

Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Matemáticas

**RESULTADOS TEÓRICOS Y
PRÁCTICOS SOBRE
PROCESAMIENTO DE SEÑALES**

Memoria presentada para optar al grado de
Doctor en Ciencias
por

Sergio Daniel Vera Rea

Dirigida por
D. Eugenio Hernández Rodríguez

Madrid 2012

A mi madre Sara ...

... porque también hizo de padre.

Por mudarse del extremo norte al extremo sur de la ciudad: *“porque quería que mi hijo estudiara en la universidad”*.

Por llevarme al museo de ciencias naturales cada vez que se lo pedía.

Por educarme en el libre pensamiento.

*Amo el canto de zenzontle
pájaro de cuatrocientas voces,
amo el color del jade
y el enervante perfume de las flores,
pero más amo a mi hermano: el hombre.
Nezahualcóyotl.*

Agradecimientos

Del lado de allá ...

*“... ahora que por fin está bastante claro/ dónde estás y dónde/ estoy/
sé por primera vez/ que tendré fuerzas/ para construir contigo/
una amistad tan piola/ que del vecino/ territorio del amor/
ese desesperado/
empezarán a mirarnos/ con envidia/ y acabaran organizando/
excursiones/ para venir a preguntarnos/ cómo hicimos./”*

Mario Benedetti, *Lovers Go Home*.

*“... hay ríos metafísicos, ella los nada como aquella golondrina está nadando en el aire,
girando alucinada entorno al campanario, dejándose caer para levantarse mejor con el
impulso. Yo describo y defino y deseo esos ríos, ella los nada. Yo los busco, los
encuentro, los miro desde el puente, ella los nada...”*

Julio Cortázar, *Rayuela*.

A Ruth Estévez, por animarme a perseguir mi sueño.

*“... pensaba para mí que el terrible esfuerzo y obligación de educarme a mí mismo
podrían serme dispensados por el destino de encontrar a tiempo un filósofo al que poder
convertir en mi educador... un liberador.”*

Friedrich Nietzsche, *Schopenhauer como educador*.

A Eugenio, por su paciencia.

*¿cuándo somos de veras lo que somos?
... para que pueda ser he de ser otro,
salir de mí, buscarme entre los otros,
los otros que no son si yo no existo,
los otros que me dan plena existencia,
no soy, no hay yo, siempre somos nosotros ...*

Octavio Paz, *Piedra de sol*.

A Keith R., de quien he aprendido mucho,
a Matteo B., por tu generosidad
y a Paco G., por tu bondad.

A mis compañeros de despacho, en especial a Pablo el argentino y Ana la sevillana: el uno porque reconciliamos nuestras diferencias de pensamiento (nunca sabremos si ese gol en fuera de lugar condicionó el partido) y la una por ser una alegría andando, os estimo mucho. A mi quinta, en especial a Paloma y Fernando. A Angélica, mi diablilla, además de algebrista excelente integradora. A Sara, otra gran compañera. Ana Primo, Marilú, Elena, Ángel, Adrián, David, Elías y demás miembros del “*Boca Seniors*” a.k.a. “Lujo en la miseria”. A Bruno B. y Pablo V., compañeros crónicos reiterados.

A mis compañer@s de piso y anex@s: Frauke, Claudie y Jessica, James y Reza (el niño persa), Pamela, Nantke, Claus y Cinzia, María (la de Paco), Pedro y Alessandro. A través de vosotros he conocido un poco el mundo.

A la familia Estévez-Gómez-Arévalo por hacerme sentir uno de los suyos.

A Dragan Vukotic, Fernando Soria, Ana Vargas.

Del lado de acá ...

“El escritor [tecleando en su computadora]: nunca después tuve amigos como aquellos que tuve cuando tenía 12 ¡Jesús! ¿alguien los tiene?”

Stephen King, *The body*.

A Leobardo Anaya, “my younger brother from another mother”.

*“Oh I get by with a little help from my friends,
Mmm, I get high with a little help from my friends,
Mmm, I’m gonna try with a little help from my friends.”*

The Beatles, *With a little help from my friends.*

A Álvaro & Susi, Juancho & Gina, Raúl, Roberto, Dani, Rubén e Isolda...
por tantos años maravillosos... y los que nos quedan!!!!

A los ahora ya señor@s: Daniel E., Luis R., Betito G., Arturo O. “el
chino”, Paquito B., Lulú M., por las grandes desveladas tratando de termi-
nar proyectos. Recuerdos con mucho cariño.

A mis prim@s: Ariel, Neto, Oscar, Roger, Cecil, Armando, Germán[†],
Emelia, Judith, Jaqueline, Isis, Yesiré, Alilat, por aquellas vacaciones cuando
éramos unos mocosos y púberes. A mis tí@s: Dora, Evangelina & Armando,
Emelia & Raymundo, Sofía[†] & Rafaél, Ofelia, y Felipe, por cuidarme de
pequeño. Y ¡cómo no! a mi abuela Margarita[†], ¡¡¡¡responsable de todo este
desmadre!!!! Los quiero un chingo.

A la familia Pérez-Vidal, en especial a Héctor, por ser mi familia hace
muchos años.

Contenido

Lista de Figuras	ix
1 Introducción	1
1.1 Percepción compresiva	1
1.2 Aproximación no lineal restringida	2
1.3 <i>Shearlets</i>	3
2 Algoritmos ambiciosos para la percepción compresiva	5
2.1 Introducción y resultados.	5
2.2 PC y geometría en espacios de grandes dimensiones.	16
2.2.1 Una condición suficiente	19
2.2.2 Concentración de la medida y RIP	23
2.3 Algoritmos ambiciosos en la percepción compresiva	26
2.3.1 Los algoritmos	27
2.3.2 Consecuencias de RIP	32
2.3.2.1 Identificación del soporte en WMP, WOMP y WGP	33
2.3.2.2 Identificación del soporte con condiciones relajadas	39
2.3.3 Convergencia	41
2.3.3.1 Convergencia de GP, WGP y RWGP	42
2.3.3.2 Convergencia de WMP y RWMP	46
2.3.3.3 Convergencia de WOMP y RWOMP	48
2.3.4 Comportamiento de las reglas de selección para algunas matrices aleatorias	50
2.3.4.1 Identificación probabilística del soporte para algoritmos relajados	53

CONTENIDO

2.3.4.2	Identificación probabilística del soporte con la regla de selección $\mathcal{J}(\alpha)$	58
2.4	Experimentos	62
3	Aproximación no lineal restringida en espacios de sucesiones	67
3.1	Introducción	67
3.2	Aproximación no lineal.	69
3.2.1	Algunos resultados en la ANL	69
3.2.2	Un algoritmo, democracia y otras propiedades	72
3.2.3	Resultados más generales sobre ANL	74
3.3	Aproximación no lineal restringida	75
3.3.1	ANL restringida en espacios de Banach	77
3.3.2	Thresholding restringido en H^r y B^p	78
3.4	ANL restringida en espacios de sucesiones	79
3.4.1	Espacios de sucesiones	80
3.4.2	Aproximación no lineal restringida en espacios de sucesiones	81
3.4.3	Dos definiciones importantes	82
3.4.4	Funciones de peso para espacios de Lorentz discretos	82
3.4.5	Espacios de Lorentz discretos generales	85
3.4.6	Desigualdades tipo Jackson	87
3.4.7	ANLR e interpolación real	91
3.4.8	Desigualdades tipo Bernstein	92
3.4.9	Espacios de sucesiones asociados a espacios de suavidad	96
3.4.10	ANLR para espacios de sucesiones de Triebel-Lizorkin	98
3.4.11	Interpolación real para espacios de Triebel-Lizorkin	105
3.5	Espacios de Orlicz	107
3.6	<i>Thresholding</i> restringido en espacios de sucesiones	118
3.6.1	La caracterización por <i>thresholding</i>	120
3.7	Cuantización restringida	123
3.8	Experimentos	127

4	<i>Shearlets</i> y nuevos espacios de distribuciones	143
4.1	Introducción y resultados	143
4.2	<i>Shearlets</i>	149
4.2.1	<i>Shearlets</i> continuas	151
4.2.2	<i>Shearlets</i> discretas	151
4.2.3	<i>Shearlets</i> discretas en el cono	152
4.3	Notación y casi ortogonalidad	155
4.3.1	Casi ortogonalidad	157
4.4	La caracterización	161
4.4.1	Espacios Triebel-Lizorkin AB -anisotrópicos inhomogéneos	161
4.4.2	Dos resultados básicos	162
4.4.3	Acotabilidad de S_ψ y T_ψ	171
4.5	La identidad con los marcos de Parseval suaves	174
4.5.1	El nuevo sistema de <i>shearlets</i> suaves	174
4.5.2	La identidad de reconstrucción en \mathcal{S}'	176
4.6	$\mathbf{F}_{p_1}^{\alpha_1, q_1}$ y $\mathbf{F}_{p_2}^{\alpha_2, q_2}(AB)$	179
4.6.1	Las inclusiones	180
4.6.2	Más relaciones	184
4.7	Pesos	186
Referencias		187

CONTENIDO

Lista de Figuras

2.1	Esquemmatización del problema (P'_1). Se muestran las restricciones en el “cono” y el “tubo”	22
2.2	Reconstrucción de señales dispersas con RWOMP, $\alpha = 0.125$. Por cada par (m, k) , 200 experimentos se generaron con una matriz gaussiana de orden $m \times 256$. Hasta k iteraciones se permitieron.	63
2.3	Reconstrucción de señales dispersas con k -RWOMP, con los mismos parámetros que para la Figura 2.2. Los resultados son mejores que en la Figura 2.2 para RWOMP. Hasta k iteraciones son permitidas.	64
2.4	k-RWOMP and k-RWGP vs CoSaMP, SWGP and GP.	66
3.1	Imagen de Lena original.	130
3.2	Imagen de Lena con ruido.	131
3.3	Aproximación de Lena a partir de la señal original haciendo <i>thresholding</i> usual con parámetros $s = 0$, $\lambda = 30$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 649$ y da un $SNR = 13.02$ con un número de coeficientes $NoC = 12266$	132
3.4	Aproximación de Lena a partir de la señal original haciendo <i>thresholding</i> restringido parámetros $s = 0$, $\lambda = 30$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 681$ y da un $SNR = 12.8$ con un número de coeficientes $NoC = 12267$	133
3.5	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> usual ($s=0$) con umbral $\lambda = 30$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 578$ que da un $SNR = 13.53$ con $NoC = 27581$	134

LISTA DE FIGURAS

3.6	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> restringido ($s=-1$) con umbral $\lambda = 3$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 567$ que da un $SNR = 13.61$ con $NoC = 24380$	135
3.7	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> restringido ($s=-2$) con umbral $\lambda = 0.35$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 630$ que da un $SNR = 13.15$ con $NoC = 26591$	136
3.8	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> usual ($s = 0$) con $\lambda = 30$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 504$ que da un $SNR = 14.12$ con $NoC = 25325$	137
3.9	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> restringido ($s = -1$) con $\lambda = 3$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 471$ que da un $SNR = 14.41$ con $NoC = 24460$	138
3.10	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> usual ($s = 0$) con $\lambda = 20$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 582$ que da un $SNR = 13.5$ con $NoC = 64707$	139
3.11	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> restringido ($s = -1$) con $\lambda = 2$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 503$ que da un $SNR = 14.13$ con $NoC = 43292$	140
3.12	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> usual ($s = 0$) con $\lambda = 40$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 496$ que da un $SNR = 14.19$ con $NoC = 11461$	141
3.13	Estimación de Lena a partir de la señal con ruido haciendo <i>thresholding</i> restringido ($s = -1$) con $\lambda = 5.9$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 525$ que da un $SNR = 13.95$ con $NoC = 11362$	142

LISTA DE FIGURAS

4.1	Esquema del recubrimiento del cono horizontal \mathcal{D}^h en el plano de frecuencias $\hat{\mathbb{R}}^2$	146
4.2	Relación de inclusión entre $\mathbf{F}_p^{\alpha_2, q}(AB)$ y $\mathbf{F}_p^{\alpha_1, q}$	150
4.3	Esquema del recubrimiento del plano \mathbb{R}^2 con paralelogramos en $\mathcal{Q}^{1,2}$. Los paralelogramos con líneas sólidas cubren el paralelogramo $Q_{j, \ell, k} = Q_{0, 1, 0}$	156

LISTA DE FIGURAS

1

Introducción

El procesamiento de señales se basa en el hecho de que muchas clases de funciones o señales “interesantes” pueden representarse con “pocos” elementos de una base (Fourier, Gabor, ondículas, entre otras transformadas tiempo-frecuencia, tiempo-escala) o de un conjunto de funciones “más sencillas” (polinomios ortogonales, splines, etcétera). A partir de esta representación las aplicaciones del procesamiento de señales son: compresión (para almacenamiento y/o transmisión), reducción de ruido, realce de algunas características, reconocimiento de voz y rostro, y un largo etcétera. En el caso más familiar de la compresión, pueden descartarse algunos elementos de esa representación sin que haya diferencia en la percepción visual o auditiva, como sucede con el fenómeno de las bandas críticas usado en los algoritmos de compresión de datos mp3 y otros, ver por ejemplo http://en.wikipedia.org/wiki/Critical_bands. A continuación presentamos un resumen desde una perspectiva más general de los temas sobre los que trata este trabajo, su importancia y una somera exposición de los resultados contenidos en este trabajo. Una descripción más detallada de cada uno de ellos se hace al comienzo de cada capítulo.

1.1 Percepción compresiva

Prácticamente todos los dispositivos con los que interactuamos a diario funcionan de modo digital. Además, dichos dispositivos poseen una memoria y resolución limitadas. Esto implica que en la práctica debemos manejar la información en forma de vectores y matrices finitas.

1. INTRODUCCIÓN

El hecho de que se descarte gran parte de los elementos de la representación plantea la cuestión ¿es posible implementar un procedimiento tal que no se tenga que realizar toda la transformada si sólo nos vamos a quedar con alrededor del 10% de la información? ¿Es posible tener las mismas garantías de error de aproximación que en la compresión tradicional? ¿Qué clase de señales cumple esto?

Realmente este es un problema ya planteado por matemáticos como Gel'fand y Kolmogorov, y parcialmente resuelto por otros como Kashin, Garnaev y Gluskin. Está relacionado con la geometría en grandes dimensiones y las n -anchuras. Los resultados son parciales porque si bien encuentran cotas para el error, no dicen cual es el par codificador-decodificador que puede realizarlo. En el Capítulo 2 haremos explícita dicha relación.

Uno de los dogmas en el procesamiento de señales es el llamado Teorema del muestreo de Whittaker–Nyquist–Kotelnikov–Shannon y establece que, para recuperar una clase de señales con un ancho de banda dado, debe muestrearse al doble de la frecuencia máxima de dicha clase. La percepción compresiva establece que, si la clase de funciones o señales tiene soporte pequeño (y esta es la hipótesis principal), entonces podemos muestrear con una frecuencia mucho más baja que la del Teorema del muestreo.

Nuestra aportación es en la forma que reconstruimos o decodificamos la señal comprimida o codificada. Lo hacemos a través de algunas modificaciones novedosas de algoritmos conocidos y proponemos uno propio. Nos basamos en un pequeño lema propio que es consecuencia de una condición suficiente para la reconstrucción en la teoría de la percepción compresiva. Los resultados se encuentran en el artículo enviado a publicación [57].

1.2 Aproximación no lineal restringida

Un área más clásica relacionada con el procesamiento de señales es la de la teoría de la aproximación. En esta teoría se trata de “resolver” una función o señal posiblemente complicada, llamada *función objetivo*, por funciones más sencillas y más fáciles de computar, llamadas *aproximantes*. El término no lineal es consecuencia de no fijar el conjunto de índices en los que se hace la aproximación. Así, si dos funciones tienen N aproximantes, los índices de éstos no necesariamente deben coincidir. Una de las

cuestiones principales en la teoría de la aproximación es la relación que existe entre la pertenencia a algún espacio de sucesiones de los coeficientes de la representación y la pertenencia a algún espacio de suavidad de la función. Por ejemplo, en las series de Fourier las discontinuidades de la función incrementan los coeficientes de las altas frecuencias. Otra cuestión importante es la velocidad con la que el error disminuye conforme usamos más y más aproximantes.

En la aproximación no lineal restringida se controlan los coeficientes de la aproximación a la representación de la función en un espacio teniendo en cuenta las propiedades de otro espacio. Dos funciones o señales pueden tener una norma L^2 muy parecida pero ser visualmente muy diferentes. Una de las consecuencias es que probamos, al menos empíricamente, que la norma L^2 no es necesariamente la mejor para medir la calidad visual. Además, la calidad visual se incrementa, en algunos casos, si nos aproximamos a la función usando los coeficientes que mejor se aproximen a la función en un espacio de suavidad.

Nuestro aporte en este capítulo es la generalización de algunos resultados de la aproximación no lineal restringida trabajando en los espacios de sucesiones. Estos pueden ser “transferidos” a los espacios de funciones. Algunos de estos resultados son la caracterización de algunos espacios de aproximación como ciertos espacios de interpolación. Se encuentran también nuevos resultados en interpolación. Algunos resultados de este capítulo se encuentran en la publicación [58].

