



UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA QUÍMICA
INSTITUTO DE MATEMÁTICA APLICADA DEL LITORAL

TESIS PRESENTADA COMO PARTE DE LOS REQUISITOS DE LA UNIVERSIDAD NACIONAL DEL
LITORAL PARA LA OBTENCIÓN DEL GRADO ACADÉMICO DE
Doctor en Matemática

EN EL CAMPO DE: **Problemas Inversos**

TÍTULO DE LA TESIS:

**Regularización wavelet-espectral y wavelet-vaguelet de
problemas inversos mal condicionados**

AUTOR:

María Florencia Acosta

INSTITUCIÓN DONDE SE REALIZÓ:

Instituto de Matemática Aplicada del Litoral
(CONICET-UNL)

DIRECTOR DE TESIS:

Dr. Ruben Daniel Spies

IMAL (CONICET) - FIQ (UNL)

CODIRECTOR DE TESIS:

Dra. Gisela Luciana Mazzieri

IMAL (CONICET) - FBCB (UNL)

JURADO COMPUESTO POR:

Dr. Hugo A. Aimar

Dra. Aurora D. Rubio

Dr. Domingo A. Tarzia

AÑO DE PRESENTACIÓN: 2019

*“Je donnerais tout ce que je sais,
pour la moitié de ce que je ne sais pas.”*

- René Descartes

Agradecimientos

Quiero agradecer a todos aquellos quienes de una u otra manera colaboraron en mi formación e hicieron posible el desarrollo de esta tesis.

A mis directores, gracias por el apoyo, este trabajo fue posible gracias su guía, empeño y esfuerzo. Fundamentalmente, gracias Rubén por el incalculable tiempo invertido en la corrección de esta tesis.

A mi familia por el apoyo incondicional y la motivación constante. Gracias Tami por soportar años de una hermana estresada.

A Adrián por todo el amor y la paciencia necesaria para convivir con una becaria doctoral en la última etapa de tesis.

A CONICET por haber apoyado económicamente la realización de esta tesis mediante una beca doctoral. Gracias también al IMAL por brindarme un lugar de trabajo con calidez humana y la infraestructura necesaria. Particularmente gracias a Marce y Diego por esas charlas a la hora del café y por su buena predisposición siempre.

A cada miembro del jurado, por haber aceptado leer este trabajo y por cada una de sus sugerencias.

A todo el hermoso grupo de Problemas Inversos, principalmente a mis hermanitos Iba y Pilin.

A mis amigos y compañeros de oficina e instituto, por haberme soportado en cada etapa. Gracias Maikel, Pablo, Penélope y Gusti!

A mis amigas y correctoras favoritas, gracias Marilina y Estefi por el tiempo invertido en escuchar, leer y sugerir.

A Mario, el jefe y amigo que me tuvo una paciencia infinita a lo largo de este proceso.

A mis docentes del profesorado que me incentivaron a seguir estudiando. Gracias Ana María, Bibiana, Liliana y Karina.

Y sobre todo gracias las mejores amigas que el profesorado me dejó... Caro y Maruka!

Índice general

| | |
|--|----|
| <i>Agradecimientos</i> | 1 |
| Resumen | 7 |
| Introducción | 9 |
| Capítulo 1. Preliminares | 15 |
| 1.1. Inversa generalizada de Moore-Penrose | 15 |
| 1.2. Expansión en valores singulares de operadores compactos y autoadjuntos | 17 |
| 1.3. Teoría espectral | 19 |
| 1.4. Análisis multiresolución y representación wavelet | 22 |
| 1.4.1. Análisis multiresolución | 23 |
| 1.4.2. Umbralado wavelet | 26 |
| 1.5. Espacios de Sobolev y de Besov | 27 |
| 1.5.1. Análisis multiresolución y los espacios de Besov débiles | 29 |
| 1.6. Ruido blanco | 30 |
| 1.6.1. Definiciones básicas | 31 |
| 1.6.2. Procesos estocásticos | 33 |
| Capítulo 2. Introducción a los problemas inversos y métodos de regularización | 35 |
| 2.1. Problemas inversos mal condicionados | 35 |
| 2.1.1. Ejemplo: la diferenciación como problema inverso | 36 |
| 2.2. Métodos de regularización | 37 |
| 2.2.1. Métodos de regularización espectrales | 39 |
| 2.2.2. Métodos de Tikhonov generalizados | 42 |
| 2.3. Convergencia y optimalidad | 45 |
| 2.4. Calificación | 50 |
| 2.5. Reglas de elección de parámetro | 52 |
| 2.5.1. Reglas de elección de parámetro <i>a priori</i> | 52 |

| | | |
|--------------------|---|-----------|
| 2.5.2. | Reglas de elección de parámetro <i>a posteriori</i> : el principio de discrepancia de Morozov | 53 |
| 2.6. | Métodos de regularización espectrales | 54 |
| 2.6.1. | El método de descomposición en valores singulares truncada (TSVD) | 55 |
| 2.6.2. | El método de Tikhonov-Phillips | 55 |
| 2.6.3. | El método de Showalter o de regularización asintótica | 56 |
| 2.6.4. | El método de Landweber | 57 |
| Capítulo 3. | Regularización wavelet-espectral | 59 |
| 3.1. | Preliminares | 60 |
| 3.2. | Análisis de convergencia | 61 |
| 3.2.1. | Estimación del error de proyección wavelet | 61 |
| 3.2.2. | Estimación del error de umbralado wavelet | 66 |
| 3.2.3. | Estimación del error de regularización | 68 |
| 3.2.4. | Convergencia del método wavelet-Landweber | 72 |
| 3.3. | Generalización a otros métodos espectrales | 76 |
| Capítulo 4. | Regularización wavelet-vaguelet | 85 |
| 4.1. | Introducción | 85 |
| 4.1.1. | Descomposición en valores singulares (SVD) | 85 |
| 4.1.2. | Representación mediante wavelets | 86 |
| 4.1.3. | Descomposición wavelet-vaguelet (DWV) | 87 |
| 4.2. | Operadores lineales débilmente invertibles | 88 |
| 4.3. | Operadores homogéneos | 91 |
| 4.4. | Regularización wavelet-vaguelet | 93 |
| 4.4.1. | Introducción | 94 |
| 4.4.2. | Vaguelets | 95 |
| 4.4.3. | Descomposición y regularización wavelet-vaguelet | 101 |
| 4.5. | Análisis de convergencia | 106 |
| 4.5.1. | Los métodos RWV y RWVU | 106 |
| 4.5.2. | Errores | 107 |
| 4.5.3. | Lemas previos | 112 |
| 4.5.4. | Convergencia del método RWV | 122 |
| 4.5.5. | Convergencia del método RWVU | 125 |

| | |
|--|------------|
| 4.6. Operadores integrales | 131 |
| Capítulo 5. Aplicaciones y resultados numéricos | 135 |
| 5.1. Aplicaciones del método de regularización wavelet-espectral | 135 |
| 5.1.1. Deconvolución 1D: restauración de señales | 136 |
| 5.1.2. Deconvolución 2D: restauración de imágenes | 156 |
| 5.2. Aplicaciones del método RWV | 168 |
| Conclusiones y trabajos futuros | 177 |
| Anexos | 179 |
| Anexo A | 179 |
| Anexo B | 180 |
| Bibliografía | 187 |

Resumen

En esta tesis se aborda el estudio de dos métodos de regularización para problemas inversos mal condicionados que utilizan descomposición wavelet. El primero de ellos (regularización wavelet-espectral) utiliza descomposición wavelet seguida de un proceso de umbralado (*wavelet thresholding*), para reducir la influencia de las componentes de alta frecuencia presentes en los datos, provenientes del ruido, previo a la aplicación de un método de regularización espectral. El segundo método (regularización wavelet-vaguelet) utiliza también descomposición wavelet, pero ahora incorporando simultáneamente una representación del operador asociado al problema, basada en su descomposición en valores singulares, SVD.

El fundamento principal para la incorporación de wavelets en el tratamiento de problemas inversos mal condicionados, en los que los datos están contaminados con algún tipo de ruido, radica en la capacidad de filtrado de la descomposición y umbralado wavelet. En un contexto general aplicado a señales o imágenes, una base multiresolución wavelet nos permite diferenciar y agrupar los rasgos más importantes que determinan “la identidad” de una señal a través de las componentes de baja y alta frecuencia. Utilizando esta ventaja, el método propone realizar una descomposición wavelet del dato seguida de umbralado para eliminar o atenuar las componentes de alta frecuencia, especialmente asociadas al ruido, con el objetivo de que este filtrado previo provea a los métodos espectrales tradicionales datos considerablemente más cercanos a los datos exactos, libres de ruido. Esto permite obtener regularizaciones construidas con parámetros de regularización significativamente más pequeños, obteniendo así soluciones más cercanas a la solución exacta. En este contexto, se realiza un análisis detallado de la convergencia para los errores de proyección, umbralado y regularización del método wavelet-espectral para los métodos de Landweber y TSVD. Posteriormente, se prueba el teorema fundamental de convergencia de orden casi-óptimo para el método combinado wavelet-Landweber, en el Teorema 3.16. Además, se generalizan estos resultados de convergencia para el caso de métodos wavelet-espectrales en los que la familia espectral posee ciertas propiedades asociadas a su calificación clásica. En particular, en el Teorema 3.22 se prueba que, bajo ciertas condiciones, el método de dos pasos converge de forma casi-óptima.

En los primeros capítulos de esta tesis se utilizan dos tipos de bases ortogonales para la representación de funciones en el caso de problemas inversos mal condicionados. El primer tipo, SVD, representa eficientemente al operador del problema $Tx = y$. El segundo tipo, las bases wavelets ortonormales, se utilizan para representar de manera eficiente a ciertos espacios de funciones que poseen algún tipo de regularidad y, por tal motivo, se las utiliza para codificar adecuadamente información previa disponible sobre la solución exacta del problema. Con el objetivo de disponer de una descomposición que permita incorporar simultáneamente tanto las ventajas de SVD como las de la descomposición wavelet se introduce el concepto de vaguelets en la resolución de problemas inversos mal condicionados en el desarrollo del método de regularización wavelet-vaguelet (RWV) en el Capítulo 4. Previo a la presentación del método se definen dos tipos de operadores especiales, operadores débilmente invertibles y operadores homogéneos y se prueban algunos resultados importantes relacionados a tales operadores, los que resultan de particular interés en varias de las demostraciones posteriores relacionadas con la convergencia del método propuesto. Luego, se presenta el concepto de vaguelet y se prueban una serie de propiedades de tales funciones. Posteriormente se realiza un análisis detallado de la convergencia del error correspondiente al método propuesto considerando dos versiones del mismo: con y sin umbralado wavelet fuerte. En particular se prueba que, bajo ciertas condiciones, el método de regularización wavelet-vaguelet y el método wavelet-vaguelet con umbralado wavelet convergen de manera óptima. Por último, en lo que respecta al método RWV, se presenta y analiza una aplicación concreta del método para el caso de operadores integrales que poseen descomposición wavelet vaguelet.

Finalmente, en el Capítulo 5 se presentan varias aplicaciones a problemas inversos mal condicionados en restauración de señales e imágenes, cuyas implementaciones numéricas permiten una mejor visualización de los resultados teóricos obtenidos en esta tesis, y muestran claramente la potencialidad de los métodos desarrollados. Estos resultados prueban que los métodos desarrollados mejoran, en algunos casos significativamente, el desempeño de los métodos de regularización tradicionales.

Introducción

En las últimas tres décadas el estudio de problemas inversos ha sido sin lugar a dudas una de las áreas de mayor y más rápido crecimiento dentro de la Matemática Aplicada. Una de las principales causas de ello ha sido el creciente número de aplicaciones de las técnicas de regularización a una amplia variedad de problemas provenientes de otras ciencias y de la industria.

En un contexto general, un problema inverso (lineal) puede formularse como la necesidad de determinar x en la ecuación de la forma

$$Tx = y, \tag{1}$$

donde T es un operador lineal acotado de un espacio normado X en un espacio normado Y . El problema de determinar x en (1), se dice que es mal condicionado cuando no satisface al menos uno de los tres postulados de Hadamard ([17]). Sin embargo, la principal causa de mal condicionamiento es la violación del tercer postulado (falta de dependencia continua) que se manifiesta en el hecho que T^\dagger , la inversa generalizada de Moore-Penrose de T , no es acotada. Esto se traduce en que todos los métodos numéricos tradicionales para aproximar la solución exacta se vuelven inestables y, en tal caso, pequeños errores en el dato y pueden resultar en errores arbitrariamente grandes en x ([41]).

Como ya mencionamos, problemas mal condicionados se presentan en una amplia variedad de áreas y aplicaciones. Por ejemplo, la determinación de la distribución de la densidad en un cuerpo mediante la emisión de rayos X, que constituye la base de la Tomografía Computada ([35]), se realiza mediante la inversión de la transformada de Radón o, en el caso de densidades radiales, de la ecuación integral de Abel. Este problema es siempre mal condicionado y la aproximación de las soluciones requiere la aplicación de métodos de regularización. La ecuación integral de Abel aparece también en Física del Plasma. Una gran cantidad de problemas en Física originan también problemas inversos mal condicionados. Entre ellos mencionamos: problemas de fluorescencia ([29]), Reología ([22]), superconductividad ([11]), problemas de potenciales inversos en Geofísica ([33], [18]) y una gran variedad de problemas provenientes de diversas ramas de la Medicina ([40], [39]).

El tratamiento de señales e imágenes es una fuente inagotable de problemas inversos mal condicionados. Dentro de esta área merecen citarse todos los problemas relacionados con las imágenes obtenidas tanto por telescopios terrestres como por telescopios orbitales, imágenes en medicina generadas por emisión de rayos X, señales de EEG, etc. ([36], [37]). En estos problemas, un modelo matemático ampliamente utilizado da lugar a una integral de convolución con un núcleo que se conoce como “función de dispersión puntual”. Estas ecuaciones integrales de Fredholm de primera clase, salvo en casos excepcionales de núcleos degenerados, resultan en problemas mal condicionados ([17]). Una aplicación muy interesante de los métodos de regularización en la reconstrucción de imágenes distorsionadas, se presenta en el artículo [28], donde se detalla como se estudiaron y reconstruyeron las imágenes enviadas durante los primeros tres años (entre 1990 y 1993) por el telescopio espacial Hubble (el que fue puesto en órbita con un espejo con fallas de manufactura). Esta reconstrucción fue posible gracias a la utilización de métodos de regularización que permitieron resolver un problema inverso mal condicionado con cerca de 800 mil variables.

Todo lo anterior revela la importancia que tiene el desarrollo de herramientas matemáticas apropiadas para resolver problemas inversos mal condicionados. Esto es así puesto que, sin regularización, los métodos numéricos tradicionales se vuelven inestables haciéndose imprescindible la implementación de estrategias, métodos o algoritmos matemáticos que permitan recuperar la mayor cantidad de información posible sobre la solución exacta del problema, manteniendo la estabilidad del mismo. Se trata entonces de un tema muy importante tanto desde la perspectiva puramente matemática como por sus potenciales aplicaciones en una inmensa variedad de problemas concretos e interesantes.

En este trabajo presentamos dos métodos de regularización que utilizan descomposición wavelet. El primero de ellos (regularización wavelet-espectral) utiliza descomposición wavelet seguida de un proceso de umbralado (*wavelet thresholding*), para reducir la influencia de las componentes de alta frecuencia presentes en el dato, provenientes del ruido, previo a la aplicación de un método de regularización espectral. El segundo método de regularización (regularización wavelet-vaguelet) utiliza también descomposición wavelet, pero ahora incorporando simultáneamente una representación del operador asociado al problema, basada en su descomposición en valores singulares, SVD.

Esta tesis consta de cinco capítulos. El primer capítulo contiene conceptos preliminares y varias herramientas que son fundamentales para el estudio de problemas inversos mal condicionados. Allí, presentamos en primer lugar el concepto de inversa generalizada de Moore-Penrose,

las principales definiciones relacionadas con la teoría de operadores compactos y sus expansiones en valores singulares. El capítulo continúa con una breve introducción a la teoría espectral, a la teoría de wavelets y al análisis multiresolución. Posteriormente, en la Sección 1.5 realizamos un breve repaso de espacios de Sobolev y de Besov, los que serán necesarios en muchos de los resultados analíticos que presentamos en los Capítulos 3 y 4. Finalmente, en la última sección del Capítulo 1 realizamos un breve repaso de algunos conceptos básicos sobre procesos estocásticos y ruido blanco.

En el Capítulo 2 presentamos una introducción al mundo de los problemas inversos mal condicionados y a los métodos de regularización. En la primera sección formalizamos el concepto de mal condicionamiento en el sentido de Hadamard presentando un ejemplo simple con el operador derivada. En la sección siguiente introducimos el concepto de método de regularización y enunciamos algunas de sus propiedades. En particular, estudiamos los métodos de regularización espectrales, de Tikhonov generalizados y repasamos algunos de los métodos tradicionales más frecuentemente utilizados. En la Sección 2.3 definimos los diferentes tipos de errores generados al regularizar un problema e introducimos la noción de optimalidad. A continuación, se presentan órdenes de convergencia y cotas para los distintos tipos de errores y se introduce el concepto de calificación. En la Sección 2.5 se introduce el concepto de regla de elección de parámetros, definiendo los conceptos de reglas *a-priori* y *a-posteriori*. Finalmente, en la última sección se retoman algunos ejemplos de métodos de regularización espectrales relacionándolos con conceptos ya desarrollados tales como descomposición en valores singulares y calificación.

En el Capítulo 3 presentamos el primero de los métodos de regularización desarrollados en esta tesis: el método de regularización wavelet-espectral. El fundamento principal para la incorporación de wavelets en el tratamiento de este tipo de problemas, en los que los datos están contaminados con algún tipo de ruido, radica en la capacidad de filtrado de la descomposición y umbralado wavelet. En un contexto general aplicado a señales o imágenes, una base multiresolución wavelet nos permite diferenciar y agrupar los rasgos más importantes que determinan “la identidad” de una señal a través de las componentes de baja y alta frecuencia. Utilizando esta ventaja, el método propone realizar una descomposición wavelet del dato seguida de umbralado para atenuar o eliminar las componentes de alta frecuencia, especialmente asociadas al ruido, con el objetivo de que este filtrado previo provea a los métodos espectrales tradicionales de datos considerablemente más cercanos a los datos exactos, libres de ruido. Esto permite obtener regularizaciones construidas con parámetros de regularización significativamente más pequeños, obteniendo así soluciones más cercanas a la solución exacta.

Así, el Capítulo 3 comienza con una sección de preliminares donde presentamos el problema a abordar, las principales hipótesis sobre el operador y los espacios funcionales con los que trabajaremos. En la segunda sección hacemos un análisis detallado de la convergencia para los errores de proyección, umbralado y regularización, lo que nos permitirá posteriormente probar el teorema fundamental de convergencia de orden casi-óptimo para el método combinado wavelet-Landweber, en el Teorema 3.16. Por último, en la Sección 3.3 generalizamos esos resultados de convergencia para el caso de métodos wavelet-espectrales en los que la familia espectral posee ciertas propiedades asociadas a su calificación clásica (ver [?], [17], [43]). En particular, en el Teorema 3.22 probamos que, bajo ciertas condiciones, el método de dos pasos converge de forma casi-óptima.

En el Capítulo 3 se utilizan dos tipos de bases ortogonales para la representación de funciones en problemas inversos mal condicionados. El primer tipo, SVD, representa eficientemente al operador del problema $Tx = y$. El segundo tipo, las bases wavelets ortonormales, se utilizan para representar de manera eficiente ciertos espacios de funciones que poseen algún tipo de regularidad y por tal motivo se las utiliza para codificar adecuadamente información previa disponible sobre la solución exacta del problema. A la hora de resolver problemas inversos, sería deseable disponer de una descomposición que permita incorporar simultáneamente tanto las ventajas de SVD como las de la descomposición wavelet. Esto nos permitiría contar con una representación que incorpore de manera eficiente tanto información relevante sobre el operador asociado al problema como sobre los supuestos de regularidad de la solución exacta del mismo. Con este objetivo en mente introducimos el concepto de “vaguelets” en la resolución de problemas inversos mal condicionados en el desarrollo del método de regularización wavelet-vaguelet (RWV) que presentamos en el Capítulo 4.

Así, en la primera sección del Capítulo 4 describimos brevemente la motivación que origina el surgimiento del método y presentamos además un resumen de los elementos que intervienen en su construcción. En las Secciones 4.2 y 4.3 definimos dos tipos de operadores especiales, operadores débilmente invertibles y operadores homogéneos y probamos algunos resultados importantes relacionados a tales operadores, los que resultarán de interés en demostraciones posteriores relacionadas con la convergencia del método propuesto. En la sección siguiente presentamos el método de regularización wavelet-vaguelet. Para ello definimos previamente el concepto de vaguelet y probamos una serie de propiedades de las mismas. En la Sección 4.5 realizamos un análisis detallado de la convergencia del error correspondiente al método propuesto considerando dos versiones del mismo: con y sin umbralado wavelet fuerte. En particular probamos que, bajo

ciertas condiciones, el método de regularización wavelet-vaguelet y el método wavelet-vaguelet con umbralado wavelet convergen de manera óptima. Por último, en la Sección 4.6 presentamos y analizamos una aplicación concreta del método para el caso de operadores integrales que poseen descomposición wavelet vaguelet.

Finalmente, en el Capítulo 5, implementamos y evaluamos en problemas concretos los métodos y resultados desarrollados en los Capítulos 3 y 4. Se presentan varias aplicaciones a problemas inversos mal condicionados en restauración de señales e imágenes, cuyas implementaciones numéricas permiten una mejor visualización de los resultados teóricos obtenidos y muestran claramente la potencialidad de los métodos desarrollados. Estos resultados prueban que dichos métodos mejoran, en algunos casos significativamente, el desempeño de los métodos de regularización tradicionales.

CAPÍTULO 1

Preliminares

A lo largo de este capítulo definiremos los conceptos básicos requeridos para la comprensión de esta tesis y enunciaremos varios resultados previos y clásicos que se utilizarán en el desarrollo de la misma. En particular, recordaremos los conceptos de inversa generalizada de Moore-Penrose, ruido blanco, espacios de Sobolev y espacios de Besov. A su vez, haremos una breve introducción a los conceptos básicos de la teoría espectral e introduciremos varios espacios y operadores asociados a la teoría de wavelets.

1.1. Inversa generalizada de Moore-Penrose

En esta sección proporcionamos la teoría básica para definir las mejores soluciones aproximadas de ecuaciones con operadores lineales y la inversa generalizada de Moore-Penrose para operadores lineales en espacios de Hilbert.

Consideremos la ecuación

$$Tx = y \tag{1.1}$$

donde $T : X \rightarrow Y$ es un operador lineal acotado entre dos espacios de Hilbert. Un elemento $x_0 \in X$ se denomina *solución de mínimos cuadrados* (s.m.c.) de (1.1) si

$$\|Tx_0 - y\| = \inf_{z \in X} \|Tz - y\|.$$

Si además x_0 satisface

$$\|x_0\| = \inf\{\|z\| : z \text{ es s.m.c de } Tx = y\},$$

entonces x_0 recibe el nombre de *mejor solución aproximada* de $Tx = y$. Así, la mejor solución aproximada de (1.1) es la solución de mínimos cuadrados de mínima norma.

Para definir la inversa generalizada de Moore-Penrose del operador T , comenzamos definiendo el operador \tilde{T} como

$$\tilde{T} \doteq T |_{N(T)^\perp} : \mathcal{N}(T)^\perp \rightarrow \mathcal{R}(T),$$

donde $\mathcal{N}(\cdot)$ denota el núcleo de un operador, $\mathcal{R}(\cdot)$ denota el rango y “ \perp ” denota complemento ortogonal. Claramente \tilde{T} es un operador biyectivo. La inversa generalizada de Moore-Penrose se define entonces de la siguiente manera.

Definición 1.1. La *inversa generalizada de Moore-Penrose* T^\dagger de T se define como la única extensión lineal de \tilde{T}^{-1} al dominio $\mathcal{D}(T^\dagger) \doteq \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ que satisface $\mathcal{N}(T^\dagger) = \mathcal{R}(T)^\perp$.

A continuación enunciamos algunas propiedades importantes de la inversa generalizada de Moore-Penrose que utilizaremos más adelante. Debido a que estos resultados son relativamente bien conocidos, no incluimos aquí las correspondientes demostraciones. El lector interesado puede ver por ejemplo las citas [17] y [25].

Proposición 1.2. Sean P y Q las proyecciones ortogonales de X sobre $\mathcal{N}(T)$ y de Y sobre $\overline{\mathcal{R}(T)}$, respectivamente. Entonces $\mathcal{R}(T^\dagger) = \mathcal{N}(T)^\perp$ y se satisfacen las siguientes cuatro identidades, denominadas ecuaciones de Moore-Penrose:

$$TT^\dagger T = T, \quad (1.2)$$

$$T^\dagger TT^\dagger = T^\dagger, \quad (1.3)$$

$$T^\dagger T = I - P, \quad (1.4)$$

$$TT^\dagger = Q|_{\mathcal{D}(T^\dagger)}. \quad (1.5)$$

Resulta evidente que las ecuaciones (1.4) y (1.5) implican (1.2) y (1.3). Más aún, se puede probar que las ecuaciones de Moore-Penrose caracterizan unívocamente al operador T^\dagger .

Proposición 1.3. El operador T^\dagger es cerrado. Además, T^\dagger es acotado si y solo si $\mathcal{R}(T)$ es cerrado.

El siguiente teorema es un resultado muy importante que relaciona las soluciones de mínimos cuadrados de (1.1) con la inversa generalizada de Moore-Penrose de T .

Teorema 1.4. Si $y \in \mathcal{D}(T^\dagger)$, entonces (1.1) tiene una única mejor solución aproximada dada por $x^\dagger \doteq T^\dagger y$. El conjunto S de todas las soluciones de mínimos cuadrados de (1.1) está dado por $S = x^\dagger + \mathcal{N}(T)$.

Para finalizar esta sección presentamos un teorema que caracteriza las soluciones de mínimos cuadrados de (1.1) mediante una ecuación a la que haremos referencia con frecuencia más adelante.

Teorema 1.5. Si $y \in \mathcal{D}(T^\dagger)$, entonces $z \in X$ es una solución de mínimos cuadrados de $Tx = y$ si y solo si satisface la “ecuación normal”

$$T^*Tz = T^*y, \tag{1.6}$$

donde T^* es el operador adjunto de T .

A partir de este resultado obtenemos que $T^\dagger y$ es también la mejor solución aproximada del problema $T^*Tz = T^*y$, es decir

$$T^\dagger = (T^*T)^\dagger T^*.$$

Se puede probar además que si $y \notin \mathcal{D}(T^\dagger)$, entonces no existen soluciones de mínimos cuadrados de $Tx = y$. Trabajar con la ecuación normal (1.6) en lugar de (1.1), suele tener varias ventajas, especialmente cuando T no es autoadjunto, pues T^*T si lo es. Además, en su versión discretizada, la ecuación (1.6) resulta en una matriz cuadrada.

En la siguiente sección presentaremos una herramienta que nos permitirá estudiar varias propiedades de la inversa generalizada de Moore-Penrose en el caso de operadores compactos: la expansión en valores singulares.

1.2. Expansión en valores singulares de operadores compactos y autoadjuntos

Sea K un operador lineal sobre un espacio de Hilbert X , compacto y autoadjunto. Entonces existe un *autosistema* $\{(\lambda_n; u_n)\}_{n \in \mathbb{N}}$ asociado a K , donde λ_n son los autovalores no nulos de K (los autovalores son reales por ser K autoadjunto) y u_n sus correspondientes autovectores. Por medio de este autosistema, el operador K se puede “diagonalizar” de la siguiente manera:

$$Kx = \sum_{n=1}^{\infty} \lambda_n \langle x, u_n \rangle u_n, \quad \forall x \in X. \tag{1.7}$$

Si el operador K no es autoadjunto, entonces podría no tener autovalores y por lo tanto, podría no existir un autosistema asociado a K . En tal caso, un proceso de diagonalización análogo a (1.6) puede hacerse utilizando el concepto de “sistema singular”, que definimos a continuación.

Definición 1.6. Sea $K : X \rightarrow Y$ un operador lineal compacto (no necesariamente autoadjunto). Un *sistema singular asociado a K* es una familia de la forma $\{(\sigma_n; u_n; v_n)\}_{n \in \mathbb{N}}$ donde:

- $\{\sigma_n^2\}_{n \in \mathbb{N}}$ son los autovalores no nulos de K^*K (y de KK^*), escritos en orden decreciente con su multiplicidad,

- $\{u_n\}_{n \in \mathbb{N}} \subset X$ es un sistema ortonormal completo de autovectores de K^*K (que expande $\overline{\mathcal{R}(K^*)} = \overline{\mathcal{R}(K^*K)}$) y
- $\{v_n\}_{n \in \mathbb{N}} \subset Y$ es un sistema ortonormal completo de autovectores de KK^* (que expande $\overline{\mathcal{R}(K)} = \overline{\mathcal{R}(KK^*)}$). Puede probarse que

$$v_n = \frac{Ku_n}{\|Ku_n\|}.$$

Observemos que K^*K es no negativo, pues $\langle K^*Kx, x \rangle = \|Kx\|^2 \geq 0$ para todo x y por ello, sus autovalores son no negativos. Además, como K es compacto, el operador K^*K también lo es y por lo tanto tiene a lo sumo una cantidad numerable de autovalores que se pueden acumular únicamente en cero (ver [50], Teorema 6.7).

Si $\{(\sigma_n; u_n; v_n)\}_{n \in \mathbb{N}}$ es el sistema singular para K , se puede probar que se satisfacen las siguientes igualdades:

$$Ku_n = \sigma_n v_n \quad \forall n \in \mathbb{N}, \quad (1.8)$$

$$K^*v_n = \sigma_n u_n \quad \forall n \in \mathbb{N}, \quad (1.9)$$

$$Kx = \sum_{n=1}^{\infty} \sigma_n \langle x, u_n \rangle v_n, \quad \forall x \in X, \quad (1.10)$$

$$K^*y = \sum_{n=1}^{\infty} \sigma_n \langle y, v_n \rangle u_n, \quad \forall y \in Y, \quad (1.11)$$

dónde, en el caso de infinitos valores singulares, las series convergen en las normas de los espacios de Hilbert X e Y , según corresponda. Las expresiones (1.10) y (1.11) se denominan “expansiones en valores singulares” (de K y K^* , respectivamente) y son las análogas, en dimensión infinita, a la bien conocida descomposición en valores singulares (SVD) de una matriz.

Utilizando el sistema singular asociado a un operador compacto K podemos obtener una representación en serie de su inversa generalizada de Moore-Penrose. En efecto, se tiene el siguiente resultado.

Teorema 1.7. Sean $K : X \rightarrow Y$ un operador lineal compacto con sistema singular $\{(\sigma_n; u_n; v_n)\}_{n \in \mathbb{N}}$, e $y \in Y$. Entonces,

$$y \in \mathcal{D}(K^\dagger) \iff \sum_{n=1}^{\infty} \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty. \quad (1.12)$$

Además, para $y \in \mathcal{D}(K^\dagger)$,

$$K^\dagger y = \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle}{\sigma_n} u_n. \quad (1.13)$$

La condición (1.12) es conocida como Criterio de Picard y provee una condición necesaria y suficiente para la existencia de una mejor solución aproximada para $y \in Y$. Si existe, entonces esa mejor solución aproximada tiene la representación dada por (1.13).

1.3. Teoría espectral

Como veremos en el resto de la tesis, la teoría espectral de operadores en espacios de Hilbert y el cálculo funcional proveen un sólido andamiaje matemático para el desarrollo de la teoría de regularización de problemas inversos mal condicionados. En esta sección presentaremos algunos conceptos y resultados básicos cuyas demostraciones pueden encontrarse, por ejemplo, en [3], [21] y [17].

Un concepto central en esta teoría es el de *función de un operador autoadjunto*, el cual definiremos en primer lugar para el caso de operadores lineales compactos autoadjuntos y luego estudiaremos el caso de operadores lineales autoadjuntos en general.

Sea $\{(\sigma_n; u_n; v_n)\}_{n \in \mathbb{N}}$ el sistema singular asociado al operador lineal compacto $K : X \rightarrow Y$. Como $\{(\sigma_n^2; u_n)\}_{n \in \mathbb{N}}$ es un autosistema para el operador compacto autoadjunto K^*K , se sigue de (1.7) que

$$K^*Kx = \sum_{n=1}^{\infty} \sigma_n^2 \langle x, u_n \rangle u_n, \quad \forall x \in X.$$

Definimos ahora el concepto de *familia espectral* para operadores compactos.

Definición 1.8. Sea $K : X \rightarrow Y$ un operador lineal, continuo y compacto, y $\{(\sigma_n, u_n, v_n)\}_{n \in \mathbb{N}}$ el sistema singular asociado a K . Para $\lambda \in \mathbb{R}$ y $x \in X$, definimos

$$E_\lambda x \doteq \begin{cases} \sum_{n: \sigma_n^2 < \lambda} \langle x, u_n \rangle u_n & \text{si } \lambda \leq 0 \\ \sum_{n: \sigma_n^2 < \lambda} \langle x, u_n \rangle u_n + Px & \text{si } \lambda > 0 \end{cases}$$

donde P es la proyección ortogonal de X sobre $\mathcal{N}(K^*K)$. El operador ortogonal E_λ es la proyección ortogonal sobre $X_\lambda \doteq \overline{\text{span}\{u_n : n \in \mathbb{N}, \sigma_n^2 < \lambda\}}$ y la colección de proyecciones $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ se denomina *familia espectral* de (o asociada a) K^*K .

La proyección P se agrega en la definición de E_λ solo para $\lambda > 0$ para considerar el caso en que 0 es autovalor de K^*K . Caso contrario, $\mathcal{N}(K^*K) = \{0\}$ y P es el operador nulo.

A continuación definiremos familias espectrales generales en espacios de Hilbert y, posteriormente, veremos que existe una correspondencia biunívoca entre el conjunto de familias

espectrales en un espacio de Hilbert X y el conjunto de operadores lineales autoadjuntos sobre X .

Definición 1.9. Una familia $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ de proyecciones ortogonales sobre X se denomina *familia espectral* o *resolución de la identidad* si satisface las siguientes condiciones:

- 1) $E_\lambda E_\mu = E_{\min\{\lambda, \mu\}}$, para $\lambda, \mu \in \mathbb{R}$;
- 2) $E_{-\infty} = 0$, $E_{+\infty} = I$, donde $E_{\pm\infty}x \doteq \lim_{\lambda \rightarrow \pm\infty} E_\lambda x$, para todo $x \in X$;
- 3) $E_{\lambda-} = E_\lambda$, donde $E_{\lambda-}x = \lim_{\varepsilon \rightarrow 0^+} E_{\lambda-\varepsilon}x$, para todo $x \in X$.

La siguiente proposición muestra que es posible definir una integral con respecto a una familia espectral arbitraria $\{E_\lambda\}_{\lambda \in \mathbb{R}}$.

Proposición 1.10. Sean $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ una familia espectral sobre X , $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continua y $x \in X$. Entonces para cualesquiera $a, b \in \mathbb{R}$, $a < b$, existe la integral $\int_a^b f(\lambda) dE_\lambda x$ definida como el límite de las sumas de Riemann $\sum_{i=1}^n f(\xi_i)(E_{\lambda_i} - E_{\lambda_{i-1}})x$, donde $a = \lambda_0 < \dots < \lambda_n = b$, $\xi_i \in (\lambda_{i-1}, \lambda_i]$, para $\max_{1 \leq i \leq n} |\lambda_i - \lambda_{i-1}| \rightarrow 0$.

Las correspondientes integrales impropias, si existen, se definen a partir de estas integrales en la forma usual, tomando límites.

Definición 1.11. Sean $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ una familia espectral sobre X , $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continua y $x \in X$. La integral $\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda x$ se define como el límite en X , si existe, de $\int_a^b f(\lambda) dE_\lambda x$ cuando $a \rightarrow -\infty$ y $b \rightarrow +\infty$.

Puesto que la condición 1) de la Definición 1.9 implica que $\langle E_\lambda x, x \rangle \leq \langle E_\mu x, x \rangle$ para $\lambda \leq \mu$ y para todo $x \in X$, se sigue entonces que la función $\lambda \rightarrow \langle E_\lambda x, x \rangle = \|E_\lambda x\|^2$ (para $x \in X$ fijo) es monótona no decreciente. Además la condición 3) implica que dicha función es continua por izquierda. Luego, esta función define una medida sobre \mathbb{R} que denotamos con $d\|E_\lambda x\|^2$ y a la que nos referiremos como “medida generada por la familia espectral E_λ para $x \in X$ ”. A continuación presentamos una condición necesaria y suficiente que garantiza la existencia de la integral $\int f(\lambda) dE_\lambda x$ en términos de una integral con respecto a la medida $d\|E_\lambda x\|^2$.

Proposición 1.12. Sean $\{E_\lambda\}_{\lambda \in \mathbb{R}}$ una familia espectral sobre X , $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continua y $x \in X$. Entonces

$$\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda x$$

existe si y solo si

$$\int_{-\infty}^{+\infty} f^2(\lambda) d\|E_\lambda x\|^2 < \infty. \quad (1.14)$$

Notar que el lado izquierdo en (1.14) es una integral impropia de Riemann-Stieltjes de $f^2(\lambda)$ con respecto a la función (real) $F(\lambda) \doteq \|E_\lambda x\|^2$.

La proposición a continuación exhibe la conexión que existe entre familias espectrales y operadores lineales autoadjuntos definidos sobre espacios de Hilbert. Si A es un operador lineal autoadjunto sobre X , entonces existe una única familia espectral con respecto a la cual A se representa mediante la función identidad $f(\lambda) = \lambda$. Más precisamente, se tiene el siguiente resultado:

Proposición 1.13. *Sea A un operador lineal autoadjunto sobre un espacio de Hilbert X . Entonces existe una única familia espectral $\{E_\lambda^A\}_{\lambda \in \mathbb{R}}$ sobre X tal que*

$$\mathcal{D}(A) = \left\{ x \in X : \int_{-\infty}^{+\infty} \lambda^2 d\|E_\lambda^A x\|^2 < \infty \right\}$$

y

$$Ax = \int_{-\infty}^{+\infty} \lambda dE_\lambda^A x, \quad \forall x \in \mathcal{D}(A). \quad (1.15)$$

Simbólicamente escribimos

$$A = \int_{-\infty}^{+\infty} \lambda dE_\lambda^A. \quad (1.16)$$

La familia espectral de la proposición anterior se denomina *familia espectral de A* y nos referiremos a (1.15) ó (1.16) como la *descomposición o representación espectral de A* .

En virtud de dicha proposición podemos ahora definir funciones de un operador autoadjunto las cuales también pueden representarse mediante integrales espectrales.

Definición 1.14. Sea A un operador autoadjunto en X con familia espectral $\{E_\lambda^A\}_{\lambda \in \mathbb{R}}$ y denotemos con \mathcal{M}_0 al conjunto de funciones medibles con respecto a la medida $d\|E_\lambda^A x\|^2$ para todo $x \in X$. Para $f \in \mathcal{M}_0$, se define el operador $f(A)$ mediante

$$f(A)x \doteq \int_{-\infty}^{+\infty} f(\lambda) dE_\lambda^A x \quad \text{para } x \in \mathcal{D}(f(A)),$$

donde

$$\mathcal{D}(f(A)) = \left\{ x \in X : \int_{-\infty}^{+\infty} f^2(\lambda) d\|E_\lambda^A x\|^2 < \infty \right\}.$$

Observar que \mathcal{M}_0 contiene, en particular, a todas las funciones continuas por tramos. Esta observación será de particular interés cuando utilicemos funciones de operadores autoadjuntos para construir métodos de regularización como aproximaciones a la inversa generalizada del operador asociado al problema en cuestión.

A continuación presentamos algunas propiedades de las funciones de operadores autoadjuntos. Las mismas serán utilizadas en los capítulos siguientes principalmente para el caso de operadores compactos.

- Si A es un operador autoadjunto estrictamente positivo, es decir si existe $\gamma > 0$ tal que

$$\langle Ax, x \rangle \geq \gamma \|x\|^2 \text{ para todo } x \in \mathcal{D}(A),$$

con familia espectral $\{E_\lambda^A\}_{\lambda \in \mathbb{R}}$, entonces para toda función $f \in \mathcal{M}_0$ se tiene que

$$\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda^A x = \int_\gamma^{+\infty} f(\lambda) dE_\lambda^A x.$$

En tal caso, es suficiente con que la función f esté definida solo en el intervalo $[\gamma, +\infty)$.

- Si $T : X \rightarrow Y$ es un operador lineal acotado y $A \doteq T^*T$ con familia espectral $\{E_\lambda^{T^*T}\}_{\lambda \in \mathbb{R}}$, entonces para toda $f \in \mathcal{M}_0$ y $\forall x \in X$, se tiene que

$$\int_{-\infty}^{+\infty} f(\lambda) dE_\lambda^{T^*T} x = \int_0^{\|T\|^2 + \varepsilon} f(\lambda) dE_\lambda^{T^*T} x = \lim_{\varepsilon \rightarrow 0} \int_0^{\|T\|^2 + \varepsilon} f(\lambda) dE_\lambda^{T^*T} x.$$

Luego, la función f se puede restringir al intervalo $[0, \|T\|^2 + \varepsilon]$ para cualquier $\varepsilon > 0$. Más aún, bajo las mismas hipótesis sobre el operador T , se puede probar que valen las siguientes identidades, $\forall x \in X$:

$$\|f(T^*T)x\|^2 = \int f^2(\lambda) d\|E_\lambda^{T^*T} x\|^2 \quad (1.17)$$

y

$$f(T^*T)T^* = T^*f(TT^*), \quad (1.18)$$

para toda función f continua por tramos.

1.4. Análisis multiresolución y representación wavelet

En esta sección presentaremos el concepto de análisis multiresolución a partir de bases ortonormales wavelet y definiremos algunos operadores relacionados con las descomposiciones en bases wavelet. En los capítulos siguientes, estos conceptos se incorporarán al tratamiento de problemas inversos mal condicionados, donde los datos están contaminados con algún nivel de ruido. Probaremos, en el Capítulo 3, que aplicar umbralado wavelet a los datos, previo al proceso de regularización, constituye una herramienta eficaz para mejorar el desempeño de todos los métodos de regularización espectrales.

En 1988 Stéphane Mallat desarrolló una teoría que permite descomponer una función mediante otras funciones, denominadas de aproximación y detalle, utilizando bases ortonormales wavelet en $L^2(\mathbb{R}^d)$ (ver § 2.12 [10], [30]). En un contexto general aplicado a señales o imágenes,

los rasgos más importantes que determinan la identidad de una señal se sitúan en las bajas frecuencias, mientras que las altas frecuencias están generalmente asociadas con el ruido y con características correspondientes a ciertos detalles mas finos. Por este motivo, en el análisis multiresolución se habla de aproximación y detalle para referirse a las componentes de baja frecuencia y a las de altas frecuencias, respectivamente.

1.4.1. Análisis multiresolución

Por simplicidad introducimos la noción de análisis multiresolución en $L^2(\mathbb{R})$ propuesta por Mallat en [30], para luego realizar la extensión correspondiente a dominios en \mathbb{R}^d .

Definición 1.15. Un *análisis multiresolución* en $L^2(\mathbb{R})$ es una sucesión de subespacios cerrados $\{V_j\}_{j \in \mathbb{Z}}$, tales que se satisfacen las siguientes propiedades:

- La sucesión es creciente, es decir, para todo $j \in \mathbb{Z}$ vale que $V_j \subset V_{j+1}$;
- $\overline{\cup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$ y $\cap_{j \in \mathbb{Z}} V_j = \{0\}$;
- Los espacios V_j son autosimilares, es decir, $f(\cdot) \in V_0 \Leftrightarrow f(2^j \cdot) \in V_j$;
- Existe una función θ tal que $\{\theta(\cdot - k)\}_{k \in \mathbb{Z}}$ es una base de Riesz de V_0 . Es decir, existen constantes $0 < C_1 < C_2$ tales que

$$C_1 \sum_{k \in \mathbb{Z}} |a_k|^2 \leq \left\| \sum_{k \in \mathbb{Z}} a_k \theta(\cdot - k) \right\|^2 \leq C_2 \sum_{k \in \mathbb{Z}} |a_k|^2$$

para toda sucesión $\{a_k\}_{k \in \mathbb{N}} \in \ell^2$.

Teorema 1.16. Si $\{V_j\}_{j \in \mathbb{Z}}$ es un análisis multiresolución en $L^2(\mathbb{R})$ y ϕ es una función, denominada de escala, cuya transformada de Fourier es

$$\hat{\phi}(\omega) = \frac{\hat{\theta}(\omega)}{(\sum_{n=-\infty}^{\infty} |\hat{\theta}(\omega + 2\pi n)|^2)^{\frac{1}{2}}},$$

entonces el conjunto $\{\phi_{j,k}(\cdot) = 2^{j/2} \phi(2^j \cdot - k)\}_{k \in \mathbb{Z}}$ es una base ortonormal del espacio vectorial V_j para todo $j \in \mathbb{Z}$.

Demostración. La demostración de este teorema se puede encontrar, por ejemplo, en el Apéndice B de [30]. ■

Por otro lado, se puede probar que si $\{V_j\}_{j \in \mathbb{Z}}$ es un análisis multiresolución, entonces existe una base ortonormal wavelet

$$\{\psi_{j,k}(\cdot) = 2^{j/2} \psi(2^j \cdot - k)\}_{j,k \in \mathbb{Z}} \quad (1.19)$$

de $L^2(\mathbb{R})$ tal que para cada j fijo $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ es una base ortonormal del complemento ortogonal de V_j en V_{j+1} , al que denotamos con W_j . En efecto, para cada $j \in \mathbb{Z}$ se verifica que $V_{j+1} = W_j \oplus V_j$. De esta forma podemos descomponer el espacio $L^2(\mathbb{R})$ en subespacios mutuamente ortogonales: $L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$. Luego, podemos expresar cualquier función $f \in L^2(\mathbb{R})$ como

$$f(x) = \sum_{j,k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x).$$

Por otro lado, dado $j_0 \in \mathbb{Z}$ fijo, la descomposición $L^2(\mathbb{R}) = V_{j_0} \oplus \bigoplus_{j \geq j_0} W_j$ nos permite obtener la denominada “representación multiescala”

$$f(x) = \sum_{k \in \mathbb{Z}} \langle f, \phi_{j_0,k} \rangle \phi_{j_0,k}(x) + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x),$$

donde para cada $k \in \mathbb{Z}$, $\langle f, \phi_{j_0,k} \rangle$ es el coeficiente de aproximación de f en la escala 2^{-j_0} y $\langle f, \psi_{j,k} \rangle$ para $j \geq j_0$ son los coeficientes de detalle.

Dado que en la práctica se trabaja siempre con un número finito de detalles, resulta pertinente definir el operador proyección ortogonal $P_j : L^2(\mathbb{R}) \xrightarrow{\perp} V_j$, el cual se define de manera natural como

$$P_j f(x) = \sum_{k \in \mathbb{Z}} \langle f, \phi_{j,k} \rangle \phi_{j,k}(x),$$

o utilizando la representación multiescala, donde $j_0 \leq j$, como

$$P_j f(x) = \sum_{k \in \mathbb{Z}} \langle f, \phi_{j_0,k} \rangle \phi_{j_0,k}(x) + \sum_{i=j_0}^j \sum_{k \in \mathbb{Z}} \langle f, \psi_{i,k} \rangle \psi_{i,k}(x).$$

Por simplicidad en la notación, en general consideraremos $j_0 = 0$.

En esta tesis estamos particularmente interesados en funciones definidas en dominios acotados contenidos en \mathbb{R}^d , por lo cual definimos la versión multivariable para las funciones ϕ y ψ que nos permitirán obtener bases ortonormales para los correspondientes espacios V_j y W_j .

Para Ω un dominio acotado en \mathbb{R}^d , definimos a la función de escala multivariada como

$$\phi(\mathbf{x}) \doteq \prod_{i=1}^d \phi(x_i) \quad \text{para todo } \mathbf{x} = (x_1, \dots, x_d) \in \Omega.$$

Para $j \in \mathbb{Z}$, el espacio V_j correspondiente será entonces el generado por las funciones

$$\phi_{j,\mathbf{k}}(\mathbf{x}) = 2^{\frac{j d}{2}} \phi(2^j \mathbf{x} - \mathbf{k}), \quad \forall \mathbf{k} \in \mathbb{Z}^d, \forall \mathbf{x} \in \Omega,$$

y el operador proyección ortogonal $P_j : L_2(\Omega) \xrightarrow{\perp} V_j$ queda entonces definido por

$$P_j f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \phi_{j,\mathbf{k}} \rangle \phi_{j,\mathbf{k}}(\mathbf{x}).$$

Luego, para todo $j_0 \leq j$ y para toda $f \in V_j$ se verifica que $P_j f = P_{j_0} f + \sum_{j_0 \leq \ell \leq j-1} g_\ell$ donde $g_\ell = P_{\ell+1} f - P_\ell f$. Además, tomando límite para $j \rightarrow \infty$ obtenemos que para toda $f \in L^2(\Omega)$ se cumple que $f = \lim_{j \rightarrow \infty} \left[P_{j_0} f + \sum_{j_0 \leq \ell \leq j-1} g_\ell \right]$ (ver **Observación 2.6.2** en [10]).

Para el caso de la función ψ , la extensión a \mathbb{R}^d resulta un poco más compleja pues necesitamos de $2^d - 1$ funciones “wavelets madre” para representar cada W_j . En este contexto denotamos a cada una de estas funciones por

$$\psi^{(\epsilon)}(\mathbf{x}) \doteq \prod_{i=1}^d \psi^{\epsilon_i}(x_i), \quad \epsilon = (\epsilon_1, \dots, \epsilon_d) \in \Gamma \doteq \{0, 1\}^d - \{\mathbf{0}\},$$

donde $\psi^0(x) \doteq \phi(x)$ y $\psi^1(x) \doteq \psi(x)$. Notemos que se excluye el caso en el que ϵ es el vector nulo pues tendríamos $\psi^0(\mathbf{x}) = \phi(\mathbf{x})$.

Utilizando esta notación, la versión multivariable para la representación multiescala de una función $f \in L^2(\Omega)$ viene dada por

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \phi_{j_0, \mathbf{k}} \rangle \phi_{j_0, \mathbf{k}}(\mathbf{x}) + \sum_{j \geq j_0} \sum_{\epsilon \in \Gamma} \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{j, \mathbf{k}}^{(\epsilon)} \rangle \psi_{j, \mathbf{k}}^{(\epsilon)}(\mathbf{x}). \quad (1.20)$$

Observemos que definiendo $\Gamma^0 \doteq \Gamma \cup \{\mathbf{0}\} = \{0, 1\}^d$ y $\psi_{j, \mathbf{k}}^0 \doteq \phi_{j_0, \mathbf{k}}$ es posible escribir la expresión anterior de manera mas simple como una única sumatoria, en la forma:

$$f(\mathbf{x}) = \sum_{j \geq j_0} \sum_{\epsilon \in \Gamma^0} \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{j, \mathbf{k}}^{(\epsilon)} \rangle \psi_{j, \mathbf{k}}^{(\epsilon)}(\mathbf{x}). \quad (1.21)$$

Para un desarrollo más extenso sobre la versión multivariable de bases wavelet ver, por ejemplo, [10] § 2.12.

Por simplicidad de aquí en adelante utilizaremos la notación $\mathbf{x} = x$ y $\mathbf{k} = k$, entendiendo que en todos los casos donde se utilice esta notación quedará explícito si se trata de dominios en \mathbb{R}^d para $d \geq 2$ ó $d = 1$. Además, escribiremos la representación multiescala (1.21) en la forma mas simple:

$$f(x) = \sum_{\lambda \in \Lambda_{(j_0)}} \langle f, \psi_\lambda \rangle \psi_\lambda(x), \quad (1.22)$$

donde $\lambda = (j, k, \epsilon)$ y $\Lambda_{(j_0)} \doteq \{(j, k, \epsilon) \in \mathbb{Z} \times \mathbb{Z}^d \times \Gamma^0 : j \geq j_0 \wedge \text{sop}(\psi_\lambda) \subset \Omega\}$ para $j_0 \in \mathbb{Z}$ fijo. Dado que j_0 será prefijado al definir la descomposición wavelet a utilizar (en la mayoría de los casos consideramos $j_0 = 0$), a los efectos de simplificar aún mas la notación, obviaremos especificar tal dependencia denotando $\Lambda_{(j_0)}$ simplemente con Λ .

Por último introduciremos algunos conceptos que serán necesarios más adelante: el grado de un polinomio y el grado de reproducción polinomial u orden de exactitud polinomial de la

función de escala ϕ . En \mathbb{R}^d decimos que un polinomio tiene grado m cuando es combinación lineal de las funciones base $x_1^{k_1} \dots x_d^{k_d}$ para $k_1 + \dots + k_d \leq m$ y al menos uno de los coeficientes asociados a un factor $x_1^{k_1^*} \dots x_d^{k_d^*}$ para el cual $k_1^* + \dots + k_d^* = m$, es distinto de cero. A su vez, cuando hacemos referencia al grado de reproducción polinomial decimos que, en términos generales, una función ϕ tiene grado de reproducción polinomial m si todos los polinomios de grado menor o igual a m pertenecen al espacio V_0 correspondiente a ϕ . Más precisamente:

Definición 1.17. Sea $\phi \in L^1(\mathbb{R})$ una función de soporte compacto tal que $\|\phi\|_{L^1(\mathbb{R})} = 1$. Decimos que ϕ tiene *grado de reproducción polinomial* m si para todo $q = 0, 1, \dots, m$ podemos expresar el monomio x^q como

$$x^q = \sum_{k \in \mathbb{Z}} [k^q + p_{q-1}(k)] \phi(x - k),$$

donde p_{q-1} es un polinomio de grado $q - 1$.

1.4.2. Umbralado wavelet

Como mencionamos al comienzo de esta sección, el procedimiento conocido como umbralado wavelet es comúnmente utilizado en la reconstrucción de funciones afectadas por algún tipo de perturbación o ruido. Consiste en la modificación de los coeficientes wavelet de la función cuyos valores absolutos son cercanos a cero. Estos coeficientes se corresponden con las componentes de alta frecuencia, comúnmente asociadas al ruido. La ventaja de este procedimiento respecto de un filtrado por bandas de frecuencia reside en que se obtiene una función casi libre de ruido, modificando mínimamente las características de la función original. Resulta ser una herramienta más potente que los métodos tradicionales de suavizado, que solo consiguen eliminar el ruido a costa de suavizar también la función original.

En la literatura son bien conocidas dos clases de umbralado wavelet:

- El *umbralado suave (o débil)*, el cual descarta (hace cero) los coeficientes wavelet cuyos valores absolutos son menores o iguales que el umbral y sustrae el umbral de los coeficientes cuyo valor absoluto es mayor.
- El *umbralado fuerte*, en cambio, simplemente descarta (hace cero) los coeficientes menores que el umbral, en la expansión en base wavelet en términos de los elementos de V_j .

Con el umbralado suave, todas las componentes de la señal sufren atenuación, independientemente del valor del umbral. Por lo tanto, este criterio parece menos útil que el de umbralado

fuerte cuando el objetivo es solo eliminar ruido de alta frecuencia. A lo largo de este trabajo utilizaremos entonces umbralado fuerte.

Dado un umbral $\gamma \in \mathbb{R}^+$ denotamos con $S_\gamma : V_j \rightarrow V_j$ al operador de umbralado wavelet fuerte definido para cada nivel j mediante

$$S_\gamma f(x) = \sum_{k \in \mathbb{Z}^d} \beta_{j,k} \chi_{\{|\beta_{j,k}| \geq \gamma\}}(\beta_{j,k}) \phi_{j,k}(x), \quad (1.23)$$

donde $\beta_{j,k} \doteq \langle f, \phi_{j,k} \rangle$ y χ denota la función característica o indicatriz.

El desempeño de este proceso de filtrado depende, como es lógico, de la elección del umbral. El mismo debe elegirse para cada nivel de resolución y en su elección deben tenerse en cuenta tanto las características de la representación wavelet de la función, como el nivel de ruido que esta posee.

1.5. Espacios de Sobolev y de Besov

Como es usual en problemas inversos mal condicionados, para estudiar la convergencia de los métodos de regularización que propondremos en los siguientes capítulos, es necesario hacer ciertos supuestos de suavidad con respecto a la solución exacta de los mismos. En este contexto, y para los métodos que desarrollaremos en los Capítulos 3 y 4, resultan apropiados dos subespacios vectoriales de $L^p(\mathbb{R}^d)$: los espacios de Sobolev y de Besov que presentaremos a continuación.

Existe una gran variedad de formas diferentes de medir la suavidad de una función f . La más natural es, quizás, el orden de diferenciabilidad, es decir, el mayor índice $m \in \mathbb{N}$ para el cual la derivada $f^{(m)}$ es continua, o con más generalidad $\partial^\alpha f$ con $|\alpha| = \alpha_1 + \dots + \alpha_d \leq m$ para una función multivariable definida en \mathbb{R}^d . Esta medida particular de suavidad se asocia a la clase de espacios funcionales denotados con $C^m(\mathbb{R}^d)$, el espacio de funciones continuas que tienen derivadas parciales $\partial^\alpha f$ acotadas y continuas, para $|\alpha| \leq m$. Otra manera de cuantificar las propiedades de suavidad de una función (en un sentido menos estricto), es a través de los espacios de Sobolev, los cuales introducimos a continuación.

En términos generales un espacio de Sobolev $W^{m,p}(\Omega)$, para $\Omega \subset \mathbb{R}^d$, es el subespacio de $L^p(\Omega)$, formado por todas las funciones tales que sus derivadas de hasta orden m pertenecen también a $L^p(\Omega)$.

Definición 1.18. Sean $\Omega \subset \mathbb{R}^d$, $1 \leq p \leq \infty$ y $m \in \mathbb{N}$. El *espacio de Sobolev* $W^{m,p}(\Omega)$ se define como el conjunto de todas las funciones $f \in L^p(\Omega)$ tales que

$$\|f\|_{W^{m,p}(\Omega)} \doteq \|f\|_{L^p(\Omega)} + |f|_{W^{m,p}(\Omega)} < \infty,$$

donde $|f|_{W^{m,p}(\Omega)} \doteq \sum_{|\alpha|=m} \|\partial^\alpha f\|_{L^p(\Omega)}$.

Para el caso particular en el que $p = 2$, los espacios de Sobolev se dotan de manera natural de la estructura de espacio de Hilbert al igual que el espacio L^2 y es común en la literatura utilizar la notación $H^m(\Omega) \equiv W^{m,2}(\Omega)$.

Así como los espacios de Sobolev permiten de alguna forma cuantificar ciertas propiedades de suavidad de una función, existe una gran variedad de espacios más generales que incluyen otros parámetros a tener en cuenta. A lo largo de este trabajo nos interesará considerar espacios que se originan mediante “técnicas de interpolación” generales entre espacios funcionales de suavidad entera, espacios que pueden ser caracterizados exactamente por la tasa de error de aproximación multiresolución, así como por propiedades de tamaño de los coeficientes wavelet. Los espacios que utilizaremos, que poseen estas propiedades, son los espacios de Besov $B_{p,q}^s$ que utilizan un parámetro extra q y se pueden definir a través de diferencias finitas. En efecto, el espacio de Besov $B_{p,q}^s(\Omega)$, o equivalentemente $B_q^s(L^p(\Omega))$, es un espacio quasi-normado completo que resulta ser un espacio de Banach en Ω cuando $0 < p, q \leq \infty$. Generalizan espacios de funciones más elementales como los espacios de Sobolev y resultan ser más efectivos para cuantificar propiedades de regularidad de las funciones. Para definir estos espacios necesitaremos introducir previamente cierta notación y algunas definiciones, las que presentamos a continuación.

Sean $f \in L^p(\Omega)$ y $r \in \mathbb{N}$. Denotemos con $\omega_r(f, t)_p$ al r -ésimo módulo de continuidad de f definido como

$$\omega_r(f, t)_p \doteq \omega_r(f, t, \Omega)_p \doteq \sup_{|h| \leq t} \|\Delta_h^r(f, \cdot, \Omega)\|_{L^p(\Omega)}, \quad t > 0,$$

donde $\Delta_h^r(f, \cdot, \Omega)$ es la r -ésima diferencia positiva de f relativa a Ω , definida por

$$\Delta_h^r(f, x, \Omega) \doteq \begin{cases} \Delta_h^r(f, x) & \text{si } x, x+h, \dots, x+rh \in \Omega, \\ 0 & \text{en otro caso,} \end{cases}$$

donde $\Delta_h^r(f, x) \doteq \sum_{\ell=0}^r (-1)^{r-\ell} \binom{r}{\ell} f(x + \ell h)$.

Consideremos ahora $s > 0$, $r = \lfloor s + 1 \rfloor$ (la parte entera de $s + 1$) y $0 < p, q \leq \infty$. Definimos la siguiente seminorma

$$|f|_{B_{p,q}^s(\Omega)} \doteq \left(\sum_{k=1}^{\infty} [2^{ks} \omega_r(f, 2^{-k}, \Omega)_p]^q \right)^{\frac{1}{q}},$$

con las correspondiente modificaciones para el caso $q = \infty$, es decir, $|f|_{B_{p,\infty}^s(\Omega)} \doteq \sup_{k \in \mathbb{N}} 2^{ks} \omega_r(f, 2^{-k}, \Omega)_p$.

Utilizando esta notación y definiciones estamos ahora en condiciones de definir los espacios de Besov.

Definición 1.19. Sean $s > 0$, $r = \lfloor s + 1 \rfloor$ y $0 < p, q \leq \infty$. El *espacio de Besov* $B_{p,q}^s(\Omega)$ se define como el conjunto de todas las funciones $f \in L^p(\Omega)$ para las cuales la siguiente norma es finita

$$\|f\|_{B_{p,q}^s(\Omega)} \doteq \|f\|_{L^p(\Omega)} + |f|_{B_{p,q}^s(\Omega)}.$$

Es oportuno e importante resaltar la relación entre los espacios que acabamos de definir. Si bien en este caso definimos los espacios de Sobolev solo para $m \in \mathbb{N}_0$, existe una generalización $W^{s,p}(\Omega)$ para $m < s < m + 1$, a partir de la cual se puede probar que $W^{s,p}(\Omega) = B_{p,q}^s(\Omega)$ cuando s no es un entero (ver, por ejemplo, [10], § 3.2). En el caso particular con el que trabajaremos más adelante, en el que s es entero, esta igualdad se da solo para el caso $p = 2$ ($B_{2,2}^s(\Omega) = W^{s,2}(\Omega) = H^s(\Omega)$). Más aún se puede probar que los espacios de Sobolev $W^{s,p}(\Omega)$ no son espacios de Besov para valores de s enteros si $p \neq 2$ (ver [46]).

1.5.1. Análisis multiresolución y los espacios de Besov débiles

En las últimas dos décadas, autores como DeVore ([12]), Donoho y Johnstone ([16]) y Cohen et. al. ([9]) encontraron relaciones entre los espacios ℓ^p débiles, denotados por $w\ell^p$, y la sucesión de subespacios cerrados de un análisis multiresolución en L^2 . Particularmente, Cohen y otros autores ([9]) vincularon el enfoque de la aproximación no lineal no solo a los espacios débiles ℓ^p sino también a los espacios débiles de Besov que pueden ser vistos como espacios débiles ponderados ℓ^p (ver [8]). Mas aún, mostraron que los espacios débiles de Besov aparecen en la caracterización de la aproximación del umbral wavelet, en el proceso de umbralado que describimos en la sección anterior. Siguiendo esta línea, en el análisis multiresolución y en el estudio de órdenes de convergencia para aproximaciones realizadas mediante umbralado wavelet trabajaremos con espacios de Besov débiles. Por esta razón, es necesario introducir los siguientes conceptos y resultados básicos sobre el tema.

Supongamos que tenemos una base ortogonal generada por las funciones wavelet ψ y de escala ϕ , como en la sección anterior, de manera que podemos expresar cualquier función $f \in L^2(\Omega)$ como

$$f(x) = \sum_{\lambda \in \Lambda} \beta_\lambda \psi_\lambda(x), \quad (1.24)$$

donde $\beta_\lambda = \langle f, \psi_\lambda \rangle$, $\lambda = \lambda(j, k, \epsilon)$ y $\Lambda \doteq \{(j, k, \epsilon) \in \mathbb{Z} \times \mathbb{Z}^d \times \Gamma^0 : j \geq j_0 \text{ y } \text{supp}(\psi_\lambda) \subset \Omega\}$ para algún j_0 fijo.

El siguiente lema es un resultado importante que caracteriza a los espacios de Besov en términos de propiedades de sus coeficientes wavelet.

Lema 1.20. *Dados $0 < p, q \leq \infty$ y $0 < s < \infty$, si $\psi \in L^r(\Omega)$ para algún $r \in [1, \infty]$ tal que $0 < p \leq r$ entonces la función f (como en (1.24)) pertenece a $B_{p,q}^s(\Omega)$ si, y solo si, los coeficientes $\{\beta_\lambda\}_\lambda$ verifican*

$$\left(\sum_{j_0 \leq j} 2^{jq(s + \frac{d}{2} - \frac{d}{p})} \left(\sum_{\lambda \in \Lambda_j} |\beta_\lambda|^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}} < \infty, \quad (1.25)$$

con las obvias modificaciones para $p = \infty$ y $q = \infty$.

Demostración. Ver Teorema 30.7 en [7]. ■

Para simplificar la notación, de aquí en más denotaremos con $\Theta_{p,q}^s = \Theta_{p,q}^s(C)$ (para C una constante positiva dada) al conjunto:

$$\Theta_{p,q}^s(C) \doteq \left\{ \{\beta_\lambda\} : \sum_{j_0 \leq j} 2^{jq(s + \frac{d}{2} - \frac{d}{p})} \left(\sum_{\lambda \in \Lambda_j} |\beta_\lambda|^p \right)^{\frac{q}{p}} \leq C^q \right\}. \quad (1.26)$$

Definición 1.21. Sean $0 < s, p, q < \infty$. Decimos que f (como en (1.24)) pertenece al *espacio de Besov débil* $WB_{p,q}^s(\Omega)$ si el siguiente supremo es finito

$$\sup_{\sigma > 0} \left[\sigma^q \sum_{j \geq j_0} 2^{jq(s + \frac{d}{2} - \frac{d}{p})} N(j, \sigma)^{\frac{q}{p}} \right],$$

donde $N(j, \sigma) \doteq \#\{\beta_{j,k}^{(\epsilon)} \mid k \in \mathbb{Z}^d, \forall \epsilon \in \Gamma^0 : |\beta_{j,k}^{(\epsilon)}| > \sigma\}$.

Se puede probar fácilmente, utilizando la desigualdad de Markov (ver [23] pág. 123) y la caracterización de espacios de Besov (dada por (1.25)), que el espacio de Besov está contenido en el correspondiente espacio débil de Besov, es decir $B_{p,q}^s(\Omega) \subset WB_{p,q}^s(\Omega)$.

1.6. Ruido blanco

En esta sección presentamos brevemente el andamiaje teórico necesario para definir el proceso estocástico denominado ruido blanco. Este concepto será fundamental a la hora de estudiar el modelo estocástico que supondremos para los datos observados en los problemas inversos mal condicionados que consideramos en esta tesis. En primer lugar, recordamos algunos conceptos básicos relacionados con espacios de probabilidad y variables aleatorias. Posteriormente, analizamos conceptos asociados a distribuciones condicionales, especificando en particular que se entiende por esperanza y varianza condicionales. Finalizamos la sección introduciendo los conceptos básicos relacionados a un proceso estocástico y, en particular, el proceso de movimiento browniano y su relación con el ruido blanco.

1.6.1. Definiciones básicas

Comenzamos la sección definiendo los conceptos de σ -álgebra y medida de probabilidad que nos permitirán posteriormente presentar el concepto de variable aleatoria en un espacio de probabilidad.

Definición 1.22. Dado un conjunto no vacío Ω , decimos que una familia \mathcal{A} de subconjuntos de Ω es una σ -álgebra sobre Ω si satisface las siguientes condiciones:

1. $\Omega \in \mathcal{A}$;
2. Si $A \in \mathcal{A}$ entonces $A^c \in \mathcal{A}$;
3. Si $A_n \in \mathcal{A}$ para $n = 1, 2, 3, \dots$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

En el caso que \mathcal{A} sea una σ -álgebra sobre Ω , llamaremos al par (Ω, \mathcal{A}) *espacio de medida*.

Definición 1.23. Sea Ω un conjunto no vacío y \mathcal{A} una σ -álgebra sobre Ω . Decimos que una función $P : \mathcal{A} \rightarrow [0, 1]$ es una *medida de probabilidad* en o sobre \mathcal{A} (o simplemente una *probabilidad* en \mathcal{A}) si satisface las siguientes condiciones, llamadas axiomas de Kolmogorov:

1. $P(A) \geq 0, \forall A \in \mathcal{A}$ (no negatividad);
2. $P(\Omega) = 1$ (homogeneidad);
3. Si $A_1, A_2, \dots \in \mathcal{A}$ son disjuntos de a pares, entonces $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$ (σ -aditividad).

Definición 1.24. Un *espacio de probabilidad* es una terna (Ω, \mathcal{A}, P) donde

- ★ Ω es un conjunto no vacío,
- ★ \mathcal{A} es una σ -álgebra de subconjuntos de Ω , y
- ★ P es una medida de probabilidad sobre \mathcal{A} .

La σ -álgebra \mathcal{A} suele referirse como “la σ -álgebra de eventos aleatorios”.

Estamos ahora en condiciones de definir el concepto de variable aleatoria.

Definición 1.25. Una *variable aleatoria* X en un espacio de probabilidad (Ω, \mathcal{A}, P) es una función real definida sobre Ω tal que $\forall x \in \mathbb{R} [X \leq x] \doteq \{\omega \in \Omega : X(\omega) \leq x\}$ es un evento aleatorio para todo $x \in \mathbb{R}$. Es decir, $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria si $[X \leq x] \in \mathcal{A}$ para todo $x \in \mathbb{R}$.

Los siguientes conceptos nos permitirán caracterizar y estudiar algunas propiedades de variables aleatorias que utilizaremos más adelante.

Definición 1.26. La *función de distribución acumulada o acumulativa* de una variable aleatoria X , $F_X : \mathbb{R} \rightarrow [0, 1]$, se define por $F_X(x) = P([X \leq x]) \doteq P(X \leq x)$, $x \in \mathbb{R}$.

Definición 1.27. Sea X una variable aleatoria cualquiera y F_X su función de distribución acumulativa. La *esperanza* de X se define

$$E(X) \doteq \int_{-\infty}^{\infty} x dF_X(x),$$

siempre que esta integral esté bien definida. En el caso que esta integral sea finita decimos que X es integrable.

Sean X una variable aleatoria, $b \in \mathbb{R}$ y $k \in \mathbb{N}$. El valor $E[(X - b)^k]$, si existe, se llama *k-ésimo momento de X alrededor de b* . Si X es integrable entonces el *k-ésimo momento de X alrededor de la media*, $E[(X - E(X))^k]$, se denomina *k-ésimo momento central de X* . El segundo momento central de X se denomina *varianza* de X :

$$Var(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2.$$

Para $k > 0$, $E(|X|^k)$ es el *k-ésimo momento absoluto de X* .

A continuación recordamos algunos conceptos relacionados con distribuciones condicionales.

Definición 1.28. Sean X e Y variables aleatorias definidas en un mismo espacio de probabilidad (Ω, \mathcal{A}, P) . Una función $P(X \in B|Y = y)$, definida para B boreliano e $y \in \mathbb{R}$ se llamará *distribución condicional (regular) de X dada Y* si:

- ★ para todo $y \in \mathbb{R}$ fijo, $P(X \in B|Y = y)$ define una probabilidad en \mathcal{B} , la σ -álgebra de Borel en la recta, y
- ★ para todo $B \in \mathcal{B}$ fijo, $P(X \in B|Y = y)$ es una función medible de y y

$$\int_{-\infty}^y P(X \leq x|Y = t) dF_Y(t) = P(X \leq x|Y \leq y) \quad \forall (x, y) \in \mathbb{R}^2.$$

Definición 1.29. Sean X e Y variables aleatorias en (Ω, \mathcal{A}, P) . La *esperanza condicional* de X dado $Y = y$, es la esperanza de la distribución condicional de X dado $Y = y$, si esta esperanza existe. Es decir,

$$E(X|Y = y) = \int x dF_X(x|Y = y).$$

La *varianza condicional* es el momento de segundo orden condicional centrado en la media condicional, es decir:

$$Var(X|Y) \doteq E\left(\left(X - E(X|Y)\right)^2|Y\right) = E(X^2|Y) - (E(X|Y))^2.$$

1.6.2. Procesos estocásticos

En esta sección definiremos el concepto de *proceso estocástico* y dos casos particulares que nos resultarán de utilidad en adelante (Capítulos 3 y 4): el movimiento browniano y el ruido blanco.

