michael Hintermüller J.C. De los Reyes. *A duality based semismooth Newton framework for solving variational inequalities of the second kind*, volume 13, pp 671-692. Interfaces and Free Boundaries, 2011.

[18] R Tenam I Ekeland. *Convex Analysis and Variational Problems*. Classics in Appl. Math. 28, 1999.

[19] J. Zowe H. Maurer. *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, volume 16, pp 98-110. Mathematical Programming, 1979.

[20] Juan Carlos De los Reyes. *Constrained Optimal Control of Stationary Viscous Incompressible Fluids by Primal-Dual Active Set Methods*. Ph.D. Thesis, Karl-Franzens University of Graz, 2003.

[21] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.

[22] John B Conway. *A Course in Functional Analysis*. Springer, 1990.

[23] JC De Los Reyes. *Análisis Numérico para la resolución de ecuaciones diferenciales parciales*. Unidad de publicaciones de la Escuela Politécnica Nacional, 2012.

[24] C. Carstensen J. Alberty and S. A. Funken. *Remarks around 50 lines of Matlab: short finite element implementation*, volume 20, pp 117–137. Numerical Algorithms 20, 1999.

[25] Juan Carlos de los Reyes. *Optimization of mixed variational inequalities arising in flow of viscoplastic materials*, volume Comput Optim Appl (2012) 52:757–784. Springer Science+Business Media, LLC, 2011.

[26] JC De Los Reyes. *Numerical Methods for PDE-constrained Optimization*. Unidad de publicaciones de la Escuela Politécnica Nacional, 2012.

[27] Andrea Braides. *Introduction to Homogenization and Gamma Convergence*. School on Homogenization at ICTP Trieste, 1993. lectures abstract.

[28] G. Dal Maso. *An Introduction to Gamma-Convergence*. Birkhäuser, Boston, 1993.

[29] Heinz H. Bauschke and Yves Lucet. *What is a Fenchel Conjugate?*, volume 59 Number 1, pp 44-46. Notices of the AMS, 2012.

ANEXO A

Algunos Resultados sobre Relajación, Γ-convergencia y Conjugada de Fenchel

A.0.1 Relajación de un Problema

Las siguientes definiciones y resultados son necesarias para entender el por qué de algunos procedimientos vistos en la sección 1.3 referentes a la Relajación del Problema de Control de Imágenes en el espacio de Variación Acotada y en H_0^1.

Los resultados fueron tomados de [27] y [28], donde se pueden estudiar a mayor detalle.

Definición A.1 (Envoltura Semicontinua Inferior). *Para toda función $F : X \to \bar{\mathbb{R}}$, la Envoltura Semicontinua Inferior (o Función Relajada) de F se define como:*

$$sc^- F(x) = \sup_{G \in \mathcal{G}(F)} G(x)$$

donde $\mathcal{G}(F)$ es el conjunto de todas las funciones G semicontinuas inferiormente en X tales que:

$$G(y) \leq F(y) \quad \forall y \in X$$

Observación A.1. *$sc^- F : X \to \bar{\mathbb{R}}$ es semicontinua inferiormente en X. Por definición, $sc^- F \leq$ y $sc^- F \geq G$ para toda función G semicontinua inferiomente tal que $G \leq F$. Luego, $sc^- F$ es la función semicontinua inferiormente más grande mayorada por F.*

Ejemplo A.1. *Sea E un subconjunto de X y $\chi_E : X \to \{0, 1\}$ la función Indicatriz, definida como:*

$$\chi_E(x) := \begin{cases} 1 & si\ x \in A \\ 0 & si\ x \notin A \end{cases}$$

Luego, $sc^- \chi_E = \chi_{\bar{E}}$ donde $\bar{E}$ es la clausura de E en X.

Teorema A.1 (Teorema 3,8 en [28]). *Sea $F : X \to \bar{\mathbb{R}}$ coerciva, entonces se cumplen las siguientes propiedades:*

- *$sc^- F$ es coerciva y semicontinua inferiormente*
- *$sc^- F$ tiene un mínimo en X*

- $\min_{x\in X}(sc^{-}F)(x) = \inf_{x\in X} F(x)$
- *Todo punto de acumulación de una sucesión minimizante de F es un mínimo de $sc^{-}F$ en X*
- *Si X satisface el primer axioma de contabilidad, entonces cada punto mínimo para $sc^{-}F$ es el límite para una sucesión minimizante para F en X*

Dado que todo espacio métrico cumple el primer axioma de contabilidad, se puede decir que virtualmente la existencia de cualquier problema de minimización, cuya demostración nos remite a encontrar el límite de una sucesión minimizante, puede ser resuelta a través de su envoltura semicontinua inferior.

A este procedimiento se le llama "Relajación de un Problema"

A.0.2 Γ-convergencia

Definición A.2 (Γ-convergencia). *Sea X un espacio topológico, y $F_n : X \to [0, +\infty)$ una sucesión de funcionales en X. Luego, F_n se dice que Γ-converge al Γ-límite $F : X \to [0, +\infty)$ si se cumplen las siguientes condiciones:*

- *Desigualdad de acotación inferior: Para toda sucesión $(x_n) \in X$ tal que $x_n \to x$ cuando $n \to \infty$,*
$$F(x) \leq \liminf_{n\to\infty} F_n(x_n)$$
- *Desigualdad de acotación superior: Para todo $x \in X$, existe una sucesión (x_n) que converge a x tal que*
$$F(x) \geq \limsup_{n\to\infty} F_n(x_n)$$

La primera condición significa que F es una cota inferior asintótica para todos los F_n, y gracias a la segunda condición tenemos que esta cota inferior es óptima.

Teorema A.2 (Proposición 5,7 en [28]). *Si (F_h) es una sucesión decreciente que converge a F puntualmente, entonces (F_h) Γ-converge a la envoltura semicontinua inferior de F.*

A.1 La conjudada de Fenchel

En la sección 1.4 se estudia el modelo de control de imágenes en el espacio de Hilbert H_0^1. Como se explica en detalle en [14], es necesaria una regularización de conjuntos activo e inactivo, para poder caracterizar el problema a través de un sistema de optimalidad adecuado y obtener una aproximación numérica del tipo Newton Semismooth. Dicha regularización es necesaria debido a un inconveniente presentado con el problema dual y, más específicamente, con la propiedad de dualidad de Fenchel [14] [29], que se introducirá a continuación.