1.3 *Shearlets*

Las diversas transformadas de Fourier son muy poderosas y son usadas y estudiadas no sólo en matemáticas sino en la ingeniería, la física y otras ciencias. Una desventaja, sin embargo, es que las funciones exponenciales complejas no tienen un soporte compacto en el espacio, lo cual impide un análisis local fino. El descubrimiento y desarrollo de las ondículas a mediados de los 80's permitió sortear este problema al demostrarse que podían crearse bases para L^2 (o marcos de forma más general) con soporte compacto en el espacio. Más aún, dependiendo de la regularidad y decaimiento de dichas ondículas, éstas generan bases para muchos otros espacios de funciones a partir de un análisis multirresolución (AMR). A pesar del gran éxito que han tenido las ondículas en campos como las EDP's o el procesamiento de señales, un análisis más geométrico

1. INTRODUCCIÓN

multi dimensional está limitado por el hecho que las ondículas en dimensiones mayores a 1 generalmente se obtienen a partir de productos tensoriales. Esto implica que las ondículas “ven” las singularidades sin “describir” por completo su geometría. En las aplicaciones que nos ocupan (procesamiento de señales), es deseable poder percibir con más precisión la orientación de dicha singularidad. En los últimos años se han desarrollado variantes de las ondículas que son más direccionales. Algunos de éstos incluyen los bancos de filtros, las *curvelets* y las *contourlets*, para nombrar unas pocas. En cuanto a las *contourlets*, como son construidas en un ámbito discreto y finito, carecen de flexibilidad de diseño y, para algunas aplicaciones, suponen que existen funciones suaves con soporte compacto espacial que aproximan a una partición del plano de las frecuencias como en la Subsección 4.2.3. Por otra parte, las *curvelets* se construyen en coordenadas polares, así que su implementación es más bien difícil.

En [55], Guo, Lim, Labate, Weiss y Wilson, introdujeron las ondículas con dilatación compuesta. Este tipo de representación aprovecha la teoría de los sistemas afines en \mathbb{R}^n , por lo que provee una transición natural de la representación continua al ámbito discreto parecido a una base (como es el caso de las ondículas). Relacionado con las ondículas de dilatación compuesta se tiene el sistema de *shearlets* que provee marcos de Parseval para $L^2(\mathbb{R}^2)$ o subespacios de éste (dependiendo del muestreo discreto de los parámetros, ver Subsecciones 4.2.2 y 4.2.3). Pueden consultarse una larga cantidad de aplicaciones no sólo para el procesamiento digital de imágenes en <http://www.shearlet.org>. Las *shearlets* producen representaciones óptimas de funciones en $C^2(\mathbb{R}^2)$ excepto en discontinuidades a lo largo de curvas C^2 , ver [53].

En este trabajo demostramos que, como en el caso de las ondículas, se pueden caracterizar espacios (altamente anisotrópicos) del tipo Triebel-Lizorkin usando los coeficientes de las “*shearlets* en el cono”. Relacionamos también estos nuevos espacios con los clásicos. La línea de argumentación sigue la transformada φ de Frazier y Jawerth en [41]. Puede consultarse una descripción detallada de los resultados obtenidos en la primera sección del Capítulo 4. Los resultados descritos en el Capítulo 4 son la base de la publicación [104].

2

Algoritmos ambiciosos para la percepción compresiva

Pluralitas non est ponenda sine necessitate.

William de Ockham.

... Unwin preguntó por qué, dócilmente.

-Por diversas razones -fue la respuesta-.

En primer lugar, esa casa es un laberinto.

En segundo lugar, la vigilaban un escalvo y un león.

En tercer lugar, se desvaneció un tesoro secreto.

En cuarto lugar, el asesino estaba muerto cuando el asesinato ocurrió.

En quinto lugar...

Unwin, cansado, lo detuvo.

- No multipliques los misterios -le dijo-. Éstos deben ser simples...

Jorge Luis Borges, *Abenjacán el Bojarí, muerto en su laberinto.*

2.1 Introducción y resultados.

Sea $\Psi \in \mathbb{C}^{N \times N}$ una matriz hermítica. Como las columnas forman una base de \mathbb{C}^N , cualquier elemento $\mathbf{x} \in \mathbb{C}^N$ puede representarse como $\mathbf{x} = \sum_{n=1}^N \langle \mathbf{x}, \psi_n \rangle \psi_n$, donde $\{\psi_n\}_{n=1}^N$ son las columnas (átomos) de la matriz Ψ . Si los elementos $y_i = \langle \mathbf{x}, \psi_i \rangle$ de $\mathbf{y} = \Psi^* \mathbf{x}$ (donde $(\cdot)^*$ significa transpuesta conjugada) reordenados de forma decreciente en magnitud cumplen $|y_i| \leq R i^{-1/p}$, $0 < p < 1$, se dice que \mathbf{x} es compresible (en la representación Ψ) y el ínfimo de los R tal que esto sucede para toda i es la (cuasi-)

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

norma débil $w\ell^p$ de \mathbf{y} . Esta condición es equivalente a $|\{i : |y_i| \geq \lambda\}| \lambda^p \leq R^p$ para todo $\lambda > 0$. Si sólo k elementos de \mathbf{y} son no nulos, se dice que \mathbf{y} es k -disperso.

La **compresión tradicional** de un vector $\mathbf{x} \in \mathbb{C}^N$ se hace, *grosso modo*, en dos pasos. A saber:

- CT1) Se realiza la operación lineal $\mathbf{y} = \mathbf{\Psi}^* \mathbf{x} \in \mathbb{C}^N$ donde $\mathbf{\Psi} \in \mathbb{C}^{N \times N}$ es la matriz de una transformada (Fourier, coseno, ondícula, ...) generalmente hermítica.
- CT2) Se buscan y guardan los k -elementos más grandes en valor absoluto de \mathbf{y} junto con sus correspondientes índices (*transform coding*).

Las **desventajas** de la compresión tradicional son:

- 1) Se realizan N productos internos (“mediciones”) pero sólo se guardan $k \ll N$ elementos¹.
- 2) Es un algoritmo adaptivo en el sentido de que los índices que se guardan dependen de la señal.

Para estimar (*i.e.*, posible reconstrucción exacta en el caso k -disperso o aproximación en el caso compresible) la señal original \mathbf{x} se realiza la transformada inversa $\tilde{\mathbf{x}} = \mathbf{\Psi}^{*-1} \mathbf{y}_k$, donde $\mathbf{y}_k \in \mathbb{C}^N$ es el vector \mathbf{y} restringido a sus k -mayores (en módulo) elementos y el resto de elementos de \mathbf{y}_k son cero, y en donde $\mathbf{\Psi}^{*-1} = \mathbf{\Psi}$ puesto que es una matriz hermítica. Usando la igualdad de Parseval y la definición de la norma débil $w\ell^p$ se ve que el error de estimación, tomando k elementos de la base, en la norma ℓ^2 es

$$\begin{aligned} \|\mathbf{x} - \tilde{\mathbf{x}}\|_{\ell^2} &= \|\mathbf{y} - \mathbf{y}_k\|_{\ell^2} = \left(\sum_{i=k+1}^{\infty} |y_i|^2 \right)^{1/2} \\ &\leq \left(\sum_{i=k+1}^{\infty} \|\mathbf{y}\|_{w\ell^p}^2 i^{-2/p} \right)^{1/2} = c_p \|\mathbf{y}\|_{w\ell^p} k^{\frac{1}{2} - \frac{1}{p}}. \end{aligned} \quad (2.1)$$

Nuestra **desiderata** es encontrar un procedimiento tal que:

- D1) El número de “mediciones” (productos internos) m sea comparable con el tamaño k de la información a guardar, es decir, la matriz de mediciones $\mathbf{\Phi}$ es de tamaño $m \times N$.

¹Dependiendo de las aplicaciones la información comprimida es de alrededor del 10% de los datos originales.

D2) Sea no adaptivo, es decir, que no dependa de la señal original.

D3) Sea práctico y que el error de estimación sea comparable con el tradicional.

Para la **percepción compresiva** (PC) supondremos de ahora en adelante que el vector $\mathbf{x} \in \mathbb{C}^N$ ya es k -disperso o compresible ya que, como veremos más adelante, la compresión se hará mediante cierta matriz “especial”, obteniendo $\mathbf{b} = \mathbf{A}\mathbf{y} = \mathbf{A}\Psi\mathbf{x} = \Phi\mathbf{x} \in \mathbb{C}^m$, $m \ll N$.

Sea $\Phi \in \mathbb{C}^{m \times N}$ una matriz con $m \ll N$. Dada una señal $\mathbf{x} \in \mathbb{C}^N$ obtenemos las mediciones $\mathbf{y} = \Phi\mathbf{x} \in \mathbb{C}^m$ mediante la matriz de compresión Φ . En la teoría de la **percepción compresiva** (PC) el problema es reconstruir o aproximar la señal original \mathbf{x} a partir de las mediciones \mathbf{y} . Es claro que las propiedades que tenga la matriz Φ han de influir en la posibilidad de hallar una “buena” aproximación de la señal original, o en la posibilidad de recuperarla completamente.

El sistema de ecuaciones lineales $\mathbf{y} = \Phi\mathbf{x}$, con incógnitas las componentes de \mathbf{x} , tiene al menos una solución (la señal original). Pero, como $m \ll N$, tiene infinitas soluciones. Todas sus soluciones estarán en el subespacio

$$\mathcal{F}(\mathbf{y}) := \{\mathbf{x} \in \mathbb{R}^N : \Phi\mathbf{x} = \mathbf{y}\} = \mathbf{x}_0 + \mathcal{N}$$

donde $\mathbf{x}_0 \in \mathcal{F}(\mathbf{y})$ y $\mathcal{N} = \mathcal{N}(\Phi) = \{\mathbf{z} \in \mathbb{R}^N : \Phi\mathbf{z} = \mathbf{0}\}$ es el núcleo de Φ .

Una forma de hallar un $\tilde{\mathbf{x}} \in \mathcal{F}(\mathbf{y})$ es resolviendo el siguiente problema de optimización cuadrática:

$$(P_2) \quad \min_{\mathbf{z} \in \mathbb{R}^N} \|\mathbf{z}\|_{\ell^2} \quad \text{sujeto a} \quad \Phi\mathbf{z} = \mathbf{y}.$$

Es fácil ver que la solución de (P_2) es $\tilde{\mathbf{x}} = \Phi^*(\Phi\Phi^*)^{-1}\mathbf{y}$ cuando $\Phi\Phi^*$ es no singular (de hecho, $\Phi\tilde{\mathbf{x}} = \Phi\Phi^*(\Phi\Phi^*)^{-1}\mathbf{y} = \mathbf{y}$ y $\tilde{\mathbf{x}}$ es perpendicular a \mathcal{N} ya que $\langle \tilde{\mathbf{x}}, \mathbf{z} \rangle = \langle \Phi^*(\Phi\Phi^*)^{-1}\mathbf{y}, \mathbf{z} \rangle = \langle (\Phi\Phi^*)^{-1}\mathbf{y}, \Phi\mathbf{z} \rangle = 0$ cuando $\mathbf{z} \in \mathcal{N}$). Desafortunadamente, cuando la señal \mathbf{x} es k -dispersa ($k \ll N$) el problema (P_2) casi nunca encuentra la solución correcta.

Puesto que (P_2) no encuentra soluciones k -dispersas, una posible alternativa es buscar el vector más disperso en $\mathcal{F}(\mathbf{y})$; esto es, $\tilde{\mathbf{x}}$ que resuelva

$$(P_0) \quad \min_{\mathbf{z} \in \mathbb{R}^N} \|\mathbf{z}\|_{\ell^0} \quad \text{sujeto a} \quad \Phi\mathbf{z} = \mathbf{y}. \quad (2.2)$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Desafortunadamente, resolver (P_0) requiere estudiar todas las $\binom{N}{k}$ posibles combinaciones de posiciones no nulas de \mathbf{x} y es un problema NP-completo.

La idea seminal de la percepción compresiva se encuentra, principalmente, en dos trabajos independientes de Donoho [36] y de Candès-Romberg-Tao [17]. En ambos artículos se propone el problema de optimización ¹

$$(P_1) \quad \min_{\mathbf{z} \in \mathbb{R}^N} \|\mathbf{z}\|_{\ell^1} \quad \text{sujeto a} \quad \Phi \mathbf{z} = \mathbf{y}, \quad (2.3)$$

para hallar una estimación $\tilde{\mathbf{x}}$ de \mathbf{x} . La matriz $\Phi : \mathbb{R}^N \mapsto \mathbb{R}^m$, $m \ll N$, es una matriz “especial” que cumple tres condiciones en [36], y en [17] es una matriz que se obtiene tomando aleatoriamente m filas de la matriz de Fourier discreta \mathcal{F} de tamaño $N \times N$, esto es

$$\mathcal{F} = \frac{1}{\sqrt{N}} \begin{pmatrix} e^{2\pi \frac{nl}{N}} \\ n = 0, \dots, N-1 \\ l = 0, \dots, N-1 \end{pmatrix}.$$

Uno de los resultados en [17] establece que: si $m = O(k \log N)$ el algoritmo (P_1) logra reconstruir **exactamente** con **gran probabilidad** (mayor que $1 - O(N^{-c})$, para un parámetro c) cualquier vector k -disperso, $k \ll N$.

Señalamos que (P_1) es un problema de optimización convexa que puede reducirse a un algoritmo de programación lineal (conocido como *basis pursuit*) cuya complejidad computacional es $O(N^3)$.

Además, si $\tilde{\mathbf{x}}$ es la solución de (P_1) y $\mathbf{x} \in w\ell^p$, $0 < p < 1$, Candès-Tao prueban en [19] que con el mismo número de mediciones $m = O(k \log N)$ se tiene

$$\|\mathbf{x} - \tilde{\mathbf{x}}\|_{\ell^2(\mathbb{R}^N)} \leq C_p \|\mathbf{x}\|_{w\ell^p} \left(\frac{m}{\log N}\right)^{\frac{1}{2} - \frac{1}{p}}. \quad (2.4)$$

Volvamos de nuevo a las condiciones que debe satisfacer la matriz Φ para que el problema (P_1) tenga solución. La que se usa comunmente es la propiedad de que $\Phi \in \mathbb{C}^{m \times N}$ satisfaga la **propiedad de la isometría restringida** (RIP). Dado $k \in \mathbb{Z}_+$, $k \leq m$, se dice que Φ satisface RIP con parámetro δ_k , $0 < \delta_k < 1$, si se cumple que

$$(1 - \delta_k) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2 \leq (1 + \delta_k) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2, \quad (2.5)$$

para todos los vectores $\mathbf{x} \in \mathbb{R}^N$ que sean k -dispersos. La condición RIP (2.5) mide cómo de cerca los vectores columna $\phi_i, i = 1, \dots, N$, de Φ se comportan como un

¹Usado ya en sismología, estadística y procesamiento de señales y también conocido como *Basis Pursuit* cuando la matriz Φ es una base o, en general, un marco (sobrerrepresentación).

sistema ortonormal en \mathbb{R}^N , pero sólo cuando nos restringimos a combinaciones lineales dispersas formadas a lo sumo por k vectores.

El Teorema 1.4 de [20] muestra que si $\Phi \in \mathbb{C}^{m \times N}$ es una matriz que satisface $\delta_k + \delta_{2k} + \delta_{3k} < 1$ la solución del problema (P_1) dado en (2.3) reconstruye de manera exacta cualquier vector k -disperso $\mathbf{x} \in \mathbb{R}^N$ a partir de las mediciones $\mathbf{y} = \Phi \mathbf{x}$.

En [20] y [36] se prueba que si Φ es una matriz aleatoria cuyos elementos se obtienen a partir de una distribución gaussiana de media 0 y desviación típica $1/\sqrt{m}$, o bien a partir de una distribución de Bernoulli (con valores $\{-1, 1\}$ con probabilidad 0.5 en cada caso), y si $m \geq Ck \log(N/k)$, la citada matriz satisface RIP de orden k con gran probabilidad. Para el caso de matrices $m \times N$ obtenidas a partir de la matriz discreta de Fourier de orden $N \times N$ seleccionando m filas aleatoriamente, se debe tener $m \geq Ck(\log N)^6$ (leves mejoras de esta cota pueden encontrarse en [92]).

Una forma elegante de probar RIP para matrices gaussianas y de Bernoulli (por supuesto, en probabilidad) es hacer uso de desigualdades de **concentración de la mediana** para estas matrices (ver (2.32)), lo que se prueba en [27] para las matrices gaussianas y en [1] para las de Bernoulli. A partir de aquí, Baraniuk-Davenport-DeVore-Wakin prueban en [6] que se verifica RIP para estas matrices con gran probabilidad. La demostración de este resultado se detalla en la Sección 2.2.2.

Se pueden construir matrices deterministas Φ que satisfagan RIP, como en DeVore [30], pero el rango de valores válidos de k es como \sqrt{m} , lejos de los valores proporcionados por las matrices aleatorias anteriormente descritas.

Otra forma de hallar una estimación $\tilde{\mathbf{x}}$ de una señal \mathbf{x} , conocidas Φ e \mathbf{y} ($\mathbf{y} = \Phi \mathbf{x}$), es usar algoritmos ambiciosos. Estos algoritmos se han usado en la teoría de aproximación no lineal, incluso antes de que comenzara a desarrollarse la teoría de la percepción compresiva (ver [64], [7] y, más recientemente, [34], [96], [97], [98]). En los algoritmos ambiciosos puros se seleccionan sucesivamente, y uno por uno, índices con alguna propiedad de mayoración (por ejemplo, que los elementos asociados a esos índices se ordenen de mayor a menor en alguna norma). En nuestro caso vamos a permitir una regla de selección más débil, que permitirá seleccionar más índices en cada paso.