Definición 1.30. Sean $D \subseteq \mathbb{R}$ y (Ω, \mathcal{A}, P) un espacio de probabilidad. Un *proceso estocástico* sobre D en (Ω, \mathcal{A}, P) es una función $X : D \times \Omega \rightarrow \mathbb{R}$ tal que para todo $t \in D$ la función $X(t, \cdot) : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria sobre (Ω, \mathcal{A}, P) .

Un caso particular de proceso estocástico es el llamado *movimiento Browniano*. Si bien existen varias definiciones e interpretaciones físicas del movimiento browniano, a lo largo de este trabajo utilizaremos la siguiente definición formal:

Definición 1.31. Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , un proceso estocástico $W : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ se denomina *movimiento browniano estándar* si satisface las siguientes condiciones:

1. $P(W(0, \cdot) = 0) = 1$;
2. $W(\cdot, \omega)$ tiene incrementos independientes. Es decir, para todos $0 \leq t_0 < t_1 < t_2 < \dots < t_n$, las siguientes variables aleatorias son independientes:

$$W(t_1, \cdot) - W(t_0, \cdot), W(t_2, \cdot) - W(t_1, \cdot), \dots, W(t_n, \cdot) - W(t_{n-1}, \cdot); \quad (1.27)$$

3. Para cualesquiera $0 < s < t$ la variable aleatoria $W(t, \cdot) - W(s, \cdot)$ tiene distribución normal. Más aún

$$W(t, \cdot) - W(s, \cdot) \sim N(0, t - s); \quad (1.28)$$

4. Dado $\omega \in \Omega$ fijo $P(\{\omega : \text{la función } t \rightarrow W(t, \omega) \text{ es continua}\}) = 1$.

A continuación introducimos el concepto de ruido blanco.

Definición 1.32. Un proceso de ruido blanco X es la “derivada formal” de un movimiento browniano W , donde $X = \frac{\partial W}{\partial t}$ debe ser considerado como un funcional que actúa sobre funciones continuamente diferenciables como sigue:

$$\int_a^b f(t) X(t, \cdot) dt = \int_a^b f(t) dW(t) = f(t) W(t, \cdot) \Big|_a^b - \int_a^b W(t, \cdot) df(t). \quad (1.29)$$

El lector interesado puede ver detalles formales por ejemplo en las referencias [24] y [27].

A continuación, probamos un lema que nos permitirá en el Capítulo 4 estudiar los efectos del agregado de ruido blanco a los datos observados de problemas inversos mal condicionados.

Lema 1.33. *Dados $g \in L^2(\Omega)$ y dW ruido blanco, entonces $W_g = \langle g, dW \rangle$ es una variable aleatoria normal con media nula.*

Demostración. La prueba de que W_g tiene distribución normal se sigue esencialmente de la hipótesis de incrementos independientes del movimiento browniano (ver, por ejemplo, [14]).

Para probar que W_g tiene media nula notemos en primer lugar que

$$\begin{aligned} W_g &= \int_0^\infty g(t) dW(t, \omega) dt \\ &= g(t)W(t, \omega) \Big|_0^\infty - \int_0^\infty W(t, \omega) dg(t) \\ &= - \int_0^\infty W(t, \omega) dg(t). \end{aligned} \tag{1.30}$$

Luego, de (1.30) y dado que el movimiento browniano tiene media nula, tenemos que

$$\begin{aligned} E(W_g) &= \int_\Omega \omega \int_0^\infty g(t) dW(t, \omega) dt d\omega \\ &= \int_\Omega \omega \left(- \int_0^\infty W(t, \omega) dg(t) \right) d\omega \\ &= - \int_0^\infty \int_\Omega \omega W(t, \omega) d\omega dg(t) \\ &= 0. \end{aligned} \quad \blacksquare$$

CAPÍTULO 2

Introducción a los problemas inversos y métodos de regularización

A lo largo de este capítulo presentaremos definiciones y resultados básicos que nos permitirán introducirnos al mundo de los problemas inversos mal condicionados y los métodos de regularización. En la Sección 2.1 recordaremos el concepto de mal condicionamiento en el sentido de Hadamard y desarrollaremos brevemente un ejemplo correspondiente a la diferenciación. En la sección siguiente presentamos el concepto de regularización y enunciamos algunas de sus propiedades. En particular, estudiaremos los métodos de regularización espectrales, de Tikhonov generalizados y mencionamos algunos ejemplos tradicionales en la literatura. En la Sección 2.3 definiremos algunos tipos de errores que surgen al regularizar un problema e introducimos la noción de optimalidad. En la Sección 2.4 presentaremos órdenes de convergencia y cotas para los distintos tipos de errores e introduciremos el concepto clásico de calificación necesario para la correcta interpretación de hipótesis y resultados en los próximos capítulos. En la Sección 2.5 presentaremos los conceptos de reglas de elección de parámetros *a priori* y *a posteriori*, analizando en particular el principio de discrepancia de Morozov, que utilizaremos en repetidas oportunidades a lo largo de esta tesis. Finalmente, en la última sección se retoman algunos ejemplos de métodos de regularización espectrales relacionándolos con conceptos ya desarrollados tales como descomposición en valores singulares y calificación.

2.1. Problemas inversos mal condicionados

En un contexto bastante general, se puede presentar un problema inverso como la necesidad de determinar x en una ecuación de la forma

$$Tx = y, \tag{2.1}$$

donde T es un operador lineal y acotado entre dos espacios de Hilbert X e Y , e y es el dato conocido. Por simplicidad, aquí consideraremos solo problemas lineales, aunque muchas de las definiciones, conceptos y resultados que se presentan a continuación siguen siendo válidos o tienen un correspondiente correlato en el caso de problemas no lineales de la forma $Fx = y$, en

lugar de (2.1). Con respecto al buen o mal condicionamiento de un problema, esto hace referencia al cumplimiento de los tres postulados de Hadamard de buen condicionamiento (*well-posedness*) que enunciamos a continuación:

- ★ Existencia: Para todo dato admisible, existe una solución.
- ★ Unicidad: Para todo dato admisible, la solución es única.
- ★ Dependencia continua: La solución depende continuamente de los datos.

Cualquier problema matemático que viole al menos uno de estos postulados se dice “mal condicionado” (*ill-posed*). Es importante señalar que los problemas inversos, especialmente en dimensión infinita, frecuentemente no satisfacen al menos uno de tales postulados.

En el contexto del problema (2.1) se dice que el dato y es alcanzable si $y \in \mathcal{R}(T)$. Luego, el primer postulado es equivalente a que todo $y \in Y$ sea alcanzable o, equivalentemente, que $\mathcal{R}(T) = Y$. El segundo postulado, en cambio, se satisface si y solo si $\mathcal{N}(T) = \{0\}$. De la definición del operador inversa generalizada de Moore-Penrose dada en el capítulo anterior, es evidente que el incumplimiento de cualquiera de los dos primeros postulados puede solucionarse, en algún sentido, obteniendo una solución generalizada a través de este operador.

En la práctica, y por las razones antes mencionadas, la principal causa del mal condicionamiento de un problema inverso, es la violación del tercer postulado. Este incumplimiento del supuesto de dependencia continua de los datos se pone de manifiesto en la no acotación de la inversa generalizada de Moore-Penrose y se traduce en inestabilidad cuando se emplean métodos numéricos tradicionales para aproximar las soluciones. Para restablecer la estabilidad es necesaria la aplicación de *métodos de regularización*, los cuales consisten, esencialmente, en aproximar el problema mal condicionado por una familia convenientemente construida de problemas bien condicionados. Sin embargo, se debe tener siempre presente que ningún truco matemático puede hacer estable un problema que es intrínseca e inherentemente inestable. Todo lo que un método de regularización puede hacer es recuperar información parcial acerca de la solución, de modo tan estable como sea posible. Ciertamente el “arte” de aplicar métodos de regularización consiste en hallar el compromiso justo entre exactitud (confiabilidad) y estabilidad ([17]).

2.1.1. Ejemplo: la diferenciación como problema inverso

Integración y derivación son probablemente dos de los operadores, considerados inversos, más conocidos por su simplicidad y la gran variedad de aplicaciones al estudio de funciones. En principio, no es claro cuál de estos operadores corresponde al problema directo y cuál al

inverso. Sin embargo, como veremos en el siguiente ejemplo, es el operador diferenciación el que no satisface el supuesto de dependencia continua de los datos. En efecto, sean $y \in C^1([0, 1])$, $\delta \in (0, 1)$ y $n \in \mathbb{N}$ tal que $n \geq 2$. Definimos

$$y_n^\delta(t) = y(t) + \delta \operatorname{sen}\left(\frac{nt}{\delta}\right), \quad t \in [0, 1]. \quad (2.2)$$

Luego,

$$(y_n^\delta)'(t) = y'(t) + n \cos\left(\frac{nt}{\delta}\right), \quad t \in [0, 1]. \quad (2.3)$$

Claramente, $\|y - y_n^\delta\|_\infty = \delta$ pero $\|y' - (y_n^\delta)'\|_\infty = n$. Puesto que δ y n son arbitrarios, suponiendo que y e y_n^δ representan el dato exacto y el dato con ruido, respectivamente, se tiene que, para un nivel de error arbitrariamente pequeño (δ) en los datos, es posible obtener errores arbitrariamente grandes en las derivadas. A partir de este ejemplo se puede observar claramente cómo el operador derivada no depende continuamente de los datos con respecto a la norma infinito.

En relación a la formulación de los problemas directos e inversos, asociados a la ecuación $Tx = y$, notemos que en este caso, y' resuelve la clásica ecuación integral

$$(Tx)(t) = \int_0^t x(s) ds = y(t) - y(0),$$

en $C[0, 1]$ siempre que $y \in C^1[0, 1]$. El problema directo correspondiente consiste entonces en encontrar y a partir del dato x , es decir, integrando. Éste resulta ser un proceso estable en $C[0, 1]$. Notemos que la integración es un proceso de suavizado, donde los errores altamente oscilatorios en y (como, por ejemplo, errores de la forma $n \cos\left(\frac{nt}{\delta}\right)$ como el definido en (2.3)) son fuertemente atenuados (obteniendo $\delta \operatorname{sen}\left(\frac{nt}{\delta}\right)$) y tienen un efecto despreciable en los datos a los efectos del problema inverso. Este efecto de “suavizado” es el principal responsable de que errores de pequeña amplitud, pero de alta frecuencia en los datos, generen grandes oscilaciones en la solución del problema inverso. Es importante destacar que este tipo de errores de reconstrucción no se limitan a esta aplicación concreta: cada vez que un problema directo tiene propiedades de suavizado, se espera la aparición de oscilaciones provenientes de pequeñas perturbaciones de los datos (de alta frecuencia) en la solución del problema inverso. Este efecto será más evidente cuanto más fuerte sea el efecto de suavizado del operador asociado al problema directo.

2.2. Métodos de regularización

En términos generales, como mencionamos en la primera sección de este capítulo, regularizar un problema mal condicionado significa esencialmente aproximarlos por una sucesión de

problemas bien condicionados. En esta sección introduciremos rigurosamente el concepto de *regularización* y presentaremos algunas de sus propiedades básicas.

A partir del problema $Tx = y$ nos proponemos aproximar la mejor solución aproximada o *solución exacta* $x^\dagger \doteq T^\dagger y$ para un dato específico y . Para aplicaciones concretas, en general, el dato no se conoce de manera exacta puesto que proviene de mediciones. En cambio, se conoce solo un *dato aproximado* o *dato con ruido*, al que denotaremos con y^δ , y supondremos que satisface $\|y - y^\delta\| \leq \delta$, donde δ es el nivel de ruido asociado a los errores de medición. De la Sección 2.1 se tiene que, en el caso de un problema mal condicionado, la existencia de $T^\dagger y^\delta$, no implica que $T^\dagger y^\delta$ sea una buena aproximación a $T^\dagger y$ puesto que el operador T^\dagger no es acotado. Buscamos entonces una aproximación x_α^δ de x^\dagger que tenga dependencia continua del dato y^δ para obtener convergencia a la solución exacta x^\dagger cuando el nivel de ruido δ tienda a cero y el parámetro de regularización α se elija de manera adecuada. Intuitivamente, una regularización de T^\dagger consiste en aproximar el operador no acotado T^\dagger por una familia uniparamétrica de operadores continuos $\{R_\alpha\}$, con α en cierto conjunto de índices que precisaremos luego. De esta manera, para R_α y y^δ , una aproximación de x^\dagger viene dada por $x_\alpha^\delta \doteq R_\alpha y^\delta$. Un requerimiento para el parámetro de regularización α radica en que si el nivel de ruido tiende a cero entonces la *solución regularizada* x_α^δ debe converger a x^\dagger .

Definición 2.1. Sean $\alpha \in (0, \alpha_0)$ con $\alpha_0 \in (0, \infty]$ y $R_\alpha : Y \rightarrow X$ un operador continuo. Decimos que la familia $\{R_\alpha\}_{\alpha \in (0, \alpha_0)}$ es una *familia de operadores de regularización para T^\dagger* si, para todo $y \in \mathcal{D}(T^\dagger)$, existe una *regla de elección de parámetro* $\alpha = \alpha(\delta, y^\delta)$ tal que

$$\lim_{\delta \rightarrow 0^+} \sup_{\substack{y^\delta \in Y \\ \|y^\delta - y\| \leq \delta}} \|R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y\| = 0. \quad (2.4)$$

La regla de elección de parámetro α es una función $\alpha : \mathbb{R}^+ \times Y \rightarrow (0, \alpha_0)$ tal que

$$\lim_{\delta \rightarrow 0^+} \sup_{\substack{y^\delta \in Y \\ \|y^\delta - y\| \leq \delta}} \alpha(\delta, y^\delta) = 0. \quad (2.5)$$

Dado $y \in \mathcal{D}(T^\dagger)$, si se satisfacen (2.4) y (2.5), entonces el par $(\{R_\alpha\}, \alpha(\delta, y^\delta))$ se denomina *método de regularización (convergente)* para resolver el problema $Tx = y$.

De esta forma, un método de regularización consiste esencialmente en una familia de operadores de regularización $\{R_\alpha\}$ y una regla de elección de parámetro $\alpha(\delta, y^\delta)$. Es convergente en el sentido que, si el parámetro de regularización se elige de acuerdo a esa regla, entonces las soluciones regularizadas convergen a la solución exacta x^\dagger a medida que el nivel de ruido tiende

a cero. Notar que no requerimos que la familia de operadores de regularización sea una familia de operadores lineales.

Es oportuno mencionar aquí que existen dos tipos de reglas de elección de parámetro α :

- ★ *a priori*: α sólo depende del nivel de ruido δ , es decir $\alpha = \alpha(\delta)$;
- ★ *a posteriori*: α depende tanto del nivel de ruido δ como del dato y^δ , es decir: $\alpha = \alpha(\delta, y^\delta)$.

La siguiente proposición permite afirmar que se pueden construir operadores de regularización como aproximaciones puntuales de la inversa generalizada de Moore-Penrose de T .

Proposición 2.2. *Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$ y R_α un operador continuo para todo $\alpha > 0$. Si $R_\alpha y \rightarrow T^\dagger y$ para todo $y \in \mathcal{D}(T^\dagger)$ cuando $\alpha \rightarrow 0^+$, entonces $\{R_\alpha\}$ es una familia de operadores de regularización para T^\dagger . En este caso, para cada $y \in \mathcal{D}(T^\dagger)$ existe una regla de elección de parámetro a-priori, $\alpha(\delta)$, tal que $(\{R_\alpha\}, \alpha(\delta))$ es un método de regularización convergente para resolver $Tx = y$.*

Demostración. Ver [17], Proposición 3.4. ■

2.2.1. Métodos de regularización espectrales

En esta sección se presentan brevemente algunos resultados conocidos acerca de los métodos de regularización espectrales (cuyas demostraciones pueden encontrarse por ejemplo en [17]), los cuales serán de gran utilidad para probar algunos de los resultados de convergencia, en los Capítulos 3 y 4 de esta tesis.

Dados X, Y espacios de Hilbert y $T \in \mathcal{L}(X, Y)$, T^\dagger la inversa generalizada de Moore-Penrose de T , puede expresarse en términos de una integral con respecto a la familia espectral $\{E_\lambda^{T^*T}\}_{\lambda \in \mathbb{R}}$ asociada al operador autoadjunto T^*T (ver Proposición 1.13). Recordemos que denotamos con \mathcal{M}_0 al conjunto de funciones medibles con respecto a la medida $d\|E_\lambda^{T^*T}x\|^2$ para todo $x \in X$. Utilizando argumentos básicos de teoría espectral se puede probar que para todo $y \in \mathcal{D}(T^\dagger)$ se tiene que x^\dagger satisface la ecuación normal (1.6) y x^\dagger se puede representar entonces como

$$x^\dagger \doteq T^\dagger y = (T^*T)^{-1}T^*y = \int \frac{1}{\lambda} dE_\lambda^{T^*T}T^*y. \quad (2.6)$$

Con el objetivo de hallar una aproximación estable de la solución del problema inverso mal condicionado (2.1), la propuesta consiste en reemplazar el integrando no acotado $\frac{1}{\lambda}$ en (2.6) por una familia de funciones $\{g_\alpha(\lambda)\}$ continuas a trozos en $[0, \|T\|^2]$ y tal que para todo $\lambda \in (0, \|T\|^2]$

se tenga que $\lim_{\alpha \rightarrow 0^+} g_\alpha(\lambda) = \frac{1}{\lambda}$. Luego, definimos R_α como

$$R_\alpha \doteq \int_0^{\|T\|^2 +} g_\alpha(\lambda) dE_\lambda^{T^* T} = g_\alpha(T^* T) T^*. \quad (2.7)$$

Por construcción, el operador R_α resulta continuo y, por lo tanto, la solución regularizada a partir de un dato con ruido y^δ , se puede calcular ahora de manera estable mediante

$$x_\alpha^\delta \doteq R_\alpha y^\delta \doteq \int_0^{\|T\|^2 +} g_\alpha(\lambda) dE_\lambda^{T^* T} y^\delta.$$

Este operador R_α tiene su correlato en el contexto de operadores compactos, donde es posible expresar la integral en (2.7) en términos de sumas utilizando el sistema singular $\{(\sigma_n; u_n; v_n)\}$ de T (ver (1.10) en la Sección 1.2). En tal caso se tiene

$$x_\alpha^\delta = R_\alpha y^\delta = \sum_{n=1}^{\infty} g_\alpha(\sigma_n^2) \sigma_n \langle y^\delta, v_n \rangle u_n. \quad (2.8)$$

Las siguientes condiciones sobre $\{g_\alpha\}$ aseguran que la familia de operadores $\{R_\alpha\}$ dada por (2.7) es una familia de operadores de regularización para T^\dagger (ver Teorema 2.5).

Definición 2.3. Sean $\alpha_0 > 0$ y $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ una familia paramétrica de funciones $g_\alpha : (0, \alpha_0) \rightarrow \mathbb{R}$. Diremos que $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ es un *método de regularización espectral* (MRE), si satisface las siguientes condiciones:

(H1) Para todo $\alpha \in (0, \alpha_0)$, $g_\alpha(\lambda)$ es continua por tramos para $\lambda \in [0, \infty)$ y continua por derecha en los puntos de discontinuidad;

(H2) Existe una constante $C > 0$ (independiente de α) tal que $|\lambda g_\alpha(\lambda)| \leq C$ para todo $\lambda \in [0, \infty)$, para todo $\alpha \in (0, \alpha_0)$;

(H3) Para todo $\lambda \in (0, \infty)$, $\lim_{\alpha \rightarrow 0^+} g_\alpha(\lambda) = \frac{1}{\lambda}$.

Observación 2.4. Dado un operador T lineal y acotado, es suficiente con que $g_\alpha(\lambda)$ esté definida y satisfaga (H1)-(H3) en $[0, \|T\|^2]$ o, con más generalidad, en cualquier intervalo que contenga al espectro de $T^* T$.

El siguiente teorema afirma que si $\{g_\alpha\}$ es un MRE entonces R_α definido como en (2.7) es una regularización para T^\dagger . Bajo estas condiciones diremos que la familia $\{R_\alpha\}$ es una familia de regularización espectral para T^\dagger , puesto que los operadores R_α se definen en términos de una integral con respecto a la familia espectral $\{E_\lambda^{T^* T}\}_{\lambda \in \mathbb{R}}$ asociada al operador $T^* T$.

Teorema 2.5. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$ y $\{g_\alpha\}$ un MRE. Entonces, para todo $y \in \mathcal{D}(T^\dagger)$ se tiene que

$$\lim_{\alpha \rightarrow 0^+} g_\alpha(T^* T) T^* y = T^\dagger y.$$

Si $y \notin \mathcal{D}(T^\dagger)$, entonces $\lim_{\alpha \rightarrow 0^+} \|g_\alpha(T^*T)T^*y\| = \infty$.

A continuación enunciamos un teorema que proporciona una estimación para el error de regularización en términos de una acotación para la función r_α definida por

$$r_\alpha(\lambda) \doteq 1 - \lambda g_\alpha(\lambda), \quad (2.9)$$

bajo una cierta condición fuente.

Teorema 2.6. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$, $\{g_\alpha\}$ un MRE y r_α definida por (2.9). Si $x^\dagger \in X_{\mu^*, \rho} \doteq \{x \in X : x = (T^*T)^\mu w, \|w\| \leq \rho\}$ para $\mu^*, \rho > 0$ y $\omega_\mu : (0, \alpha_0) \rightarrow \mathbb{R}^+$ es una función que satisface, para $\mu = \mu^*$ y $\mu = \mu^* + \frac{1}{2}$,

$$\lambda^\mu |r_\alpha(\lambda)| \leq \omega_\mu(\alpha) \quad \forall \alpha \in (0, \alpha_0), \forall \lambda \geq 0, \quad (2.10)$$

entonces

$$\|x_\alpha - x^\dagger\| \leq \omega_{\mu^*}(\alpha)\rho \quad (2.11)$$

y

$$\|Tx_\alpha - Tx^\dagger\| \leq \omega_{\mu^* + \frac{1}{2}}(\alpha)\rho. \quad (2.12)$$

En el siguiente teorema se proporcionan cotas (en X y en Y) para el error asociado al ruido cuando se utiliza un MRE. Para $\alpha > 0$ definimos

$$G_\alpha \doteq \sup_{\lambda \in [0, \|T\|^2]} |g_\alpha(\lambda)|, \quad (2.13)$$

donde $\{g_\alpha\}$ es un MRE.

Teorema 2.7. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$, $\{g_\alpha\}$ un MRE y G_α como en (2.13). Si $y, y^\delta \in Y$ son tales que $\|y^\delta - y\| \leq \delta$, entonces

$$\|Tx_\alpha - Tx_\alpha^\delta\| \leq C\delta$$

y

$$\|x_\alpha - x_\alpha^\delta\| \leq \delta\sqrt{CG_\alpha},$$

donde C es la constante de la hipótesis (H2) en la Definición 2.3, $x_\alpha \doteq R_\alpha y$ y $x_\alpha^\delta \doteq R_\alpha y^\delta$.

En la última sección de este capítulo presentaremos algunos de los métodos de regularización espectral más conocidos y utilizados en las aplicaciones. Allí se retomarán conceptos que serán desarrollados en las siguientes secciones, tales como el concepto de “calificación clásica” de un método de regularización. A continuación presentamos brevemente el método de regularización

espectral de Tikhonov-Phillips, a los efectos de presentar su versión generalizada en la sección siguiente.

El método de Tikhonov-Phillips

En el *método de regularización de Tikhonov-Phillips* ([38], [45], [44]), quizás el método de regularización más popular entre los espectrales, la familia de funciones $\{g_\alpha\}$ está dada por

$$g_\alpha(\lambda) \doteq \frac{1}{\lambda + \alpha} \quad \text{para } \alpha \in (0, \alpha_0), \lambda \in [0, \|T\|^2]. \quad (2.14)$$

Se puede ver inmediatamente que $\{g_\alpha\}$ es efectivamente un MRE y que satisface (H2) con $C = 1$. Con este método, la solución regularizada a partir del dato con ruido y^δ viene entonces dada por

$$x_\alpha^\delta \doteq g_\alpha(T^*T)T^*y^\delta = (T^*T + \alpha I)^{-1}T^*y^\delta, \quad (2.15)$$

es decir, x_α^δ está definida mediante la ecuación lineal $T^*Tx_\alpha^\delta + \alpha x_\alpha^\delta = T^*y^\delta$.

La regularización de Tikhonov-Phillips puede formularse además como un problema de optimización, como lo enuncia el siguiente teorema.

Teorema 2.8. *Sea x_α^δ definida como en (2.15). Entonces x_α^δ es el único minimizante global del funcional de Tikhonov-Phillips J_α dado por*

$$J_\alpha(x) \doteq \|Tx - y\|^2 + \alpha \|x\|^2. \quad (2.16)$$

Demostración. Ver [17], Teorema 5.1. ■

En (2.16) el término $\|x\|^2$ se denomina penalizante o término de penalización del funcional J_α . El Teorema anterior permite asegurar que para cada parámetro α fijo, la solución regularizada obtenida con el método de Tikhonov-Phillips calculada con la familia de operadores de regularización $\{R_\alpha\}$ (asociada a la familia $\{g_\alpha\}$ dada por (2.14)) coincide con el único mínimo global del funcional J_α .

2.2.2. Métodos de Tikhonov generalizados

Los métodos de regularización espectral para problemas inversos mal condicionados, introducidos en la sección anterior, poseen la restricción de no admitir soluciones discontinuas o no regulares en general. En algunos problemas, como por ejemplo en ciertas aplicaciones en procesamiento y restauración de imágenes, esta restricción suele ser muy fuerte o indeseable y la aplicación de dichos métodos resulta en aproximaciones pobres o de baja calidad, especialmente cerca de puntos o regiones en los que la solución exacta no es regular. En ciertas situaciones,

puede ser deseable el diseño de métodos de regularización que permitan soluciones “no regulares”, por ejemplo imponiendo que sólo sean de variación acotada en lugar de suaves ([1], [48]). Así expresado, este parecería ser un objetivo contradictorio pues “regularizar” en el contexto de problemas inversos implica precisamente “suavizar” (de allí surge precisamente el nombre “métodos de regularización”). Sin embargo es importante señalar aquí que el objetivo primordial de un método para resolver un problema inverso mal condicionado no es el de “regularizar” el problema sino el de inducir estabilidad en el cálculo de las aproximaciones a su solución. Sucede que, como es bien sabido, los métodos tradicionales (tales como Tikhonov-Phillips, Expansión en valores singulares truncada, etc.) consiguen tal propósito filtrando o atenuando de diferentes maneras las componentes de alta frecuencia, de las cuales se sirven los errores (en los datos y de discretización) para originar la inestabilidad que, como vimos, es propia en los problemas mal condicionados. El resultado es el de aproximaciones “suaves” o “regularizadas”. Por esta razón, en el área de problemas inversos es usual hablar de “métodos de regularización”, aún cuando sería más adecuado hablar de *métodos de estabilización*. Esta observación sugiere que en los casos en los que no es razonable esperar soluciones exactas regulares, podría resultar ventajoso e interesante el diseño de estrategias que induzcan estabilidad sin “regularizar”. La utilización de penalizantes basados en este tipo de estrategias da lugar a los llamados métodos de Tikhonov-Phillips generalizados ([17]).

La modelización de un gran número de problemas en Física, Mecánica o procesamiento de imágenes, por ejemplo, requiere la introducción de un penalizante que permita soluciones discontinuas. Es por ello que, con el objetivo de comparar el desempeño de los distintos métodos que se presentarán a lo largo de esta tesis, introducimos el método de regularización de penalización por variación total ([1], [4], [49] y [48]).

El método de penalización por variación total

En lo que resta de esta sección supondremos que Ω una región convexa y acotada en \mathbb{R}^d , $d = 1, 2$ o 3 con frontera $\partial\Omega$ Lipschitz continua.

Definición 2.9. La *variación total* de una función $u \in L^1(\Omega)$, denotada por $J_0(u)$, se define como

$$J_0(u) \doteq \sup_{\vec{v} \in \mathcal{V}} \int_{\Omega} u \operatorname{div} \vec{v} \, dx,$$

donde \mathcal{V} , denominado el espacio de funciones de prueba, está dado por $\mathcal{V} \doteq \{\vec{v} \in C_0^1(\Omega, \mathbb{R}^d) : |\vec{v}(x)| \leq 1 \, \forall x \in \Omega\}$.

La siguiente proposición nos brinda una expresión simplificada para la variación total de una función $u \in C^1(\Omega)$.

Proposición 2.10. Si $u \in C^1(\Omega)$, entonces $J_0(u) = \int |\nabla u| dx$.

Demostración. Sean $u \in C^1(\Omega)$ y $\vec{v} \in C_0^1(\Omega, \mathbb{R}^d)$ entonces,

$$\begin{aligned}
 \int_{\Omega} u \operatorname{div} \vec{v} dx &= \int_{\Omega} u \sum_{i=1}^d \frac{\partial v_i}{\partial x_i} dx \\
 &= \sum_{i=1}^d \int_{\Omega} u \frac{\partial v_i}{\partial x_i} dx \\
 &= \sum_{i=1}^d \left\{ - \int_{\Omega} \frac{\partial u}{\partial x_i} v_i dx + \int_{\partial\Omega} u v_i \nu^i dS \right\} \\
 &= - \sum_{i=1}^d \int_{\Omega} \frac{\partial u}{\partial x_i} v_i dx && \text{(pues } v_i \text{ es de soporte compacto)} \\
 &= - \int_{\Omega} \sum_{i=1}^d \frac{\partial u}{\partial x_i} v_i dx \\
 &= - \int_{\Omega} \nabla u \cdot \vec{v} dx.
 \end{aligned}$$

Luego,

$$\begin{aligned}
 J_0(u) &= \sup_{\vec{v} \in \mathcal{V}} \int_{\Omega} u \operatorname{div} \vec{v} dx = \sup_{\vec{v} \in \mathcal{V}} \left(- \int_{\Omega} \nabla u \cdot \vec{v} dx \right) \\
 &= \sup_{\vec{v} \in \mathcal{V}} \int_{\Omega} \nabla u \cdot \vec{v} dx = \int_{\Omega} \sup_{\vec{v} \in \mathcal{V}} \nabla u \cdot \vec{v} dx \\
 &= \int_{\Omega} |\nabla u| dx. \quad \blacksquare
 \end{aligned}$$

Estamos ahora en condiciones de definir el funcional de variación acotada.

Definición 2.11. Para $u \in L^1(\Omega)$ se define el *funcional o norma de variación total*

$$\|u\|_{VT(\Omega)} \doteq \|u\|_{L^1(\Omega)} + J_0(u). \quad (2.17)$$

El espacio de funciones de variación total acotada sobre Ω , denotado por $VT(\Omega)$, se define entonces por el conjunto de todas las funciones $u \in L^1(\Omega)$ tales que $\|u\|_{VT(\Omega)} < \infty$. Se puede probar, que este funcional recién definido es una norma sobre $VT(\Omega)$ y, con esta norma, $VT(\Omega)$ es un espacio de Banach. El funcional variación total $J_0(\cdot)$ es una seminorma sobre este mismo espacio.

A partir de las definiciones anteriores estamos en condiciones de definir lo que se entiende por regularización por variación total. *La solución regularizada por variación total* del problema

$Tx = y$ se define como

$$x_\alpha^{VT} \doteq \arg \min_{u \in VT(\Omega)} \|Tu - y\|^2 + \alpha \|u\|_{VT(\Omega)}.$$

2.3. Convergencia y optimalidad

Al hacer referencia a la tasa de convergencia de un método de regularización ($\{R_\alpha\}, \alpha$), se piensa en la tasa con la cual tiende a cero el *error de regularización* o *error de aproximación*

$$\left\| (R_\alpha - T^\dagger)y \right\| = \|x_\alpha - x^\dagger\|, \quad (2.18)$$

cuando $\alpha \rightarrow 0^+$, o en la tasa con la cual tiende a cero el *error total* (que surge de comparar la solución regularizada con la solución exacta)

$$\left\| R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y \right\| = \left\| x_{\alpha(\delta, y^\delta)}^\delta - x^\dagger \right\|, \quad (2.19)$$

cuando $\delta \rightarrow 0^+$ (donde $\|y^\delta - y\| \leq \delta$). La tasa de convergencia a cero de (2.18) depende del operador de regularización R_α y del dato exacto y , pues el error de regularización es independiente del ruido. Por otro lado la tasa de convergencia a cero del error total (2.19) depende de la regla de elección del parámetro de regularización. Sin embargo, ambos errores están relacionados puesto que

$$\left\| x_{\alpha(\delta, y^\delta)}^\delta - x^\dagger \right\| \leq \left\| x_{\alpha(\delta, y^\delta)} - x^\dagger \right\| + \left\| x_{\alpha(\delta, y^\delta)} - x_{\alpha(\delta, y^\delta)}^\delta \right\|. \quad (2.20)$$

El término $\left\| x_{\alpha(\delta, y^\delta)} - x_{\alpha(\delta, y^\delta)}^\delta \right\| = \left\| R_{\alpha(\delta, y^\delta)}(y - y^\delta) \right\|$ es el *error asociado al ruido* y corresponde a la propagación del error en los datos. Notar que este término es cero si $\delta = 0$.

Si R_α convergiera a T^\dagger uniformemente, i.e. en la norma de los operadores (ya sabemos que esto no es posible), entonces como $\|x_\alpha - x^\dagger\| \leq \|R_\alpha - T^\dagger\| \|y\|$, se podría encontrar una tasa uniforme en y para la convergencia a cero del error de regularización (2.18). Sin embargo, veremos más adelante que para problemas mal condicionados nunca es posible obtener una tasa de convergencia uniforme. Las tasas de convergencia óptimas para métodos de regularización bajo supuestos sobre la solución (o equivalentemente sobre el dato exacto) están muy relacionados con el módulo de continuidad de T^\dagger sobre ciertos subconjuntos definidos por supuestos *a priori*.

Definición 2.12. Para $\mathcal{M} \subseteq X$ y $\delta > 0$, se define el *módulo de continuidad* de T^\dagger

$$\Omega_T(\delta, \mathcal{M}) \doteq \sup\{\|x\| : x \in \mathcal{M}, \|Tx\| \leq \delta\}. \quad (2.21)$$

En general, $\Omega_T(\delta, \mathcal{M})$ será infinito; este es el caso, por ejemplo, cuando $\mathcal{M} \cap \mathcal{N}(T) \neq \{0\}$ y \mathcal{M} es un subespacio de X . Más aún, si $\mathcal{M} \cap \mathcal{N}(T) = \{0\}$, entonces $\Omega_T(\delta, \mathcal{M})$ es finito si y sólo si T^\dagger es continuo sobre $T(\mathcal{M})$.

Más adelante veremos que $\Omega_T(\delta, \mathcal{M})$ está fuertemente relacionado con la tasa de convergencia óptima que puede alcanzar un método de regularización para $Tx = y$, bajo la hipótesis de que $x \in \mathcal{M}$.

Una solución regularizada del problema $Tx = y$ tiene la forma $x = Ry$, donde $R : Y \rightarrow X$ es un operador continuo no necesariamente lineal (R pertenece a alguna familia de operadores de regularización $\{R_\alpha\}$). Si en lugar del dato exacto y se conoce un dato con ruido y^δ , entonces el *peor error posible* que se puede cometer usando el operador R , bajo la información de que $\|y - y^\delta\| \leq \delta$ y bajo el supuesto *a priori* de que $x^\dagger \in \mathcal{M}$, está dado por

$$\Delta(\delta, \mathcal{M}, R) \doteq \sup_{\substack{x \in \mathcal{M}, y^\delta \in Y \\ \|Tx - y^\delta\| \leq \delta}} \|Ry^\delta - x\|. \quad (2.22)$$

En el siguiente resultado veremos que la precisión de cualquier regularización no puede ser mejor que $\Omega_T(\delta, \mathcal{M})$.

Proposición 2.13. *Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$, $\mathcal{M} \subseteq X$, $\delta > 0$ y $R : Y \rightarrow X$ un operador continuo, no necesariamente lineal, con $R(0) = 0$. Entonces*

$$\Delta(\delta, \mathcal{M}, R) \geq \Omega_T(\delta, \mathcal{M}). \quad (2.23)$$

Demostración. Sean $\delta > 0$ y $x \in \mathcal{M}$ tales que $\|Tx\| \leq \delta$. Si $y^\delta = 0$ entonces se sigue de (2.22) que $\Delta(\delta, \mathcal{M}, R) \geq \|R(0) - x\| = \|x\|$, pues por hipótesis $R(0) = 0$. Tomando supremo sobre todos los $x \in X$ tales que $\|Tx\| \leq \delta$, resulta que

$$\Delta(\delta, \mathcal{M}, R) \geq \sup\{\|x\| : x \in \mathcal{M}, \|Tx\| \leq \delta\} = \Omega_T(\delta, \mathcal{M}). \quad \blacksquare$$

Por otro lado, si $\mathcal{R}(T)$ no es cerrado, no puede existir una tasa de convergencia uniforme para un método de regularización, es decir, la convergencia puede ser arbitrariamente lenta. Más precisamente, se tiene el siguiente resultado fundamental:

Proposición 2.14. *Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$ con $\mathcal{R}(T)$ no cerrado, $\{R_\alpha\}$ una familia de operadores de regularización para T^\dagger con $R_\alpha(0) = 0$ para todo α y sea $\alpha(\delta, y^\delta)$ una regla de elección de parámetro. Entonces, no existe ninguna función $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ con $\lim_{\delta \rightarrow 0^+} f(\delta) = 0$ tal que*

$$\|R_{\alpha(\delta, y^\delta)} y^\delta - T^\dagger y\| \leq f(\delta)$$

para todo $y \in \mathcal{D}(T^\dagger)$ con $\|y\| \leq 1$ y $\delta > 0$.

Demostración. Ver [17], Proposición 3.11. \blacksquare

Así, sólo se pueden obtener tasas uniformes de convergencia sobre subespacios o subconjuntos propios de $\mathcal{D}(T^\dagger)$. Estos subconjuntos se construyen usualmente a partir de supuestos *a priori* sobre el dato exacto o sobre la solución exacta. Para formular tales supuestos acerca de la solución exacta consideraremos subconjuntos de X de la forma

$$\{x \in X : x = Bw, \|w\| \leq \rho\}, \quad (2.24)$$

donde B es un operador lineal acotado definido sobre X y $\rho > 0$. Cuando $B = (T^*T)^\mu$ para algún $\mu > 0$, denotamos el conjunto en (2.24) con

$$X_{\mu, \rho} \doteq \{x \in X : x = (T^*T)^\mu w, \|w\| \leq \rho\} \quad (2.25)$$

y definimos

$$X_\mu \doteq \bigcup_{\rho > 0} X_{\mu, \rho} = \mathcal{R}(T^*T)^\mu. \quad (2.26)$$

Estos conjuntos clásicos son denominados *conjuntos fuente* y una condición de la forma $x \in X_{\mu, \rho}$ se denomina *condición fuente* o *representación fuente* y cuantifica el grado de regularidad de la solución medida con respecto a la que induce el operador T . En efecto, si $x \in X_{\mu, \rho}$, entonces existe $w \in X$ con $\|w\| \leq \rho$ tal que $x = (T^*T)^\mu w$. Luego, el problema $Tx = y$ se puede plantear como $T(T^*T)^\mu w = y$. Si bien bajo este supuesto ambos problemas son equivalentes, tienen distinto grado de mal condicionamiento pues $T(T^*T)^\mu$ “suaviza” más que T . En consecuencia, como problema inverso, el segundo es peor condicionado que el primero.

Para el caso de operadores compactos, el conjunto X_μ se puede caracterizar en forma precisa utilizando los valores singulares del operador. En efecto, se tiene el siguiente resultado.

Proposición 2.15. *Sean X, Y espacios de Hilbert y $K : X \rightarrow Y$ un operador lineal compacto con sistema singular $\{(\sigma_n; u_n; v_n)\}$. Entonces, para todo $\mu > 0$,*

$$K^\dagger y \in \mathcal{R}((K^*K)^\mu) \iff \sum_n \frac{|\langle y, v_n \rangle|^2}{\sigma_n^{2+4\mu}} < \infty. \quad (2.27)$$

Demostración. Para $\mu > 0$ y $y \in \mathcal{D}(K^\dagger)$, la mejor solución aproximada $K^\dagger y \in \mathcal{R}((K^*K)^\mu)$ si y sólo si existe $w \in X$ tal que

$$K^\dagger y = (K^*K)^\mu w = \sum_n \sigma_n^{2\mu} \langle w, u_n \rangle u_n. \quad (2.28)$$

Comparando (2.28) con la expansión en valores singulares de la inversa generalizada de Moore-Penrose dada en (1.13) resulta que

$$\frac{\langle y, v_n \rangle}{\sigma_n} = \sigma_n^{2\mu} \langle w, u_n \rangle \quad \forall n \in \mathbb{N}.$$

Ahora bien, $w \in X$ si y sólo si $\{\langle w, u_n \rangle\} \in \ell^2$, es decir,

$$\sum_n \frac{|\langle y, v_n \rangle|^2}{\sigma_n^{2+4\mu}} < \infty,$$

con lo cual queda demostrada la proposición. \blacksquare

Observación 2.16. Notar que para $\mu = 0$, (2.27) se reduce al criterio de Picard (1.12). Es importante observar que la condición (2.27) es más restrictiva cuanto peor condicionado sea el problema, es decir, cuanto más rápido tiendan a cero los valores singulares σ_n . Esta condición también se hace más restrictiva a medida que crece μ .

En los dos resultados siguientes veremos que la cota *a priori* para la precisión de cualquier regularización, $\Omega_T(\delta, \mathcal{M})$ dada en (2.23), cuando \mathcal{M} es el conjunto fuente X_μ , no puede ser mejor que $O\left(\delta^{\frac{2\mu}{2\mu+1}}\right)$ y que este orden de convergencia es óptimo en el sentido que lo haremos preciso en la Proposición 2.18.

Proposición 2.17. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$ y $\mu, \rho > 0$. Entonces para cualquier $\delta > 0$,

$$\Omega_T(\delta, X_{\mu, \rho}) \leq \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}. \quad (2.29)$$

Demostración. Sea $x \in X_{\mu, \rho}$. Entonces existe $w \in X$ con $\|w\| \leq \rho$ tal que $x = (T^*T)^\mu w$. Sea $\{E_\lambda\}$ la familia espectral asociada al operador autoadjunto T^*T . Entonces, del hecho que $x = (T^*T)^\mu w$ y de la desigualdad de interpolación (ver Proposición 2.17 en [17] con $r = \mu$ y $q = \mu + 1/2$) se sigue que

$$\begin{aligned} \|x\| &= \|(T^*T)^\mu w\| \leq \left\| (T^*T)^{\mu+\frac{1}{2}} w \right\|^{\frac{2\mu}{2\mu+1}} \|w\|^{\frac{1}{2\mu+1}} \\ &= \left\| (T^*T)^{\frac{1}{2}} x \right\|^{\frac{2\mu}{2\mu+1}} \|w\|^{\frac{1}{2\mu+1}} = \|Tx\|^{\frac{2\mu}{2\mu+1}} \|w\|^{\frac{1}{2\mu+1}} \\ &\leq \|Tx\|^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}, \end{aligned}$$

puesto que $\|w\| \leq \rho$. Luego, si $\|Tx\| \leq \delta$, entonces $\|x\| \leq \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$. Tomando supremo sobre los $x \in X_{\mu, \rho}$ tales que $\|Tx\| \leq \delta$, en virtud de (2.21) se sigue que $\Omega_T(\delta, X_{\mu, \rho}) \leq \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$, lo cual concluye la demostración. \blacksquare

El resultado anterior nos dice entonces que $\Omega_T(\delta, X_{\mu, \rho})$ tiende a cero al menos tan rápido como $O\left(\delta^{\frac{2\mu}{2\mu+1}}\right)$ cuando $\delta \rightarrow 0^+$. De la siguiente proposición se sigue que, cuando el operador es compacto con rango no cerrado, esta estimación es óptima, al menos asintóticamente; es decir, en (2.29) vale la igualdad para una sucesión $\{\delta_k\}_{k \in \mathbb{Z}}$ que converge a cero.

Proposición 2.18. Sean X, Y espacios de Hilbert y $K : X \rightarrow Y$ un operador lineal compacto con rango no cerrado. Entonces, para cualesquiera $\mu, \rho > 0$, existe una sucesión $\{\delta_k\}_{k \in \mathbb{Z}}$ que converge a cero cuando $k \rightarrow \infty$ tal que

$$\Omega_K(\delta_k, X_{\mu, \rho}) = \delta_k^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}.$$

Demostración. Ver [17], Proposición 3.15. ■

Observación 2.19. Si el operador T es acotado pero no es necesariamente compacto y tiene rango no cerrado, el resultado anterior permanece esencialmente válido en el sentido que se puede encontrar una sucesión $\{\delta_k\}_{k \in \mathbb{Z}}$ que converge a cero, cuando $k \rightarrow \infty$, tal que $\Omega_T(\delta_k, X_{\mu, \rho}) = O\left(\delta_k^{\frac{2\mu}{2\mu+1}}\right)$ pero no es $o\left(\delta_k^{\frac{2\mu}{2\mu+1}}\right)$, es decir, no tiende a cero más rápido que $\delta_k^{\frac{2\mu}{2\mu+1}}$.

De las Proposiciones 2.13 y 2.17 y de la Observación 2.19 se deduce que, si $\mathcal{R}(T)$ no es cerrado, entonces ningún método de regularización puede ser tal que el error total converja a cero más rápido que $\delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}$ cuando $\delta \rightarrow 0^+$, bajo el supuesto *a priori* de que $x^\dagger \in X_{\mu, \rho}$, o más rápido que $O\left(\delta^{\frac{2\mu}{2\mu+1}}\right)$ bajo el supuesto de que $x^\dagger \in X_\mu$. Esto significa que de todas las tasas de la forma δ^s , el valor $s = \frac{2\mu}{2\mu+1}$ produce la mejor tasa posible a la que se puede aspirar bajo estos supuestos *a priori*. Esta observación induce la noción de optimalidad de un método de regularización sobre los conjuntos fuente $X_{\mu, \rho}$.

Definición 2.20. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$ con $\mathcal{R}(T)$ no cerrado, $\{R_\alpha\}$ una familia de operadores de regularización para T^\dagger y α^* una regla de elección de parámetro.

i) Se dice que $(\{R_\alpha\}, \alpha^*)$ es (un método de regularización) *óptimo en $X_{\mu, \rho}$* si

$$\Delta(\delta, X_{\mu, \rho}, R_{\alpha^*}) = \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}, \quad \forall \delta > 0.$$

ii) Se dice que $(\{R_\alpha\}, \alpha^*)$ es (un método de regularización) *de orden óptimo en $X_{\mu, \rho}$* si existe una constante $C \geq 1$ tal que

$$\Delta(\delta, X_{\mu, \rho}, R_{\alpha^*}) \leq C \delta^{\frac{2\mu}{2\mu+1}} \rho^{\frac{1}{2\mu+1}}, \quad \forall \delta > 0.$$

Notar que los conjuntos fuente crecen a medida que μ tiende a cero y la convergencia óptima se vuelve más lenta.

En lo que sigue enunciamos un teorema que nos provee un orden de convergencia óptimo para un método de regularización en el cual los errores en el dato y^δ se suponen modelados por un proceso estocástico con distribución normal. Dicho teorema será de utilidad a la hora de comparar algunos resultados que analizaremos en los Capítulos 3 y 4.

Teorema 2.21. Sean $s > 0$, $x \in H^s(\mathbb{R}^d)$ y y^δ modelado a partir de un proceso estocástico con distribución normal y nivel de error $\delta > 0$ para ruido aditivo. Sea, además, $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ un operador acotado con $\mathcal{R}(T)$ no cerrado, que satisface una propiedad de “suavizado” de orden t sobre los espacios de Sobolev, es decir, si $x \in H^s(\mathbb{R}^d)$ entonces $Tx \in H^{s+t}(\mathbb{R}^d)$.

Dada una familia de operadores de regularización $(\{R_\alpha\}, \alpha)$ para T^\dagger y α^* una regla de elección de parámetros, decimos que el método de regularización tiene orden óptimo si alcanza el siguiente orden de convergencia

$$E(\|R_{\alpha^*} y^\delta - x\|) = O(\delta^{\frac{t}{s+t+d/2}}). \quad (2.30)$$

Demostración. Ver [31], Lemma 3.2. ■

2.4. Calificación

En esta sección introducimos un concepto fuertemente vinculado a ciertos órdenes óptimos de convergencia que pueden alcanzarse mediante la aplicación de métodos de regularización, bajo ciertos supuestos *a priori*: el concepto de “calificación” de un método de regularización.

En muchos casos la función ω_μ en (2.10) se elige en términos de potencias de α como

$$\omega_\mu(\alpha) \doteq c\alpha^\mu, \quad (2.31)$$

donde c es una constante positiva independiente de α . En tal caso, en virtud del Teorema 2.6, la hipótesis “ $x^\dagger \in X_\mu$ ” resultará en un mejor orden de convergencia del error de regularización cuanto mayor sea μ (esto es, cuanto más fuerte sea la hipótesis sobre el grado de regularidad de x^\dagger). El mayor valor de μ para el cual se cumple (2.10) está relacionado con el concepto de *calificación clásica* de un MRE.

Definición 2.22. Sea $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ un MRE. Denotemos con $\mathcal{I}(g_\alpha)$ al conjunto

$$\mathcal{I}(g_\alpha) \doteq \{\mu \geq 0 : \forall \lambda \in [0, \infty), \exists k > 0 \text{ tal que } \lambda^\mu |r_\alpha(\lambda)| \leq k \alpha^\mu, \forall \alpha \in (0, \alpha_0)\}$$

y sea $\mu_0 \doteq \sup_{\mu \in \mathcal{I}(g_\alpha)} \mu$. Si $0 < \mu_0 < \infty$, decimos que $\{g_\alpha\}$ posee *calificación clásica de orden* μ_0 .

Es oportuno señalar que $\mathcal{I}(g_\alpha)$ es siempre no vacío puesto que $0 \in \mathcal{I}(g_\alpha)$ en virtud de (H2).

Si un MRE posee calificación clásica, se tiene el siguiente resultado con respecto al error de regularización:

Corolario 2.23. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$ y $\{g_\alpha\}$ un MRE con calificación clásica de orden μ . Si $x^\dagger \in X_\mu$, entonces $\|x_\alpha - x^\dagger\| = O(\alpha^\mu)$ cuando $\alpha \rightarrow 0^+$, donde $x^\dagger \doteq T^\dagger y$, $x_\alpha \doteq R_\alpha y$ y X_μ es como en (2.26).

Demostración. Es una consecuencia inmediata de la Definición 2.22 y del Teorema 2.6. ■

Notar que si $\{g_\alpha\}$ es un MRE para el cual $\mu_0 = \infty$, donde μ_0 está dado por la Definición 2.22, entonces la condición (2.10) con ω_μ dado por (2.31) se satisface para todo $\mu > 0$. De esta forma, se deduce del Teorema 2.6 que, para cualquier $\mu > 0$, la hipótesis “ $x^\dagger \in X_\mu$ ” implica que la tasa de convergencia a cero del error de regularización es $O(\alpha^\mu)$. En este caso, cuanto mayor regularidad tenga x^\dagger , mejor será el orden de convergencia del error de regularización. Sin embargo, no ocurre lo mismo cuando $\{g_\alpha\}$ tiene calificación clásica de orden μ_0 (i.e. cuando $0 < \mu_0 < \infty$), pues la mejor tasa de convergencia que se puede obtener para el error de regularización es $O(\alpha^{\mu_0})$, la cual se obtiene bajo el supuesto *a priori* de que $x^\dagger \in X_{\mu_0}$. Si se agregan hipótesis adicionales de regularidad sobre x^\dagger (por ejemplo, $x^\dagger \in X_{\mu^*}$ con $\mu^* > \mu_0$), la tasa de convergencia del error de regularización no se puede mejorar, i.e., sigue siendo $O(\alpha^{\mu_0})$.

Mathé y Pereverzev, en el año 2003 ([32]), introdujeron por primera vez la siguiente definición de calificación de un MRE, con lo cual formalizaron y extendieron la noción clásica de este concepto.

Definición 2.24. Sean $a > 0$ y $\rho : (0, a] \rightarrow (0, \infty)$ una función creciente. Se dice que el MRE $\{g_\alpha\}$ posee *calificación* ρ si existe una constante $\gamma \in (0, \infty)$ tal que

$$\sup_{\lambda \in (0, a]} |r_\alpha(\lambda)| \rho(\lambda) \leq \gamma \rho(\alpha) \quad \forall \alpha \in (0, a], \quad (2.32)$$

donde $r_\alpha(\lambda) \doteq 1 - \lambda g_\alpha(\lambda)$.

Observación 2.25. Es importante notar que, según esta definición, la calificación de un MRE es una función (de α) y no un número como ocurre con la calificación clásica (el orden μ_0). Por otra parte, se puede ver inmediatamente que si un MRE tiene calificación clásica de orden μ_0 , entonces la función $\rho(\alpha) \doteq \alpha^\mu$ es calificación según la Definición 2.24 para todo $\mu \in (0, \mu_0]$. Incluso se puede probar que si $\omega_\mu(\alpha)$, dada por (2.31), satisface (2.10) para algún μ^* , entonces también satisface

$$\sup_{0 < \sigma \leq \infty} |1 - \sigma g_\alpha(\sigma)| \sigma^\mu = o(\alpha^\mu),$$

para todo $\mu \in [0, \mu^*]$.

En 2009, Spies y Temperini ([42]) generalizaron los conceptos de calificación clásica y calificación según Mathé, para métodos de regularización arbitrarios introduciendo tres niveles de calificación: débil, fuerte y óptima. Allí mostraron que la calificación débil extiende la calificación de Mathé y Perverzev. También mostraron que algunos métodos tradicionales de regularización que no poseen calificación clásica (tales como los métodos de Landweber, Showalter y TSVD

que se detallan en la próxima sección) sí tienen calificación generalizada, lo que lleva a órdenes de convergencia óptimos para el error de regularización.

2.5. Reglas de elección de parámetro

En esta sección introduciremos reglas de elección de parámetro que forman parte de métodos de regularización de orden óptimo sobre conjuntos fuente del tipo X_μ o $X_{\mu,\rho}$, $\mu, \rho > 0$. Presentaremos, en primer lugar, reglas de elección de parámetro *a priori*, que dependen sólo del nivel de ruido δ y por ello, pueden definirse “antes” de la obtención del dato y^δ . En segundo lugar, presentaremos una de las reglas de elección de parámetro *a posteriori* más frecuentemente utilizada, conocida como el “principio de discrepancia de Morozov”.

2.5.1. Reglas de elección de parámetro *a priori*

Las reglas de elección de parámetro *a priori* dependen sólo del nivel de ruido δ y no del dato y^δ . Una forma sencilla en que se suelen definir este tipo de reglas es como potencias de δ , es decir,

$$\alpha(\delta) = p \delta^s, \quad (2.33)$$

donde p y s son constantes positivas.

Sean $\alpha(\delta)$ una regla de elección de parámetro *a priori* y $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$. Si existen constantes positivas c_1 y c_2 tales que

$$c_1 h(\delta) \leq \alpha(\delta) \leq c_2 h(\delta),$$

para δ en un intervalo a la derecha de 0, entonces diremos que α es del orden de $h(\delta)$ (o que α y h tienen el mismo orden) para $\delta \rightarrow 0$ y lo denotaremos con $\alpha \sim h(\delta)$ para $\delta \rightarrow 0$.

Recordemos que una regla de elección de parámetro, a diferencia de una familia de operadores de regularización, se define para un dato y fijo. En algunos casos, como veremos a continuación, cuando se tiene algún tipo de información *a priori* acerca de la solución exacta $x^\dagger \doteq T^\dagger y$, por ejemplo cuando se conocen los valores exactos de μ y ρ en la condición fuente $x^\dagger \in X_{\mu,\rho}$, es posible definir reglas de elección de parámetros como en (2.33), donde las constantes p y s se determinan a partir de μ y ρ , de tal manera que, si el MRE satisface determinadas condiciones, dichas reglas forman parte de un método de regularización de orden óptimo en $X_{\mu,\rho}$. En efecto:

Proposición 2.26. Sean X, Y espacios de Hilbert, $T \in \mathcal{L}(X, Y)$, $\{g_\alpha\}$ un MRE tal que G_α , definida como en (2.13), satisfice

$$G_\alpha = O\left(\frac{1}{\alpha}\right) \quad \text{cuando } \alpha \rightarrow 0^+. \quad (2.34)$$

Sean $\mu, \rho > 0$. Supongamos que existe una constante $c > 0$ independiente de α tal que $\omega_\mu : (0, \alpha_0) \rightarrow \mathbb{R}^+$ definida como en (2.31) satisfice (2.10). Si $\alpha^*(\delta)$ es una regla de elección de parámetro a-priori que satisfice

$$\alpha^* \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}, \quad (2.35)$$

entonces el método de regularización $(\{R_\alpha\}, \alpha^*(\delta))$ es de orden óptimo en $X_{\mu, \rho}$, donde $\{R_\alpha\}$ es la familia de operadores definidos por (2.7).

Demostración. Sean $x^\dagger \doteq T^\dagger y \in X_{\mu, \rho}$, $y^\delta \in Y$ tal que $\|y^\delta - y\| \leq \delta$, $x_\alpha \doteq R_\alpha y$ y $x_\alpha^\delta \doteq R_\alpha y^\delta$. De los Teoremas 2.6 y 2.7 (recordar que por hipótesis $\omega_\mu(\alpha) = c\alpha^\mu$ satisfice (2.10)) y de la condición (2.34) sobre G_α se sigue que

$$\|x_{\alpha^*}^\delta - x^\dagger\| \leq \|x_{\alpha^*} - x^\dagger\| + \|x_{\alpha^*} - x_{\alpha^*}^\delta\| \leq c\alpha^{*\mu}\rho + k\delta\sqrt{\frac{C}{\alpha^*}}, \quad (2.36)$$

donde k es una constante positiva independiente de α y C es la constante de la hipótesis (H2). Utilizando (2.35) es fácil probar que el lado derecho de la última desigualdad en (2.36) se puede acotar por $\hat{c}\delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}}$, donde \hat{c} es una constante positiva independiente de δ . Luego, (ver Definición 2.20) el método de regularización $(\{R_\alpha\}, \alpha^*(\delta))$ es de orden óptimo en $X_{\mu, \rho}$, como queríamos probar. ■

Observación 2.27. A menudo sucede que no se conoce el valor de ρ . En tal caso, si consideramos una regla de elección de parámetro $\alpha \sim \delta^{\frac{2}{2\mu+1}}$, se obtiene una cota para el error total que es al menos óptima con respecto a la potencia de δ , es decir, de orden óptimo en X_μ : $\|x_\alpha^\delta - x^\dagger\| = O\left(\delta^{\frac{2\mu}{2\mu+1}}\right)$. Si también μ es desconocido, entonces en tal caso no es apropiado usar una regla de elección de parámetro a priori y se deberá considerar alguna regla a posteriori como las que veremos en la siguiente sección.

2.5.2. Reglas de elección de parámetro a posteriori: el principio de discrepancia de Morozov

En esta sección presentamos una regla de elección de parámetro a posteriori muy utilizada en aplicaciones y problemas concretos, conocida como el principio de discrepancia de Morozov ([34]). Como señalamos en la Observación 2.27, este tipo de reglas de elección de parámetro

es particularmente apropiada en los casos en los que no se dispone de ninguna información *a priori* acerca de la regularidad de la solución exacta, en particular, cuando no se conoce ningún valor de $\mu > 0$ tal que $x^\dagger \in X_\mu$.

Dado un MRE $\{g_\alpha\}$, el principio de discrepancia de Morozov consiste esencialmente en elegir el parámetro de regularización mediante una comparación entre la “*discrepancia*” (de allí el nombre) $\|Tx_\alpha^\delta - y^\delta\| = \|(TR_\alpha - I)y^\delta\|$ y una cota asociada con el nivel de ruido δ , donde R_α está dado por (2.7). Más precisamente, supongamos que queremos resolver el problema $Tx = y$, pero en lugar del dato exacto y sólo se conoce un dato con ruido y^δ tal que $\|y - y^\delta\| \leq \delta$. En este caso, claramente no tiene sentido tratar de encontrar una solución aproximada \tilde{x} que produzca una discrepancia $\|T\tilde{x} - y^\delta\|$ que sea *estrictamente menor* que δ . Lo mejor que podríamos obtener es una discrepancia de orden δ . Por otra parte, para un nivel de ruido fijo $\delta > 0$, el error total aumenta a medida que el parámetro de regularización α tiende a cero, con lo cual la estabilidad disminuye. Este simple razonamiento nos permite concluir que lo mejor que podemos hacer es elegir el mayor parámetro de regularización posible que resulte en una discrepancia “de orden δ ”. Es precisamente la formalización de este razonamiento lo que constituye el principio de discrepancia de Morozov. Sea τ una constante positiva tal que

$$\tau > \sup_{\alpha > 0, \lambda \in [0, \|T\|^2]} |r_\alpha(\lambda)|,$$

donde $r_\alpha(\lambda)$ es como en (2.9). Notar que $\tau > 1$ pues $r_\alpha(0) = 1$. El parámetro de regularización definido mediante el principio de discrepancia está dado por

$$\alpha_d(\delta, y^\delta) \doteq \sup \left\{ \alpha > 0 : \|Tx_\alpha^\delta - y^\delta\| \leq \tau\delta \right\}. \quad (2.37)$$

Es fácil probar que, si para cada $\lambda > 0$, el mapeo $\alpha \rightarrow g_\alpha(\lambda)$ es continuo por izquierda, entonces el funcional $\alpha \rightarrow \|Tx_\alpha^\delta - y^\delta\|$ también es continuo por izquierda. Luego, se alcanza el supremo en (2.37) y así,

$$\left\| Tx_{\alpha_d(\delta, y^\delta)}^\delta - y^\delta \right\| \leq \tau\delta.$$

2.6. Métodos de regularización espectrales

En esta sección presentamos algunos de los métodos de regularización espectrales más utilizados en aplicaciones y problemas concretos muy conocidos en la literatura. En cada caso definiremos la correspondiente familia de funciones $\{g_\alpha\}$, la representación para la solución regularizada en el caso de operadores compactos y analizaremos si dichos métodos poseen calificación clásica.

2.6.1. El método de descomposición en valores singulares truncada (TSVD)

El método de descomposición en valores singulares truncada o TSVD se obtiene anulando los valores del espectro de T^*T que se encuentran por debajo de un cierto umbral dado por el parámetro de regularización α . En este caso entonces, para $\alpha \in (0, \alpha_0)$ la familia $\{g_\alpha\}$ se define por

$$g_\alpha(\lambda) \doteq \begin{cases} \frac{1}{\lambda} & \text{si } \lambda \in [\alpha, \infty), \\ 0 & \text{si } \lambda \in [0, \alpha). \end{cases} \quad (2.38)$$

Es fácil ver que $\{g_\alpha\}$ es un MRE y que la hipótesis (H2) se verifica con $C = 1$. La condición (2.10) se satisface con $\omega_\mu(\alpha) = \alpha^\mu$ para todo $\mu > 0$, de lo cual se sigue que $\mu_0 = \infty$, con μ_0 como en la Definición 2.22. Luego, este método no posee calificación clásica.

La solución regularizada a partir del dato con ruido y^δ está dada por

$$x_\alpha^\delta \doteq g_\alpha(T^*T)T^*y^\delta = \int_\alpha^{\|T\|^2} \frac{1}{\lambda} dE_\lambda T^*y^\delta. \quad (2.39)$$

Observar que en el caso en que T es compacto, con sistema singular $\{(\sigma_n; u_n; v_n)\}$, entonces x_α^δ se puede calcular usando (2.8) y (1.8) en la siguiente forma:

$$x_\alpha^\delta = \sum_{n:\sigma_n^2 \geq \alpha} \frac{1}{\sigma_n^2} \langle T^*y^\delta, u_n \rangle u_n = \sum_{n:\sigma_n^2 \geq \alpha} \frac{1}{\sigma_n^2} \langle y^\delta, Tu_n \rangle u_n = \sum_{n:\sigma_n^2 \geq \alpha} \frac{1}{\sigma_n} \langle y^\delta, v_n \rangle u_n,$$

la cual puede verse como una versión “truncada” de la descomposición en valores singulares de $T^\dagger y$ dada en (1.13). Esto justifica el nombre del método.

2.6.2. El método de Tikhonov-Phillips

Como vimos en la Sección 2.2 el método de regularización de Tikhonov-Phillips se define a partir de la familia de funciones $\{g_\alpha\}$ dada por

$$g_\alpha(\lambda) = \frac{1}{\lambda + \alpha} \quad \text{para } \alpha \in (0, \alpha_0), \lambda \in [0, \|T\|^2]. \quad (2.40)$$

Para determinar si este método espectral posee calificación clásica, y en caso afirmativo hallar su orden, se debe encontrar una función $\omega_\mu(\alpha)$ del tipo dado en (2.31), i.e. $\omega_\mu(\alpha) = c\alpha^\mu$, de tal manera que sea cota superior de la función

$$h_\mu(\lambda) \doteq \lambda^\mu |r_\alpha(\lambda)| = \lambda^\mu \frac{\alpha}{\alpha + \lambda}.$$

Observar que para $\mu < 1$, se tiene que $h_\mu(\lambda)$ alcanza su máximo en $\lambda = \frac{\mu\alpha}{1-\mu}$ y entonces $h_\mu(\lambda) \leq \mu^\mu(1-\mu)^{1-\mu}\alpha^\mu \leq \alpha^\mu$. Para $\mu \geq 1$, $h_\mu(\lambda)$ es estrictamente creciente y alcanza su mayor valor en $\|T\|^2$. Entonces, definimos ω_μ como

$$\omega_\mu(\alpha) \doteq \begin{cases} \alpha^\mu, & \text{para } \mu \leq 1 \\ \|T\|^{2\mu-2}\alpha, & \text{para } \mu > 1. \end{cases}$$

Luego, se deduce que $\{g_\alpha\}$ tiene calificación clásica de orden $\mu_0 = 1$, pues (2.10) se satisface con $\omega_\mu(\alpha) = \alpha^\mu$ para todo $\mu \in [0, 1]$.

Finalmente, es importante observar que para un operador compacto T con sistema singular $\{(\sigma_n; u_n; v_n)\}$, la ecuación (2.15) tiene la forma

$$x_\alpha^\delta = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2 + \alpha} \langle y^\delta, v_n \rangle u_n.$$

Una comparación con la ecuación que define a la inversa generalizada de Moore-Penrose (1.13) permite ver claramente cómo Tikhonov-Phillips estabiliza las aproximaciones: los errores en $\langle y^\delta, v_n \rangle$ ya no son amplificadas por los factores $1/\sigma_n$, sino por los factores $\sigma_n/(\sigma_n^2 + \alpha)$, los cuales permanecen acotados cuando $n \rightarrow \infty$ y, más aún, tienden a 0.

2.6.3. El método de Showalter o de regularización asintótica

El método de regularización asintótica o método de Showalter se construye a partir de la solución $u_\delta : \mathbb{R}_0^+ \rightarrow X$ del problema de valores iniciales

$$\begin{cases} u'_\delta(t) + T^*Tu_\delta(t) = T^*y^\delta, & t \geq 0, \\ u_\delta(0) = 0, \end{cases} \quad (2.41)$$

definiendo $x_\alpha^\delta \doteq u_\delta(\frac{1}{\alpha})$. Si se reemplaza y^δ por y en (2.41), entonces escribimos u y x_α en lugar de u_δ y x_α^δ . Veamos que

$$v(t) \doteq \int \frac{(1 - e^{-\lambda t})}{\lambda} dE_\lambda T^*y^\delta, \quad t \geq 0$$

es solución de (2.41), donde $\{E_\lambda\}$ es la familia espectral asociada a T^*T . Como

$$v'(t) = \int e^{-\lambda t} dE_\lambda T^*y^\delta$$

(esta derivación puede justificarse rigurosamente mediante el Teorema de la Convergencia Dominada), se tiene que

$$T^*Tv(t) = \int \lambda \left(\frac{1 - e^{-\lambda t}}{\lambda} \right) dE_\lambda T^*y^\delta = T^*y^\delta - v'(t),$$

lo cual junto al hecho obvio de que $v(0) = 0$, permite afirmar que $v(t)$ es efectivamente solución del problema de valores iniciales (2.41). Luego,

$$x_\alpha^\delta = v\left(\frac{1}{\alpha}\right) = \int \frac{1 - e^{-\frac{\lambda}{\alpha}}}{\lambda} dE_\lambda T^* y^\delta.$$

Como la función

$$g_\alpha(\lambda) = \frac{(1 - e^{-\frac{\lambda}{\alpha}})}{\lambda} = \int_0^{\frac{1}{\alpha}} e^{-\lambda s} ds \quad (2.42)$$

es continua para todo $\lambda \geq 0$, resulta que g_α pertenece a \mathcal{M}_0 , el conjunto de funciones medibles con respecto a la medida $d\|E_\lambda x\|^2$, y así, $x_\alpha^\delta = g_\alpha(T^*T)T^*y^\delta$. Se puede ver fácilmente que g_α satisface las hipótesis (H1)-(H3) de la Definición 2.3 y por lo tanto es un MRE. La hipótesis (H2) se verifica con $C = 1$.

Este método suele llamarse también de *regularización asintótica* porque en (2.42) no se integra hasta ∞ , sino sólo hasta $1/\alpha$, donde α actúa como parámetro de regularización. Cuando $\alpha \rightarrow 0^+$, este extremo superior de integración se hace más grande y así, la exactitud $\|x_\alpha - x^\dagger\|$ mejora (el error de regularización tiende a 0), pero eventualmente el error asociado al ruido $\|x_\alpha^\delta - x_\alpha\|$ empeora (recordar que se sigue del Teorema 2.7 que $\|x_\alpha^\delta - x_\alpha\| \leq \delta\sqrt{CG_\alpha} = \delta\sqrt{1/\alpha}$).

Analicemos ahora si este MRE posee calificación clásica. La función

$$h_\mu(\lambda) \doteq \lambda^\mu |r_\alpha(\lambda)| = \lambda^\mu e^{-\frac{\lambda}{\alpha}},$$

alcanza su máximo en $\lambda = \alpha\mu$ y entonces $h_\mu(\lambda) \leq \alpha^\mu \mu^\mu e^{-\mu}$ para todo $\mu > 0$. Así, la condición (2.10) se satisface con $\omega_\mu(\alpha) \doteq \alpha^\mu \mu^\mu e^{-\mu}$ para todo $\mu > 0$, de lo cual se deduce que $\mu_0 = \infty$. Luego, este método no posee calificación clásica.

Del Teorema 2.5 y del hecho que $x_\alpha = u(t)$ para $t = 1/\alpha$ se sigue que si $y \in \mathcal{D}(T^\dagger)$, entonces

$$\lim_{t \rightarrow \infty} u(t) = \lim_{\alpha \rightarrow 0^+} x_\alpha = T^\dagger y, \quad (2.43)$$

mientras que si $y \notin \mathcal{D}(T^\dagger)$ entonces $\lim_{t \rightarrow \infty} \|u(t)\| = \infty$.

Resulta interesante observar que, puesto que $v'(t) \rightarrow 0$ y $x_\alpha \rightarrow x^\dagger$ cuando $t \rightarrow \infty$, la solución exacta x^\dagger se puede interpretar como el estado estacionario de la solución del problema de valores iniciales (2.41). Asimismo, es importante observar que la ecuación diferencial en (2.41) se convierte en la ecuación normal $T^*Tx^\dagger = T^*y$ para “ $t = \infty$ ”.

2.6.4. El método de Landweber

El *método de Landweber* es, en principio, un método iterativo que consiste esencialmente en calcular la sucesión de aproximaciones $\{x_k^\delta\}_{k \in \mathbb{N}}$ de forma recursiva, a partir de un valor inicial

$x^* = x_0^\delta$, mediante

$$x_k^\delta = x_{k-1}^\delta + T^*(y^\delta - Tx_{k-1}^\delta).$$

Notar que en este caso la regla de elección de parámetro α , característica de los métodos espectrales se ve reemplazada por el criterio de parada en las iteraciones.