A.1.1 Definición

La conjugada de Fenchel es, en análisis convexo, la contraparte de la transformada de Fourier, y es vista como una generalización de la transformada de Legendre [29]

Definición A.3 (Conjugada de Fenchel). *Sea X un espacio de Hilbert real y $f : x \to]-\infty, +\infty[$ una función propia (es decir que $domf = \{x \in X / f(x) \in \mathbb{R}\} \neq 0$). La conjugada de Fenchel f^* de f en $u \in X$ se define como:*

$$f^*(u) = \sup_{x \in X} (\langle x, u \rangle - f(x))$$

Una consecuencia inmediata de esta definición es la desigualdad de Fenchel-Young:

$$f(x) + f^*(u) \geq \langle x, u \rangle \tag{A.1}$$

Además, f^* es convexa y semicontinua inferiormente por ser el supremo de la familia de funciones afines continuas: $(\langle x, \cdot \rangle - f(x))_{x \in X}$.

Por lo que, podemos concluir la siguiente equivalencia:

$$f(x) = f^{**}(x) \Leftrightarrow \begin{cases} \text{f es convexa} \\ \text{y semicontinua inferiormente} \end{cases}$$

A.1.2 Ejemplos

Ejemplo A.2. $(\alpha\ exp)^*(u) = \begin{cases} +\infty & si \quad u < 0 \\ 0 & si \quad u = 0 \\ u \ln(u/\alpha) - u & si \quad u > 0 \end{cases}$

Ejemplo A.3. *Sea $1 < p < +\infty$:*

$$\left(\frac{1}{p}|\cdot|^p\right)^* = \frac{1}{q}|\cdot|^q$$

donde $\frac{1}{p} + \frac{1}{q} = 1$

A.1.3 Caracterización de subgradientes

El dominio de la Conjugada de Fenchel es Γ_X: el cono de funciones convexas, semicontinuas inferiormente y propias en X.

A la conjugada de Fenchel se encuentra fuertemente ligada la definición de operador subdiferencial ∂f:

Definición A.4 (Operador Subdiferencial). *El Operador Subdiferencial ∂f es el conjunto de valores de una aplicación en X llamados Subgradientes. Esto es:*

$$u \in \partial f \Leftrightarrow (\forall h \in X): \qquad f(x) + \langle h, u \rangle \leq f(x + h)$$

Así, la desigualdad (A.1) caracteriza subgradientes (elementos del subdiferencial):

$$f(x) + f^\star(u) = \langle x, u\rangle \Leftrightarrow u \in \partial f(x)$$
$$\Leftrightarrow x \in \partial f^\star(u)$$

Así, cuando f es continua, se tiene sencillamente que: $\partial f(x) = \{\nabla f(x)\}$

El operador subdiferencial es una manera de derivada generalizada que además conserva la propiedad crítica de la optimización:

$$0 \in \partial f(x) \Leftrightarrow x \text{ es un minimizador global de } f$$

A.1.4 Teorema de la Dualidad de Fenchel

El Teorema de la dualidad de Fenchel es un resultado que envuelve a problemas de optimización en funciones convexas. El problema de Fenchel nos dice que ambos problemas, el primal y el dual, tienen la misma solución bajo ciertas condiciones.

Obviamente, esto posibilita la solución de muchos problemas de optimización a través de su problema dual.

Teorema A.3 (Teorema de la dualidad de Fenchel). *Sean X, Y espacios de Banach, $f : X \to \mathbb{R} \cup \{+\infty\}$ y $g : Y \to \mathbb{R} \cup \{+\infty\}$ funciones convexas y $A : X \to Y$ una aplicación lineal acotada, luego los problemas:*

$$p^\star = \inf_{x \in X} \{f(x) + g(Ax)\} \tag{P}$$

$$d^\star = \sup_{y \in Y} \{-f^\star(A^\star y^\star) - g^\star(-y^\star)\} \tag{D}$$

satisfacen $p^\star \geq d^\star$ donde $f^\star$ y $g^\star$ son conjugadas convexas de f y g respectivamente, y $A^\star$ es el operador adjunto.

Si además f, g y A cumplen una de las siguientes propiedades:

- *f y g son semicontinuas inferiormente y $0 \in \ker(dom\ g - Adom\ f)$, donde $dom\ f = \{z : f(z) < \infty\}$ o,*
- *$A\ dom\ f + cont\ g \neq \varnothing$ donde cont es el conjunto de los puntos donde la función es continua.*

Entonces, $p^\star = d^\star$

ANEXO B

Códigos en Matlab

B.1 Códigos para el problema (2.17) correspondiente al caso de un parámetro de ruido Real