Hacemos precisas las afirmaciones anteriores. Conocemos la matriz $\Phi \in \mathbb{C}^{m \times N}$, $m < N$, y la observación $\mathbf{y} \in \mathbb{R}^m$. Queremos hallar una estimación $\tilde{\mathbf{x}}$ de $\mathbf{x} \in \mathbb{R}^N$ que

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

satisface

$$\mathbf{y} = \Phi \mathbf{x}$$

o, si fuera posible, identificar \mathbf{x} . Se comienza, en el paso $n = 0$, con las estimaciones $\mathbf{y}^0 = 0$, $\mathbf{x}^0 = 0$ y se pone como residuo $\mathbf{r}^0 = \mathbf{y}$. Mediante algún algoritmo (más adelante expondremos varios) se van obteniendo

$$\mathbf{y}^n, \mathbf{x}^n, \mathbf{r}^n, \quad n = 1, 2, 3, \dots$$

las distintas actualizaciones. Hay varias formas de terminar el algoritmo: o bien cuando $\mathbf{r}^n = 0$, o bien cuando $\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq \varepsilon$ (fijado), o bien después de un número prefijado de iteraciones acorde con el algoritmo.

Para calcular las actualizaciones utilizaremos dos reglas de selección débiles. En la primera de ellas, para $n = 1, 2, 3, \dots$, elegimos, para $0 < \alpha \leq 1$,

$$\mathcal{J}_n(\alpha) = \{i \in \{1, \dots, N\} : |\langle \phi_i, \mathbf{r}^{n-1} \rangle| \geq \alpha \sup_{j \in \{1, \dots, N\}} |\langle \phi_j, \mathbf{r}^{n-1} \rangle|\}, \quad (2.6)$$

donde ϕ_i son las columnas de Φ . Observar que

$$\sup_{j \in \{1, \dots, N\}} |\langle \phi_j, \mathbf{r}^{n-1} \rangle| = \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}. \quad (2.7)$$

En la regla de selección (2.6) siempre se tiene $\mathcal{J}_n(\alpha) \neq \emptyset$ para $\mathbf{r}^{n-1} \neq \mathbf{0}$.

En la segunda regla de selección, para $n = 1, 2, 3, \dots$, y $0 < \tilde{\alpha} \leq 1$ elegimos

$$\tilde{\mathcal{J}}_n(\tilde{\alpha}) = \{i \in \{1, \dots, N\} : |\langle \phi_i, \mathbf{r}^{n-1} \rangle| \geq \tilde{\alpha} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}\}. \quad (2.8)$$

En este caso puede probarse (ver Sección 2.3.2.2) que si

$$\tilde{\alpha} \leq \frac{(1 - \delta_k)^{1/2}}{\sqrt{k}} \quad (2.9)$$

entonces $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ en cada iteración ($\mathbf{r}^{n-1} \neq \mathbf{0}$), siempre que Φ cumpla RIP con parámetro δ_k y los residuos $\mathbf{r}^n \in \text{span}(\Phi_\Gamma) = \{\Phi_\Gamma \mathbf{z} : \mathbf{z} \in \mathbb{R}^N, \text{sop}(\mathbf{z}) \subset \Gamma\}$ y $|\Gamma| \leq k$.

A continuación describimos los algoritmos de tipo ambicioso que consideraremos en este capítulo. El primero es del tipo **Matching Pursuit** (MP) o **búsqueda de correspondencia**. Con cualquiera de las reglas (2.6) o (2.8) se selecciona un conjunto de índices \mathcal{J}_n ($\mathcal{J}_n(\alpha)$ ó $\tilde{\mathcal{J}}_n(\tilde{\alpha})$). Las mediciones se actualizan con

$$\mathbf{y}^n = \mathbf{y}^{n-1} + \sum_{i \in \mathcal{J}_n} \langle \phi_i, \mathbf{r}^{n-1} \rangle \phi_i,$$

la estimación de la señal \mathbf{x} se hace con $\mathbf{x}^n = (x_i^n)$ donde

$$x_i^n = \begin{cases} x_i^{n-1} + \langle \phi_i, \mathbf{r}^{n-1} \rangle & \text{si } i \in \mathcal{J}_n \\ x_i^{n-1} & \text{si } i \notin \mathcal{J}_n, \end{cases}$$

y el residuo se actualiza con

$$\mathbf{r}^n = \mathbf{r}^{n-1} + \sum_{i \in \mathcal{J}_n} \langle \phi_i, \mathbf{r}^{n-1} \rangle \phi_i.$$

Es fácil probar, por recursión, que $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$. Cuando se usa la regla de selección (2.6) el algoritmo se llama WMP (Weak Matching Pursuit) y si se usa (2.8) se llama RWMP (Relaxed Weak Matching Pursuit).

El segundo es del tipo **Orthogonal Matching Pursuit** (OMP) o **búsqueda de correspondencia ortogonal**. Una vez obtenido el conjunto de índices \mathcal{J}_n ($\mathcal{J}_n(\alpha)$ o $\tilde{\mathcal{J}}_n(\tilde{\alpha})$) las mediciones se actualizan con

$$\mathbf{y}^n = P_{\Gamma^n} \mathbf{y}$$

donde P_{Γ^n} es la proyección ortogonal de \mathbf{y} sobre $V_n = \{\Phi_{\Gamma^n} \mathbf{x} : \mathbf{x} \in \mathbb{R}^N\}$, y $\Gamma^n = \cup_{j=1}^n \mathcal{J}_j$ y $\Phi_{\Gamma^n} \in \mathbb{C}^{m \times |\Gamma^n|}$ denota la submatriz de Φ cuyas columnas tienen índices de Γ^n . Es un ejercicio sencillo de álgebra lineal probar que

$$\mathbf{y}^n = \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y}$$

donde $\Phi_{\Gamma^n}^\dagger = (\Phi_{\Gamma^n}^* \Phi_{\Gamma^n})^{-1} \Phi_{\Gamma^n}^*$ es la pseudo-inversa de Φ_{Γ^n} , siempre que exista. La actualización de la estimación de la señal \mathbf{x} se hace con \mathbf{x}^n que es el vector cuyas coordenadas con índices en Γ^n son $\mathbf{x}_{\Gamma^n}^n = \Phi_{\Gamma^n}^\dagger \mathbf{y}$ y son nulas para el resto de los índices. Los residuos se actualizan con

$$\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n.$$

Si se usa la regla de selección (2.6) el algoritmo se llama WOMP (Weak Orthogonal matching Pursuit) y si se usa (2.8) se llama RWOMP (Relaxed Weak Orthogonal Matching Pursuit).

Una de las ventajas de los algoritmos OMP es que, debido a la ortogonalidad, los índices que se seleccionan en el paso n son distintos de los previamente seleccionados. Por otro lado, la complejidad del cálculo es mayor que en los algoritmos MP ya que es necesario calcular la pseudo-inversa que requiere invertir y multiplicar matrices.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

El tercero, y último, es del tipo **Gradient Pursuit** (GP) o **búsqueda del gradiente** (ver Sección 2.3.1 para una explicación de esta nomenclatura). Una vez obtenido el conjunto de índices \mathcal{J}_n ($\mathcal{J}_n(\alpha)$ ó $\tilde{\mathcal{J}}_n(\tilde{\alpha})$) se considera

$$\Gamma^n = \cup_{j=1}^n \mathcal{J}_j.$$

Sean

$$\mathbf{d}_{\Gamma^n}^n = \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}, \quad a^n = \frac{\langle \mathbf{r}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n \rangle}{\|\Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n\|_{\ell^2(\mathbb{R}^m)}^2}$$

(ver su interpretación en la Sección 2.3.1). La estimación de la señal \mathbf{x}^n es cero fuera de Γ^n y en Γ^n ,

$$\mathbf{x}_{\Gamma^n}^n = \mathbf{x}^{n-1} + a^n \mathbf{d}_{\Gamma^n}^n.$$

La aproximación a la observación es

$$\mathbf{y}^n = \mathbf{y}^{n-1} + a^n \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n,$$

mientras que los residuos se actualizan con

$$\mathbf{r}^n = \mathbf{r}^{n-1} - a^n \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n.$$

Obsérvese que para $i \in \Gamma^n$

$$x_i^n = x_i^{n-1} + a^n \langle \phi_i, \mathbf{r}^{n-1} \rangle.$$

Puede probarse por recursión que $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$. Si se usa la regla de selección (2.6) el algoritmo se llama WGP (Weak Gradient Pursuit) y si se usa (2.8) se llama RWGP (Relaxed Weak Gradient Pursuit).

Después de dar algunas consecuencias de RIP en la Sección 2.3.2, nos adentramos en los resultados principales de este capítulo sobre señales k -dispersas ($\text{sop}(\mathbf{x}) \subset \Gamma^\sharp$ con $|\Gamma^\sharp| \leq k$). Por un lado, en la Sección 2.3.2 se estudian condiciones suficientes para que los seis algoritmos anteriormente descritos identifiquen índices del soporte de un vector k -disperso $\mathbf{x} \in \mathbb{R}^N$. Por otro lado, en la Sección 2.3.3 se estudia la convergencia de estos algoritmos para señales k -dispersas.

Con respecto a la identificación de elementos en el soporte de \mathbf{x} y para los algoritmos WMP, WOMP y WGP (los que se hacen con la regla de selección $\mathcal{J}_n(\alpha)$ dada en (2.6))

el principal resultado es el Teorema 2.3.5. En él se prueba que si Φ es una matriz que satisface RIP con parámetro δ_{k+1} , la condición

$$\delta_{k+1} < \alpha \frac{1 - \delta_k}{\sqrt{k}} \quad (2.10)$$

es suficiente para asegurar que $\mathcal{J}_n(\alpha) \subset \Gamma^\sharp = \text{sop}(\mathbf{x})$, por lo que estos algoritmos identifican siempre índices de Γ^\sharp .

En el caso de WOMP este resultado permite identificar Γ^\sharp en como mucho $|\Gamma^\sharp| = k$ iteraciones ya que, como se ha observado anteriormente, los elementos de $\mathcal{J}_n(\alpha)$ son distintos de los previamente seleccionados. Además, encontrado Γ^\sharp , puede probarse que

$$\mathbf{x}_{\Gamma^\sharp} = \mathbf{x}_{\Gamma^\sharp}^n = \Phi_{\Gamma^\sharp}^\dagger \mathbf{y} = \Phi_{\Gamma^\sharp}^\dagger \mathbf{y}$$

con lo que se identifica \mathbf{x} completamente.

Con respecto a los algoritmos RWMP, RWOMP y RWGP (los que se hacen con la regla de selección $\tilde{\mathcal{J}}(\tilde{\alpha})$ dada en (2.8)) y para identificar elementos de $\Gamma^\sharp = \text{sop}(\mathbf{x})$, el principal resultado es el Teorema 2.3.12. Se prueba que $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ y

$$\delta_{k+1} < \tilde{\alpha}(1 - \delta_k)^{1/2} \quad (2.11)$$

es suficiente para asegurar $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \subset \Gamma^\sharp$ en estos algoritmos.

En la Nota 2.3.13 se prueba que la condición

$$\tilde{\alpha} \leq \frac{(1 - \delta_k)^{1/2}}{\sqrt{k}} \quad (2.12)$$

implica $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ para todo n . Junto con (2.11), tenemos

$$\frac{\delta_{k+1}}{(1 - \delta_k)^{1/2}} < \frac{(1 - \delta_k)^{1/2}}{\sqrt{k}}$$

lo que produce la restricción

$$\sqrt{k} \leq \frac{1 - \delta_k}{\delta_{k+1}}.$$

En la Sección 2.3.3 se estudia la convergencia de estos algoritmos cuando \mathbf{x} es una señal k -dispersa. Para los algoritmos del tipo GP (WGP y RWGP) el resultado principal es el Teorema 2.3.17. Con la condición (2.10) para WGP y (2.11) junto con $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ para RWGP se tiene

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq C_k \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \quad (2.13)$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

con $C_k = (1 - \frac{1-\delta_k}{k(1+\delta_k)})^{1/2} < 1$, y $\text{sop}(\mathbf{x}) = \Gamma^\sharp$ con $|\Gamma^\sharp| \leq k$; por tanto la energía de los residuos tiende exponencialmente a cero:

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq C_k^n \|\mathbf{r}^0\|_{\ell^2(\mathbb{R}^m)} = C_k^n \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}. \quad (2.14)$$

De aquí se deduce, usando las propiedades RIP de Φ , que

$$\|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)} \leq \left(\frac{1-\delta_k}{1+\delta_k}\right)^{n/2} C_k^n \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}, \quad n = 1, 2, 3, \dots \quad (2.15)$$

para toda señal k -dispersa \mathbf{x} . Para que haya convergencia a 0 de la parte derecha de (2.15) se ha de tener $\delta_k < \frac{1}{2k+1}$.

Para los algoritmos del tipo MP (WMP y RWMP) el resultado principal es el Teorema 2.3.23. Con la condición (2.10) para WMP y (2.11) junto con $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ para RWMP se tiene

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq C'_k \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \quad (2.16)$$

siempre que \mathbf{x} sea k -dispersa, con

$$C'_k = \left(1 - \frac{(1-\delta_k)^2}{k}\right)^{1/2} < 1.$$

Iterando (2.16) y usando las propiedades que se deducen de RIP, se tiene

$$\|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)} \leq \left(\frac{1-\delta_k}{1+\delta_k}\right)^{n/2} C_k'^n \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}, \quad n = 1, 2, 3, \dots \quad (2.17)$$

para toda señal k -dispersa \mathbf{x} . En este caso, para que haya convergencia a 0 de la parte derecha de (2.17) se ha de tener $\delta_k < \frac{1}{k+2}$.

Para WOMP y RWOMP sabemos que bajo las hipótesis (2.10) para el primero de ellos y (2.11) junto con $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ para el segundo, $\Gamma^k = \Gamma^\sharp$, con lo que $\mathbf{r}^k = \mathbf{0}$ para señales k -dispersas.

Las matrices $\Phi \in \mathbb{R}^{m \times N}$ que permiten recuperar una señal k -dispersa \mathbf{x} con el problema de optimización (P1) son matrices aleatorias con $m \geq Ck \log(N/k)$ y la recuperación se obtiene con gran probabilidad. En la Sección 2.3.4 usamos las matrices aleatorias $\Phi(\omega) \in \mathbb{R}^{m \times N}$ que satisfacen

(M1) Las columnas de $\Phi(\omega)$ son estadísticamente independientes.

(M2) Para cada columna $\phi_j(\omega)$, $j = 1, \dots, N$, de $\Phi(\omega)$ se tiene $\mathbb{E} \left\{ \|\phi_j(\omega)\|_{\ell^2(\mathbb{R}^m)}^2 \right\} = 1$.

(M3) Sea $\mathbf{u} \in \mathbb{R}^m$ un vector con $\|\mathbf{u}\|_{\ell^2(\mathbb{R}^m)} \leq 1$. Si $\phi(\omega)$ es una columna de $\Phi(\omega)$ independiente de \mathbf{u} ,

$$\mathbb{P}\{|\langle \phi(\omega), \mathbf{u} \rangle| \geq \varepsilon\} \leq q_1 e^{-c_1 \varepsilon^2 m},$$

con q_1, c_1 constantes, $q_1 \geq 1$.

(M4) Para todo conjunto $\Gamma \subset \{1, \dots, N\}$ con $|\Gamma| \leq k < N$ y para todo $\mathbf{r} \in \text{span}(\Phi_\Gamma(\omega))$ se tiene

$$\mathbb{P}\left\{\|\Phi_\Gamma^*(\omega)\mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)} \geq \frac{1}{2}\|\mathbf{r}\|_{\ell^2(\mathbb{R}^m)}\right\} \geq 1 - q_2 D^k e^{-c_2 m},$$

con q_2, D y c_2 constantes, $q_2, D > 1$.

Estas condiciones, (M1), (M2), (M3) y (M4) son satisfechas por las matrices aleatorias gaussianas $\Phi(\omega)$ cuyas entradas $\phi_{i,j}(\omega)$ son una realización independiente de una variable aleatoria gaussiana $N(0, \sqrt{m})$, es decir, de media 0 y desviación estándar $1/\sqrt{m}$ (ver [27]). También son satisfechas por las matrices aleatorias cuyas entradas son v. a. independientes de Bernoulli con valores $\{-1/\sqrt{m}, 1/\sqrt{m}\}$ (ver [1]).