Por otro lado, es importante observar y se puede probar fácilmente que este método de regularización iterativo es también un método espectral asociado a la familia espectral discreta

$$g_k(\lambda) \doteq \sum_{j=0}^{k-1} (1-\lambda)^j \text{ con } 0 < \lambda < 2.$$

Luego, el método iterativo de Landweber puede expresarse de forma no recursiva como

$$R_k y = \sum_{j=0}^{k-1} \sum_{\sigma_n > 0} (1 - \sigma_n^2)^j \sigma_n \langle y, v_n \rangle u_n. \quad (2.44)$$

Con respecto a la calificación de este MRE, observar que en este caso se tiene que

$$h_\mu(\lambda) \doteq \lambda^\mu |r_k(\lambda)| \leq \begin{cases} (k+1)^{-\mu}, & \text{para } \mu \leq 1, \\ (\mu)^\mu (k+1)^{-\mu}, & \text{para } \mu > 1. \end{cases}$$

Luego, la condición (2.10) se satisface con

$$\omega_\mu(\alpha) \doteq \begin{cases} (k+1)^{-\mu}, & \text{para } \mu \leq 1, \\ (\mu)^\mu (k+1)^{-\mu}, & \text{para } \mu > 1, \end{cases}$$

de lo cual se sigue que $\mu_0 = \infty$. Concluimos entonces que este método tampoco posee calificación clásica.

CAPÍTULO 3

Regularización wavelet-espectral

En este capítulo presentaremos un método de regularización wavelet-espectral para resolver problemas inversos lineales mal condicionados. El fundamento principal para la incorporación de wavelets en el tratamiento de este tipo de problemas en los que los datos están contaminados con algún nivel de ruido, radica en lo ya expuesto en el Capítulo 1. Principalmente destacamos que aplicar umbralado wavelet a estos datos constituye una herramienta eficaz para reconstruir un dato casi libre de ruido sin modificar demasiado las características principales del dato original, comparando con métodos tradicionales de filtrado por bandas de frecuencia. En un contexto general aplicado a señales o imágenes, una base multiresolución wavelet nos permite diferenciar y agrupar los rasgos más importantes que determinan “la identidad” de una señal través de las componentes de alta y baja frecuencia. Utilizando esta ventaja proponemos realizar un umbralado wavelet para atenuar o eliminar componentes de alta frecuencia especialmente asociadas al ruido con el objetivo de que este filtrado previo provea a los métodos espectrales tradicionales datos considerablemente más cercanos a los datos exactos. Esto nos debería permitir obtener regularizaciones con parámetros de regularización significativamente mas pequeños, obteniendo así soluciones más cercanas a la solución exacta.

El método wavelet-espectral que presentaremos consta de dos pasos. Una expansión wavelet seguida del umbralado fuerte constituyen el primero de ellos, mientras que el segundo consiste en la aplicación de un método de regularización espectral clásico, con ciertas propiedades que pueden expresarse en términos de su calificación clásica (ver [17],[43]). El principal objetivo de este capítulo es estudiar su convergencia y orden de convergencia. Además de presentar el método detalladamente, en particular, probaremos que bajo determinadas condiciones, el método es óptimo según (2.30) (o más precisamente casi-óptimo en un sentido que detallaremos rigurosamente cuando sea necesario). La utilización de wavelets en métodos de regularización remite primariamente a un trabajo de Klann, Maass y Ramlau ([26]) donde los autores proponen un método combinado wavelet-Tikhonov y prueban algunos resultados de convergencia en términos del nivel de ruido.

Comenzamos este capítulo con una sección de preliminares donde presentaremos el problema a abordar, las principales hipótesis sobre el operador y los espacios funcionales con los que trabajamos. En la Sección 3.2 haremos un análisis detallado de la convergencia para los errores de proyección, umbralado y regularización, lo que nos permitirá probar un teorema fundamental de convergencia de orden casi-óptimo para el método combinado wavelet-Landweber. Por último, en la Sección 3.3 generalizaremos esos resultados de convergencia para el caso de métodos wavelet-espectrales en los que la familia espectral posee ciertas propiedades asociadas a su calificación clásica (ver [17], [43]). En particular probaremos que bajo ciertas condiciones, el método de dos pasos converge de forma óptima.

3.1. Preliminares

Consideraremos problemas inversos de la forma

$$Tf = g \tag{3.1}$$

donde T es un operador lineal y de rango no cerrado entre dos espacios de Hilbert X e Y y g es el dato que se supone conocido, quizás con un cierto grado de error. En lo que sigue, y por razones de simplicidad supondremos además que T es compacto, sin embargo, todos los resultados siguen siendo válidos para el caso de operadores no compactos.

A lo largo de las siguientes secciones Ω denotará un dominio acotado contenido en \mathbb{R}^d , $1 \leq d < \infty$. En ciertas ocasiones, por simplicidad utilizaremos $\Omega = Q$, el cubo unitario $Q \doteq [0, 1]^d$, no obstante todos los resultados pueden extenderse a dominios acotados Ω generales.

A continuación describiremos las hipótesis que impondremos sobre el operador T , sobre la solución exacta de nuestro problema y también sobre el error de observación. En relación al operador $T : L^2(\Omega) \rightarrow L^2(\Omega)$ supondremos que satisface una propiedad a la que llamaremos de t -suavizado sobre espacios de Sobolev. Dicha propiedad consiste en la existencia de una constante $t = t(T) > 0$ tal que $Ran(T) \subseteq H^t(\Omega)$, donde $H^t(\Omega)$ es, como definimos en la Sección 1.1, el espacio de Sobolev usual de orden t . Como es usual en problemas inversos mal condicionados, con respecto a la solución exacta supondremos que esta satisface una condición fuente del tipo $f \in Ran((T^*T)^{\nu/2})$, para algún $\nu > 0$. Se puede probar que, bajo la condición de t -suavizado del operador T , esta condición es equivalente a $f \in H^{\nu t}(\Omega)$. Con más generalidad, se puede probar que $\forall s \in \mathbb{R}$,

$$f \in Ran((T^*T)^{s/2t}) \Leftrightarrow f \in H^s(\Omega). \tag{3.2}$$

Por último, para el error de observación supondremos un modelo estocástico. Más precisamente, nuestra observación estará dada por $g^\delta = g + \delta dW = Tf + \delta dW$, donde dW es ruido blanco y $\delta > 0$ es el nivel de error.

De aquí en más, si T posee la propiedad de t -suavizado, en virtud de (3.2) identificaremos $f \in H^s(\Omega)$ con el elemento x_f de $L^2(\Omega)$ tal que $f = (T^*T)^{\frac{s}{2t}} x_f$ y utilizaremos la equivalencia de las normas $\|f\|_{H^s(\Omega)}$ y $\|(T^*T)^{-\frac{s}{2t}} f\|$. Observar que si $\{(\sigma_n; u_n; v_n)\}$ es un sistema singular para el operador T , esta equivalencia se traduce en el hecho que $\forall s > 0$

$$\|f\|_{H^s(\Omega)}^2 \simeq \sum_{\sigma_n > 0} \sigma_n^{-\frac{2s}{t}} |\langle f, u_n \rangle|^2. \quad (3.3)$$

A menudo y por razones de simplicidad notacional, utilizaremos (3.3) como una igualdad.

Como en la Sección 1.4 denotaremos con ψ y ϕ a la función wavelet y su función de escala asociada, respectivamente, ambas en $L^2(\Omega)$. Con $P_j : L^2(\Omega) \rightarrow V_j$ denotaremos a la proyección ortogonal sobre V_j y, para $\gamma > 0$, $S_\gamma : V_j \rightarrow V_j$ denotará al operador de umbralado wavelet.

A continuación procederemos a formalizar nuestro método de regularización wavelet-espectral. En un primer paso, aplicaremos el operador proyección P_j seguido del operador de umbralado S_γ , con el objetivo de atenuar la influencia de las componentes de alta frecuencia del ruido y recuperar ciertas condiciones de suavidad del dato. En el segundo paso aplicaremos el operador de regularización espectral, al que denotamos con R_α . Así, la solución regularizada del problema (3.1) con la utilización de este método estará dada por

$$f_{reg}^\delta \doteq R_\alpha S_\gamma P_j g^\delta.$$

3.2. Análisis de convergencia

En esta sección estudiaremos la convergencia y órdenes de convergencia del método de dos pasos descrito anteriormente. Para ello comenzaremos realizando estimaciones para cada uno de los errores de proyección, umbralado y regularización, individualmente. Finalmente, utilizando todas estas estimaciones en la última sección de este capítulo, presentaremos un teorema en el que probaremos que el método combinado converge y veremos además que el orden de convergencia para el error total es casi-óptimo.

3.2.1. Estimación del error de proyección wavelet

En esta sección analizaremos la convergencia de las proyecciones sobre los espacios wavelets, bajo ciertas condiciones sobre el dato y la función de escala.

Para probar el resultado principal de esta sección (Teorema 3.5) necesitaremos previamente de una serie de lemas y teoremas auxiliares que presentamos a continuación. Dado un dominio $D \in \mathbb{R}^d$, de aquí en más denotaremos con $\Pi_m(D)$ al conjunto de todos los polinomios de grado menor o igual que m , definidos sobre D . Tal como vimos en el capítulo 1, con polinomio de grado m hacemos referencia a la combinación lineal de las funciones base $x_1^{k_1} \dots x_d^{k_d}$ para $k_1 + \dots + k_d \leq m$.

Lema 3.1. Sean $1 \leq p \leq +\infty$, $\ell \in \mathbb{N}_0$, $j \in \mathbb{Z}$ y V_j como antes. Entonces si $\phi \in W^{\ell,p}(\Omega)$ existe una constante C independiente de j tal que para toda $f \in V_j$ se satisface

$$\|f\|_{W^{\ell,p}(\Omega)} \leq C 2^{\ell j} \|f\|_{L^p(\Omega)}.$$

Demostración. Ver Teorema 3.4.1 en [10]. ■

Lema 3.2. Sean $1 \leq p \leq +\infty$, $m \in \mathbb{N}$ y Ω un dominio acotado. Entonces existe una constante C que depende de d , m , p y Ω , tal que para toda $f \in W^{m+1,p}(\Omega)$ se verifica

$$\inf_{g \in \Pi_m(\Omega)} \|f - g\|_{L^p(\Omega)} \leq C |f|_{W^{m+1,p}(\Omega)}.$$

Demostración. Ver Teorema 3.2.1 en [10]. ■

El resultado del Lema 3.2 tiene una consecuencia directa sobre el error de aproximación polinomial local para funciones en $W^{m+1,p}(\Omega)$. En efecto

Corolario 3.3. Sean $1 \leq p \leq \infty$, $m \in \mathbb{N}$, $f \in W^{m+1,p}(Q)$, $I_h \subseteq Q$, donde I_h es un cubo en \mathbb{R}^d de lados de longitud h y paralelos a los ejes coordenados. Entonces existe una constante C independiente de h y de f , tal que

$$\inf_{g \in \Pi_m(I_h)} \|f - g\|_{L^p(I_h)} \leq C h^{m+1} |f|_{W^{m+1,p}(I_h)}. \quad (3.4)$$

Demostración. Consideremos en primer lugar el caso $1 \leq p < \infty$. Sean $h \in (0, 1]$, $a \in \mathbb{R}^d$ y $T_h : \mathbb{R}^d \rightarrow \mathbb{R}^d$ la transformación afín definida por $T_h x = hx + a$ tal que $T_h(Q) = I_h$. Sea $f \in W^{m+1,p}(Q)$. Veremos que la estimación (3.4) se sigue de aplicar el Lema 3.2 a $f_h \doteq f \circ T_h$. En efecto, notemos en primer lugar que

$$\begin{aligned} |f_h|_{W^{m+1,p}(Q)} &= \sum_{|\alpha|=m+1} \left(\int_Q |\partial_x^\alpha f(hx + a)|^p dx \right)^{1/p} \\ &= \sum_{|\alpha|=m+1} \left(\int_Q |h^{m+1} f^{(\alpha)}(hx + a)|^p dx \right)^{1/p} \end{aligned}$$

$$\begin{aligned}
&= h^{m+1} \sum_{|\alpha|=m+1} \left(\int_{I_h} |f^{(\alpha)}(z)|^p h^{-d} dz \right)^{1/p} \\
&= h^{m+1-\frac{d}{p}} |f|_{W^{m+1,p}(I_h)}. \tag{3.5}
\end{aligned}$$

De manera exactamente análoga, se prueba que para todo $g \in \Pi_m(Q)$

$$\|f - g\|_{L^p(I_h)} = h^{\frac{d}{p}} \|f_h - g \circ T_h\|_{L^p(Q)}. \tag{3.6}$$

Luego, se sigue que

$$\begin{aligned}
\inf_{g \in \Pi_m(I_h)} \|f - g\|_{L^p(I_h)} &\leq \inf_{g \in \Pi_m(Q)} \|f - g\|_{L^p(I_h)} && \text{(pues } \Pi_m(Q) \subset \Pi_m(I_h)) \\
&= h^{\frac{d}{p}} \inf_{g \in \Pi_m(Q)} \|f_h - g \circ T_h\|_{L^p(Q)} && \text{(en virtud de (3.6))} \\
&= h^{\frac{d}{p}} \inf_{\tilde{g} \in \Pi_m(I_h)} \|f_h - \tilde{g}\|_{L^p(Q)} && (\tilde{g} \doteq g \circ T_h) \\
&\leq h^{\frac{d}{p}} \inf_{\tilde{g} \in \Pi_m(Q)} \|f_h - \tilde{g}\|_{L^p(Q)}. && \text{(por } \Pi_m(Q) \subset \Pi_m(I_h)) \tag{3.7}
\end{aligned}$$

Por otro lado, puesto que $f \in W^{m+1,p}(Q)$ del Lema 3.2 se tiene que

$$\inf_{g \in \Pi_m(Q)} \|f_h - g\|_{L^p(Q)} \leq C |f_h|_{W^{m+1,p}(Q)}. \tag{3.8}$$

Finalmente (3.4) se sigue concatenando las estimaciones (3.7), (3.8) y (3.5).

Supongamos ahora que $p = \infty$ y sea $f \in W^{m+1,\infty}(Q)$. Queremos probar que existe C independiente de h y de f tal que

$$\inf_{g \in \Pi_m(I_h)} \sup_{x \in I_h} |f(x) - g(x)| \leq C h^{m+1} \sum_{|\alpha|=m+1} \sup_{x \in I_h} |\partial_x^\alpha f(x)|. \tag{3.9}$$

Las ecuaciones (3.5) y (3.6) tienen su correlato en el caso $p = \infty$. En efecto, siguiendo pasos análogos se prueba fácilmente que

$$|f_h|_{W^{m+1,\infty}(Q)} = \sum_{|\alpha|=m+1} \sup_{x \in Q} |\partial_x^\alpha f_h(x)| = h^{m+1} \sum_{|\alpha|=m+1} \sup_{z \in I_h} |f^\alpha(z)| = h^{m+1} |f|_{W^{m+1,\infty}(I_h)} \tag{3.10}$$

y también la identidad análoga a (3.6) para el caso $p = \infty$, es decir

$$\|f - g\|_{L^\infty(I_h)} = \sup_{z \in I_h} |f(z) - g(z)| = \sup_{x \in Q} |f_h(x) - g \circ T_h(x)| = \|f_h - g \circ T_h\|_{L^\infty(Q)}. \tag{3.11}$$

Tomando ahora ínfimo para $g \in \Pi_m(I_h)$ en (3.11) y siguiendo pasos análogos a los que condujeron a (3.7), obtenemos

$$\inf_{g \in \Pi_m(I_h)} \sup_{z \in I_h} |f(z) - g(z)| \leq \inf_{g \in \Pi_m(Q)} \sup_{x \in Q} |f_h(x) - g(x)|. \tag{3.12}$$

Por otro lado, del Lema 3.2 con $p = \infty$ se sigue que existe una constante C tal que

$$\inf_{g \in \Pi_m(Q)} \sup_{x \in Q} |f_h(x) - g(x)| \leq C \sum_{|\alpha|=m+1} \sup_{x \in Q} |\partial^\alpha f_h(x)|. \quad (3.13)$$

Finalmente, la desigualdad (3.9) se sigue inmediatamente concatenando (3.12), (3.13) y (3.10). ■

El corolario anterior nos permitirá demostrar el teorema que presentamos a continuación, el cual puede verse como una versión generalizada del Teorema 3.2.2 en [10]. Este teorema permite obtener una cota para el error de proyección wavelet sobre V_j en términos de seminormas en los espacios de Sobolev $H^t(\Omega)$, para $t \leq n + 1$, donde n es el grado de reproducción polinomial en V_j .

Teorema 3.4. *Sean $\phi \in L^2(\Omega)$, V_j como antes, n el grado de reproducción polinomial en V_j y $t \in \mathbb{N}$ tal que $t \leq n + 1$. Entonces existe una constante C independiente de j tal que para toda $f \in H^n(\Omega)$ se verifica que $\|f - P_j f\|_{L^2(\Omega)} \leq C 2^{-tj} |f|_{H^t(\Omega)}$.*

Demostración. Por simplicidad haremos la demostración solo para el caso $\Omega = Q = [0, 1]^d$. El caso general se sigue de considerar un cubo \tilde{Q} que contenga al dominio Ω y utilizar una transformación afín T tal que $T(\tilde{Q}) = Q$.

Sea $j \in \mathbb{N}$ fijo. Para $k \in \mathbb{Z}^d$ definimos los cubos diádicos $I_{j,k} \doteq 2^{-j}(k + Q)$ y el conjunto $\mathcal{K} \subset \mathbb{Z}^d : \bigcup_{k \in \mathcal{K}} I_{j,k} = Q$. Por definición de ínfimo, existe un polinomio $p_{j,k} \in \Pi_{t-1}(I_{j,k})$ tal que

$$\|f - p_{j,k}\|_{L^2(I_{j,k})} \leq 2 \inf_{p \in \Pi_{t-1}(I_{j,k})} \|f - p\|_{L^2(I_{j,k})} \leq C_1 2^{-tj} |f|_{H^t(I_{j,k})}, \quad (3.14)$$

con C_1 independiente de f , j y k , donde la última desigualdad se sigue del Corolario 3.3.

Luego

$$\begin{aligned} \|f - P_j f\|_{L^2(I_{j,k})} &\leq \|f - p_{j,k}\|_{L^2(I_{j,k})} + \|p_{j,k} - P_j f\|_{L^2(I_{j,k})} \\ &= \|f - p_{j,k}\|_{L^2(I_{j,k})} + \|P_j(p_{j,k} - f)\|_{L^2(I_{j,k})} \quad (t-1 \leq n-1) \\ &\leq 2\|f - p_{j,k}\|_{L^2(I_{j,k})} \\ &\leq 2C_1 2^{-tj} |f|_{H^t(I_{j,k})}. \quad (\text{por 3.14}) \end{aligned}$$

Finalmente, el teorema se sigue de la desigualdad anterior sumando sobre todos los índices $k \in \mathcal{K}$. ■

A continuación demostramos el teorema principal de esta sección, el que será de suma importancia para varios de los resultados que presentaremos más adelante.

Teorema 3.5 (Orden de convergencia del error de proyección). Sean $t, s \in \mathbb{N}_0$ con $0 \leq s \leq t$, $\phi \in H^s(\Omega)$ y $n \in \mathbb{N}$ tal que $t < n$ donde $n - 1$ es el grado de reproducción polinomial en V_j . Entonces existe una constante C independiente de j , tal que para toda $f \in H^t(\Omega)$ se verifica

$$\|f - P_j f\|_{H^s(\Omega)} \leq C 2^{-j(t-s)} \|f\|_{H^t(\Omega)}.$$

Demostración. El caso $s = t$ se sigue trivialmente, supongamos entonces que $s < t$. Tal como vimos en el Capítulo 1, utilizando la descomposición multiescala de f podemos expresar el error de proyección sobre V_j de la siguiente manera

$$f - P_j f = \sum_{\ell=j}^{\infty} g_\ell = \sum_{\ell=j}^{\infty} P_{\ell+1} f - P_\ell f.$$

Tomando norma en $H^s(\Omega)$ en ambos miembros de la igualdad y aplicando desigualdad triangular en el lado derecho obtenemos

$$\|f - P_j f\|_{H^s(\Omega)} \leq \sum_{\ell=j}^{\infty} \|P_{\ell+1} f - P_\ell f\|_{H^s(\Omega)}. \quad (3.15)$$

Por otro lado

$$\begin{aligned} \|P_{\ell+1} f - P_\ell f\|_{H^s(\Omega)} &\leq C_1 2^{s\ell} \|P_{\ell+1} f - P_\ell f\|_{L^2(\Omega)} && \text{(en virtud del Lema 3.1)} \\ &\leq C_1 2^{s\ell} (\|P_{\ell+1} f - f\|_{L^2(\Omega)} + \|f - P_\ell f\|_{L^2(\Omega)}) \\ &\leq C_1 C_2 2^{s\ell} \left(2^{-\ell t} |f|_{H^t(\Omega)} + 2^{-(\ell+1)t} |f|_{H^t(\Omega)} \right) && \text{(por Teorema 3.4)} \\ &\leq C_1 C_2 2^{s\ell} 2^{-\ell t} |f|_{H^t(\Omega)} (1 + 2^{-t}) \\ &\leq 2 C_1 C_2 2^{-\ell(t-s)} |f|_{H^t(\Omega)}. \end{aligned}$$

Finalmente, reemplazando esta desigualdad en (3.15) y considerando $C_3 = 2 C_1 C_2$ obtenemos la tesis

$$\begin{aligned} \|f - P_j f\|_{H^s(\Omega)} &\leq C_3 \sum_{\ell=j}^{\infty} 2^{-\ell(t-s)} |f|_{H^t(\Omega)} \\ &= C_3 |f|_{H^t(\Omega)} \sum_{\ell=0}^{\infty} \left(\frac{1}{2^{t-s}} \right)^{\ell+j} \\ &= C_3 |f|_{H^t(\Omega)} \left(\frac{1}{2^{t-s}} \right)^j \left(\frac{1}{1 - \frac{1}{2^{t-s}}} \right) \\ &\leq 2 C_3 2^{-(t-s)} |f|_{H^t(\Omega)}. \quad \blacksquare \end{aligned}$$

3.2.2. Estimación del error de umbralado wavelet

En esta sección estudiaremos la convergencia para el error de umbralado wavelet. Enunciamos a continuación el teorema principal que nos proporciona condiciones suficientes que garantizan un orden de convergencia “apropiado” para la esperanza del error de umbralado.

Teorema 3.6 (Orden de convergencia del error de umbralado). *Sean $\theta, \eta \in \mathbb{N}$ tales que $0 < \eta \leq \theta$, sean además $g \in L^2(\Omega)$ y $g^\delta \doteq g + \delta dW$, donde dW es ruido blanco. Supongamos que la función de escala ϕ satisface $\phi \in H^\eta(\Omega)$ y sean, a su vez, P_j el correspondiente operador proyección sobre V_j y S_γ el operador de umbralado. Entonces, eligiendo $\gamma \doteq C\delta\sqrt{|\log \delta|}$ (C una constante positiva y arbitraria pero fija) y $j \leq j_0 \doteq \lfloor -\frac{\log(\delta^2|\log \delta|)}{\log 2} \rfloor$, se tiene que*

$$E(\|S_\gamma P_j g^\delta - P_j g\|_{H^\eta(\Omega)}^2) = O\left((\delta\sqrt{|\log \delta|})^{\frac{4(\theta-\eta)}{(2\theta+d)}}\right).$$

Para demostrar este teorema serán necesarios algunos resultados previos que también enunciamos a continuación.

Lema 3.7. *Sean $\theta, \eta \in \mathbb{N}$ tales que $0 < \eta \leq \theta$, $q \doteq 2(2\eta + d)/(2\theta + d)$ y supongamos que la función de escala $\phi \in B_{2,2}^\eta(\Omega)$, Entonces, para cada $j \in \mathbb{N}$, $V_j = P_j(L^2(\Omega)) \subseteq B_{q,q}^\theta(\Omega)$.*

Demostración. Sean $j \in \mathbb{N}$ y $g \in L^2(\Omega)$. Dado que Ω es un dominio acotado, el espacio $V_j \doteq \text{span}\{\phi_{j,k} : k \in \mathbb{Z}^d\}$ es finito dimensional. Por lo tanto, la proyección de g sobre V_j , dada por $P_j g = \sum_k \beta_{j,k} \phi_{j,k}$, donde $\beta_{j,k} = \langle g, \phi_{j,k} \rangle$, tiene un número finito de coeficientes no nulos y en consecuencia $\{\beta_{j,k}\}_{k \in \mathbb{Z}^d} \in \ell^q$.

Por otro lado, puesto que $\phi \in B_{2,2}^\eta(\Omega)$ y $q < 2$ el espacio de funciones cuyos coeficientes wavelets $\{\beta_{j,k}\}_{k \in \mathbb{Z}^d} \in \ell^q$ coincide con el espacio de Besov $B_{q,q}^s(\Omega)$ donde s está definido por $s \doteq \frac{2(2\eta+d)}{2q} - \frac{d}{2}$ (ver [8], § 3). Ahora, para $q = 2(2\eta + d)/(2\theta + d) < 2$ se tiene que $s = \theta$, y por lo tanto $P_j g \in B_{q,q}^\theta(\Omega)$. ■

Corolario 3.8. *Bajo las condiciones del lema anterior se tiene que si $g \in L^2(\Omega)$ entonces $P_j g \in WB_{q,q}^\theta$ para todo $j \in \mathbb{N}$.*

Demostración. Como vimos en la Sección 1.5.1, este corolario se sigue inmediatamente del hecho que $B_{q,q}^\theta(\Omega) \subset WB_{q,q}^\theta(\Omega)$. ■

Teorema 3.9. *Sean $0 < p < \infty$, $\alpha \in (0, 1)$ y $q \doteq (1 - \alpha)p$. Supongamos que $\phi \in H^\eta(\Omega)$ para algún $0 < \eta$ y sean P_j el operador proyección sobre V_j y S_γ el operador de umbralado. Entonces eligiendo $\gamma = \gamma(\delta) \doteq C_0 \delta\sqrt{|\log \delta|}$ (C_0 una constante positiva y arbitraria pero fija) y $j_0 = \lfloor -\frac{\log(\delta^2|\log \delta|)}{\log 2} \rfloor$, se tiene que para todo $j \leq j_0$ el espacio de las funciones $g \in B_{p,p}^\eta(\Omega)$ tales*

que

$$E(\|S_\gamma P_j g^\delta - g\|_{B_{p,p}^\eta(\Omega)}^p) = O\left((\delta\sqrt{|\log \delta|})^{\alpha p}\right) \quad (3.16)$$

coincide con el espacio débil $WB_{q,q}^s(\Omega)$ con $s = \frac{p(2\eta+d)}{2q} - \frac{d}{2}$.

Demostración. Utilizaremos el Teorema 2 en [8] con $\eta = \sigma$, $\delta \approx \frac{1}{\sqrt{n}}$, $\gamma = \gamma(\delta) \approx c(n)$, $\gamma_n = \emptyset$, $p_j = 2^{j(\eta+d/2-d/p)}$ y $\Lambda_n = \Lambda_\delta = \{j \in \mathbb{N} : 2^{j(\eta+d/2-d/p)} \leq \gamma^{-2}\}$.

Para ello deberemos probar que se satisfacen las siguientes condiciones:

- i) $\lim_{n \rightarrow \infty} c(n) = 0$ y $\limsup_{n \rightarrow \infty} \frac{c(n)}{c(n+1)} < \infty$.
- ii) Existe $\tau > -1$ tal que para todo $r > 0$ se satisface $\#\{j \in \mathbb{N} : 2^{j(\eta+d/2-d/p)} \leq r\} \leq Cr^\tau$, donde C es una constante positiva independiente de r y τ .
- iii) Existe una constante $C > 0$ tal que para todo $j \in \Lambda_\delta$ se verifica que $E\left(|\beta_{j,k}^\delta - \beta_{j,k}|^{2p}\right) \leq C\gamma^{2p}$, donde $\beta_{j,k}^\delta = \langle g^\delta, \phi_{j,k} \rangle$ y $\beta_{j,k} = \langle g, \phi_{j,k} \rangle$ son los coeficientes wavelet de g^δ y g respectivamente.
- iv) Existen constantes positivas κ , C y $\gamma \geq \alpha p$, tales que $P\left(|\beta_{j,k}^\delta - \beta_{j,k}| \geq \frac{\kappa\gamma}{2}\right) \leq C\gamma^\gamma$ para todo $j \in \Lambda_\delta$ y para todo $0 < \gamma \leq e^{-1}$.

Observar que las condiciones ii)-iv) corresponden, respectivamente, a las hipótesis (30), (a) y (b) en [8].

La condición i) se sigue inmediatamente de la definición de $c(n)$.

Veremos ahora que ii) se satisface para todo $\tau \geq 1$, en particular para $\tau = 1$ y $C \doteq [(\log 2)(\eta + d/2 - d/p)]^{-1}$. En efecto, Si $j \in \mathbb{N}$ es tal que $2^{j(\eta+d/2-d/p)} \leq r$, entonces $j \leq \frac{\log r}{(\log 2)(\eta+d/2-d/p)} \leq \frac{r}{(\log 2)(\eta+d/2-d/p)}$. Por lo tanto $\#\{j : 2^{j(\eta+d/2-d/p)} \leq r\} \leq \frac{r}{(\log 2)(\eta+d/2-d/p)} = Cr$.

Para probar iii) observemos en primer lugar que puesto que para cada $j \in \mathbb{N}$ y $k \in \mathbb{Z}^d$ $\beta_{j,k}^\delta - \beta_{j,k} = \delta \langle dW, \phi_{j,k} \rangle$, se sigue que $\beta_{j,k}^\delta - \beta_{j,k} \sim N(0, \delta^2)$ y por lo tanto $X \doteq \delta^{-1}(\beta_{j,k}^\delta - \beta_{j,k}) \sim N(0, 1)$. Por lo tanto,

$$\begin{aligned} E(|\beta_{j,k}^\delta - \beta_{j,k}|^{2p}) &= E(|\delta X|^{2p}) = \delta^{2p} E(|X|^{2p}) = \frac{2\delta^{2p}}{\sqrt{2\pi}} \int_0^\infty x^{2p} e^{-\frac{x^2}{2}} dx \\ &= \frac{2^p \delta^{2p}}{\sqrt{\pi}} \int_0^\infty t^{(p+\frac{1}{2})-1} e^{-t} dt = \frac{2^p \Gamma(p + \frac{1}{2})}{\sqrt{\pi}} \delta^{2p} \\ &\leq \frac{2^p \Gamma(p + \frac{1}{2})}{\sqrt{\pi}} (\delta\sqrt{|\log \delta|})^{2p}, \end{aligned} \quad (\text{para todo } 0 < \delta \leq e^{-1}).$$

Luego $E\left(|\beta_{j,k}^\delta - \beta_{j,k}|^{2p}\right) \leq C\gamma^{2p}$ donde $C \doteq \frac{2^p \Gamma(p+1/2)}{\sqrt{\pi}}$.

Finalmente para probar iv) utilizaremos un resultado clásico de probabilidad (ver Apéndice A, proposición 5.1), que asegura que si $Z \sim N(0, 1)$, entonces para todo $\omega > 0$ se verifica que

$$P(|Z| > \omega) \leq 3e^{-\frac{\omega^2}{2}}. \quad (3.17)$$

Elijamos κ cualquier contante tal que $\kappa \geq \frac{2\sqrt{2\alpha p}}{C_0}$ y sea $\gamma \doteq \frac{C_0^2 \kappa^2}{8}$. Entonces para todos $j \in \Lambda_\delta$ y $0 < \delta \leq e^{-1}$ se tiene que

$$\begin{aligned} P\left(|\beta_{j,k}^\delta - \beta_{j,k}|\right) &\geq \frac{\kappa\gamma}{2} = P\left(|X| \geq \frac{\kappa\gamma}{2\delta}\right) = P\left(|X| \geq \frac{\kappa\sqrt{|\log \delta|}}{2}\right) && \text{(por (3.17))} \\ &\leq 3e^{-\frac{\kappa^2|\log \delta|}{8}} = 3e^{-\frac{\kappa^2 \log \delta}{8}} = 3\delta^\gamma \\ &\leq 3(\delta\sqrt{|\log \delta|})^\gamma = 3\gamma^\gamma. && \left(\gamma \in (0, e^{-1}] \Rightarrow \delta \in (0, e^{-1}]\right) \end{aligned}$$

Luego se verifica iv) con κ y γ como antes y $C = 3$. ■

Estamos ahora en condiciones de demostrar el teorema principal de esta sección.

Demostración del Teorema 1.6. La demostración se sigue ahora inmediatamente del Corolario 3.8 y del Teorema 3.9. En efecto, dados $q = \frac{2(2\eta+d)}{2\theta+d}$, $\phi \in B_{2,2}^\eta(\Omega)$ y $g \in L^2(\Omega)$ por el corolario podemos afirmar que $P_j g \in WB_{q,q}^\theta(\Omega)$. Luego, eligiendo $\alpha = \frac{2(\theta-\eta)}{2\theta+d}$ y $p = 2$ en el Teorema 3.9 obtenemos que $s = \frac{2(2\eta+d)}{2q} - \frac{d}{2} = \theta$. Finalmente, se sigue del Teorema 3.9 que $P_j g \in B_{2,2}^\eta(\Omega)$ y, de (3.16), que

$$E(\|S_\gamma P_j g^\delta - P_j g\|_{B_{2,2}^\eta(\Omega)}^2) = O\left((\delta\sqrt{|\log \delta|})^{\frac{4(\theta-\eta)}{2\theta+d}}\right). \quad \blacksquare$$

3.2.3. Estimación del error de regularización

En esta sección estudiaremos la convergencia de los errores de regularización para dos métodos de regularización espectral. En particular, nos proponemos probar resultados sobre la convergencia de errores al utilizar el método iterativo de Landweber y el método TSVD. Estos resultados permitirán probar que los respectivos métodos en dos pasos alcanzan un orden de convergencia casi-óptimo, en el sentido ya descrito.

Para ambos métodos de regularización probaremos dos lemas relacionados a la convergencia y los correspondiente órdenes de convergencia de los errores generados por la restauración obtenida por dichos métodos. En primer lugar encontraremos una cota para la norma del operador de regularización R_α visto como un operador en $\mathcal{L}(H^\eta(\Omega); L_2(\Omega))$. Luego, estimaremos una cota para el error de regularización concreto generado por el método elegido.

Cotas para el método iterativo de Landweber

Como vimos en la Sección 2.6 se puede probar que el método de regularización iterativo de Landweber es también un método espectral asociado a la familia espectral discreta $g_k(\lambda) \doteq \sum_{j=0}^{k-1} (1-\lambda)^j$, con $0 < \lambda < 2$.

En este caso, supondremos sin pérdida de generalidad que $\|T\| \leq 1$. Si ese no fuera el caso podemos redefinir al operador dividiéndolo por su norma. A continuación, probaremos un lema que nos provee una cota para la norma del operador de regularización.

Lema 3.10. Sean $\eta \in \mathbb{N}$ y $t \in \mathbb{R}^+$ tales que $0 < \eta < t$. Sean, además, $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador lineal, acotado con $\|T\| \leq 1$, compacto, con la propiedad de t -suavizado y $R_k = g_k(T^*T)T^*$ el operador correspondiente a la k -ésima iteración de Landweber. Entonces existe una constante C , independiente de k , que verifica

$$\|R_k\|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))} \leq C k^{\frac{t-\eta}{2t}}.$$

Demostración. Sea $\{(\sigma_n; u_n; v_n)\}$ el sistema singular asociado al operador T . Como vimos en el Capítulo 2, en la ecuación (2.44), el operador de regularización asociado al método iterativo de Landweber puede expresarse de forma no recursiva como

$$R_k y = \sum_{j=0}^{k-1} \sum_{\sigma_n > 0} (1 - \sigma_n^2)^j \sigma_n \langle y, v_n \rangle u_n.$$

Luego, para todo $y \in H^\eta(\Omega)$ se sigue que

$$\begin{aligned} \|R_k y\|_{L_2(\Omega)}^2 &= \sum_{\sigma_n > 0} \left(\sum_{j=0}^{k-1} (1 - \sigma_n^2)^j \sigma_n \right)^2 |\langle y, v_n \rangle|^2 \sigma_n^{\frac{2\eta}{t}} \sigma_n^{-\frac{2\eta}{t}} \\ &\leq \sup_{0 < \sigma \leq 1} \left(\sum_{j=0}^{k-1} (1 - \sigma^2)^j \sigma^{(1+\frac{\eta}{t})} \right)^2 \sum_{\sigma_n > 0} \sigma_n^{-\frac{2\eta}{t}} |\langle y, v_n \rangle|^2 \\ &= \sup_{0 < \sigma \leq 1} \left(\sum_{j=0}^{k-1} (1 - \sigma^2)^j \sigma^{\frac{t+\eta}{t}} \right)^2 \|y\|_{H^\eta(\Omega)}^2 \quad (\text{en virtud de (3.3)}) \\ &\leq \left(\sum_{j=0}^{k-1} \left(\frac{2jt}{2jt + t + \eta} \right)^j \left(\frac{t + \eta}{2jt + t + \eta} \right)^{\frac{t+\eta}{2t}} \right)^2 \|y\|_{H^\eta(\Omega)}^2 \quad \left(\text{para } \sigma^2 \doteq \frac{t + \eta}{2jt + t + \eta} \right) \\ &\leq \left(\sum_{j=0}^{k-1} \left(\frac{t + \eta}{2jt + t + \eta} \right)^{\frac{t+\eta}{2t}} \right)^2 \|y\|_{H^\eta(\Omega)}^2 \\ &\leq \|y\|_{H^\eta(\Omega)}^2 \left(\frac{t + \eta}{t} \right)^{\frac{t+\eta}{t}} \left(\sum_{j=0}^{k-1} \left(\frac{1}{j+1} \right)^{\frac{t+\eta}{2t}} \right)^2 \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{t+\eta}{t}\right)^{\frac{t+\eta}{t}} \|y\|_{H^\eta(\Omega)}^2 \left(\sum_{j=1}^k \left(\frac{1}{j}\right)^{\frac{t+\eta}{2t}}\right)^2 \\
&\leq \left(\frac{t+\eta}{t}\right)^{\frac{t+\eta}{t}} \|y\|_{H^\eta(\Omega)}^2 \left(\int_0^k x^{-\frac{t+\eta}{2t}} dx\right)^2 \\
&\leq \left(\frac{t+\eta}{t}\right)^{\frac{t+\eta}{t}} \|y\|_{H^\eta(\Omega)}^2 \left(\frac{t-\eta}{2t}\right)^{-2} k^{\frac{t-\eta}{t}} \\
&\leq C^2 \|y\|_{H^\eta(\Omega)}^2 k^{\frac{t-\eta}{t}},
\end{aligned}$$

donde $C^2 \doteq \left(\frac{t+\eta}{t}\right)^{\frac{t+\eta}{t}} \left(\frac{2t}{t-\eta}\right)^2$.

Se sigue entonces que $\|R_k\|_{\mathcal{L}(H^\eta(\Omega);L_2(\Omega))} \leq C k^{\frac{t-\eta}{2t}}$. \blacksquare

Demostraremos ahora un lema que nos provee una cota para el error de regularización utilizando este método iterativo de Landweber.

Lema 3.11. Sean $t \geq 0$, $T : L^2(\Omega) \rightarrow L^2(\Omega)$ y R_k como en el lema anterior. Entonces para todo $s \in [0, t]$ existe una constante C , que depende de s , tal que

$$\|R_k T - I\|_{\mathcal{L}(H^s(\Omega);L_2(\Omega))} \leq C k^{-\frac{s}{2t}}. \quad (3.18)$$

Demostración. Sean $f \in H^s(\Omega)$, $g \doteq T f$ y $\{(\sigma_n; u_n; v_n)\}$ el sistema singular asociado al operador T . Entonces por definición del operador R_k asociado al método iterativo de Landweber se tiene que

$$R_k g - f = (I - T^* T)^k f = \sum_{\sigma_n > 0} (1 - \sigma_n^2)^k \langle f, u_n \rangle u_n.$$

Luego, se sigue que

$$\begin{aligned}
\|R_k g - f\|_{L_2(\Omega)}^2 &= \sum_{\sigma_n > 0} (1 - \sigma_n^2)^{2k} |\langle f, u_n \rangle|^2 \sigma^{\frac{2s}{t}} \sigma^{-\frac{2s}{t}} \\
&\leq \sup_{0 < \sigma \leq 1} (1 - \sigma^2)^{2k} \sigma^{\frac{2s}{t}} \|f\|_{H^s(\Omega)}^2 && \text{(por (3.3))} \\
&= \left(\frac{2kt}{2kt + s}\right)^{2k} \left(\frac{s}{2kt + s}\right)^{\frac{s}{t}} \|f\|_{H^s(\Omega)}^2 && \left(\text{donde } \sigma^2 = \frac{s}{2kt + s}\right) \\
&\leq C^2 (1/k)^{\frac{s}{t}} && \left(\text{para } C^2 \doteq s^{\frac{s}{t}} \|f\|_{H^s(\Omega)}^2\right)
\end{aligned}$$

Esto implica que $\|R_k g - f\|_{L_2(\Omega)} \leq C k^{-\frac{s}{2t}}$, de donde se sigue la acotación (3.18). \blacksquare

Cotas para el método de descomposición en valores singulares truncada

A continuación probaremos dos resultados homólogos a los Lemas 3.10 y 3.11, pero ahora con el método de descomposición en valores singulares truncada.

Lema 3.12. Sean $\eta \in \mathbb{N}$ y $t \in \mathbb{R}^+$ tales que $0 < \eta < t$. Sean además $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador lineal, acotado, compacto y con la propiedad de t -suavizado y para $\alpha > 0$, R_α el operador correspondiente al método TSVD (ver (2.39)). Entonces

$$\| R_\alpha \|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))} \leq \alpha^{\frac{\eta-t}{2t}}.$$

Demostración. Dado $\{(\sigma_n; u_n; v_n)\}$ el sistema singular asociado al operador T , como vimos en el Capítulo 2, el operador correspondiente al método TSVD para $\alpha > 0$ puede expresarse como

$$R_\alpha y = \sum_{n: \sigma_n^2 \geq \alpha} \frac{1}{\sigma_n} \langle y, v_n \rangle u_n.$$

Luego, para todo $y \in H^\eta(\Omega)$ se tiene que

$$\begin{aligned} \| R_\alpha y \|_{L_2(\Omega)}^2 &= \sum_{n: \sigma_n^2 \geq \alpha} \frac{1}{\sigma_n^2} |\langle y, v_n \rangle|^2 \sigma_n^{\frac{2\eta}{t}} \sigma_n^{\frac{-2\eta}{t}} \leq \sup_{\sigma \geq \sqrt{\alpha}} \sigma^{\frac{2\eta}{t}-2} \sum_{n: \sigma_n^2 \geq \alpha} \sigma_n^{\frac{-2\eta}{t}} |\langle y, v_n \rangle|^2 \\ &\leq \| y \|_{H^\eta(\Omega)}^2 \sup_{\sigma \geq \sqrt{\alpha}} \sigma^{\frac{2\eta}{t}-2} \leq \| y \|_{H^\eta(\Omega)}^2 \alpha^{\frac{\eta-t}{t}}. \end{aligned}$$

De donde se sigue que $\| R_\alpha \|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))} \leq \alpha^{\frac{\eta-t}{2t}}$. ■

A continuación probamos un lema que nos provee una cota para el error de regularización generado utilizando el método TSVD.

Lema 3.13. Sean $t \geq 0$, $T : L^2(\Omega) \rightarrow L^2(\Omega)$ y R_α como en el lema anterior. Entonces para todo $s \in [0, t]$ se verifica que

$$\| R_\alpha T - I \|_{\mathcal{L}(H^s(\Omega); L_2(\Omega))} \leq \alpha^{\frac{s}{2t}}. \quad (3.19)$$

Demostración. Sean $f \in H^s(\Omega)$, $g \doteq Tf$ y $\{(\sigma_n; u_n; v_n)\}$ el sistema singular asociado al operador T . Entonces por definición del operador R_α asociado al método TSVD se tiene

$$\begin{aligned} R_\alpha g &= \sum_{\sigma_n \geq \sqrt{\alpha}} \frac{1}{\sigma_n} \langle g, v_n \rangle u_n = \sum_{\sigma_n > \sqrt{\alpha}} \frac{1}{\sigma_n^2} \langle T^* g, u_n \rangle u_n \\ &= \sum_{\sigma_n > \sqrt{\alpha}} \frac{1}{\sigma_n^2} \langle T^* T f, u_n \rangle u_n = \sum_{\sigma_n > \sqrt{\alpha}} \langle f, u_n \rangle u_n. \end{aligned}$$

Luego,

$$\| R_\alpha g - f \|_{L_2(\Omega)}^2 = \sum_{0 < \sigma_n < \sqrt{\alpha}} |\langle f, u_n \rangle|^2 \sigma_n^{\frac{2s}{t}} \sigma_n^{\frac{-2s}{t}} \leq \sup_{0 < \sigma < \sqrt{\alpha}} \sigma^{\frac{2s}{t}} \| f \|_{H^s(\Omega)}^2 = \| f \|_{H^s(\Omega)}^2 \alpha^{\frac{s}{t}}.$$

Lo cual implica que $\| R_\alpha T f - f \|_{L_2(\Omega)} \leq \alpha^{\frac{s}{2t}}$, de donde se sigue (3.19). ■

3.2.4. Convergencia del método wavelet-Landweber

En esta sección probaremos uno de los teoremas principales de este capítulo. El resultado muestra que el método de regularización wavelet-Landweber converge con orden casi-óptimo, en el sentido de (2.30). Este teorema constituye el punto de partida para la generalización que presentaremos en la siguiente sección a otros métodos espectrales.

Realizamos la demostración utilizando el método de regularización de Landweber, puesto que nos parece interesante exponer los resultados para un método iterativo que no depende de un parámetro de regularización continuo α , como los demás métodos espectrales. La demostración para el método TSVD se sigue de manera análoga en virtud de los resultados con órdenes de convergencia equivalentes, obtenidos en la subsección precedente.

Como vimos en la Sección 2.3, al hacer referencia a la tasa de convergencia de un método de regularización (R_α, α) nos referimos a la velocidad con la cual tiende a cero el error de aproximación $\|R_\alpha g - f\|$ cuando $\alpha \rightarrow 0^+$, o a la velocidad con la cual tiende a cero el error total $\|R_\alpha g^\delta - f\|$ cuando $\delta \rightarrow 0^+$. En dicha sección hemos definido también el error asociado al ruido $\|R_\alpha g^\delta - R_\alpha g\|$ y vimos que estos tres errores están relacionados por la desigualdad

$$\|R_\alpha g^\delta - f\| \leq \|R_\alpha(g^\delta - g)\| + \|R_\alpha g - f\|.$$

Sin embargo, en el método de dos pasos wavelet-Landweber la solución regularizada implica un paso más, debido al umbralado wavelet. Recordemos que la aproximación del método está dada por $f_{reg}^\delta = R_k S_\gamma P_j g^\delta$. En este caso, para el error total se prueba fácilmente que

$$\|R_k S_\gamma P_j g^\delta - f\| \leq \|R_k(S_\gamma P_j g^\delta - P_j g)\| + \|R_k(P_j g - g)\| + \|R_k g - f\|. \quad (3.20)$$

Nos referiremos a los tres errores en el lado derecho de (3.20) como errores de umbralado, de proyección y de regularización, respectivamente. Para llegar al teorema principal de esta sección, acotaremos previamente cada uno de estos errores.

Los siguientes dos lemas nos brindan cotas para los errores de proyección y umbralado, respectivamente.

Lema 3.14. Sean $s \geq 0$ y $t > 0$, tales que $s + t \in \mathbb{N}$. Sean, además, T un operador con las mismas hipótesis que en el Lema 3.10, $g \in H^{s+t}(\Omega)$, $\eta \in \mathbb{N}_0$ tal que $\eta < t$ y $\phi \in H^\eta(\Omega)$. Entonces, para R_k el k -ésimo operador de Landweber, existe una constante C , independiente de k , tal que se tiene la siguiente cota para el error de proyección:

$$\|R_k(P_j g - g)\|_{L_2(\Omega)}^2 \leq C k^{\frac{t-\eta}{t}} 2^{-2j(s+t-\eta)}.$$

Demostración. El resultado se sigue del Lema 3.10 y el Teorema 3.5. En efecto

$$\begin{aligned}
\| R_k(P_j g - g) \|_{L_2(\Omega)}^2 &\leq \| R_k \|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))}^2 \| P_j g - g \|_{L_2(\Omega)}^2 \\
&\leq \| R_k \|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))}^2 \| P_j g - g \|_{H^\eta(\Omega)}^2 \\
&\leq C_1 k^{\frac{t-\eta}{t}} \| P_j g - g \|_{H^\eta(\Omega)}^2 \quad (\text{por el Lema 3.10}) \\
&\leq C_2 k^{\frac{t-\eta}{t}} 2^{-2j(s+t-\eta)} \| g \|_{H^{s+t}(\Omega)}^2 \quad (\text{por el Teorema 3.5}) \\
&\leq C k^{\frac{t-\eta}{t}} 2^{-2j(s+t-\eta)},
\end{aligned}$$

donde C es independiente de k . ■

Lema 3.15. Sean $\eta \in \mathbb{N}_0$, $s \geq 0$ y $t > 0$ tales que $\eta \leq t$ y $s + t \in \mathbb{N}$. Sean, además, el operador T como en los lemas anteriores y R_k el k -ésimo operador de Landweber. Supongamos que $g \in L^2(\Omega)$, $\phi \in H^\eta(\Omega)$, $g^\delta \doteq g + \delta dW$ y denotemos con P_j al correspondiente operador proyección sobre V_j y S_γ el operador de umbralado wavelet. Entonces, eligiendo $\gamma \doteq K\delta\sqrt{|\log \delta|}$ (K una constante positiva y arbitraria pero fija) y $j \leq j_0 \doteq \lfloor -\frac{\log(\delta^2 |\log \delta|)}{\log 2} \rfloor$, existe una constante C independiente de k y δ tal que

$$E(\| R_k(S_\gamma P_j g^\delta - P_j g) \|_{L_2(\Omega)}^2) \leq C k^{\frac{t-\eta}{t}} (\delta\sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}.$$

Demostración. El lema es consecuencia directa del Lema 3.10 y el Teorema 3.6 con $\theta = s + t$. En efecto

$$\begin{aligned}
E(\| R_k(S_\gamma P_j g^\delta - P_j g) \|_{L_2(\Omega)}^2) &\leq \| R_k \|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))}^2 E(\| S_\gamma P_j g^\delta - P_j g \|_{L_2(\Omega)}^2) \\
&\leq \| R_k \|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))}^2 E(\| S_\gamma P_j g^\delta - P_j g \|_{H^\eta(\Omega)}^2) \\
&\leq C_1 k^{\frac{t-\eta}{t}} E(\| S_\gamma P_j g^\delta - P_j g \|_{H^\eta(\Omega)}^2) \quad (\text{por el Lema 3.10}) \\
&\leq C k^{\frac{t-\eta}{t}} (\delta\sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}, \quad (\text{por el Teorema 3.6})
\end{aligned}$$

donde C es una constante independiente de k y de δ . ■

Estamos finalmente en condiciones de demostrar el teorema principal de esta sección.

Teorema 3.16 (Orden de convergencia para el error del método wavelet-Landweber).

Sean $\eta \in \mathbb{N}_0$, $s \geq 0$ y $t > 0$ tales que $0 \leq \eta \leq t$ y $s + t \in \mathbb{N}$. Sean además $f \in H^s(\Omega)$, $g^\delta \doteq Tf + \delta dW$ y $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador lineal, acotado con $\|T\| \leq 1$, compacto y con la propiedad de t -suavizado. Sean además

- i) $\phi \in H^\eta(\Omega)$ la función de escala correspondiente a una wavelet con grado de reproducción polinomial m en V_j tal que $m + 1 > t$,

ii) S_γ el operador de umbralado wavelet con $\gamma \doteq C \delta \sqrt{|\log \delta|}$ y

iii) R_k el k -ésimo operador iterativo de Landweber.

Si el nivel de resolución wavelet j satisface $2^{-j} \leq (\delta \sqrt{|\log \delta|})^{1/(\eta+d/2)}$ y

$$\begin{aligned} k &\simeq (\delta \sqrt{|\log \delta|})^{\frac{-4t}{2s+2t+d}} \quad \text{si } \eta < t, \text{ y} \\ k &\simeq (\delta \sqrt{|\log \delta|})^{-2} \quad \text{si } \eta = t, \end{aligned}$$

entonces

$$E(\| R_k S_\gamma P_j g^\delta - f \|_{L^2(\Omega)}^2) = O\left((\delta \sqrt{|\log \delta|})^{\frac{4s}{2s+2t+d}} \right).$$

Demostración. Notemos en primer lugar que

$$\| R_k S_\gamma P_j g^\delta - f \|_{L^2(\Omega)} \leq \| R_k (S_\gamma P_j g^\delta - P_j g) \|_{L^2(\Omega)} + \| R_k (P_j g - g) \|_{L^2(\Omega)} + \| R_k g - f \|_{L^2(\Omega)}.$$

Luego,

$$\| R_k S_\gamma P_j g^\delta - f \|_{L^2(\Omega)}^2 \leq 3 \left(\| R_k (S_\gamma P_j g^\delta - P_j g) \|_{L^2(\Omega)}^2 + \| R_k (P_j g - g) \|_{L^2(\Omega)}^2 + \| R_k g - f \|_{L^2(\Omega)}^2 \right).$$

Tomando esperanza se sigue entonces que

$$\begin{aligned} E(\| R_k S_\gamma P_j g^\delta - f \|_{L^2(\Omega)}^2) &\leq 3E(\| R_k (S_\gamma P_j g^\delta - P_j g) \|_{L^2(\Omega)}^2) \\ &\quad + 3 \| R_k (P_j g - g) \|_{L^2(\Omega)}^2 + 3 \| R_k g - f \|_{L^2(\Omega)}^2 \\ &\doteq 3(EU^2 + EP^2 + ER^2). \end{aligned} \tag{3.21}$$

Esta desigualdad nos permite entonces dividir el error total del método de dos pasos en lo que denominaremos los errores de umbralado (EU), de proyección (EP) y de regularización (ER).

Caso I: Supongamos en primer lugar que $\eta < t$. Entonces una cota para el primer término en (3.21) correspondiente a EU^2 , se sigue de la aplicación directa del Lema 3.15. En efecto

$$EU^2 = E(\| R_k (S_\gamma P_j g^\delta - P_j g) \|_{L^2(\Omega)}^2) \leq C_1 k^{\frac{t-\eta}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}},$$

para todo

$$j \leq j_0 \doteq \lfloor -\frac{\log(\delta^2 |\log \delta|)}{\log 2} \rfloor. \tag{3.22}$$

Por otro lado, una cota para el segundo término en (3.21) correspondiente a EP^2 , se sigue del Lema 3.14, de donde obtenemos

$$EP^2 = \| R_k (P_j g - g) \|_{L^2(\Omega)}^2 \leq C_2 k^{\frac{t-\eta}{t}} 2^{-2j(s+t-\eta)}.$$

Igualando los órdenes obtenidos para EU^2 y EP^2 (sin considerar las constantes) obtenemos una estimación para el nivel de proyección wavelet. En efecto, de

$$k^{\frac{t-\eta}{t}} 2^{-2j(s+t-\eta)} = k^{\frac{t-\eta}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}},$$

de donde se sigue que $2^{-j} \approx (\delta \sqrt{|\log \delta|})^{2/(2s+2t+d)}$ (consideramos una aproximación pues $j \in \mathbb{N}$).

Suponiendo que $\delta \sqrt{|\log \delta|} \leq 1$, resulta claro que se verifica la condición (3.22) y que además la elección propuesta del nivel de resolución wavelet j es menor que la estimación obtenida

$$2^{-j} \leq (\delta \sqrt{|\log \delta|})^{\frac{1}{\eta+d/2}} \leq (\delta \sqrt{|\log \delta|})^{\frac{2}{2s+2t+d}}.$$

Si bien esta cota para j nos permite realizar el umbralado wavelet con un nivel de menor resolución que el propuesto en la hipótesis de este teorema, utilizaremos $2^{-j} \approx (\delta \sqrt{|\log \delta|})^{\frac{1}{\eta+d/2}}$. De este modo priorizamos realizar una elección de parámetros en la que el nivel de resolución sea independiente del nivel de suavidad de la solución y de la propiedad de t -suavizado del operador. Reemplazando ahora el nivel de proyección en la cota para EP^2 obtenemos

$$\begin{aligned} EU^2 + EP^2 &= E(\| R_k(S_\gamma P_j g^\delta - P_j g) \|_{L_2(\Omega)}^2) + \| R_k(P_j g - g) \|_{L_2(\Omega)}^2 \\ &\leq (C_1 + C_2) k^{\frac{t-\eta}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}. \end{aligned}$$

Resta considerar la cota para el error de regularización ER^2 . Del Lema 3.11, se tiene que

$$ER^2 = \| R_k g - f \|_{L_2(\Omega)}^2 \leq C_3 k^{-\frac{s}{t}}.$$

Ahora igualando los órdenes de convergencia de las estimaciones obtenidas para los errores de umbralado y regularización EU^2 y ER^2 , podemos estimar el número de iteraciones k . En efecto:

$$k^{-\frac{\eta-t}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}} = k^{-\frac{s}{t}},$$

de donde se tiene que $k \approx (\delta \sqrt{|\log \delta|})^{\frac{-4t}{2s+2t+d}}$, ($k \in \mathbb{N}$).

Utilizando las estimaciones obtenidas para los parámetros j y k se tiene finalmente la acotación buscada

$$E(\| R_k S_\gamma P_j g^\delta - f \|_{L_2(\Omega)}^2) \leq 3(EU^2 + EP^2 + ER^2) \leq 3(C_1 + C_2 + C_3) (\delta \sqrt{|\log \delta|})^{\frac{4s}{2s+2t+d}}.$$

Caso II: Supongamos ahora $\eta = t$. Notemos en primer lugar que,

$$\begin{aligned} \| R_k y \|_{L_2(\Omega)}^2 &= \sum_{\sigma_n > 0} \left(\sum_{j=0}^{k-1} (1 - \sigma_n^2)^j \sigma_n \right)^2 |\langle y, v_n \rangle|^2 \sigma_n^2 \sigma_n^{-2} \\ &\leq \sup_{0 < \sigma \leq 1} \left(\sum_{j=0}^{k-1} (1 - \sigma^2)^j \sigma^2 \right)^2 \sum_{n: \sigma_n > 0} \sigma_n^{-2} |\langle y, v_n \rangle|^2 \end{aligned}$$

$$\begin{aligned}
&= \sup_{0 < \sigma \leq 1} \sigma^4 \left(\sum_{j=0}^{k-1} (1 - \sigma^2)^j \right)^2 \|y\|_{H^t(\Omega)}^2 \quad (\text{por (3.3) puesto que } \eta = t) \\
&\leq \sup_{0 < \sigma \leq 1} \sigma^4 \left(\frac{1}{1 - (1 - \sigma^2)} \right)^2 \|y\|_{H^t(\Omega)}^2 \\
&\leq \sup_{0 < \sigma \leq 1} \|y\|_{H^t(\Omega)}^2 \\
&= \|y\|_{H^t(\Omega)}^2.
\end{aligned}$$

Se sigue entonces que

$$\|R_k\|_{\mathcal{L}(H^t(\Omega); L_2(\Omega))} \leq 1. \quad (3.23)$$

Luego,

$$\begin{aligned}
E(\|R_k(S_\gamma P_j g^\delta - P_j g)\|_{L_2(\Omega)}^2) &\leq \|R_k\|_{\mathcal{L}(H^t(\Omega); L_2(\Omega))}^2 E(\|S_\gamma P_j g^\delta - P_j g\|_{L_2(\Omega)}^2) \\
&\leq E(\|S_\gamma P_j g^\delta - P_j g\|_{L_2(\Omega)}^2) \quad (\text{por (3.23)}) \\
&\leq (\delta \sqrt{|\log \delta|})^{\frac{4s}{2s+2t+d}},
\end{aligned}$$

donde la última desigualdad se sigue del Teorema 3.6 con $\theta = s + t$. ■

3.3. Generalización a otros métodos espectrales

En esta sección nos proponemos generalizar algunos de los resultados de la sección anterior en la cual demostramos convergencia y se obtuvieron ordenes de convergencia para el método en dos pasos wavelet-Landweber. Nuestro objetivo es generalizar esos resultados a otros métodos espectrales buscando condiciones que garanticen la optimalidad de los órdenes de convergencia. Para ello presentaremos condiciones suficientes sobre una familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$, para que el correspondiente método de dos pasos wavelet-espectral asociado a esa familia posea un orden de convergencia casi-óptimo. A tal efecto probaremos en primer lugar, resultados análogos a los Lemas 3.10 y 3.11 para familias espectrales generales. Una vez probados estos dos resultados, procederemos a demostrar acotaciones equivalentes a las obtenidas en los Lemas 3.14 y 3.15 para finalmente proceder a enunciar y demostrar el teorema principal de este capítulo.

Consideremos entonces una familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ que satisface las siguientes dos condiciones:

(CE1) Existen constantes positivas C_1 y D independientes de α tal que para todo $\sigma \in (0, D]$ y para todo $\alpha \in (0, \alpha_0)$

$$\sup_{0 < \sigma \leq D} |\sigma^{\frac{1}{2}} g_\alpha(\sigma)| \leq C_1 \alpha^{-\frac{1}{2}}. \quad (3.24)$$

(CE2) Existe una constante $C_2 > 0$ tal que para todo $\alpha \in (0, \alpha_0)$ se satisface

$$\sup_{0 < \sigma < \infty} |1 - \sigma g_\alpha(\sigma)| \sigma \leq C_2 \alpha. \quad (3.25)$$

En relación a la condición (CE1) es oportuno señalar que la mayoría de los métodos espectrales clásicos las satisfacen. En efecto, para los métodos TSVD, Tikhonov y Landweber, se puede probar fácilmente que la acotación (3.24) se satisface con valores de $C_1 = 1$ para los primeros dos y $C_1 = 2$ para Landweber. La segunda condición (CE2) está íntimamente vinculada con el concepto de calificación clásica de un método espectral que definimos en la Sección 2.4. En efecto, utilizando la definición de calificación clásica se puede probar fácilmente que todo método espectral que posea calificación clásica $\mu^* \geq 1$ satisface la condición (CE2).

A continuación, probaremos que la condición (CE1) implica una acotación similar para exponentes de la forma $\frac{a-1}{2}$ para todo $a \in [0, 1]$. Observar que la condición (3.24) corresponde al caso $a = 0$.

Lema 3.17. *Sea $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ una familia espectral que satisface (CE1). Entonces para todo $\sigma \in (0, D]$ y para todo $a \in [0, 1]$ se verifica que*

$$\sup_{0 < \sigma \leq D} |\sigma^{\frac{1+a}{2}} g_\alpha(\sigma)| \leq C_1 \alpha^{\frac{a-1}{2}}. \quad (3.26)$$

Demostración. Consideraremos dos casos posibles:

Caso I: Supongamos en primer lugar que $\frac{\alpha}{\sigma} \geq 1$. Entonces para todo $a \in [0, 1]$ se tiene que $(\frac{\alpha}{\sigma})^{\frac{1-a}{2}} \leq (\frac{\alpha}{\sigma})^{\frac{1}{2}}$ y por lo tanto, para todo $\sigma \in (0, D]$ se verifica

$$\alpha^{\frac{1-a}{2}} |\sigma^{\frac{1+a}{2}} g_\alpha(\sigma)| = \left(\frac{\alpha}{\sigma}\right)^{\frac{1-a}{2}} |\sigma g_\alpha(\sigma)| \leq \left(\frac{\alpha}{\sigma}\right)^{\frac{1}{2}} |\sigma g_\alpha(\sigma)| \leq \alpha^{\frac{1}{2}} |\sigma^{\frac{1}{2}} g_\alpha(\sigma)| \leq C_1,$$

donde la última desigualdad se sigue en virtud de (CE1).

La estimación (3.26) se sigue finalmente multiplicando la desigualdad anterior por $\alpha^{\frac{a-1}{2}}$ y tomando supremo.

Caso II: Supongamos ahora que $\frac{\alpha}{\sigma} < 1$ y sea $a \in [0, 1]$ como en el enunciado. Puesto que $\frac{1-a}{2} \geq 0$ se sigue que $(\frac{\alpha}{\sigma})^{\frac{1-a}{2}} < 1$. Luego, al igual que en el caso anterior, para todo $\sigma \in (0, D]$ se satisface que

$$\alpha^{\frac{1-a}{2}} |\sigma^{\frac{1+a}{2}} g_\alpha(\sigma)| = \left(\frac{\alpha}{\sigma}\right)^{\frac{1-a}{2}} |\sigma g_\alpha(\sigma)| \leq |\sigma g_\alpha(\sigma)| \leq C_1,$$

donde la última desigualdad se sigue del hecho que $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ es una familia espectral y por definición satisface $|\sigma g_\alpha(\sigma)| \leq C$. Nuevamente, multiplicando por $\alpha^{\frac{a-1}{2}}$ y tomando supremo se obtiene la estimación (3.26). ■

En relación a la condición (CE2) notar que si la familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ satisface (3.25) entonces satisface

$$\sup_{0 < \sigma \leq \infty} |1 - \sigma g_\alpha(\sigma)| \sigma^\mu = C_2 \alpha^\mu, \quad (3.27)$$

para todo $\mu \in [0, 1]$, lo cual se sigue inmediatamente de la Observación 2.25.

A continuación, probaremos un resultado análogo al obtenido en el Lema 3.10 para operadores de regularización asociados a familias espectrales que satisfacen la condición (CE1).

Lema 3.18. *Sean $\eta \in \mathbb{N}$ y $t \in \mathbb{R}^+$ tales que $0 < \eta < t$. Sean, además, $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador lineal, acotado, compacto que satisface con la propiedad de t -suavizado, $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ una familia espectral que satisface (CE1) y $R_\alpha = g_\alpha(T^*T)T^*$ el correspondiente operador de regularización asociado. Entonces existe una constante $C > 0$ tal que se verifica*

$$\|R_\alpha\|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))} \leq C \alpha^{\frac{\eta-t}{2t}}.$$

Demostración. Sea $\{(\sigma_n; u_n; v_n)\}$ el sistema singular asociado al operador T . Como vimos en el Capítulo 2 el operador de regularización asociado a $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ se define por

$$R_\alpha y = \sum_{\sigma_n > 0} \sigma_n g_\alpha(\sigma_n^2) \langle y, v_n \rangle u_n, \quad y \in L^2(\Omega).$$

Puesto que $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ satisface (3.24) entonces, en virtud del Lema 3.17 se tiene que también satisface (3.26). Luego, para todo $y \in H^\eta(\Omega)$ se verifica que

$$\begin{aligned} \|R_\alpha y\|_{L_2(\Omega)}^2 &= \sum_{\sigma_n > 0} (\sigma_n g_\alpha(\sigma_n^2))^2 |\langle y, v_n \rangle|^2 \sigma_n^{\frac{2\eta}{t}} \sigma_n^{\frac{-2\eta}{t}} \\ &\leq \sup_{0 < \sigma \leq \|T\|^2} \left(g_\alpha(\sigma^2) \sigma^{\frac{\eta+t}{t}} \right)^2 \sum_{\sigma_n > 0} \sigma_n^{\frac{-2\eta}{t}} |\langle y, v_n \rangle|^2 \\ &\leq C \alpha^{\frac{\eta-t}{t}} \|y\|_{H^\eta(\Omega)}^2. \end{aligned} \quad (\text{por (3.26)})$$

Se sigue entonces que $\|R_\alpha\|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))} \leq C \alpha^{\frac{\eta-t}{2t}}$ donde C es la constante en (3.24). \blacksquare

En el siguiente lema probamos una cota para el error de regularización que se obtiene utilizando un operador R_α asociado a una familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ que satisface la condición (CE2). Notar que este resultado es análogo al Lema 3.11 donde se lo probó para el caso del operador de regularización asociado al método de Landweber.

Lema 3.19. *Sean $t \geq 0$ y $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ una familia espectral que satisface (CE1) y (CE2). Sean, además, $T : L^2(\Omega) \rightarrow L^2(\Omega)$ y $R_\alpha = g_\alpha(T^*T)T^*$ como en el Lema 3.18. Entonces para todo $s \in [0, t]$ existe una constante K , que depende de s , tal que*

$$\|R_\alpha T - I\|_{\mathcal{L}(H^s(\Omega); L_2(\Omega))} \leq K \alpha^{\frac{s}{2t}}. \quad (3.28)$$

Demostración. Sean $f \in H^s(\Omega)$, $g \doteq Tf$ y $\{(\sigma_n; u_n; v_n)\}$ el sistema singular asociado al operador T . Entonces por definición de R_α se tiene que

$$R_\alpha g = \sum_{\sigma_n > 0} \sigma_n g_\alpha(\sigma_n^2) \langle g, v_n \rangle u_n = \sum_{\sigma_n > 0} \sigma_n g_\alpha(\sigma_n^2) \langle Tf, v_n \rangle u_n = \sum_{\sigma_n > 0} \sigma_n^2 g_\alpha(\sigma_n^2) \langle f, u_n \rangle u_n,$$

y por lo tanto, $R_\alpha g - f = \sum_{\sigma_n > 0} \left(1 - \sigma_n^2 g_\alpha(\sigma_n^2)\right) \langle f, u_n \rangle u_n$.

Por otro lado, puesto que $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ satisface (3.25) sabemos de la Observación 2.25 que se verifica (3.27). Luego, para todo $f \in H^s(\Omega)$ se sigue que

$$\begin{aligned} \|R_\alpha g - f\|_{L^2(\Omega)}^2 &= \sum_{\sigma_n > 0} |1 - \sigma_n^2 g_\alpha(\sigma_n^2)|^2 |\langle f, u_n \rangle|^2 \sigma_n^{\frac{2s}{t}} \sigma_n^{-\frac{2s}{t}} \\ &\leq \sup_{0 < \sigma \leq \|T\|^2} \left(|1 - \sigma^2 g_\alpha(\sigma^2)| \sigma^{\frac{s}{t}} \right)^2 \|f\|_{H^s(\Omega)}^2 \\ &\leq K \alpha^{\frac{s}{t}} \|f\|_{H^s(\Omega)}^2, \end{aligned} \quad (\text{por (3.27)})$$

de donde se sigue la acotación (3.28). ■

En lo que sigue utilizaremos el mismo razonamiento que el utilizado en la sección anterior, considerando la desigualdad que nos permitió acotar el error de regularización del método wavelet-espectral, es decir:

$$\|R_\alpha S_\gamma P_j g^\delta - f\|_{L^2(\Omega)} \leq \|R_\alpha(S_\gamma P_j g^\delta - P_j g)\|_{L^2(\Omega)} + \|R_\alpha(P_j g - g)\|_{L^2(\Omega)} + \|R_\alpha g - f\|_{L^2(\Omega)}.$$

En el Lema 3.19 obtuvimos una cota para el error de estimación del tercer término. Resta probar entonces el correlato de los Lemas 3.14 y 3.15 para una familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ que satisface las condiciones (CE1) y (CE2).