El primer codigo, resuelve el algoritmo 2 para encontrar la solución de la ecuación de estado del problema (2.17)

```
function U=NSS(n,lamb,gamma,eeps)   %Programa que resuelve la ecuación de estado
%para un mallado de n puntos intermedios, un valor lambda dado,
%y los parámetros gamma y epsilon escogidos.
[B A]=matricesmr(n);            %Programa que construye las matrices de masa (B)
%y rigidez(A), dado un mallado con n puntos intermedios
[Q1 Q2 D1 D2]=matricesqg(n);     %Programa que devuelve las matrices A2
%y la discertización del gradiente en dif. finitas
 %Archivo en el que se encuentra almacenada la imagen (original y con
 %ruido):
storedstructure=load('imagingdata.mat');
U1=storedstructure.noised;
Uo=storedstructure.original;
%A continuación, tomamos los puntos intermedios de la imagen, dado el
%mallado n:
nn=int32((n+1)/2);
U0=U1(88-nn + rem(n+1,2):88+nn,88-nn + rem(n+1,2):88+nn);   %imagen con ruido
Io=Uo(88-nn+ rem(n+1,2):88+nn,88-nn+ rem(n+1,2):88+nn); %imagen original
%Reordenamos los datos obtenidos hasta el momento:
fo=reshape(Io,(n+2)^2,1);
f=reshape(U0,(n+2)^2,1);
Q=[Q2 Q1];
G=[D2;D1];
%%%%%%%%%%%Datos de inicialización:
t=2*((n+1)^2);   %numero de triangulos
q=sparse(2*t,1);  %vector q
```

```
y=sparse((n+2)^2,1); %vector solución y
dy=sparse((n+2)^2,1); %dirección de la iteración del método NSS
dq=sparse(2*t,1); %dirección de q
cont=0; %contador de iteraciones
norma=1;
%Inicio de las iteraciones del método NSS
while norma>1e-4 && cont<12
cont=cont+1;
Gy=(G*y);
nGy=xi(Gy);  %norma del gradiente de y
%Definimos los conjuntos activo, activo débil e inactivo:
act1=gamma*nGy-1-(1/(2*gamma));
act=spones(max(0,act1));  %Conjunto activo
I1=spdiags(act,0,2*t,2*t);
act3=1-(1/(2*gamma))-gamma*nGy;
Act3=spones(max(0,act3)); %conjunto inactivo
I3=spdiags(Act3,0,2*t,2*t);
I2=speye(2*t)-I1-I3; %conjunto activo débil
%Construcción de las matrices P=Dy/|Dy| y Dmi=1/max(1, gamma |Dy|) :
g1=spones(Gy);
g2=1-g1;
g3=(1/2)*g2;
Gy=Gy+g3;
n1=spones(nGy);
n2=1-n1;
nGy=nGy+n2;
Gy1=spdiags(Gy(1:t),0,t,t);
Gy2=spdiags(Gy(t+1:2*t),0,t,t);
R=[kron(ones(2,1),Gy1) kron(ones(2,1),Gy2)];
Nin=spdiags(1./nGy,0,2*t,2*t);
P=Nin*R;
mk=max(1,gamma*nGy);
Dm=spdiags(mk,0,2*t,2*t);
Dmi=spdiags(1./mk,0,2*t,2*t);
%Construcción de la matriz Dqp=q/max(1,|q|):
nqp=xi(q);
mqk=max(1,nqp);
Dmqk=spdiags(mqk,0,2*t,2*t);
Dmqkin=inv(Dmqk);
qp=Dmqkin*q;
Dqp=spdiags(qp,0,2*t,2*t);
Vgy=spdiags(Gy,0,2*t,2*t);
```

```
%%%%%%%%%Definimos h para el conjunto activo débil:
if max(max(I2))>0.2
be=1-(gamma*nGy)+(1/(2*gamma));
Be=spdiags(be,0,2*t,2*t);
R2= Nin*(speye(2*t)-(gamma/2)*(Be.^2)) - (Nin * R' * (Nin*Dqp + R*gamma^2 * Be));
else
    Be=sparse(2*t,2*t);
    R2=sparse(2*t,2*t);
end
%Resolución de la ecuación de incremento:
Hess=[((eeps*A)+(lamb*B)) Q
    -(I1*gamma*(speye(2*t,2*t) - (Nin * R'*Dqp))+I3*gamma+I2*Dm*R2)*G Dm];
eta2=[-((eeps*A)+(lamb*B))*y-Q*q+(lamb*B*f)
   -Dm*q+gamma*Gy];
dx=Hess\eta2;
dy=dx(1:(n+2)^2); %tamaño del paso de la iteración del NSS
dq=dx((n+2)^2+1:size(dx,1));
y=y+dy;
q=q+dq;
norma=norm(dy,2);
end
U=reshape(y,n+2,n+2); %Solución
end
```

El siguiente programa, resuelve la ecuación adjunta del problema (2.17)

```
function p=ecadj4(n,U,lambda,gamma,eeps) %Programa que calcula la solución de
%la ecuación adjunta, dado el número de puntos del mallado, la solución
%de la ecuación de estado, el parámetro lambda y las cantidades gamma y epsilon
[B A]=matricesmr(n);             %Programa que construye las matrices de masa (B)
%y rigidez(A), dado un mallado con n puntos intermedios
[Q1 Q2 D1 D2]=matricesqg(n);     %Programa que devuelve las matrices A2
%y la discertización del gradiente en dif. finitas
 %Archivo en el que se encuentra almacenada la imagen (original y con
 %ruido):
storedstructure=load('imagingdata.mat');
U1=storedstructure.noised;
Uo=storedstructure.original;
%A continuación, tomamos los puntos intermedios de la imagen, dado el
%mallado n:
nn=int32((n+1)/2);
U0=U1(88-nn + rem(n+1,2):88+nn,88-nn + rem(n+1,2):88+nn);   %imagen con ruido
Io=Uo(88-nn+ rem(n+1,2):88+nn,88-nn+ rem(n+1,2):88+nn); %imagen original
```

```
%Reordenamos los datos obtenidos hasta el momento:
fo=reshape(Io,(n+2)^2,1);
f=reshape(U0,(n+2)^2,1);
Q=[Q2 Q1];
G=[D2;D1];
t=2*((n+1)^2);   %numero de triangulos
q=sparse(2*t,1);  %vector q
y=sparse((n+2)^2,1); %vector solución y
dy=sparse((n+2)^2,1); %dirección de la iteración del método NSS
dq=sparse(2*t,1); %dirección de q
Gy=(G*y);
nGy=xi(Gy);  %norma del gradiente de y
%Definimos los conjuntos activo, activo débil e inactivo:
act1=gamma*nGy-1-(1/(2*gamma));
act=spones(max(0,act1));  %Conjunto activo
I1=spdiags(act,0,2*t,2*t);
act3=1-(1/(2*gamma))-gamma*nGy;
Act3=spones(max(0,act3)); %conjunto inactivo
I3=spdiags(Act3,0,2*t,2*t);
I2=speye(2*t)-I1-I3; %conjunto activo débil
%Construcción de las matrices P=Dy/|Dy| y Dmi=1/max(1, gamma |Dy|) :
g1=spones(Gy);
g2=1-g1;
g3=(1/2)*g2;
Gy=Gy+g3;
n1=spones(nGy);
n2=1-n1;
nGy=nGy+n2;
Gy1=spdiags(Gy(1:t),0,t,t);
Gy2=spdiags(Gy(t+1:2*t),0,t,t);
R=[kron(ones(2,1),Gy1) kron(ones(2,1),Gy2)];
Nin=spdiags(1./nGy,0,2*t,2*t);
P=Nin*R;
mk=max(1,gamma*nGy);
Dm=spdiags(mk,0,2*t,2*t);
Dmi=spdiags(1./mk,0,2*t,2*t);
%Construcción de la matriz Dqp=q/max(1,|q|):
nqp=xi(q);
mqk=max(1,nqp);
Dmqk=spdiags(mqk,0,2*t,2*t);
Dmqkin=inv(Dmqk);
qp=Dmqkin*q;
```

```
Dqp=spdiags(qp,0,2*t,2*t);
Vgy=spdiags(Gy,0,2*t,2*t);
%%%%%%%%Definimos h para el conjunto activo débil:
if max(max(I2))>0.2
be=1-(gamma*nGy)+(1/(2*gamma));
Be=spdiags(be,0,2*t,2*t);
R2= Nin*(speye(2*t)-(gamma/2)*(Be.^2)) - (Nin * R' * (Nin*Dqp + R*gamma^2 * Be));
else
    Be=sparse(2*t,2*t);
    R2=sparse(2*t,2*t);
end
%Resolución de la ecuación lineal:
 Hess=[((eeps*A)+(lambda*B)) Q
     -(I1*gamma*(speye(2*t,2*t) - (Nin * R'*Dqp))+I3*gamma+I2*Dm*R2)*G Dm];
eta2=[-B*(reshape(U,(n+2)^2,1)-fo)
    sparse(2*t,1)];
x=Hess\eta2;
pd=x(1:(n+2)^2);
p=reshape(pd,n+2,n+2); %Solución
end
```