En esta sección se prueba (ver Teorema 2.3.28) que si $\tilde{\alpha} \leq \frac{1}{2\sqrt{k}}$, $1 \leq \ell < N$ y

$$m \geq \max\{C \ln l(N-k), Ck\}$$

se tiene que, dada una señal k -dispersa \mathbf{x} , la probabilidad de que las matrices aleatorias que satisfacen (M1)-(M4) identifiquen elementos de $\Gamma^\sharp = \text{sop}(\mathbf{x})$ en las primeras l iteraciones con los algoritmos RWMP, RWOMP y RWGP (aquellos que usan la regla de selección $\tilde{I}_n(\tilde{\alpha})$ dada en (2.8)) es mayor o igual a

$$1 - Cl(N-k)e^{-c_1 \tilde{\alpha} m}. \quad (2.18)$$

Para algoritmos del tipo RWOMP sabemos que, si $\tilde{I}_n(\tilde{\alpha}) \neq \emptyset$, siempre se identifican nuevos índices de $\Gamma^\sharp = \text{sop}(\mathbf{x})$. Por tanto, con este algoritmo, el soporte Γ^\sharp se identifica en como mucho k iteraciones por lo que la señal \mathbf{x} puede recuperarse. Como $k(N-k) \leq N^2/4$, la probabilidad de que las matrices aleatorias que satisfacen (M1)-(M4) recuperen una señal k -dispersa \mathbf{x} en las primeras k iteraciones es mayor o igual a

$$1 - CN^2 e^{-c_2 \tilde{\alpha} m}. \quad (2.19)$$

Este resultado se encuentra en el Corolario 2.3.30.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Se estudiará también en esta sección el comportamiento de las matrices aleatorias con respecto a la regla de selección $J_n(\alpha)$ dada por (2.6). Los resultados, que se describen en el Teorema 2.3.32 y Corolario 2.3.33 son similares a los obtenidos en (2.18) y (2.19).

Finalmente, la Sección 2.4 está dedicada a mostrar los resultados de algunos experimentos tanto para señales vectoriales como para imágenes. A juzgar por estos resultados, el algoritmo RWGP parece ser el que da mejores aproximaciones cuando se aplica a imágenes.

2.2 PC y geometría en espacios de grandes dimensiones.

En [36] se relaciona la percepción compresiva con el problema geométrico de las anchuras de Gel'fand y Kolmogorov de la siguiente forma. Sea $K \subset \mathbb{R}^N$, K compacto, $I : K \mapsto \mathbb{R}^m$ el operador de información (codificador) y $\Delta : \mathbb{R}^m \mapsto \mathbb{R}^N$ un algoritmo de estimación (decodificador) de $\mathbf{x} \in K$. Se define el error minimax en ℓ^p , como

$$E_m(K, \ell^p(\mathbb{R}^N)) = \inf_{\Delta, I} \sup_{\mathbf{x} \in K} \|\mathbf{x} - \Delta(I(\mathbf{x}))\|_{\ell^p(\mathbb{R}^N)},$$

en donde Δ, I son operadores no necesariamente lineales. En el caso de percepción compresiva I es la aplicación lineal dada por la matriz Φ y Δ es el problema de optimización (P_1).

La anchura de Gel'fand de K de orden m con respecto a la norma $\ell^p(\mathbb{R}^N)$ se define como

$$d^m(K, \ell^p(\mathbb{R}^N)) := \inf_Y \sup_{\mathbf{x} \in K \cap Y} \|\mathbf{x}\|_{\ell^p(\mathbb{R}^N)}, \quad (2.20)$$

donde el ínfimo se toma sobre subespacios lineales Y de codimensión menor o igual a m , es decir, Y es el complemento ortogonal (con respecto al producto interno euclídeo) de un subespacio lineal de dimensión m . El siguiente es un resultado ya establecido y relaciona el error minimax con la anchura de Gel'fand (ver [22]).

Lema 2.2.1 *Sea $K \subset \mathbb{R}^N$, K compacto, tal que $K = -K$ y para el cual existe un $c > 1$ tal que $K + K \subset cK$. Entonces,*

$$d^m(K, \ell^p(\mathbb{R}^N)) \leq E_m(K, \ell^p(\mathbb{R}^N)) \leq cd^m(K, \ell^p(\mathbb{R}^N)). \quad (2.21)$$

2.2 PC y geometría en espacios de grandes dimensiones.

Demostración. Obsérvese que $Y^\perp = \mathcal{N}(\Psi)$ es un subespacio de codimensión menor o igual a m . Además, dado un subespacio lineal Y^\perp de codimensión m , se puede asociar su complemento ortogonal Y de dimensión m con la matriz Φ de $m \times N$ cuyas filas se forman con cualquier base de Y . A través de esta identificación se ve que

$$d^m(K, \ell^p(\mathbb{R}^N)) = \inf_{\Phi} \sup_{\eta \in K \cap \mathcal{N}} \|\eta\|_{\ell^p(\mathbb{R}^N)},$$

en donde el ínfimo se toma sobre todas las matrices $m \times N$.

Sea (Φ, Δ) cualquier par codificador-decodificador y $z = \Delta(0)$. Para cualquier $\eta \in \mathcal{N}$ también sucede que $-\eta \in \mathcal{N}$. Se tiene entonces que $\|\eta - z\|_{\ell^p(\mathbb{R}^N)} \geq \|\eta\|_{\ell^p(\mathbb{R}^N)}$, o bien $\|-\eta - z\|_{\ell^p(\mathbb{R}^N)} \geq \|\eta\|_{\ell^p(\mathbb{R}^N)}$, por la desigualdad triangular. Como $K = -K$, se concluye que

$$d^m(K, \ell^p(\mathbb{R}^N)) \leq \sup_{\eta \in \mathcal{N} \cap K} \|\eta - \Delta(\Phi(\eta))\|_{\ell^p(\mathbb{R}^N)}.$$

Tomando el ínfimo sobre todos los pares (Δ, Φ) se tiene la desigualdad izquierda de (2.21).

Para probar la desigualdad derecha se elige una matriz Φ tal que optime la definición de $d^m(K, \ell^p(\mathbb{R}^N))$. Definimos un decodificador Δ de la siguiente manera. Dado \mathbf{z} en el rango de Φ escribimos $\mathbf{F}(\mathbf{z})$ para el conjunto de \mathbf{x} tal que $\Phi(\mathbf{x}) = \mathbf{z}$. Ahora, si $\mathbf{F}(\mathbf{z}) \cap K \neq \emptyset$ tomamos cualquier $\mathbf{x}'(\mathbf{z}) \in \mathbf{F}(\mathbf{z}) \cap K$ y $\Delta(\mathbf{z}) = \mathbf{x}'(\mathbf{z})$. Si $\mathbf{F}(\mathbf{z}) \cap K = \emptyset$, se define $\Delta(\mathbf{z})$ como cualquier elemento de $\mathbf{F}(\mathbf{z})$. Con esto y como $\mathbf{x} \in \mathbf{F}(\mathbf{z}) \cap K$ implica que $\mathbf{x} - \mathbf{x}' \in \mathcal{N}$ y $\mathbf{x} - \mathbf{x}' \in cK$ se tiene

$$\begin{aligned} E_m(K, \ell^p(\mathbb{R}^N)) &\leq \sup_{\mathbf{x} \in K} \|\mathbf{x} - \Delta(\Phi(\mathbf{x}))\|_{\ell^p(\mathbb{R}^N)} \leq \sup_{\mathbf{x}, \mathbf{x}' \in \mathbf{F}(\mathbf{z}) \cap K} \|\mathbf{x} - \mathbf{x}'\|_{\ell^p(\mathbb{R}^N)} \\ &\leq \sup_{\eta \in c[K \cap \mathcal{N}]} \|\eta\|_{\ell^p(\mathbb{R}^N)} \leq C d^m(K, \ell^p(\mathbb{R}^N)), \end{aligned}$$

por la elección de Φ . Esto termina la demostración. ■

Para terminar de establecer la relación entre la geometría de los espacios de Banach y el problema de la percepción compresiva debemos definir previamente las anchuras de Kolmogorov. Para un $K \subset \mathbb{R}^N$ acotado, la anchura de Kolmogorov de K de orden m con respecto a la norma $\ell^p(\mathbb{R}^N)$ se define como

$$d_m(K, \ell^p(\mathbb{R}^N)) := \inf_Y \sup_{\mathbf{x} \in K} \inf_{\mathbf{z} \in Y} \|\mathbf{x} - \mathbf{z}\|_{\ell^p(\mathbb{R}^N)}, \quad (2.22)$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

en donde el ínfimo es sobre todos los subespacios lineales Y de dimensión m de \mathbb{R}^N , ver [36]. El siguiente resultado de dualidad (ver [90]) relaciona las anchuras de Gel'fand con las de Kolmogorov, cuando los K son las bolas unitarias de espacios $\ell^p(\mathbb{R}^N)$, establece que

$$d_m(B^p(\mathbb{R}^N), \ell^q(\mathbb{R}^N)) = d^m(B^{q'}(\mathbb{R}^N), \ell^{p'}(\mathbb{R}^N)),$$

y en particular

$$d_m(B^2(\mathbb{R}^N), \ell^\infty(\mathbb{R}^N)) = d^m(B^1(\mathbb{R}^N), \ell^2(\mathbb{R}^N)).$$

Un resultado de Kashin ([65]) y mejorado por Garnaev y Gluskin ([45]) sobre las anchuras de Kolmogorov es el siguiente:

$$d_m(B^2(\mathbb{R}^N), \ell^\infty(\mathbb{R}^N)) \leq C \left(\frac{\log(N/m)}{m} \right)^{1/2}.$$

Este resultado permite estimar el error en percepción compresiva (2.4) a través del Lema 2.2.1 y de la dualidad entre anchuras como

$$E_m(B^1(\mathbb{R}^N), \ell^2(\mathbb{R}^N)) \simeq \left(\frac{\log(N/m)}{m} \right)^{1/2},$$

que es el resultado en el contexto de percepción compresiva del Teorema 1 en [36] y del Teorema 1.1 en [19] cuando $\mathbf{x} \in \ell^1(\mathbb{R}^N)$ (ver (2.4)). Así que, se puede afirmar que el error de la percepción compresiva es casi óptimo: salvo un factor logarítmico de m . Sin embargo, este resultado no se deriva de la misma forma cuando $p \neq 1$, $0 < p < 1$ (ver [36]). Además, el resultado teórico de las anchuras no dice cuáles son aquellos operadores de compresión de información (codificadores) y algoritmos de estimación (decodificadores) óptimos. La teoría de la percepción compresiva resuelve este problema para las condiciones anteriores.

Respecto a D3) ya se dijo anteriormente que el problema de optimación en ℓ^0 da soluciones dispersas, sin embargo, es un problema **NP**. Al relajar esta condición con la norma ℓ^1 el problema se vuelve en uno convexo y para el cual existen algoritmos numéricos “rápidos”. Estas soluciones “rápidas” convergen en, efectivamente, un tiempo polinómico en el tamaño N del problema. Sin embargo, son todavía demasiado lentos para las aplicaciones en tiempo real. La manera de sortear este problema es el tema de la Subsección 2.3 y en donde se muestran los resultados propios. Continuando el comentario respecto a D3), se tiene que, para $m = O(k \log N)$, el error de la percepción compresiva (2.4) es comparable con el error de la aproximación no lineal (2.1).

2.2.1 Una condición suficiente

En esta subsección daremos cuenta, breve y someramente, del desarrollo de una condición suficiente para que la solución \mathbf{x}^\sharp del problema de optimación (2.3) reconstruya o se aproxime a un vector \mathbf{x} con las condiciones de dispersión o compresibilidad. Nos concentraremos en los resultados de Candès-Romberg-Tao [17], [19], [20] y [18].

Sean $\mathbf{x} \in \mathbb{R}^N$ un vector k -disperso con soporte en Γ (i.e., $|\Gamma| = k$) y $\Omega \subset \mathbb{Z}_N$ un conjunto aleatorio (de tamaño $m = \mathbb{E}\{|\Omega|\}$) de renglones de la matriz de Fourier discreta \mathcal{F} de tamaño $N \times N$. La matriz de percepción compresiva es $\Phi = \mathcal{F}_\Omega$ y las observaciones son $\mathbf{y} = \Phi \mathbf{x} \in \mathbb{R}^m$. Ya se ha dicho que el resultado principal de [17] es que: si $m = \mathbb{E}\{|\Omega|\} = O(k \log N)$ entonces, para **una** realización de $\Phi = \mathcal{F}_\Omega$ y **una** observación \mathbf{y} , la probabilidad de que $\mathbf{x}^\sharp = \mathbf{x}$ será al menos $1 - O(N^{-c})$ (c un parámetro de exactitud). La prueba se basa en que una condición necesaria y suficiente para la solución de (2.3) es la existencia de un polinomio trigonométrico P cuya transformada de Fourier tenga soporte en Ω , sea igual a $\text{sgn}(\mathbf{x})$ en el soporte Γ de \mathbf{x} , y tenga magnitud estrictamente menor a 1 en cualquier otra parte (Lemma 2.1 en [17]). Se propone el polinomio $P := F_\Omega^* F_{\Gamma \rightarrow \Omega} (F_{\Gamma \rightarrow \Omega}^* F_{\Gamma \rightarrow \Omega})^{-1} \varrho^* \text{sgn}(\mathbf{x})$, en donde $\varrho : \ell^2(\Gamma) \mapsto \ell^2(\mathbb{Z}_N)$ extiende el vector en Γ a un vector en \mathbb{Z}_N colocando ceros fuera de Γ y ϱ^* es la restricción a Γ . El resto de la prueba se basa en probar la invertibilidad de $F_{\Gamma \rightarrow \Omega}^* F_{\Gamma \rightarrow \Omega}$ y probar la condición de magnitud de P en Γ^c en probabilidad. Esto se consigue estimando en la norma de Froebenius los momentos de las matrices y aprovechando las cancelaciones de los átomos (exponenciales complejas $e^{2\pi i t \cdot \omega}$) de \mathcal{F} fuera de la diagonal de dichas potencias de matrices.

En [19], Candès y Tao, además de extender el problema al caso en que \mathbf{x} es compresible ($\mathbf{x} \in w\ell^p$, $0 < p \leq 1$), establecen dos condiciones suficientes sobre la matriz Φ para que el error de reconstrucción (2.4) se verifique con “gran probabilidad”. Una de esas condiciones (*Exact Reconstruction Principle*: ERP) es la existencia de un polinomio trigonométrico como el descrito en [17]. La otra condición (*Uniform Uncertainty Principle*: UUP) está relacionada con la acotación por arriba y por abajo de los valores propios de **una** matriz Φ de tamaño $m \times N$, es decir,

$$\frac{1}{2} \frac{m}{N} \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2 \leq \frac{3}{2} \frac{m}{N} \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2,$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

para **todo** vector k -disperso \mathbf{x} tal que $m = O(k \log N)$ con probabilidad mayor a $1 - O(N^{-\alpha})$, $\alpha = \alpha(k, m, N)$, y en donde \mathbb{R}^Ω es el subespacio de \mathbb{R}^N con coordenadas significativas en el conjunto $\Omega \subset \{1, 2, \dots, N\}$ de tamaño m .

En [20], Candès y Tao introducen una condición muy parecida a UUP.

Definición 2.2.2 Dado $k \in \mathbb{N}$, una matriz $\Phi \in \mathbb{C}^{m \times N}$, $m \geq k$, cumple la **propiedad de la isometría restringida RIP** con parámetro δ_k , $0 < \delta_k < 1$, si se cumple que

$$(1 - \delta_k) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2 \leq (1 + \delta_k) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2, \quad (2.23)$$

para todos los vectores $\mathbf{x} \in \mathbb{R}^N$ que sean k -dispersos.

El ínfimo de las constantes δ_k que satisfacen (2.23) se llama **constante de la isometría restringida**.

También en [20] se introducen las siguientes constantes. Dados $k, k' \in \mathbb{N}$, la **constante de ortogonalidad restringida**, $\theta_{k,k'}$, es el menor de los números que satisfacen

$$|\langle \Phi \mathbf{x}, \Phi \mathbf{x}' \rangle| \leq \theta_{k,k'} \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)} \|\mathbf{x}'\|_{\ell^2(\mathbb{R}^N)} \quad (2.24)$$

para todos los vectores $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^N$ con $\text{sop } \mathbf{x} \cap \text{sop } \mathbf{x}' = \emptyset$ y $|\text{sop } \mathbf{x}| \leq k$, $|\text{sop } \mathbf{x}'| \leq k'$, donde $\text{sop } \mathbf{x} = \{i \in \{1, \dots, N\} : x_i \neq 0\}$.

Observar que los números δ_k y $\theta_{k,k'}$ miden cómo de cerca los vectores columna ϕ_i , $i = 1, \dots, N$, de Φ se comportan como un sistema ortonormal en \mathbb{R}^N , pero sólo cuando nos restringimos a combinaciones lineales (dispersas) formadas como mucho por k vectores.

El Lema 1.2 de [20] prueba que para todo k, k' se cumple

$$\theta_{k,k'} \leq \delta_{k+k'} \leq \theta_{k,k'} + \max(\delta_k, \delta_{k'}). \quad (2.25)$$

El Teorema 1.4, también de [20], muestra que si $\Phi \in \mathbb{C}^{m \times N}$ es una matriz que satisface

$$\delta_k + \theta_{k,k'} + \theta_{k,2k} < 1, \quad (2.26)$$

el problema de optimización (P_1) dado en (2.3) encuentra de manera exacta cualquier vector k -disperso $\mathbf{x} \in \mathbb{R}^N$ a partir de las mediciones $\mathbf{y} = \Phi \mathbf{x}$.

Una cuestión importante es sobre la existencia de dichas matrices. Se ha probado que las matrices gaussianas, de Bernoulli y parcial aleatoria de Fourier, verifican RIP con alta probabilidad (debido a que satisfacen la propiedad de la concentración de la

2.2 PC y geometría en espacios de grandes dimensiones.

medida, ver (2.32)) siempre y cuando el número de mediciones $m \geq Ck \log(N/k)$ para las dos primeras (ver [19], [20]) y $m \geq Ck \log^5(N)$ para la matriz parcial aleatoria de Fourier (ver [19], [92]).