Lema 3.20. Sean $s \geq 0$ y $t > 0$ tales que $s + t \in \mathbb{N}$. Sean además, $g \in H^{s+t}(\Omega)$ y $\eta \in \mathbb{N}$ tal que $\eta < t$ y $\phi \in H^\eta(\Omega)$. Entonces, para R_α el operador correspondiente a la familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ que satisface (3.24), existe una constante C independiente de α y de j tal que

$$\|R_\alpha(P_j g - g)\|_{L^2(\Omega)}^2 \leq C \alpha^{\frac{\eta-t}{t}} 2^{-2j(s+t-\eta)}.$$

Demostración. El resultado de sigue de la aplicación del Lema 3.18 y el Teorema 3.5. En efecto

$$\begin{aligned} \|R_\alpha(P_j g - g)\|_{L^2(\Omega)}^2 &\leq \|R_\alpha\|_{\mathcal{L}(H^\eta(\Omega); L^2(\Omega))}^2 \|P_j g - g\|_{L^2(\Omega)}^2 \\ &\leq \|R_\alpha\|_{\mathcal{L}(H^\eta(\Omega); L^2(\Omega))}^2 \|P_j g - g\|_{H^\eta(\Omega)}^2 \\ &\leq C_1 \alpha^{\frac{\eta-t}{t}} \|P_j g - g\|_{H^\eta(\Omega)}^2 \end{aligned} \quad (\text{por Lema 3.18})$$

$$\begin{aligned} &\leq C_2 \alpha^{\frac{\eta-t}{t}} 2^{-2j(s+t-\eta)} \|g\|_{H^{s+t}(\Omega)}^2, && \text{(por Teorema 3.5)} \\ &\leq C \alpha^{\frac{\eta-t}{t}} 2^{-2j(s+t-\eta)}, \end{aligned}$$

donde C es independiente de α y de j . ■

Lema 3.21. Sean $\eta \in \mathbb{N}_0$, $s \geq 0$ y $t > 0$ tales que $\eta \leq t$ y $s+t \in \mathbb{N}$. Sean además, el operador T como en los lemas anteriores y R_α el operador correspondiente a una familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ que satisface la condición (CE2) dada por (3.25). Supongamos que $g \in L^2(\Omega)$, $\phi \in H^\eta(\Omega)$, $g^\delta \doteq g + \delta dW$ y denotemos con P_j al correspondiente operador proyección sobre V_j y S_γ al operador de umbralado wavelet. Entonces, para toda elección de $\gamma \doteq K\delta\sqrt{|\log \delta|}$ (K una constante positiva arbitraria) y $j \leq j_0 \doteq \lfloor -\frac{\log(\delta^2|\log \delta|)}{\log 2} \rfloor$ existe una constante C , independiente de α y de δ , tal que se tiene la siguiente cota para el error de umbralado wavelet:

$$E(\|R_\alpha(S_\gamma P_j g^\delta - P_j g)\|_{L_2(\Omega)}^2) \leq C \alpha^{\frac{\eta-t}{t}} (\delta\sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}.$$

Demostración. El resultado es consecuencia del Lema 3.18 y del Teorema 3.6. En efecto, notar que

$$\begin{aligned} E(\|R_\alpha(S_\gamma P_j g^\delta - P_j g)\|_{L_2(\Omega)}^2) &\leq \|R_\alpha\|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))}^2 E(\|S_\gamma P_j g^\delta - P_j g\|_{L_2(\Omega)}^2) \\ &\leq \|R_\alpha\|_{\mathcal{L}(H^\eta(\Omega); L_2(\Omega))}^2 E(\|S_\gamma P_j g^\delta - P_j g\|_{H^\eta(\Omega)}^2) \\ &\leq C_1 \alpha^{\frac{\eta-t}{t}} E(\|S_\gamma P_j g^\delta - P_j g\|_{H^\eta(\Omega)}^2) && \text{(por Lema 3.18)} \\ &\leq C \alpha^{\frac{\eta-t}{t}} (\delta\sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}, && \text{(por Teorema 3.6 con } \theta = s+t) \end{aligned}$$

con C una constante independiente de δ y α . ■

Finalmente, estamos en condiciones de demostrar el teorema principal del capítulo. Este resultado nos provee condiciones suficientes sobre una familia espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$, para que el correspondiente método de dos pasos wavelet-espectral asociado a esa familia posea un orden de convergencia casi-óptimo.

Teorema 3.22 (Orden de convergencia para el error total del método wavelet-espectral). Sean $\eta \in \mathbb{N}_0$, $s \geq 0$ y $t > 0$ tales que $\eta \leq t$ y $s+t \in \mathbb{N}$. Supongamos que $T : L^2(\Omega) \rightarrow L^2(\Omega)$ es un operador lineal, acotado, compacto y con la propiedad de t -suavizado, $f \in H^s(\Omega)$, $g = Tf$, $\delta > 0$ y $g^\delta \doteq g + \delta dW$. Sean además

- i) $\phi \in H^\eta(\Omega)$ la función de escala correspondiente a una wavelet con grado de reproducción polinomial m en V_j , tal que $m+1 > t$,
- ii) S_γ el operador de umbralado wavelet con $\gamma \doteq C\delta\sqrt{|\log \delta|}$ y

iii) R_α el operador asociado a la familia de regularización espectral $\{g_\alpha\}_{\alpha \in (0, \alpha_0)}$ que satisface las condiciones (CE1) y (CE2) dadas por (3.24) y (3.25).

Si el nivel de resolución wavelet j satisface $2^{-j} \leq (\delta \sqrt{|\log \delta|})^{1/(\eta+d/2)}$ y

$$\begin{aligned} \alpha &\simeq (\delta \sqrt{|\log \delta|})^{\frac{4t}{2s+2t+d}} \quad \text{si } \eta < t, \text{ y} \\ \alpha &\simeq (\delta \sqrt{|\log \delta|})^2 \quad \text{si } \eta = t, \end{aligned}$$

entonces el error total de regularización del método wavelet-espectral es casi óptimo en el sentido que satisface

$$E(\| R_\alpha S_\gamma P_j g^\delta - f \|_{L^2(\Omega)}^2) = O\left((\delta \sqrt{|\log \delta|})^{\frac{4s}{2s+2t+d}} \right).$$

Demostración. Observemos en primer lugar que, siguiendo acotaciones análogas a las utilizadas en la demostración del Teorema 3.16, se tiene que

$$\begin{aligned} E(\| R_\alpha S_\gamma P_j g^\delta - f \|_{L^2(\Omega)}^2) &\leq 3E(\| R_\alpha(S_\gamma P_j g^\delta - P_j g) \|_{L^2(\Omega)}^2) \\ &\quad + 3\| R_\alpha(P_j g - g) \|_{L^2(\Omega)}^2 + 3\| R_\alpha g - f \|_{L^2(\Omega)}^2 \\ &= 3(EU^2 + EP^2 + ER^2), \end{aligned}$$

donde, como antes, EU, EP y ER denotan los errores de umbralado, proyección y regularización, respectivamente.

Caso I: Consideremos en primer lugar el caso $\eta < t$. Una cota para el primer término correspondiente a EU^2 se sigue del Lema (3.21). En efecto

$$EU^2 = E(\| R_\alpha(S_\gamma P_j g^\delta - P_j g) \|_{L^2(\Omega)}^2) \leq C_1 \alpha^{\frac{\eta-t}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}, \quad (3.29)$$

para todo $j \leq j_0 \doteq \lfloor -\frac{\log(\delta^2 |\log \delta|)}{\log 2} \rfloor$.

Por otro lado, una cota para el error de proyección EP^2 se obtiene del Lema 3.20:

$$EP^2 = \| R_\alpha(P_j g - g) \|_{L^2(\Omega)}^2 \leq C_2 \alpha^{\frac{\eta-t}{t}} 2^{-2j(s+t-\eta)}. \quad (3.30)$$

Igualando los órdenes obtenidos para EU^2 y EP^2 en (3.29) y (3.30), respectivamente, podemos estimar el nivel de proyección wavelet. En efecto, de

$$\alpha^{\frac{\eta-t}{t}} 2^{-2j(s+t-\eta)} = \alpha^{\frac{\eta-t}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}},$$

se sigue que $2^{-j} \approx (\delta \sqrt{|\log \delta|})^{2/(2s+2t+d)}$ (consideramos una aproximación pues $j \in \mathbb{N}$).

Suponiendo $\delta \sqrt{|\log \delta|} \leq 1$ se observa que la elección propuesta del nivel de resolución j es menor a la estimación obtenida

$$2^{-j} \leq (\delta \sqrt{|\log \delta|})^{\frac{1}{\eta+d/2}} < (\delta \sqrt{|\log \delta|})^{\frac{2}{2s+2t+d}}.$$

Para una elección de parámetros en la que el nivel de resolución sea independiente del nivel de suavidad de la solución y de la propiedad de t -suavizado del operador, como en el Teorema 3.16, proponemos utilizar $2^{-j} \approx (\delta \sqrt{|\log \delta|})^{\frac{1}{\eta+d/2}}$. Reemplazando el nivel de resolución wavelet estimado en la cota para EP^2 obtenemos

$$\begin{aligned} EU^2 + EP^2 &= E(\| R_\alpha(S_\gamma P_j g^\delta - P_j g) \|_{L_2(\Omega)}^2) + \| R_\alpha(P_j g - g) \|_{L_2(\Omega)}^2 \\ &\leq (C_1 + C_2) \alpha^{\frac{\eta-t}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}}. \end{aligned}$$

Por último, resta estimar una cota para el error de regularización. Del Lema (3.19) se tiene directamente

$$ER^2 = \| R_\alpha g - f \|_{L_2(\Omega)} \leq C_3 \alpha^{\frac{s}{t}}.$$

Igualando ahora los órdenes de convergencia correspondientes a los errores EU^2 y ER^2 , es posible estimar el parámetro de regularización α . En efecto, de

$$\alpha^{\frac{\eta-t}{t}} (\delta \sqrt{|\log \delta|})^{\frac{4(s+t-\eta)}{2s+2t+d}} = \alpha^{\frac{s}{t}},$$

se obtiene la estimación $\alpha = (\delta \sqrt{|\log \delta|})^{\frac{4t}{2s+2t+d}}$.

Finalmente, utilizando las estimaciones obtenidas para los parámetros j y α se sigue la acotación del error total correspondiente al método de dos pasos:

$$E(\| R_\alpha S_\gamma P_j g^\delta - f \|)^2 \leq 3(EU^2 + EP^2 + ER^2) \leq 3(C_1 + C_2 + C_3) (\delta \sqrt{|\log \delta|})^{\frac{4s}{2s+2t+d}}.$$

Caso II: Consideremos ahora el caso $\eta = t$. Notemos en primer lugar que

$$\begin{aligned} \| R_\alpha y \|_{L_2(\Omega)}^2 &= \sum_{\sigma_n > 0} \sigma_n^2 |g_\alpha(\sigma_n^2)|^2 |\langle y, v_n \rangle|^2 \sigma_n^{-2} \sigma_n^2 \\ &\leq \sup_{0 < \sigma \leq \|T\|^2} \sigma^4 |g_\alpha(\sigma^2)|^2 \sum_{\sigma_n > 0} \sigma_n^{-2} |\langle y, v_n \rangle|^2 \\ &= \sup_{0 < \sigma \leq \|T\|^2} (\sigma^2 |g_\alpha(\sigma^2)|)^2 \|y\|_{H^t(\Omega)}^2 \quad (\text{por (3.3), puesto que } \eta = t) \\ &= C_1 \|y\|_{H^t(\Omega)}^2. \quad (\text{por Lema 3.17 con } a = 1) \end{aligned}$$

Se sigue entonces que

$$\| R_\alpha \|_{\mathcal{L}(H^t(\Omega); L_2(\Omega))} \leq C_1, \quad (3.31)$$

donde C_1 es una constante independiente de α .

Luego,

$$\begin{aligned} E(\| R_\alpha(S_\gamma P_j g^\delta - P_j g) \|_{L_2(\Omega)}^2) &\leq \| R_\alpha \|_{\mathcal{L}(H^t(\Omega); L_2(\Omega))}^2 E(\| S_\gamma P_j g^\delta - P_j g \|_{L_2(\Omega)}^2) \\ &\leq C_1^2 E(\| S_\gamma P_j g^\delta - P_j g \|_{L_2(\Omega)}^2) \quad (\text{por (3.31)}) \end{aligned}$$

$$\leq C(\delta\sqrt{|\log \delta|})^{\frac{4s}{2s+2t+d}},$$

donde la última desigualdad se sigue del Teorema 3.6 con $\theta = s + t$.

Esto concluye la demostración del teorema. ■

CAPÍTULO 4

Regularización wavelet-vaguelet

En este capítulo presentamos un método de regularización wavelet-vaguelet para resolver problemas inversos lineales mal condicionados. Comenzamos con una sección introductoria donde describimos la motivación que origina el surgimiento del método y presentamos además un breve resumen de los elementos que intervienen en su construcción. En las Secciones 4.2 y 4.3 definimos dos tipos de operadores especiales, operadores débilmente invertibles y operadores homogéneos y probamos algunos resultados importantes relacionados a tales operadores, los que resultarán de interés en demostraciones posteriores relacionadas con la convergencia del método propuesto. En la sección siguiente presentamos el método de regularización wavelet-vaguelet. Para ello definimos previamente el concepto de vaguelet y probamos una serie de propiedades. En la Sección 4.5 realizamos un análisis detallado de la convergencia del error correspondiente al método propuesto considerando dos versiones del mismo: con y sin umbralado wavelet. Por último, en la Sección 4.6 presentamos y analizamos una aplicación concreta del método para el caso de operadores integrales que poseen descomposición wavelet vaguelet.

4.1. Introducción

En esta sección repasamos brevemente los métodos de representación de funciones analizados en los Capítulos 1 y 3 (a saber SVD y wavelets), resaltamos sus limitaciones y a partir de ellas fundamentamos la necesidad de introducir una nueva descomposición que resulte más eficiente a la hora de resolver problemas inversos mal condicionados.

4.1.1. Descomposición en valores singulares (SVD)

Como vimos en la Sección 1.2 la descomposición en valores singulares es una herramienta adecuada para el estudio de ciertas propiedades de la inversa generalizada de Moore-Penrose en el caso de operadores compactos. En particular de la Definición 1.6 tenemos que dados X, Y espacios de Hilbert y $T : X \rightarrow Y$ un operador lineal y compacto, un *sistema singular asociado a T* es una familia $\{(\sigma_n; u_n, v_n)\}_{n \in \mathbb{N}}$ donde:

- $\{\sigma_n^2\}_{n \in \mathbb{N}}$ son los autovalores no nulos de T^*T (y de TT^*), considerados en orden decreciente, con su multiplicidad,
- $\{u_n\}_{n \in \mathbb{N}}$ es el correspondiente sistema ortonormal completo de autovectores de T^*T (que expande $\overline{\mathcal{R}(T^*)} = \overline{\mathcal{R}(T^*T)}$) y,
- $\{v_n\}_{n \in \mathbb{N}}$ es el correspondiente sistema ortonormal completo de autovectores de TT^* (que expande $\overline{\mathcal{R}(T)} = \overline{\mathcal{R}(TT^*)}$); se tiene que $v_n \doteq \frac{Tu_n}{\|Tu_n\|}$.

Utilizando este sistema singular asociado a T se puede obtener una representación en serie de la inversa generalizada de Moore-Penrose de dicho operador:

$$T^\dagger y = \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle}{\sigma_n} u_n, \quad \text{con } \mathcal{D}(T^\dagger) = \left\{ y \in Y : \sum_{n=1}^{\infty} \frac{|\langle y, v_n \rangle|^2}{\sigma_n^2} < \infty \right\}. \quad (4.1)$$

Como vimos anteriormente en la Sección 1.2, el principal problema de esta representación, en el caso de problemas mal condicionados, radica en el hecho que los autovalores tienden a cero y por lo tanto es necesario algún tipo de regularización.

Tal como vimos en el Capítulo 2, una regularización de T^\dagger consiste esencialmente en aproximar adecuadamente el operador no acotado T^\dagger por una familia paramétrica $\{R_\alpha\}$ de operadores lineales y continuos. De esta manera, dados $\{R_\alpha\}$ y y^δ con $\|y - y^\delta\|_Y \leq \delta$, una aproximación de la solución de mínimos cuadrados de mínima norma del problema $Tx = y$ está dada por $x_\alpha^\delta \doteq R_\alpha y^\delta$. Algunos métodos de regularización particularmente relacionados a SVD son los métodos de filtro estudiados en la Sección 2.2.

4.1.2. Representación mediante wavelets

En los últimos años se ha despertado mucho interés en el uso de bases ortonormales wavelets para representar funciones y se han descubierto muchas ventajas en la utilización de tales bases, en particular en relación a la representación rala (*sparse*) de ciertas funciones. Así, por ejemplo, existe evidencia empírica de que transformaciones wavelets pueden proveer representaciones ralas robustas de imágenes reales. En efecto, como puede verse en los trabajos de DeVore y Frazier (ver [13] y [19]), en algunos casos de procesamiento de imágenes es posible obtener una reconstrucción razonablemente buena de la imagen con tan solo un muy pequeño porcentaje de los coeficientes en la transformada wavelet discreta (1 o 2%).

Como vimos en la Sección 1.4 dada una wavelet $\psi \in L^2(\mathbb{R}^d)$ y su función de escala asociada $\phi \in L^2(\mathbb{R}^d)$ definimos para todo $\mathbf{x} \in \mathbb{R}^d$ las transformaciones diádicas:

$$\psi_{j,\mathbf{k}}^{(\epsilon)}(\mathbf{x}) \doteq 2^{\frac{j d}{2}} \psi^{(\epsilon)}(2^j \mathbf{x} - \mathbf{k}) \quad \text{y} \quad \phi_{j,\mathbf{k}}(\mathbf{x}) \doteq 2^{\frac{j d}{2}} \phi(2^j \mathbf{x} - \mathbf{k}) \quad \forall \mathbf{k} \in \mathbb{Z}^d, \forall \epsilon \in \Gamma \doteq \{0, 1\}^d - \{\mathbf{0}\},$$

y los respectivos espacios generados para un nivel de resolución j :

$$W_j \doteq \text{span}\{\psi_{j,\mathbf{k}}^{(\epsilon)} : \mathbf{k} \in \mathbb{Z}^d \wedge \epsilon \in \Gamma\} \text{ y } V_j \doteq \text{span}\{\phi_{j,\mathbf{k}} : \mathbf{k} \in \mathbb{Z}^d\}.$$

Como vimos anteriormente (Sección 1.4), para todo $j_0 \in \mathbb{Z}$ fijo se tiene que $L^2(\mathbb{R}) = V_{j_0} \bigoplus_{j \geq j_0} W_j$ y para toda $f \in L^2(\mathbb{R}^d)$ se tiene la siguiente representación multiescala

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \phi_{j_0,\mathbf{k}} \rangle \phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{j \geq j_0} \sum_{\epsilon \in \Gamma} \sum_{\mathbf{k} \in \mathbb{Z}^d} \langle f, \psi_{j,\mathbf{k}}^{(\epsilon)} \rangle \psi_{j,\mathbf{k}}^{(\epsilon)}(\mathbf{x}).$$

Definiendo $\Gamma^0 \doteq \{0, 1\}^d$, $\psi_{j,\mathbf{k}}^0 \doteq \phi_{j_0,\mathbf{k}}$, $\lambda = (j, k, \epsilon)$ y $\Lambda \doteq \{(j, k, \epsilon) \in \mathbb{Z} \times \mathbb{Z}^d \times \Gamma^0 : j \geq j_0\}$ se tiene la representación wavelet en su versión simplificada que ya hemos utilizado en los Capítulos 1 y 3

$$f(x) = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \psi_\lambda(x).$$

Recordar que Λ depende de j_0 pero, dado que j_0 será prefijado en un comienzo, a los efectos de simplificar la notación obviaremos especificar tal dependencia.

4.1.3. Descomposición wavelet-vaguelet (DWV)

Hasta aquí hemos visto dos tipos de bases ortogonales para la representación de funciones en el caso de problemas inversos mal condicionados. El primer tipo, SVD, representa eficientemente al operador del problema $Tx = y$. El segundo tipo, las bases wavelets ortonormales, se utilizan para representar de manera eficiente ciertos espacios de funciones que poseen algún tipo de regularidad, y por tal motivo se las utiliza para codificar adecuadamente información previa disponible sobre la solución exacta del problema. A la hora de resolver problemas inversos, sería muy deseable disponer de una descomposición que permita incorporar simultáneamente tanto las ventajas de SVD como de la descomposición wavelet, con el objeto de disponer de una representación que incorpore de manera eficiente tanto información relevante sobre el operador asociado al problema como así también sobre los supuestos de regularidad de la solución exacta del mismo. Desafortunadamente esto es, en muchos casos, imposible, aunque existe un caso particular de problemas inversos donde pueden representarse de manera adecuada, en algún sentido, tanto el operador y como los espacios de funciones asociados a las condiciones fuente. No obstante los detalles sobre este tipo de problemas y la correspondiente descomposición se analizarán en profundidad en las Sección 4.4.2, presentamos brevemente a continuación un bosquejo de los elementos subyacentes en la idea que da origen a este novedoso enfoque.

Sean X, Y espacios de Hilbert y $T : X \rightarrow Y$ un operador lineal. La descomposición wavelet-vaguelet utiliza tres conjuntos de funciones $\{\psi_\lambda\}_{\lambda \in \Lambda}$, $\{u_\lambda\}_{\lambda \in \Lambda}$ y $\{v_\lambda\}_{\lambda \in \Lambda}$ tales que:

- a) $\{\psi_\lambda\}_{\lambda \in \Lambda}$ una base ortonormal de X (ver Sección 1.4),
 b) $\{u_\lambda\}_{\lambda \in \Lambda}$ y $\{v_\lambda\}_{\lambda \in \Lambda}$ son conjuntos casi-ortogonales en Y , es decir, para toda sucesión $\{a_\lambda\}_{\lambda \in \Lambda} \in \ell^2$ se verifica que

$$\left\| \sum_{\lambda \in \Lambda} a_\lambda u_\lambda \right\|_Y \equiv \|a_\lambda\|_{\ell^2(\Lambda)} \text{ y } \left\| \sum_{\lambda \in \Lambda} a_\lambda v_\lambda \right\|_Y \equiv \|a_\lambda\|_{\ell^2(\Lambda)},$$

donde $a \equiv b$ significa que existen constantes $C_1, C_2 > 0$ tales que $C_1 b \leq a \leq C_2 b$,

- c) existen valores casi-singulares k_j que dependen solo del nivel de resolución wavelet j y satisfacen

$$T\psi_\lambda = k_j v_\lambda \quad \text{y} \quad T^* u_\lambda = k_j \psi_\lambda \quad \forall \lambda \in \Lambda, \quad \text{y}$$

- d) $\{u_\lambda\}_{\lambda \in \Lambda}$ y $\{v_\lambda\}_{\lambda \in \Lambda}$ verifican la relación de bi-ortonormalidad

$$\langle u_\lambda, v_\mu \rangle = \delta_{\lambda, \mu} \quad \forall \lambda, \mu \in \Lambda.$$

En el contexto de la aplicación de la descomposición wavelet-vaguelet a la construcción de métodos de regularización para problemas inversos mal condicionados será de particular importancia el hecho que, bajo ciertas hipótesis generales sobre el operador T y la función f , ésta última puede representarse en la forma

$$f = \sum_{\lambda \in \Lambda} k_j^{-1} \langle Tf, u_\lambda \rangle \psi_\lambda.$$

4.2. Operadores lineales débilmente invertibles

En esta sección definimos el concepto de operadores débilmente invertibles y probamos algunos resultados que necesitaremos más adelante al estudiar condiciones suficientes para que un operador admita descomposición wavelet-vaguelet.

Definición 4.1. Sean X, Y espacios de Hilbert, $T : \mathcal{D}(T) \subset X \rightarrow Y$ un operador lineal y sea $\{w_\ell\}_{\ell \in \mathbb{N}}$ una base ortonormal en X . Decimos que T es débilmente invertible con respecto a la base $\{w_\ell\}_{\ell \in \mathbb{N}}$ si para todo $\ell \in \mathbb{N}$ existe $c_\ell \in Y^*$ tal que

$$c_\ell(Tf) = \langle f, w_\ell \rangle \quad \forall f \in \mathcal{D}(T).$$

Observación 4.2. Si un operador T definido entre espacios de Hilbert es débilmente invertible con respecto a una base ortonormal $\{w_\ell\}_{\ell \in \mathbb{N}}$ entonces es inyectivo. En efecto, supongamos que existen $f_1, f_2 \in \mathcal{D}(T)$, tales que $Tf_1 = Tf_2$. Como T es débilmente invertible con respecto a la base $\{w_\ell\}_{\ell \in \mathbb{N}}$ entonces para todo $\ell \in \mathbb{N}$ existe $c_\ell \in Y^*$ tal que $c_\ell(Tf_1) = \langle f_1, w_\ell \rangle$. Luego

$$f_1 = \sum_{\ell \in \mathbb{N}} \langle f_1, w_\ell \rangle w_\ell = \sum_{\ell \in \mathbb{N}} c_\ell(Tf_1) w_\ell = \sum_{\ell \in \mathbb{N}} c_\ell(Tf_2) w_\ell = \sum_{\ell \in \mathbb{N}} \langle f_2, w_\ell \rangle w_\ell = f_2.$$

En el contexto del problema inverso asociado a la ecuación $Tf = g$, si T es débilmente invertible con respecto a una cierta base, es posible encontrar una representación de la función f en términos del dato g y de esa base, que resulta apropiada para la implementación de métodos de regularización. En efecto, dado un operador T débilmente invertible con respecto a una base ortonormal $\{w_\ell\}_{\ell \in \mathbb{N}}$, a partir de los funcionales $\{c_\ell\}_{\ell \in \mathbb{N}}$ se tiene que si $Tf = g$ entonces

$$f = \sum_{\ell \in \mathbb{N}} \langle f, w_\ell \rangle w_\ell = \sum_{\ell \in \mathbb{N}} c_\ell(Tf) w_\ell = \sum_{\ell \in \mathbb{N}} c_\ell(g) w_\ell \doteq T^{-1}g. \quad (4.2)$$

Lema 4.3. *Sea $T : X \rightarrow Y$ un operador lineal con X, Y Hilbert y sea $\{w_\ell\}_{\ell \in \mathbb{N}}$ una base ortonormal de X . Si $\mathcal{D}(T)$ es denso en X , entonces T es débilmente invertible con respecto a $\{w_\ell\}_{\ell \in \mathbb{N}}$ si y solo si $\{w_\ell\}_{\ell \in \mathbb{N}} \subset \mathcal{R}(T^*)$.*

Demostración. Por un lado, observar que si T es débilmente invertible con respecto a la base $\{w_\ell\}_{\ell \in \mathbb{N}}$, se sigue inmediatamente de la definición de operador adjunto que los funcionales c_ℓ de la Definición 4.1 satisfacen $c_\ell \in \mathcal{D}(T^*) \forall \ell \in \mathbb{N}$ y, más aún, $T^*c_\ell = w_\ell \forall \ell \in \mathbb{N}$.

Recíprocamente, si $\{w_\ell\}_{\ell \in \mathbb{N}} \subset \mathcal{R}(T^*)$ entonces existe $\{c_\ell\}_{\ell \in \mathbb{N}} \subset \mathcal{D}(T^*)$ tal que $T^*c_\ell = w_\ell \forall \ell \in \mathbb{N}$ y por lo tanto se tiene que $\forall f \in \mathcal{D}(T)$

$$\langle f, w_\ell \rangle = \langle f, T^*c_\ell \rangle = \langle Tf, c_\ell \rangle = c_\ell(Tf) \quad \forall \ell \in \mathbb{N}.$$

Luego, T es débilmente invertible con respecto a la base $\{w_\ell\}_{\ell \in \mathbb{N}}$. ■

Observación 4.4. Sea T un operador lineal, acotado, inyectivo, con $\mathcal{D}(T) = X$. Si $\mathcal{R}(T^*)$ es cerrado entonces el operador T es débilmente invertible con respecto a cualquier base ortonormal. Este resultado se sigue directamente del lema anterior considerando que al ser T un operador acotado e inyectivo, $\mathcal{R}(T^*)$ es denso en X . En efecto, dado que T es acotado tenemos que $\mathcal{D}(T^*) = Y$ y por lo tanto existe T^{**} y $T^{**} = T$. Luego, $\mathcal{N}(T) = \mathcal{N}(T^{**}) = \mathcal{R}(T^*)^\perp$ y por lo tanto $\{0\}^\perp = \mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$.

El siguiente lema permite deducir la invertibilidad débil de un operador T a partir de la análoga propiedad del operador autoadjunto T^*T .

Lema 4.5. *Sea $T : X \rightarrow Y$ un operador lineal y acotado, con $\mathcal{D}(T) = X$ y sea además $\{w_\ell\}_{\ell \in \mathbb{N}}$ una base ortonormal de X . Si T^*T es débilmente invertible con respecto a $\{w_\ell\}_{\ell \in \mathbb{N}}$ entonces T también lo es con respecto a la misma base.*

Demostración. Como el operador T es acotado, tenemos que $\mathcal{D}(T^*T) = \mathcal{D}(T) = X$. Por otro lado, dado que T^*T es débilmente invertible, para todo $\ell \in \mathbb{N}$ existe $s_\ell \in X^*$ tal que

$s_\ell(T^*Tf) = \langle f, w_\ell \rangle$ para toda $f \in \mathcal{D}(T^*T)$. Luego, en virtud del Teorema de Riesz, $\forall \ell \in \mathbb{N}$ existe $s_\ell \in X$ tal que

$$\langle f, w_\ell \rangle = s_\ell(T^*Tf) = \langle T^*Tf, s_\ell \rangle = \langle Tf, Ts_\ell \rangle.$$

Definiendo $c_\ell \doteq Ts_\ell$ tenemos entonces que $c_\ell(Tf) = \langle f, w_\ell \rangle$ para todo $\ell \in \mathbb{N}$, para toda $f \in X$ y, por lo tanto, T es débilmente invertible con respecto a la base $\{w_\ell\}_{\ell \in \mathbb{N}}$. ■

Presentamos a continuación un ejemplo de operador integral débilmente invertible. Otros ejemplos de este tipo serán desarrollados con mayor detalle en la Sección 4.6 para núcleos de cuadrado integrable o débilmente singulares.

Ejemplo 4.6. Consideremos el operador integral $T : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ definido por

$$Tf(t) = \int_{-\infty}^t f(u)du, \quad (4.3)$$

con $\mathcal{D}(T) = \{f \in L^2(\mathbb{R}) : \frac{\hat{f}(\omega)}{|\omega|} \in L^2(\mathbb{R})\}$. Se puede probar fácilmente que $\mathcal{R}(T) = \{f \in L^2(\mathbb{R}) : |\omega|\hat{f}(\omega) \in L^2(\mathbb{R})\}$ y $T^* = -T$.

En el análisis de este ejemplo nos proponemos utilizar el andamiaje metodológico desarrollado hasta el momento para obtener una representación adecuada de la inversa del operador T en términos del dato g , utilizando los funcionales lineales asociados a la condición de invertibilidad débil del operador. A tales efectos consideremos una base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ asociada a una madre wavelet $\psi \in \mathcal{R}(T^*)$. Se tiene entonces que $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{R}(T^*)$ y, por lo tanto, $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{R}(T)$, puesto que $T^* = -T$. Así, del Lema 4.3 se sigue que el operador T definido por (4.3) es débilmente invertible con respecto a la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$. Y puesto que $\mathcal{D}(T)$ es denso en $L^2(\mathbb{R})$ los funcionales $\{c_\lambda\}_{\lambda \in \Lambda}$ de la Definición 4.1 satisfacen $T^*c_\lambda = \psi_\lambda$. Se sigue entonces que $c_\lambda = -T^{-1}\psi_\lambda = -\psi'_\lambda$. De esta forma, dada la representación wavelet diádica $\psi_\lambda(t) = \psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k) \forall \lambda \in \Lambda$ se puede obtener una expresión formal para la familia $\{c_\lambda\}_{\lambda \in \Lambda}$. A saber,

$$c_\lambda(t) = c_{j,k}(t) = -\psi'_{j,k}(t) = -2^{\frac{3j}{2}}\psi'(2^j t - k) \text{ y } \|c_\lambda\|_{L^2(\mathbb{R})} = 2^j \|\psi'\|_{L^2(\mathbb{R})} \forall \lambda \in \Lambda.$$

De este modo para toda $f \in \mathcal{D}(T)$ tenemos que

$$f = \sum_{\lambda \in \Lambda} \langle f, \psi_\lambda \rangle \psi_\lambda = \sum_{\lambda \in \Lambda} c_\lambda(Tf) \psi_\lambda = \sum_{\lambda \in \Lambda} \langle Tf, c_\lambda \rangle \psi_\lambda. \quad (4.4)$$

Esta es la correspondiente representación asociada a (4.2), la cual será de suma importancia en el desarrollo del método de regularización que definiremos más adelante. Observar que en este caso particular T no es acotado y tampoco lo es su inversa generalizada. En la Sección 4.6

analizaremos otros ejemplos en un contexto más general de operadores integrales con núcleos de cuadrado integrable y débilmente singulares.

4.3. Operadores homogéneos

En esta sección definimos los conceptos de funciones y operadores homogéneos y probamos algunos resultados que necesitaremos más adelante al analizar condiciones necesarias y suficientes sobre el núcleo de un operador de convolución para que este admita una descomposición wavelet-vaguelet. Estos resultados están relacionados a los correspondientes operadores duales e inversos y a la transformada de Fourier de una función homogénea.

Iniciamos esta sección introduciendo el concepto de operador dilatación y, a partir del mismo, definimos luego el concepto de operador homogéneo.

Definición 4.7. Dados $d \in \mathbb{N}$ y $a > 0$ definimos el *operador dilatación* $D_a : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ como $D_a f(t) \doteq f(at)$.

Definición 4.8. Decimos que un operador $T : \mathcal{D}(T) \subset L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ es *homogéneo* con *grado de homogeneidad* β si para todo $a > 0$ se verifica que $D_a \mathcal{D}(T) \subset \mathcal{D}(T)$ (es decir, $D_a f \in \mathcal{D}(T) \forall f \in \mathcal{D}(T)$ y $\forall a > 0$, o sea $\mathcal{D}(T)$ es cerrado por dilataciones) y

$$TD_a = a^{-\beta d} D_a T. \quad (4.5)$$

Los siguientes dos lemas relacionan el grado de homogeneidad de un operador con el de su inverso y el de su adjunto.

Lema 4.9. *Sea $T : \mathcal{D}(T) \subset L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ un operador inyectivo. Si T tiene grado de homogeneidad β entonces el operador T^{-1} es también homogéneo de grado $-\beta$.*

Demostración. Dado que el operador T es homogéneo de grado β tenemos, por definición, que para todo $a > 0$ se verifica $TD_a = a^{-\beta d} D_a T$. Por otro lado, al ser T inyectivo existe $T^{-1} : \mathcal{R}(T) \subset L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$. Luego, sobre $\mathcal{R}(T)$ se tiene que

$$D_a T^{-1} = T^{-1} T D_a T^{-1} = T^{-1} (a^{-\beta d} D_a T) T^{-1} = a^{-\beta d} T^{-1} D_a$$

y, por lo tanto, $T^{-1} D_a = a^{-(\beta)d} D_a T^{-1}$, lo que nos dice que T^{-1} es homogéneo de grado $-\beta$. ■

Lema 4.10. *Sea $T : \mathcal{D}(T) \subset L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ con $\mathcal{D}(T)$ denso. Si T es homogéneo, entonces el operador adjunto T^* es también homogéneo y tiene el mismo grado de homogeneidad.*

Demostración. Supongamos que T es homogéneo de grado β , entonces $\mathcal{D}(T)$ es cerrado por dilataciones. Probaremos en primer lugar que $\mathcal{D}(T^*)$ es también cerrado por dilataciones. En efecto, observar en primer lugar que $\forall f, g \in L^2(\mathbb{R}^d)$ y $\forall a > 0$ se tiene que

$$\langle D_a f, g \rangle = \langle f, a^{-d} D_{\frac{1}{a}} g \rangle. \quad (4.6)$$

Sean $g \in \mathcal{D}(T^*)$ y $a > 0$. Queremos probar que $D_a g \in \mathcal{D}(T^*)$. Como $g \in \mathcal{D}(T^*)$ entonces existe una constante K_g , con $0 < K_g < \infty$, tal que

$$|\langle T f, g \rangle| \leq K_g \|f\|_{L^2(\mathbb{R}^d)} \quad \forall f \in \mathcal{D}(T). \quad (4.7)$$

Luego,

$$\begin{aligned} |\langle T f, D_a g \rangle| &= a^{-d} |\langle D_{\frac{1}{a}} T f, g \rangle| && \text{(por (4.6))} \\ &= a^{-d(\beta+1)} |\langle T D_{\frac{1}{a}} f, g \rangle| && \text{(pues } T \text{ es homogéneo de grado } \beta) \\ &\leq a^{-d(\beta+1)} K_g \|D_{\frac{1}{a}} f\|_{L^2(\mathbb{R}^d)} && \text{(por (4.7))} \\ &= a^{-d(\beta+\frac{3}{2})} K_g \|f\|_{L^2(\mathbb{R}^d)}. && \left(\text{pues } \|D_b f\|_{L^2(\mathbb{R}^d)} = b^{\frac{d}{2}} \|f\|_{L^2(\mathbb{R}^d)} \right) \end{aligned}$$

Esto implica que el funcional $f \rightarrow \langle T f, D_a g \rangle$ es continuo sobre $\mathcal{D}(T)$ y por lo tanto $D_a g \in \mathcal{D}(T^*)$.

Para probar que T^* es homogéneo de grado β tenemos que verificar la identidad $T^* D_a = a^{-\beta d} D_a T^* \quad \forall a > 0$. Para ello probaremos primero que se satisfacen cada una de las siguientes identidades para todos $f \in \mathcal{D}(T)$, $g \in \mathcal{D}(T^*)$ y $a > 0$:

- i) $\langle T D_a f, g \rangle = \langle f, a^{-d} D_{\frac{1}{a}} T^* g \rangle$,
- ii) $\langle f, a^{-\beta d - d} T^* D_{\frac{1}{a}} g \rangle = \langle f, a^{-d} D_{\frac{1}{a}} T^* g \rangle$ y
- iii) $\langle a^{-\beta d} D_a T f, g \rangle = \langle f, a^{-\beta d - d} T^* D_{\frac{1}{a}} g \rangle$.

La identidad i) se sigue del hecho que para $f \in \mathcal{D}(T)$, $g \in \mathcal{D}(T^*)$ y $a > 0$ se tiene que

$$\begin{aligned} \langle T D_a f, g \rangle &= \langle D_a f, T^* g \rangle && \text{(por definición de } T^*) \\ &= \langle f, a^{-d} D_{\frac{1}{a}} T^* g \rangle. && \text{(por (4.6))} \end{aligned}$$

Asimismo, para ii) observar que para $f \in \mathcal{D}(T)$, $g \in \mathcal{D}(T^*)$ y $a > 0$

$$\begin{aligned} \langle f, a^{-\beta d - d} T^* D_{\frac{1}{a}} g \rangle &= \langle T f, a^{-\beta d - d} D_{\frac{1}{a}} g \rangle && \text{(por definición de } T^*) \\ &= \langle a^{-\beta d} D_a T f, g \rangle && \text{(por (4.6))} \\ &= \langle T D_a f, g \rangle && \text{(pues } T \text{ es homogéneo de grado } \beta) \\ &= \langle D_a f, T^* g \rangle && \text{(por definición de } T^*) \end{aligned}$$

$$= \langle f, a^{-d} D_{\frac{1}{a}} T^* g \rangle. \quad (\text{por (4.6)})$$

La identidad iii) se sigue del hecho que

$$\begin{aligned} \langle a^{-\beta d} D_a T f, g \rangle &= a^{-\beta d} \langle T f, a^{-d} D_{\frac{1}{a}} g \rangle && (\text{por (4.6)}) \\ &= a^{-\beta d} \langle f, a^{-d} T^* D_{\frac{1}{a}} g \rangle && (\text{pues } D_a g \in \mathcal{D}(T^*)) \\ &= \langle f, a^{-\beta d - d} T^* D_{\frac{1}{a}} g \rangle. \end{aligned}$$

Finalmente, de las identidades i), ii) y iii) se sigue que $\langle f, a^{-d} D_{\frac{1}{a}} T^* g \rangle = \langle f, a^{-\beta d - d} T^* D_{\frac{1}{a}} g \rangle$ para todas $f \in \mathcal{D}(T)$, $g \in \mathcal{D}(T^*)$. Puesto que $\mathcal{D}(T)$ es denso en $L^2(\mathbb{R}^d)$ se sigue entonces que $a^{-d} D_{\frac{1}{a}} T^* g = a^{-\beta d - d} T^* D_{\frac{1}{a}} g$ para toda $g \in \mathcal{D}(T^*)$. Con $b = \frac{1}{a}$ se obtiene finalmente que

$$T^* D_b = b^{-\beta d} D_b T^* \quad \forall b > 0.$$

Luego, T^* es homogéneo de grado β . ■

Finalizamos esta sección introduciendo el concepto de función homogénea y probamos un lema sobre la transformada de Fourier de funciones homogéneas.

Definición 4.11. Decimos que $f : \mathbb{R}^d \rightarrow \mathbb{R}$ es *homogénea de grado β* si $\forall a > 0$ se verifica

$$f(ax) = a^{-\beta d} f(x), \quad \forall x \in \mathbb{R}^d.$$

Recordar que para $f \in L^1(\mathbb{R}^d)$ su transformada de Fourier $\hat{f} \in L^1(\mathbb{R}^d)$ se define por

$$\hat{f}(\omega) \doteq \int_{\mathbb{R}^d} e^{-it \cdot \omega} f(t) dt, \quad \omega \in \mathbb{R}^d.$$

Lema 4.12. Sea $f \in L^2(\mathbb{R}^d; \mathbb{R})$ homogénea de grado β . Entonces su transformada de Fourier \hat{f} es también homogénea con grado de homogeneidad $1 - \beta$.

Demostración. Se sigue inmediatamente del hecho que para todo $a > 0$ y $\omega \in \mathbb{R}^d$ se verifica

$$\hat{f}(a\omega) = \int_{\mathbb{R}^d} e^{-it \cdot a\omega} f(t) dt = \int_{\mathbb{R}^d} e^{-i\tau \cdot \omega} f\left(\frac{\tau}{a}\right) \frac{d\tau}{a^d} = \frac{1}{a^d} \frac{1}{a^{-\beta d}} \int_{\mathbb{R}^d} e^{-i\tau \cdot \omega} f(\tau) d\tau = a^{-(1-\beta)d} \hat{f}(\omega). \quad \blacksquare$$

4.4. Regularización wavelet-vaguelet

En esta sección presentaremos un método de regularización wavelet-vaguelet para resolver problemas inversos lineales mal condicionados. Para ello introducimos en primer lugar los elementos fundamentales del problema y esbozamos un bosquejo del método en función de las herramientas y resultados estudiados en las secciones anteriores. Posteriormente, definimos el

concepto de vaguelet y analizamos algunas de sus propiedades fundamentales para finalmente presentar el método de regularización wavelet-vaguelet a través de la descomposición homóloga.

4.4.1. Introducción

En esta sección y las que restan de este capítulo, consideramos nuevamente problemas inversos de la forma

$$Tf = g, \quad (4.8)$$

donde $T : X \rightarrow Y$ es un operador lineal, acotado, de rango no cerrado y X, Y son espacios de Hilbert. En general, supondremos que el dato g no se conoce de manera exacta. Más precisamente, consideraremos un modelo de observación con ruido aditivo de la forma

$$g^\delta = g + \delta dW, \quad (4.9)$$

donde dW es ruido blanco y $\delta > 0$ una constante que cuantifica el nivel de error. En lo que resta de este capítulo, a menos que lo especifiquemos de otro modo, trabajaremos exclusivamente con los espacios $X = Y = L^2(\Omega)$, donde $\Omega \subset \mathbb{R}^d$ es un cierto dominio acotado.

Al igual que en la Sección 1.4 denotamos con ψ y ϕ a la función wavelet y su función de escala asociada, respectivamente, ambas en $L^2(\Omega)$. Sean además j_0, Γ^0 y la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ como en (1.20). Dado $\gamma \doteq \{\gamma_j\}_{j \in \mathbb{Z}}$, con $\gamma_j \geq 0 \forall j \in \mathbb{Z}$ (de ahora en más nos referiremos a una tal sucesión γ como un “umbral”), definimos el operador umbralado wavelet fuerte $S_\gamma : L^2(\Omega) \rightarrow L^2(\Omega)$ asociado al umbral γ y a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$, como sigue. Dada $f \in L^2(\Omega)$, escribimos $f(x) = \sum_{\lambda \in \Lambda} \alpha_\lambda \psi_\lambda(x)$, donde $\alpha_\lambda \doteq \langle f, \psi_\lambda \rangle$ y $\Lambda \doteq \{(j, k, \epsilon) \in \mathbb{Z} \times \mathbb{Z}^d \times \Gamma^0 : j \geq j_0 \wedge \text{sop}(\psi_\lambda) \subset \Omega\}$. El operador umbralado wavelet fuerte S_γ se define entonces como

$$[S_\gamma f(\cdot)](x) \doteq \sum_{\lambda \in \Lambda} \alpha_\lambda \chi_{\{|\alpha_\lambda| \geq \gamma_j\}}(\alpha_\lambda) \psi_\lambda(x). \quad (4.10)$$

El método de regularización que propondremos en este capítulo consiste, esencialmente, en elegir adecuadamente una base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ con respecto a la cual el operador T asociado al problema (4.8) sea débilmente invertible. En tal caso, imponiendo condiciones que garanticen la inyectividad del operador adjunto, se sigue inmediatamente, en virtud del Lema 4.3, que los funcionales c_λ de la Definición 4.1, están en este caso dados por $c_\lambda = (T^*)^{-1} \psi_\lambda \forall \lambda \in \Lambda$. Veremos que si el operador T es, además, homogéneo, la colección de funcionales $\{c_\lambda\}_{\lambda \in \Lambda}$ satisface ciertas propiedades que la transforman en una familia a la que denominaremos de “vaguelets”. En la siguiente sección introduciremos formalmente el concepto de vaguelet y probaremos algunos de los resultados principales asociados a las mismas. Como veremos más adelante (Secciones 4.4.3

y 4.6), el uso de vaguelets en problemas del tipo (4.8), será particularmente apropiado en los casos en que el operador T^* posea un comportamiento más estable que el operador T en el sentido del mal condicionamiento asociado al problema inverso.

Recordemos nuevamente que en el modelo (4.8), suponemos que g es desconocido y solo se dispone de una versión g^δ contaminada por ruido, dada por (4.9). La invertibilidad débil de T con respecto a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$, junto con la representación (4.2) nos lleva inmediatamente a definir nuestra primera aproximación, en términos de g^δ , de la solución exacta del problema (4.8):

$$f_c^\delta(x) \doteq \sum_{\lambda \in \Lambda} c_\lambda(g^\delta) \psi_\lambda(x), \quad x \in \Omega. \quad (4.11)$$

El subíndice “c” en f_c^δ en (4.11) hace referencia a la familia de funcionales $\{c_\lambda\}_{\lambda \in \Lambda}$ asociada a la invertibilidad débil de T con respecto a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$. Una segunda aproximación se obtiene a partir de (4.11) mediante la aplicación del operador umbralado wavelet fuerte S_γ definido en (4.10) (con un umbral γ convenientemente elegido):

$$f_{c,\gamma}^\delta(x) \doteq S_\gamma f_c^\delta(x), \quad x \in \Omega. \quad (4.12)$$

En la siguiente sección introduciremos el concepto de vaguelets y estudiaremos algunas de sus principales propiedades, las que serán de fundamental importancia para el posterior desarrollo de la descomposición wavelet-vaguelet y el correspondiente método de regularización asociado, para el problema (4.8).

4.4.2. Vaguelets

Esta sección tiene como objetivo introducir formalmente el concepto de vaguelet. La incorporación de vaguelets en el contexto de problemas inversos lineales mal condicionados a través de la llamada descomposición wavelet-vaguelet (DWV), que estudiaremos en la Sección 4.4.3, da origen a un método de regularización que en algunos casos resulta ser muy eficiente. Esto es fundamentalmente consecuencia del hecho que como veremos las vaguelets resultan ser una herramienta eficaz para representar de manera adecuada tanto al operador asociado al problema como a ciertos espacios funcionales asociados a condiciones fuente de la solución exacta. Comenzaremos definiendo el concepto de vaguelet y, posteriormente, probaremos algunos resultados que permiten demostrar la propiedad de casi-ortogonalidad de las mismas.

Definición 4.13. Una colección de funciones $\{v_{j,k}\}_{j,k} \subset C(\mathbb{R}^d; \mathbb{R})$ con $j \in \mathbb{Z}$, $k \in \mathbb{Z}^d$ se dice que constituye una familia de *vaguelets* si existen constantes positivas C , α y β , con $\beta < \alpha < 1$,

tales que $\forall j \in \mathbb{Z}$ y $\forall k \in \mathbb{Z}^d$ se verifican las siguientes condiciones:

$$|v_{j,k}(x)| \leq C 2^{\frac{d}{2}j} (1 + |2^j x - k|)^{-d-\alpha}, \quad \forall x \in \mathbb{R}^d, \quad (4.13)$$

$$\int_{\mathbb{R}^d} v_{j,k}(x) dx = 0 \quad \text{y} \quad (4.14)$$

$$|v_{j,k}(x') - v_{j,k}(x)| \leq C 2^{j(\frac{d}{2}+\beta)} |x' - x|^\beta, \quad \forall x, x' \in \mathbb{R}^d. \quad (4.15)$$

Notar que (4.13) implica que cada una de las funciones $v_{j,k}(x)$ concentra su energía en un entorno de $x = 2^{-j}k$, mientras que (4.15) es una condición Hölder de orden β .

Lema 4.14 (Lema de Schur). Sean $M \doteq \{m_{i,j}\}_{i,j \in \mathbb{N}}$ una matriz infinita y $\{w_i\}_{i \in \mathbb{N}}$ una sucesión de números positivos. Si $\forall i \in \mathbb{N}$

$$\sum_{j \in \mathbb{N}} |m_{i,j}| w_j \leq w_i, \quad (4.16)$$

y, simétricamente, para todo $j \in \mathbb{N}$

$$\sum_{i \in \mathbb{N}} |m_{i,j}| w_i \leq w_j, \quad (4.17)$$

entonces $M : \ell^2(\mathbb{N}) \rightarrow \ell^2(\mathbb{N})$ es un operador acotado y su norma es menor o igual que 1.

Demostración. Sean $a \doteq \{a_j\}_{j \in \mathbb{N}} \in \ell^2(\mathbb{N})$ y $b \doteq Ma$, es decir $b_i \doteq \sum_{j \in \mathbb{N}} m_{i,j} a_j$. Escribiendo $|m_{i,j}| |a_j| = |m_{i,j}|^{\frac{1}{2}} w_j^{\frac{1}{2}} w_j^{-\frac{1}{2}} |m_{i,j}|^{\frac{1}{2}} |a_j|$ y aplicando la desigualdad de Cauchy-Schwarz se tiene que

$$\begin{aligned} |b_i|^2 &\leq \left(\sum_{j \in \mathbb{N}} |m_{i,j}| w_j \right) \left(\sum_{j \in \mathbb{N}} |m_{i,j}| w_j^{-1} |a_j|^2 \right) \\ &\leq w_i \left(\sum_{j \in \mathbb{N}} |m_{i,j}| w_j^{-1} |a_j|^2 \right). \end{aligned} \quad (\text{por (4.16)})$$

Luego, sumando para todo $i \in \mathbb{N}$, se sigue que

$$\begin{aligned} \|Ma\|_{\ell^2(\mathbb{N})}^2 &= \sum_{i \in \mathbb{N}} |b_i|^2 \\ &\leq \sum_{i \in \mathbb{N}} w_i \left(\sum_{j \in \mathbb{N}} |m_{i,j}| w_j^{-1} |a_j|^2 \right) \\ &= \sum_{j \in \mathbb{N}} w_j^{-1} |a_j|^2 \sum_{i \in \mathbb{N}} |m_{i,j}| w_i \\ &\leq \sum_{j \in \mathbb{N}} w_j^{-1} |a_j|^2 w_j \quad (\text{por (4.17)}) \\ &= \sum_{j \in \mathbb{N}} |a_j|^2 \\ &= \|a_j\|_{\ell^2(\mathbb{N})}^2. \end{aligned}$$

Así, finalmente obtenemos $\|M\|_{\mathcal{L}(\ell^2(\mathbb{N}))} \leq 1$. ■

El siguiente teorema constituye un resultado importante sobre propiedades de las vaguelets.

Teorema 4.15. *Sea $\{v_{j,k}\}_{j,k}$ una familia de vaguelets con las correspondientes constantes asociadas α , β y C_1 . Dada una sucesión $a \doteq \{a_{j,k}\}_{j,k} \in \ell^2(\mathbb{Z} \times \mathbb{Z}^d)$ se define*

$$[Q(a)](x) \doteq \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}^d} a_{j,k} v_{j,k}(x), \quad x \in \mathbb{R}^d.$$

Entonces Q mapea $\ell^2(\mathbb{Z} \times \mathbb{Z}^d)$ en $L^2(\mathbb{R}^d)$ y, más aún, existe una constante $C = C(\alpha, \beta, C_1)$ tal que

$$\|Q(a)\|_{L^2(\mathbb{R}^d)} \leq C \|a\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)}, \quad \forall a \in \ell^2(\mathbb{Z} \times \mathbb{Z}^d), \quad (4.18)$$

es decir, la transformación $Q : \ell^2(\mathbb{Z} \times \mathbb{Z}^d) \rightarrow L^2(\mathbb{R}^d)$ es continua.

Demostración. Para probar (4.18) debemos estimar la siguiente norma

$$\|Q(a)\|_{L^2(\mathbb{R}^d)}^2 = \int_{\mathbb{R}^d} \left(\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} v_{j,k}(x) \right)^2 dx = \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} \sum_{j' \in \mathbb{Z}, k' \in \mathbb{Z}^d} a_{j,k} a_{j',k'} \int_{\mathbb{R}^d} v_{j,k}(x) v_{j',k'}(x) dx. \quad (4.19)$$

El primer paso de esta demostración consistirá entonces en buscar cotas para los términos $|I_{(j,k,j',k')}| \doteq \left| \int_{\mathbb{R}^d} v_{j,k}(x) v_{j',k'}(x) dx \right|$. Claramente, utilizando un argumento de simetría es posible reducir el caso general al caso $j \leq j'$. Bajo tal supuesto se tiene que

$$\begin{aligned} |I_{(j,k,j',k')}| &\leq \int_{\mathbb{R}^d} |v_{j,k}(x)| |v_{j',k'}(x)| dx \\ &\leq C_1^2 2^{(j-j')\frac{d}{2}} \int_{\mathbb{R}^d} \frac{2^{j'd}}{(1 + |2^j x - k|)^{d+\alpha} (1 + |2^{j'} x - k'|)^{d+\alpha}} dx \quad (\text{por (4.13)}) \\ &= C_1^2 2^{(j-j')\frac{d}{2}} \int_{\mathbb{R}^d} \frac{1}{(1 + \frac{2^j}{2^{j'}} |2^{j'}(2^{-j}k - 2^{-j'}k') - u|)^{d+\alpha} (1 + |u|)^{d+\alpha}} du. \quad (\text{con } u = 2^{j'}x - k') \end{aligned} \quad (4.20)$$

Definiendo $b = \frac{2^j}{2^{j'}}$ y $z = 2^{j'}|k2^{-j} - k'2^{-j'}|$ se tiene que la integral en (4.20) puede escribirse como

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{1}{(1 + \frac{2^j}{2^{j'}} |2^{j'}(2^{-j}k - 2^{-j'}k') - u|)^{d+\alpha} (1 + |u|)^{d+\alpha}} du &= \int_{\mathbb{R}^d} \frac{1}{(1 + b|z - u|)^{d+\alpha} (1 + |u|)^{d+\alpha}} du \\ &= (f_b * f_1)(z), \end{aligned} \quad (4.21)$$

donde $f_b(x) \doteq \frac{1}{(1 + b|x|)^{\alpha+d}}$, $x \in \mathbb{R}^d$. Puesto que $j' \geq j$ se tiene que $0 < b \leq 1$ y, por lo tanto, el Lema 5.3 (ver Apéndice) implica entonces que existe una constante $\tilde{C} = \tilde{C}(d, \alpha)$ tal que

$$(f_b * f_1)(z) \leq \tilde{C} f_b(z)$$

$$= \tilde{C} \frac{1}{(1 + 2^j |2^{-j'} k' - 2^{-j} k|)^{d+\alpha}}. \quad (4.22)$$

Más precisamente, la constante \tilde{C} está dada por $\tilde{C} \doteq \frac{2\pi^{d/2}}{\alpha\Gamma(d/2)} (2^{d+2+\alpha} + 1)$.

con (4.21) y (4.22) en (4.20) se sigue que existe una constante $\tilde{C}_1 = \tilde{C}_1(d, \alpha) \doteq \tilde{C} C_1^2$ para la cual se verifica que

$$\begin{aligned} |I_{(j,k,j',k')}| &\leq \tilde{C}_1 2^{-|j-j'|\frac{d}{2}} \frac{1}{(1 + 2^j |2^{-j'} k' - 2^{-j} k|)^{d+\alpha}} \\ &= \tilde{C}_1 2^{-|j-j'|\frac{d}{2}} \left(\frac{2^{-j} + 2^{-j'}}{2^{-j} + 2^{-j'} + (1 + 2^{j-j'}) |2^{-j'} k' - 2^{-j} k|} \right)^{d+\alpha} \\ &\leq \tilde{C}_1 2^{-|j-j'|\frac{d}{2}} \left(\frac{2^{-j} + 2^{-j'}}{2^{-j} + 2^{-j'} + |2^{-j'} k' - 2^{-j} k|} \right)^{d+\alpha}. \end{aligned} \quad (4.23)$$

Por otro lado, utilizando las tres condiciones que definen a una familia de vaguelets, con α y β como en (4.13) y (4.15), respectivamente, obtenemos una segunda estimación para $|I_{(j,k,j',k')}|$, como sigue:

$$\begin{aligned} |I_{(j,k,j',k')}| &= \left| \int_{\mathbb{R}^d} (v_{j,k}(x) - v_{j,k}(2^{-j'} k')) v_{j',k'}(x) dx \right| && \text{(por (4.14))} \\ &\leq C_1 2^{j(\frac{d}{2}+\beta)} \int_{\mathbb{R}^d} |x - 2^{-j'} k'|^\beta |v_{j',k'}(x)| dx && \text{(por (4.15))} \\ &\leq C_1 C_1 2^{j(\frac{d}{2}+\beta)} 2^{d\frac{j'}{2}} \int_{\mathbb{R}^d} \frac{2^{-j'\beta} |2^{j'} x - k'|^\beta}{(1 + |2^{j'} x - k'|)^{d+\alpha}} dx && \text{(por (4.13))} \\ &\leq C_1^2 2^{j(\frac{d}{2}+\beta)} 2^{d\frac{j'}{2}} 2^{-j'\beta} \int_{\mathbb{R}^d} \frac{1}{(1 + |2^{j'} x - k'|)^{d+\alpha-\beta}} dx \\ &= C_1^2 2^{\frac{j'd}{2}} 2^{\frac{j'd}{2}} 2^{(j-j')\beta} \int_{\mathbb{R}^d} \frac{2^{-j'd}}{(1 + |u|)^{d+\alpha-\beta}} du && \text{(haciendo } u = 2^{j'} x - k') \\ &\leq C_1^2 2^{-|j-j'|(\beta+\frac{d}{2})} \frac{2\pi^{\frac{d}{2}}}{(\alpha-\beta)\Gamma(d/2)} && \text{(por Proposición (5.5))} \\ &= \tilde{C}_2 2^{-|j-j'|(\beta+\frac{d}{2})}, \end{aligned} \quad (4.24)$$

con $\tilde{C}_2 = \tilde{C}_2(d, \alpha, \beta) \doteq C_1^2 \frac{2\pi^{\frac{d}{2}}}{(\alpha-\beta)\Gamma(d/2)}$.

Finalmente, definiendo $\theta \doteq \frac{\alpha}{\alpha+\beta+d}$, escribiendo $|I| = |I|^{1-\theta} |I|^\theta$ y utilizando (4.23) para acotar el primer factor y (4.24) para acotar el segundo obtenemos la siguiente estimación

$$\begin{aligned} |I_{(j,k,j',k')}| &= |I_{(j,k,j',k')}|^{1-\theta} |I_{(j,k,j',k')}|^\theta \\ &\leq \tilde{C}_3 \left(2^{-|j-j'|\frac{d}{2}} \right)^{1-\theta} \left(\frac{2^{-j} + 2^{-j'}}{2^{-j} + 2^{-j'} + |k' 2^{-j'} - k 2^{-j}|} \right)^{(d+\alpha)(1-\theta)} \left(2^{-|j-j'|(\beta+\frac{d}{2})} \right)^\theta \\ &= \tilde{C}_3 2^{-|j-j'|(\gamma+\frac{d}{2})} \left(\frac{2^{-j} + 2^{-j'}}{2^{-j} + 2^{-j'} + |k' 2^{-j'} - k 2^{-j}|} \right)^{d+\gamma} \end{aligned}$$

$$\doteq m_{j,k,j',k'} \quad (4.25)$$

donde $\gamma = \beta\theta \doteq \frac{\beta\alpha}{\alpha + \beta + d}$ y $\tilde{C}_3 = \tilde{C}_3(d, \alpha, C_1, \beta) \doteq \tilde{C}_1^{1-\theta} \tilde{C}_2^\theta$.

Para simplificar la notación escribiremos $[Q(a)](x) = \sum_{\ell \in L} a_\ell v_\ell(x)$, donde $\ell = (j, k)$ y $L = \{(j, k) \in \mathbb{Z} \times \mathbb{Z}^d\}$, con lo que (4.25) toma la forma $|I_{\ell, \ell'}| \leq m_{\ell, \ell'}$. Luego, con (4.25) en (4.19) tenemos entonces que

$$\begin{aligned} \|Q(a)\|_{L^2(\mathbb{R}^d)}^2 &\leq \sum_{\ell \in L} \sum_{\ell' \in L} a_\ell a_{\ell'} |I_{\ell, \ell'}| && \text{(por (4.19))} \\ &\leq \sum_{\ell \in L} a_\ell \sum_{\ell' \in L} a_{\ell'} m_{\ell, \ell'} && \text{(por (4.25))} \\ &= \langle a, Ma \rangle_{\ell^2(L)} \\ &\leq \|a\|_{\ell^2(L)} \|Ma\|_{\ell^2(L)}, \end{aligned} \quad (4.26)$$

donde M es la matriz infinita $M \doteq \{m_{\ell, \ell'}\}_{\ell, \ell' \in L}$.

Observando (4.26) vemos que sería suficiente con probar que el operador $M : \ell^2(L) \rightarrow \ell^2(L)$ es acotado, para lo cual utilizaremos el *Lema de Schur*. Para ello será suficiente con probar que $\{m_{\ell, \ell'}\}_{\ell, \ell' \in L}$ verifica las condiciones (4.16) y (4.17) para alguna sucesión $\{w_\ell\}_{\ell \in L}$ de números positivos. Veremos que ambas condiciones se cumplen para $w_\ell \doteq 2^{-j \frac{d}{2}}$. Separaremos la demostración en dos casos.

Caso 1: $j' \geq j$. Definimos $n \doteq j' - j$ y obtenemos la siguiente estimación

$$\begin{aligned} \sum_{j' \geq j} \sum_{k' \in \mathbb{Z}^d} w_{j', k'} m_{j, k, j', k'} &= \sum_{j' \geq j} \sum_{k' \in \mathbb{Z}^d} 2^{-j' \frac{d}{2}} \tilde{C}_3 2^{-|j-j'|(\gamma + \frac{d}{2})} \left(\frac{2^{-j} + 2^{-j'}}{2^{-j} + 2^{-j'} + |k' 2^{-j'} - k 2^{-j}|} \right)^{d+\gamma} \\ &\leq \tilde{C}_3 \sum_{j' \geq j} 2^{-j' \frac{d}{2}} 2^{-n(\gamma + \frac{d}{2})} \sum_{k' \in \mathbb{Z}^d} \left(\frac{2}{1 + |k - 2^{-n} k'|} \right)^{d+\gamma} \\ &\leq \tilde{C}_3 2^{d+\gamma} 2^{-j \frac{d}{2}} \sum_{j' \geq j} 2^{-j' \frac{d}{2}} 2^{j \frac{d}{2}} 2^{-n(\gamma + \frac{d}{2})} \sum_{k' \in \mathbb{Z}^d} \left(\frac{1}{1 + 2^{-n} |k'|} \right)^{d+\gamma} \\ &\leq \tilde{C}_3 2^{d+\gamma} 2^{-j \frac{d}{2}} \sum_{n \geq 0} 2^{-nd} 2^{-n\gamma} \sum_{k' \in \mathbb{Z}^d} \left(\frac{1}{1 + 2^{-n} |k'|} \right)^{d+\gamma} \\ &\leq \tilde{C}_3 2^{d+\gamma} 2^{-j \frac{d}{2}} \sum_{n \geq 0} 2^{-nd} 2^{-n\gamma} (1 + 2 \cdot 2^{nd} \gamma^{-1}) \quad \text{(por el Lema 5.4)} \\ &\leq \tilde{C}_3 (1 + 2 \gamma^{-1}) 2^{d+\gamma} 2^{-j \frac{d}{2}} \sum_{n \geq 0} 2^{-n\gamma} \\ &= \tilde{C}_4 2^{-j \frac{d}{2}} \\ &= \tilde{C}_4 w_{j, k}. \end{aligned}$$

donde la primer desigualdad se sigue el hecho que $\frac{2^{-j}}{2^{-j} + 2^{-j'}} \geq \frac{1}{2}$ y $\tilde{C}_4 = \tilde{C}_4(\tilde{C}_3, d, \gamma)$.

Caso 2: $j > j'$. Aquí definimos $n \doteq j - j'$ y obtenemos

$$\begin{aligned}
\sum_{j' < j} \sum_{k' \in \mathbb{Z}^d} w_{j',k'} m_{j,k,j',k'} &= \sum_{j' \geq j} \sum_{k' \in \mathbb{Z}^d} 2^{-j' \frac{d}{2}} \tilde{C}_3 2^{-|j-j'|(\gamma + \frac{d}{2})} \left(\frac{2^{-j} + 2^{-j'}}{2^{-j} + 2^{-j'} + |k'2^{-j'} - k2^{-j}|} \right)^{d+\gamma} \\
&\leq \tilde{C}_3 \sum_{j' < j} 2^{-j' \frac{d}{2}} \sum_{k' \in \mathbb{Z}^d} 2^{-n(\gamma + \frac{d}{2})} \left(\frac{2}{1 + |2^{-n}k - k'|} \right)^{d+\gamma} \\
&\leq \tilde{C}_3 \sum_{j' < j} 2^{-j' \frac{d}{2}} 2^{-n(\gamma + \frac{d}{2})} 2^{d+\gamma} \sum_{k' \in \mathbb{Z}^d} \left(\frac{1}{1 + |k'|} \right)^{d+\gamma} \\
&\leq \tilde{C}_5 \sum_{j' < j} 2^{-j' \frac{d}{2}} 2^{-(j-j')(\gamma + \frac{d}{2})} \\
&= \tilde{C}_5 2^{-j \frac{d}{2}} \sum_{n > 0} 2^{n\gamma} \\
&\leq \tilde{C}_6 2^{-j \frac{d}{2}} \\
&= \tilde{C}_6 w_{j,k}.
\end{aligned} \tag{4.27}$$

Luego, con $c = [\max\{\tilde{C}_4, \tilde{C}_6\}]^{-1}$ se tiene que los elementos de la matriz infinito dimensional cM satisfacen (4.16). La desigualdad (4.17) se prueba de manera análoga, por simetría. Se sigue del Lema de Schur que $cM : \ell^2(L) \rightarrow \ell^2(L)$ es acotado y por lo tanto M también lo es. \blacksquare

Estamos ahora en condiciones de demostrar la propiedad de casi-ortogonalidad de dos conjuntos de vaguelets.

Lema 4.16. Sean $\{v_{j,k}\}_{j,k}$ y $\{u_{j,k}\}_{j,k}$ dos conjuntos bi-ortonormales de vaguelets. Entonces

$\forall \{a_{j,k}\}_{j,k} \in \ell^2(\mathbb{Z} \times \mathbb{Z}^d)$

$$\left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} u_{j,k} \right\|_{L^2(\mathbb{R}^d)} \equiv \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)} \quad y \quad \left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} v_{j,k} \right\|_{L^2(\mathbb{R}^d)} \equiv \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)},$$

es decir, existen constantes positivas M_u, m_u, M_v y m_v tales que

$$m_u \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)} \leq \left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} u_{j,k} \right\|_{L^2(\mathbb{R}^d)} \leq M_u \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)} \quad y \tag{4.28}$$

$$m_v \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)} \leq \left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} v_{j,k} \right\|_{L^2(\mathbb{R}^d)} \leq M_v \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)}. \tag{4.29}$$

Demostración. Puesto que $\{u_{j,k}\}_{j,k}$ es una familia de vaguelets, el Teorema 4.15 implica que existe una constante $M_u > 0$ tal que $\forall \{a_{j,k}\}_{j,k} \subset \ell^2(\mathbb{Z} \times \mathbb{Z}^d)$ se verifica que

$$\left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} u_{j,k} \right\|_{L^2(\mathbb{R}^d)} \leq M_u \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)}.$$

Por otro lado,

$$\begin{aligned} \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)}^2 &= \left\langle \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} v_{j,k}, \sum_{j' \in \mathbb{Z}, k' \in \mathbb{Z}^d} a_{j',k'} u_{j',k'} \right\rangle_{L^2(\mathbb{R}^d)} \quad (\text{por la bi-ortogonalidad}) \\ &\leq \left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} v_{j,k} \right\|_{L^2(\mathbb{R}^d)} \left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} u_{j,k} \right\|_{L^2(\mathbb{R}^d)}. \quad (\text{por desigualdad de Schwartz}) \end{aligned} \quad (4.30)$$

Ahora, puesto que $\{v_{j,k}\}_{j,k}$ es una familia de vaguelets, el Teorema 4.15 implica que existe una constante positiva C tal que

$$\left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} v_{j,k} \right\|_{L^2(\mathbb{R}^d)} \leq C \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)}. \quad (4.31)$$

De (4.30) y (4.31), con $m_v = C^{-1}$ se sigue que $m_v \|a_{j,k}\|_{\ell^2(\mathbb{Z} \times \mathbb{Z}^d)} \leq \left\| \sum_{j \in \mathbb{Z}, k \in \mathbb{Z}^d} a_{j,k} u_{j,k} \right\|_{L^2(\mathbb{R}^d)}$, lo cual concluye la demostración de (4.28). La prueba de (4.29) se sigue de manera análoga. ■

4.4.3. Descomposición y regularización wavelet-vaguelet

El objetivo de esta sección radica, principalmente, en desarrollar la teoría necesaria para la utilización de descomposiciones wavelet-vaguelet de un operador en el contexto de problemas inversos. A lo largo de esta sección definimos el concepto de descomposición wavelet-vaguelet de un operador con respecto a una base ortonormal wavelet. Posteriormente, demostramos un teorema que nos brinda condiciones suficientes sobre el operador y la madre wavelet generadora de la base, para que dicho operador posea una DWV. Finalmente, realizamos una serie de observaciones sobre la aplicación de esta descomposición en el caso de los problemas inversos y sobre la implicancia de cada una de las condiciones impuestas en el teorema mencionado. Haremos hincapié principalmente en que todas ellas se satisfacen para operadores integrales, bajo ciertas condiciones no demasiado restrictivas que especificamos en la Sección 4.6.

Al igual que en la descomposición wavelet, en la descomposición wavelet-vaguelet partimos de una función desconocida representada en términos de una expansión wavelet (ver (1.22)).

Dados $\Omega \subseteq \mathbb{R}^d$, $j_0 \in \mathbb{Z}$ fijo, $\lambda = (j, k, \epsilon)$, $\Lambda \doteq \{(j, k, \epsilon) \in \mathbb{Z} \times \mathbb{Z}^d \times \Gamma^0 : j \geq j_0 \wedge \text{sup}(\psi_\lambda) \subset \Omega\}$ y una base ortonormal wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ en $L^2(\Omega) \forall x \in \mathcal{D}(T)$ se tiene la siguiente representación

$$x = \sum_{\lambda \in \Lambda} \langle x, \psi_\lambda \rangle \psi_\lambda. \quad (4.32)$$

El punto crucial de la descomposición wavelet-vaguelet reside en que para algunos operadores existe una sucesión de constantes que permiten encontrar dos familias de vaguelets asociadas por dualidad.

Definición 4.17. Sea $\Omega \subseteq \mathbb{R}^d$. Decimos que un operador $T : L^2(\Omega) \rightarrow L^2(\Omega)$ lineal, inyectivo, cerrado, densamente definido, con $\mathcal{R}(T)$ denso en $L^2(\Omega)$, posee una *descomposición wavelet-vaguelet* con respecto a una base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{D}(T) \cap \mathcal{R}(T^*)$, si existe una sucesión de constantes $\{\kappa_\lambda\}_{\lambda \in \Lambda} \subset \ell^2(\Lambda)$ tal que las sucesiones $\{u_\lambda\}_{\lambda \in \Lambda}$ y $\{v_\lambda\}_{\lambda \in \Lambda}$ definidas por $u_\lambda(t) \doteq \kappa_\lambda (T^{-1})^* \psi_\lambda(t)$ y $v_\lambda(t) \doteq \kappa_\lambda^{-1} T \psi_\lambda(t)$ son vaguelets.