Finalmente, se tiene el codigo de resolución del algoritmo 3 de un método cuasi Newton, con la regla de armijo para encontrar el parámetro óptimo real del problema (2.17)

```
function l=gradientdes15(n,gamma,eeps)
betha=1e-10; %parámetro betha
lambda=1; %valor de lambda con el que se inicia el algoritmo
b=[lambda;0]; %vector que almacena el lambda actual y el "anterior", en este caso 0
it=0; %contador de iteraciones
norma1=1;
H=1; %inicialización de la matriz DFP
%Archivo en el que se encuentra almacenada la imagen (original y con ruido):
storedstructure=load('imagingdata.mat');
U1=storedstructure.noised;
Uo=storedstructure.original;
%A continuación, tomamos los puntos intermedios de la imagen, dado el
%mallado n:
nn=int32((n+1)/2);
U0=U1(88-nn + rem(n+1,2):88+nn,88-nn + rem(n+1,2):88+nn);   %imagen con ruido
Io=Uo(88-nn+ rem(n+1,2):88+nn,88-nn+ rem(n+1,2):88+nn); %imagen original
%Reordenamos los datos obtenidos hasta el momento:
fo=reshape(Io,(n+2)^2,1); %imagen original
f=reshape(U0,(n+2)^2,1); %imagen con ruido
```

```
U=NSS(n,lambda,gamma,eeps); %Primer resultado de la ecuación de estado
eca=ecadj4(n,U,lambda,gamma,eeps); %Primer resultado del problema adjunto
in=trapezoidal1((U-U0)*eca);
Fp=(2*betha*lambda+in); %Cálculo del gradiente del funcional de costo
alpha=1; %Inicialización del tamaño de la dirección de descenso
s=-H*Fp; %dirección de descenso
V=[Fp;0]; %vector que almacena el gradiente del funcional actual y el "anterior"
%en este caso 0
v=(V(1)-V(2)); %diferencia entre gradiente del funcional de costo actual
%y el anterior
%Inicio del método cuasi Newton con fórmula DFP:
while norma1*alpha>1e-8|| it<5
%datos iniciales para el algoritmo de búsqueda lineal
nlambda=lambda;
cost1=funcionaldecosto2(U,betha,lambda); %Cálculo del funcional de costo
armijo=100;
alpha=1;
p=-H*Fp; %dirección de descenso
countarm=0;
V1=V;
H1=H;
%Algoritmo de búsqueda lineal (Fórmula de Armijo):
while armijo >-alpha*(1e-4)*(H1*Fp^2) && alpha>1e-3
alpha=1/2^(countarm); %incremento del parámetro para la dirección de descenso
  nlambda=max(nlambda+alpha*-H1*Fp,0); %prueba el resultado con el alpha
%de la iteración de la búsqueda lineal actual
s1=alpha*-H1*Fp; %dirección de descenso con el alpha actual
U=NSS(n,nlambda,gamma,eeps); %resultado de la ecuación de estado
eca=ecadj4(n,U,nlambda,gamma,eeps); %resultado del problema adjunto
in=trapezoidal1((U-U0)*eca);
Fp=(2*betha*lambda+in); %Cálculo del gradiente del funcional,
%tomando en cuenta el parámetro alfa actual
V1=[Fp; V(1)]; %vector que almacena el valor del gradiente en la iteracion k+1
%y en k
v1=(V1(1)-V1(2)); %diferencia del Gradiente
H1=s1/v1; %Cálculo de la fórmula DFP
  cost2=funcionaldecosto2(U,betha,nlambda); %Cálculo del funcional de costo
 armijo=cost2-cost1;
countarm=countarm+1;
end
cost2 %Funcional de costo, luego de haber obtenido un alfa que satisfaga
%la regla de Armijo
```

```
lambda=max(lambda+alpha*p,0) %Cálculo de lambda en la iteración actual
s=alpha*p;
U=NSS(n,lambda,gamma,eeps); %Resultado de la ecuación de estado con
%los datos de la iteración actual
eca=ecadj4(n,U,lambda,gamma,eeps); %Resultado del problema adjunto
%con los datos de la iteración actual
in=trapezoidal1((U-U0)*eca);
Fp=(2*betha*lambda+in); %Gradiente del funcional de costo
V=[Fp; V(1)]; %vector que almacena el valor del gradiente en la iteracion k+1 y en k
v=(V(1)-V(2));%diferencia del Gradiente
norma1=norm(Fp)
H=s/v; %Cálculo de la fórmula DFP
it=it+1
end
l=lambda %resultado
%Cómputo de la solución en la ecuación de estado y gráficos
Uf=full(NSS(n,l,gamma,eeps));
figure(1)
imshow(Uf,[])
figure(2)
imshow(U0,[])
figure(3)
imshow(Io,[])
end
```