Finalmente, en [18], Candès-Romberg-Tao incluyen el caso en que las observaciones estén contaminadas por ruido, *i.e.*, $\mathbf{y} = \Phi \mathbf{x} + \mathbf{n}$, $\|\mathbf{n}\|_{\ell^2} \leq \epsilon$. Incluyen los casos en que \mathbf{x} es un vector k -disperso y cuando \mathbf{x} es un vector arbitrario no necesariamente compresible. Se busca ahora la solución \mathbf{x}^\sharp al problema de optimización

$$(P'_1) \quad \min_{\mathbf{z} \in \mathbb{R}^N} \|\mathbf{z}\|_{\ell^1} \quad \text{sujeto a} \quad \|\Phi \mathbf{z} - \mathbf{y}\|_{\ell^2} \leq \epsilon. \quad (2.27)$$

El siguiente es el Teorema 1 en [18] (caso k -disperso) y del cual se incluye su demostración por razones didácticas.

Teorema 2.2.3 [Candès-Romberg-Tao] *Sea Φ una matriz que cumple RIP con constantes tal que $\delta_{3k} + 3\delta_{4k} < 2$. Para \mathbf{x} con soporte en Γ tal que $|\Gamma| = k$ y cualquier perturbación \mathbf{n} con $\|\mathbf{n}\|_{\ell^2(\mathbb{R}^N)} \leq \epsilon$, la solución \mathbf{x}^\sharp al problema de optimización (P'_1) obedece*

$$\|\mathbf{x} - \mathbf{x}^\sharp\|_{\ell^2} \leq C_k \cdot \epsilon, \quad (2.28)$$

en donde C_k depende sólo de δ_{4k} .

Demostración. Se hace uso de dos hechos geométricos que son consecuencia del problema de optimización (P'_1) y que están esquematizados en la Figura 2.1. A saber:

- **Restricción cilíndrica.** Como $\mathbf{y} - \Phi \mathbf{x} = \mathbf{n}$ con $\|\mathbf{n}\|_{\ell^2(\mathbb{R}^N)} \leq \epsilon$ y \mathbf{x}^\sharp cumple la restricción cuadrática de (P'_1) , entonces la desigualdad triangular da

$$\|\Phi \mathbf{x}^\sharp - \Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)} \leq \|\Phi \mathbf{x}^\sharp - \mathbf{y}\|_{\ell^2(\mathbb{R}^m)} + \|\Phi \mathbf{x} - \mathbf{y}\|_{\ell^2(\mathbb{R}^m)} \leq 2\epsilon.$$

Es decir, $\Phi \mathbf{x}^\sharp$ está dentro de un cilindro de radio $< 2\epsilon$ alrededor del subespacio $\Phi \mathbf{x}$.

- **Restricción conoidal.** Como \mathbf{x} es, de hecho, realizable en (P_2) , se tiene que $\|\mathbf{x}^\sharp\|_{\ell^1(\mathbb{R}^N)} \leq \|\mathbf{x}\|_{\ell^1(\mathbb{R}^N)}$. Podemos descomponer $\mathbf{x}^\sharp = \mathbf{x} + \mathbf{h}$, obteniendo

$$\|\mathbf{x}\|_{\ell^1(\mathbb{R}^N)} \geq \|\mathbf{x} + \mathbf{h}\|_{\ell^1(\mathbb{R}^N)} \geq \|\mathbf{x}\|_{\ell^1(\mathbb{R}^N)} - \|\mathbf{h}_\Gamma\|_{\ell^1(\mathbb{R}^N)} + \|\mathbf{h}_{\Gamma^c}\|_{\ell^1(\mathbb{R}^N)},$$

en donde Γ^c es el complemento de Γ en $\{1, 2, \dots, N\}$. Así, \mathbf{h} obedece la restricción en el cono

$$\|\mathbf{h}_{\Gamma^c}\|_{\ell^1(\mathbb{R}^N)} \leq \|\mathbf{h}_\Gamma\|_{\ell^1(\mathbb{R}^N)},$$

es decir, los componentes de \mathbf{h} están concentrados en el conjunto Γ .

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

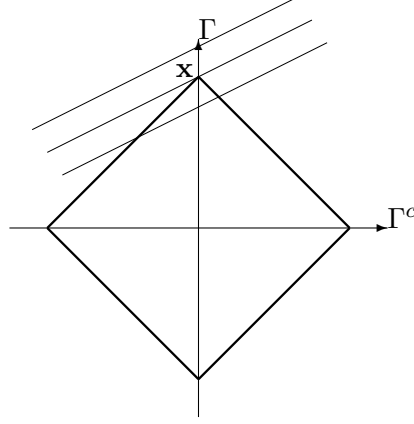


Figure 2.1: Esquematación del problema (P'_1) . Se muestran las restricciones en el “cono” y el “tubo”.

Escribimos $\Gamma_0 = \Gamma$. Dividimos Γ_0^c en subconjuntos de tamaño M (más adelante se precisará su valor) y reordenamos los elementos de Γ_0^c de manera que la magnitud de las coordenadas de \mathbf{h}_{Γ^c} decrezcan en los conjuntos Γ_j . Así, Γ_1 contiene los M índices de los coeficientes más grandes de \mathbf{h}_{Γ^c} , Γ_2 contiene M índices de los siguientes coeficientes más grandes de \mathbf{h}_{Γ^c} , y así sucesivamente (el último conjunto puede contener menos elementos).

Así, el k -ésimo elemento más grande de $\mathbf{h}_{\Gamma_0^c}$ obedece

$$|\mathbf{h}_{\Gamma_0^c}(\pi(k))| \leq \|\mathbf{h}_{\Gamma_0^c}\|_{\ell^1(\mathbb{R}^N)} / k$$

(donde $\pi(k)$ es la reordenación) y, por tanto

$$\|\mathbf{h}_{(\Gamma_0 \cup \Gamma_1)^c}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\mathbf{h}_{\Gamma_0^c}\|_{\ell^1(\mathbb{R}^N)}^2 \sum_{k=M+1}^N 1/k^2 \leq \|\mathbf{h}_{\Gamma_0^c}\|_{\ell^1(\mathbb{R}^N)}^2 / M.$$

Por la desigualdad conoidal y por las desigualdades entre normas, se tiene

$$\|\mathbf{h}_{(\Gamma_0 \cup \Gamma_1)^c}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\mathbf{h}_{\Gamma_0}\|_{\ell^1(\mathbb{R}^N)}^2 / M \leq \|\mathbf{h}_{\Gamma_0}\|_{\ell^2(\mathbb{R}^N)}^2 |\Gamma_0| / M,$$

es decir, la norma ℓ^2 de \mathbf{h} está concentrada en $\Gamma_0 \cup \Gamma_1 = \Gamma_{01}$. Así,

$$\|\mathbf{h}\|_{\ell^2(\mathbb{R}^N)}^2 = \|\mathbf{h}_{\Gamma_{01}}\|_{\ell^2(\mathbb{R}^N)}^2 + \|\mathbf{h}_{\Gamma_{01}^c}\|_{\ell^2(\mathbb{R}^N)}^2 \leq (1 + |\Gamma_0| / M) \|\mathbf{h}_{\Gamma_{01}}\|_{\ell^2(\mathbb{R}^N)}^2. \quad (2.29)$$

Ahora,

$$\begin{aligned}
 \|\Phi \mathbf{h}\|_{\ell^2(\mathbb{R}^m)} &= \left\| \Phi \mathbf{h}_{\Gamma_{01}} + \sum_{j \geq 2} \Phi \mathbf{h}_{\Gamma_j} \right\|_{\ell^2(\mathbb{R}^m)} \\
 &\geq \|\Phi \mathbf{h}_{\Gamma_{01}}\|_{\ell^2(\mathbb{R}^m)} - \sum_{j \geq 2} \|\Phi \mathbf{h}_{\Gamma_j}\|_{\ell^2(\mathbb{R}^m)} \\
 &\geq \sqrt{1 - \delta_{M+|\Gamma_0|}} \|\mathbf{h}_{\Gamma_{01}}\|_{\ell^2(\mathbb{R}^N)} - \sqrt{1 + \delta_M} \sum_{j \geq 2} \|\mathbf{h}_{\Gamma_j}\|_{\ell^2(\mathbb{R}^N)}, \quad (2.30)
 \end{aligned}$$

usando la propiedad RIP (2.23).

Por otra parte, por construcción de los conjuntos Γ_j se tiene $|h_i| \leq \|\mathbf{h}_{\Gamma_j}\|_{\ell^1(\mathbb{R}^N)} / M$ para todo $i \in \Gamma_{j+1}$. Así,

$$\|\mathbf{h}_{\Gamma_{j+1}}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\mathbf{h}_{\Gamma_j}\|_{\ell^1(\mathbb{R}^N)}^2 / M,$$

de donde se sigue que

$$\begin{aligned}
 \sum_{j \geq 2} \|\mathbf{h}_{\Gamma_j}\|_{\ell^2(\mathbb{R}^N)} &\leq \sum_{j \geq 1} \|\mathbf{h}_{\Gamma_j}\|_{\ell^1(\mathbb{R}^N)} / \sqrt{M} = \|\mathbf{h}_{\Gamma_0}\|_{\ell^1(\mathbb{R}^N)} / \sqrt{M} \\
 &\leq \|\mathbf{h}_{\Gamma_0}\|_{\ell^1(\mathbb{R}^N)} / \sqrt{M} \leq \sqrt{|\Gamma_0| / M} \|\mathbf{h}_{\Gamma_0}\|_{\ell^2(\mathbb{R}^N)},
 \end{aligned}$$

usando la restricción conoidal en la segunda desigualdad. Aplicando esto a (2.30) se tiene

$$\|\Phi \mathbf{h}\|_{\ell^2(\mathbb{R}^m)} \geq C_{|\Gamma_0|, M} \|\mathbf{h}_{\Gamma_{01}}\|_{\ell^2(\mathbb{R}^N)}, \quad C_{|\Gamma_0|, M} := \sqrt{1 - \delta_{|\Gamma_0|+M}} - \sqrt{|\Gamma_0| / M} \sqrt{1 + \delta_M}.$$

Haciendo $\rho = |\Gamma_0| / M$, se sigue de (2.29) y $\|\Phi \mathbf{h}\|_{\ell^2(\mathbb{R}^m)} \leq 2\epsilon$ (restricción cilíndrica) que

$$\|\mathbf{h}\|_{\ell^2(\mathbb{R}^N)} \leq \sqrt{1 + \rho} \|\mathbf{h}_{\Gamma_{01}}\|_{\ell^2(\mathbb{R}^N)} \leq \frac{\sqrt{1 + \rho}}{C_{|\Gamma_0|, M}} \|\Phi \mathbf{h}\|_{\ell^2(\mathbb{R}^m)} \leq 2 \frac{\sqrt{1 + \rho}}{C_{|\Gamma_0|, M}} \epsilon,$$

siempre y cuando $C_{|\Gamma_0|, M}$ sea positivo. Si $M = 3|\Gamma_0|$, entonces $C_{|\Gamma_0|, M} > 0$ si $\delta_{3|\Gamma_0|} + 3\delta_{4|\Gamma_0|} < 2$. Así termina la demostración. ■

2.2.2 Concentración de la medida y RIP

Las demostraciones de que las matrices gaussianas y de Bernoulli cumplen RIP (o UUP) se basan en resultados asintóticos sobre la probabilidad de las acotaciones de los auto

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

valores máximos y mínimos de dichas matrices. Las proyecciones aleatorias han sido usadas como herramientas fundamentales en la teoría asintótica de espacios normados de dimensión finita y en la teoría de la aproximación, como ya se ha mencionado en la Sección 2.2. Una más de las relaciones entre la percepción compresiva y los espacios de altas dimensiones se da con el lema de Johnson-Lindenstrauss o, más precisamente, con su demostración. El resultado original de dicho lema se lee así:

Lema 2.2.4 [Johnson-Lindenstrauss] *Sea $\epsilon \in (0, 1)$. Para todo conjunto P de $|P|$ puntos en \mathbb{R}^N y un número positivo $m > m_0 = O(\ln(|P|)/\epsilon^2)$, existe una aplicación Lipschitz $f : \mathbb{R}^N \mapsto \mathbb{R}^m$ tal que*

$$(1 - \epsilon) \|u - v\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|f(u) - f(v)\|_{\ell^2(\mathbb{R}^m)}^2 \leq (1 + \epsilon) \|u - v\|_{\ell^2(\mathbb{R}^N)}^2, \quad (2.31)$$

para todo $u, v \in P$.

Nótese el parecido con la propiedad RIP (2.23). Este es un resultado de existencia y no sabemos cómo construir f . Si permitimos que (2.31) se cumpla en probabilidad se ha probado que puede tomarse $f = \Phi$ como una matriz aleatoria Gaussiana o de Bernoulli de tamaño $m \times N$. La demostración se basa en probar que

$$\mathbb{E} \left\{ \|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2 \right\} = \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2,$$

y, posteriormente, que $\|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2$ está fuertemente concentrado (**concentración de la medida**) alrededor de su valor esperado, es decir,

$$\mathbb{P} \left\{ \left| \|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2 - \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2 \right| \geq \epsilon \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2 \right\} \leq 2e^{-mc_0(\epsilon)}, \quad 0 < \epsilon < 1, \quad (2.32)$$

en donde $c_0(\epsilon)$ depende sólo de ϵ y es tal que para todo $\epsilon \in (0, 1)$, $c_0(\epsilon) > 0$. En [27], Dasgupta-Gupta lo prueban para las matrices gaussianas, y en [1], Achlioptas para las de Bernoulli.

A partir de esto, Baraniuk-Davenport-DeVore-Wakin prueban en [6] que la propiedad RIP se verifica para las matrices gaussianas y de Bernoulli. Primero construyen una red de puntos que cubre la intersección de la bola unitaria $\ell^1(\mathbb{R}^N)$ y un subespacio k -dimensional (por coordenadas), se aplica (2.32) a todos estos puntos acotando la unión. Posteriormente se aplica a todos los vectores k -dimensionales en \mathbb{R}^N posibles, es decir, se acota con las combinaciones de N elementos tomados de k en k . A continuación hacemos explícitos estos dos resultados.

2.2 PC y geometría en espacios de grandes dimensiones.

Lema 2.2.5 *Sea Φ una matriz aleatoria de tamaño $m \times N$ que cumpla con la desigualdad de concentración de la medida (2.32). Entonces, para un conjunto Γ tal que $|\Gamma| = k < N$ y cualquier $0 < \delta < 1$, se tiene*

$$(1 - \delta) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)} \leq \|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)} \leq (1 + \delta) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}, \quad \text{para todo } \mathbf{x} \in \mathbb{R}^\Gamma, \quad (2.33)$$

con probabilidad

$$\geq 1 - 2(12/\delta)^k e^{-c_0(\delta/2)^m}. \quad (2.34)$$

Demostración. Es suficiente probar (2.33) para el caso $\|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)} = 1$ ya que Φ es lineal. Del capítulo 15 de [74], dado $\delta > 0$, se puede encontrar un conjunto de puntos $Q_\Gamma \subsetneq \mathbb{R}^\Gamma$ tal que para todo $\mathbf{q} \in Q_\Gamma$, $\|\mathbf{q}\|_{\ell^2(\mathbb{R}^N)} \leq 1$, y para todo $\mathbf{x} \in \mathbb{R}^\Gamma$ con $\|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)} \leq 1$, se tiene

$$\min_{\mathbf{q} \in Q_\Gamma} \|\mathbf{x} - \mathbf{q}\|_{\ell^2(\mathbb{R}^N)} \leq \delta/4$$

y, además, $|Q_\Gamma| \leq (12/\delta)^k$. Haciendo $\epsilon = \delta/2$ y aplicando (2.32) a cada uno de los $|Q_\Gamma|$ puntos $\mathbf{q} \in Q_\Gamma$, se tiene (con probabilidad $\geq 1 - 2(12/\delta)^k e^{-c_0(\delta/2)^m}$)

$$(1 - \delta/2) \|\mathbf{q}\|_{\ell^2(\mathbb{R}^N)}^2 \leq \|\Phi \mathbf{q}\|_{\ell^2(\mathbb{R}^m)}^2 \leq (1 + \delta/2) \|\mathbf{q}\|_{\ell^2(\mathbb{R}^N)}^2, \quad \text{para todo } \mathbf{q} \in Q_\Gamma,$$

lo que implica

$$(1 - \delta/2) \|\mathbf{q}\|_{\ell^2(\mathbb{R}^N)} \leq \|\Phi \mathbf{q}\|_{\ell^2(\mathbb{R}^m)} \leq (1 + \delta/2) \|\mathbf{q}\|_{\ell^2(\mathbb{R}^N)}, \quad \text{para todo } \mathbf{q} \in Q_\Gamma.$$

Defínase A como el menor número tal que

$$\|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)} \leq (1 + A) \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}, \quad \text{para todo } \mathbf{x} \in \mathbb{R}^\Gamma,$$

queremos mostrar que $A \leq \delta$. Dado \mathbf{x} con soporte en Γ y con $\|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)} \leq 1$ podemos elegir $\mathbf{q} \in Q_\Gamma$ tal que $\|\mathbf{x} - \mathbf{q}\| \leq \delta/4$. Entonces, se tiene

$$\|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)} \leq \|\Phi \mathbf{q}\|_{\ell^2(\mathbb{R}^m)} + \|\Phi(\mathbf{x} - \mathbf{q})\|_{\ell^2(\mathbb{R}^m)} \leq 1 + \delta/2 + (1 + A)\delta/4.$$

Por la definición de A se tiene $A \leq \delta/2 + (1 + A)\delta/4$. Es decir, $A \leq \frac{3\delta/4}{1 - \delta/4} \leq \delta$. Esto prueba la desigualdad por arriba. La estimación por abajo se sigue de esto ya que

$$\|\Phi \mathbf{x}\|_{\ell^2(\mathbb{R}^m)} \geq \|\Phi \mathbf{q}\|_{\ell^2(\mathbb{R}^m)} - \|\Phi(\mathbf{x} - \mathbf{q})\|_{\ell^2(\mathbb{R}^m)} \geq 1 - \delta/2 - (1 + \delta)\delta/4 \geq 1 - \delta.$$

■

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Teorema 2.2.6 *Sea Φ una matriz aleatoria $m \times N$ que cumple la concentración de la medida (2.32). Entonces, existen constantes $c_1, c_2 > 0$, que dependen sólo de δ , tales que se verifica RIP (2.23) para $k \leq c_1 m / \log(N/k)$ con probabilidad mayor a $1 - e^{-c_2 m}$.*

Demostración. Sabemos que para cada uno de los subespacios \mathbb{R}^F de dimensión k , la matriz Φ fallará en satisfacer (2.33) con probabilidad menor a $2(12/\delta)^k e^{-c_0(\delta/2)m}$. Existen $\binom{N}{k} \leq (eN/k)^k$ de tales subespacios. Por consiguiente, la propiedad RIP (2.23) no se verificará con probabilidad

$$\leq 2(eN/k)^k (12/\delta)^k e^{-c_0(\delta/2)m} = e^{-c_0(\delta/2)m + k[\log(eN/k) + \log(12/\delta)] + \log(2)}.$$

Para un c_1 suficientemente pequeño y por la hipótesis $k \leq c_1 m / \log(N/k)$, se puede encontrar un $c_2 > 0$ tal que $c_2 m < c_0(\delta/2)m - c_1 m [1 + (1 + \log(12/\delta)) / \log(N/k)]$, con lo que se prueba el teorema. ■

Las demostraciones de que las matrices parciales de Fourier cumplen RIP son más elaboradas.