En tal caso nos referiremos a $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ como una descomposición wavelet-vaguelet de T con respecto a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ y a $\{\kappa_\lambda\}_{\lambda \in \Lambda}$ como los valores casi-singulares de dicha descomposición.

Observación 4.18. Notar que si $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ es una descomposición wavelet-vaguelet del operador T con respecto a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ entonces todo $y \in \mathcal{R}(T)$ puede escribirse de la forma

$$y = \sum_{\lambda \in \Lambda} \langle y, u_\lambda \rangle v_\lambda. \quad (4.33)$$

En efecto, si $y \in \mathcal{R}(T)$ entonces $Tx = y$ para algún $x \in \mathcal{D}(T)$. Luego,

$$\begin{aligned} y &= \sum_{\lambda \in \Lambda} \langle x, \psi_\lambda \rangle T \psi_\lambda && \text{(por (4.32) y } T \text{ cerrado)} \\ &= \sum_{\lambda \in \Lambda} \langle x, \kappa_\lambda \psi_\lambda \rangle \kappa_\lambda^{-1} T \psi_\lambda \\ &= \sum_{\lambda \in \Lambda} \langle T^{-1} y, \kappa_\lambda \psi_\lambda \rangle \kappa_\lambda^{-1} T \psi_\lambda \\ &= \sum_{\lambda \in \Lambda} \langle y, \kappa_\lambda (T^{-1})^* \psi_\lambda \rangle \kappa_\lambda^{-1} T \psi_\lambda \\ &= \sum_{\lambda \in \Lambda} \langle y, u_\lambda \rangle v_\lambda. \end{aligned}$$

Observación 4.19. Si $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ es una descomposición wavelet-vaguelet del operador T con respecto a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ entonces T es débilmente invertible con respecto a esa

base. En efecto, de la observación anterior se tiene que $\forall x \in \mathcal{D}(T)$

$$Tx = \sum_{\lambda \in \Lambda} \langle Tx, u_\lambda \rangle \kappa_\lambda^{-1} T\psi_\lambda.$$

Y por lo tanto,

$$\begin{aligned} x &= \sum_{\lambda \in \Lambda} \langle Tx, u_\lambda \rangle \kappa_\lambda^{-1} \psi_\lambda && \text{(pues } T^{-1} \text{ es cerrado)} \\ &= \sum_{\lambda \in \Lambda} \langle Tx, c_\lambda \rangle \psi_\lambda && \text{(donde } c_\lambda \doteq \kappa_\lambda^{-1} u_\lambda \text{).} \end{aligned}$$

Luego, $\langle Tx, c_\lambda \rangle = \langle x, \psi_\lambda \rangle$ y por lo tanto T es débilmente invertible con respecto a la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$ (ver Definición 4.1).

Introducimos ahora el concepto de operador traslación con el objetivo de simplificar la notación en lo que resta de esta sección.

Definición 4.20. Dados $d \in \mathbb{N}$ y $a \in \mathbb{R}^d$ definimos el *operador traslación* $E_a : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ como $E_a f(t) \doteq f(t - a)$.

Notemos que los operadores dilatación y traslación no conmutan, más aún, una sencilla cuenta prueba que dados $b > 0$ y $a \in \mathbb{R}^d$ se verifica que $E_a D_b = D_b E_{ab}$.

A continuación enunciamos y demostramos un teorema que provee condiciones suficientes para que un operador posea una descomposición wavelet vaguelet con respecto a una base wavelet dada. Como consecuencia y en función de la observación 4.19, estas condiciones implican además la invertibilidad débil de T con respecto a dicha base y la existencia de la correspondiente sucesión de funcionales $\{c_\lambda\}_{\lambda \in \Lambda}$.

Teorema 4.21 (Descomposición wavelet-vaguelet de un operador). *Sea $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ un operador lineal, inyectivo, cerrado, densamente definido, homogéneo de grado α , con $\mathcal{R}(T)$ denso en $L^2(\mathbb{R}^d)$. Sea, además, ψ una madre wavelet con promedio nulo, y supongamos que se satisfacen las siguientes hipótesis:*

H1) $\psi \in \mathcal{D}(T) \cap \mathcal{R}(T^*)$;

H2) T conmuta con traslaciones, es decir, $\forall a \in \mathbb{R}^d$ se verifica que $TE_a = E_a T$;

H3) Existen constantes $b_1, b_2 > 0$ y $\nu_1, \nu_2 > 0$, tales que $\forall x \in \mathbb{R}^d$

$$|[T\psi(\cdot)](x)| \leq \frac{b_1}{(1 + |x|)^{d+\nu_1}} \quad \text{y} \quad |[(T^{-1})^* \psi(\cdot)](x)| \leq \frac{b_2}{(1 + |x|)^{d+\nu_2}};$$

H4) $T\psi$ y $(T^{-1})^* \psi$ son Hölder continuas, es decir, existen constantes $b_3, b_4 > 0$ y $\beta_1, \beta_2 \in \mathbb{R}$ con $0 < \beta_1 < \nu_1$ y $0 < \beta_2 < \nu_2$ tales que $\forall x, x' \in \mathbb{R}^d$

$$|[T\psi(\cdot)](x) - [T\psi(\cdot)](x')| \leq b_3 |x - x'|^{\beta_1} \quad \text{y}$$

$$\left| [(T^{-1})^* \psi(\cdot)](x) - [(T^{-1})^* \psi(\cdot)](x') \right| \leq b_4 |x - x'|^{\beta_2};$$

H5) $T\psi$ y $(T^{-1})^ \psi$ tienen promedio nulo.*

Entonces la familia $\{(2^{-jd\alpha}, 2^{-jd\alpha}(T^{-1})^* \psi_\lambda, 2^{jd\alpha} T \psi_\lambda)\}_{\lambda \in \Lambda}$ es una descomposición wavelet-vaguelet del operador T con respecto a la base ortonormal wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ y $\{2^{-j\alpha d}\}_{j \geq j_0}$ es la correspondiente sucesión de valores casi-singulares de dicha descomposición.

Demostración. En primer lugar observar que puesto que T es homogénea y satisface *H2*, la hipótesis *H1* implica que $\psi_\lambda \in \mathcal{D}(T) \cap \mathcal{R}(T^*)$ para todo $\lambda \in \Lambda$. En efecto, como T es homogéneo, $\mathcal{D}(T)$ es cerrado por dilataciones y una cuenta sencilla (ver Apéndice, Proposición 5.6) prueba que en tal caso $\mathcal{R}(T^*)$ también lo es. Por otro lado, si $\mathcal{D}(T)$ es cerrado por traslaciones y T conmuta con ellas, es decir, satisface la hipótesis *H2*, entonces otra sencilla cuenta (ver Apéndice, Proposición 5.7) prueba que $\mathcal{D}(T^*)$ también es cerrado por traslaciones y T^* conmuta con ellas. Luego, para todo $\lambda \in \Lambda$ están bien definidas $u_\lambda \doteq 2^{-jd\alpha}(T^{-1})^* \psi_\lambda$ y $v_\lambda \doteq 2^{jd\alpha} T \psi_\lambda$.

Debemos probar que u_λ y v_λ satisfacen las tres condiciones de la Definición 4.13, $\forall \lambda \in \Lambda$. En lo que sigue, realizaremos la prueba de (4.13), (4.14) y (4.15) para demostrar que $\{v_\lambda\}_{\lambda \in \Lambda}$ es una familia de vaguelets. La demostración correspondiente a $\{u_\lambda\}_{\lambda \in \Lambda}$ se puede obtener de manera análoga (ver Observación 4.22).

Notemos en primer lugar que $\forall \lambda \in \Lambda$ se tiene que

$$\begin{aligned} v_\lambda(x) &= 2^{jd\alpha} [T\psi_\lambda(\cdot)](x) \\ &= 2^{jd\alpha} 2^{\frac{dj}{2}} [T\psi(2^j \cdot - k)](x) \\ &= 2^{jd\alpha} 2^{\frac{dj}{2}} [TD_{2^j} \psi(\cdot - k2^{-j})](x) \\ &= 2^{\frac{dj}{2}} D_{2^j} [T\psi(\cdot - k2^{-j})](x) && \text{(pues } T \text{ es homogéneo de grado } \alpha) \\ &= 2^{\frac{dj}{2}} D_{2^j} [TE_{k2^{-j}} \psi(\cdot)](x) \\ &= 2^{\frac{dj}{2}} D_{2^j} E_{k2^{-j}} [T\psi(\cdot)](x) && \text{(por } H2) \\ &= 2^{\frac{dj}{2}} [T\psi(\cdot)](2^j x - k). \end{aligned} \tag{4.34}$$

De (4.34) se sigue entonces que

$$\begin{aligned} |v_\lambda(x)| &= 2^{\frac{dj}{2}} |[T\psi(\cdot)](2^j x - k)| \\ &\leq 2^{\frac{dj}{2}} \frac{b_1}{(1 + |2^j x - k|)^{d+\nu_1}}. \end{aligned} \tag{por } H3$$

Luego, v_λ satisface la condición (4.13) para todo $\lambda \in \Lambda$.

Probaremos a continuación que $\{v_\lambda\}_{\lambda \in \Lambda}$ satisface (4.15). Para ello notar que para todo $\lambda \in \Lambda$ y $\forall x, x' \in \mathbb{R}^d$

$$\begin{aligned} |v_\lambda(x) - v_\lambda(x')| &= 2^{\frac{dj}{2}} |[T\psi(\cdot)](2^j x - k) - [T\psi(\cdot)](2^j x' - k)| && \text{(por (4.34))} \\ &\leq b_3 2^{\frac{dj}{2}} |2^j x - 2^j x'|^{\beta_1} && \text{(por } H_4) \\ &= b_3 2^{j(\frac{d}{2} + \beta_1)} |x - x'|^{\beta_1}. \end{aligned}$$

Finalmente, (4.14) se sigue trivialmente de la condición *H5*. En efecto, para todo $\lambda \in \Lambda$

$$\int_{\mathbb{R}^d} v_\lambda dx = 2^{\frac{dj}{2}} \int_{\mathbb{R}^d} [T\psi](2^j x - k) dx = 2^{-\frac{dj}{2}} \int_{\mathbb{R}^d} [T\psi](u) du = 0. \quad \blacksquare$$

Observación 4.22. La condición *H2* pide que el operador T conmute con traslaciones. Supuestas las condiciones *H1* y *H2* se sigue que $\mathcal{R}(T)$ es cerrado por traslaciones y $(T^{-1})^*$ conmuta con ellas (ver Apéndice, Proposición 5.7 y Lema 5.8). Algunos operadores que satisfacen esta condición son, por ejemplo, los operadores diferenciales y operadores de convolución. En efecto, si T es un operador de convolución con núcleo $k(\cdot)$, entonces

$$\begin{aligned} [TE_a\psi](x) &= [T\psi(\cdot - a)](x) \\ &= \int_{\mathbb{R}^d} k(x - t)\psi(t - a) dt \\ &= \int_{\mathbb{R}^d} k((x - a) - u)\psi(u) du && \text{(haciendo } u = t - a) \\ &= [T\psi(\cdot)](x - a) \\ &= E_a [T\psi](x). \end{aligned}$$

Observación 4.23. La hipótesis *H3* claramente impone condiciones de acotación e integrabilidad de $T\psi$ y $(T^{-1})^* \psi$.

Observación 4.24. La condición *H4* impone que las imágenes de la madre wavelet ψ por los operadores T y $(T^{-1})^*$ sean Hölder continuas. En el caso en el que el operador T sea un operador de convolución con núcleo $k \in L^2(\mathbb{R}^d)$, una condición suficiente que garantiza *H4* es que k sea Hölder continuo. Para el caso de operadores integrales de tipo convolución con núcleo $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ una sencilla cuenta (ver apéndice, Proposición 5.9) permite probar que, si k es Hölder continuo (de orden ν) con respecto a la primera variable, es decir, existe $c : \mathbb{R}^d \rightarrow \mathbb{R}$ tal que $|k(x, t) - k(x', t)| \leq c(t)|x - x'|^\nu$, y si además $c(\cdot) \in L^1(\mathbb{R}^d)$, entonces $T\psi$ también es Hölder continua de orden ν .

Observación 4.25. La condición *H5* implica que los operadores T y $(T^{-1})^*$ mantienen el promedio nulo de la madre wavelet. Esta hipótesis en particular es satisfecha por los operadores integrales de convolución con núcleo $k \in L^1(\mathbb{R}^d)$. En efecto,

$$\begin{aligned} \int_{\mathbb{R}^d} [T\psi](x) dx &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k(x-t)\psi(t) dt dx \\ &= \int_{\mathbb{R}^d} \psi(t) \int_{\mathbb{R}^d} k(x-t) dx dt \\ &= C \int_{\mathbb{R}^d} \psi(t) dt && (C \doteq \int_{\mathbb{R}^d} k(x) dx) \\ &= 0. \end{aligned}$$

4.5. Análisis de convergencia

En esta sección estudiaremos la convergencia de los métodos de regularización propuestos en la Sección 4.4.1: el método de regularización wavelet-vaguelet (RWV) y el método de regularización wavelet-vaguelet con umbralado (RWVU). Para ello, luego de recordar brevemente en que consiste cada método, definiremos varias medidas de error entre las que se destacan en particular las de tipo minimax. Posteriormente, enunciaremos y demostraremos varios resultados que nos permitirán, al finalizar la sección, encontrar cotas para los ordenes de convergencia de los métodos RWV y RWVU.

4.5.1. Los métodos RWV y RWVU

Como expresamos anteriormente, en este capítulo consideramos el problema inverso de aproximar f en

$$Tf = g, \tag{4.35}$$

donde $T : L^2(\Omega) \rightarrow L^2(\Omega)$ es un operador lineal, inyectivo, cerrado, densamente definido, con $\mathcal{R}(T)$ denso en $L^2(\Omega)$, $\Omega \subset \mathbb{R}^d$. Suponemos además un modelo de observación con ruido aditivo de la forma $g^\delta = g + \delta dW$, donde dW es ruido blanco y $\delta > 0$. En la sección anterior propusimos el método RWV para encontrar una aproximación a la solución exacta del problema (4.35). Como vimos, este método consiste en elegir adecuadamente una base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{D}(T) \cap \mathcal{R}(T^*)$ con respecto a la cual el operador T posea una DWV $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ (ver Definición 4.17). De la Observación 4.19 se sigue entonces que T es débilmente invertible con respecto a dicha base. Luego, el método RWV consiste en proponer una solución aproximada para el problema (4.35)

de la forma:

$$f_c^\delta(x) \doteq \sum_{\lambda \in \Lambda} c_\lambda(g^\delta) \psi_\lambda(x) = \sum_{\lambda \in \Lambda} \kappa_\lambda^{-1} \langle g^\delta, u_\lambda \rangle \psi_\lambda(x), \quad x \in \Omega. \quad (4.36)$$

Por otro lado, dado un umbral $\gamma > 0$ convenientemente elegido, el método RWVU consiste en aplicar el operador de umbralado wavelet fuerte S_γ (ver definición (4.10)) a la aproximación propuesta por el método RWV. Así, la solución aproximada para el problema (4.35) propuesta por el método RWVU queda definida por

$$f_{c,\gamma}^\delta(x) \doteq S_\gamma f_c^\delta(x), \quad x \in \Omega. \quad (4.37)$$

4.5.2. Errores

A los efectos de evaluar la performance de los métodos RWV y RWVU, en esta sección definiremos una serie de conceptos que nos permitirán luego cuantificar los diferentes errores involucrados. En primer lugar, consideraremos el caso ya definido y desarrollado, donde los datos provienen de un modelo continuo. A los efectos de tener en cuenta este tipo de problemas inversos con un enfoque que destaque el interés práctico, continuaremos analizando el caso del correspondiente modelo discreto. Culminaremos la sección incluyendo un enfoque bayesiano que nos permitirá adoptar una visión estocástica en los modelos propuestos.

Modelo continuo

A continuación definimos una medida de error de tipo minimax, la cual resultará fundamental para encontrar cotas para los órdenes de convergencia de los métodos RWV y RWVU. A lo largo de esta sección $\mathcal{F} \subset L^2(\mathbb{R}^d)$ denotará un conjunto fuente para f y \mathcal{M} un conjunto de funciones admisibles para las aproximaciones obtenidas con el método utilizado.

Definición 4.26 (Riesgo minimax cuadrático). Sea $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ el operador asociado al problema de aproximar f en (4.35). Suponemos que $f \in \mathcal{F}$ y que el modelo observacional es estocástico con ruido aditivo: $g^\delta = Tf + \delta dW$, donde $\delta > 0$. Para este problema, el *riesgo minimax cuadrático* (RMC) asociado al conjunto fuente \mathcal{F} y al conjunto admisible \mathcal{M} se define como

$$\mathcal{R}(\delta, \mathcal{F}, \mathcal{M}) \doteq \inf_{f^\delta \in \mathcal{M}} \sup_{f \in \mathcal{F}} E \|f^\delta - f\|_{L^2(\mathbb{R}^d)}^2. \quad (4.38)$$

Con el objetivo de simplificar la notación, en el caso en que \mathcal{M} sea el conjunto de aproximaciones obtenidas utilizando el método de RWV, denotaremos a tal conjunto mediante \mathcal{M}_c y, en el caso en que \mathcal{M} sea el conjunto de aproximaciones obtenidas utilizando el método de RWVU, lo denotaremos con $\mathcal{M}_{c,\gamma}$. Adicionalmente, nos referiremos al riesgo minimax cuadrático asociado a dichos métodos, es decir a $\mathcal{R}(\delta, \mathcal{F}, \mathcal{M}_c)$ y $\mathcal{R}(\delta, \mathcal{F}, \mathcal{M}_{c,\gamma})$, como riesgo minimax

cuadrático wavelet-vaguelet (RMC-WV) y riesgo minimax cuadrático wavelet-vaguelet con umbralado (RMC-WVU), respectivamente.

Para el caso en que las soluciones aproximadas del problema (4.35) sean las correspondientes a los métodos RWV y RWVU, nos resultará útil considerar a sus respectivos riesgos minimax cuadráticos en términos de los coeficientes wavelet de las funciones involucradas. Para ello, notar que, en este caso, los coeficientes wavelet de la solución aproximada f^δ , a los que denotamos con $\alpha_\lambda^\delta \doteq c_\lambda(g^\delta)$, pueden escribirse en términos de los coeficientes $\alpha_\lambda = \langle f, \psi_\lambda \rangle$ de la solución exacta f y de la DWV. En efecto,

$$c_\lambda(g^\delta) = \langle Tf + \delta dW, c_\lambda \rangle = \langle Tf, c_\lambda \rangle + \delta \langle dW, c_\lambda \rangle = \alpha_\lambda + \delta \kappa_\lambda^{-1} \langle dW, u_\lambda \rangle. \quad (4.39)$$

Obtenemos así una relación entre los coeficientes de las descomposiciones wavelet de f y f^δ :

$$\alpha_\lambda^\delta = \alpha_\lambda + \delta \sigma_\lambda W_\lambda \quad \forall \lambda \in \Lambda, \quad (4.40)$$

donde $\sigma_\lambda \doteq \kappa_\lambda^{-1}$ y $W_\lambda = \langle dW, u_\lambda \rangle$ es una variable aleatoria normal con media cero (ver Sección 1.6).

Estamos interesados, en particular, en el caso en que \mathcal{F} es el espacio de Besov $B_{p,q}^s(\mathbb{R}^d)$ ($s > 0$, $0 < p, q \leq \infty$) (ver Sección 1.5). Como vimos en el Lema 1.20, si $\psi \in L^r(\mathbb{R}^d)$ para algún $r \geq \max\{1, p\}$, entonces $f \in B_{p,q}^s(\mathbb{R}^d)$ si y solo si sus correspondientes coeficientes wavelet α_λ satisfacen $\{\alpha_\lambda\}_{\lambda \in \Lambda} \in \Theta_{p,q}^s$ (ver (1.26)). A partir de esta caracterización, podemos derivar la siguiente expresión para el RMC-WV asociado a un nivel de ruido δ y la condición fuente $B_{p,q}^s(\mathbb{R}^d)$:

$$\begin{aligned} \mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_c) &= \inf_{f^\delta \in \mathcal{M}_c} \sup_{f \in B_{p,q}^s} E \|f^\delta - f\|_{L^2(\Omega)}^2 \\ &= \inf_{\alpha_\lambda^\delta: f^\delta \in \mathcal{M}_c} \sup_{\{\alpha_\lambda\} \in \Theta_{p,q}^s} E \|\alpha_\lambda^\delta - \alpha_\lambda\|_{\ell^2(\Lambda)}^2. \end{aligned} \quad (4.41)$$

De aquí en más, en virtud de (4.41) utilizaremos indistintamente $\mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_c)$ o $\mathcal{R}(\delta, \Theta_{p,q}^s, \mathcal{M}_c)$ para denotar al RMC-WV asociado a δ y $B_{p,q}^s$. Idéntica notación adoptaremos para el RWC-WVU.

Modelo discreto

Hasta aquí, el modelo asociado a RMC-WV y RMC-WVU se corresponde con un modelo continuo. Sin embargo, como anticipamos, este no es el único modelo que consideraremos. Nos interesa particularmente diferenciar los RMC asociados al problema cuando el dato se obtiene de un modelo continuo y cuando se obtiene a partir de un modelo discreto. Una discretización estándar (cualquiera) del espacio $L^2(\Omega)$, sumado al hecho de que T es un operador lineal y

$\{u_\lambda\}_{\lambda \in \Lambda}$ es una familia de vaguelets, originan un modelo discreto análogo a (4.40):

$$\alpha_\lambda^\delta = \alpha_\lambda + \delta \sigma_\lambda \tilde{W}_\lambda \quad \forall \lambda \in \Lambda, \quad (4.42)$$

donde \tilde{W}_λ son realizaciones de ruido blanco discreto (V.A. i.i.d. $\sim N(0, 1)$). En el caso de este nuevo modelo discretizado, denotaremos con $\mathcal{R}_D(\delta, \Theta_{p,q}^s, \mathcal{M}_c)$ y $\mathcal{R}_D(\delta, \Theta_{p,q}^s, \mathcal{M}_{c,\gamma})$ a los errores RMC-WV y RMC-WVU, respectivamente, asociados a δ y $\Theta_{p,q}^s$.

Enfoque bayesiano

Este enfoque parte de suponer que el objeto de interés no es una función $f \in L^2(\mathbb{R}^d)$ sino más bien una función aleatoria F (con valores en $L^2(\mathbb{R}^d)$), de la cual f puede considerarse como una realización. En este caso, claramente no tiene sentido imponer la condición fuente de la pertenencia de F en el espacio de Besov $B_{p,q}^s(\mathbb{R}^d)$. No obstante ello, sí es posible concebir que F puede expandirse (de manera análoga a su contraparte determinística) en términos de la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ en la forma

$$F(x) = \sum_{\lambda \in \Lambda} A_\lambda \psi_\lambda(x), \quad x \in \mathbb{R}^d, \quad (4.43)$$

donde las A_λ son variables aleatorias (sobre $L^2(\mathbb{R}^d)$) con una cierta distribución. En esta formulación, la equivalencia $f \in B_{p,q}^s(\mathbb{R}^d) \iff \{\alpha_\lambda\}_{\lambda \in \Lambda} \in \Theta_{p,q}^s$, tiene un claro e inmediato correlato en función de los coeficientes aleatorios A_λ de la expansión de F en (4.43). En efecto, reemplazando $|\alpha_\lambda|^p$ por $E(|A_\lambda|^p)$, la correspondiente condición resulta de imponer a la distribución conjunta Π_A de la sucesión de variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$ la siguiente condición sobre sus p -momentos absolutos:

$$\left\{ \left[E(|A_\lambda|^p) \right]^{\frac{1}{p}} \right\}_{\lambda \in \Lambda} \in \Theta_{p,q}^s. \quad (4.44)$$

En lo que sigue denotaremos con $\Xi_{p,q}^s$ al conjunto de variables aleatorias (o equivalentemente al de sus correspondientes distribuciones conjuntas) que satisfacen (4.44). Es decir

$$\Xi_{p,q}^s \doteq \left\{ \{X_j\} : X_j \text{ variable aleatoria } \forall j \wedge \left\{ \left[E_X(|X_j|^p) \right]^{\frac{1}{p}} \right\}_j \in \Theta_{p,q}^s \right\}.$$

Es conveniente concebir a $\Xi_{p,q}^s$ como la versión bayesiana de $\Theta_{p,q}^s$.

Bajo estos supuestos, y con un modelo observacional con ruido aditivo gaussiano, la relación (4.40) del caso determinístico que relaciona los coeficientes α_λ y α_λ^δ de la solución exacta f y de la solución aproximada f_c^δ obtenida mediante el método RWV, respectivamente, resultan ahora en una idéntica relación sobre sus contrapartes estocásticas A_λ y A_λ^δ :

$$A_\lambda^\delta \doteq A_\lambda + \delta \sigma_\lambda Z_\lambda, \quad \text{con } Z_\lambda \text{ i.i.d. } \sim N(0, 1), \quad \forall \lambda \in \Lambda. \quad (4.45)$$

Notar que los métodos RWV y RWVU siguen siendo aplicables en este contexto bayesiano. Más aún, a los efectos de mantener la notación lo más simple posible, seguiremos denotando con \mathcal{M}_c y $\mathcal{M}_{c,\gamma}$ al conjunto de aproximaciones obtenidas mediante estos dos métodos, respectivamente.

Con el objeto de analizar la performance de los métodos RWV y RWVU en este contexto bayesiano será necesario definir varias medidas de error.

Definición 4.27 (Riesgo de Bayes). Dados $\delta > 0$, una sucesión $\{A_\lambda\}_{\lambda \in \Lambda}$ de variables aleatorias con distribución conjunta Π_A , y el modelo estocástico $A_\lambda^\delta = A_\lambda + \delta Z_\lambda$, con Z_λ i.i.d. $\sim N(0, 1)$ $\forall \lambda \in \Lambda$, definimos el *riesgo de Bayes* (asociado a Π_A y δ) como

$$B(\Pi_A, \delta) \doteq \sum_{\lambda \in \Lambda} E_{\Pi_A} \left[\left(E(A_\lambda | \{A_\sigma^\delta\}_{\sigma \in \Lambda}) - A_\lambda \right)^2 \right]. \quad (4.46)$$

Tal como en el caso determinístico, en este enfoque bayesiano también se diferencian los modelos continuos y discretos. Por este motivo denotaremos con B_D al riesgo de Bayes asociado al modelo discreto.

A continuación introduciremos además el concepto de riesgo de Bayes maximal. Para ello necesitaremos definir previamente el concepto de “distribución menos favorable”. Comenzamos denotando con \mathcal{D} al conjunto de las distribuciones de probabilidad de vectores aleatorios ∞ -dimensionales. Dados $\delta > 0$ y Π_1, Π_2 en \mathcal{D} , decimos que Π_1 es menos favorable que Π_2 a nivel δ (o que Π_2 precede a Π_1 a nivel δ) si

$$B(\Pi_1, \delta) \geq B(\Pi_2, \delta). \quad (4.47)$$

Notar que (4.47) introduce un orden total en \mathcal{D} .

Dados $p > 0$ y una sucesión de números reales positivos $\{\tau_j\}_{j \in \Lambda}$, denotamos con $\mathcal{V}(p, \{\tau_j\})$ al conjunto de las distribuciones en \mathcal{D} tales que su j -ésimo p -momento absoluto está acotado por τ_j , $\forall j$, es decir:

$$\mathcal{V}(p, \{\tau_j\}) \doteq \left\{ \Pi_X \in \mathcal{D} : E_{\Pi_X} (|X_j|^p) \leq \tau_j^p \right\}.$$

A los efectos de mantener la notación lo más simple posible, denotaremos con $\mathcal{V}(p, \Theta_{p,q}^s)$ al conjunto

$$\mathcal{V}(p, \Theta_{p,q}^s) \doteq \left\{ \Pi_X \in \mathcal{D} : \left\{ E_{\Pi_X} (|X_j|^p) \right\}_j \in \Theta_{p,q}^s \right\}.$$

Estamos ahora en condiciones de definir el riesgo de Bayes maximal.

Definición 4.28 (Riesgo de Bayes maximal). Dados $s \in \mathbb{N}$, $0 < p, q \leq \infty$ y $\delta > 0$, decimos que $\Pi^* = \Pi^*(p, q, s, \delta)$ es una *distribución de riesgo de Bayes maximal*, para tales parámetros, si

$$\Pi^* \in \underset{\substack{\Pi \in \mathcal{V}(p, \{\tau_j\}) \\ \{\tau_j\} \in \Theta_{p,q}^s}}{\operatorname{argmax}} B(\Pi, \delta) = \underset{\Pi \in \mathcal{V}(p, \Theta_{p,q}^s)}{\operatorname{argmax}} B(\Pi, \delta). \quad (4.48)$$

El concepto de RMC (4.38), y en particular los de RMC-WV (4.41) y RMC-WVU, tienen también su correlato en este contexto bayesiano. Definimos a continuación la versión bayesiana del RMC-WV.

Definición 4.29 (Riesgo minimax cuadrático bayesiano wavelet-vaguelet). Sea $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ el operador asociado al problema de aproximar f en (4.35). Suponemos que el modelo observacional es una realización (g^δ) de una variable aleatoria estocástica con ruido aditivo $G^\delta = G + \delta dW = TF + \delta dW$ con $\delta > 0$ y que los coeficientes wavelet $\{A_\lambda\}_{\lambda \in \Lambda}$ de la solución exacta F son variables aleatorias que satisfacen $\{A_\lambda\}_{\lambda \in \Lambda} \in \Xi_{p,q}^s$. El *riesgo minimax cuadrático bayesiano wavelet-vaguelet* (RMCB-WV) asociado a $\Xi_{p,q}^s$ y al conjunto admisible \mathcal{M}_c se define como

$$\mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) \doteq \inf_{\{A_\lambda^\delta\}: F_c^\delta \in \mathcal{M}_c} \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} E \left(\|A_\lambda^\delta - A_\lambda\|_{\ell^2(\Lambda)}^2 \right), \quad (4.49)$$

donde F_c^δ denota la solución aproximada del problema $TF = G$ obtenida a partir del método RWV.

Tal como en el caso determinístico, en este enfoque bayesiano también se diferencian los modelos continuos y discretos. Por este motivo denotaremos a los riesgos RMCB-WV y RMCB-WVU asociados al modelo discreto con $\mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c)$ y $\mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_{c,\gamma})$, respectivamente. Notemos que, por definición de RMC-WV y RMCB-WV, se sigue inmediatamente que

$$\mathcal{R}(\delta, \Theta_{p,q}^s, \mathcal{M}_c) \leq \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c). \quad (4.50)$$

De igual modo

$$\mathcal{R}(\delta, \Theta_{p,q}^s, \mathcal{M}_{c,\gamma}) \leq \mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_{c,\gamma}). \quad (4.51)$$

Finalizamos esta sección introduciendo un concepto que resultará muy útil en el caso que la distribución conjunta Π_A sea tal que haga que las variables aleatorias A_λ sean independientes. Para ello, dados $0 < p \leq \infty$ y $\tau > 0$ denotamos con $\mathcal{U}(p, \tau)$ al conjunto de variables aleatorias con p -momento absoluto acotado por τ , es decir:

$$\mathcal{U}(p, \tau) \doteq \{Y \text{ variable aleatoria} : E_Y(|Y|^p) \leq \tau^p\}.$$

Definición 4.30 (Riesgo escalar de Bayes). Sean $\tau, \delta \in \mathbb{R}^+$, $p \in \mathbb{N}$, X una variable aleatoria y $X^\delta \doteq X + \delta Z$ con $Z \sim N(0, 1)$. Se define el *riesgo escalar de Bayes*, asociado a p , τ y δ , como

$$\rho(p, \tau, \delta) \doteq \sup_{X \in \mathcal{U}(p, \tau)} E_X \text{Var}(X|X^\delta). \quad (4.52)$$

4.5.3. Lemas previos

En esta sección enunciamos y demostramos una serie de resultados que serán fundamentales a la hora de probar la existencia de cotas para los órdenes de convergencia de los métodos RWV y RWVU. Demostramos en primer lugar una propiedad vinculada al riesgo scalar de Bayes e.

Lema 4.31. *Dados $\tau, \delta \in \mathbb{R}^+$ y $p \in \mathbb{N}$, el riesgo escalar de Bayes ρ satisface:*

$$\rho(p, \tau, \delta) = a^{-2} \rho(p, a\tau, a\delta), \quad \forall a > 0. \quad (4.53)$$

Demostración. Para demostrar este lema utilizaremos la siguiente propiedad (bastante intuitiva) de la esperanza condicional (ver Apéndice, Lema 5.10). Dadas X e Y variables aleatorias y $a \neq 0$ se tiene que

$$E(X|aY) = E(X|Y). \quad (4.54)$$

En primer lugar notar que por definición de ρ (4.52),

$$\rho(p, a\tau, a\delta) = \sup_{X \in \mathcal{U}(p, a\tau)} E_X \text{Var}(X|X^{a\delta}) = \sup_{X \in \mathcal{U}(p, a\tau)} E_X \text{Var}(X|X + a\delta Z).$$

Luego, para probar (4.53), bastaría con probar que $\forall a > 0$ y $\forall X \in \mathcal{U}(p, \tau)$ existe $\tilde{X} \in \mathcal{U}(p, a\tau)$ que verifica $E_X \text{Var}(X|X + \delta Z) = a^{-2} E_{\tilde{X}} \text{Var}(\tilde{X}|\tilde{X} + a\delta Z)$. Veamos que para $\tilde{X} \doteq aX$ se satisfacen ambas condiciones. En primer lugar, notar que:

$$\begin{aligned} E_X \text{Var}(X|X + \delta Z) &= E_X \left(E(X^2|X + \delta Z) - [E(X|X + \delta Z)]^2 \right) \\ &= E_X \left(a^{-2} E(a^2 X^2|a(X + \delta Z)) - [a^{-1} E(aX|a(X + \delta Z))]^2 \right) \quad (\text{por (4.54)}) \\ &= a^{-2} E_{\tilde{X}} \left(E(\tilde{X}^2|\tilde{X} + a\delta Z) - [E(\tilde{X}|\tilde{X} + a\delta Z)]^2 \right) \quad (\tilde{X} = aX) \\ &= a^{-2} E_{\tilde{X}} \text{Var}(\tilde{X}|\tilde{X} + a\delta Z). \end{aligned}$$

Por otro lado, dado que $X \in \mathcal{U}(p, \tau)$ tenemos que $E_X(|X|^p) \leq \tau^p$ lo cual implica $E_{\tilde{X}}(|\tilde{X}|^p) = E_X(|aX|^p) \leq a^p \tau^p$ y por lo tanto $\tilde{X} \in \mathcal{U}(p, a\tau)$. ■

A continuación, demostramos una importante propiedad del riesgo de Bayes relacionada a la independencia de las distribuciones involucradas. Para ello denotamos con Λ_j al subconjunto de índices de Λ correspondientes a j fijo, es decir $\Lambda_j \doteq \{(j, k, \epsilon) \in \Lambda : k \in \mathbb{Z} \wedge \epsilon \in \{0, 1\}^d\}$.

Lema 4.32. *Dada una sucesión de variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$ con distribución conjunta μ existe una sucesión de variables aleatorias $\{\bar{A}_\lambda\}_{\lambda \in \Lambda}$ con distribución conjunta $\bar{\mu}$ tal que:*

1. $\bar{\mu}$ es menos favorable que μ a nivel δ , $\forall \delta > 0$, y
2. Las variables aleatorias $\{\bar{A}_\lambda\}_{\lambda \in \Lambda}$ son independientes (ó equivalentemente $\bar{\mu}$ factoriza) y además, para cada j fijo, $j \geq j_0$, las variables aleatorias $\{\bar{A}_\lambda\}_{\lambda \in \Lambda_j}$ están idénticamente distribuidas.

Demostración. Dada la distribución conjunta μ de la sucesión de variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$, para $\lambda \in \Lambda$, denotaremos con μ_λ a la correspondiente distribución λ -marginal de μ . Asimismo, para $\lambda = (j, k, \epsilon) \in \Lambda$ denotaremos con $j(\lambda)$ a $j(\lambda) = \Pi_1(\lambda) = j$ donde $\Pi_1 : \Lambda \rightarrow \mathbb{Z}$ es la proyección sobre la primer componente. A continuación para cada $\lambda \in \Lambda$ definimos $\bar{\mu}_\lambda \doteq \sum_{\sigma \in \Lambda_{j(\lambda)}} \frac{\mu_\sigma}{\#\Lambda_{j(\lambda)}}$ (observar que puesto que Ω es acotado, $\#\Lambda_{j(\lambda)} < \infty$ para todo $\lambda \in \Lambda$). Notar que si $\lambda_1, \lambda_2 \in \Lambda$ y $\Pi_1(\lambda_1) = \Pi_1(\lambda_2)$ entonces $\bar{\mu}_{\lambda_1} = \bar{\mu}_{\lambda_2}$. Por lo tanto, puesto que $\Pi_1(\Lambda_j) = j$, se tiene que $\bar{\mu}_\lambda$ es la misma distribución $\forall \lambda \in \Lambda_j$. Denotaremos a esta distribución común con $\bar{\mu}_{\lambda(j)}$. Finalmente definimos $\bar{\mu} \doteq \bigotimes_{\lambda \in \Lambda} \bar{\mu}_\lambda$ y denotaremos con \bar{A}_λ a una variable aleatoria con distribución $\bar{\mu}_\lambda$. Se sigue inmediatamente de lo anterior que las variables aleatorias $\{\bar{A}_\lambda\}_{\lambda \in \Lambda}$ son independientes y que para todo j fijo, $j \geq j_0$, las variables aleatorias $\{\bar{A}_\lambda\}_{\lambda \in \Lambda_j}$ están idénticamente distribuidas.

Resta probar que $\bar{\mu}$ es menos favorable que μ a nivel δ , $\forall \delta > 0$. Para ello comenzamos definiendo, para $\lambda \in \Lambda$, $b_\delta(\mu_\lambda) \doteq E_{\mu_\lambda} \left[(E(A_\lambda | A_\lambda^\delta) - A_\lambda)^2 \right]$. Observar entonces que:

$$\begin{aligned}
B(\bar{\mu}, \delta) &= \sum_{\lambda \in \Lambda} E_{\bar{\mu}} \left[(E(\bar{A}_\lambda | \{\bar{A}_\sigma^\delta\}_{\sigma \in \Lambda}) - \bar{A}_\lambda)^2 \right] \\
&= \sum_{\lambda \in \Lambda} E_{\bar{\mu}_\lambda} \left[(E(\bar{A}_\lambda | \bar{A}_\lambda^\delta) - \bar{A}_\lambda)^2 \right] \quad (\text{las } \bar{A}_\lambda \text{ son indep.} \Rightarrow \text{las } \bar{A}_\lambda^\delta \text{ son indep.}) \\
&= \sum_{\lambda \in \Lambda} b_\delta(\bar{\mu}_\lambda) \\
&= \sum_{j \geq j_0} \sum_{\lambda \in \Lambda_j} b_\delta(\bar{\mu}_\lambda) \\
&= \sum_{j \geq j_0} \#\Lambda_j b_\delta(\bar{\mu}_{\lambda(j)}). \quad (\text{pues } \bar{\mu}_\lambda = \bar{\mu}_{\lambda(j)} \forall \lambda \in \Lambda_j) \quad (4.55)
\end{aligned}$$

Por otro lado, observemos que puesto que el error cuadrático de la estimación de la variable aleatoria A_λ dada la sucesión $\{A_\sigma^\delta\}_{\sigma \in \Lambda}$ es menor o igual que el mismo cuando solo se conoce A_λ^δ , se sigue que $\forall \lambda \in \Lambda, \forall \delta > 0$:

$$E_\mu \left[(E(A_\lambda | \{A_\sigma^\delta\}_{\sigma \in \Lambda}) - A_\lambda)^2 \right] \leq E_\mu \left[(E(A_\lambda | A_\lambda^\delta) - A_\lambda)^2 \right]$$

$$\begin{aligned}
&= E_{\mu_\lambda} \left[\left(E(A_\lambda | A_\lambda^\delta) - A_\lambda \right)^2 \right] \\
&= b_\delta(\mu_\lambda),
\end{aligned}$$

lo cual implica que

$$B(\mu, \delta) = \sum_{\lambda \in \Lambda} E_\mu \left[\left(E(A_\lambda | \{A_\sigma^\delta\}_{\sigma \in \Lambda}) - A_\lambda \right)^2 \right] \leq \sum_{\lambda \in \Lambda} b_\delta(\mu_\lambda). \quad (4.56)$$

Por otro lado, puesto que $\forall \delta > 0$ la función b_δ es cóncava, para cada j fijo, $j \geq j_0$, se tiene que

$$\sum_{\lambda \in \Lambda_j} \frac{b_\delta(\mu_\lambda)}{\#\Lambda_j} \leq b_\delta \left(\sum_{\lambda \in \Lambda_j} \frac{\mu_\lambda}{\#\Lambda_j} \right) = b_\delta(\bar{\mu}_{\lambda(j)}). \quad (4.57)$$

Finalmente,

$$\begin{aligned}
B(\mu, \delta) &\leq \sum_{\lambda \in \Lambda} b_\delta(\mu_\lambda) && \text{(por (4.56))} \\
&= \sum_{j \geq j_0} \sum_{\lambda \in \Lambda_j} b_\delta(\mu_\lambda) \\
&= \sum_{j \geq j_0} \#\Lambda_j \left(\sum_{\lambda \in \Lambda_j} \frac{b_\delta(\mu_\lambda)}{\#\Lambda_j} \right) && (\#\Lambda_j < \infty) \\
&\leq \sum_{j \geq j_0} \#\Lambda_j b_\delta(\bar{\mu}_{\lambda(j)}) && \text{(por (4.57))} \\
&= B(\bar{\mu}, \delta), && \text{(por (4.55))}
\end{aligned}$$

de lo cual se sigue que $\bar{\mu}$ es menos favorable que μ a nivel δ , como queríamos probar. \blacksquare

En los siguientes dos lemas, demostraremos propiedades relacionadas a una variable aleatoria definida por el producto interno entre ruido blanco y una función en $L^2(\Omega)$. Estos resultados serán de utilidad a la hora de estudiar la convergencia de los métodos de regularización propuestos en la sección anterior (ver definición de W_λ en (4.40)).

Lema 4.33. Sean (S, \mathcal{A}, P) un espacio de probabilidad y $\Omega \subset \mathbb{R}^d$ acotado. Sean además $f \in L^2(\Omega)$ y $dW : S \times \Omega \rightarrow \mathbb{R}$ ruido blanco gaussiano (ver Definición 1.32). Entonces $\text{Var}(\langle dW, f \rangle) = \|f\|_{L^2(\Omega)}^2$.

Demostración. Por simplicidad realizaremos la prueba solo para el caso $d = 1$. El caso general, para $d \in \mathbb{N}$, se sigue de manera similar. Supondremos entonces que Ω es un intervalo $I = [a, b] \subset \mathbb{R}$. Consideramos primero el caso en que f es una función simple en $L^2(\Omega)$, es decir f es combinación lineal de funciones características. Sea entonces $f(x) = \sum_{i=1}^n \alpha_i \chi_{[x_i, x_{i+1})}(x)$

donde $A_n \doteq \{a = x_1, x_2, \dots, x_n = b\}$ es una partición de I y $\alpha_1, \dots, \alpha_n$ son coeficientes reales.

Entonces, $\forall s \in S$

$$\begin{aligned} \langle dW, f \rangle(s) &= \int_I f(x) dW(x, s) dx \\ &= \int_I \sum_{i=1}^n \alpha_i \chi_{[x_i, x_{i+1})}(x) dW(x, s) dx \\ &= \sum_{i=1}^n \alpha_i \int_{x_i}^{x_{i+1}} dW(x, s) dx \\ &= \sum_{i=1}^n \alpha_i (W(x_{i+1}, s) - W(x_i, s)). \end{aligned}$$

Tomando varianza en la identidad anterior y utilizando propiedades básicas del movimiento browniano, obtenemos:

$$\begin{aligned} \text{Var}(\langle dW, f \rangle(s)) &= \text{Var} \left(\sum_{i=1}^n \alpha_i (W(x_{i+1}, s) - W(x_i, s)) \right) \\ &= \sum_{i=1}^n \alpha_i^2 \text{Var}(W(x_{i+1}, s) - W(x_i, s)) \quad (\text{por (1.27)}) \\ &= \sum_{i=1}^n \alpha_i^2 |x_{i+1} - x_i| \quad (\text{por (1.28)}) \\ &= \|f\|_{L^2(\Omega)}^2. \end{aligned}$$

La demostración para el caso general de $f \in L^2(\Omega)$, se puede realizar utilizando un argumento de densidad de las funciones simples en $L^2(\Omega)$ y usando la hipótesis de incrementos independientes del movimiento browniano, lo cual permite justificar rigurosamente cada una de las siguientes igualdades:

$$\begin{aligned} \|f\|_{L^2(\Omega)}^2 &= \int_{\Omega} f^2(x) \text{Var} \left(\frac{dW(x, \cdot)}{dx} \right) dx = \int_{\Omega} \text{Var} \left(f(x) \frac{dW(x, \cdot)}{dx} \right) dx \\ &\stackrel{(\star)}{=} \text{Var} \left(\int_{\Omega} f(x) \frac{dW(x, \cdot)}{dx} dx \right) = \text{Var}(\langle f, dW \rangle). \end{aligned}$$

En particular, señalamos que la igualdad (\star) es consecuencia de la hipótesis de incrementos independientes del movimiento browniano. ■

Definición 4.34. Dado Δ un conjunto numerable, diremos que una sucesión $\{W_i\}_{i \in \Delta}$ de variables aleatorias de media cero es (ó constituye) *ruido casi-independiente* (RCI) si existen

constantes positivas m y M tales que, para todo $i \in \Delta$,

$$m \leq E \operatorname{Var}(W_i | \{W_k\}_{k \neq i}) \leq \operatorname{Var}(W_i) \leq M. \quad (4.58)$$

Observación 4.35. En el caso particular en que las W_i tengan distribución normal, nos referiremos a $\{W_i\}_{i \in \Delta}$ como *ruido gaussiano casi-independiente* (RGCI). Notar que si las w_i son independientes e idénticamente distribuidas con varianza común σ^2 , entonces (4.58) vale con igualdades y $m = M = \sigma^2$.

El siguiente teorema provee condiciones suficientes para que una sucesión de variables aleatorias normales, generadas a partir del producto interno de ruido blanco con una familia de funciones en $L^2(\Omega)$, constituya RGCI.

Teorema 4.36. Sean Δ un conjunto numerable y $\{W_i\}_{i \in \Delta}$ una sucesión de variables aleatorias definidas por $W_i = \langle u_i, dW \rangle$ para todo $i \in \Delta$, donde $\{u_i\}_{i \in \Delta}$ es una sucesión de funciones en $L^2(\Omega)$ y dW es ruido blanco. Supongamos que existen constantes positivas C_1 y C_2 tales que

$$C_1 \|a_i\|_{\ell^2(\Delta)} \leq \left\| \sum_{i \in \Delta} a_i u_i \right\|_{L^2(\Omega)} \leq C_2 \|a_i\|_{\ell^2(\Delta)} \quad \forall \{a_i\} \in \ell^2(\Delta). \quad (4.59)$$

Entonces $\{W_i\}_{i \in \Delta}$ es RGCI, más precisamente

$$C_1^2 \stackrel{(I)}{\leq} E \operatorname{Var}(W_i | \{W_k\}_{k \neq i}) \stackrel{(II)}{\leq} \operatorname{Var}(W_i) \stackrel{(III)}{\leq} C_2^2 \quad \forall i \in \Delta. \quad (4.60)$$

Demostración. Para probar (4.60) demostraremos cada una de las desigualdades (I), (II) y (III) individualmente. Notemos en primer lugar que la desigualdad (II), o sea $E \operatorname{Var}(W_i | (W_h)_{h \neq i}) \leq \operatorname{Var}(W_i)$, resulta trivial por definición de varianzas condicionales.

Para la desigualdad (III) notemos que dada u_i en $L^2(\Omega)$, aplicando el Lema 4.33 se tiene que

$$\operatorname{Var}(W_i) = \operatorname{Var}(\langle u_i, dW \rangle) = \|u_i\|_{L^2(\Omega)}^2. \quad (4.61)$$

Por otro lado, para cada $j \in \Delta$, observemos que eligiendo $a_i = \delta_{ij} \forall i \in \Delta$ en (4.59), se obtiene

$$C_1 \leq \|u_j\|_{L^2(\Omega)} \leq C_2. \quad (4.62)$$

Luego, de (4.61) y (4.62) se sigue entonces que $\operatorname{Var}(W_i) \leq C_2^2, \forall i \in \Delta$.

Por último, la desigualdad (I) se prueba utilizando la siguiente propiedad de varianzas condicionales para variables aleatorias normales de media nula (ver apéndice, Lema 5.11)

$$E \operatorname{Var}(W_i | \{W_k\}_{k \neq i}) = \inf_{a_k \in \mathbb{R}^+} \operatorname{Var} \left(W_i - \sum_{k \neq i} a_k W_k \right). \quad (4.63)$$

En efecto,

$$\begin{aligned}
E \operatorname{Var}(W_i | \{W_k\}_{k \neq i}) &= \inf_{a_k \in \mathbb{R}^+} \left\{ \operatorname{Var} \left(\sum_{k \in \Delta} a_k W_k \right) : a_i = 1 \right\} && \text{(por (4.63))} \\
&= \inf_{a_k \in \mathbb{R}^+} \left\{ \left\| \sum_{k \in \Delta} a_k u_k \right\|_{L^2(\Omega)}^2 : a_i = 1 \right\} && \text{(por (4.61))} \\
&\geq \inf_{a_k \in \mathbb{R}^+} \left\{ C_1^2 \sum_{k \in \Delta} a_k^2 : a_i = 1 \right\} && \text{(por (4.59))} \\
&= C_1^2. && \blacksquare
\end{aligned}$$

A continuación enunciamos dos proposiciones que serán necesarias para la demostración del Teorema 4.39, que relaciona los RMC-B continuo y discreto, \mathcal{B} y \mathcal{B}_D (ver Definición 4.29).

Proposición 4.37. *Dada una sucesión de variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$ con distribución conjunta Π , existe una sucesión de variables aleatorias independientes $\{\bar{A}_\lambda\}_{\lambda \in \Lambda}$ con distribución conjunta $\bar{\Pi}$ tal que*

1. $\bar{\Pi}$ posee las mismas distribuciones marginales que Π , y
2. $\bar{\Pi}$ es menos favorable que Π a nivel δ , $\forall \delta > 0$.

Demostración. La demostración se obtiene esencialmente siguiendo los mismos pasos que en la demostración del Lema 4.32, por lo que omitimos detalles de la misma. \blacksquare

Proposición 4.38. *Sean $0 < p, q \leq \infty$, $s \in \mathbb{N}$ y $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ el operador asociado al problema de aproximar f en $Tf = g$. Supongamos que el modelo observacional es una realización (g^δ) de una variable aleatoria estocástica con ruido aditivo, $G^\delta = G + \delta dW = TF + \delta dW$ con $\delta > 0$ y que los coeficientes wavelet $\{A_\lambda\}_{\lambda \in \Lambda}$ de la solución exacta F son variables aleatorias con distribución conjunta Π , que satisfacen $\{A_\lambda\}_{\lambda \in \Lambda} \in \Xi_{p,q}^s$. Entonces,*

$$\mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) = \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B(\Pi, \delta), \quad y \quad (4.64a)$$

$$\mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c) = \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B_D(\Pi, \delta), \quad (4.64b)$$

donde B y B_D (ver Definición 4.27) son los correspondientes riesgos de Bayes asociados a los modelos continuo y discreto, respectivamente.

Demostración. La demostración utiliza el Teorema Minimax de la Teoría de Decisión Estadística, detalles de la misma pueden encontrarse, por ejemplo, en [6], Teorema 1, Capítulo 2. \blacksquare

Estamos ahora en condiciones de presentar el siguiente teorema, que provee condiciones suficientes bajo las cuales los RMC-B continuo y discreto (\mathcal{B} y \mathcal{B}_D) poseen el mismo orden de convergencia.

Teorema 4.39. Sean $\Omega \subset \mathbb{R}^d$ acotado, $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\Omega)$ una base wavelet y $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador homogéneo de grado α , que posee una descomposición wavelet vaguelet $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ con respecto a la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$, con valores casi-singulares $\kappa_\lambda = 2^{-\alpha j(\lambda)} \forall j(\lambda) \geq j_0$ y supongamos que $f \in B_{p,q}^s(\Omega)$ para algún $s \in \mathbb{N}$ y $0 < p, q \leq \infty$. Sea además $W_\lambda \doteq \langle u_\lambda, dW \rangle$, donde dW es ruido blanco. Si $\{W_\lambda\}_{\lambda \in \Lambda}$ es RGCI entonces, $\forall \delta > 0$, los RMC-B continuo y discreto \mathcal{B} y \mathcal{B}_D , asociados a $g = Tf$ y $g^\delta \doteq g + \delta dW$ satisfacen

$$\mathcal{B}_D(\sqrt{m} \delta, \Xi_{p,q}^s, \mathcal{M}_c) \leq \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) \leq \mathcal{B}_D(\sqrt{M} \delta, \Xi_{p,q}^s, \mathcal{M}_c), \quad (4.65)$$

donde m y M son las constantes correspondientes al RGCI en (4.58).

Demostración. La demostración rigurosa de este teorema se sustenta en varios resultados previos y es sumamente técnica y extensa, razón por la cual nos limitaremos a esbozar los pasos principales que conducen a la misma. Si bien algunos de estos resultados previos son bastante conocidos, todos ellos son referenciados apropiadamente.

En primer lugar, se puede probar que las hipótesis del teorema implican la siguiente cadena de desigualdades:

$$B_D(\bar{\Pi}, \sqrt{m} \delta) \stackrel{(I)}{\leq} B(\Pi, \delta) \stackrel{(II)}{\leq} B_D(\bar{\Pi}, \sqrt{M} \delta). \quad (4.66)$$

En el caso la desigualdad (II), por un lado, el Lema 9 en [14] afirma que para toda cadena de Markov (Y_0, Y_1, Y_2) se verifica que $E \text{Var}(Y_2|Y_1) \leq E \text{Var}(Y_2|Y_0)$. Por otro lado, dado (Y_2, Y_1, Y_0) un vector de procesos estocásticos con incrementos independientes, el Lema 10 en [14] provee hipótesis bajo las cuales el vector (Y_0, Y_1, Y_2) es una cadena de Markov. La desigualdad (II) se sigue entonces definiendo $Y_2 \doteq A_\lambda$, $Y_1 \doteq A_\lambda^\delta$ y $Y_0 \doteq \bar{A}_\lambda^\delta$, donde $\{\bar{A}_\lambda^\delta\}_{\lambda \in \Lambda}$ es la sucesión asociada a la distribución conjunta $\bar{\mu}$, obtenida a partir de la distribución conjunta μ de la sucesión $\{A_\lambda^\delta\}$, como en la Proposición 4.37.

La desigualdad (I) se puede probar utilizando la Proposición 4.37 y un resultado (ver Lema 11 en [14]) que afirma que dada una sucesión $\{Z_\lambda\}_{\lambda \in \Lambda}$ de variables aleatorias normales con media cero, para todo $\lambda \in \Lambda$ existe una variable aleatoria X_λ normal e independiente de $\{Z_\sigma\}_{\sigma \neq \lambda}$ tal que

$$Z_\lambda = E(Z_\lambda | \{Z_\sigma\}_{\sigma \neq \lambda}) + X_\lambda.$$

Además, $\tilde{Z}_\lambda \doteq E(Z_\lambda | \{Z_\sigma\}_{\sigma \neq \lambda})$ es también una variable aleatoria normal que puede escribirse como una combinación lineal de $\{Z_\sigma\}_{\sigma \neq \lambda}$. La desigualdad (I) se sigue entonces definiendo $Y_2 \doteq A_\lambda$, $Y_1 \doteq A_\lambda + \delta\sigma_j X_\lambda$ y $Y_0 \doteq A_\lambda^\delta = A_\lambda + \delta\sigma_j Z_\lambda = A_\lambda + \delta\sigma_j(\tilde{Z}_\lambda + X_\lambda)$ y utilizando nuevamente los Lemas 9 y 10 en [14].

El resto de la demostración consiste en probar que (4.66) implica (4.65). Para ello observemos por un lado que

$$\begin{aligned} \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) &= \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B(\Pi, \delta) && \text{(por (4.64a))} \\ &\geq \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B(\bar{\Pi}, \delta) \\ &\geq \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B_D(\bar{\Pi}, \sqrt{m}\delta) && \text{(por (I))} \\ &= \mathcal{B}_D(\sqrt{m}\delta, \Xi_{p,q}^s, \mathcal{M}_c), && \text{(por (4.64b))} \end{aligned}$$

donde la primera desigualdad se sigue trivialmente del hecho que el conjunto de las distribuciones conjuntas asociadas a una sucesión de variables aleatorias independientes está contenido en el conjunto de distribuciones conjuntas sin esta restricción. Esto prueba la primera desigualdad en (4.65).

Para la segunda desigualdad en (4.65), notar que

$$\begin{aligned} \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) &= \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B(\Pi, \delta) && \text{(por (4.64a))} \\ &\leq \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B_D(\bar{\Pi}, \sqrt{M}\delta) && \text{(por (II))} \\ &= \mathcal{B}_D(\sqrt{M}\delta, \Xi_{p,q}^s, \mathcal{M}_c). && \text{(por (4.64b))} \end{aligned}$$

Esto concluye la demostración del teorema. ■

A los efectos de poder finalmente abordar la demostración del resultado principal de este capítulo introduciremos previamente tres definiciones y enunciaremos y demostraremos dos lemas que simplificarán el desarrollo de la demostración de ese teorema principal (Teorema 4.45).

Definición 4.40. Dadas constantes $\alpha \in \mathbb{R}$, p, q, δ y $C \in \mathbb{R}^+$ y $s \in \mathbb{N}$, definimos

$$P_{p,q,s,\alpha}(\delta, C) \doteq \sup_{\{\tau_j\}^+ : \sum_{j>j_0} 2^{jq(s+d/2)} \tau_j^q \leq C^q} \sum_{j>j_0} 2^{jd} \rho(p, \tau_j, 2^{\alpha j} \delta), \quad \text{y} \quad (4.67)$$

$$Q_{p,q,s,\alpha}(\delta, C) \doteq \sup_{\{\tau_j\}^+ : \sum_{j \in \mathbb{Z}} 2^{jq(s+d/2)} \tau_j^q \leq C^q} \sum_{j \in \mathbb{Z}} 2^{jd} \rho(p, \tau_j, 2^{\alpha j} \delta), \quad (4.68)$$

donde $\rho(\cdot, \cdot, \cdot)$ es el riesgo escalar de Bayes (ver Definición 4.30) y $\{\tau_j\}^+$ denota una sucesión de números reales no negativos.

Definición 4.41. Sean p, q, s, α, δ y ρ como en la Definición 4.40. Definimos dos funcionales sobre ℓ^2 , $J_{p,\alpha,\delta}$ y $G_{s,q,\alpha}$, de la siguiente manera

$$J_{p,\alpha,\delta}(\{x_j\}) \doteq \delta^2 \sum_{j \in \mathbb{Z}} 2^{j(d+2\alpha)} \rho\left(p, \frac{x_j}{\delta}, 1\right),$$

y

$$G_{s,q,\alpha}(\{x_j\}) \doteq \sum_{j \in \mathbb{Z}} 2^{jq(s+d/2+\alpha)} x_j^q.$$

Definición 4.42. Dadas $a \in \mathbb{R}^+$ y $h \in \mathbb{Z}$ definimos el operador $\mathcal{U}_{a,h} : \ell^2 \rightarrow \ell^2$ como $\mathcal{U}_{a,h}(\{x_j\}) \doteq \{ax_{j-h}\}$.

Las siguientes identidades se siguen a partir de las Definiciones 4.40, 4.41 y 4.42:

$$\begin{aligned} \text{ID1)} \quad & J_{p,\alpha,\delta}(\mathcal{U}_{\delta,h}(\{x_j\})) = \delta^2 2^{h(d+2\alpha)} J_{p,\alpha,1}(\{x_j\}), \\ \text{ID2)} \quad & G_{s,q,\alpha}(\mathcal{U}_{\delta,h}(\{x_j\})) = \delta^q 2^{qh(s+d/2+\alpha)} G_{s,q,\alpha}(\{x_j\}), \\ \text{ID3)} \quad & Q_{p,q,s,\alpha}(\delta, C) = \sup_{\{v_j\}^+ : G_{s,q,\alpha}(\{v_j\}) \leq C^q} J_{p,\alpha,\delta}(\{v_j\}). \end{aligned}$$

Las demostraciones de estas identidades son inmediatas. En particular ID3) se sigue definiendo $v_j \doteq \frac{\tau_j}{2^{j\alpha}} \forall j \in \mathbb{Z}$.

El siguiente lema establece el orden de $Q_{p,q,s,\alpha}(\delta, C)$ para $\delta \rightarrow 0^+$.

Lema 4.43. Sean α, p, q, s, δ como antes y $C > 0$. Entonces

$$Q_{p,q,s,\alpha}(\delta, C) \equiv \delta^{\frac{4s}{2s+d+2\alpha}} \text{ para } \delta \rightarrow 0^+. \quad (4.69)$$

Demostración. Observar que

$$\begin{aligned} Q_{p,q,s,\alpha}(\delta, C) &= \sup_{\{u_j\}^+ : G_{s,q,\alpha}(\{u_j\}) \leq C^q} J_{p,\alpha,\delta}(\{u_j\}) && \text{(por ID3)} \\ &= \sup_{\mathcal{U}_{\delta,h}(\{v_j\}^+) : G_{s,q,\alpha}(\mathcal{U}_{\delta,h}(\{v_j\})) \leq C^q} J_{p,\alpha,\delta}(\mathcal{U}_{\delta,h}(\{v_j\})) \\ &= \sup_{\{v_j\}^+ : G_{s,q,\alpha}(\{v_j\}) \leq (C\delta^{-1}2^{-h(s+d/2+\alpha)})^q} \delta^2 2^{h(d+2\alpha)} J_{p,\alpha,1}(\{v_j\}) && \text{(ID1) e ID2)} \\ &= \delta^2 2^{h(d+2\alpha)} Q_{p,q,s,\alpha}(1, C^*), && \text{(por (4.68))} \end{aligned} \quad (4.70)$$

donde $C^* \doteq C\delta^{-1}2^{-h(s+d/2+\alpha)}$.

En particular para $\delta = \delta_h \doteq 2^{-h(s+d/2+\alpha)}$, con $h \in \mathbb{Z}$, se tiene que $C^* = C$ y

$$Q_{p,q,s,\alpha}(\delta_h, C) = \delta_h^{\frac{4s}{2s+d+2\alpha}} Q_{p,q,s,\alpha}(1, C) = O\left(\delta_h^{\frac{4s}{2s+d+2\alpha}}\right) \text{ para } \delta_h \rightarrow 0^+. \quad (4.71)$$

Ahora, para todo $\delta > 0$ existe $h \in \mathbb{Z}$ tal que $\delta_{h+1} < \delta \leq \delta_h$ y por lo tanto, de (4.70) se sigue que

$$Q_{p,q,s,\alpha}(\delta_{h+1}, C) \leq Q_{p,q,s,\alpha}(\delta, C) \leq Q_{p,q,s,\alpha}(\delta_h, C). \quad (4.72)$$

Finalmente, (4.69) se sigue inmediatamente de (4.71) y (4.72). ■

Finalizamos esta sección enunciando y demostrando un resultado que será fundamental para encontrar órdenes de convergencia para los métodos RWV y RWVU. Este resultado prueba que $P_{p,q,s,\alpha}(\delta, C)$ y de $Q_{p,q,s,\alpha}(\delta, C)$ tienen el mismo orden para $\delta \rightarrow 0^+$.

Lema 4.44. *Sean p, q, s, α, C y δ como en la Definición 4.40. Entonces*

$$P_{p,q,s,\alpha}(\delta, C) \equiv Q_{p,q,s,\alpha}(\delta, C) \text{ para } \delta \rightarrow 0^+.$$

Demostración. En primer lugar observemos que para toda sucesión $\{\tau_j\}_{j>j_0}^+$ que satisface $\sum_{j_0 < j} 2^{jq(s+d/2)} \tau_j^q \leq C^q$, existe otra sucesión $\{\bar{\tau}_j\}_{j \in \mathbb{Z}}^+$, donde $\bar{\tau}_j \doteq \tau_j \chi_{\{j>j_0\}}(j)$ la cual satisface $\sum_{j \in \mathbb{Z}} 2^{jq(s+d/2)} \bar{\tau}_j^q \leq C^q$ y claramente

$$\sum_{j>j_0} 2^{jd} \rho(p, \tau_j, 2^{j\alpha} \delta) \leq \sum_{j \in \mathbb{Z}} 2^{jd} \rho(p, \bar{\tau}_j, 2^{j\alpha} \delta).$$

Tomando supremo se sigue entonces que

$$P_{p,q,s,\alpha}(\delta, C) \leq Q_{p,q,s,\alpha}(\delta, C). \quad (4.73)$$

Para la desigualdad opuesta, supongamos que $\{\tau_j\}_{j \in \mathbb{Z}}^+$ satisface $\sum_{j \in \mathbb{Z}} 2^{jq(s+d/2)} \tau_j^q \leq C^q$. Definiendo $\bar{\tau}_j \doteq \tau_j$ para $j > j_0$ se tiene entonces que $\sum_{j>j_0} 2^{jq(s+d/2)} \bar{\tau}_j^q \leq C^q$. Por otro lado, de la definición del riesgo escalar de Bayes (Definición 4.30) se sigue inmediatamente que $\rho(p, \tau_j, 2^{j\alpha} \delta) \leq (2^{j\alpha} \delta)^2$ para todo $j \in \mathbb{Z}$. Entonces

$$\sum_{j \leq j_0} 2^{jd} \rho(p, \tau_j, 2^{j\alpha} \delta) \leq \delta^2 \sum_{j \leq j_0} 2^{j(d+2\alpha)} = C_{j_0} \delta^2,$$

donde C_{j_0} es una constante finita que depende de j_0 , α y d .

Luego, $\forall \{\tau_j\}_{j \in \mathbb{Z}}^+$ tal que $\sum_{j \in \mathbb{Z}} 2^{jq(s+d/2)} \tau_j^q \leq C^q$, se tiene que

$$\sum_{j \in \mathbb{Z}} 2^{jd} \rho(p, \tau_j, 2^{j\alpha} \delta) \leq \sum_{j>j_0} 2^{jd} \rho(p, \tau_j, 2^{j\alpha} \delta) + C_{j_0} \delta^2.$$

Tomando supremo se sigue finalmente que

$$Q_{p,q,s,\alpha}(\delta, C) \leq P_{p,q,s,\alpha}(\delta, C) + C_{j_0} \delta^2. \quad (4.74)$$

De (4.69) y (4.74) concluimos entonces que existe una constante C' tal que

$$P_{p,q,s,\alpha}(\delta, C) \geq C' Q_{p,q,s,\alpha}(\delta, C) \text{ para } \delta \rightarrow 0^+. \quad (4.75)$$

El resultado del lema se sigue finalmente de (4.73) y (4.75). \blacksquare

4.5.4. Convergencia del método RWV

Probaremos aquí uno de los resultados principales de este capítulo, según el cuál el método de regularización wavelet-vaguelet propuesto en la Sección 4.4.3 provee un orden de convergencia óptimo (ver Teorema 2.21) para el error de estimación total.

Teorema 4.45. Sean $\Omega \subset \mathbb{R}^d$ acotado, $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\Omega)$ una base wavelet y $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador homogéneo de grado α , que posee una descomposición wavelet vaguelet $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ con respecto a la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$, con valores casi-singulares $\kappa_\lambda = 2^{-\alpha j(\lambda)} \forall j(\lambda) \geq j_0$ y supongamos que $f \in B_{p,q}^s(\Omega)$ para algún $s \in \mathbb{N}$ y $0 < p, q \leq \infty$. Entonces, para todo $\delta > 0$, el riesgo minimax cuadrático asociado a $g = Tf$ y $g^\delta \doteq g + \delta dW$ satisface

$$\mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_c) = O\left(\delta^{\frac{4s}{2s+2\alpha+d}}\right) \text{ para } \delta \rightarrow 0^+. \quad (4.76)$$

Demostración. Por simplicidad supondremos que $\Omega = [0, 1]^d$. El caso general para Ω acotado se sigue cubriendo Ω con un hipercubo y teniendo en cuenta la correspondiente modificación en el valor de $\#\Lambda_j$. Sean $\{(2^{-\alpha j(\lambda)}, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ la descomposición wavelet-vaguelet del operador T con respecto a la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ y dW ruido blanco. Puesto que $\{u_\lambda\}_{\lambda \in \Lambda}$ es una sucesión de vaguelets, el Teorema 4.36 asegura que la sucesión $\{W_\lambda\}_{\lambda \in \Lambda}$, donde $W_\lambda \doteq \langle u_\lambda, dW \rangle$, es ruido gaussiano casi-independiente (RGCI, ver Definición 4.34 y Observación 4.35). Del Teorema 4.39 se sigue entonces que existen constantes positivas C_1 y C_2 tales que

$$\mathcal{B}_D(C_1 \delta, \Xi_{p,q}^s, \mathcal{M}_c) \leq \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) \leq \mathcal{B}_D(C_2 \delta, \Xi_{p,q}^s, \mathcal{M}_c). \quad (4.77)$$

Probaremos ahora que

$$\mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c) = O(\delta^{\frac{2s}{s+\alpha+d/2}}) \text{ para } \delta \rightarrow 0^+.$$

Recordemos en primer lugar que, por (4.64b) se tiene que

$$\mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c) = \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} B_D(\Pi, \delta)$$

$$= \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E_\Pi \left[\left(E(A_\lambda | \{A_\sigma^\delta\}_{\sigma \in \Lambda}) - A_\lambda \right)^2 \right], \quad (\text{por Def. 4.27}) \quad (4.78)$$

donde Π es la distribución conjunta de las variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$.

Ahora, en virtud del Lema 4.32, podemos restringir el supremo en (4.78) a sucesiones de variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$ que sean independientes e idénticamente distribuidas para cada nivel j . Luego,

$$\sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E_\Pi \left[\left(E(A_\lambda | \{A_\sigma^\delta\}_{\sigma \in \Lambda}) - A_\lambda \right)^2 \right] = \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E_\Pi \text{Var}(A_\lambda | A_\lambda^\delta). \quad (4.79)$$

De (4.78) y (4.79) se tiene entonces que

$$\mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c) = \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E_\Pi \text{Var}(A_\lambda | A_\lambda^\delta). \quad (4.80)$$

Por otro lado, notar que dados $\{A_\lambda\}_{\lambda \in \Lambda} \in \Xi_{p,q}^s$ y el modelo estocástico $A_\lambda^\delta = A_\lambda + 2^{\alpha j(\lambda)} \delta Z_\lambda$, $\forall \lambda \in \Lambda$, descrito en (4.45), se tiene obviamente que

$$E_\Pi \text{Var}(A_\lambda | A_\lambda^\delta) \leq \sup_{B_\lambda: E(|B_\lambda|^p) \leq E(|A_\lambda|^p)} E_\Pi \text{Var}(B_\lambda | B_\lambda^\delta), \quad \forall \lambda \in \Lambda,$$

donde $B_\lambda^\delta = B_\lambda + 2^{\alpha j(\lambda)} \delta Z_\lambda$, $\forall \lambda \in \Lambda$. Se sigue entonces que

$$\sum_{\lambda \in \Lambda} E_\Pi \text{Var}(A_\lambda | A_\lambda^\delta) \leq \sum_{\lambda \in \Lambda} \sup_{B_\lambda: E(|B_\lambda|^p) \leq E(|A_\lambda|^p)} E_\Pi \text{Var}(B_\lambda | B_\lambda^\delta).$$

Luego, tomando supremo obtenemos que

$$\begin{aligned} \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E_\Pi \text{Var}(A_\lambda | A_\lambda^\delta) &\leq \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} \sup_{B_\lambda: E(|B_\lambda|^p) \leq E(|A_\lambda|^p)} E_\Pi \text{Var}(B_\lambda | B_\lambda^\delta) \\ &= \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \sup_{B_\lambda: E(|B_\lambda|^p) \leq \tau_\lambda^p} E_\Pi \text{Var}(B_\lambda | B_\lambda^\delta), \end{aligned} \quad (4.81)$$

donde la última igualdad se obtiene definiendo $\tau_\lambda \doteq [E(|A_\lambda|^p)]^{1/p} \forall \lambda \in \Lambda$, y recordando que, puesto que las variables aleatorias A_λ son i.i.d. para cada nivel j , se tiene que $\tau_{\lambda_1} = \tau_{\lambda_2}$ si $j(\lambda_1) = j(\lambda_2)$.