B.2 Códigos para el problema (2.33) correspondiente al caso de un parámetro de ruido polinomial

El primer codigo, resuelve el algoritmo 2 para encontrar la solución de la ecuación de estado del problema (2.33)

```
function U=NSSp(n,lambda,gamma,eeps)%Programa que resuelve la ecuación de estado
%para un mallado de n puntos intermedios, el vector lambda, de los coeficientes
%del polinomio cuadrático A,B,C,D,E,F, dados, y los parámetros gamma y epsilon.
[B A]=matricesmr(n); %Programa que construye las matrices de masa (B)
%y rigidez(A), dado un mallado con n puntos intermedios
[Q1 Q2 D1 D2]=matricesqg(n); %Programa que devuelve las matrices A2 y la
%discertización del gradiente en dif. finitas
 %Archivo en el que se encuentra almacenada la imagen (original y con
 %ruido):
storedstructure=load('imagingdata.mat');
U1=storedstructure.noised3;
```

```
Uo=storedstructure.original;
%A continuación, tomamos los puntos intermedios de la imagen, dado el
%mallado n:
nn=int32((n+1)/2);
U0=U1(88-nn + rem(n+1,2):88+nn,88-nn + rem(n+1,2):88+nn);   %imagen con ruido
Io=Uo(88-nn+ rem(n+1,2):88+nn,88-nn+ rem(n+1,2):88+nn); %imagen original
%Reordenamos los datos obtenidos hasta el momento:
fo=reshape(Io,(n+2)^2,1);
f=reshape(U0,(n+2)^2,1);
Q=[Q2 Q1];
G=[D2;D1];

%%%%%%%%%%%Datos de inicialización:
t=2*((n+1)^2);   %numero de triangulos
q=sparse(2*t,1);  %vector q
y=sparse((n+2)^2,1); %vector solución y
dy=sparse((n+2)^2,1); %dirección de la iteración del método NSS
dq=sparse(2*t,1); %dirección de q
cont=0; %contador de iteraciones
norma=1;
%Construcción de la matriz con los valores del polinomio lambda en cada
%punto del mallado:
h=1/(1+n);
l1=lambda(1); %Parámetro A en el polinomio cuadrático lambda
l2=lambda(2); %Parámetro B en el polinomio cuadrático lambda
l3=lambda(3); %Parámetro C en el polinomio cuadrático lambda
l4=lambda(4);%Parámetro D en el polinomio cuadrático lambda
l5=lambda(5); %Parámetro E en el polinomio cuadrático lambda
l6=lambda(6); %Parámetro F en el polinomio cuadrático lambda
xv=0:h:1;
yv=0:h:1;
sz=length(xv);
for i=1:sz
    for j=1:sz
 lamb1(i,j)=l1*xv(i)^2 + l2*yv(j)^2 + l3*xv(i)*yv(j) + l4*xv(i) + l5*yv(j) + l6;
    end
end
lamb=spdiags(reshape(lamb1',sz^2,1),0,sz^2,sz^2);
%Inicio de las iteraciones del método NSS
while norma>1e-5 && cont<10
cont=cont+1;
Gy=(G*y);
```

```
nGy=xi(Gy);  %norma del gradiente de y
%Definimos los conjuntos activo, activo débil e inactivo:
act1=gamma*nGy-1-(1/(2*gamma));
act=spones(max(0,act1));  %Conjunto activo
I1=spdiags(act,0,2*t,2*t);
act3=1-(1/(2*gamma))-gamma*nGy;
Act3=spones(max(0,act3)); %conjunto inactivo
I3=spdiags(Act3,0,2*t,2*t);
I2=speye(2*t)-I1-I3; %conjunto activo débil
%Construcción de las matrices P=Dy/|Dy| y Dmi=1/max(1, gamma |Dy|) :
g1=spones(Gy);
g2=1-g1;
g3=(1/2)*g2;
Gy=Gy+g3;
n1=spones(nGy);
n2=1-n1;
nGy=nGy+n2;
Gy1=spdiags(Gy(1:t),0,t,t);
Gy2=spdiags(Gy(t+1:2*t),0,t,t);
R=[kron(ones(2,1),Gy1) kron(ones(2,1),Gy2)];
Nin=spdiags(1./nGy,0,2*t,2*t);
P=Nin*R;
mk=max(1,gamma*nGy);
Dm=spdiags(mk,0,2*t,2*t);
Dmi=spdiags(1./mk,0,2*t,2*t);
%Construcción de la matriz Dqp=q/max(1,|q|):
nqp=xi(q);
mqk=max(1,nqp);
Dmqk=spdiags(mqk,0,2*t,2*t);
Dmqkin=inv(Dmqk);
qp=Dmqkin*q;
Dqp=spdiags(qp,0,2*t,2*t);
Vgy=spdiags(Gy,0,2*t,2*t);
%%%%%%%%Definimos h para el conjunto activo débil:
if max(max(I2))>0.2
be=1-(gamma*nGy)+(1/(2*gamma));
Be=spdiags(be,0,2*t,2*t);
R2= Nin*(speye(2*t)-(gamma/2)*(Be.^2)) - (Nin * R' * (Nin*Dqp + R*gamma^2 * Be));
else
    Be=sparse(2*t,2*t);
    R2=sparse(2*t,2*t);
end
```

```
Mm=(lamb.*B); %Producto puntual de la matriz de masa por el polinomio lambda
%Resolución de la ecuación de incremento:
Hess=[((eeps*A)+Mm) Q
    -(I1*gamma*(speye(2*t,2*t) - (Nin * R'*Dqp))+I3*gamma+I2*Dm*R2)*G Dm];
eta2=[-((eeps*A)+Mm)*y-Q*q+(Mm*f)
   -Dm*q+gamma*Gy];
dx=Hess\eta2;
dy=dx(1:(n+2)^2);  %tamaño del paso de la iteración del NSS
dq=dx((n+2)^2+1:size(dx,1));
y=y+dy;
q=q+dq;
norma=norm(dy,2);
end
U=reshape(y,n+2,n+2);  %Solución
end
```