2.3 Algoritmos ambiciosos en la percepción compresiva

Algunas consecuencias de la propiedad de la isometría restringida (2.23) para matrices de percepción compresiva han permitido el desarrollo de algoritmos que aceleran el proceso de encontrar una solución exacta o aproximada de \mathbf{x} conocidas las observaciones $\mathbf{y} = \Phi \mathbf{x}$. Aunque dichas consecuencias están implícitas en las pruebas de [20] (ver Lemma 1.2 y Lemma 2.1), algunas de estas fueron explotadas por primera vez por Needell-Vershynin en [84] para desarrollar un algoritmo de búsqueda basado en el algoritmo OMP (*Orthogonal Matching Pursuit*). Desde entonces se han propuesto algoritmos basados en dichas consecuencias.

Existen otros algoritmos de reconstrucción basados en matrices aleatorias cuyas garantías de probabilidad de reconstrucción dependen de un incremento en el número de mediciones m , incluso en un factor cuadrático de la dispersión del vector y polilogarítmico en la dimensión ambiente N , es decir, $m = O(k^2(\log N)^n)$. Recordemos que RIP se cumple para matrices de Fourier parciales, gaussianas y de Bernoulli, con $m = O(k \log(N))$.

2.3 Algoritmos ambiciosos en la percepción compresiva

Supongamos que la matriz Φ verifica RIP y que conocemos el soporte Γ^\sharp del vector k -disperso $\mathbf{x} \in \mathbb{R}^N$ y que las observaciones son $\mathbf{y} = \Phi \mathbf{x} := \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp}$. Para reconstruir \mathbf{x} sería suficiente aplicar la pseudo-inversa de Φ restringida al conjunto Γ^\sharp (que existe por (2.23)), ya que

$$(\Phi_{\Gamma^\sharp})^\dagger \mathbf{y} = (\Phi_{\Gamma^\sharp}^* \Phi_{\Gamma^\sharp})^{-1} \Phi_{\Gamma^\sharp}^* \mathbf{y} = (\Phi_{\Gamma^\sharp}^* \Phi_{\Gamma^\sharp})^{-1} \Phi_{\Gamma^\sharp}^* \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} = \mathbf{x}. \quad (2.35)$$

Obsérvese la acción prominente de $\Phi_{\Gamma^\sharp}^* : \mathbb{R}^m \mapsto \mathbb{R}^{\Gamma^\sharp}$. Además, se tiene $(\Phi_{\Gamma^\sharp}^* \Phi_{\Gamma^\sharp})^{-1} : \mathbb{R}^{\Gamma^\sharp} \mapsto \mathbb{R}^{\Gamma^\sharp}$. Supongamos que el residuo en la primera iteración es $\mathbf{r}^0 = \mathbf{y}$ y de alguna forma calculamos diferentes residuos $\mathbf{r}^1, \mathbf{r}^2, \dots, \mathbf{r}^n$. Entonces, los mayores elementos de $\mathbf{g}^n = (g_1^n, \dots, g_N^n)^t = \Phi^* \mathbf{r}^{n-1}$ ($g_i^n = \langle \phi_i, \mathbf{r}^{n-1} \rangle$, donde ϕ_i son las columnas de Φ) nos dan una idea del soporte de \mathbf{x} ya que se verifica RIP: el cuadrado de la energía de los k elementos de \mathbf{x} y las observaciones $\mathbf{y} = \Phi \mathbf{x}$ no difieren más que δ_k .

A nuestro juicio, los parámetros más importantes a analizar son los de convergencia y probabilidad de reconstrucción exacta (o error de aproximación). En este trabajo describiremos el funcionamiento de algunos algoritmos ambiciosos, las consecuencias de RIP y su uso en la identificación del soporte, la velocidad de convergencia y la probabilidad de reconstrucción exacta.

2.3.1 Los algoritmos

Los **algoritmos ambiciosos**¹ seleccionan un (conjunto de) elemento(s) en cada iteración identificando el(los) mayor(es) productos internos $|\langle \phi_i, \mathbf{r}^{n-1} \rangle|$. La principal diferencia entre estos algoritmos es la dirección de búsqueda y la forma de actualizar la aproximación y/o estimación. A continuación describimos algunos de ellos.

WEAK MATCHING PURSUIT (WMP) BÚSQUEDA DE CORRESPONDENCIA DÉBIL

¹Al *greedy algorithm* se le conoce en castellano generalmente como algoritmo avaricioso. La traducción que hace el *Concise Oxford Dictionary* de la palabra *greed* es codicia o avaricia. Según el diccionario de la Real Academia Española se tiene que: Avaricia: 1. f. Afán **desordenado** de poseer y adquirir riquezas para atesorarlas. También se tiene que, según la RAE: Codicia: 1. f. Afán excesivo de riquezas. 2. f. Deseo vehemente de algunas cosas buenas. Finalmente, de nuevo según la RAE: Ambicioso, sa.: 1. adj. Que tiene ambición. U. t. c. s. 2. adj. Que tiene ansia o deseo vehemente de algo. U. t. c. s.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

El algoritmo MP fue el primero en usarse en el procesamiento de señales (ver [79]). El parámetro de “debilidad” $\alpha \in (0, 1]$ es usado para seleccionar uno solo o varios elementos y se introdujo por primera vez, hasta donde sabemos, en [64] y ha sido extensamente usado en la aproximación no lineal (ver [34], [96], [97], [98] y las referencias ahí contenidas).

En el inicio establecemos: en la iteración $n = 0$ el residuo es $\mathbf{r}^0 = \mathbf{y}$, la aproximación a la observación es $\mathbf{y}^0 = 0$ y la estimación a la señal es $\mathbf{x}^0 = 0$. Recuerde que ϕ_i denota la i -ésima columna de Φ . El bucle hasta que se llega a algún criterio es:

- $\mathbf{g}^n = \Phi^* \mathbf{r}^{n-1}$, el representante de la señal.
- $\mathcal{J}_n := \mathcal{J}_n(\alpha) := \left\{ i : |g_i^n| \geq \alpha \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)} \right\}$, la regla de selección débil por etapas con $0 < \alpha \leq 1$.
- $\mathbf{y}^n = \mathbf{y}^{n-1} + \sum_{i \in \mathcal{J}_n} g_i^n \phi_i = \mathbf{y}^{n-1} + \sum_{i \in \mathcal{J}_n} \langle \phi_i, \mathbf{r}^{n-1} \rangle \phi_i$, la aproximación a la observación.
- $x_i^n = x_i^{n-1} + g_i^n = x_i^{n-1} + \langle \phi_i, \mathbf{r}^{n-1} \rangle$, $i \in \mathcal{J}_n$ y $x_j^n = x_j^{n-1}$ si $j \notin \mathcal{J}_n$, la estimación de la señal.
- $\mathbf{r}^n = \mathbf{r}^{n-1} - \sum_{i \in \mathcal{J}_n} g_i^n \phi_i = \mathbf{r}^{n-1} - \sum_{i \in \mathcal{J}_n} \langle \phi_i, \mathbf{r}^{n-1} \rangle \phi_i$, el residuo.

Por recursión y la definición de \mathbf{y}^n y \mathbf{r}^n se tiene que $\mathbf{r}^n = \mathbf{r}^{n-1} - (\mathbf{y}^n - \mathbf{y}^{n-1}) = \dots = \mathbf{y} - \mathbf{y}^n$.

WEAK ORTHOGONAL MATCHING PURSUIT (WOMP)

BÚSQUEDA DE CORRESPONDENCIA ORTOGONAL DÉBIL

El algoritmo WOMP es similar a WMP. Una vez que \mathcal{J}_n ha sido seleccionado en WMP, la aproximación $\mathbf{y}^n = \mathbf{y}^{n-1} + \sum_{i \in \mathcal{J}_n} g_i^n \phi_i$ puede no ser la mejor aproximación a \mathbf{y} en la norma $\ell^2(\mathbb{R}^m)$ por elementos del subespacio $V_n := \{\Phi_{\Gamma^n} \mathbf{x} : \mathbf{x} \in \mathbb{R}^N\}$, donde $\Gamma^n = \cup_{i=1}^n \mathcal{J}_i$. En WOMP se hace $\mathbf{y}^n = P_{\Gamma^n} \mathbf{y}$, donde P_{Γ^n} denota la proyección ortogonal sobre el subespacio V_n . Se puede obtener una expresión para \mathbf{y}^n en términos de Φ_{Γ^n} . Como $\mathbf{y}^n \in V_n$ podemos escribir $\mathbf{y}^n = \Phi_{\Gamma^n} \mathbf{x}^n$ para algún $\mathbf{x}^n \in \mathbb{R}^N$, y como \mathbf{y}^n realiza la mínima distancia (en $\ell^2(\mathbb{R}^m)$) de \mathbf{y} a V_n , el elemento $\mathbf{y} - \mathbf{y}^n$ ha de ser perpendicular a V_n . Entonces,

$$\mathbf{y} - \mathbf{y}^n = \mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}^n \perp \Phi_{\Gamma^n} \mathbf{x}, \forall \mathbf{x} \in \mathbb{R}^N.$$

2.3 Algoritmos ambiciosos en la percepción compresiva

Así pues, para todo $\mathbf{x} \in \mathbb{R}^N$,

$$0 = \langle \mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}^n, \Phi_{\Gamma^n} \mathbf{x} \rangle = \langle \Phi_{\Gamma^n}^* \mathbf{y} - \Phi_{\Gamma^n}^* \Phi_{\Gamma^n} \mathbf{x}^n, \mathbf{x} \rangle.$$

Por tanto, $\Phi_{\Gamma^n}^* \mathbf{y} - \Phi_{\Gamma^n}^* \Phi_{\Gamma^n} \mathbf{x}^n = 0$, de donde se deduce

$$\mathbf{x}^n = (\Phi_{\Gamma^n}^* \Phi_{\Gamma^n})^{-1} \Phi_{\Gamma^n}^* \mathbf{y} = \Phi_{\Gamma^n}^\dagger \mathbf{y} \quad \text{e} \quad \mathbf{y}^n = \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y} \quad (2.36)$$

donde $\Phi_{\Gamma^n}^\dagger = (\Phi_{\Gamma^n}^* \Phi_{\Gamma^n})^{-1} \Phi_{\Gamma^n}^*$ es la pseudo-inversa de Φ_{Γ^n} . Observar que (2.36) da no sólo la aproximación a la estimación sino también la aproximación a la señal.

En este algoritmo el residuo $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$ es ortogonal a las columnas de Φ_{Γ^n} ya que

$$\begin{aligned} \Phi_{\Gamma^n}^* \mathbf{r}^n &= \Phi_{\Gamma^n}^* (\mathbf{y} - \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y}) \\ &= \Phi_{\Gamma^n}^* \mathbf{y} - \Phi_{\Gamma^n}^* \Phi_{\Gamma^n} (\Phi_{\Gamma^n}^* \Phi_{\Gamma^n})^{-1} \Phi_{\Gamma^n}^* \mathbf{y} = 0. \end{aligned} \quad (2.37)$$

Así, es seguro que en cada iteración la regla de selección produce nuevos elementos. En el inicio establecemos: en la iteración $n = 0$ el residuo es $\mathbf{r}^0 = \mathbf{y}$, la aproximación a la observación es $\mathbf{y}^0 = 0$ y la estimación a la señal es $\mathbf{x}^0 = 0$. Recuerde que ϕ_i denota la i -ésima columna de Φ . El bucle hasta que se llega a algún criterio es:

- $\mathbf{g}^n = \Phi^* \mathbf{r}^{n-1}$, el representante de la señal.
- $\mathcal{J}_n := \mathcal{J}_n(\alpha) := \left\{ i : |g_i^n| \geq \alpha \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)} \right\}$, la regla de selección débil con $0 < \alpha \leq 1$.
- $\Gamma^n = \Gamma^{n-1} \cup \mathcal{J}_n$, actualización de los elementos seleccionados.
- $\mathbf{x}^n := \mathbf{x}^n|_{\Gamma^n} = \Phi_{\Gamma^n}^\dagger \mathbf{y}$, estimación de la señal.
- $\mathbf{y}^n = \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y}$, aproximación a la observación.
- $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$, el residuo.

Observar que \mathbf{y}^n resuelve el problema $\min_{\mathbf{z} \in V_n} \|\mathbf{y} - \mathbf{z}\|_{\ell^2(\mathbb{R}^m)}$ y por tanto \mathbf{x}^n es el elemento de \mathbb{R}^N que resuelve el problema

$$\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}, \quad \text{sop } \mathbf{x} \subset \Gamma^n. \quad (2.38)$$

Conviene observar que la notación $\mathbf{x}^n := \mathbf{x}_{\Gamma^n}^n$ que hemos escrito para la estimación de la señal significa que $x_i^n = 0$ si $i \notin \Gamma^n$. Por tanto

$$\Phi \mathbf{x}^n = \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^n = \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y} = \mathbf{y}^n.$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

WEAK GRADIENT PURSUIT (WGP) BÚSQUEDA DE GRADIENTE DÉBIL

Este algoritmo ambicioso es, tal vez, el menos conocido de los que se describen en este trabajo (ver [85]). Blumensath y Davies usan este algoritmo en el contexto de las matrices aleatorias en [11] y [12]. En la iteración $n = 0$ se establecen el residuo $\mathbf{r}^0 = \mathbf{y}$, el soporte $\Gamma^0 = \emptyset$ y, por tanto, la aproximación $\mathbf{y}^0 = 0$ y la estimación $\mathbf{x}^0 = 0$. El bucle básico es:

- $\mathbf{g}^n = \Phi^* \mathbf{r}^{n-1}$, el “representante” (“proxy”) de la señal.
- $\mathcal{J}_n(\alpha) := \left\{ i : |g_i^n| \geq \alpha \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)} \right\}$, la regla de selección.
- $\Gamma^n = \Gamma^{n-1} \cup \mathcal{J}_n$, actualización de los elementos seleccionados.
- $\mathbf{d}_{\Gamma^n}^n = \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}$, actualización de la dirección.
- $a^n = \frac{\langle \mathbf{r}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n \rangle}{\|\Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n\|_{\ell^2(\mathbb{R}^m)}^2}$, el paso optimado.
- $\mathbf{x}^n := \mathbf{x}_{\Gamma^n}^n = \mathbf{x}^{n-1} + a^n \mathbf{d}_{\Gamma^n}^n$, la estimación de la señal.
- $\mathbf{y}^n = \mathbf{y}^{n-1} + a^n \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n$, la aproximación a la observación.
- $\mathbf{r}^n = \mathbf{r}^{n-1} - a^n \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n$, el residuo.