Por lo tanto,

$$\begin{aligned} \mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c) &= \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E_\Pi \text{Var}(A_\lambda | A_\lambda^\delta) \quad (\text{por (4.80)}) \\ &\leq \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \sup_{B_\lambda: E(|B_\lambda|^p) \leq \tau_\lambda^p} E_\Pi \text{Var}(B_\lambda | B_\lambda^\delta) \quad (\text{por (4.81)}) \\ &= \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \rho(p, \tau_\lambda, 2^{\alpha j(\lambda)} \delta), \end{aligned} \quad (4.82)$$

donde la última igualdad se sigue de la definición del riesgo escalar de Bayes (Definición 4.30) y del hecho que $B_\lambda^\delta = B_\lambda + 2^{\alpha j(\lambda)} \delta Z_\lambda, \forall \lambda \in \Lambda$.

Ahora, por definición de $\Theta_{p,q}^s, \{\tau_\lambda\}_{\lambda \in \Lambda} \in \Theta_{p,q}^s$ si y solo si

$$\sum_{j \geq j_0}^{\infty} 2^{jq(s + \frac{d}{2} - \frac{d}{p})} \left(\sum_{\lambda \in \Lambda_j} |\tau_\lambda|^p \right)^{\frac{q}{p}} \leq C^q.$$

Pero como $\tau_{\lambda_1} = \tau_{\lambda_2} \forall j(\lambda_1) = j(\lambda_2) \forall j \geq j_0$, se tiene que $\sum_{\lambda \in \Lambda_j} \tau_\lambda^p = \#\Lambda_j \cdot \tau_j^p$, y como $\Omega = [0, 1]^d$, $\#\Lambda_j = 2^{jd}$ para todo $j \geq j_0$. Luego

$$\begin{aligned} & \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \rho(p, \tau_\lambda, 2^{\alpha j(\lambda)} \delta) \\ &= \sup_{\{\tau_j\}^+ : \sum_{j \geq j_0} 2^{jq(s+d/2-d/p)} (\#\Lambda_j)^{\frac{q}{p}} \tau_j^q \leq C^q} \sum_{j \geq j_0} \#\Lambda_j \rho(p, \tau_j, 2^{\alpha j} \delta) \\ &= \sup_{\{\tau_j\}^+ : \sum_{j \geq j_0} 2^{jq(s+d/2)} \tau_j^q \leq C^q} \left(2^{j_0 d} \rho(p, \tau_{j_0}, 2^{\alpha j_0} \delta) + \sum_{j > j_0} 2^{jd} \rho(p, \tau_j, 2^{\alpha j} \delta) \right) \\ &\leq \sup_{\{\tau_j\}^+ : \sum_{j \geq j_0} 2^{jq(s+d/2)} \tau_j^q \leq C^q} \left(2^{j_0(d+2\alpha)} \delta^2 + \sum_{j > j_0} 2^{jd} \rho(p, \tau_j, 2^{\alpha j} \delta) \right) \\ &= 2^{j_0(d+2\alpha)} \delta^2 + \sup_{\{\tau_j\}^+ : \sum_{j > j_0} 2^{jq(s+d/2)} \tau_j^q \leq C^q} \sum_{j > j_0} 2^{jd} \rho(p, \tau_j, 2^{\alpha j} \delta) \\ &= 2^{j_0(d+2\alpha)} \delta^2 + P_{p,q,s,\alpha}(\delta, C) \\ &\leq 2^{j_0(d+2\alpha)} \delta^2 + Q_{p,q,s,\alpha}(\delta, C). \quad (\text{por (4.73)}) \quad (4.83) \end{aligned}$$

De (4.82) y (4.83) obtenemos entonces que

$$\begin{aligned} \mathcal{B}_D(\delta, \Xi_{p,q}^s, \mathcal{M}_c) &\leq 2^{j_0(d+2\alpha)} \delta^2 + Q_{p,q,s,\alpha}(\delta, C) \\ &= O\left(\delta^{\frac{2s}{s+d/2+\alpha}}\right), \quad \text{para } \delta \rightarrow 0^+, \quad (4.84) \end{aligned}$$

donde la última igualdad es consecuencia inmediata del Lema 4.43.

Finalmente,

$$\begin{aligned} \mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_c) &\leq \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_c) \quad (\text{por (4.50)}) \\ &\leq \mathcal{B}_D(C_2 \delta, \Xi_{p,q}^s, \mathcal{M}_c) \quad (\text{por (4.77)}) \\ &= O\left(\delta^{\frac{2s}{s+\alpha+d/2}}\right), \quad \text{para } \delta \rightarrow 0^+. \quad (\text{por (4.84)}) \end{aligned}$$

Esto concluye la demostración de este teorema. ■

Observación 4.46. Notar que el orden de convergencia mejora con la suavidad del dato s y, en efecto, tiende a $O(\delta)$ para $s \rightarrow \infty$. Del mismo modo, el orden empeora con la dimensión d y con el grado de homogeneidad α del operador T .

4.5.5. Convergencia del método RWVU

En esta sección nos proponemos demostrar un resultado análogo al obtenido en el Teorema 4.45 pero ahora para el método de regularización wavelet-vaguelet con umbralado wavelet fuerte. En particular probaremos que el método RWVU resulta también en un orden de convergencia óptimo para el error total.

Como vimos en la Sección 1.4.2, el procedimiento conocido como umbralado wavelet fuerte consiste, esencialmente, en anular los coeficientes cuyos valores son menores a un umbral prefijado. A menudo este procedimiento resulta más adecuado que muchos otros métodos tradicionales de suavizado o de eliminación de ruido, que sólo consiguen reducir ruido a costa de suavizar también la función original. Recordemos que dado $\gamma \in \mathbb{R}^+$ (el umbral) denotamos con $S_\gamma : V_j \rightarrow V_j$ al operador de umbralado sobre el espacio de nivel j definido por

$$S_\gamma f(x) = \sum_{k \in \mathbb{Z}^d} \beta_{j,k} \chi_{\{|\beta_{j,k}| \geq \gamma\}}(\beta_{j,k}) \phi_{j,k}(x), \quad f \in V_j,$$

dónde $\beta_{j,k} \doteq \langle f, \phi_{j,k} \rangle$.

En este caso, a diferencia del umbralado utilizado en el capítulo anterior, suponemos $f \in L^2(\mathbb{R})$ y utilizamos la descomposición $L^2(\mathbb{R}) = V_{j_0} \oplus \bigoplus_{j \geq j_0} W_j$. Consideraremos que el umbral γ depende de cada nivel j y por lo tanto, ocasionalmente lo denotaremos con $\gamma = \gamma(j(\lambda))$ o simplemente $\gamma = \gamma(\lambda)$. Como vimos en las secciones 4.4 y 4.5, las soluciones aproximadas por el método RWVU, se obtienen a partir del siguiente modelo estocástico que surge de utilizar un enfoque bayesiano:

$$A_\lambda^{\delta, \gamma(\lambda)} \doteq S_{\gamma(\lambda)} A_\lambda^\delta = S_{\gamma(\lambda)} (A_\lambda + \delta \sigma_\lambda Z_\lambda). \quad (4.85)$$

Para probar el resultado fundamental sobre la convergencia del método RWVU, necesitaremos introducir previamente el concepto de *riesgo escalar de umbralado de Bayes* y dos resultados previos (Teorema 4.48 y Teorema 4.49) que facilitarán el desarrollo de la demostración del resultado principal (Teorema 4.51).

Definición 4.47 (Riesgo escalar de umbralado de Bayes). Sean $\tau, \delta, \gamma \in \mathbb{R}^+$, $p \in \mathbb{N}$, X una variable aleatoria y $X^{\delta, \gamma} \doteq S_\gamma X^\delta = S_\gamma (X + \delta Z)$, con $Z \sim N(0, 1)$. Se define el *riesgo escalar*

de umbralado de Bayes, asociado a p , γ , τ y δ , como

$$\rho_U(p, \tau, \delta) \doteq \inf_{\gamma > 0} \sup_{X \in \mathcal{U}(p, \tau)} E_X((X^{\delta, \gamma} - X)^2). \quad (4.86)$$

El siguiente teorema relaciona el riesgo escalar de umbralado de Bayes que acabamos de definir con el riesgo escalar de Bayes asociado al método RWV (ver Definición 4.30).

Teorema 4.48. Sean $\tau, \delta, \gamma \in \mathbb{R}^+$, $p \in \mathbb{N}$ y los riesgos escalares de Bayes ρ y ρ_U definidos por (4.52) y (4.86), respectivamente. Entonces, existe una constante $C = C(p)$, tal que

$$\rho_U(p, \tau, \delta) \leq C(p) \rho(p, \tau, \delta) \quad \forall \tau > 0 \text{ y } \forall \delta > 0. \quad (4.87)$$

Demostración. Utilizando un razonamiento análogo al del Lema 4.31 se puede probar inmediatamente que la identidad (4.53) vale también para ρ_U , es decir

$$\rho_U(p, \tau, \delta) = a^{-2} \rho_U(p, a\tau, a\delta), \quad \forall a > 0. \quad (4.88)$$

Luego, de (4.53) y (4.88) se sigue que

$$\begin{aligned} \sup_{\tau > 0, \delta > 0} \frac{\rho_U(p, \tau, \delta)}{\rho(p, \tau, \delta)} &= \sup_{\tau > 0, \delta > 0} \frac{\rho_U(p, \tau/\delta, 1)}{\rho(p, \tau/\delta, 1)} \\ &= \sup_{\nu > 0} \frac{\rho_U(p, \nu, 1)}{\rho(p, \nu, 1)} \\ &\doteq \sup_{\nu > 0} h_p(\nu). \end{aligned} \quad (4.89)$$

En virtud de (4.89), resulta que para probar (4.87) es suficiente con probar que $\exists C = C(p)$ tal que $\sup_{\nu > 0} h_p(\nu) \leq C$. A partir de las definiciones de ρ y ρ_U se puede probar, sin mayores dificultades, que ambas son continuas como función de ν , para ν en $(0, \infty)$, y en consecuencia también lo es h_p . Por lo tanto será suficiente con probar que la función $h_p(\nu)$ se mantiene acotada tanto para $\nu \rightarrow 0^+$ como para $\nu \rightarrow \infty$.

Notar que para $\nu \rightarrow \infty$ la condición $X \in \mathcal{U}(p, \nu)$ que aparece en ambas definiciones de ρ y ρ_U se desvanece, al igual que, se desvanece la condición sobre el umbral γ , para $\gamma \rightarrow 0^+$. Por lo tanto $\rho_U(p, \nu, 1) \leq \rho(p, \nu, 1)$ para $\nu \rightarrow \infty$. Así, $\limsup_{\nu \rightarrow \infty} h_p(\nu) \leq 1$. Más aún, con un argumento similar se puede probar que existe $\lim_{\nu \rightarrow \infty} h_p(\nu) = 1$.

La demostración de que $h_p(\nu)$ se mantiene acotada para $\nu \rightarrow 0^+$ es significativamente más engorrosa, por lo cual omitiremos aquí mayores detalles limitándonos, a continuación, a bosquejar los pasos fundamentales que conducen a la misma. En primer lugar, definiendo $\gamma = \gamma(\nu) \doteq (2 \log \nu^{-p} + a \log(2 \log \nu^{-p}))^{1/2}$, para $\nu < 1$, con $a > p - 1$, se puede probar que

existe una constante positiva \tilde{C} tal que

$$\sup_{X \in \mathcal{U}(p, \tau)} E_X((X^{\delta, \gamma(\nu)} - X)^2) \leq \tilde{C} \nu^p (\gamma(\nu))^{2-p} \text{ para } \nu \rightarrow 0^+,$$

y por lo tanto

$$\rho_U(p, \nu, 1) = \inf_{\gamma > 0} \sup_{X \in \mathcal{U}(p, \tau)} E_X((X^{\delta, \gamma} - X)^2) \leq \tilde{C} \nu^p (\gamma(\nu))^{2-p} \text{ para } \nu \rightarrow 0. \quad (4.90)$$

Por otro lado, se puede probar que para $\nu \rightarrow 0^+$

$$\rho(p, \nu, 1) \equiv \begin{cases} \nu^2 & \text{si } p \geq 2 \\ \nu^p (2 \log \nu^{-p})^{\frac{2-p}{2}} & \text{si } 0 < p < 2 \end{cases}, \quad (4.91)$$

donde $f \equiv g$ significa que existen constante C_1 y C_2 positivas tales que $C_1 g \leq f \leq C_2 g$ (ver, por ejemplo, Proposición 18 en [15]).

Luego, de (4.90) y (4.91), para $p \geq 2$ existe una constante positiva \hat{C} tal que

$$\frac{\rho_U(p, \nu, 1)}{\rho(p, \nu, 1)} \leq \hat{C} \frac{\nu^p (\gamma(\nu))^{2-p}}{\nu^2} = \hat{C} \left(\frac{\nu}{\gamma(\nu)} \right)^{p-2} \leq \hat{C} \text{ para } \nu \rightarrow 0^+, \quad (4.92)$$

pues $\gamma(\nu) \rightarrow \infty$ para $\nu \rightarrow 0^+$.

De igual modo y también como consecuencia de (4.90) y (4.91), para $0 < p < 2$ existe otra constante positiva \bar{C} tal que

$$\frac{\rho_U(p, \nu, 1)}{\rho(p, \nu, 1)} \leq \bar{C} \left(\frac{(\gamma(\nu))^2}{2 \log \nu^{-p}} \right)^{\frac{2-p}{2}} = \bar{C} \left(1 + \frac{a \log(2 \log \nu^{-p})}{2 \log \nu^{-p}} \right)^{\frac{2-p}{2}} \leq 2 \bar{C}, \text{ para } \nu \rightarrow 0^+, \quad (4.93)$$

donde la última desigualdad se sigue del hecho que $2 - p > 0$ y $\log \nu^{-p} \rightarrow \infty$ para $\nu \rightarrow 0^+$.

Finalmente, de (4.92) y (4.93) se sigue que $\limsup_{\nu \rightarrow 0^+} h_p(\nu) < \infty$ para todo $p > 0$. Esto completa la demostración del teorema. ■

El siguiente teorema es en principio un resultado técnico que esencialmente nos dice que, bajo ciertas condiciones, un ínfimo y un supremo pueden ser intercambiados. El mismo, será de suma importancia en la demostración del teorema principal. Para simplificar la notación, denotaremos con V al conjunto de distribuciones de variables aleatorias con p momentos en $\Theta_{p,q}^s$, es decir $V(p, q, s) \doteq \{\Pi_A \in \mathcal{D} : \{(E|A_\lambda|^p)^{1/p}\}_{\lambda \in \Lambda} \in \Theta_{p,q}^s\}$.

Teorema 4.49. Sean $p, q, s \in \mathbb{N}$ y $\delta \in \mathbb{R}_0^+$. Entonces

$$\inf_{\gamma(\lambda) \geq 0} \sup_{\mu_A \in V(p, q, s)} \sum_{\lambda \in \Lambda} E_{\mu_A}(A_\lambda^{\delta, \gamma(\lambda)} - A_\lambda)^2 = \sup_{\mu_A \in V(p, q, s)} \inf_{\gamma(\lambda) \geq 0} \sum_{\lambda \in \Lambda} E_{\mu_A}(A_\lambda^{\delta, \gamma(\lambda)} - A_\lambda)^2, \quad (4.94)$$

donde $\{A_\lambda\}_{\lambda \in \Lambda}$ denota una sucesión de variables aleatorias con distribución conjunta μ_A y $\{A_\lambda^{\delta, \gamma(\lambda)}\}_{\lambda \in \Lambda}$ es la sucesión de variables aleatorias definida como en (4.85).

Demostración. Sea $\mu_A \in V(p, q, s)$ la distribución asociada a la sucesión de variables aleatorias $\{A_\lambda\}_{\lambda \in \Lambda}$. Para cada $\lambda \in \Lambda$ definimos $h_\lambda : \mathbb{R}_0^+ \times V(p, q, s) \rightarrow \mathbb{R}_0^+$ y $g_\lambda : V(p, q, s) \rightarrow \mathbb{R}_0^+$ como

$$h_\lambda(\gamma(\lambda), \mu_A) \doteq E_{\mu_A} \left((A_\lambda^{\delta, \gamma(\lambda)} - A_\lambda)^2 \right) \quad \text{y} \quad g_\lambda(\mu_A) \doteq \inf_{\gamma(\lambda) \geq 0} h_\lambda(\gamma(\lambda), \mu_A). \quad (4.95)$$

Luego, el lado derecho de (4.94) se puede escribir como

$$\begin{aligned} \sup_{\mu_A \in V(p, q, s)} \inf_{\gamma(\lambda) \geq 0} \sum_{\lambda \in \Lambda} h_\lambda(\gamma(\lambda), \mu_A) &= \sup_{\mu_A \in V(p, q, s)} \sum_{\lambda \in \Lambda} \inf_{\gamma(\lambda) \geq 0} h_\lambda(\gamma(\lambda), \mu_A) \\ &= \sup_{\mu_A \in V(p, q, s)} \sum_{\lambda \in \Lambda} g_\lambda(\mu_A), \end{aligned} \quad (4.96)$$

donde la primera igualdad se sigue del hecho que $\#\Lambda_j < \infty$ y los términos de la suma están desacoplados en $\{\gamma(\lambda)\}_{\lambda \in \Lambda}$ para todo λ_1 y λ_2 tal que $j(\lambda_1) \neq j(\lambda_2)$. La segunda igualdad se sigue de (4.95).

Usando un argumento de compacidad se puede probar que el supremo en (4.96) se alcanza para alguna $\mu^* \in V(p, q, s)$, la que, por lo tanto, es una medida menos favorable con respecto al error de umbralado en el sentido de (4.47). De igual modo, se puede probar que para cada $\lambda \in \Lambda$ existe $\gamma^* = \gamma^*(\lambda, \mu) \geq 0$ tal que

$$g_\lambda(\mu) = h_\lambda(\gamma^*(\lambda), \mu), \quad \forall \mu \in V(p, q, s). \quad (4.97)$$

A continuación definimos $H : \ell_{0,+}^2 \times V(p, q, s) \rightarrow \mathbb{R}_0^+$ y $G : V(p, q, s) \rightarrow \mathbb{R}_0^+$ como

$$H(\{\gamma(\lambda)\}, \mu) \doteq \sum_{\lambda \in \Lambda} h_\lambda(\gamma(\lambda), \mu) \quad \text{y} \quad G(\mu) \doteq \sum_{\lambda \in \Lambda} g_\lambda(\mu).$$

Luego,

$$\begin{aligned} H(\gamma^*(\lambda, \mu^*), \mu^*) &= G(\mu^*) && \text{(por def. de } \gamma^*) \\ &= \sup_{\mu_A \in V(p, q, s)} \sum_{\lambda \in \Lambda} \inf_{\gamma(\lambda) \geq 0} h_\lambda(\gamma(\lambda), \mu_A). && \text{(por def. de } \mu^*) \end{aligned}$$

Probaremos ahora que el par $(\{\gamma^*(\lambda, \mu^*)\}, \mu^*)$ es un punto silla para la función H , lo que claramente implica (4.94). Para ello, dados $\mu \in V(p, q, s)$ y $t \in [0, 1]$ definimos $\mu_t \in V(p, q, s)$ como $\mu_t \doteq (1-t)\mu^* + t\mu$. Notar que cualquiera sea $\mu \in V(p, q, s)$, $\mu_0 = \mu^*$. De la definición de μ^* se sigue que

$$\sum_{\lambda \in \Lambda} g_\lambda(\mu^*) \geq \sum_{\lambda \in \Lambda} g_\lambda(\mu) \quad \forall \mu \in V(p, q, s).$$

Por lo tanto $\forall \mu \in V(p, q, s)$ se tiene que

$$0 \geq \left[\frac{d}{dt} \sum_{\lambda \in \Lambda} g_\lambda(\mu_t) \right] \Big|_{t=0}$$

$$\begin{aligned}
&= \sum_{\lambda \in \Lambda} \left[\frac{d}{dt} h_\lambda(\gamma^*(\lambda, \mu_t), \mu_t) \right] \Big|_{t=0} \quad (\text{por def. de } g_\lambda) \\
&= \sum_{\lambda \in \Lambda} \left\{ \left[\frac{\partial}{\partial t} h_\lambda(\gamma^*(\lambda, \mu_t), \mu^*) \right] \Big|_{t=0} + \left[\frac{\partial}{\partial t} h_\lambda(\gamma^*(\lambda, \mu^*), \mu_t) \right] \Big|_{t=0} \right\} \quad \left(\begin{array}{l} \text{pues } \mu_t|_{t=0} = \mu^* \text{ y por} \\ \text{def. de } \gamma^*(\lambda, \mu^*) \end{array} \right) \\
&\geq \sum_{\lambda \in \Lambda} \left[\frac{\partial}{\partial t} h_\lambda(\gamma^*(\lambda, \mu^*), \mu_t) \right] \Big|_{t=0} \quad \left(\begin{array}{l} \left[\frac{\partial}{\partial t} h_\lambda(\gamma^*(\lambda, \mu_t), \mu^*) \right] \Big|_{t=0} \geq 0 \text{ pues } \forall \mu \in V(p, q, s) \\ h_\lambda(\gamma^*(\lambda, \mu^*), \mu^*) \leq h_\lambda(\gamma^*(\lambda, \mu), \mu^*) \end{array} \right) \\
&= \sum_{\lambda \in \Lambda} (h_\lambda(\gamma^*(\lambda, \mu^*), \mu) - h_\lambda(\gamma^*(\lambda, \mu^*), \mu^*)). \quad (4.98)
\end{aligned}$$

La última igualdad se sigue inmediatamente de la definición de μ_t , puesto que $h_\lambda(\gamma(\lambda), \mu_t) = (1-t)h_\lambda(\gamma(\lambda), \mu^*) + t h_\lambda(\gamma(\lambda), \mu)$.

Luego, $\sum_{\lambda \in \Lambda} h_\lambda(\gamma^*(\lambda, \mu^*), \mu^*) \leq \sum_{\lambda \in \Lambda} h_\lambda(\gamma^*(\lambda, \mu), \mu)$ para todo $\mu \in V(p, q, s)$ y por lo tanto μ^* es un máximo para $G(\gamma^*(\mu), \mu)$. Se sigue entonces que $(\{\gamma^*(\lambda, \mu^*)\}, \mu^*)$ es un punto silla para la función H , lo que completa la demostración. ■

Observación 4.50. Con el objetivo de simplificar la demostración del teorema precedente, hemos intencionalmente omitido justificar la t -diferenciabilidad de todas las funciones que aparecen en (4.98). Si bien tal justificación puede hacerse de manera rigurosa, consideramos que la misma es irrelevante para los objetivos de la demostración del teorema, él que, como ya mencionamos, es en sí mismo de naturaleza sumamente técnica.

Estamos finalmente en condiciones de probar el resultado principal de este capítulo, según el cuál, el método de regularización wavelet-vaguelet con umbralado propuesto en la Sección 4.4.3, provee un orden de convergencia óptimo para el error de estimación total.

Teorema 4.51. Sean $\Omega \subset \mathbb{R}^d$ acotado, $\{\psi_\lambda\}_{\lambda \in \Lambda} \subset L^2(\Omega)$ una base wavelet y $T : L^2(\Omega) \rightarrow L^2(\Omega)$ un operador homogéneo de grado α , que posee una descomposición wavelet vaguelet $\{(\kappa_\lambda, u_\lambda, v_\lambda)\}_{\lambda \in \Lambda}$ con respecto a la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$, con valores casi-singulares $\kappa_\lambda = 2^{-\alpha j(\lambda)} \forall j(\lambda) \geq j_0$ y supongamos que $f \in B_{p,q}^s(\Omega)$ para algún $s \in \mathbb{N}$ y $0 < p \leq q \leq \infty$.

Entonces, para todo $\delta > 0$, el riesgo minimax cuadrático wavelet-vaguelet con umbralado (RMC-WVU) asociado a $g = Tf$ y $g^\delta \doteq g + \delta dW$ satisface

$$\mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_{c,\gamma}) = O\left(\delta^{\frac{2s}{s+\alpha+d/2}}\right), \quad \text{para } \delta \rightarrow 0^+.$$

Demostración. Notemos que de la definición del RMC-WVU y RMCB-WVU (ver Definición 4.26 y Definición 4.29) se tiene que

$$\mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_{c,\gamma}) \leq \mathcal{B}(\delta, \Xi_{p,q}^s, \mathcal{M}_{c,\gamma}) \quad (\text{por (4.51)})$$

$$\begin{aligned}
&= \inf_{\{\gamma(\lambda)\}^+} \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} E(\|A_\lambda^{\delta,\gamma} - A_\lambda\|^2) \quad (\text{por (4.49)}) \\
&= \inf_{\{\gamma(\lambda)\}^+} \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right). \quad (4.99)
\end{aligned}$$

Por otro lado, utilizando un argumento análogo a la demostración de (4.81) en la prueba del Teorema 4.45 obtenemos

$$\sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right) \leq \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \sup_{A_\lambda: E(|A_\lambda|^p) \leq \tau_\lambda^p} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right), \quad (4.100)$$

para $\tau_\lambda \doteq [E(|A_\lambda|^p)]^{1/p} \forall \lambda \in \Lambda$.

Luego,

$$\begin{aligned}
\mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_{c,\gamma}) &\leq \inf_{\{\gamma(\lambda)\}^+} \sup_{\{A_\lambda\} \in \Xi_{p,q}^s} \sum_{\lambda \in \Lambda} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right) \quad (\text{por (4.99)}) \\
&\leq \inf_{\{\gamma(\lambda)\}^+} \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \sup_{A_\lambda: E(|A_\lambda|^p) \leq \tau_\lambda^p} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right) \quad (\text{de (4.100)}) \\
&\leq \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \inf_{\{\gamma(\lambda)\}^+} \sum_{\lambda \in \Lambda} \sup_{A_\lambda: E(|A_\lambda|^p) \leq \tau_\lambda^p} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right) \quad (\text{por Teor. 4.49}) \\
&= \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \inf_{\gamma(\lambda) > 0} \sup_{A_\lambda: E(|A_\lambda|^p) \leq \tau_\lambda^p} E\left((A_\lambda^{\delta,\gamma} - A_\lambda)^2\right) \begin{pmatrix} \gamma(\lambda) = \gamma(j(\lambda)) \\ y \notin \Lambda_j < \infty \\ \forall j \geq j_0 \end{pmatrix} \\
&= \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \rho_U(p, \tau_\lambda, 2^{\alpha j(\lambda)} \delta) \quad \left(\begin{array}{l} \text{por Definición 4.47 y} \\ A_\lambda^{\delta,\gamma} = S_\gamma(A_\lambda + 2^{\alpha j(\lambda)} \delta Z_\lambda) \end{array} \right) \\
&\leq C'(p) \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \rho(p, \tau_\lambda, 2^{\alpha j(\lambda)} \delta). \quad (\text{por Teor. 4.48}) \quad (4.101)
\end{aligned}$$

Finalmente, puesto que las variables aleatorias A_λ se pueden suponer i.i.d. para cada nivel j , vale la desigualdad (4.83) de la demostración del Teorema 4.45 y, por lo tanto, se tiene que

$$\begin{aligned}
\mathcal{R}(\delta, B_{p,q}^s, \mathcal{M}_{c,\gamma}) &\leq C'(p) \sup_{\substack{\{\tau_\lambda\}^+ \in \Theta_{p,q}^s: \\ \tau_{\lambda_1} = \tau_{\lambda_2} \text{ si } j(\lambda_1) = j(\lambda_2)}} \sum_{\lambda \in \Lambda} \rho(p, \tau_\lambda, 2^{\alpha j(\lambda)} \delta) \quad (\text{por (4.101)}) \\
&\leq C'(p) \left(2^{j_0(d+2\alpha)} \delta^2 + Q_{p,q,s,\alpha}(\delta, C) \right) \quad (\text{por (4.83)}) \\
&= O\left(\delta^{\frac{2s}{s+d/2+\alpha}}\right), \quad \text{para } \delta \rightarrow 0^+. \quad (\text{por Lema 4.43})
\end{aligned}$$

Esto concluye la demostración del teorema. ■

Al igual que para el método RWV, notar que aquí también el orden de convergencia mejora con la suavidad del dato s y tiende a $O(\delta)$ para $s \rightarrow \infty$. Del mismo modo, el orden empeora con la dimensión d y con el grado de homogeneidad α del operador T .

4.6. Operadores integrales

En esta sección presentamos un tipo particular de operadores que satisfacen las condiciones necesarias para tener una descomposición wavelet-vaguelet. Este caso particular será retomado en la sección de aplicaciones numéricas, como ejemplos de aplicación a señales e imágenes. Comenzamos la sección definiendo operadores de tipo convolución. Luego, enunciamos y demostramos una serie de proposiciones y lemas, los que son necesarios para construir la descomposición wavelet-vaguelet. Cerramos la sección y el capítulo con la formalización de la descomposición wavelet-vaguelet para el caso de operadores integrales de convolución.

Definición 4.52. Decimos que $T : L^2(\mathbb{R}^d; \mathbb{R}) \rightarrow L^2(\mathbb{R}^d; \mathbb{R})$ es un *operador integral de tipo convolución asociado al núcleo* $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ si

$$(Tf)(s) \doteq \int_{\mathbb{R}^d} f(t)k(s, t) dt.$$

Un caso especial de operadores integrales, con los que trabajaremos a lo largo de este capítulo, lo constituyen los operadores de convolución, en los que el núcleo $k(s, t)$ depende de s y t solo a través de la diferencia, $s - t$.

Definición 4.53. Denotamos por $T_k : L^2(\mathbb{R}^d; \mathbb{R}) \rightarrow L^2(\mathbb{R}^d; \mathbb{R})$ al operador de convolución asociado al núcleo $k : \mathbb{R}^d \rightarrow \mathbb{R}$ dado por

$$(T_k f)(s) \doteq \int_{\mathbb{R}^d} f(t)k(s - t) dt, \quad \text{para } f \in L^2(\mathbb{R}^d; \mathbb{R}).$$

La siguiente proposición es un resultado clásico sobre condiciones suficientes para que un operador integral de tipo convolución sea compacto.

Teorema 4.54. Sean $\Omega \subset \mathbb{R}^d$ compacto y $k : \Omega \times \Omega \rightarrow \mathbb{R}$ una función que satisface al menos una de las siguientes condiciones:

- $k \in L^2(\Omega \times \Omega; \mathbb{R})$, o
- k es débilmente singular, es decir, k es continua sobre $\{(s, t) \in \Omega \times \Omega : s \neq t\}$ y existen $\epsilon > 0$ y $M > 0$ tal que, para todo $s \neq t \in \Omega$ se verifica

$$|k(s, t)| \leq \frac{M}{|s - t|^{d-\epsilon}}.$$

Entonces el operador de tipo convolución $T : L^2(\Omega; \mathbb{R}) \rightarrow L^2(\Omega; \mathbb{R})$, con núcleo k , es compacto.

Demostración. La prueba de este teorema puede encontrarse, por ejemplo, en el capítulo 2 (§ 2) en [17]. ■

Proposición 4.55. Sea $T_k : L^2(\mathbb{R}^d; \mathbb{R}) \rightarrow L^2(\mathbb{R}^d; \mathbb{R})$ un operador de convolución con núcleo $k : \mathbb{R}^d \rightarrow \mathbb{R}$, entonces su adjunto está dado por $T^* = T_{k_*}$ donde $k_* : \mathbb{R}^d \rightarrow \mathbb{R}$ está dado por $k_*(x) = k(-x) \forall x \in \mathcal{D}(T_k)$.

Demostración. La prueba se sigue trivialmente verificando que $\langle g, T_k f \rangle = \langle T_{k_*} g, f \rangle \forall f, g \in \mathcal{D}(T_k)$. ■

Observación 4.56. De la proposición anterior se sigue que un operador de convolución es autoadjunto si y solo si su núcleo es par. Un ejemplo de operador que satisface esta condición y será utilizado en la sección siguiente es el operador convolución con núcleo gaussiano.

Lema 4.57. Sea $T_k : L^2(\mathbb{R}^d; \mathbb{R}) \rightarrow L^2(\mathbb{R}^d; \mathbb{R})$ un operador de convolución con núcleo $k : \mathbb{R}^d \rightarrow \mathbb{R}$, entonces:

- T_k es inyectivo si y solo si $\hat{k} \neq 0$ c.t.p. y
- si $\hat{k} \neq 0$ c.t.p. el operador inverso puede escribirse formalmente en términos de \hat{k} como

$$(T_k^{-1} g_{(\omega)})(t) = \mathcal{F}^{-1} \left\{ \frac{\hat{g}(\omega)}{\hat{k}(\omega)} \right\} (t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\omega t} \frac{\hat{g}(\omega)}{\hat{k}(\omega)} d\omega. \quad (4.102)$$

Demostración. a) \Rightarrow) Sea $f \in L^2(\mathbb{R}^d; \mathbb{R})$ tal que $T_k f = 0$ en c.t.p., esto implica que $\mathcal{F}\{T_k f\} = 0$ c.t.p. y por definición del operador esto implica $\mathcal{F}\{f * k\} = 0$ c.t.p.. Luego, $\hat{f} \cdot \hat{k} = 0$ c.t.p. y como por hipótesis $\hat{k} \neq 0$ c.t.p. entonces debe ser $\hat{f} = 0$ c.t.p. y por lo tanto $f = 0$. Luego el operador es inyectivo.

\Leftarrow) Supongamos por el absurdo que $\hat{k}(\omega) = 0$ para todo $\omega \in A$, donde A es un conjunto de medida positiva, y definamos $\hat{f}(\omega) \doteq \chi_A(\omega)$ donde χ_A es la función característica sobre el conjunto A . Bajo estas condiciones tenemos que $0 = \hat{f}(\omega) \cdot \hat{k}(\omega) = (T_k f)(\omega) \forall \omega$. Luego, $T_k f(t) = 0$ para f no nula y por lo tanto el operador no es inyectivo.

- Del inciso anterior tenemos que si $\hat{k} \neq 0$ c.t.p. el operador es inyectivo y por lo tanto podemos definir su inversa $T_k^{-1} : \mathcal{R}(T_k) \rightarrow \mathcal{D}(T_k)$. En tal caso, la identidad (4.102), es obvia. ■

Observar que la condición $\hat{k} \neq 0$ c.t.p. en el lema anterior en principio es una condición técnica y estrictamente necesaria, y no es claro que condiciones implican sobre k . Sin embargo, un teorema de Wiener del año 1932 provee una condición necesaria y suficiente para que la transformada de Fourier de una función sea distinta de cero en casi todo punto. En efecto,

Teorema 4.58. *Sea $f \in L^2(\mathbb{R})$, entonces $\hat{f} \neq 0$ c.t.p. si y solo si el conjunto de todas las traslaciones de f es denso en $L^2(\mathbb{R})$.*

Demostración. Ver [51], Capítulo 1, Sección 1. ■

Finalizamos este capítulo con un teorema que provee condiciones suficientes para que un operador de convolución posea una descomposición wavelet-vaguelet con respecto a una base wavelet dada.

Teorema 4.59. *Sea $T_k : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ un operador de convolución homogéneo de grado α con núcleo k , tal que $k \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ es Hölder continuo y $\hat{k} \neq 0$ c.t.p.. Sea además, $\psi \in L^2(\mathbb{R}^d)$ una wavelet con promedio nulo que verifica $\psi \in \mathcal{D}(T_k) \cap \mathcal{R}(T_k^*)$ y supongamos que existen constantes $b_1, b_2 > 0$ y $\nu_1, \nu_2 > 0$, tales que $\forall x \in \mathbb{R}^d$*

$$|[T_k \psi(\cdot)](x)| \leq \frac{b_1}{(1+|x|)^{d+\nu_1}} \quad y \quad |[(T_k^{-1})^* \psi(\cdot)](x)| \leq \frac{b_2}{(1+|x|)^{d+\nu_2}}.$$

Entonces la familia $\{(2^{-jd\alpha}, 2^{-jd\alpha}(T_k^{-1})^ \psi_\lambda, 2^{jd\alpha} T_k \psi_\lambda)\}_{\lambda \in \Lambda}$ es una descomposición wavelet vaguelet del operador T_k con respecto a la base ortonormal wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ y $\{2^{-j\alpha d}\}_{j \geq j_0}$ es la correspondiente sucesión de valores casi-singulares de dicha descomposición.*

Demostración. La prueba de este teorema consiste esencialmente en verificar que el operador T_k y la wavelet ψ satisfacen las hipótesis del Teorema 4.21. Notemos en primer lugar que puesto que $\hat{k} \neq 0$ c.t.p., por el Lema 4.57 se sigue que T_k es inyectivo.

De la Observación 4.22 se tiene que T_k conmuta con traslaciones, es decir, $\forall a \in \mathbb{R}^d$ se verifica que $T_k E_a = E_a T_k$, y por lo tanto se verifica la hipótesis $H2)$ del Teorema 4.21.

Por otro lado, dado que el núcleo k es Hölder continuo, de la Observación 4.24 se sigue que se satisface la hipótesis $H4)$, es decir, $T_k \psi$ y $(T_k^{-1})^* \psi$ son Hölder continuas.

Finalmente, puesto que ψ tiene promedio nulo y $k \in L^1(\mathbb{R})$ de la Observación 4.25 se tiene que $T_k \psi$ y $(T_k^{-1})^* \psi$ tienen promedio nulo. Luego, se satisface la hipótesis $H5)$ y esto concluye la demostración. ■

CAPÍTULO 5

Aplicaciones y resultados numéricos

Los resultados presentados en esta tesis han sido hasta aquí predominantemente teóricos. Sin embargo, consideramos importante implementar y evaluar, en aplicaciones concretas, los diferentes métodos de regularización estudiados en este trabajo. Este capítulo se centra en los resultados de las implementaciones numéricas obtenidas a partir de la resolución de problemas inversos particulares. A lo largo del capítulo utilizaremos tanto el método de regularización wavelet-espectral, desarrollado en el Capítulo 3, como el método de regularización wavelet-vaguelet, desarrollado en el Capítulo 4, para aproximar soluciones a problemas de deconvolución. En la Sección 5.1 se utiliza el método wavelet-espectral para resolver problemas de deconvolución en una dimensión, para restauración de señales, y en dos dimensiones, para restauración de imágenes. Por otro lado, la Sección 5.2 está dedicada a la implementación del método wavelet-vaguelet, estudiado en el Capítulo 4, nuevamente para el caso de problemas de deconvolución en restauración de señales.

5.1. Aplicaciones del método de regularización wavelet-espectral

En el Capítulo 3 presentamos el método de regularización wavelet-espectral para problemas inversos mal condicionados. Dicho método, involucra esencialmente dos pasos. Un primer paso de descomposición y umbralado wavelet y un segundo paso de regularización espectral. Con el objetivo de mostrar y cuantificar las ventajas del método propuesto, en las siguientes secciones presentaremos ejemplos numéricos utilizando este método, para problemas inversos de la forma $Tf = g$, donde el operador T es un operador integral de convolución.

Recordemos que un operador de convolución $T_k : L^2(\mathbb{R}^d; \mathbb{R}) \rightarrow L^2(\mathbb{R}^d; \mathbb{R})$ asociado a un núcleo $k : \mathbb{R}^d \rightarrow \mathbb{R}$ está dado por

$$(T_k f)(s) \doteq \int_{\mathbb{R}^d} f(t)k(s-t) dt. \quad (5.1)$$

Nos proponemos encontrar una solución aproximada a la solución exacta f en

$$g(s) \doteq \int_{\mathbb{R}^d} f(t)k(s-t) dt = (T_k f(\cdot))(s).$$

Por simplicidad supondremos que $\text{supp}(f)$ es compacto y k es un núcleo gaussiano. Supondremos, que el dato g no se conoce de manera exacta. Más precisamente, y continuando con los lineamientos de los capítulos anteriores, consideraremos un modelo de observación con ruido aditivo de la forma $g^\delta = g + \delta dW$, donde dW es ruido blanco y $\delta > 0$ una constante que cuantifica el nivel de error.

5.1.1. Deconvolución 1D: restauración de señales

En muchas ocasiones las perturbaciones en ciertas señales pueden modelarse satisfactoriamente a través de un proceso de convolución. En tales casos, mediante una correcta “deconvolución”, sería posible entonces reconstruir la señal original afectada por este proceso. El problema de deconvolución fue estudiado por primera vez como una herramienta para el análisis de medidas sísmicas (ver, por ejemplo [33], [18]). En la actualidad el uso de técnicas de deconvolución se ha extendido a una amplia gama de aplicaciones, tales como control automático y filtrado digital de sistemas, entre otros.

Dados el núcleo k y la función real f con soporte contenido en un intervalo I (supondremos $I = [0, 1]$ por simplicidad) la convolución de f con k define nuestro dato (limpio) g :

$$g(s) \doteq \int_0^1 f(t)k(s-t) dt. \quad (5.2)$$

El problema inverso (deconvolución) consiste en determinar f dados el núcleo k y la observación g . Por lo general el núcleo k es conocido, como lo será en todos los ejemplos numéricos que abordaremos a continuación. Sin embargo es oportuno mencionar que en ocasiones también el núcleo es desconocido. Estos problemas son llamados de “deconvolución ciega” y su tratamiento es considerablemente más complejo. Como caso emblemático de uno de estos, citamos uno relacionado a la restauración de las imágenes enviadas por el Telescopio espacial Hubble durante sus primeros tres años en órbita (1990-1993). Debido a problemas con el diseño de uno de sus espejos, las imágenes iniciales enviadas por el telescopio estaban distorsionadas en grado extremo. Afortunadamente, al detectar el problema, los científicos pudieron resolverlo (hasta el cambio de ese espejo en el año 1993 durante una de las misiones del transbordador espacial) a través de un problema de deconvolución ciega con casi 800.000 variables. Detalles de este interesante problema y su solución pueden encontrarse en el artículo [2]. Por otro lado, el dato g suele estar disponible solo a través de mediciones con ruido, usualmente sobre un conjunto finito de puntos s_1, s_2, \dots, s_m :

$$g_i \doteq g(s_i) + e_i, \quad i = 1, \dots, m,$$

donde e_i denota el ruido de medición asociado a la i -ésima observación. El ruido generalmente se supone con una cierta distribución estadística, que puede ser conocida o desconocida. En nuestro caso supondremos ruido blanco gaussiano con varianza δ^2 .

La implementación numérica de este problema de deconvolución requiere obviamente de un proceso previo de discretización. Existe una gran variedad de métodos disponibles para la discretización de ecuaciones integrales, los cuales permiten obtener, a partir de la ecuación integral, un sistema de ecuaciones lineales que, en principio, se pueden “resolver” numéricamente para proporcionar una solución discreta aproximada. En este trabajo utilizaremos una regla de cuadratura, mediante la cual es posible aproximar la integral en (5.2) para el caso unidimensional que nos proponemos trabajar, mediante una suma finita de términos. De esta manera, la integral de convolución puede tratarse como una suma finita de convolución entre el núcleo k y la función f . A partir de una discretización del dominio de integración, definida por $t_j \doteq \frac{j-1/2}{n}$ para $j = 1, \dots, n$, se tiene entonces la siguiente estimación

$$g(s) = \int_0^1 k(s-t)f(t) dt \simeq \sum_{j=1}^n \frac{1}{n} k(s-t_j)f(t_j). \quad (5.3)$$

En este capítulo nos concentramos en analizar aplicaciones a problemas inversos mal condicionados donde el núcleo k es gaussiano. Como veremos en la siguiente sección, los problemas de deconvolución con este núcleo son muy utilizados en procesamiento de imágenes ([47], [5]). En particular, estos núcleos modelan satisfactoriamente la distorsión en las imágenes captadas por telescopios terrestres debido a las diferentes densidades de las capas de la atmósfera. Se lo conoce como “turbulencia atmosférica”. Suponemos entonces $k : \mathbb{R} \rightarrow \mathbb{R}$ definido por

$$k(s) \doteq \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(s-\mu)^2}{2\sigma^2}}, \quad (5.4)$$

donde $\mu \in \mathbb{R}^d$ y $\sigma > 0$. En cada uno de los ejemplos siguientes supondremos $\mu = 0$ y expresaremos la desviación estándar σ en términos de un porcentaje ℓ de la norma $\|f\|_\infty$, es decir

$$\sigma = \frac{\ell}{100} \|f\|_\infty. \quad (5.5)$$

A partir de esta elección del núcleo, de las ecuaciones (5.3) y (5.4) se tiene que la discretización para el problema de deconvolución con núcleo gaussiano en los puntos de colocación s_1, s_2, \dots, s_m está dada por

$$g(s_i) = \sum_{j=1}^n \frac{1}{n\sigma\sqrt{2\pi}} e^{-\frac{(s_i-t_j)^2}{2\sigma^2}} f(t_j), \text{ con } i = 1, \dots, m \text{ y } j = 1, 2, \dots, n. \quad (5.6)$$

Notemos aquí que m no necesariamente debe ser igual a n . Sin embargo a lo largo de este trabajo, por simplicidad supondremos $m = n$ y $s_i = t_i \forall i$. El problema (5.6) puede escribirse en notación matricial como

$$Ax = y, \quad (5.7)$$

donde A es una matriz de $n \times n$ y $x, y \in \mathbb{R}^n$ definidos por

$$a_{i,j} \doteq \frac{1}{n\sigma\sqrt{2\pi}} e^{-\frac{(s_i-t_j)^2}{2\sigma^2}}, \quad y_i \doteq g(s_i) \quad \text{y} \quad x_j \doteq f(t_j),$$

para $i = 1, \dots, n, j = 1, \dots, n$. En todos los ejemplos siguientes consideraremos $n = 100$.

Como vimos en la Sección 4.6, los operadores de convolución asociados a núcleos que satisfacen las hipótesis del Teorema 4.54 (e.g. núcleos gaussianos) son compactos. Por lo tanto, el correspondiente problema inverso asociado es mal condicionado en el sentido descrito en la Sección 2.1. Si bien desde un punto de vista matemático riguroso el mal condicionamiento de un problema inverso no puede ocurrir nunca en dimensión finita, la discretización de un problema mal condicionado resulta en un problema de la forma (5.7) en el que la matriz A posee un número de condición muy grande, lo que torna inestable el proceso de inversión y hace estrictamente necesaria la utilización de algún tipo de regularización.

A continuación presentaremos una serie de resultados numéricos que tienen como objetivo mostrar el desempeño del método de regularización wavelet-espectral descrito en el Capítulo 3, para diferentes tipos de soluciones exactas, con diferentes grados de regularidad, para diferentes grados de mal condicionamiento y distintos niveles de ruido.

Nos proponemos analizar el desempeño del método wavelet-espectral en el problema $T_k f = g^\delta$, donde T_k es el operador de convolución dado en (5.1), discretizado como en (5.6), y resultante en (5.7). Utilizaremos las funciones f_1, f_2 y f_3 que se ilustran en la Figura 5.1 consistentes en funciones suave, constante a trozos y mixta, respectivamente. Asimismo, utilizaremos operadores con distintos grados de mal condicionamiento. En nuestro caso, para operadores de convolución con núcleos gaussianos, el grado de mal condicionamiento está determinado por la desviación estándar del núcleo correspondiente. Por tal motivo consideraremos 4 valores de σ_i ($i = 1, 2, 3, 4$) para la desviación estándar, correspondientes a 4 valores de ℓ como en (5.5) dados por $\ell_1 = 1, \ell_2 = 2, \ell_3 = 3$ y $\ell_4 = 4$.

En las Figuras 5.2 (a), (b), (c) y (d) se muestran las gráficas de f_3 y $T_k f_3$ para estos 4 valores de ℓ . Por otro lado, en la Figura 5.2 (e) se muestran los autovalores de la matriz A correspondientes a las discretizaciones de T_k para los 4 valores de σ . Notar que, como es de esperar para valores de ℓ grandes los autovalores decaen rápidamente a cero indicando un

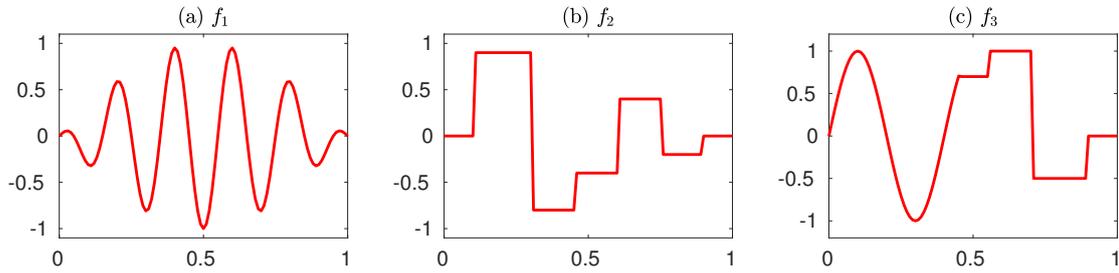


FIGURA 5.1. Soluciones exactas suave, constante a trozos y “mixta”.

fuerte grado de mal condicionamiento. En todos los ejemplos siguientes consideramos un núcleo gaussiano con desviación estándar igual al 3% de $\|f\|_\infty$, es decir $\ell = 3$.

Puesto que el método wavelet-espectral propuesto en el Capítulo 3 permite utilizar distintos métodos espectrales, en esta sección compararemos los métodos wavelet-espectrales obtenidos con Tikhonov Phillips (TP) y Landweber (L) y, además, implementaremos un método de Tikhonov generalizado: el método de penalización por Variación Total (VT). Denotaremos los respectivos métodos de dos pasos con W-TP, W-L y W-VT, respectivamente.

En todos los ejemplos de este capítulo utilizaremos como regla de elección de parámetro el Principio de Discrepancia de Morozov descrito en la Sección 2.5.2, con $\tau = 1.1$. En todos los casos presentaremos las gráficas de las reconstrucciones obtenidas con cada método junto con la solución exacta. Además, a los efectos de cuantificar rigurosamente el desempeño de cada uno de ellos, presentaremos el índice ISNR (Improvement in Signal-to-Noise Ratio) promedio correspondiente a 100 repeticiones del método (asociado a 100 realizaciones diferentes del correspondiente vector de ruido blanco aditivo).

El paso de umbralado wavelet en el método wavelet-espectral introduce nuevos parámetros, que deben ser fijados, previo a la implementación de los mismos. En primer lugar, se debe definir la wavelet a utilizar y el nivel del umbral λ el que, tal como vimos en el Capítulo 3, depende del nivel de ruido. Por otro lado, luego de realizar el paso de umbralado wavelet es necesario re-estimar el nivel de error δ del nuevo dato umbralado. En nuestro caso y a menos que se especifique otra cosa utilizamos la Wavelet Daubechies con cuatro momentos nulos. Los parámetros λ y δ se eligieron de forma empírica comparando distintos valores posibles del umbral y promediando el nuevo nivel de error $\|g - S_\lambda g^\delta\| \leq \delta_\lambda$. Si bien resulta evidente que estos parámetros dependen no solo del nivel de ruido inicial, sino también de las características de la función solución y del operador, para los ejemplos propuestos, utilizamos $\lambda \doteq 2.5 \delta$ y $\delta_\lambda \doteq 0.5 \delta$. Esta elección de tales

parámetros ha resultado en reconstrucciones muy satisfactorias, no es claro como elegir las de manera óptima. Claramente más investigación es necesaria en esta dirección.

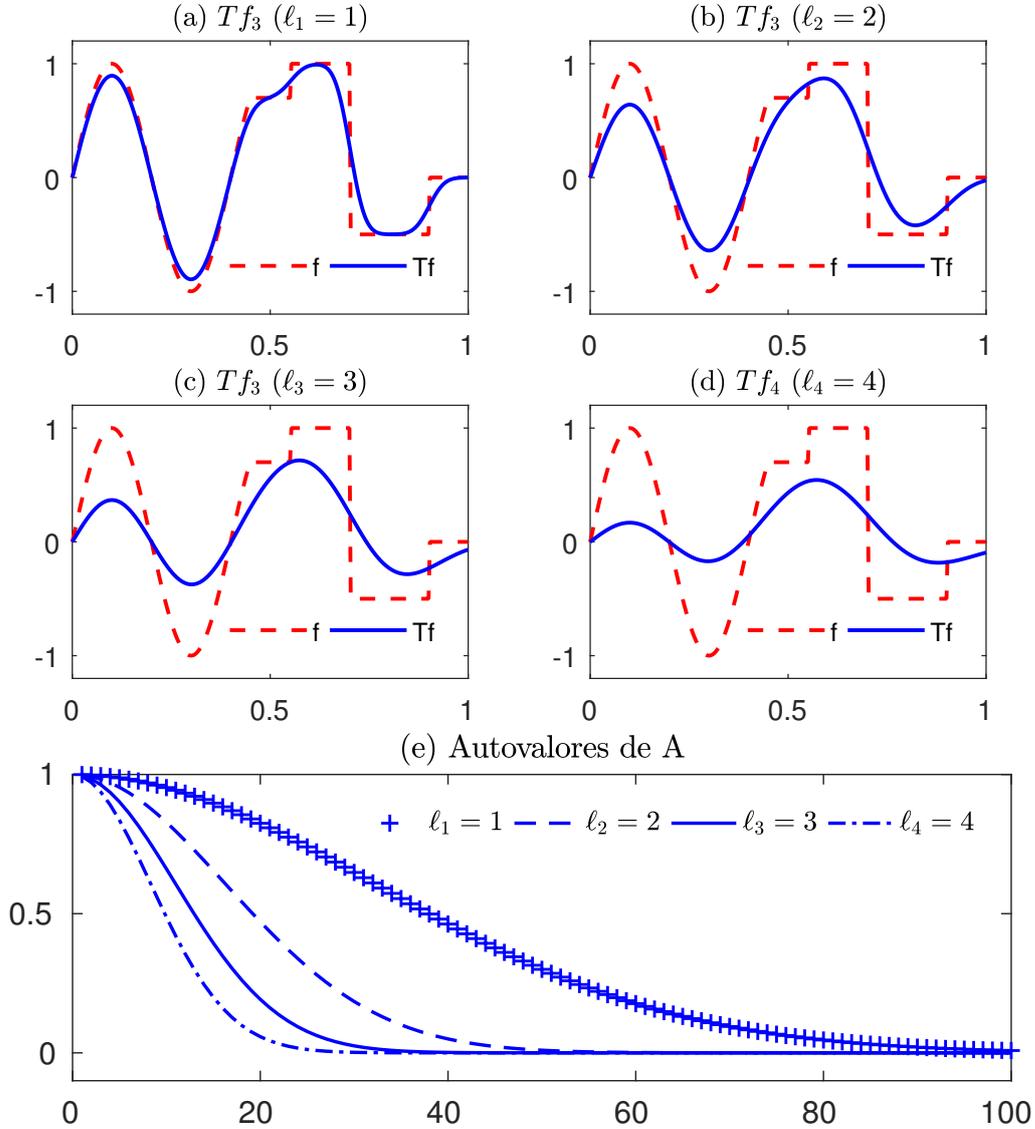


FIGURA 5.2. Gráficas de f_3 y $g = T_k f_3$ para desviaciones estándar del 1% (a), 2% (b), 3% (c) y 4% (d) de $\|f\|_\infty$ y autovalores de la matriz A (e).

A continuación presentamos una serie de ejemplos de restauración de las señales f_1 , f_2 y f_3 , previamente descritas, bajo diferentes condiciones (nivel de ruido, grado de mal condicionamiento, nivel de umbrado, etc.). Para cada una de estas condiciones presentaremos un caso, donde mostraremos varios ejemplos utilizando diferentes funciones solución y modificando los valores de ciertos parámetros. El objetivo de mostrar los resultados de cada uno de estos casos será analizar el desempeño del método wavelet-espectral en comparación con otros métodos clásicos de regularización identificando la influencia de cada uno de los parámetros involucrados.

CASO I:

En este caso nos proponemos analizar el desempeño de los métodos de dos pasos con umbralado wavelet para funciones con diferentes grados de regularidad. Para ello presentamos tres ejemplos utilizando las funciones f_1 , f_2 y f_3 presentadas anteriormente.

Ejemplo I.1. En este ejemplo hemos aplicado el operador T_k ($\sigma = 0.03 \|f\|_\infty$) a la función f_1 , utilizando un nivel de ruido del 5%, es decir $\delta = 0.05 \|g\|_\infty$.

En la Figura 5.3 (a) y (b) se muestran las gráficas de $g^\delta = T_k f_1 + \delta dW$ y $S_\lambda g^\delta$, respectivamente. Observar como el paso de umbralado wavelet reduce significativamente las componentes de alta frecuencia originadas por el ruido.

La Figura 5.4 (a) - (f) muestra las restauraciones obtenidas con los métodos L, W-L, TP, W-TP, VT, W-VT, respectivamente, como así también la solución exacta a restaurar, f_1 . El análisis visual de estas restauraciones permite derivar las siguientes conclusiones:

- (i) El paso previo de descomposición wavelet y umbralado, previo a la regularización, en todos los casos mejora significativamente el desempeño del método.
- (ii) La mejora señalada en el ítem anterior es mucho más notoria para los métodos L y TP los que, como es bien sabido, tienen buen desempeño para el caso de soluciones suaves, como lo es f_1 .
- (iii) El paso de umbralado wavelet previo a la aplicación de la regularización por VT, si bien mejora el desempeño del método, sigue produciendo restauraciones constantes a trozos, característica de dicho método (el llamado “efecto escalera”).

Todas las observaciones precedentes se reflejan claramente en los índices ISNR promedio, calculados para 100 realizaciones (del ruido blanco aditivo) con cada uno de los métodos. Estos se muestran en la Tabla 1.

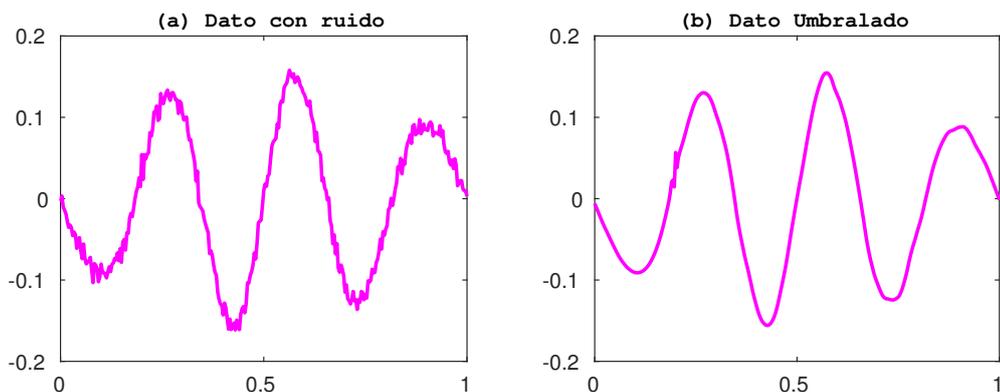


FIGURA 5.3. Dato g^δ (a) y dato umbralado $S_\lambda g^\delta$ (b).

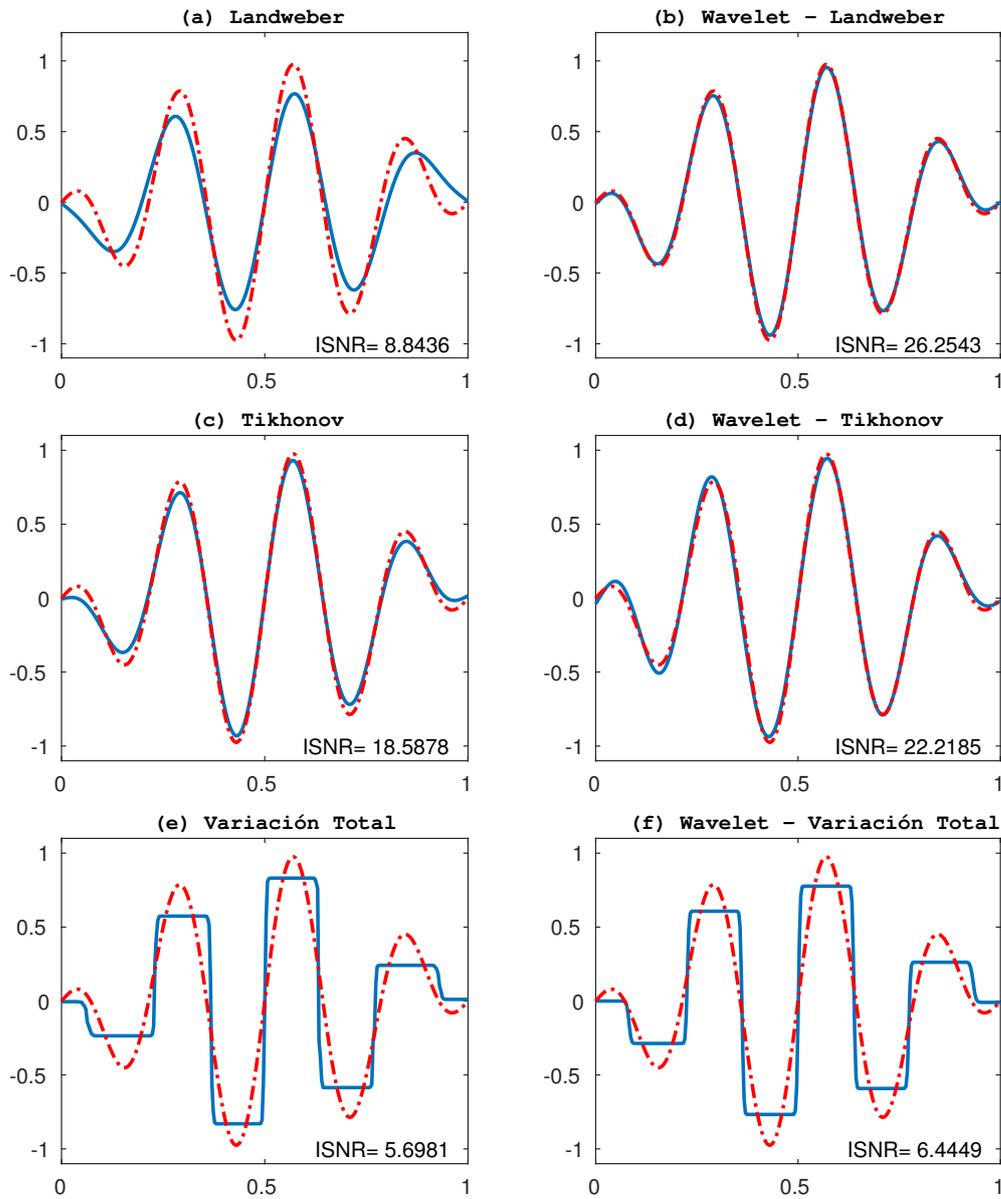


FIGURA 5.4. Restauraciones de f_1 con $\delta = 0.05 \|g\|_\infty$ utilizando los métodos L (a), W-L (b), TP (c), W-TP (d), VT (e) y W-VT (f).

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|---------|-------------------------|---------|
| Landweber | 9.7554 | Wavelet-Landweber | 25.6914 |
| Tikhonov | 16.1468 | Wavelet-Tikhonov | 22.5496 |
| Variación Total | 4.9697 | Wavelet-Variación Total | 6.8548 |

TABLA 1. Índices ISNR promedio correspondientes a 100 restauraciones de cada uno de los métodos del Ejemplo I.1.

Ejemplo I.2. En este ejemplo repetimos el experimento realizado en el Ejemplo I.1 pero ahora utilizando la función constante a trozos f_2 .

En la Figura 5.5 se muestran el dato g^δ y su descomposición wavelet umbralada $S_\lambda g^\delta$. Observar nuevamente como este paso reduce significativamente la influencia de las componentes de alta frecuencia provenientes del ruido aditivo. En la Figura 5.6 se muestran las restauraciones obtenidas con los seis métodos L, W-L, TP, W-TP, VT, W-VT. El análisis visual de las restauraciones nos permite derivar las siguientes conclusiones:

- (i) En todos los casos, el paso de descomposición wavelet seguido de umbralado mejora el desempeño del método de regularización aplicado en el segundo paso.
- (ii) La mejora en el desempeño señalado en el ítem anterior es ahora significativamente más notoria para el método de VT ((e) y (f)). Esto es consistente con el hecho que, como es bien sabido, este método muestra un muy buen desempeño para el caso de soluciones exactas que son constantes a trozos, como es el caso de la función f_2 .
- (iii) Para el caso de los métodos de L y TP ((a), (b), (c) y (d)) el paso wavelet-umbralado, si bien introduce una leve mejora en el desempeño del método, las restauraciones obtenidas siguen siendo relativamente pobres principalmente cerca de los puntos de discontinuidad. Esto se debe a que ambos métodos regularizan el proceso de inversión atenuando significativamente las componentes de alta frecuencia de la solución exacta, las que, como es bien sabido, son estrictamente necesarias para capturar razonablemente discontinuidades.

Todas las observaciones precedentes realizadas mediante el análisis visual de las restauraciones están apoyadas, además, por los índices ISNR promedio correspondientes a 100 restauraciones de cada uno de los seis métodos (correspondientes a 100 realizaciones del ruido aditivo) presentados en la Tabla 2.

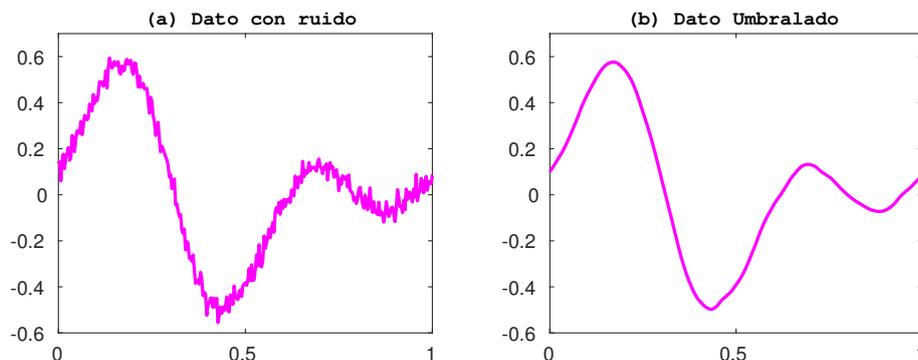


FIGURA 5.5. Dato g^δ (a) y dato umbralado $S_\lambda g^\delta$ (b).

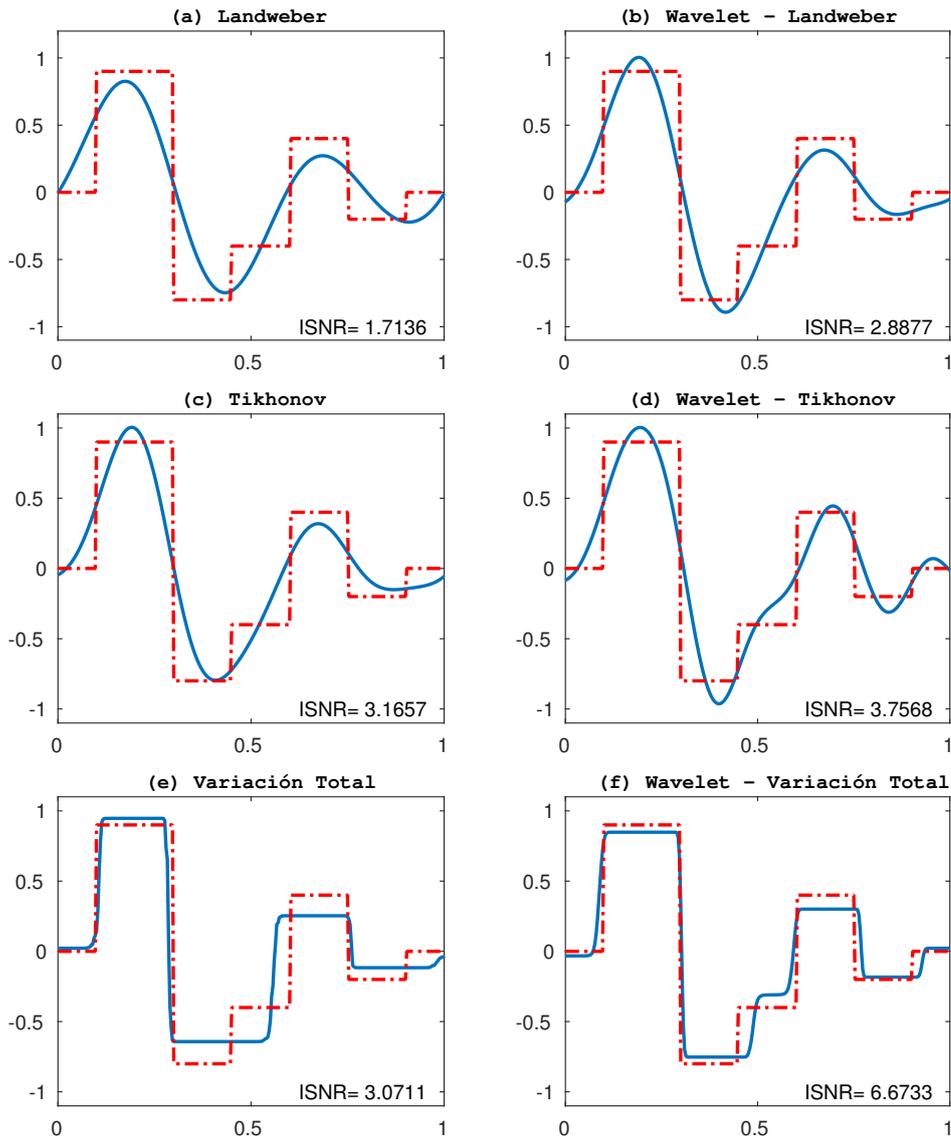


FIGURA 5.6. Restauraciones de f_2 con $\delta = 0.05 \|g\|_\infty$ utilizando los métodos L (a), W-L (b), TP (c), W-TP (d), VT (e) y W-VT (f).

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|--------|-------------------------|--------|
| Landweber | 1.8568 | Wavelet-Landweber | 3.0465 |
| Tikhonov | 2.9611 | Wavelet-Tikhonov | 3.8357 |
| Variación Total | 3.4614 | Wavelet-Variación Total | 6.2795 |

TABLA 2. Índices ISNR promedio correspondientes a 100 restauraciones de cada uno de los métodos del Ejemplo I.2.

Ejemplo I.3. En este ejemplo repetimos los experimentos realizados en los Ejemplos I.1 e I.2 utilizando ahora la solución exacta f_3 , la cual es suave en un intervalo y constante a trozos en su complemento.

En la Figura 5.7 se muestran el dato g^δ y su correspondiente descomposición wavelet umbralada $S_\lambda g^\delta$. En la Figura 5.8 se presentan las restauraciones de f_3 obtenidas con los métodos L, W-L, TP, W-TP, VT y W-VT. Nuevamente el análisis visual de las seis restauraciones nos permite concluir que:

- (i) Al igual que en los ejemplos precedentes la aplicación del paso wavelet seguido de umbralado previo a la regularización, siempre mejora el desempeño del método.
- (ii) Como puede observarse, la mejora introducida por la descomposición wavelet más umbralado es significativamente notoria para los métodos L y TP, en la región donde f_3 es suave.
- (iii) Para el método VT la mejora que introduce este paso previo, como es de esperar, es particularmente significativa en la región en la que f_3 es constante a trozos. No obstante ello también se observa un mejor desempeño del método en la región donde f_3 es suave.

Todas las observaciones anteriores están apoyadas por los índices ISNR promedio correspondientes a 100 restauraciones utilizando de cada uno de los seis métodos (correspondientes a 100 realizaciones del ruido aditivo) presentados en la Tabla 3.

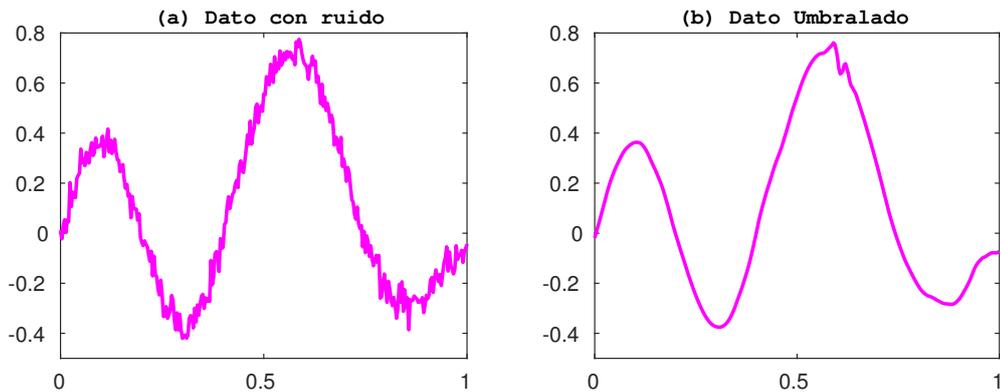


FIGURA 5.7. Dato g^δ (a) y dato umbralado $S_\lambda g^\delta$ (b).