El siguiente programa, resuelve la ecuación adjunta del problema (2.33)

```
function p=ADJp(n,U,lambda,gamma,eeps) %Programa que calcula la solución de
%la ecuación adjunta, dado el número de puntos del mallado,
%la solución de la ecuación de estado, el parámetro lambda y las cantidades gamma y epsilon
[B A]=matricesmr(n);             %Programa que construye las matrices de masa (B)
% y rigidez(A), dado un mallado con n puntos intermedios
[Q1 Q2 D1 D2]=matricesqg(n);     %Programa que devuelve las matrices A2
%y la discertización del gradiente en dif. finitas
 %Archivo en el que se encuentra almacenada la imagen (original y con
 %ruido):
storedstructure=load('imagingdata.mat');
U1=storedstructure.noised;
Uo=storedstructure.original;
%A continuación, tomamos los puntos intermedios de la imagen, dado el
%mallado n:
nn=int32((n+1)/2);
U0=U1(88-nn + rem(n+1,2):88+nn,88-nn + rem(n+1,2):88+nn);   %imagen con ruido
Io=Uo(88-nn+ rem(n+1,2):88+nn,88-nn+ rem(n+1,2):88+nn); %imagen original
%Reordenamos los datos obtenidos hasta el momento:
fo=reshape(Io,(n+2)^2,1);
f=reshape(U0,(n+2)^2,1);
Q=[Q2 Q1];
G=[D2;D1];
%Construcción de la matriz con los valores del polinomio lambda en cada
%punto del mallado:
h=1/(1+n);
```

```
l1=lambda(1); %Parámetro A en el polinomio cuadrático lambda
l2=lambda(2); %Parámetro B en el polinomio cuadrático lambda
l3=lambda(3); %Parámetro C en el polinomio cuadrático lambda
l4=lambda(4);%Parámetro D en el polinomio cuadrático lambda
l5=lambda(5); %Parámetro E en el polinomio cuadrático lambda
l6=lambda(6); %Parámetro F en el polinomio cuadrático lambda
xv=0:h:1;
yv=0:h:1;
sz=length(xv);
for i=1:sz
    for j=1:sz
lamb1(i,j)=l1*xv(i)^2 + l2*yv(j)^2 + l3*xv(i)*yv(j) + l4*xv(i) + l5*yv(j) + l6;
    end
end
lambdav=spdiags(reshape(lamb1',sz^2,1),0,sz^2,sz^2);
t=2*((n+1)^2);   %numero de triangulos
q=sparse(2*t,1);  %vector q
y=sparse((n+2)^2,1); %vector solución y
dy=sparse((n+2)^2,1); %dirección de la iteración del método NSS
dq=sparse(2*t,1); %dirección de q
Gy=(G*y);
nGy=xi(Gy);  %norma del gradiente de y
%Definimos los conjuntos activo, activo débil e inactivo:
act1=gamma*nGy-1-(1/(2*gamma));
act=spones(max(0,act1));  %Conjunto activo
I1=spdiags(act,0,2*t,2*t);
act3=1-(1/(2*gamma))-gamma*nGy;
Act3=spones(max(0,act3)); %conjunto inactivo
I3=spdiags(Act3,0,2*t,2*t);
I2=speye(2*t)-I1-I3; %conjunto activo débil
%Construcción de las matrices P=Dy/|Dy| y Dmi=1/max(1, gamma |Dy|) :
g1=spones(Gy);
g2=1-g1;
g3=(1/2)*g2;
Gy=Gy+g3;
n1=spones(nGy);
n2=1-n1;
nGy=nGy+n2;
Gy1=spdiags(Gy(1:t),0,t,t);
Gy2=spdiags(Gy(t+1:2*t),0,t,t);
R=[kron(ones(2,1),Gy1) kron(ones(2,1),Gy2)];
Nin=spdiags(1./nGy,0,2*t,2*t);
```

```
P=Nin*R;
mk=max(1,gamma*nGy);
Dm=spdiags(mk,0,2*t,2*t);
Dmi=spdiags(1./mk,0,2*t,2*t);
%Construcción de la matriz Dqp=q/max(1,|q|):
nqp=xi(q);
mqk=max(1,nqp);
Dmqk=spdiags(mqk,0,2*t,2*t);
Dmqkin=inv(Dmqk);
qp=Dmqkin*q;
Dqp=spdiags(qp,0,2*t,2*t);
Vgy=spdiags(Gy,0,2*t,2*t);
%%%%%%%%Definimos h para el conjunto activo débil:
if max(max(I2))>0.2
be=1-(gamma*nGy)+(1/(2*gamma));
Be=spdiags(be,0,2*t,2*t);
R2= Nin*(speye(2*t)-(gamma/2)*(Be.^2)) - (Nin * R' * (Nin*Dqp + R*gamma^2 * Be));
else
    Be=sparse(2*t,2*t);
    R2=sparse(2*t,2*t);
end
Mm=(lambdav.*B); %Producto puntual de la matriz de masa por el polinomio lambda
%Resolución de la ecuación lineal:
 Hess=[((eeps*A)+(Mm)) Q
     -(I1*gamma*(speye(2*t,2*t) - (Nin * R'*Dqp))+I3*gamma+I2*Dm*R2)*G Dm];
eta2=[-B*(reshape(U,(n+2)^2,1)-fo)
    sparse(2*t,1)];
x=Hess\eta2;
pd=x(1:(n+2)^2);
p=reshape(pd,n+2,n+2); %solución
end
```