Se tiene nuevamente que $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$. En efecto, si $n = 0$ se tiene $\mathbf{r}^0 = \mathbf{y} = \mathbf{y} - \mathbf{y}^0$. Suponiendo que $\mathbf{r}^{n-1} = \mathbf{y} - \mathbf{y}^{n-1}$, $n \geq 1$, se deduce

$$\mathbf{r}^n = \mathbf{r}^{n-1} - (\mathbf{y}^n - \mathbf{y}^{n-1}) = (\mathbf{y} - \mathbf{y}^{n-1}) - (\mathbf{y}^n - \mathbf{y}^{n-1}) = \mathbf{y} - \mathbf{y}^n. \quad (2.39)$$

Recordar que $\mathbf{x}^n := \mathbf{x}_{\Gamma^n}^n$ significa que las coordenadas de \mathbf{x}^n que no corresponden a índices de Γ^n son nulas. Con esta notación $\Phi \mathbf{x}^n = \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^n$ y $\mathbf{x}^{n-1} = \mathbf{x}_{\Gamma^{n-1}}^{n-1} = \mathbf{x}_{\Gamma^n}^{n-1}$.

Al igual que en WOMP, también se cumple $\mathbf{y}^n = \Phi \mathbf{x}^n$. En efecto, si $n = 1$, $\mathbf{y}^1 = a^1 \Phi_{\Gamma^1} \mathbf{d}_{\Gamma^1}^1 = \Phi_{\Gamma^1} \mathbf{x}_{\Gamma^1}^1 = \Phi \mathbf{x}^1$. Supongamos que $\mathbf{y}^{n-1} = \Phi \mathbf{x}^{n-1} = \Phi_{\Gamma^{n-1}} \mathbf{x}_{\Gamma^{n-1}}^{n-1} = \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^{n-1}$. Entonces

$$\mathbf{y}^n = \mathbf{y}^{n-1} + a^n \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n = \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^{n-1} + \Phi_{\Gamma^n} (\mathbf{x}_{\Gamma^n}^n - \mathbf{x}_{\Gamma^n}^{n-1}) = \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^n = \Phi \mathbf{x}^n. \quad (2.40)$$

2.3 Algoritmos ambiciosos en la percepción compresiva

El término “gradiente” en el nombre de este algoritmo se debe a que la actualización de la dirección, $\mathbf{d}_{\Gamma^n}^n = \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}$ es el gradiente de la función de coste

$$C_n(\mathbf{x}) = -\frac{1}{2} \|\mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2, \quad \text{sop } \mathbf{x} \in \Gamma^n \quad (2.41)$$

evaluado en \mathbf{x}^{n-1} . En efecto, un simple cálculo produce $\nabla_{\mathbf{x}} C_n(\mathbf{x}_{\Gamma^n}) = \Phi_{\Gamma^n}^* (\mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n})$ y cuando $\mathbf{x} = \mathbf{x}^{n-1}$

$$\nabla_{\mathbf{x}} C_n(\mathbf{x}^{n-1}) = \Phi_{\Gamma^n}^* (\mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^{n-1}) = \Phi_{\Gamma^n}^* (\mathbf{y} - \mathbf{y}^{n-1}) = \Phi_{\Gamma^n}^* \mathbf{r}^{n-1} = \mathbf{d}_{\Gamma^n}^n$$

debido a (2.40) y (2.39). Observar que el mínimo de $C_n(\mathbf{x})$ se alcanza en $\tilde{\mathbf{x}}^n$ tal que $\nabla_{\mathbf{x}} C_n(\tilde{\mathbf{x}}^n) = 0$, y por tanto $\Phi_{\Gamma^n}^* (\mathbf{y} - \Phi_{\Gamma^n} \tilde{\mathbf{x}}_{\Gamma^n}^n) = 0$. De aquí se deduce $\tilde{\mathbf{x}}_{\Gamma^n}^n = (\Phi_{\Gamma^n}^* \Phi_{\Gamma^n})^{-1} \Phi_{\Gamma^n}^* \mathbf{y} = \Phi_{\Gamma^n}^\dagger \mathbf{y}$ que es la aproximación de la señal en WOMP.

Una vez calculada la actualización de la dirección $\mathbf{d}_{\Gamma^n}^n = \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}$, se muestra en [51], p. 521, que para minimizar el cuadrado del error de la solución en las direcciones actualizadas se ha de tomar

$$a^n = \frac{\langle \mathbf{r}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n \rangle}{\|\Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n\|_{\ell^2(\mathbb{R}^m)}^2},$$

lo que explica la elección de a^n en el algoritmo.

Para información del lector mencionaremos que hay otras formas de actualizar las direcciones. Por ejemplo, Blumensath y Davies ([11], [12]) proponen que las nuevas direcciones sean de la forma

$$\mathbf{d}_{\Gamma^n}^n = \mathbf{g}_{\Gamma^n}^n + \omega^n \mathbf{d}_{\Gamma^n}^{n-1}$$

donde $\mathbf{g}_{\Gamma^n}^n = \Phi_{\Gamma^n}^* (\mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}^n)$ coincide con el gradiente de la función de coste evaluado en \mathbf{x}^n y ω^n es un escalar que mantiene la propiedad llamada de “conjugación” entre $\mathbf{d}_{\Gamma^n}^n$ y $\mathbf{d}_{\Gamma^n}^{n-1}$, es decir,

$$\langle \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^n \rangle = 0,$$

donde $\mathbf{d}_{\Gamma^n}^{n-1}$ indica $\mathbf{d}_{\Gamma^{n-1}}^{n-1}$ con ceros añadidos en los nuevos índices. Es fácil comprobar que debe elegirse

$$\omega^n = -\frac{\langle \Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^{n-1}, \Phi_{\Gamma^n} \mathbf{g}_{\Gamma^n}^n \rangle}{\|\Phi_{\Gamma^n} \mathbf{d}_{\Gamma^n}^{n-1}\|_{\ell^2}^2}.$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

2.3.2 Consecuencias de RIP

Las siguientes son algunas propiedades de las matrices de percepción compresiva que se usarán para analizar algunos algoritmos. Estos resultados se usan de manera implícita en algunas pruebas (ver Lemma 1.2 y Lemma 2.1) en [20]. En los algoritmos presentados en [84], [83], [11], [12] y [23], entre otros, se usan algunas de estas consecuencias. Una demostración completa se da en [23].

Lema 2.3.1 *Supongamos que $\Phi \in \mathbb{R}^{m \times N}$ satisface RIP (2.23) con δ_k , $|\Gamma| \leq k$, $\text{supp}(\mathbf{u}) = \Gamma$.*

$$\|\Phi_\Gamma\|_{\ell^2(\mathbb{R}^\Gamma) \rightarrow \ell^2(\mathbb{R}^m)} = \|\Phi_\Gamma^*\|_{\ell^2(\mathbb{R}^m) \rightarrow \ell^2(\mathbb{R}^\Gamma)} \leq (1 + \delta_k)^{1/2}, \quad (2.42)$$

$$(1 - \delta_k) \|\mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)} \leq \|\Phi_\Gamma^* \Phi_\Gamma \mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)} \leq (1 + \delta_k) \|\mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)}, \quad (2.43)$$

$$(1 + \delta_k)^{-1} \|\mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)} \leq \|(\Phi_\Gamma^* \Phi_\Gamma)^{-1} \mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)} \leq (1 - \delta_k)^{-1} \|\mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)}. \quad (2.44)$$

Para conjuntos disjuntos Γ', Γ tales que $|\Gamma' \cup \Gamma| \leq k$

$$\|\Phi_{\Gamma'}^* \Phi_\Gamma \mathbf{u}\|_{\ell^2(\mathbb{R}^{\Gamma'})} \leq \delta_k \|\mathbf{u}\|_{\ell^2(\mathbb{R}^\Gamma)}. \quad (2.45)$$

Desarrollamos a continuación otra consecuencia de RIP que usaremos para probar algunos resultados propios.

Lema 2.3.2 *Sea Φ que verifica RIP con parámetro δ_k y sea $|\Gamma| \leq k$. Para todo $\mathbf{r} \in \text{span } \Phi_\Gamma$*

$$\|\Phi_\Gamma^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)} \geq (1 - \delta_k)^{\frac{1}{2}} \|\mathbf{r}\|_{\ell^2(\mathbb{R}^m)}. \quad (2.46)$$

Demostración. Primero acotamos la pseudo-inversa

$$\left\| \Phi_\Gamma^\dagger \mathbf{r} \right\|_{\ell^2(\mathbb{R}^\Gamma)} = \|(\Phi_\Gamma^* \Phi_\Gamma)^{-1} \Phi_\Gamma^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)} \leq (1 - \delta_k)^{-1} \|\Phi_\Gamma^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)},$$

en donde hemos usado (2.44). A continuación usamos esto y la desigualdad de Schwarz para obtener

$$\begin{aligned} \|\Phi_\Gamma^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)}^2 &\geq (1 - \delta_k) \left\| \Phi_\Gamma^\dagger \mathbf{r} \right\|_{\ell^2(\mathbb{R}^\Gamma)} \|\Phi_\Gamma^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)} \\ &\geq (1 - \delta_k) \langle \Phi_\Gamma^\dagger \mathbf{r}, \Phi_\Gamma^* \mathbf{r} \rangle = (1 - \delta_k) (\Phi_\Gamma^\dagger \mathbf{r})^* \Phi_\Gamma^* \mathbf{r} \\ &= (1 - \delta_k) \mathbf{r}^* \Phi_\Gamma (\Phi_\Gamma^* \Phi_\Gamma)^{-1} \Phi_\Gamma^* \mathbf{r} \\ &= (1 - \delta_k) \mathbf{y}_\Gamma^* \Phi_\Gamma^* \Phi_\Gamma (\Phi_\Gamma^* \Phi_\Gamma)^{-1} \Phi_\Gamma^* \Phi_\Gamma \mathbf{y}_\Gamma \\ &= (1 - \delta_k) \mathbf{y}_\Gamma^* \Phi_\Gamma^* \Phi_\Gamma \mathbf{y}_\Gamma \\ &= (1 - \delta_k) \|\mathbf{r}\|_{\ell^2(\mathbb{R}^m)}^2, \end{aligned}$$

ya que \mathbf{r} está en el espacio generado por las columnas de Φ_Γ . Es decir, se puede escribir $\mathbf{r} = \Phi_\Gamma \mathbf{y}_\Gamma$ para algún \mathbf{y}_Γ y se concluye la prueba. \blacksquare

2.3.2.1 Identificación del soporte en WMP, WOMP y WGP

Damos ahora condiciones suficientes sobre la matriz de percepción compresiva Φ para que los algoritmos WMP, WOMP y WGP seleccionen elementos en el soporte Γ^\sharp de un vector k -disperso \mathbf{x} . El algoritmo WOMP seleccionará nuevas columnas en cada iteración debido a la ortogonalidad con los residuos previos. Por tanto, la convergencia para reconstrucción exacta de vectores k -dispersos está garantizada en, a lo mucho, k iteraciones cuando la condición se verifica. La regla de selección es

$$\mathcal{J}_n(\alpha) := \left\{ i : |g_i^n| \geq \alpha \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)} \right\}, \quad (2.47)$$

para algún $\alpha \in (0, 1]$.

Lema 2.3.3 *Supongamos que $\mathcal{J}_1(\alpha), \dots, \mathcal{J}_n(\alpha) \subset \Gamma^\sharp = \text{sop}(\mathbf{x})$. Entonces, el residuo \mathbf{r}^n en cada uno de los algoritmos WMP, WOMP y WGP descritos en la Sección 2.3.1 satisface $\mathbf{r}^n \in \text{span}(\Phi_{\Gamma^\sharp})$.*

Demostración. Tenemos que probar que existe $\tilde{\mathbf{x}}^n$ con $\text{sop } \tilde{\mathbf{x}}^n \subset \Gamma^\sharp$ tal que $\mathbf{r}^n = \Phi_{\Gamma^\sharp} \tilde{\mathbf{x}}^n$ para cada uno de los algoritmos.

Para WOMP, $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n = \mathbf{y} - \Phi_{\Gamma^n} \mathbf{x}^n = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} - \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^n$ donde $\Gamma^n = \cup_{k=1}^n \mathcal{J}_k(\alpha) \subset \Gamma^\sharp$. escribiendo $\mathbf{x}_{\Gamma^\sharp}^n$ para denotar el elemento que coincide con $\mathbf{x}_{\Gamma^n}^n$ y tiene ceros en los índices de $\Gamma^\sharp \setminus \Gamma^n$ se tiene $\mathbf{r}^n = \Phi_{\Gamma^\sharp} (\mathbf{x}_{\Gamma^\sharp} - \mathbf{x}_{\Gamma^\sharp}^n)$ con $\text{sop}(\mathbf{x}_{\Gamma^\sharp} - \mathbf{x}_{\Gamma^\sharp}^n) \subset \Gamma^\sharp$.

Para WGP la demostración es similar ya que, según (2.39) y (2.40), $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} - \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^n$ con $\Gamma^n = \cup_{k=1}^n \mathcal{J}_k(\alpha)$.

Para WMP procedemos por inducción. Es claro que $\mathbf{r}^0 = \mathbf{y} = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp}$ lo satisface. Para \mathbf{r}^1 tenemos $\mathbf{r}^1 = \mathbf{y} - \mathbf{y}^1 = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} - \sum_{i \in \mathcal{J}_1} \langle \phi_i, \mathbf{r}^0 \rangle \phi_i = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} - \Phi_{\mathcal{J}_1} \Phi_{\mathcal{J}_1}^* \mathbf{y}$. Con $\mathbf{z}^1 = \Phi_{\mathcal{J}_1}^* \mathbf{y}$ se tiene $\mathbf{r}^1 = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} - \Phi_{\mathcal{J}_1} \mathbf{z}^1$ con $\text{sop } \mathbf{z}^1 \subset \mathcal{J}_1 \subset \Gamma^\sharp$. Entonces, $\mathbf{r}^1 = \Phi_{\Gamma^\sharp} (\mathbf{x}_{\Gamma^\sharp} - \mathbf{z}_{\Gamma^\sharp}^1)$ con $\mathbf{z}_{\Gamma^\sharp}^1$ obtenido de \mathbf{z}^1 añadiendo ceros en las posiciones $\Gamma^\sharp \setminus \mathcal{J}_1$. En este caso $\tilde{\mathbf{x}}^1 = \mathbf{x}_{\Gamma^\sharp} - \mathbf{z}_{\Gamma^\sharp}^1$.

Supongamos que $\mathbf{r}^k = \Phi_{\Gamma^\sharp} \tilde{\mathbf{x}}^k$ con $\text{sop } \tilde{\mathbf{x}}^k \subset \Gamma^\sharp$ para todo $k = 1, 2, \dots, n-1$. Entonces, con $\mathbf{z}^n = \Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1}$,

$$\mathbf{r}^n = \mathbf{r}^{n-1} - \Phi_{\mathcal{J}_n} \Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1} = \Phi_{\Gamma^\sharp} \tilde{\mathbf{x}}^{n-1} - \Phi_{\mathcal{J}_n} \mathbf{z}^n = \Phi_{\Gamma^\sharp} (\tilde{\mathbf{x}}^{n-1} - \mathbf{z}_{\Gamma^\sharp}^n)$$

con $\text{sop}(\tilde{\mathbf{x}}^{n-1} - \mathbf{z}_{\Gamma^\sharp}^n) \subset \Gamma^\sharp$ ya que $\text{sop } \mathbf{z}^n \subset \mathcal{J}_n \subset \Gamma^\sharp$. ■

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Lema 2.3.4 *Sea Φ una matriz de tamaño $m \times N$ que satisface RIP (2.23) con parámetro δ_{k+1} . Supongamos que*

$$\frac{\sqrt{k}\delta_{k+1}}{1 - \delta_k} < \alpha \quad (2.48)$$

y que $\mathbf{r} \in \text{span}(\Phi_{\Gamma^\sharp})$ con $\Gamma^\sharp \subset \{1, \dots, N\}$ y $|\Gamma^\sharp| \leq k$. Definir

$$\mathcal{J}(\alpha) = \{i : |\langle \phi_i, \mathbf{r} \rangle| \geq \alpha \|\Phi^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^N)}\}$$

donde $\phi_i, i = 1, 2, \dots, N$, denotan las columnas de Φ . Se tiene que

$$\mathcal{J}(\alpha) \subset \Gamma^\sharp.$$

Demostración. Tomando complementarios, basta probar que

$$\|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})} < \alpha \|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}.$$

Puesto que $\mathbf{r} \in \text{span}(\Phi_{\Gamma^\sharp})$ podemos escribir $\mathbf{r} = \Phi_{\Gamma^\sharp} \mathbf{z}$ con $\text{sop } \mathbf{z} \subset \Gamma^\sharp$. Sea λ un índice de $\Gamma^{\sharp c}$ para el que

$$|\langle \phi_\lambda, \mathbf{r} \rangle| = \|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})}.$$

Observar que $|\langle \phi_\lambda, \mathbf{r} \rangle| = |\Phi_\lambda^* \mathbf{r}| = \|\Phi_\lambda^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^\lambda)} = \|\Phi_\lambda^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\lambda)}$ donde Φ_λ es la matriz de tamaño $m \times 1$ formada por la columna λ de Φ . Entonces,

$$\begin{aligned} \frac{\|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})}}{\|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}} &= \frac{\|\Phi_\lambda^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\lambda)}}{\|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}} \leq \frac{\sqrt{k} \|\Phi_\lambda^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\lambda)}}{\|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}} \\ &= \frac{\sqrt{k} \|\Phi_\lambda^* \Phi_{\Gamma^\sharp} \mathbf{z}\|_{\ell^2(\mathbb{R}^\lambda)}}{\|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}} \leq \frac{\sqrt{k}\delta_{k+1} \|\mathbf{z}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}}{(1 - \delta_k)^{1/2} \|\Phi_{\Gamma^\sharp} \mathbf{z}\|_{\ell^2(\mathbb{R}^m)}} \\ &\leq \frac{\sqrt{k}\delta_{k+1} \|\mathbf{z}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}}{(1 - \delta_k) \|\mathbf{z}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}} = \frac{\sqrt{k}\delta_{k+1}}{1 - \delta_k} < \alpha \end{aligned}$$

como queríamos probar. La primera desigualdad se debe a que $\|w\|_{\ell^2(\mathbb{R}^d)} \leq \sqrt{d} \|w\|_{\ell^\infty(\mathbb{R}^d)}$. La segunda desigualdad es debida a (2.45) en el numerador y (2.46) en el denominador (Lemma 2.3.2) ya que $\mathbf{r} \in \text{span}(\Phi_{\Gamma^\sharp})$. La tercera desigualdad es debida a la parte izquierda de la condición RIP (2.23). \blacksquare

Teorema 2.3.5 *Sea Φ una matriz de tamaño $m \times N$ que satisface RIP (2.23) con δ_{k+1} . Si*

$$\frac{\sqrt{k}\delta_{k+1}}{1 - \delta_k} < \alpha \quad (2.49)$$

2.3 Algoritmos ambiciosos en la percepción compresiva

los algoritmos WMP, WOMP y WGP descritos en la Sección 2.3.1 cuya regla de selección $\mathcal{J}_n(\alpha)$ está dada por (2.47) satisfacen $\mathcal{J}_n(\alpha) \subset \text{sop}(\mathbf{x}) = \Gamma^\sharp$ para toda señal \mathbf{x} que sea k -dispersa (es decir, los algoritmos identifican elementos del soporte Γ^\sharp de la señal k -dispersa \mathbf{x}).