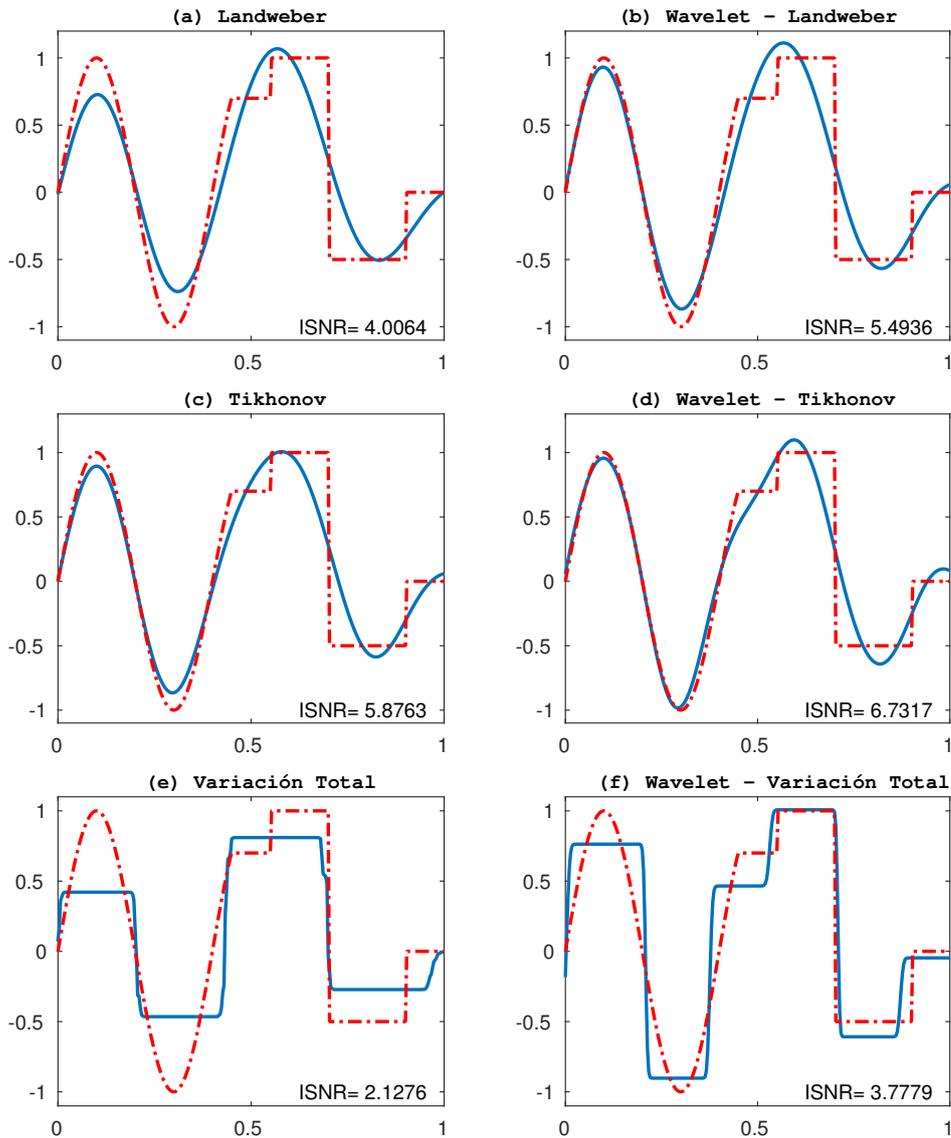


FIGURA 5.8. Restauraciones de f_3 con $\delta = 0.05 \|g\|_\infty$ utilizando los métodos L (a), W-L (b), TP (c), W-TP (d), VT (e) y W-VT (f).

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|--------|-------------------------|--------|
| Landweber | 3.9756 | Wavelet-Landweber | 5.7425 |
| Tikhonov | 5.4711 | Wavelet-Tikhonov | 6.3784 |
| Variación Total | 2.3563 | Wavelet-Variación Total | 3.5712 |

TABLA 3. Índices ISNR promedio correspondientes a 100 restauraciones de cada uno de los métodos del Ejemplo I.3.

CASO II:

Nos proponemos ahora analizar la influencia del nivel de ruido en el desempeño del método de dos pasos. Para ello consideraremos dos ejemplos en los que presentamos las restauraciones del problema $T_k f = g^\delta$ para $\sigma = 0.03 \|f\|_\infty$ con niveles de ruido δ del 1, 5, 10 y 20 % de $\|g\|_\infty$.

Ejemplo II.1. En este ejemplo hemos aplicado el operador T_k (con $\sigma = 0.03 \|f\|_\infty$) a la función f_1 . Con el objetivo de comparar el efecto de los diferentes niveles de ruido utilizamos solamente el método TP, el que, como es bien sabido, tiene un buen desempeño para el caso de funciones suaves como lo es f_1 . La Figura 5.9 muestra el dato g^δ para estos cuatro niveles de ruido y la Figura 5.10 presenta las restauraciones obtenidas con los métodos TP y W-TP para cada nivel de ruido. El análisis visual de las ocho restauraciones nos permite concluir que:

- (i) La aplicación del paso wavelet seguido de umbralado previo a la regularización siempre mejora el desempeño del método espectral, para todos los niveles de error considerados.
- (ii) En particular, la mejora introducida por la descomposición wavelet-umbralado es más notoria para niveles altos de ruido. Observar que, por ejemplo, para un nivel de ruido del 1 % esta mejora es solo del 1.19 %, mientras que para un nivel de ruido del 20 % supera el 38 %.
- (iii) Es importante notar que el desempeño del método de dos pasos, para un nivel de ruido del 20 % es comparable (y aún mejor) con el que se obtiene sin aplicar el paso previo wavelet-umbralado para un nivel de ruido del 10 % ((e) y (h)).

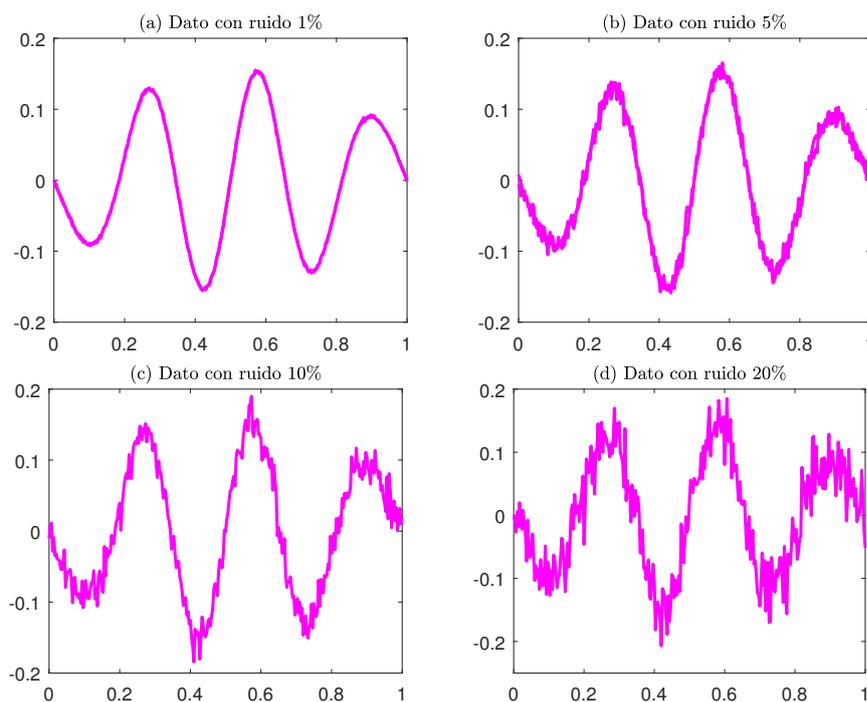


FIGURA 5.9. Dato g^δ para los cuatro niveles de ruido del Ejemplo II.1.

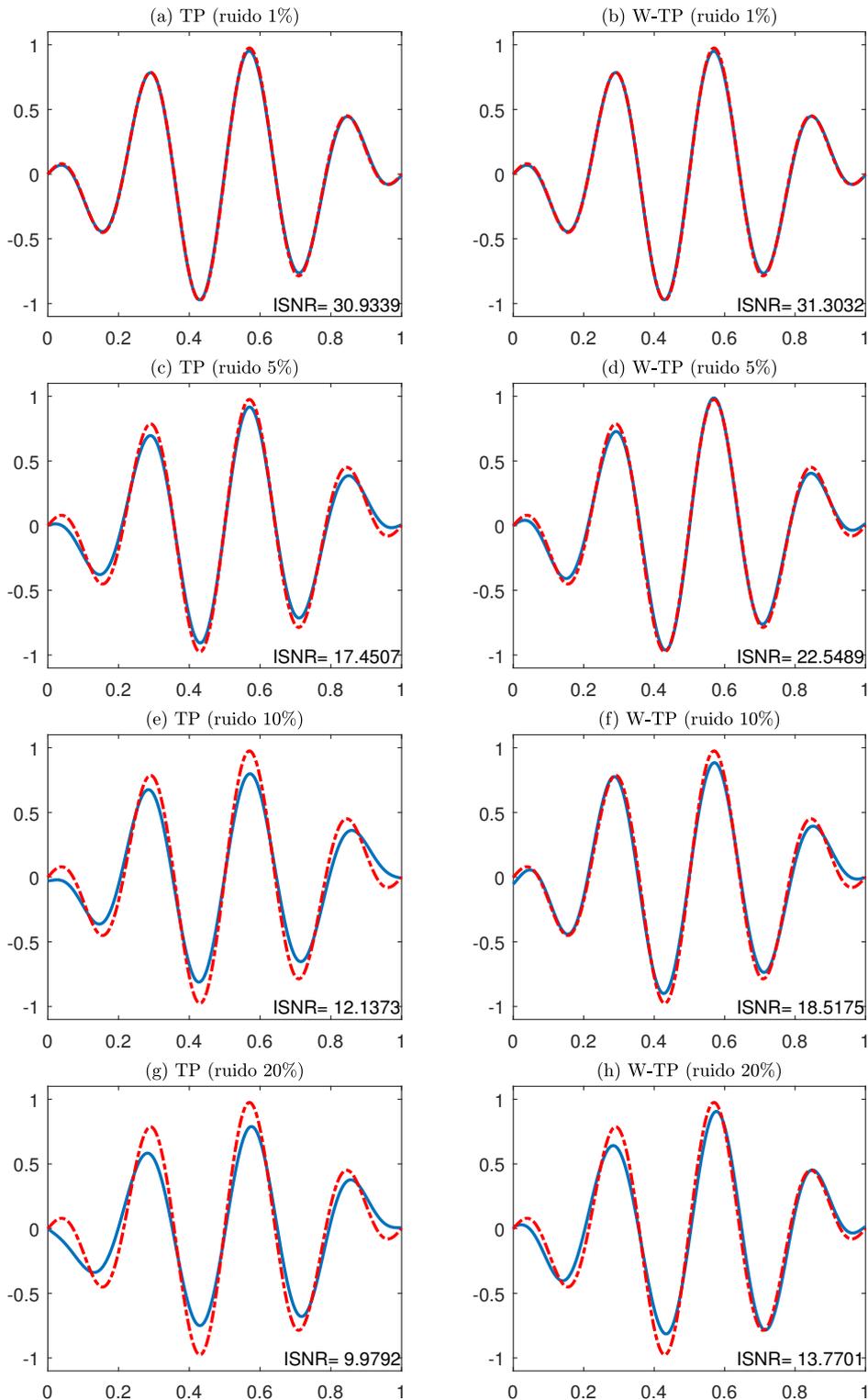


FIGURA 5.10. Restauración de f_1 , utilizando los métodos TP y W-TP con niveles de ruido del 1, 5, 10 y 20% de $\|g\|_\infty$.

Ejemplo II.2. En este ejemplo repetimos el experimento realizado en el ejemplo anterior pero ahora para la función f_2 . En este caso, dado que f_2 es una función constante a trozos realizamos las restauraciones con el método de regularización VT.

En la Figura 5.11 se presentan los correspondientes datos g^δ mientras que en la Figura 5.12 se muestran las restauraciones obtenidas con los métodos VT y W-VT para cada uno de los cuatro niveles de ruido. Nuevamente el análisis visual nos permite confirmar que la aplicación del paso previo wavelet-umbralado mejora el desempeño del método, en todos los casos. Por otro lado, esta mejora es significativamente notoria para niveles de ruido altos. Observar en particular que la aplicación, para un nivel de ruido dado, del paso wavelet-umbralado permite obtener desempeños similares a los que se obtendrían sin la aplicación de dicho paso, con la mitad de ese ruido (comparar las gráficas (c) y (f), (e) y (h) de la Figura 5.12).

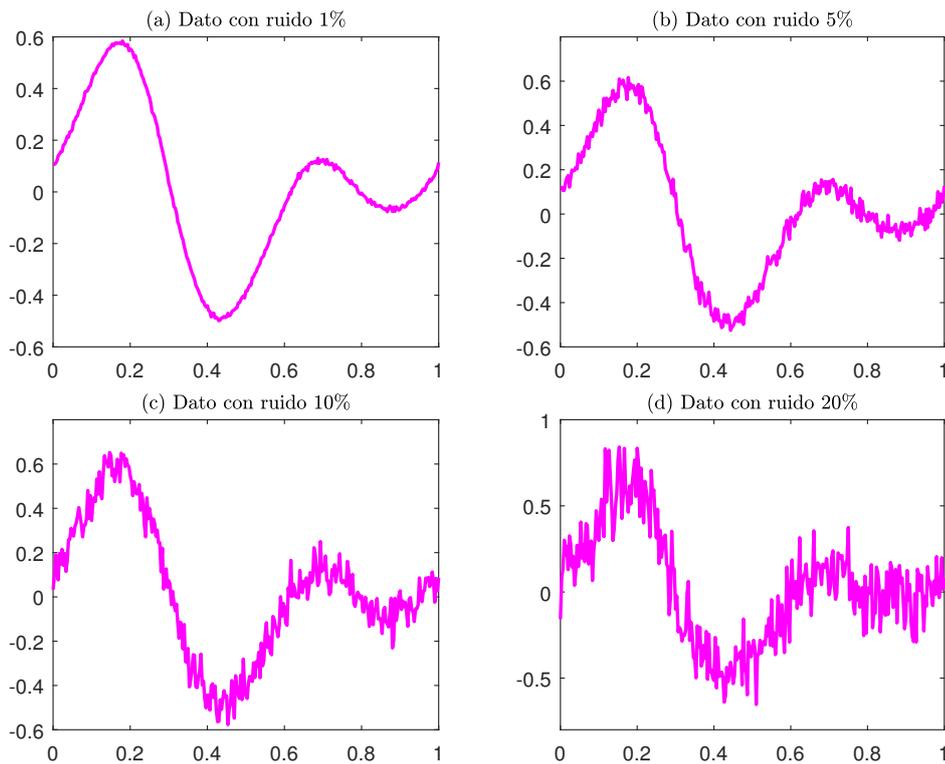


FIGURA 5.11. Dato g^δ para los cuatro niveles de ruido del Ejemplo II.2.

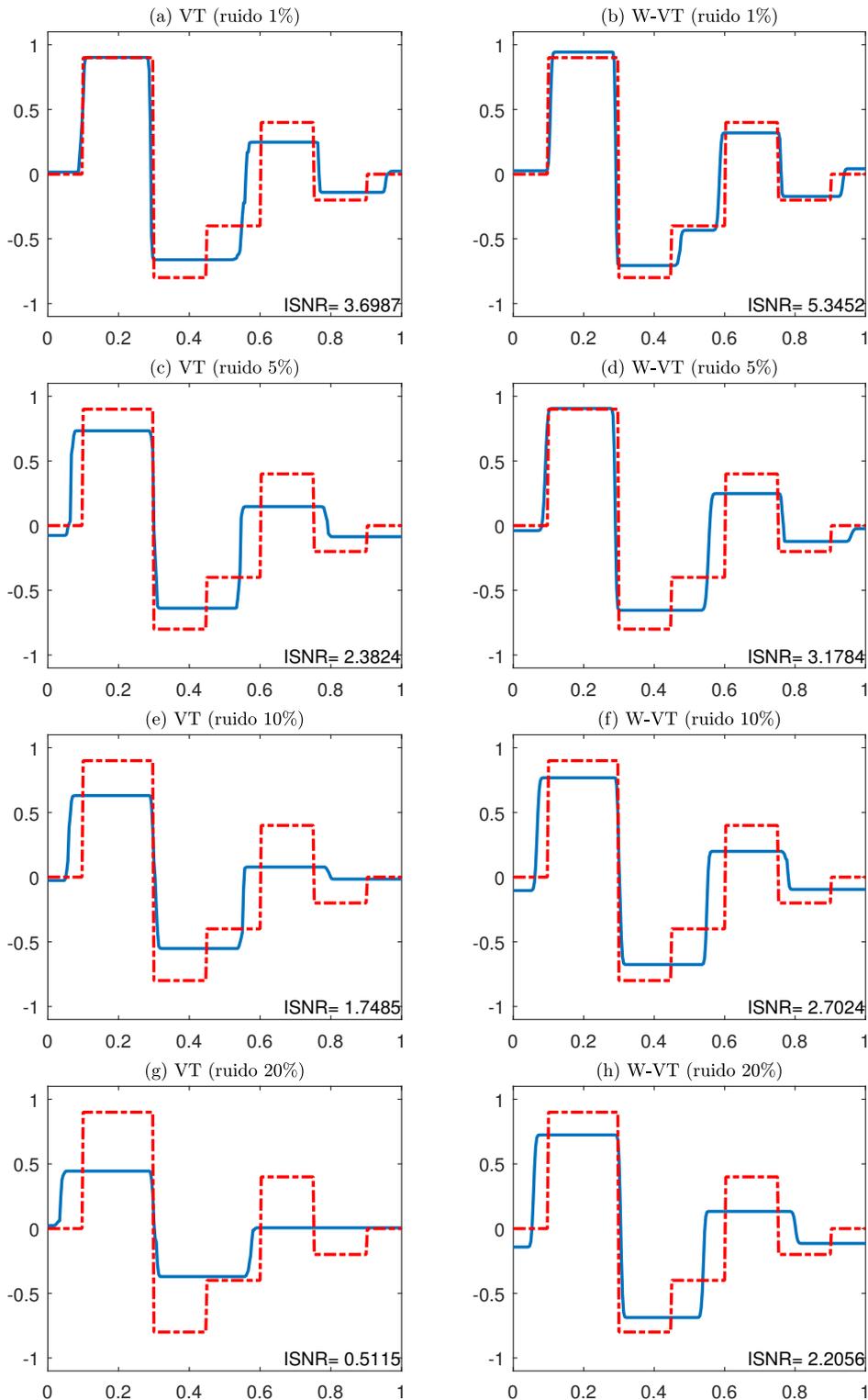


FIGURA 5.12. Restauración de f_2 , utilizando los métodos TP y W-TP con niveles de ruido del 1, 5, 10 y 20% de $\|g\|_\infty$.

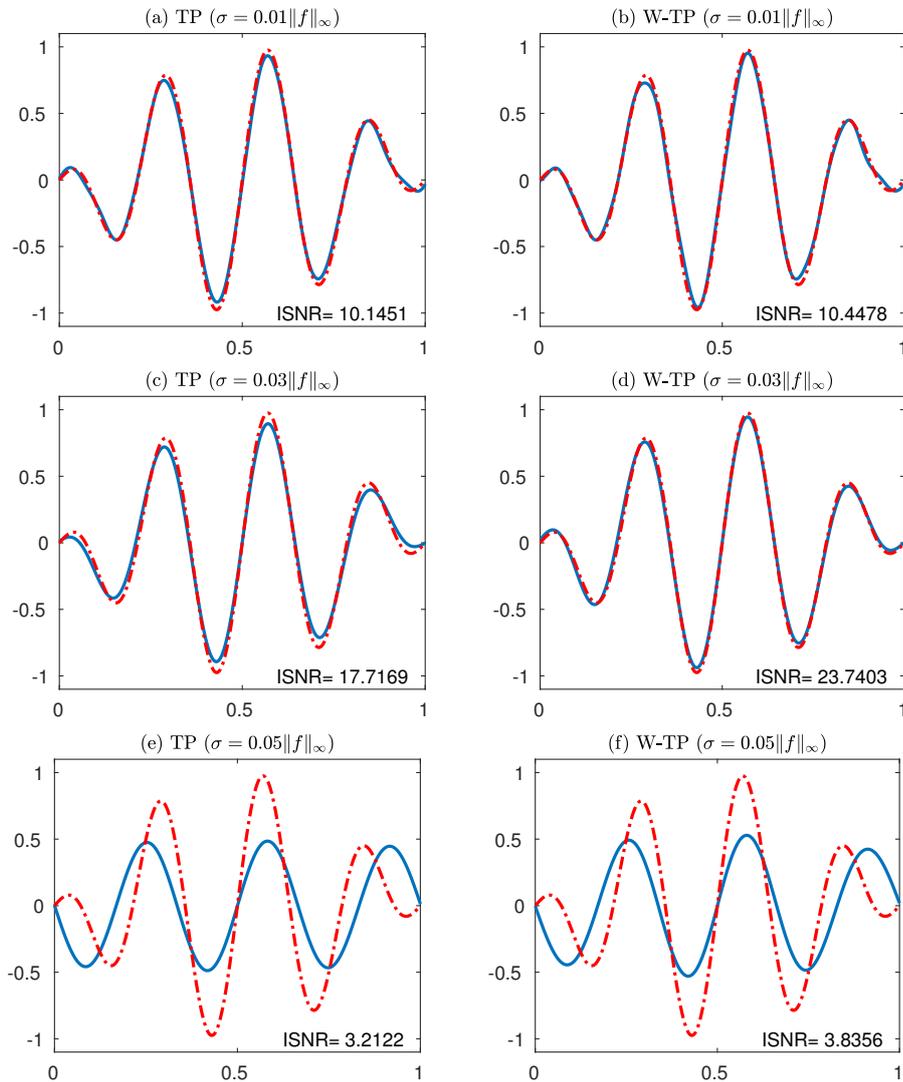


FIGURA 5.13. Restauración de f_1 , utilizando los métodos de TP y W-TP con nivel de ruido $\delta = 0.05\|g\|_\infty$ y σ del 1, 3 y 5% de $\|f\|_\infty$.

CASO III:

En este caso nos proponemos analizar el desempeño del método de dos pasos con umbralado wavelet para problemas inversos con distintos grados de mal condicionamiento. Para ello presentamos dos ejemplos, correspondientes a restauraciones de las funciones f_1 y f_2 , utilizando diferentes valores de σ (1, 3 y 5% de $\|f\|_\infty$) para el operador T_k , con nivel de ruido fijo del 5%.

Ejemplo III.1 En este ejemplo hemos aplicado el operador T_k a la función f_1 . Dadas las características de la solución exacta, nuevamente realizamos las restauraciones utilizando el método TP.

La Figura 5.13 muestra las restauraciones obtenidas con los métodos TP y W-TP para cada valor de σ considerado. Analizando esta figura podemos ver que, al igual que en todos los ejemplos anteriores, la aplicación del paso previo wavelet-umbralado siempre resulta en una mejora

en el desempeño del método espectral aplicado a continuación. Si bien sería necesario un análisis más profundo y detallado, parece razonable conjeturar que tal mejora es significativamente superior para valores intermedios de σ correspondientes a operadores con grado de mal condicionamiento intermedio. Notar que mientras que para $\sigma = 0.01\|f\|_\infty$ y $\sigma = 0.05\|f\|_\infty$ las mejoras en el ISNR son del 3.9% y 19.3%, respectivamente, para $\sigma = 0.03\|f\|_\infty$ la correspondiente mejora es del 34%.

Ejemplo III.2 Aquí repetimos el experimento del ejemplo anterior pero ahora para la solución exacta f_2 y el método VT. La Figura 5.14 muestra las restauraciones obtenidas con los métodos VT y W-VT para cada valor de σ utilizado. Aquí, nuevamente valen todas las observaciones realizadas en el ejemplo anterior. En particular la mejora resultante de la aplicación del paso wavelet-umbralado refuerza también la conjetura que realizamos en el Ejemplo III.1.

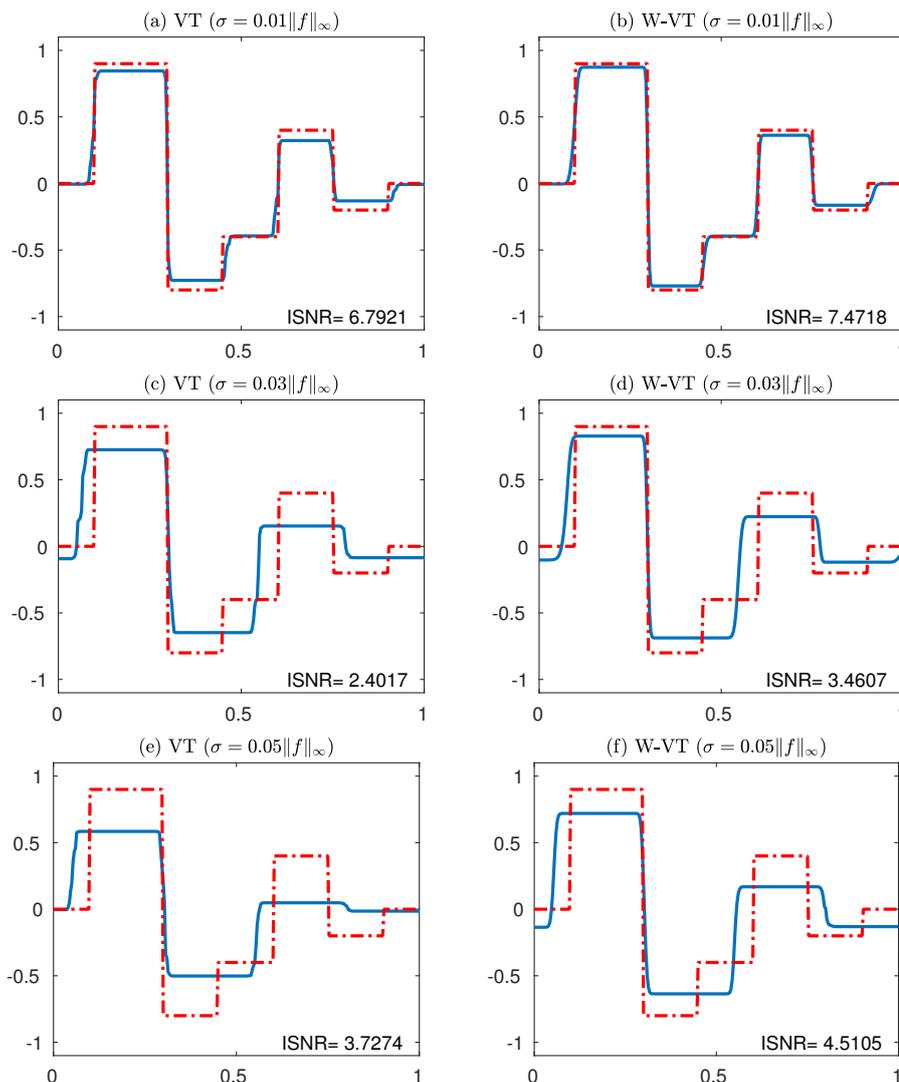


FIGURA 5.14. Restauración de f_2 , utilizando los métodos de VT y W-VT con nivel de ruido $\delta = 0.05\|g\|_\infty$ y σ del 1, 3 y 5% de $\|f\|_\infty$.

CASO IV:

En este último caso, nos proponemos analizar el efecto que tiene, en el desempeño del método de dos pasos, el valor del umbral λ . Para ello hemos aplicado al dato g^δ , obtenido a partir del operador T_k con $\sigma = 0.03 \|f\|_\infty$ y $\delta = 0.05 \|g\|_\infty$, el paso wavelet-umbralado con tres niveles de umbral diferentes ($\lambda = \delta$, $\lambda = 2.5 \delta$ y $\lambda = 5 \delta$).

Ejemplo IV.1. En este ejemplo consideramos la restauración de la función f_1 y el método de regularización TP.

En la Figura 5.15 (a) y (b) se muestran las gráficas del dato exacto g y el dato con ruido g^δ , respectivamente. Las gráficas restantes en dicha figura presentan, de a pares, el dato umbralado y la correspondiente restauración con el método W-TP para cada uno de los valores de λ .

Ejemplo IV.2. En este último ejemplo de la sección repetimos el experimento realizado en el Ejemplo IV.1, ahora considerando como solución exacta a la función f_2 y el método de regularización VT.

En la Figura 5.16 (a) y (b) se presentan las gráficas del dato exacto g y el dato con ruido g^δ , respectivamente. Las gráficas restantes, dispuestas en pares, muestran el dato umbralado y la restauración con el método W-VT correspondiente para cada valor de λ .

Analizando los dos ejemplos anteriores observamos que la mejora resultante de la aplicación del paso previo wavelet-umbralado, es más notoria para valores intermedios del umbral. Eso es bastante razonable puesto que, como puede observarse en las gráficas correspondientes, valores de λ pequeños dejan pasar componentes de alta frecuencia en el dato g^δ provenientes del ruido (Figuras 5.15 (c) y 5.16 (c)), mientras que valores de λ altos tienden a desprestigiar información importante del dato exacto g (Figuras 5.15 (g) y 5.16 (g)).

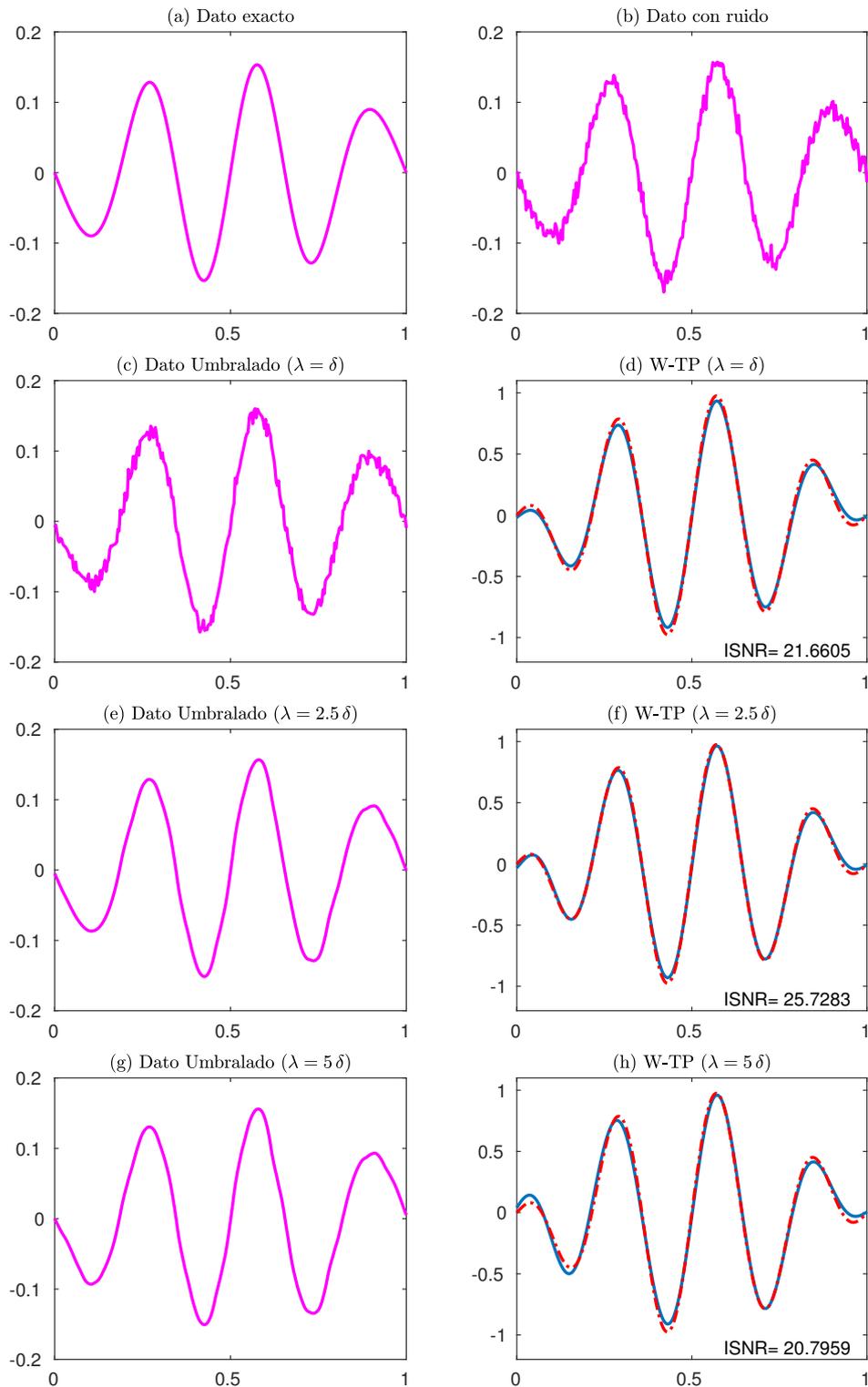


FIGURA 5.15. Dato exacto g (a), dato con ruido g^δ (b) y datos umbralados con la respectiva restauración de f_1 utilizando el método de W-TP para los umbrales $\lambda = \delta$, $\lambda = 2.5 \delta$ y $\lambda = 5 \delta$.

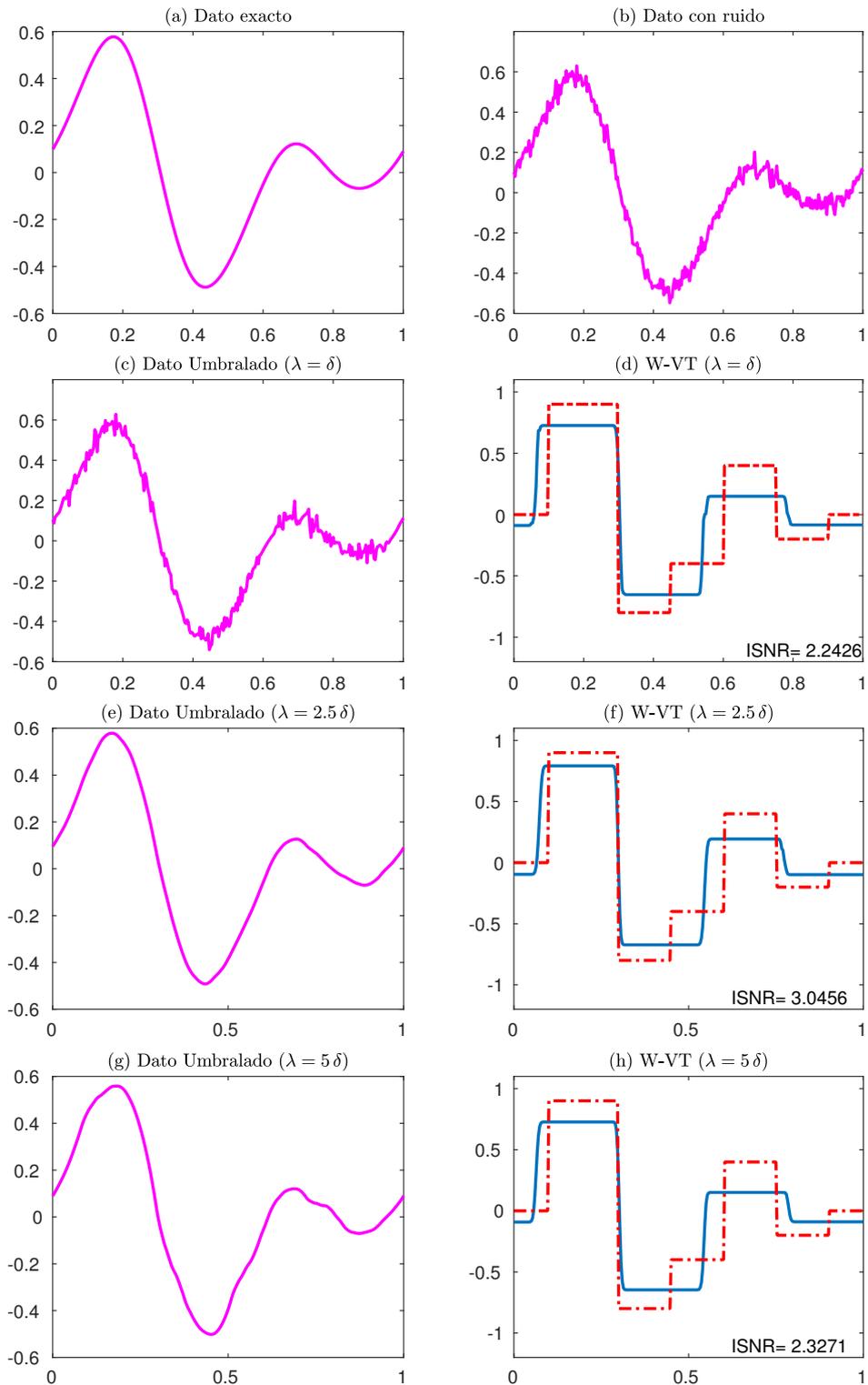


FIGURA 5.16. Dato exacto g (a), dato con ruido g^δ (b) y datos umbralados con la respectiva restauración de f_2 utilizando el método de W-VT para los umbrales $\lambda = \delta$, $\lambda = 2.5 \delta$ y $\lambda = 5 \delta$.

5.1.2. Deconvolución 2D: restauración de imágenes

En esta sección presentaremos algunos ejemplos de problemas inversos mal condicionados en restauración de imágenes digitales. A partir de los mismos, estudiaremos el desempeño del método de dos pasos wavelet-espectral desarrollado en el Capítulo 3, para el caso de problemas 2D.

La restauración de imágenes es un proceso que estudia diversos métodos que permiten recuperar una imagen original a partir de observaciones degradadas de la misma. Se diferencia de las técnicas de mejoramiento de imágenes en que estas últimas buscan manipular una imagen con el único objetivo de producir resultados “más agradables” para el observador y no suponen la existencia de una imagen original o fuente. Tal como mencionamos en la sección anterior, el problema originado con el Telescopio Espacial Hubble ([2]) es quizás el caso más conocido de un ejemplo concreto que señala la necesidad de restaurar imágenes a partir de versiones degradadas de las mismas.

El proceso de degradación (*blurring*) de una imagen puede modelarse satisfactoriamente a través de una ecuación integral de Fredholm de primera clase del tipo

$$T_k f(x, y) \doteq \int_0^1 \int_0^1 k(x, y, x', y') f(x', y') dx' dy' = g(x, y), \quad (5.8)$$

donde f y g son las imágenes original y degradada, respectivamente, y el núcleo k recibe el nombre de función de dispersión puntual (*point spread function*, PSF). Si bien existen varios problemas inversos asociados a la ecuación (5.8), el más común consiste en reconstruir o aproximar la función f a partir del conocimiento de la PSF y de la función g . Es importante señalar, sin embargo, que en la mayoría de los problemas concretos la función g se conoce solo parcialmente, en algún conjunto discreto de puntos y/o contaminada por ruido proveniente de diversas fuentes. Como lo señalamos en la Sección 5.1.1 en algunos problemas más complejos la PSF es también desconocida o solo conocida parcialmente, en cuyo caso los problemas asociados reciben el nombre de problemas de deconvolución ciega. Por el Teorema 4.54, sabemos que, bajo condiciones bastante generales sobre la PSF, el operador integral del lado izquierdo de (5.8) es compacto y por lo tanto reconstruir f es, en efecto, un problema mal condicionado.

En el problema que abordaremos a continuación, a los efectos de evaluar el desempeño del método de dos pasos wavelet-espectral desarrollado en el Capítulo 3, nos restringiremos al caso en que la PSF es conocida. Más precisamente, consideraremos el caso en que el núcleo k es

“separable” y de tipo convolución, es decir,

$$k(x, y, x', y') = \kappa(x - x') \omega(y - y'),$$

donde κ y ω son funciones reales. Esto significa que todos los puntos de la imagen poseen la misma degradación y que dicha degradación puede “separarse” en componentes vertical y horizontal. Existen muchas situaciones en las cuales un sistema de captación de imágenes puede favorecer la degradación de la calidad de las imágenes adquiridas. En general estas situaciones son el resultado de fuentes externas o internas de perturbación en el sistema de captación, enfoque defectuoso o degradaciones introducidas por el mismo medio de captación. A continuación se introducen algunos de los modelos más simples de degradación con sus correspondientes PSF asociadas [28]:

- ★ **Degradación por movimiento:** se genera cuando la cámara se mueve en el momento de captar la imagen o cuando se mueve el objeto mientras que la cámara está fija. Si el tiempo en que transcurre el movimiento es pequeño, se puede considerar que el mismo es lineal. Cuando el movimiento es horizontal, entonces la PSF asociada a este modelo viene dada por

$$k(x, y, x', y') = h_L(x - x') = \begin{cases} (2L)^{-1}, & \text{si } |x - x'| \leq L \\ 0, & \text{en otro caso} \end{cases}$$

donde L es un parámetro que cuantifica la degradación de la imagen. De manera similar, si el movimiento es vertical la PSF está dada por

$$k(x, y, x', y') = h_L(y - y') = \begin{cases} (2L)^{-1}, & \text{si } |y - y'| \leq L \\ 0, & \text{en otro caso} \end{cases}.$$

- ★ **Degradación por desenfoque uniforme o fuera de foco:** esta degradación se genera cuando la lente de la cámara está fuera de foco. Si la apertura de la cámara es circular, la imagen generada por cualquier fuente puntual será un pequeño disco llamado “*círculo de desenfoque*”. En este caso, un modelo razonable para la PSF es el siguiente

$$k(x, y, x', y') = \begin{cases} (\pi R^2)^{-1}, & \text{si } \sqrt{(x - x')^2 + (y - y')^2} \leq R \\ 0, & \text{en otro caso} \end{cases}$$

donde R es el radio del círculo de desenfoque. Es importante notar que esta función no es separable en las variables $x - x'$ e $y - y'$.

★ **Degradación de tipo turbulencia atmosférica:** se genera, por ejemplo, en astronomía debido al cambio en las condiciones de refracción de las diferentes capas de la atmósfera terrestre. La captación de imágenes realizadas por todos los telescopios terrestres son generalmente modeladas con ecuaciones de la forma (5.8) con PSFs de tipo turbulencia atmosférica. Un modelo comunmente utilizado para la PSF en este caso consiste en una función de densidad de probabilidad gaussiana bidimensional de la forma

$$k(x, y, x', y') = \frac{1}{2\pi\sigma_h\sigma_v} e^{-\frac{1}{2}\left(\frac{x-x'}{\sigma_h}\right)^2 - \frac{1}{2}\left(\frac{y-y'}{\sigma_v}\right)^2}, \quad (5.9)$$

donde las desviaciones estándar σ_h y σ_v caracterizan las degradaciones en las direcciones horizontal y vertical, respectivamente. Observar que la PSF de turbulencia atmosférica es separable.

El proceso de discretización del modelo continuo (5.8) consiste esencialmente en pixelar la función bidimensional f y a partir de ello discretizar el operador T . La imagen se discretiza de manera usual definiendo adecuadamente los píxeles, lo que transforma la imagen (continua) f en una matriz I de orden $n \times m$ donde m y n son la cantidad de píxeles en los sentidos vertical y horizontal, respectivamente. En todos los ejemplos que se presentan en esta sección, utilizaremos $n = m$.

En el caso en el que estamos interesados, en los que el operador T_k posee núcleo separable, en la discretización se utiliza esta propiedad dividiendo el proceso en dos pasos. En primer lugar, se utiliza un núcleo unidimensional para convolucionar la imagen solo en la dirección horizontal o vertical y, en un segundo paso, el núcleo unidimensional restante se usa para realizar la convolución en la otra dirección. El efecto resultante es el mismo que convolucionar con el núcleo bidimensional en un solo paso, aunque requiere menos esfuerzo computacional. Para el modelo discretizado, este proceso resulta en una matriz $A \doteq A_h \otimes A_v$, que discretiza al operador T_k donde A_h y A_v son las matrices resultantes de las discretizaciones correspondientes a la convolución con cada uno de los núcleos unidimensionales en las que se separa el núcleo de convolución k y “ \otimes ” es el producto de Kroenecker. De esta forma el modelo continuo (5.8) se puede representar en dimensión finita como

$$AI = (A_h \otimes A_v) I = B, \quad (5.10)$$

donde I y B son las discretizaciones de f y g , respectivamente. Para más detalle ver [20].

En los ejemplos que siguen utilizaremos degradación por turbulencia atmosférica dada en la ecuación (5.9) con $\sigma_h = \sigma_v \doteq \sigma$.

En la Figura 5.17 se muestran una imagen original (35×35 píxeles) (a) e imágenes degradadas de la misma para diferentes valores de σ , $\sigma = 0.02\|I\|_\infty$ (b), $\sigma = 0.04\|I\|_\infty$ (c) y $\sigma = 0.06\|I\|_\infty$ (d).

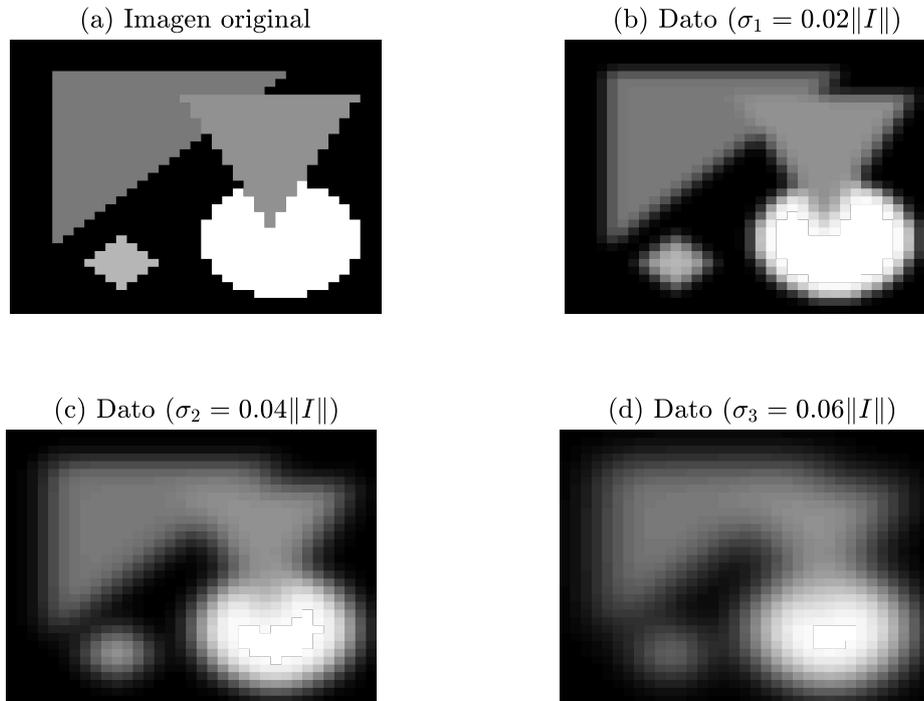


FIGURA 5.17. Imágenes original (a) y degradadas, para desviaciones estándar del 2 (b), 4 (c) y 6% (d) de $\|I\|_\infty$.

A continuación presentamos algunos ejemplos de restauración de imágenes, utilizando el método de dos pasos desarrollado en el Capítulo 3, para diferentes grados de mal condicionamiento y niveles de ruido aditivo. Utilizaremos dos imágenes con distinto grado de suavidad y distinta resolución. El objetivo de mostrar los resultados de cada uno de estos ejemplos, divididos en dos casos, será estudiar el desempeño del método wavelet-espectral en comparación con otros métodos clásicos de regularización analizando la influencia de cada uno de los parámetros considerados.

CASO I:

En este caso nos proponemos analizar el efecto que tiene el nivel de ruido en el desempeño del método de dos pasos con umbralado wavelet para el problema inverso de deblurring. La imagen original I será la presentada en la Figura 5.17 (a) y definiremos $\sigma = 0.04\|I\|_\infty$.

Ejemplo I.1. En este ejemplo hemos agregado al dato $B = AI$, obtenido con $\sigma = 0.04\|I\|_\infty$, ruido blanco aditivo de nivel $\delta = 0.02\|B\|_\infty$.

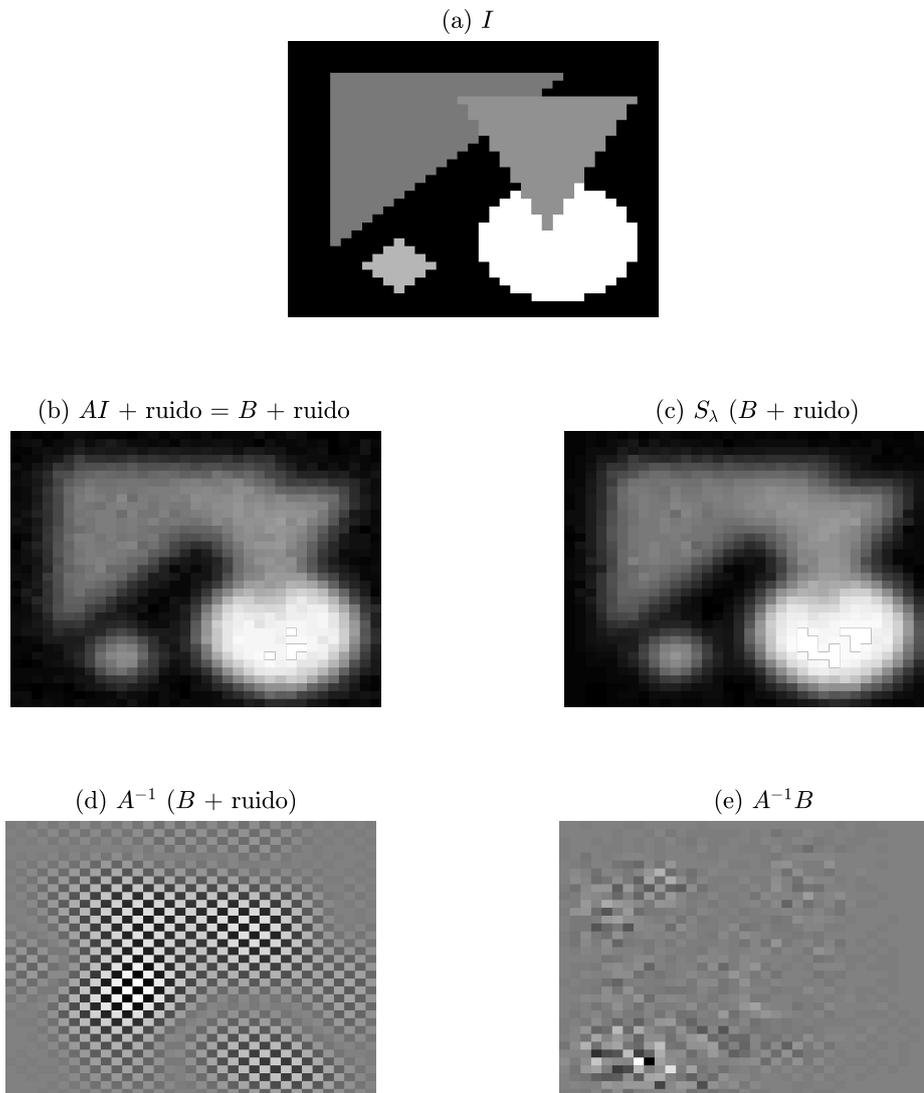


FIGURA 5.18. Imagen original (a), dato con ruido (b), dato con ruido + umbralado wavelet (c), imagen obtenida por inversión del dato con ruido (d) e imagen obtenida por inversión del dato limpio B (e).

En la Figura 5.18 se muestran la imagen original I (a), el dato con ruido (b), el dato con ruido seguido de umbralado (c), la imagen obtenida por inversión del dato con ruido (d) y la imagen obtenida por inversión del dato limpio B (e). El análisis visual de dicha figura nos permite concluir que:

- (i) El proceso de umbralado sobre el dato con ruido claramente elimina componentes de alta frecuencia en la imagen, lo que se manifiesta en bordes más suaves y difuminados (Figura 5.18 (c)).

- (ii) La aplicación directa del proceso de inversión sobre el dato (con ruido), negligiendo regularización, no puede reproducir absolutamente ningún detalle de la imagen original (ver Figura 5.18 (d)).
- (iii) Es importante señalar además que, en contra de lo que se cree usualmente, la incapacidad de reconstruir la imagen original por inversión directa que señalamos en el ítem anterior no es consecuencia del ruido sino pura y exclusivamente del mal condicionamiento del operador involucrado. Esto puede verse claramente en la Figura 5.18 (e), en la que se presenta la imagen obtenida al aplicar A^{-1} al dato limpio (sin ruido) B .

Todo lo anterior, una vez más señala claramente la necesidad de introducir regularización en el proceso de inversión en este tipo de problemas.

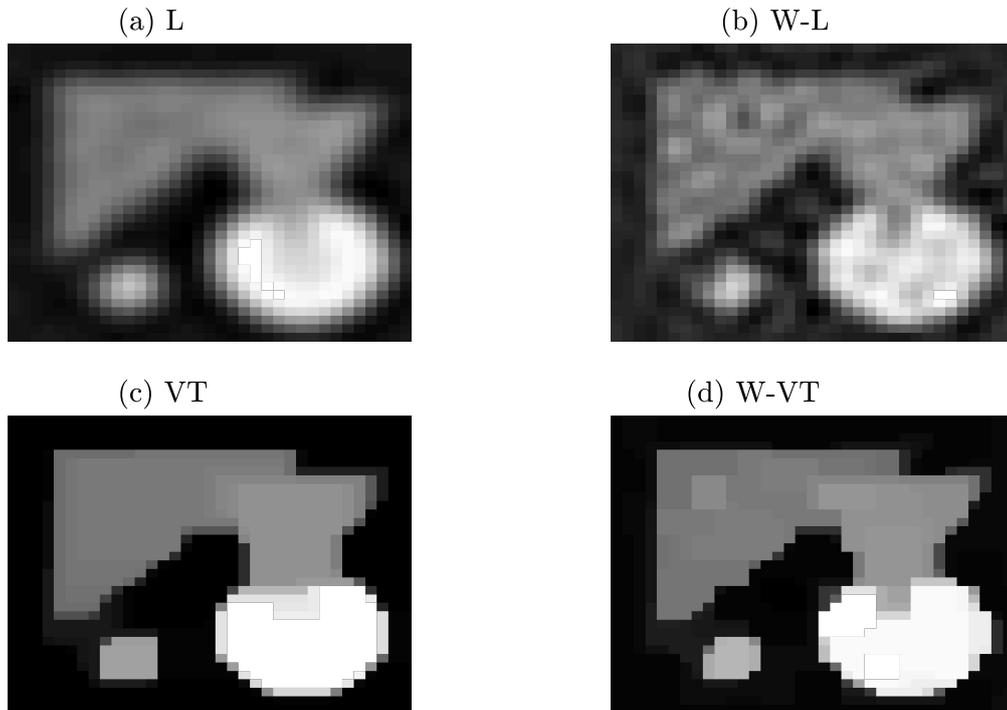


FIGURA 5.19. Restoraciones de I utilizando los métodos L (a), W-L (b), VT (c) y W-VT (d) para un dato con ruido del 2%.

Las restauraciones obtenidas con los métodos L, W-L, VT y W-VT se presentan en la Figura 5.19. Notemos en primer lugar que, para los dos métodos de regularización utilizados, el paso wavelet seguido de umbralado mejora el desempeño de ambos. Por otro lado, dadas las características de la imagen (constante por regiones) no es sorprendente que el método W-VT produzca la mejor restauración.

Todas las observaciones realizadas para este ejemplo son consistentes con los respectivos ISNRs que presentamos en la Tabla 4.

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|--------|-------------------------|--------|
| Landweber | 1.2332 | Wavelet-Landweber | 1.8860 |
| Variación Total | 2.6548 | Wavelet-Variación Total | 3.1805 |

TABLA 4. Índices ISNR correspondientes a las restauraciones de la Figura 5.19.

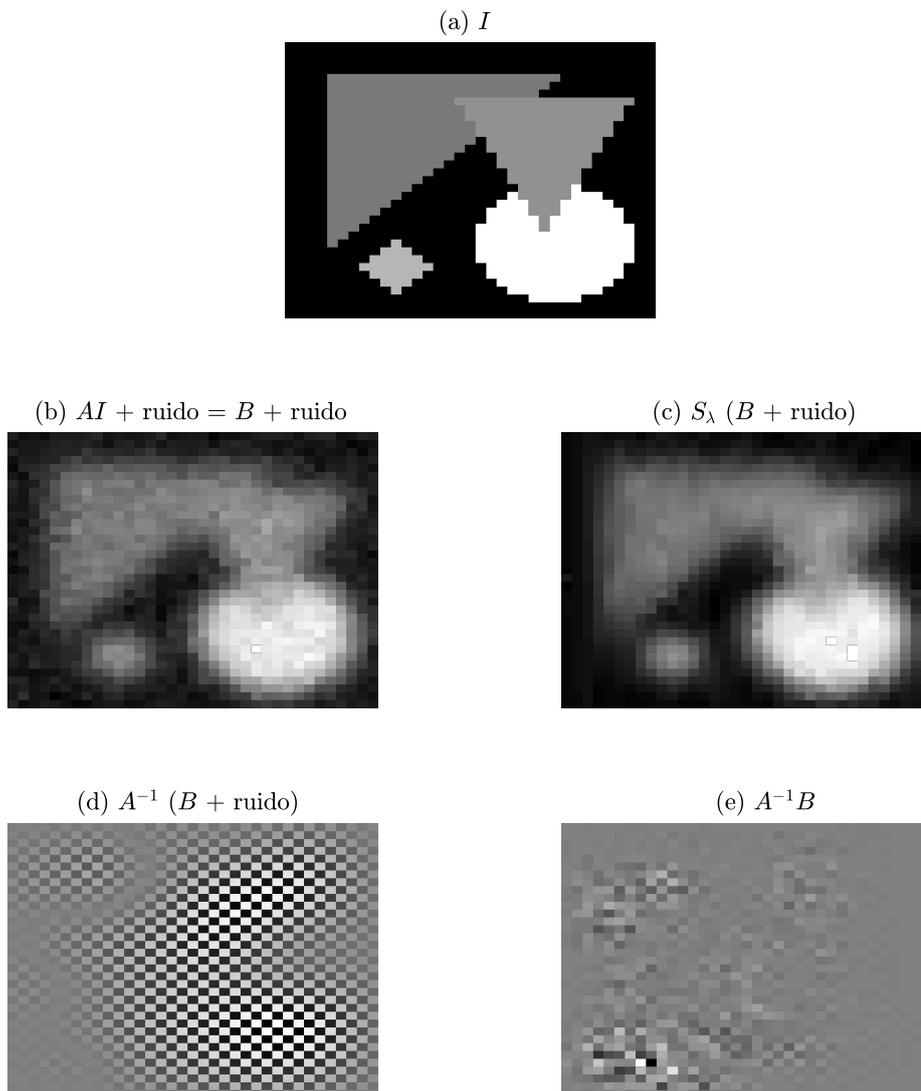


FIGURA 5.20. Imagen original (a), dato con ruido (b), dato con ruido + umbrado wavelet (c), imagen obtenida por inversión del dato con ruido (d) e imagen obtenida por inversión del dato limpio B (e).

Ejemplo I.2. En este ejemplo repetimos el experimento realizado en el ejemplo anterior pero ahora para un nivel de ruido del 4%.

En la Figura 5.20 se presentan la imagen original I (a), el dato con ruido (b), el dato con ruido seguido de umbralado (c), la imagen obtenida por inversión del dato con ruido (d) y la imagen obtenida por inversión del dato limpio B (e).

Los resultados de restauración se muestran en la Figura 5.21 y en la Tabla 5. En este caso valen las mismas observaciones realizadas en el Ejemplo I.1 con el agregado de que la mejora introducida por la aplicación del paso previo wavelet-umbralado es ahora mejor para ambos métodos (L y VT). En efecto, mientras que estas mejoras eran de 52% y 20% en el ejemplo anterior, ahora son del 130% y 28%, respectivamente. Esto pone de manifiesto una vez más que en algunos casos la eliminación de componentes de alta frecuencia introducida por el ruido, previo a la aplicación de un método de regularización, puede resultar en desempeños significativamente superiores.

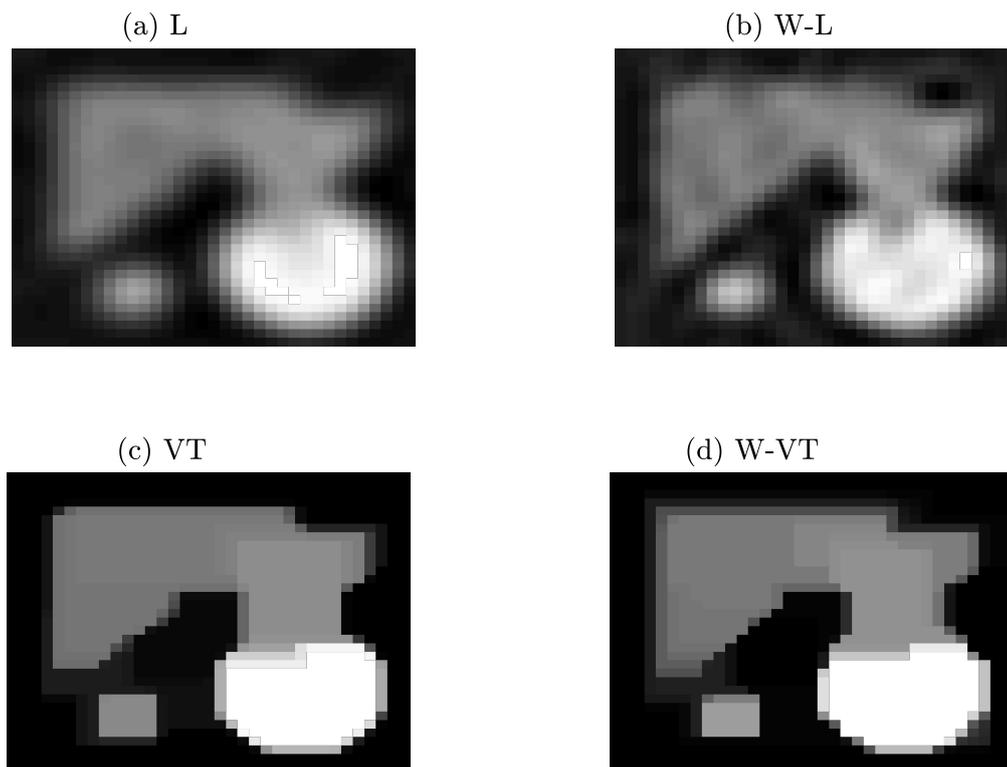


FIGURA 5.21. Restauraciones de I utilizando los métodos L (a), W-L (b), VT (c) y W-VT (d) para un dato con ruido del 4%.

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|--------|-------------------------|--------|
| Landweber | 0.6827 | Wavelet-Landweber | 1.5721 |
| Variación Total | 1.4909 | Wavelet-Variación Total | 1.9130 |

TABLA 5. Índices ISNR correspondientes a las restauraciones de la Figura 5.21.

CASO II: En este caso nos proponemos analizar el desempeño del método de dos pasos con umbralado wavelet para el problema inverso de deblurring utilizando una imagen original de mayor resolución (107×107) que la imagen del caso anterior y de diferente naturaleza, con bordes pronunciados pero también con regiones suaves (ver Figura 5.22 (a)). Para los ejemplos de este caso tomamos $\sigma = 0.01\|I\|_\infty$.

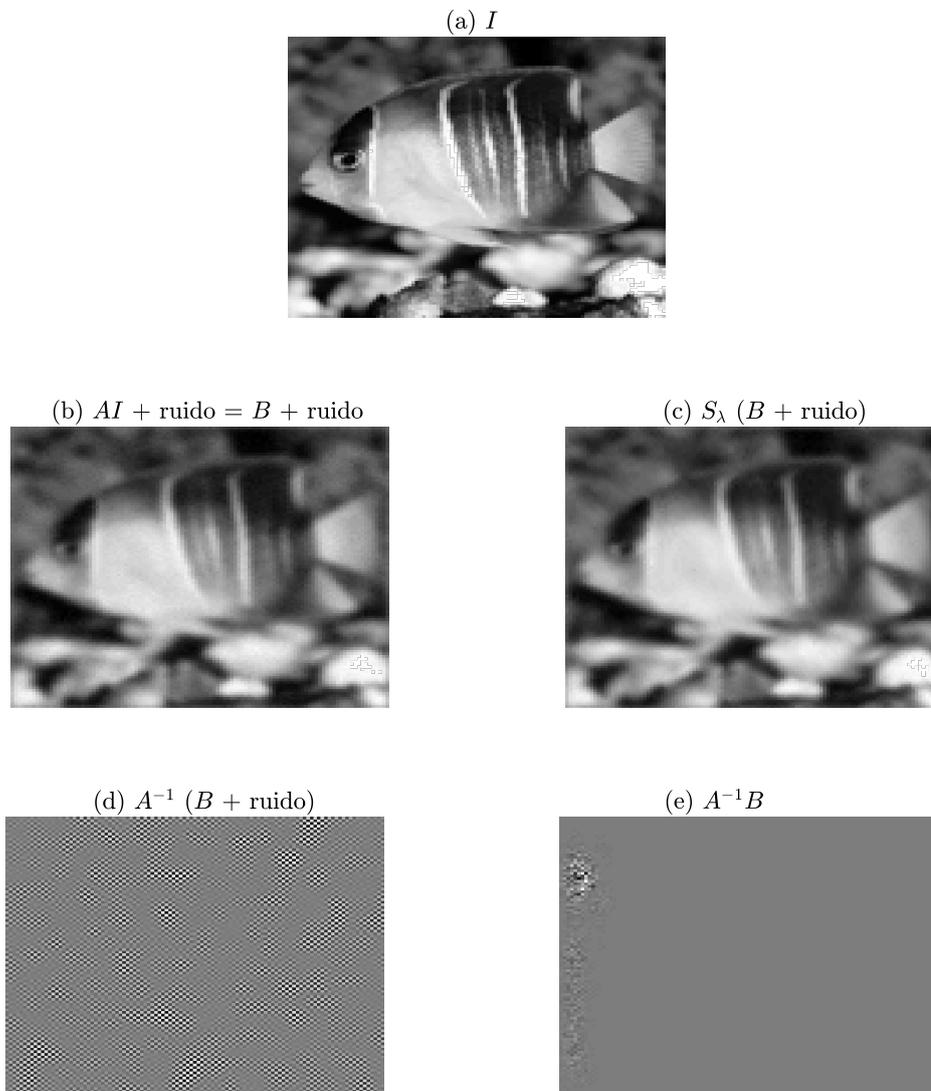


FIGURA 5.22. Imagen original (a), dato con ruido (b), dato con ruido umbralado (c), imagen obtenida por inversión del dato con ruido (d) e imagen obtenida por inversión del dato limpio B (e).

Ejemplo II.1. En este experimento presentamos el problema inverso de aproximar una imagen original I , a partir de un dato B obtenido utilizando $\sigma = 0.01\|I\|_\infty$ y un nivel de ruido aditivo del 1 %.

En la Figura 5.22 se presentan las gráficas de la imagen original I (a), el dato con ruido (b), el dato con ruido umbralado (c), la imagen obtenida por inversión del dato con ruido (d) y la imagen obtenida por inversión del dato limpio B (e).

Las restauraciones y sus correspondientes ISNRs obtenidos con los métodos L, W-L, VT y W-VT se presentan en la Figura 5.23 y la Tabla 6.

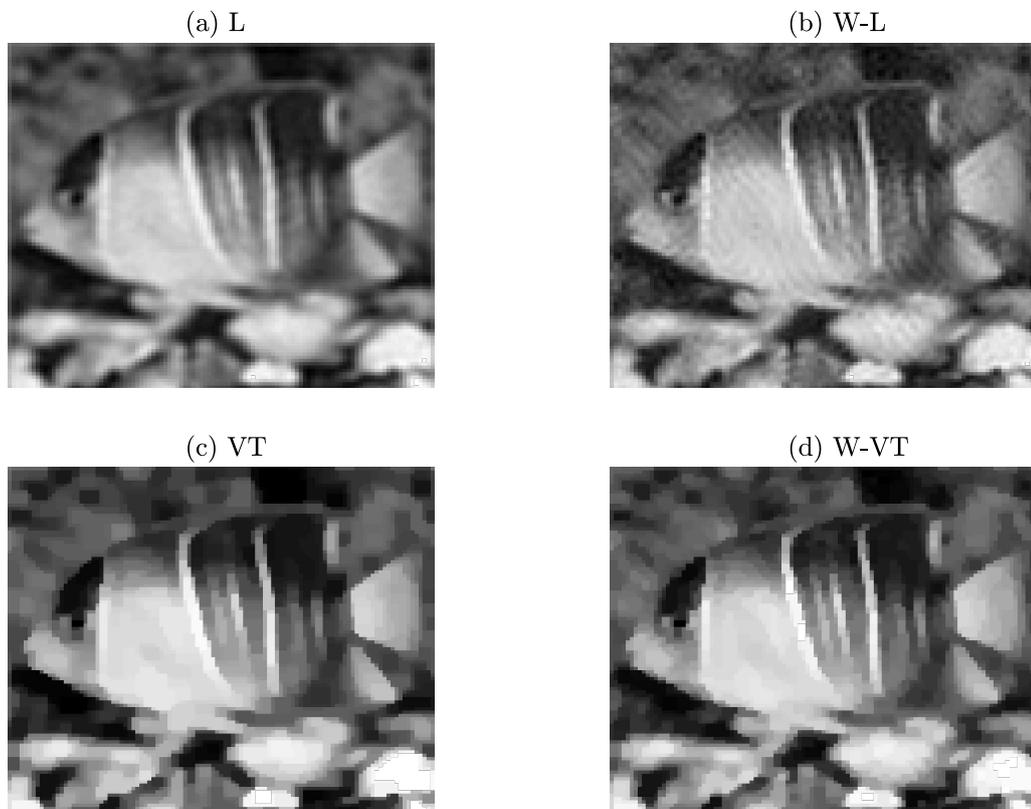


FIGURA 5.23. Restoraciones de I utilizando los métodos L (a), W-L (b), VT (c) y W-VT (d) para un dato con ruido del 1 %.

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|--------|-------------------------|--------|
| Landweber | 2.6622 | Wavelet-Landweber | 3.0261 |
| Variación Total | 2.4041 | Wavelet-Variación Total | 2.8967 |

TABLA 6. Índices ISNR correspondientes a las restauraciones de la Figura 5.22.

Ejemplo II.2. En este ejemplo repetimos el experimento realizado en el ejemplo anterior, ahora con un nivel de ruido del 2%.

En la Figura 5.24 se muestran las gráficas de la imagen original I (a), el dato con ruido (b), el dato con ruido umbralado (c), la imagen obtenida por inversión del dato con ruido (d) y la imagen obtenida por inversión del dato limpio B (e).

Además, se presentan en la Figura 5.25 y la Tabla 7 las restauraciones y sus correspondientes ISNRs obtenidos con los métodos L, W-L, VT y W-VT.