Finalmente, se tiene el codigo de resolución del algoritmo 4 de un método cuasi Newton, con la regla de armijo para encontrar el parámetro óptimo real del problema (2.33)

```
function l=quasinewP(n,gamma,eeps)
betha=1e-10;%parámetro betha
lambda=[1; 1; 1; 1; 1; 1]; %vector lambda con el que se inicia el algoritmo
it=0; %contador de iteraciones
norma1=1;
H=speye(6);%inicialización de la matriz DFP
%Archivo en el que se encuentra almacenada la imagen (original y con
%ruido):
```

```
storedstructure=load('imagingdata.mat');
U1=storedstructure.noised3;
Uo=storedstructure.original;
%A continuación, tomamos los puntos intermedios de la imagen, dado el
%mallado n:
nn=int32((n+1)/2);
U0=U1(88-nn + rem(n+1,2):88+nn,88-nn + rem(n+1,2):88+nn);   %imagen con ruido
Io=Uo(88-nn+ rem(n+1,2):88+nn,88-nn+ rem(n+1,2):88+nn); %imagen original
%Reordenamos los datos obtenidos hasta el momento:
fo=reshape(Io,(n+2)^2,1); %imagen original
f=reshape(U0,(n+2)^2,1); %imagen con ruido
%Construcción del gradiente del funcional de costo Fp
h=1/(1+n);
xv=0:h:1;
yv=0:h:1;
sz=length(xv);
for i=1:sz
    for j=1:sz
        funca(i,j)=xv(i)^2;
        funcb(i,j)=yv(j)^2;
        funcc(i,j)=xv(i)*yv(j);
        funcd(i,j)=xv(i);
        funce(i,j)=yv(j);
    end
end
U=NSSp(n,lambda,gamma,eeps); %Ecuación de estado para el cálculo del gradiente
eca=ADJp(n,U,lambda,gamma,eeps); %Problema adjunto para el cálculo del gradiente
inA=trapezoidal1(funca.*(U-U0)*eca);
inB=trapezoidal1(funcb.*(U-U0)*eca);
inC=trapezoidal1(funcc.*(U-U0)*eca);
inD=trapezoidal1(funcd.*(U-U0)*eca);
inE=trapezoidal1(funce.*(U-U0)*eca);
inF=trapezoidal1((U-U0)*eca);
Fp=[    2*(betha*(((1/5)*lambda(1)) + ((1/9)*lambda(2)) + ((1/8)*lambda(3))
+ ((1/4)*lambda(4)) + ((1/6)*lambda(5)) + ((1/3)*lambda(6)))) + inA;
        2*(betha*(((1/9)*lambda(1)) + ((1/5)*lambda(2)) + ((1/8)*lambda(3))
+ ((1/6)*lambda(4)) + ((1/4)*lambda(5)) + ((1/3)*lambda(6)))) + inB;
        2*(betha*(((1/8)*lambda(1)) + ((1/8)*lambda(2)) + ((1/9)*lambda(3))
+ ((1/6)*lambda(4)) + ((1/6)*lambda(5)) + ((1/4)*lambda(6)))) + inC;
        2*(betha*(((1/4)*lambda(1)) + ((1/6)*lambda(2)) + ((1/6)*lambda(3))
+ ((1/3)*lambda(4)) + ((1/4)*lambda(5)) + ((1/2)*lambda(6)))) + inD;
        2*(betha*(((1/6)*lambda(1)) + ((1/4)*lambda(2)) + ((1/6)*lambda(3))
```

```
+ ((1/4)*lambda(4)) + ((1/3)*lambda(5)) + ((1/2)*lambda(6)))) + inE;
        2*(betha*(((1/3)*lambda(1)) + ((1/3)*lambda(2)) + ((1/4)*lambda(3))
+ ((1/2)*lambda(4)) + ((1/2)*lambda(5)) + lambda(6))) + inF ];
alpha=1;   %Inicialización del tamaño de la dirección de descenso
s=-alpha*H*Fp;    %dirección de descenso
V=[Fp;0; 0; 0; 0; 0; 0]; %vector que almacena el gradiente del funcional actual
%y el "anterior", en este caso 0
v=(V(1:6)-V(7:12));  %diferencia entre gradiente del funcional de costo actual
%y el anterior
%Inicio del método cuasi Newton con fórmula DFP:
while norma1*alpha>1e-7 || it<5
%datos iniciales para el algoritmo de búsqueda lineal
    nlambda=lambda;
cost1=funcionaldecosto1(U,betha,lambda);
armijo=100; %Cálculo del funcional de costo
p=-H*Fp; %dirección de descenso
countarm=0;
V1=V;
countarm=0;
H1=H;
%Algoritmo de búsqueda lineal (Fórmula de Armijo):
while armijo >-alpha*(1e-4)*(Fp'*H1*Fp) && alpha>1e-3
alpha=1/2^(countarm);%incremento del parámetro para la dirección de descenso
  nlambda=nlambda-alpha*H1*Fp;%prueba el resultado con el alfa
%de la iteración de la búsqueda lineal actual
%Construcción de la matriz con los valores del polinomio lambda en cada
%punto del mallado:
  l1=nlambda(1); %Parámetro A en el polinomio cuadrático lambda
l2=nlambda(2);%Parámetro B en el polinomio cuadrático lambda
l3=nlambda(3);%Parámetro C en el polinomio cuadrático lambda
l4=nlambda(4);%Parámetro D en el polinomio cuadrático lambda
l5=nlambda(5);%Parámetro E en el polinomio cuadrático lambda
l6=nlambda(6);%Parámetro F en el polinomio cuadrático lambda
for i=1:sz
    for j=1:sz
lamb1(i,j)=l1*xv(i)^2 + l2*yv(j)^2 + l3*xv(i)*yv(j) + l4*xv(i) + l5*yv(j) + l6;
    end
end
%Proyección de la solución (el polinomio debe ser mayor a 0 en todos los
%puntos del mallado)
if min(min(lamb1))<0
    nlambda=sparse(6,1);
```

```
end
s1=-alpha*H1*Fp;%dirección de descenso con el alpha actual