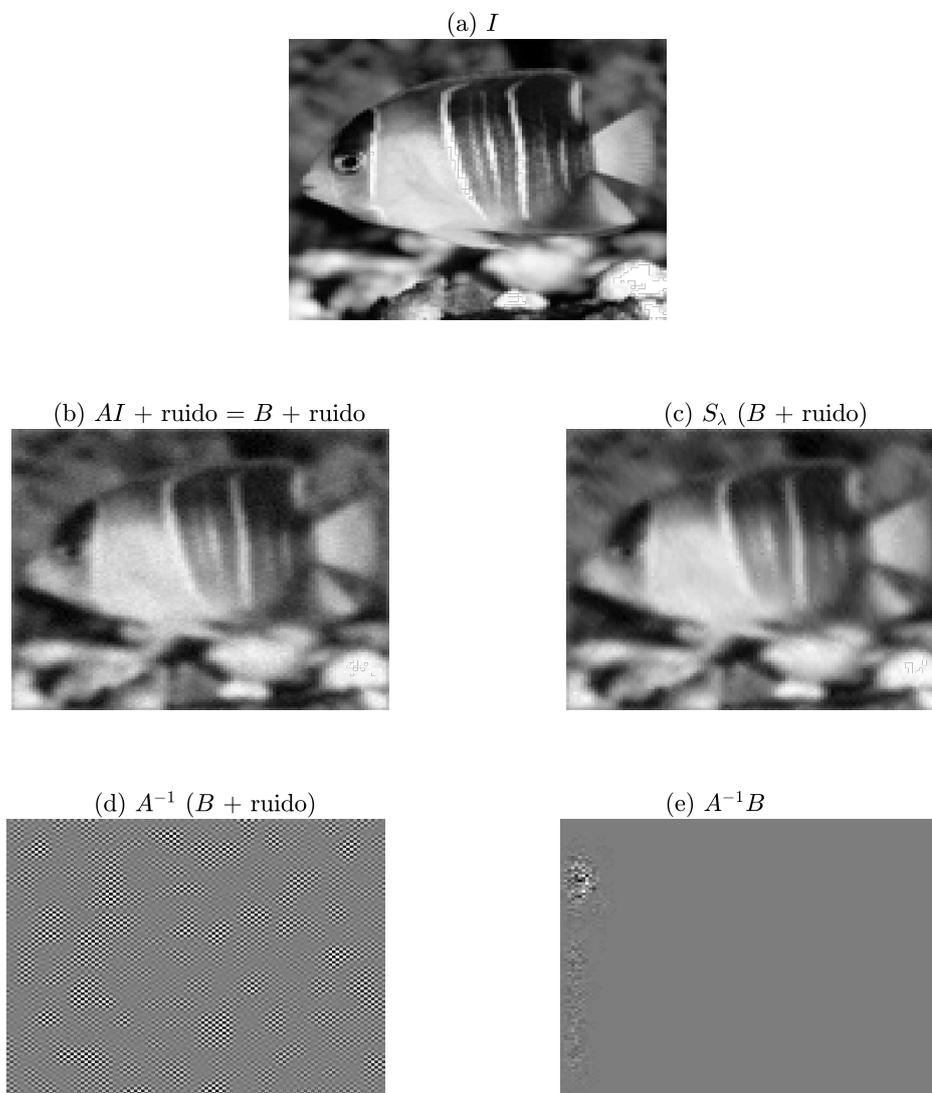


FIGURA 5.24. Imagen original (a), dato con ruido (b), dato con ruido umbralado (c), imagen obtenida por inversión del dato con ruido (d) y imagen obtenida por inversión del dato limpio B (e).

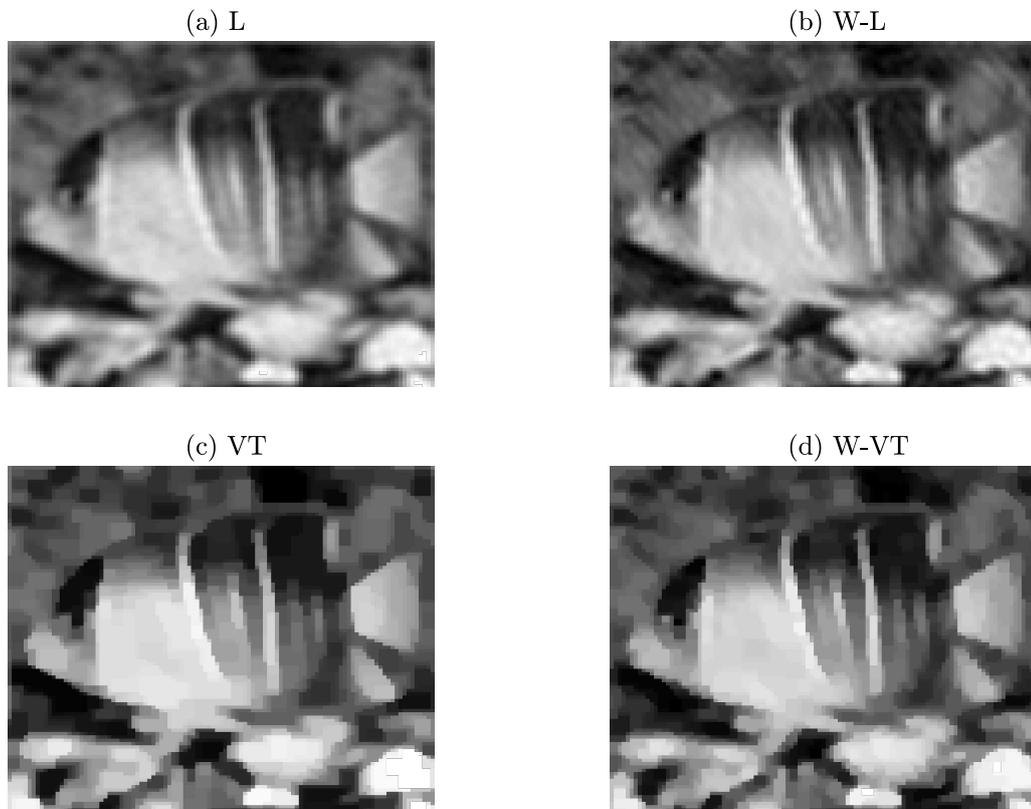


FIGURA 5.25. Restauraciones de I utilizando los métodos L (a), W-L (b), VT (c) y W-VT (d) para un dato con ruido del 2%.

| MÉTODO | ISNR | Wavelet-MÉTODO | ISNR |
|-----------------|--------|-------------------------|--------|
| Landweber | 1.2228 | Wavelet-Landweber | 2.1130 |
| Variación Total | 0.9777 | Wavelet-Variación Total | 1.4810 |

TABLA 7. Índices ISNR correspondientes a las restauraciones de la Figura 5.25.

Los dos ejemplos que acabamos de presentar (Ejemplo II.1 y II.2) ponen nuevamente de manifiesto que la aplicación del paso wavelet-umbralado resulta en mejoras significativas en el desempeño del método de regularización que se aplica a posteriori. Valen, aquí también, todas las observaciones que presentamos para los Ejemplos I.1 y I.2 de esta subsección.

5.2. Aplicaciones del método RWV

En el Capítulo 4 presentamos los métodos de regularización wavelet-vaguelet con y sin umbralado wavelet para problemas inversos mal condicionados. Como vimos, estos métodos, generan una representación que incorpora de manera eficiente información relevante tanto del operador asociado al problema como de los supuestos de regularidad de la solución exacta del mismo.

En esta sección nos proponemos analizar el desempeño del método wavelet-vaguelet con umbralado (RWVU), propuesto en la Sección 4.4.1, en el problema $T_k f = g^\delta$ donde T_k es el operador de convolución dado en (5.1), discretizado como en (5.6) y resultante en el sistema (5.7). Utilizaremos nuevamente las funciones f_1 , f_2 y f_3 que mostramos en la Figura 5.1 y, en todos los casos, presentaremos las gráficas de las reconstrucciones obtenidas con cada método junto con el correspondiente índice ISNR y la solución exacta.

Es importante resaltar que en la implementación del método RWVU el proceso de “regularización” del problema no se realiza a partir del dato g^δ , sino que, como vimos en la Sección 4.4, se realiza la regularización para cada elemento de la base wavelet elegida. Más precisamente, dado el operador T_k y la base wavelet $\{\psi_\lambda\}_{\lambda \in \Lambda}$ elegida, se debe aproximar $(T_k^{-1})^* \psi_\lambda$, para todo $\lambda \in \Lambda$, utilizando un método de regularización apropiado. De esta forma se obtiene una aproximación de la solución exacta f a partir de la identidad (ver (4.36))

$$f_c^\delta = \sum_{\lambda \in \Lambda} \kappa_\lambda^{-1} \langle g^\delta, u_\lambda \rangle \psi_\lambda(x), \quad \text{con} \quad u_\lambda = \kappa_\lambda (T_k^{-1})^* \psi_\lambda. \quad (5.11)$$

Esta metodología requiere, por lo tanto, un proceso de regularización para cada uno de los elementos de la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$. Esto, en principio, implica un aumento considerable en el tiempo de cómputo. Sin embargo, el método posee la ventaja de que este proceso se realiza solo una vez para cada operador y para cada base wavelet. Así, luego de este paso, es posible resolver cualquier problema inverso asociado al operador T_k y a la base $\{\psi_\lambda\}_{\lambda \in \Lambda}$. Esto resulta particularmente ventajoso en el caso en que una misma ecuación deba ser resuelta para varios datos disponibles, en cuyo caso el tiempo total de cómputo puede reducirse significativamente.

En esta sección, a menos que se especifique otra elección, utilizaremos la wavelet de Daubechies (‘db4’) y regularizaremos $(T_k^{-1})^* \psi_\lambda$ utilizando el método de Tikhonov-Phillips (TP) con $n = 85$. Además, para la implementación del umbralado posterior a la regularización propuesta por el método, se utilizará el mismo nivel de umbral que en la sección anterior $\lambda = 2.5 \delta$.

Con el objetivo de mostrar y cuantificar las ventajas del método propuesto, en esta sección presentaremos algunos casos similares a los desarrollados en la Sección 5.1.1, para problemas inversos de la forma $T_k f = g$, con T_k el operador integral de convolución con núcleo gaussiano.

CASO I:

En este caso nos proponemos analizar el desempeño del método RWVU para funciones con diferentes grados de regularidad. Para ello presentamos tres ejemplos utilizando las funciones f_1 , f_2 y f_3 con $\sigma = 0.04 \|f\|_\infty$ y $\delta = 0.05 \|g\|_\infty$.

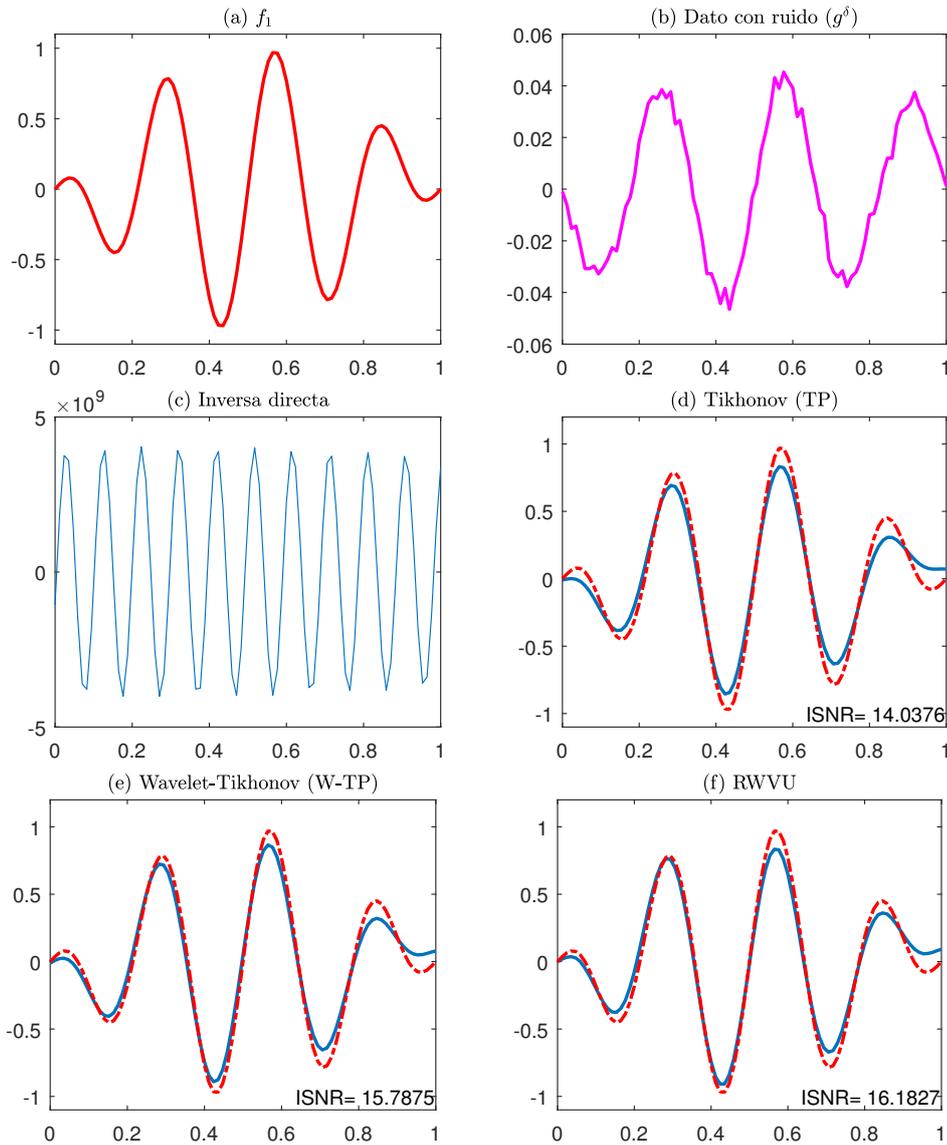


FIGURA 5.26. Solución exacta f_1 (a), dato con ruido g^δ (b), inversa del dato con ruido (c) y restauraciones de f_1 con $\delta = 0.05 \|g\|_\infty$ utilizando los métodos TP (d), W-TP (e) y RWVU (f).

Ejemplo I.1. En este ejemplo hemos aplicado el operador T_k a la función f_1 . Dadas las características de la función, comparamos el método RWVU con los métodos TP y W-TP, ya utilizados en la sección anterior.

En la Figura 5.26 se muestran las gráficas de f_1 (a), g^δ (b), la señal obtenida al aplicar la inversa del operador discretizado directamente al dato con ruido g^δ (c) y las restauraciones

obtenidas con los métodos TP (d), W-TP (e) y RWVU (f). En las Figuras 5.26 (d), (e) y (f) la gráfica en línea discontinua representa la solución exacta f_1 .

Ejemplo I.2. En este ejemplo repetimos el experimento realizado en el ejemplo anterior pero ahora para la función f_2 .

La Figura 5.27 presenta las gráficas de f_2 (a), g^δ (b), la señal obtenida al aplicar la inversa del operador discretizado directamente al dato con ruido g^δ (c) y las restauraciones obtenidas con los métodos VT (d), W-VT (e) y RWVU (f).

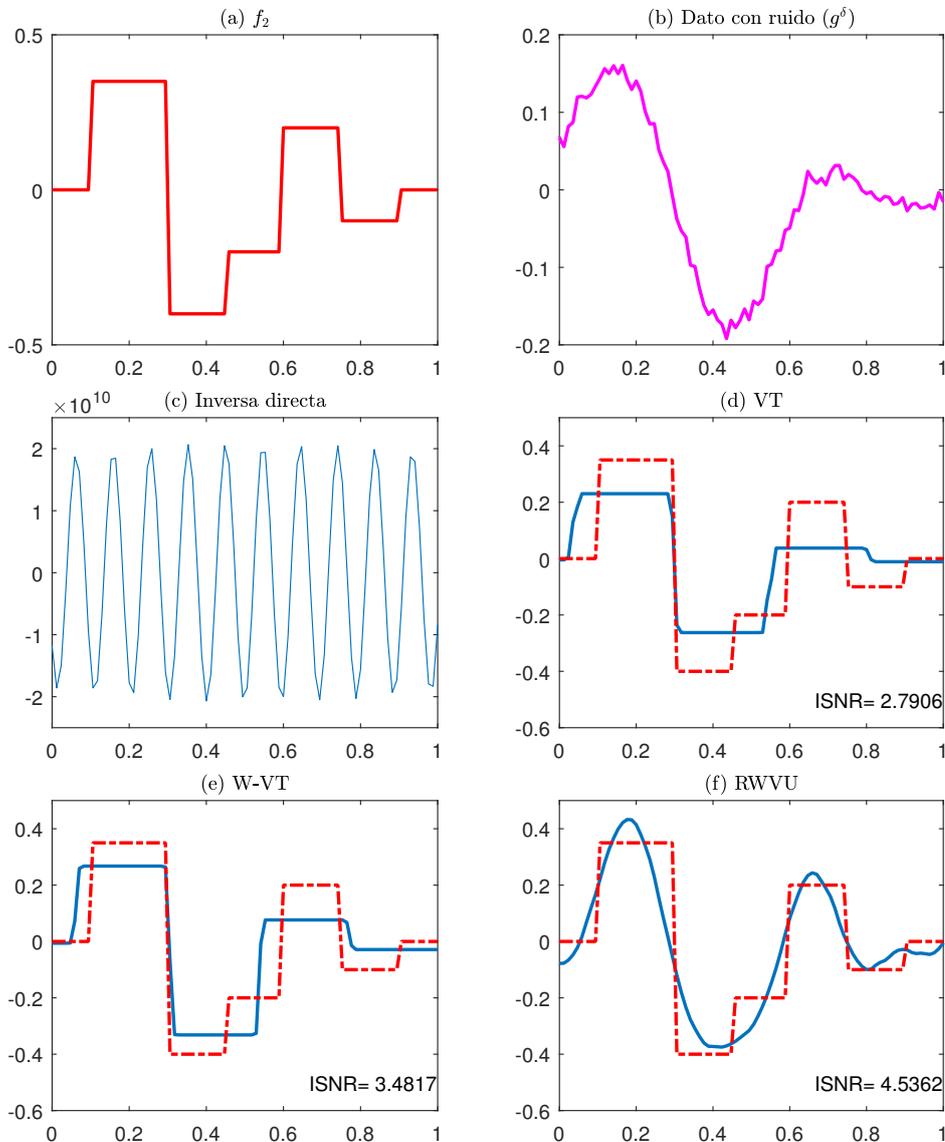


FIGURA 5.27. Dato g^δ (a), dato umbralado $S_\lambda g^\delta$ (b), inversa del dato con ruido (c) y restauraciones de f_2 con $\delta = 0.07 \|g\|_\infty$ utilizando los métodos VT (d), W-VT (e), RWVU (f).

Ejemplo I.3. En este ejemplo repetimos los experimentos realizados en los Ejemplos I.1 e I.2, ahora para la solución exacta f_3 .

La Figura 5.28 presenta las gráficas de f_3 (a), g^δ (b), la señal obtenida al aplicar la inversa del operador discretizado directamente al dato con ruido g^δ (c) y las restauraciones obtenidas con los métodos TP (d), W-TP (e) y RWVU (f).

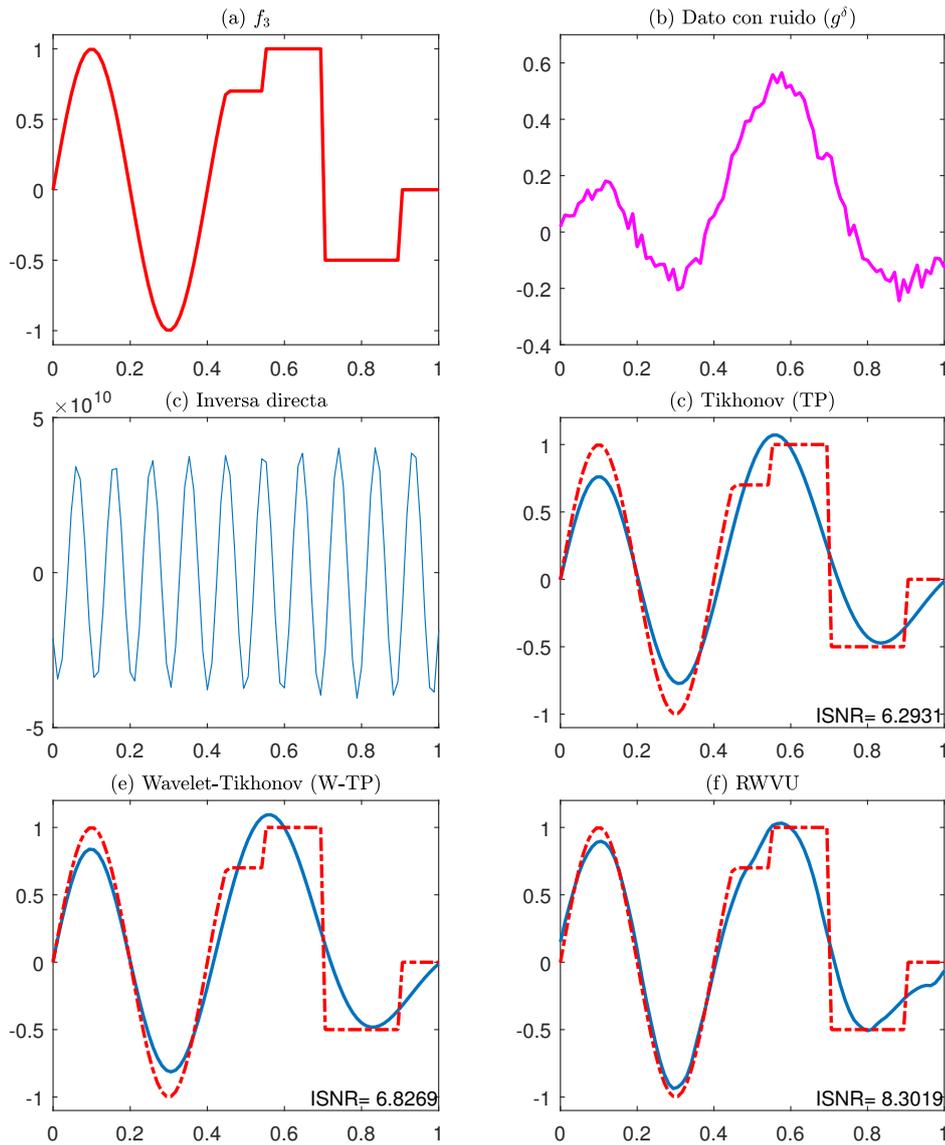


FIGURA 5.28. Dato g^δ (a), dato umbralado $S_\lambda g^\delta$ (b), inversa del dato con ruido (c) y restauraciones de f_1 con $\delta = 0.05 \|g\|_\infty$ utilizando los métodos TP (d), W-TP (e), RWVU (f).

En los tres ejemplos anteriores (I.1, I.2 e I.3) podemos observar como el método de regularización wavelet-vaguelet con umbralado (RWVU) no solamente mejora significativamente el desempeño del correspondiente método tradicional (TP o VT) sino que además mejora el desempeño de su correspondiente versión wavelet-umbralada (W-TP o W-VT). Esta última mejora,

como lo mencionamos al inicio del Capítulo 4, se debe a que el método RWVU incorpora adicionalmente las ventajas de la descomposición en valores singulares del operador.

CASO II:

Nos proponemos ahora analizar el desempeño del método RWVU utilizando dos wavelets de naturaleza muy diferente: Daubechies y Haar. En principio, parecería razonable esperar que la utilización de las wavelets de Haar mejoren el desempeño del método RWVU, especialmente en la reconstrucción de funciones contantes a trozos o con discontinuidades. Para dicho análisis consideramos dos ejemplos en los que presentamos las restauraciones de f_2 y f_3 comparando el desempeño del método RWVU al utilizar ambas wavelets, con los métodos de dos pasos utilizados en la sección anterior. Para realizar los experimentos definimos $\sigma = 0.06\|f\|_\infty$ y $\delta = 0.05\|g\|_\infty$ y denotamos con z_λ a la regularización de $(T_k^{-1})^*\psi_\lambda$ obtenida con el método de TP.

Ejemplo II.1. En este ejemplo hemos aplicado el operador T_k a la función f_2 .

En la Figura 5.29 se presentan las restauraciones obtenidas con los métodos RWVU (Daubechies) (a), RWVU (Haar) (b), W-TP (c) y W-VT (d). Las imágenes muestran también los respectivos ISNR como cuantificadores del desempeño de cada método.

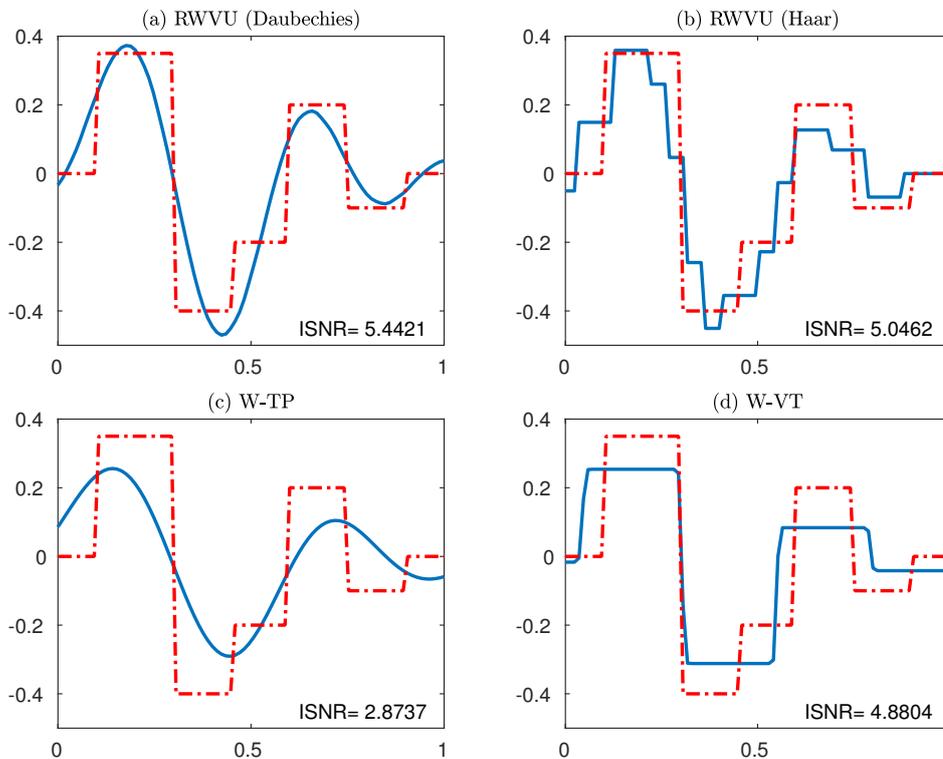


FIGURA 5.29. Restauración de f_2 utilizando los métodos de RWVU (Daubechies) (a), RWVU (Haar) (b), W-TP (c) y W-VT (d).

Ejemplo II.2. En este ejemplo repetimos el experimento realizado en el ejemplo anterior, ahora para la función f_3 . En la Figura 5.30 se presentan las restauraciones obtenidas con los métodos RWVU (Daubechies) (a), RWVU (Haar) (b), W-TP (c) y W-VT (d) junto con los respectivos ISNR.

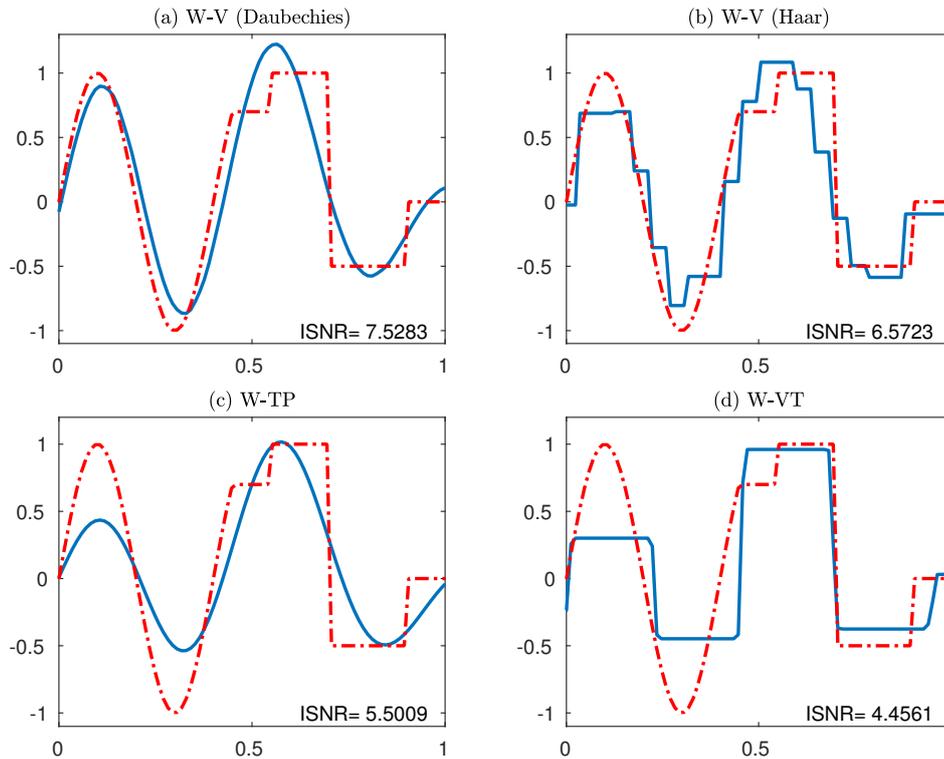


FIGURA 5.30. Restauración de f_3 utilizando los métodos de RWVU (Daubechies) (a), RWVU (Haar) (b), W-TP (c) y W-VT (d).

Un análisis de los resultados obtenidos en los últimos dos ejemplos muestra que la utilización de las wavelets de Haar en el método RWVU, si bien mejora significativamente el desempeño de los dos métodos con el paso previo wavelet-umbralado W-TP y W-VT, en ningún caso mejora el desempeño del correspondiente método RWVU utilizando wavelets de Daubechies, contrariamente a lo que mencionamos al principio y podríamos haber esperado. Un análisis cuidadoso de la ecuación (5.11) que resume el método RWVU revela inmediatamente la causa de esto. En efecto, observar que en (5.11) la utilización de las wavelets está asociada a un proceso de regularización con el operador T^* (el que, como es bien sabido imprime las mismas condiciones de suavidad que T), en el problema de aproximar z_λ en la ecuación $T^* z_\lambda = \psi_\lambda$, o sea $z_\lambda = (T_k^*)^{-1} \psi_\lambda$. Es precisamente por esta razón (las características de suavidad que induce la aplicación de T^*) que la ecuación $T^* z_\lambda = \psi_\lambda$ puede resolverse mejor para wavelets suaves que para wavelets de tipo Haar. Las Figuras 5.31 y 5.32 muestran cuatro de estas aproximaciones

(para cuatro valores de λ) para las wavelets de Daubechies y Haar, respectivamente. Allí se observa claramente que, en todos los casos, las aproximaciones para el caso de las wavelets de Daubechies son significativamente mejores que las respectivas obtenidas con las wavelets de Haar.

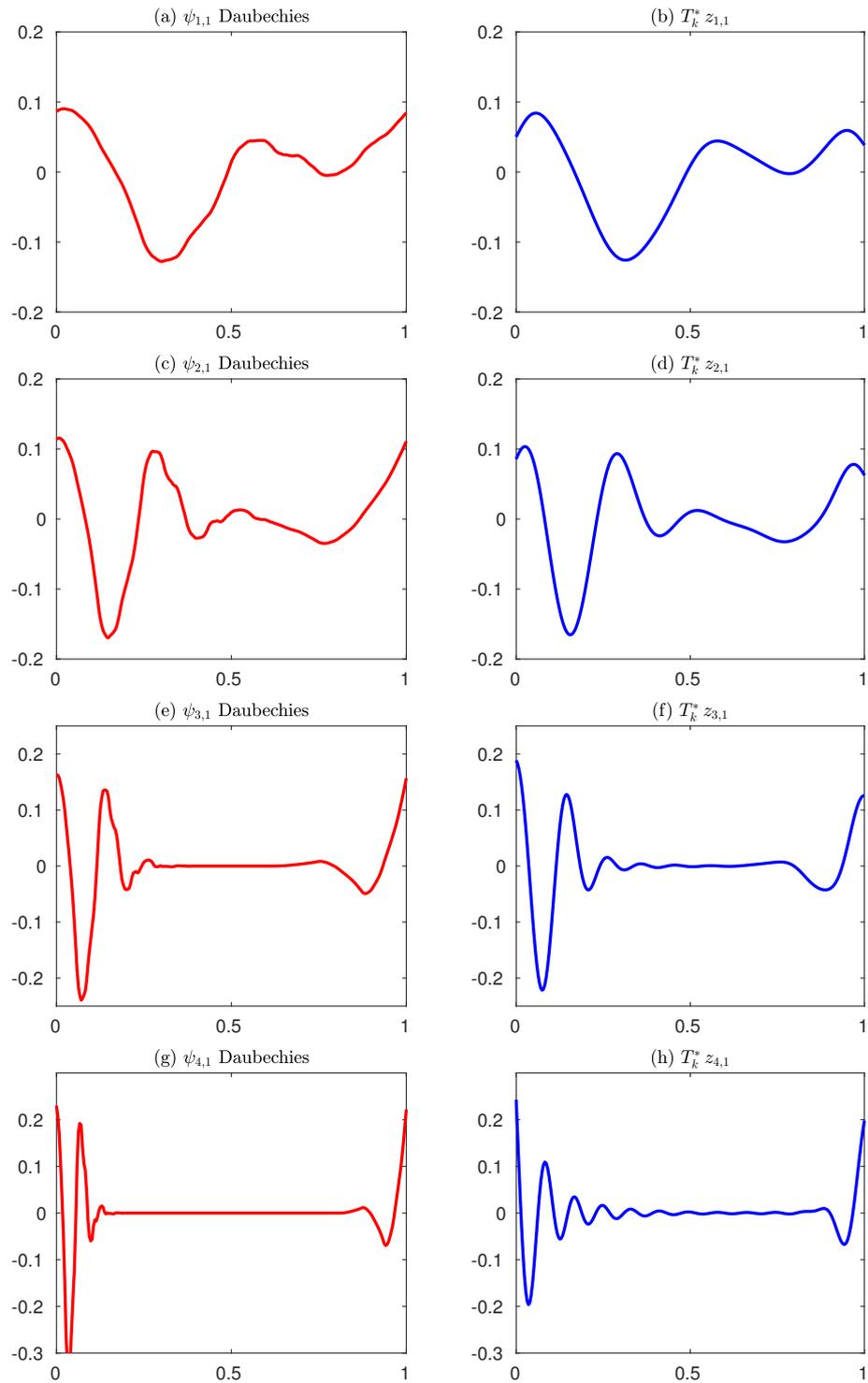


FIGURA 5.31. Wavelet Daubechies y $T_k^* z_\lambda$ para $\lambda = (1, 1)$ (a) y (b), $\lambda = (2, 1)$ (c) y (d), $\lambda = (3, 1)$ (e) y (f) y $\lambda = (4, 1)$ (g) y (h), respectivamente.

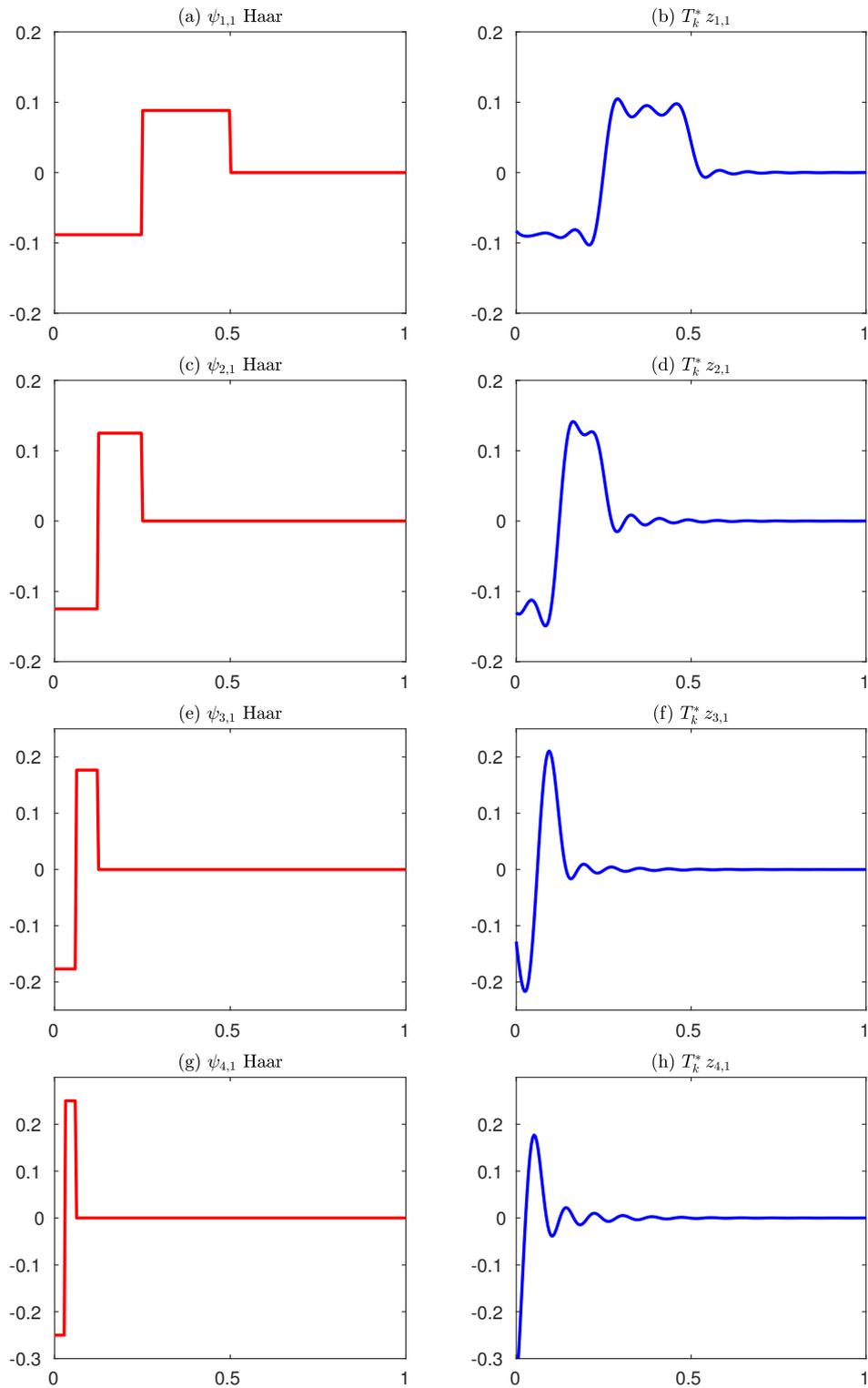


FIGURA 5.32. Wavelet Haar y $T_k^* z_{\lambda}$ para $\lambda = (1, 1)$ (a) y (b), $\lambda = (2, 1)$ (c) y (d), $\lambda = (3, 1)$ (e) y (f) y $\lambda = (4, 1)$ (g) y (h), respectivamente.

Conclusiones y trabajos futuros

En los dos primeros capítulos de esta tesis se presentaron los principales resultados sobre el estado del arte sobre el tema incluyendo varias herramientas matemáticas necesarias que nos permitieron abordar adecuadamente los problemas que se trataron en los capítulos posteriores.

En el Capítulo 3 se presentó el método de regularización wavelet-espectral. Se realizó un análisis detallado de la convergencia para los errores de proyección, umbralado y regularización del método wavelet-Landweber que concluyó en tres resultados que brindan cotas para cada uno de estos errores (ver Teoremas 3.5, 3.6 y Lema 3.11). Posteriormente, se probó el teorema fundamental de convergencia de orden casi-óptimo para el método combinado wavelet-Landweber, en el Teorema 3.16. Luego, se generalizaron estos resultados de convergencia para el caso de métodos wavelet-espectrales en los que la familia espectral posee ciertas propiedades asociadas a su calificación clásica (ver Lemas 3.17 y 3.18). En particular, en el Teorema 3.22 se probó que, bajo ciertas condiciones, el método de dos pasos converge de forma casi-óptima.

En el Capítulo 4 se presentó el método de regularización wavelet-vaguelet. En primer lugar se definieron dos tipos de operadores especiales, operadores débilmente invertibles y operadores homogéneos y se probaron algunos resultados importantes relacionados a tales operadores, los que resultaron de utilidad en demostraciones posteriores relacionadas con la convergencia del método propuesto. Luego, se introdujo el concepto de vaguelet y se demostraron varias propiedades fundamentales de estas funciones. Además, se definieron una serie de errores y riesgos correspondientes al desempeño de los métodos RWV y RWVU vinculados a los modelos continuo y discreto, y los enfoques determinístico y bayesiano. A continuación se realizó un análisis detallado de la convergencia del error correspondiente al método propuesto considerando dos versiones del mismo: con y sin umbralado wavelet fuerte. En particular, en los Teoremas 4.45 y 4.51, se probó que, bajo ciertas condiciones, el método de regularización wavelet-vaguelet y el método wavelet-vaguelet con umbralado wavelet convergen de manera óptima. Por último, en lo que respecta al método RWV, se presentó y analizó una aplicación concreta del método para el caso de operadores integrales que poseen descomposición wavelet vaguelet.

Finalmente, se presentaron varias aplicaciones a problemas inversos mal condicionados en restauración de señales e imágenes, cuyas implementaciones numéricas permitieron una mejor visualización de los resultados teóricos obtenidos y mostraron claramente la potencialidad de los métodos desarrollados. Estos resultados probaron que los métodos desarrollados mejoran, en algunos casos significativamente, el desempeño de los métodos de regularización tradicionales.

Trabajos futuros

Para dar continuidad al trabajo desarrollado en esta tesis, nos proponemos avanzar en los siguientes problemas, los que, como vimos a lo largo del desarrollo de esta tesis, requieren de más investigación.

- a) Generalización de los resultados de convergencia correspondientes al método wavelet-espectral para el caso de operadores no compactos;
- b) Estimación óptima de los parámetros λ y δ_λ correspondientes al umbralado previo a la aplicación del método de regularización wavelet-espectral;
- c) Estudio de condiciones suficientes que aseguren convergencia y estimación de órdenes de convergencia para el método de dos pasos, con el paso wavelet-umbralado previo a métodos de regularización no espectrales, tales como el método de regularización por variación total VT;
- d) Estudio del desempeño del método RWVU para otro tipo de wavelets teniendo en cuenta el grado de regularidad de la solución exacta del problema;
- e) Estimación óptima del parámetro λ correspondiente al umbralado posterior a la aplicación del método de regularización wavelet-vaguelet.

Anexos

En este anexo incluimos una serie de resultados que se utilizan a lo largo de los Capítulos 3 y 4. Se trata de resultados importantes, pero que consideramos de tipo auxiliar, tangenciales al desarrollo de esta tesis. En el Anexo A se encuentran resultados utilizados en el Capítulo 3 y en el Anexo B resultados utilizados durante el desarrollo del Capítulo 4.

Anexo A

Proposición 5.1. *Sea Z una variable aleatoria con distribución $N(0,1)$. Entonces para todo $\omega \geq 0$ se sigue que*

$$P(|Z| > \omega) \leq 3e^{-\frac{\omega^2}{2}}.$$

Demostración. Dividiremos la prueba en dos casos.

El caso $\omega \geq \frac{1}{\sqrt{2\pi}}$ se sigue inmediatamente de observar que

$$\begin{aligned} P(|X| > \omega) &= 2 \int_{\omega}^{\infty} e^{-\frac{t^2}{2}} \frac{dt}{\sqrt{2\pi}} \\ &= 2 \int_{\frac{\omega^2}{2}}^{\infty} e^{-u} \frac{du}{\sqrt{2\pi}\sqrt{2u}} && \text{(donde } u = \frac{t^2}{2}\text{)} \\ &\leq 2 \int_{\frac{\omega^2}{2}}^{\infty} e^{-u} dt && \text{(pues } \frac{1}{2\sqrt{\pi u}} \leq 1\text{)} \\ &= 2e^{-\frac{\omega^2}{2}}. \end{aligned}$$

El caso $0 \leq \omega < \frac{1}{\sqrt{2\pi}}$ en cambio, lo trabajamos como suma de dos integrales. Escribimos

$$P(|X| > \omega) = \frac{2}{\sqrt{2\pi}} \int_{\omega}^{\frac{1}{\sqrt{2\pi}}} e^{-\frac{t^2}{2}} dt + \frac{2}{\sqrt{2\pi}} \int_{\frac{1}{\sqrt{2\pi}}}^{\infty} e^{-\frac{t^2}{2}} dt \doteq \mathbf{(1)} + \mathbf{(2)}.$$

Para acotar **(1)** utilizamos que $e^{-t^2/2} \leq 1$ y podemos acotar la integral con el área del rectángulo de altura 1 y lado $\frac{1}{\sqrt{2\pi}}$. Así pues

$$\frac{2}{\sqrt{2\pi}} \int_{\omega}^{\frac{1}{\sqrt{2\pi}}} e^{-\frac{t^2}{2}} dt \leq \frac{2}{\sqrt{2\pi}} \int_0^{\frac{1}{\sqrt{2\pi}}} dt = \frac{2}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} = \frac{1}{\pi} \leq e^{-\frac{\omega^2}{2}},$$

donde la última desigualdad se sigue del hecho que $\omega^2 \leq \frac{1}{2\pi} < 2 \ln \pi$.

Para acotar (2) calculamos la integral utilizando un razonamiento similar al empleado en el caso $\omega \geq \frac{1}{\sqrt{2\pi}}$. Así se tiene que

$$2 \int_{\frac{1}{\sqrt{2\pi}}}^{\infty} e^{-\frac{t^2}{2}} \frac{dt}{\sqrt{2\pi}} = 2 \int_{\frac{1}{4\pi}}^{\infty} e^{-u} \frac{du}{\sqrt{2\pi}\sqrt{2u}} = \int_{\frac{1}{4\pi}}^{\infty} \frac{e^{-u}}{\sqrt{\pi u}} du \leq 2 \int_{\frac{1}{4\pi}}^{\infty} e^{-u} du = 2e^{-\frac{1}{4\pi}} \leq 2e^{-\frac{\omega^2}{2}}.$$

De esta forma, hemos demostrado el caso $0 \leq \omega < \frac{1}{\sqrt{2\pi}}$ pues

$$P(|X| > \omega) = \mathbf{(1)} + \mathbf{(2)} \leq e^{-\frac{\omega^2}{2}} + 2e^{-\frac{\omega^2}{2}} = 3e^{-\frac{\omega^2}{2}}. \quad \blacksquare$$

Anexo B

Lema 5.2. Sean X, Y espacios de Hilbert y $K : X \rightarrow Y$ un operador lineal, invertible con dominio $\mathcal{D}(K)$ denso en X . Entonces son equivalentes

- 1) $w \in \mathcal{R}(K^*)$;
- 2) existe un funcional lineal y acotado $c(\cdot) : Y \rightarrow \mathbb{R}$ que satisface $c(Kf) = \langle w, f \rangle$ para todo $f \in \mathcal{D}(K)$;
- 3) la desigualdad $|\langle w, f \rangle| \leq C\|Kf\|_Y$, donde C es una constante, vale para toda $f \in \mathcal{D}(K)$.

Demostración. La demostración de este lema es relativamente sencilla. Se sigue de la definición de operador adjunto y del teorema de representación de Riesz.

Consideramos las siguientes implicaciones que prueban la equivalencia de los enunciados.

- 1) \Leftrightarrow 2) Se sigue de la definición de operador adjunto; es posible definir este operador puesto que $\mathcal{D}(K)$ es denso.
- 1) \Rightarrow 3) Si $w \in \mathcal{R}(K^*)$, entonces por el Teorema de representación de Riesz, $\exists c \in Y$ tal que $\langle w, f \rangle = \langle K^*c, f \rangle = [c, Kf] \leq \|c\|_Y \|Kf\|_Y$.
- 3) \Rightarrow 2) Consideremos el funcional $c(\cdot) : \mathcal{R}(K) \subset Y \rightarrow \mathbb{R}$ dado por $c(Kf) = \langle w, f \rangle$. Este funcional está bien definido puesto que el operador K es invertible y es claramente lineal. Luego de la desigualdad $|\langle w, f \rangle| \leq C\|Kf\|_Y$ para toda $f \in \mathcal{D}(K)$ se sigue que el funcional es lineal y acotado para todo $f \in \mathcal{D}(K)$. \blacksquare

Lema 5.3. Sean $b \in (0, 1]$, $\alpha > 0$, $d \in \mathbb{N}$ y $f_b(x) \doteq \frac{1}{(1 + b|x|)^{d+\alpha}}$, $\forall x \in \mathbb{R}^d$.

Entonces $\forall z \in \mathbb{R}^d$ se satisface

$$(f_1 * f_b)(z) \leq C(d, \alpha) f_b(z) \quad \text{dónde} \quad C(d, \alpha) = \frac{2\pi^{d/2}}{\alpha\Gamma(d/2)} (2^{d+2+\alpha} + 1),$$

y donde Γ es la función Gamma definida por $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$.

Demostración. Con el objetivo de acotar la integral correspondiente a la convolución descompondremos la integral en otras dos, de la siguiente manera:

$$(f_1 * f_b)(z) = \int_{|y-z| \geq |z|} f_1(y) f_b(z-y) dy + \int_{|y-z| \leq |z|} f_1(y) f_b(z-y) dy \doteq I_1 + I_2.$$

Para realizar la acotación de I_1 utilizamos que $f_b(z) \geq f_b(z-y)$ para $|y-z| \geq |z|$. Así,

$$\begin{aligned} I_1 &= \int_{|y-z| \geq |z|} f_1(y) f_b(z-y) dy \leq f_b(z) \int_{|y-z| \geq |z|} f_1(y) dy \\ &\leq f_b(z) \int_{\mathbb{R}^d} \frac{1}{(1+|y|)^{d+\alpha}} dy = \frac{2\pi^{d/2}}{\alpha\Gamma(d/2)} f_b(z). \end{aligned}$$

Para acotar I_2 necesitaremos de algunos pasos previos. En primer lugar, notemos que $|y-z| \leq |z|$ implica que $0 \leq |y| \leq |2z|$. Luego,

$$I_2 \leq \int_0^{2|z|} f_1(y) f_b(z-y) dy. \quad (5.12)$$

Supongamos, sin pérdida de generalidad que $z \geq 0$ (una demostración análoga vale para $z < 0$). Entonces

$$\begin{aligned} \int_z^{2z} f_1(y) f_b(z-y) dy &= \int_z^{2z} \frac{dy}{(1+y)^{d+\alpha} (1+b|z-y|)^{d+\alpha}} \\ &\leq \frac{1}{(1+z)^{d+\alpha}} \int_z^{2z} \frac{dy}{(1+b|z-y|)^{d+\alpha}} \\ &= \frac{1}{(1+z)^{d+\alpha}} \int_0^z \frac{dy}{(1+b|z-y|)^{d+\alpha}} \quad \left(\begin{array}{l} \text{pues el integrando es} \\ \text{simétrico alrededor de } y = z \end{array} \right) \\ &\leq \int_0^z \frac{dy}{(1+y)^{d+\alpha} (1+b|z-y|)^{d+\alpha}} \\ &= \int_0^z f_1(y) f_b(z-y) dy. \end{aligned} \quad (5.13)$$

De (5.12) y (5.13) se sigue que

$$\begin{aligned} I_2 &\leq 2 \int_0^z f_1(y) f_b(z-y) dy \\ &= 2 f_b(z) \int_0^z \left[\frac{1+bz}{(1+y)(1+b|z-y|)} \right]^{d+\alpha} dy \\ &= 2 f_b(z) \left(\frac{1+bz}{1+b+bz} \right)^{d+\alpha} \int_0^z \left[\frac{1}{1+y} + \frac{b}{1+b|z-y|} \right]^{d+\alpha} dy \\ &\leq 2 f_b(z) \int_0^z \left[\frac{1}{1+y} + \frac{1}{1+|z-y|} \right]^{d+\alpha} dy \\ &= 2^2 f_b(z) \int_0^{z/2} \left[\frac{1}{1+y} + \frac{1}{1+|z-y|} \right]^{d+\alpha} dy \end{aligned}$$

$$\begin{aligned}
&\leq 2^2 f_b(z) \int_0^{z/2} \left[\frac{2}{1+y} \right]^{d+\alpha} dy \\
&\leq 2^{d+2+\alpha} f_b(z) \int_0^{z/2} \left[\frac{1}{1+y} \right]^{d+\alpha} dy \\
&\leq 2^{d+2+\alpha} \frac{2\pi^{d/2}}{\alpha\Gamma(d/2)} f_b(z).
\end{aligned}$$

Finalmente, de las acotaciones anteriores obtenemos

$$I_1 + I_2 \leq \frac{2\pi^{d/2}}{\alpha\Gamma(d/2)} (2^{d+2+\alpha} + 1) f_b(z) \quad \blacksquare$$

Lema 5.4. *Dados $b > 0$, $\gamma > 0$ y $d \in \mathbb{N}$ se tiene que*

$$\sum_{k \in \mathbb{Z}^d} \left(\frac{1}{1+b|k|} \right)^{d+\gamma} \leq 2\gamma^{-1}b^{-d} + 1.$$

Demostración. Por simplicidad lo probaremos para $k \in \mathbb{Z}$, el caso $k \in \mathbb{Z}^d$ se demuestra de forma similar.

En primer lugar notemos que

$$\begin{aligned}
\sum_{k \in \mathbb{Z}} \left(\frac{1}{1+b|k|} \right)^{1+\gamma} &= 2 \sum_{k \in \mathbb{N}_0} \left(\frac{1}{1+bk} \right)^{1+\gamma} - 1 \\
&= 2 \int_{\mathbb{R}_0^+} \left(\frac{1}{1+b[x]} \right)^{1+\gamma} dx + 1 \\
&\leq 2 \int_{\mathbb{R}_0^+} \left(\frac{1}{1+bx} \right)^{1+\gamma} dx + 1.
\end{aligned} \tag{5.14}$$

Por otro lado,

$$\begin{aligned}
\int_{\mathbb{R}_0^+} \left(\frac{1}{1+bx} \right)^{1+\gamma} dx &= \frac{1}{b} \int_1^\infty \left(\frac{1}{u} \right)^{1+\gamma} du \quad (\text{para } u = 1+bx) \\
&= -\frac{1}{b\gamma} (u)^{-\gamma} \Big|_1^\infty \\
&= -\frac{1}{b\gamma} (1+bx)^{-\gamma} \Big|_0^\infty \\
&= (\gamma b)^{-1}.
\end{aligned} \tag{5.15}$$

Luego, de (5.14) y (5.15) se obtiene la desigualdad buscada para el caso $d = 1$. \blacksquare

Proposición 5.5. *Sean $\alpha \in \mathbb{R}^+$, $j, k \in \mathbb{Z}$. Entonces*

$$\int_{\mathbb{R}^d} \frac{2^{jd}}{(1+|2^jx-k|)^{d+\alpha}} dx \leq \frac{2\pi^{\frac{d}{2}}}{\alpha\Gamma(d/2)}.$$

Demostración. La prueba se sigue trivialmente con un cambio de variables utilizando coordenadas polares. En efecto,

$$\begin{aligned}
\int_{\mathbb{R}^d} \frac{2^{jd}}{(1 + |2^j x - k|)^{d+\alpha}} dx &= \int_{\mathbb{R}^d} \frac{1}{(1 + |u|)^{d+\alpha}} du && \text{(haciendo } u = 2^j x - k) \\
&= \int_0^\infty \frac{r^{d-1}}{(1+r)^{d+\alpha}} \int_{S^{d-1}} d\sigma dr && \text{(usando coordenadas polares)} \\
&= \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)} \int_0^\infty \frac{r^{d-1}}{(1+r)^{d+\alpha}} dr \\
&= \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)} \int_0^\infty \frac{r^{d-1}}{(1+r)^{d-1}} \frac{1}{(1+r)^{1+\alpha}} dr \\
&\leq \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)} \int_0^\infty \frac{1}{(1+r)^{1+\alpha}} dr \\
&= \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)} \frac{(1+r)^{-\alpha}}{-\alpha} \Big|_0^\infty \\
&= \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)} \frac{1}{\alpha}. \quad \blacksquare
\end{aligned}$$

Proposición 5.6. Si $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ es un operador homogéneo densamente definido, entonces $\mathcal{R}(T^*)$ es cerrado por traslaciones.

Demostración. Dada $f \in \mathcal{R}(T^*)$, existe $g \in \mathcal{D}(T^*)$ tal que $f = T^*g$. Probaremos esta proposición demostrando que dado $a > 0$ existe $g_a \in \mathcal{D}(T^*)$ tal que $D_a f = T^*g_a$.

Supongamos que T es homogéneo de grado β . Entonces

$$\begin{aligned}
\langle a^{\beta d} D_a g, T f \rangle &= \langle g, a^{\beta d} a^{-d} D_{\frac{1}{a}} T f \rangle \\
&= \langle g, a^{-d} T D_{\frac{1}{a}} f \rangle && (T \text{ es homogéneo de grado } \beta) \\
&= a^{-d} \langle T^* g, D_{\frac{1}{a}} f \rangle && (g \in \mathcal{D}(T^*)) \\
&= \langle D_a T^* g, f \rangle.
\end{aligned}$$

De lo anterior se sigue que $a^{\beta d} D_a g \in \mathcal{D}(T)$ y además $T^* a^{\beta d} D_a g = D_a T^* g = D_a f$. Definiendo $g_a \doteq a^{\beta d} D_a g$, se completa la demostración. \blacksquare

Proposición 5.7. Si $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ es un operador densamente definido que conmuta con traslaciones, entonces $\mathcal{D}(T^*)$ es cerrado por traslaciones y también T^* conmuta con ellas.

Demostración. Sea $f \in \mathbb{R}^d$. Entonces existe $g \in \mathcal{D}(T^*)$ tal que $f = T^*g$. Sea $a \in \mathbb{R}$. Queremos ver que $E_a f(\cdot) \doteq f(\cdot - a) \in \mathcal{R}(T^*)$. Para ello, debemos ver que existe $g_a \in \mathcal{D}(T^*)$ tal

que $E_a f = T^* g_a$. Para ello, notar que

$$\begin{aligned}
 \langle D_a g, T f \rangle &= \langle g, \mathcal{D}_{-a} T f \rangle \\
 &= \langle g, T \mathcal{D}_{-a} f \rangle && \text{(puesto que } T \text{ conmuta con traslaciones)} \\
 &= \langle T^* g, \mathcal{D}_{-a} f \rangle \\
 &= \langle \mathcal{D}_a T^* g, f \rangle \quad .
 \end{aligned}$$

Por lo tanto, $D_a g \in \mathcal{D}(T)$ y además $T^* D_a g = D_a T^* g = D_a f$. Esto completa la prueba definiendo $g_a \doteq a^{\beta d} D_a g$. \blacksquare

Lema 5.8. *Si $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ es un operador inyectivo que conmuta con traslaciones, entonces $\mathcal{R}(T)$ es cerrado por traslaciones y también T^{-1} conmuta con ellas.*

Demostración. Dado que el operador T conmuta con traslaciones tenemos que $\forall a \in \mathbb{R}$ se satisface que $T E_a = E_a T$. Por otro lado, al ser T inyectivo existe $T^{-1} : \mathcal{R}(T) \subset L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$. Luego, sobre $\mathcal{R}(T)$ se tiene que

$$E_a T^{-1} = T^{-1} T E_a T^{-1} = T^{-1} (E_a T) T^{-1} = T^{-1} D_a$$

y, por lo tanto, T^{-1} conmuta con traslaciones. \blacksquare

Proposición 5.9. *Sea $T : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$ un operador integral con núcleo $k : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, Hölder continuo en la primera variable. Es decir, existen $\nu > 0$ y $c : \mathbb{R}^d \rightarrow \mathbb{R}^+$ tal que $|k(x, t) - k(x', t)| \leq c(t)|x - x'|^\nu$. Si $\psi \in L^2(\mathbb{R}^d)$ satisface que $c|\psi| \in L^1(\mathbb{R}^d)$ entonces $T\psi$ es Hölder continua con el mismo orden.*

Demostración. La prueba se sigue trivialmente de la definición del operador T y de la condición Hölder que satisface el núcleo. En efecto,

$$\begin{aligned}
 |[T\psi](x) - [T\psi](x')| &= \left| \int_{\mathbb{R}^d} (k(x, t) - k(x', t))\psi(t) dt \right| \quad \text{(por } T \text{ lineal)} \\
 &\leq \int_{\mathbb{R}^d} |k(x, t) - k(x', t)| |\psi(t)| dt \\
 &\leq |x - x'|^\nu \int_{\mathbb{R}^d} c(t) |\psi(t)| dt && \text{(puesto que } k \text{ es Hölder de orden } \nu) \\
 &\leq C |x - x'|^\nu. && \left(\text{para } C \doteq \int_{\mathbb{R}^d} c(t) |\psi(t)| dt \right) \quad \blacksquare
 \end{aligned}$$

Lema 5.10. *Dados $a \neq 0$, X e Y variables aleatorias se verifica que*

$$E(X|aY) = E(X|Y). \quad (5.16)$$

Demostración. Supondremos sin pérdida de generalidad que X e Y son variables aleatorias continuas. Una demostración análoga sirve para los casos en que alguna de las dos variables aleatorias sea discreta. Supongamos primero que $a > 0$ y sea $Z \doteq aY$. Entonces por definición de función de distribución acumulativa conjunta se tiene que $\forall x, z \in \mathbb{R}$

$$\begin{aligned} F_{X,Z}(x, z) &= P(X \leq x, Z \leq z) \\ &= P(X \leq x, aY \leq z) \\ &= P\left(X \leq x, Y \leq \frac{z}{a}\right) \\ &= F_{X,Y}\left(x, \frac{z}{a}\right). \end{aligned} \tag{5.17}$$

Entonces

$$f_{X,Z}(x, z) = \frac{\partial^2 F_{X,Z}(x, z)}{\partial x \partial z} = \frac{\partial^2 F_{X,Y}\left(x, \frac{z}{a}\right)}{\partial x \partial z} = \frac{1}{a} f_{X,Y}\left(x, \frac{z}{a}\right).$$

De igual modo, se puede probar que $f_Z(z) = \frac{1}{a} f_Y\left(\frac{z}{a}\right)$. Luego, para $z = ay$

$$f_{X|Z}(x|z) = \frac{f_{X,Z}(x, z)}{f_Z(z)} = \frac{\frac{1}{a} f_{X,Y}\left(x, \frac{z}{a}\right)}{\frac{1}{a} f_Y\left(\frac{z}{a}\right)} = \frac{f_{X,Y}(x, y)}{f_Y(y)} = f_{X|Y}(x|y),$$

y, por lo tanto

$$E(X|aY) = E(X|Z) = \int_{\mathbb{R}} x dF_{X|Z}(x|z) = \int_{\mathbb{R}} x dF_{X|Y}(x|y) = E(X|Y).$$

El caso $a < 0$ se obtiene de manera similar teniendo en cuenta el cambio en las desigualdades en (5.17). ■

Lema 5.11. Sea D un conjunto discreto y $\{X_n\}_{n \in D}$ una familia de variables aleatorias normales con media cero. Entonces $\forall n \in D$

$$E \operatorname{Var}(X_n | (X_k)_{\substack{k \in D \\ k \neq n}}) = \inf_{a_k \in \mathbb{R}^+} \operatorname{Var} \left(X_n - \sum_{\substack{k \in D \\ k \neq n}} a_k X_k \right).$$

Demostración. Probaremos este resultado para el caso $\#D = 2$. Una demostración análoga se sigue para el caso general de D numerable arbitrario teniendo en cuenta que:

I) $\operatorname{Var} X_n = E \operatorname{Var}(X_n | (X_k)_{k \neq n}) + \operatorname{Var} E(X_n | (X_k)_{k \neq n})$, y

II) dado que $\{X_n\}_{n \in D}$ es una familia de variables aleatorias normales, existe una sucesión

$$\{\alpha_k\}_{k \in D} \text{ tal que } E(X_n | (X_k)_{k \neq n}) = \sum_{k \neq n} \alpha_k X_k.$$

Sean X e Y variables aleatorias normales con media cero, denotaremos a $\operatorname{Var}(X) = \sigma_X^2$ y $\rho_{XY} = \frac{\operatorname{Cov}(X,Y)}{\sigma_X \sigma_Y}$. Por definición de varianza condicional sabemos que $\operatorname{Var}(X|Y) = \sigma_X^2(1 - \rho_{XY}^2)$.

Por otro lado,

$$\begin{aligned} \text{Var}(X - aY) &= \sigma_X^2 + a^2\sigma_Y^2 - 2a \text{Cov}(X, Y) \\ &= \sigma_X^2 + a^2\sigma_Y^2 - 2a\rho_{XY}\sigma_X\sigma_Y \\ &= \sigma_X^2 + a^2\sigma_Y^2 - 2a\rho_{XY}\sigma_X\sigma_Y + \sigma_X^2\rho_{XY}^2 - \sigma_X^2\rho_{XY}^2 \\ &= \sigma_X^2 - \sigma_X^2\rho_{XY}^2 + (a\sigma_Y - \sigma_X\rho_{XY})^2. \end{aligned} \tag{5.18}$$

Luego, de (5.18) y de la definición de varianza condicional tenemos que

$$\inf_{a \in \mathbb{R}^+} \text{Var}(X - aY) = \sigma_X^2 - \sigma_X^2\rho_{XY}^2 = \sigma_X^2(1 - \rho_{XY}^2) = \text{Var}(X|Y). \quad \blacksquare$$

Bibliografía

1. R. Acar and C. R. Vogel, *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems **10** (1994), 1217–1229.
2. H.-M Adorf, *Hubble space telescope image restoration in its fourth year*, Inverse Problems **11** (1995), 639–653.
3. N. I. Akhiezer and I. M. Glazman, *Theory of linear operators in Hilbert space. Vol. II*, Translated from the Russian by Merlynd Nestell, Frederick Ungar Publishing Co., New York, 1963.
4. H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces*, MPS/SIAM Series on Optimization, vol. 6, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006, Applications to PDEs and optimization.
5. Antoni Buades, Bartomeu Coll, and J-M Morel, *A non-local algorithm for image denoising*, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 2, IEEE, 2005, pp. 60–65.
6. Lucien Le Cam, *Asymptotic methods in statistical decision theory*, Springer Verlag, 1986.
7. Albert Cohen, *Wavelet methods in numerical analysis*, Handbook of numerical analysis **7** (2000), 417–711.
8. Albert Cohen, Ronald DeVore, Gerard Kerkycharian, and Dominique Picard, *Maximal spaces with given rate of convergence for thresholding algorithms*, Applied and Computational Harmonic Analysis **11** (2001), no. 2, 167–191.
9. Albert Cohen, Ronald A DeVore, and Reinhard Hochmuth, *Restricted nonlinear approximation*, Constructive Approximation **16** (2000), no. 1, 85–113.
10. Albert Cohhen, *Numerical analysis of wavelet methods*, vol. 32, Elsevier, 2003.
11. C. E. Creffield, E. G. Klepfish, E. R. Pike, and S. Sankar, *Spectral weight functions for the half-filled hubbord model: a singular value decomposition approach*, Phys. Rev. Lett. **75** (1995), 517–520.
12. Ronald A DeVore, *Degree of nonlinear approximation*, Approximation Theory, VI (1989), 175–201.
13. Ronald A DeVore, Björn Jawerth, and Bradley J Lucier, *Image compression through wavelet transform coding*, IEEE Transactions on information theory **38** (1992), no. 2, 719–746.
14. David L Donoho, *Nonlinear solution of linear inverse problems by wavelet–vaguelette decomposition*, Applied and computational harmonic analysis **2** (1995), no. 2, 101–126.
15. David L Donoho and Iain M Johnstone, *Minimax risk overl p -balls forl p -error*, Probability Theory and Related Fields **99** (1994), no. 2, 277–303.
16. David L Donoho, Iain M Johnstone, et al., *Neo-classical minimax problems, thresholding and adaptive function estimation*, Bernoulli **2** (1996), no. 1, 39–62.
17. H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Mathematics and its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

18. H. W. Engl, A. K. Louis, and W. Rundell, *Inverse problems in geophysics*, SIAM, Philadelphia, 1996.
19. Michael Frazier, Michael W Frazier, Björn Jawerth, and Guido L Weiss, *Littlewood-paley theory and the study of function spaces*, no. 79, American Mathematical Soc., 1991.
20. P. C. Hansen, J. G. Nagy, and D. P. O’Leary, *Deblurring images*, Fundamentals of Algorithms, vol. 3, SIAM, Philadelphia, 2006, Matrices, spectra, and filtering.
21. G. Helmsberg, *Introduction to spectral theory in Hilbert space*, North-Holland Series in Applied Mathematics and Mechanics, Vol. 6, North-Holland Publishing Co., Amsterdam, 1969.
22. J. Honerkamp, *Ill-posed problems in rheology*, Rheol. Acta **28** (1989), 363–371.
23. Barry R Barry R James, *Probabilidade: um curso em nível intermediário*, no. 519.2, 1996.
24. Dhandapani Kannan, *An introduction to stochastic processes*, 1979.
25. A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Applied Mathematical Sciences, vol. 120, Springer-Verlag, New York, 1996.
26. Esther Klann, Peter Maass, and Ronny Ramlau, *Two-step regularization methods for linear inverse problems*, Journal of Inverse and Ill-posed Problems **14** (2006), no. 6, 583–607.
27. Hui-Hsiung Kuo, *White noise distribution theory*, vol. 5, CRC press, 1996.
28. R. L. Lagendijk and J. Biemond, *Iterative identification and restoration of images*, Kluwer, Dordrecht, 1991.
29. G. Landl, T. Langthaler, H. W. Engl, and H. F. Kauffmann, *Distribution of event times in time-resolved fluorescence: the exponential series approach—algorithm, regularization, analysis*, J. Comput. Phys. **95** (1991), no. 1, 1–28.
30. S. Mallat, *A theory for multiresolution signal decomposition*, IEEE Trans. on Pattern Analysis and Machine Intelligence **11** (1989), 674–693.
31. P. Mathé and S. V. Pereverzev, *Optimal discretization of inverse problems in Hilbert scales. Regularization and self-regularization of projection methods*, SIAM J. Numer. Anal. **38** (2001), no. 6, 1999–2021 (electronic).
32. Peter Mathé and Sergei V Pereverzev, *Geometry of linear ill-posed problems in variable hilbert scales*, Inverse problems **19** (2003), no. 3, 789.
33. H. Moritz, *Advanced physical geodesy*, Wichmann, Karlsruhe, 1980.
34. V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl. **7** (1966), 414–417.
35. F. Natterer, *The mathematics of computerized tomography*, B. G. Teubner, Stuttgart, 1986.
36. Victoria Peterson, Hugo Leonardo Rufiner, and Ruben Daniel Spies, *Generalized sparse discriminant analysis for event-related potential classification*, Biomedical Signal Processing and Control **35** (2017), 70–78.
37. Victoria Peterson, Dominik Wyser, Olivier Lambercy, Ruben Spies, and Roger Gassert, *A penalized time-frequency band feature selection and classification procedure for improved motor intention decoding in multi-channel eeg*, Journal of neural engineering **16** (2019), no. 1, 016019.
38. D. L. Phillips, *A technique for the numerical solution of certain integral equations of the first kind*, J. Assoc. Comput. Mach. **9** (1962), 84–97.
39. RE Rolón, IE Gareis, LE Di Persia, RD Spies, and HL Rufiner, *Complexity-based discrepancy measures applied to detection of apnea-hypopnea events*, Complexity **2018** (2018).

-
40. Roman Emanuel Rolon, Luis Darío Larrateguy, Leandro E Di Persia, Ruben D Spies, and Hugo Leonardo Rufiner, *Discriminative methods based on sparse representations of pulse oximetry signals for sleep apnea-hypopnea detection*, Biomedical Signal Processing and Control **33** (2017), 358–367.
 41. R. D. Spies and K. G. Temperini, *Arbitrary divergence speed of the least-squares method in infinite-dimensional inverse ill-posed problems*, Inverse Problems **22** (2006), no. 2, 611–626.
 42. R. D. Spies T. Herdman and K. G. Temperini, *Generalized qualification and qualification levels for spectral regularization methods*, J. Optim. Theory Appl. **141** (2009), no. 3, 547–567.
 43. K. G. Temperini, *Regularización de problemas inversos mal condicionados, calificación generalizada y saturación global.*, Ph.D. thesis, Universidad Nacional del Litoral, 2008.
 44. Andrei Nikolajevits Tihonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. **4** (1963), 1035–1038.
 45. A. N. Tikhonov, *Regularization of incorrectly posed problems*, Soviet Math. Dokl. **4** (1963), 1624–1627.
 46. Hans Triebel, *Theory of function spaces*, Birkhauser, Basel, 1983.
 47. W Ulmer and W Kaissl, *The inverse problem of a gaussian convolution and its application to the finite size of the measurement chambers/detectors in photon and proton dosimetry*, Physics in Medicine & Biology **48** (2003), no. 6, 707.
 48. C. R. Vogel, *Computational methods for inverse problems*, Frontiers in Applied Mathematics, vol. 23, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, With a foreword by H. T. Banks.
 49. C. R. Vogel and M. Oman, *Iterative methods for total variation denoising*, Siam J. Sci. Computing **17** (1996), 227–238.
 50. J. Weidmann, *Linear operators in Hilbert spaces*, Graduate Texts in Mathematics, vol. 68, Springer-Verlag, New York, 1980, Translated from the German by Joseph Szücs.
 51. Norbert Wiener, *Tauberian theorems*, Annals of mathematics (1932), 1–100.