U=NSSp(n,nlambda,gamma,eeps); %resultado de la ecuación de estado
eca=ADJp(n,U,nlambda,gamma,eeps); %resultado del problema adjunto
%Cálculo del gradiente del funcional, tomando en cuenta el parámetro alfa
%actual
inA=trapezoidal1(funca.*(U-U0)*eca);
inB=trapezoidal1(funcb.*(U-U0)*eca);
inC=trapezoidal1(funcc.*(U-U0)*eca);
inD=trapezoidal1(funcd.*(U-U0)*eca);
inE=trapezoidal1(funce.*(U-U0)*eca);
inF=trapezoidal1((U-U0)*eca);
Fp=[    2*(betha*(((1/5)*nlambda(1)) + ((1/9)*nlambda(2)) + ((1/8)*nlambda(3))
+ ((1/4)*nlambda(4)) + ((1/6)*nlambda(5)) + ((1/3)*nlambda(6)))) + inA;
        2*(betha*(((1/9)*nlambda(1)) + ((1/5)*nlambda(2)) + ((1/8)*nlambda(3))
+ ((1/6)*nlambda(4)) + ((1/4)*nlambda(5)) + ((1/3)*nlambda(6)))) + inB;
        2*(betha*(((1/8)*nlambda(1)) + ((1/8)*nlambda(2)) + ((1/9)*nlambda(3))
+ ((1/6)*nlambda(4)) + ((1/6)*nlambda(5)) + ((1/4)*nlambda(6)))) + inC;
        2*(betha*(((1/4)*nlambda(1)) + ((1/6)*nlambda(2)) + ((1/6)*nlambda(3))
+ ((1/3)*nlambda(4)) + ((1/4)*nlambda(5)) + ((1/2)*nlambda(6)))) + inD;
        2*(betha*(((1/6)*nlambda(1)) + ((1/4)*nlambda(2)) + ((1/6)*nlambda(3))
+ ((1/4)*nlambda(4)) + ((1/3)*nlambda(5)) + ((1/2)*nlambda(6)))) + inE;
        2*(betha*(((1/3)*nlambda(1)) + ((1/3)*nlambda(2)) + ((1/4)*nlambda(3))
+ ((1/2)*nlambda(4)) + ((1/2)*nlambda(5)) + nlambda(6))) + inF ];
V1=[Fp; V1(1:6)]; %vector que almacena el valor del gradiente en la iteracion
%k+1 y en k
v1=(V1(1:6)-V1(7:12));%diferencia del Gradiente
H1=H1 -((H1*(v1*v1')*H1)/(v1'*H1*v1)) + ((s1*s1')/(v1'*s1)); %Cálculo del DFP


  cost2=funcionaldecosto1(U,betha,nlambda); %Cálculo del funcional de costo
 armijo=cost2-cost1;
countarm=countarm+1;
end
cost2 %Funcional de costo, luego de haber obtenido un alfa que satisfaga
%la regla de Armijo
lambda=lambda+alpha*p; %Cálculo de lambda en la iteración actual
%Construcción de la matriz con los valores del polinomio lambda en cada
%punto del mallado:
l1=lambda(1);
l2=lambda(2);
l3=lambda(3);
```

```
l4=lambda(4);
l5=lambda(5);
l6=lambda(6);
for i=1:sz
    for j=1:sz
lamb1(i,j)=l1*xv(i)^2 + l2*yv(j)^2 + l3*xv(i)*yv(j) + l4*xv(i) + l5*yv(j) + l6;
    end
end
%Proyección de la solución (el polinomio debe ser mayor a 0 en todos los
%puntos del mallado)
if min(min(lamb1))<0
    lambda=sparse(6,1);
end
lambda %Cálculo de lambda en la iteración actual
s=alpha*p;
U=NSSp(n,lambda,gamma,eeps);  %Resultado de la ecuación de estado
%con los datos de la iteración actual
eca=ADJp(n,U,lambda,gamma,eeps); %Resultado del problema adjunto
%con los datos de la iteración actual
%Cálculo del gradiente del funcional para la iteración actual
inA=trapezoidal1(funca.*(U-U0)*eca);
inB=trapezoidal1(funcb.*(U-U0)*eca);
inC=trapezoidal1(funcc.*(U-U0)*eca);
inD=trapezoidal1(funcd.*(U-U0)*eca);
inE=trapezoidal1(funce.*(U-U0)*eca);
inF=trapezoidal1((U-U0)*eca);
Fp=[    2*(betha*(((1/5)*lambda(1)) + ((1/9)*lambda(2)) + ((1/8)*lambda(3))
+ ((1/4)*lambda(4)) + ((1/6)*lambda(5)) + ((1/3)*lambda(6)))) + inA;
        2*(betha*(((1/9)*lambda(1)) + ((1/5)*lambda(2)) + ((1/8)*lambda(3))
+ ((1/6)*lambda(4)) + ((1/4)*lambda(5)) + ((1/3)*lambda(6)))) + inB;
        2*(betha*(((1/8)*lambda(1)) + ((1/8)*lambda(2)) + ((1/9)*lambda(3))
+ ((1/6)*lambda(4)) + ((1/6)*lambda(5)) + ((1/4)*lambda(6)))) + inC;
        2*(betha*(((1/4)*lambda(1)) + ((1/6)*lambda(2)) + ((1/6)*lambda(3))
+ ((1/3)*lambda(4)) + ((1/4)*lambda(5)) + ((1/2)*lambda(6)))) + inD;
        2*(betha*(((1/6)*lambda(1)) + ((1/4)*lambda(2)) + ((1/6)*lambda(3))
+ ((1/4)*lambda(4)) + ((1/3)*lambda(5)) + ((1/2)*lambda(6)))) + inE;
        2*(betha*(((1/3)*lambda(1)) + ((1/3)*lambda(2)) + ((1/4)*lambda(3))
+ ((1/2)*lambda(4)) + ((1/2)*lambda(5)) + lambda(6))) + inF ];
V=[Fp; V(1:6)]; %vector que almacena el valor del gradiente en la iteracion
%k+1 y en k
v=(V(1:6)-V(7:12));%diferencia entre gradiente del funcional de costo actual
%y el anterior
```

```
norma1=norm(Fp);
H=H -((H*(v*v')*H)/(v'*H*v)) + ((s*s')/(v'*s));  %Cálculo de la fórmula DFP
it=it+1
end
l=lambda %resultado
%Cómputo de la solución en la ecuación de estado y gráficos
Uf=NSSp(n,l,gamma,eeps)
figure(1)
imshow(Uf,[])
figure(2)
imshow(U0,[])
figure(3)
imshow(Io,[])
end
```