Demostración. Procedemos por inducción en n . Si $n = 1$, como $\mathbf{r}^0 = \mathbf{y} = \Phi \mathbf{x} = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp} \in \text{span}(\Phi_{\Gamma^\sharp})$ del Lema 2.3.4 deducimos $\mathcal{J}_1 = \mathcal{J}_1(\alpha) \subset \Gamma^\sharp$. Supongamos que $\mathcal{J}_1(\alpha), \dots, \mathcal{J}_{n-1}(\alpha)$ están contenidos en Γ^\sharp . Por el Lema 2.3.3, $\mathbf{r}^{n-1} \in \text{span}(\Phi_{\Gamma^\sharp})$. Aplicar ahora el Lema 2.3.4 a la regla de selección (2.47) para obtener $\mathcal{J}_n(\alpha) \subset \Gamma^\sharp$. ■

Sea Φ una matriz de tamaño $m \times N$ que satisface RIP (2.23) con parámetro δ_{k+1} y supongamos que

$$\frac{\sqrt{k}\delta_{k+1}}{1 - \delta_k} < \alpha \quad (0 < \alpha \leq 1).$$

Por el Teorema 2.3.5 si cualquiera de los algoritmos WMP, WOMP o WGP ha seleccionado k índices distintos después de varias iteraciones, entonces se ha encontrado el soporte de cualquier vector k -disperso \mathbf{x} .

Además, en el caso de WOMP, como la regla de selección siempre escoge nuevos índices debido a la ortogonalidad expresada en (2.37), el soporte de cualquier señal k -dispersa se encuentra en como mucho k -iteraciones. En este caso, se puede recuperar la señal original \mathbf{x} , que es el contenido del siguiente resultado.

Corolario 2.3.6 *Con las condiciones del Teorema 2.3.5, el algoritmo WOMP recupera cualquier señal k -dispersa \mathbf{x} en como mucho k -iteraciones.*

Demostración. Supongamos que hemos alcanzado $\Gamma^\sharp = \text{sop}(\mathbf{x}) = \Gamma^n$ en la iteración n . Sabemos que $n \leq k$ por la discusión previa al Corolario 2.3.6. Entonces

$$\mathbf{x}_{\Gamma^n}^n = \Phi_{\Gamma^n}^\dagger \mathbf{y} = \Phi_{\Gamma^\sharp}^\dagger \mathbf{y} = \mathbf{x}_{\Gamma^\sharp}$$

por la definición de \mathbf{x}^n en WOMP en la primera igualdad y por la fórmula de reconstrucción con la pseudo-inversa (2.35) en la última. ■

Para los algoritmos WMP y WGP no sabemos si se alcanza Γ^\sharp en alguna iteración. Sin embargo, el conocimiento de $\mathbf{r}^n = 0$ para algún n , nos permite recuperar la señal k -dispersa \mathbf{x} .

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Corolario 2.3.7 *Supongamos que se cumplen las condiciones del Teorema 2.3.5 y, además, que el residuo en WMP, WOMP o WGP es nulo en la iteración n ($\mathbf{r}^n = 0$). Entonces, cualquiera de estos algoritmos recupera una señal k -dispersa \mathbf{x} en n iteraciones.*

Demostración. Si $\mathbf{r}^n = 0$, en todos los algoritmos se tiene $0 = \mathbf{y} - \mathbf{y}^n$, y por tanto, $\mathbf{y} = \mathbf{y}^n$.

Para WOMP y WGP se tiene $\mathbf{y}^n = \Phi_{\Gamma^n} \mathbf{x}_{\Gamma^n}^n$ (ver (2.36) para WOMP y (2.40) para WGP). Para estos algoritmos sabemos que \mathbf{x}^n tiene soporte en Γ^n y por el Teorema 2.3.5 tenemos $\Gamma^n \subset \Gamma^\sharp$. Por tanto, $\mathbf{y}^n = \Phi_{\Gamma^\sharp} \mathbf{x}_{\Gamma^\sharp}^n$ ya que $x_i^n = 0$ si $i \in \Gamma^\sharp \setminus \Gamma^n$. Por la condición de RIP (2.23),

$$\begin{aligned} 0 &= \|\mathbf{y} - \mathbf{y}^n\|_{\ell^2(\mathbb{R}^m)}^2 = \|\Phi_{\Gamma^\sharp}(\mathbf{x}_{\Gamma^\sharp} - \mathbf{x}_{\Gamma^\sharp}^n)\|_{\ell^2(\mathbb{R}^m)}^2 \\ &\geq (1 - \delta_k) \|\mathbf{x}_{\Gamma^\sharp} - \mathbf{x}_{\Gamma^\sharp}^n\|_{\ell^2(\mathbb{R}^\sharp)}^2 = (1 - \delta_k) \|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)}^2. \end{aligned} \quad (2.50)$$

De aquí se deduce $\mathbf{x} = \mathbf{x}^n$.

Para WMP, como $\mathbf{r}^n = 0$ se tiene

$$\mathbf{r}^{n-1} = \sum_{i \in \mathcal{J}_n} g_i^n \phi_i = \Phi_{\mathcal{J}_n} \Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1}.$$

Entonces,

$$\mathbf{r}^{n-2} = \mathbf{r}^{n-1} + \sum_{i \in \mathcal{J}_{n-1}} g_i^{n-1} \phi_i = \Phi_{\mathcal{J}_n} \Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1} + \Phi_{\mathcal{J}_{n-1}} \Phi_{\mathcal{J}_{n-1}}^* \mathbf{r}^{n-2}.$$

Procediendo de manera iterativa se tiene

$$\mathbf{y} = \mathbf{r}^0 = \sum_{k=1}^n \Phi_{\mathcal{J}_k} \Phi_{\mathcal{J}_k}^* \mathbf{r}^{k-1}.$$

Sea $\tilde{\mathbf{x}}^k$ el vector de tamaño $N \times 1$ que coincide con $\Phi_{\mathcal{J}_k}^* \mathbf{r}^{k-1}$ sobre \mathcal{J}_k y es nulo fuera de \mathcal{J}_k . Sea $\Gamma^n = \cup_{k=1}^n \mathcal{J}_k$. Por el Teorema 2.3.5, $\Gamma^n \subset \Gamma^\sharp$; por tanto

$$\mathbf{y} = \sum_{k=1}^n \Phi_{\Gamma^n}(\tilde{\mathbf{x}}^k) = \Phi_{\Gamma^n} \left(\sum_{k=1}^n \tilde{\mathbf{x}}^k \right) = \Phi_{\Gamma^\sharp} \left(\sum_{k=1}^n \tilde{\mathbf{x}}^k \right)$$

ya que $\text{sop} \left(\sum_{k=1}^n \tilde{\mathbf{x}}^k \right) = \Gamma^n \subset \Gamma^\sharp$. Con el mismo argumento que en (2.50), la parte izquierda de la desigualdad de RIP (2.23) permite concluir $\mathbf{x} = \sum_{k=1}^n \tilde{\mathbf{x}}^k$. ■

2.3 Algoritmos ambiciosos en la percepción compresiva

Nota 2.3.8 Observe que $\delta_k \leq \delta_{k+1}$ ya que el conjunto de todos los vectores k -dispersos está contenido en el conjunto de todos los vectores $(k+1)$ -dispersos. Además, para lograr (2.49) las constantes RIP δ_{k+1} y δ_k deben satisfacer $\delta_{k+1} < \alpha(1 - \delta_k)/\sqrt{k}$. Esto da una restricción sobre k , la cual es $k < (\alpha(1 - \delta_k)/\delta_{k+1})^2$.

Nota 2.3.9 La demostración del Teorema 2.3.5 sigue las ideas de la demostración en [102], que considera diccionarios cuasi-incoherentes deterministas, los cuales usan el hecho de que $\mathbf{r} = P_{\Gamma^\#} \mathbf{r} = \mathbf{\Phi}_{\Gamma^\#} \mathbf{\Phi}_{\Gamma^\#}^\dagger \mathbf{r}$ (esto es, algoritmos tipo OMP) y luego acotan con desigualdades usuales entre normas. La condición obtenida es llamada Condición Exacta de Reconstrucción ERC_α para diccionarios cuasi-incoherentes ya que el resultado es sobre OMP y por tanto la reconstrucción exacta se garantiza en a lo mucho k iteraciones; un resultado similar es llamada Condición de Estabilidad en [50] ya que se aplica a MP y no hay garantía de reconstrucción exacta de señales dispersas en k iteraciones en este caso. En nuestro caso, una aplicación directa de las consecuencias de RIP nos lleva al resultado.

Nota 2.3.10 Supongamos que $\delta_{k+1} < \frac{\alpha}{\alpha + \sqrt{k}}$. Puesto que $\delta_k \leq \delta_{k+1}$ se tiene

$$\frac{\sqrt{k}\delta_{k+1}}{1 - \delta_k} < \frac{\sqrt{k} \frac{\alpha}{\alpha + \sqrt{k}}}{1 - \frac{\alpha}{\alpha + \sqrt{k}}} = \alpha.$$

Por tanto, la condición (2.49) del Teorema 2.3.5 se cumple. Por el Corolario 2.3.6, si $\delta_{k+1} < \frac{\alpha}{\alpha + \sqrt{k}}$ el algoritmo WOMP recupera cualquier señal k -dispersa \mathbf{x} en como mucho k -iteraciones. Cuando $\alpha = 1$, WOMP es conocido como OMP. Para OMP este resultado (es decir, con $\delta_{k+1} < \frac{1}{1 + \sqrt{k}}$) ha aparecido recientemente en [76] y [78] (ver también [25], [73] para resultados previos con menores cotas para δ_{k+1}). Además, se prueba en [78] que existe un vector k -disperso $\mathbf{x} \in \mathbb{R}^{k+1}$ y una matriz $\mathbf{\Phi}$ de tamaño $(k+1) \times (k+1)$ que satisface RIP con $\delta_{k+1} = \frac{1}{\sqrt{k}}$ tal que OMP no recupera \mathbf{x} en k iteraciones (probando una conjetura enunciada en [26]). Así, la cota $\delta_{k+1} < \frac{1}{1 + \sqrt{k}}$ es casi óptima.

El siguiente resultado exhibe una matriz para la que $\delta_{k+1} = \frac{\alpha}{\sqrt{k}}$, pero la regla de selección (2.47) no identifica los elementos del soporte de una señal k -dispersa. El resultado es una adaptación de un ejemplo propuesto en [78] para el algoritmo OMP (esto es, WOMP con $\alpha = 1$).

Proposición 2.3.11 Sea $0 < \alpha \leq 1$. Para todo entero $k \geq 2$ existe una señal k -dispersa \mathbf{x} y una matriz $\mathbf{\Phi}$ con $\delta_{k+1} = \frac{\alpha}{\sqrt{k}}$ tal que $\mathcal{J}_1(\alpha) \not\subset \text{sop}(\mathbf{x})$.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Demostración. Sea Φ la matriz de tamaño $(k+1) \times (k+1)$ dada por

$$\Phi = \begin{pmatrix} & & \frac{\alpha}{k} \\ & I_k & \vdots \\ 0 & \cdots & 0 & \frac{\alpha}{k} \\ & & & a \end{pmatrix}, \quad a = \sqrt{\frac{k - \alpha^2}{k}}.$$

Se tiene que

$$\Phi^* \Phi = \begin{pmatrix} & & 0 \\ & I_k & \vdots \\ \frac{\alpha}{k} & \cdots & \frac{\alpha}{k} & a \end{pmatrix} \begin{pmatrix} & & \frac{\alpha}{k} \\ & I_k & \vdots \\ 0 & \cdots & 0 & \frac{\alpha}{k} \\ & & & a \end{pmatrix} = \begin{pmatrix} & & \frac{\alpha}{k} \\ & I_k & \vdots \\ \frac{\alpha}{k} & \cdots & \frac{\alpha}{k} & 1 \end{pmatrix}.$$

El polinomio característico de la matriz $\Phi^* \Phi$ es

$$|\Phi^* \Phi - \lambda I| = -\frac{\alpha^2}{k}(1 - \lambda)^{k-1} + (1 - \lambda)^{k+1} = (1 - \lambda)^{k-1} \left[(1 - \lambda)^2 - \frac{\alpha^2}{k} \right],$$

que puede obtenerse desarrollando el determinante de la matriz $\Phi^* \Phi - \lambda I$ por la última fila (o columna). Como

$$(1 - \lambda)^2 - \frac{\alpha^2}{k} = \lambda^2 - 2\lambda + 1 - \frac{\alpha^2}{k} = 0$$

tiene como soluciones

$$\lambda = 1 - \frac{\alpha}{\sqrt{k}}, \quad \lambda = 1 + \frac{\alpha}{\sqrt{k}}$$

los autovalores de la matriz $\Phi^* \Phi$ son

$$\lambda_1 = \cdots = \lambda_{k-1} = 1, \lambda_k = 1 - \frac{\alpha}{\sqrt{k}}, \lambda_{k+1} = 1 + \frac{\alpha}{\sqrt{k}}.$$

Por tanto, la constante RIP de $\Phi^* \Phi$ de orden δ_{k+1} es $\delta_{k+1} = \frac{\alpha}{\sqrt{k}}$.

Considerar $\mathbf{x} = (\alpha, \dots, \alpha, 0) \in \mathbb{R}^{k+1}$ un vector k -disperso. Recordar que

$$J_1(\alpha) = \{i \in \{1, \dots, k+1\} : |\langle \phi_i, \mathbf{y} \rangle| \geq \alpha \|\Phi^* \mathbf{y}\|_{\ell^\infty(\mathbb{R}^{k+1})}\},$$

donde $\phi_i = \Phi e_i$ es la i -ésima columna de Φ . Si $1 \leq i \leq k$, $\phi_i = e_i$; como $\mathbf{y} = \Phi \mathbf{x} = \mathbf{x}$, se tiene que si $1 \leq i \leq k$, $\langle \phi_i, \mathbf{y} \rangle = \alpha$. Además, $\phi_{k+1} = (\frac{\alpha}{k}, \dots, \frac{\alpha}{k}, a)^t$, por lo que

$$\langle \phi_{k+1}, \mathbf{y} \rangle = \langle \phi_{k+1}, \mathbf{x} \rangle = \alpha^2.$$

Como $0 < \alpha \leq 1$, $\alpha^2 \leq \alpha$ y por tanto

$$\|\Phi^* \mathbf{y}\|_{\ell^\infty(\mathbb{R}^{k+1})} = \sup_{i=1, \dots, k+1} |\langle \phi_i, \mathbf{y} \rangle| = \alpha.$$

Como $\langle \phi_i, \mathbf{y} \rangle \geq \alpha^2$ para $i = 1, 2, \dots, k+1$ se tiene $J_1(\alpha) = \{1, \dots, k+1\}$ mientras que $\text{sop}(\mathbf{x}) = \{1, \dots, k\}$. ■

2.3.2.2 Identificación del soporte con condiciones relajadas

A continuación, consideramos otra regla de selección para recuperar índices en el soporte de una señal dispersa. Con la misma notación que en la Sección 2.3.2.1, sea $0 < \tilde{\alpha} \leq 1$ y definamos

$$\tilde{\mathcal{J}}_n := \tilde{\mathcal{J}}_n(\tilde{\alpha}) := \left\{ i : |g_i^n| \geq \tilde{\alpha} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \right\}, \quad (2.51)$$

la regla de selección *relajada débil*. La regla (2.51) compara el valor absoluto de $g_i^n = \langle \phi_i, \mathbf{r}^{n-1} \rangle$ con $\|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}$. Puede probarse (ver la prueba del Teorema 9.10 en [77], p. 422) que existe $\beta_0 > 0$ tal que para cualquier $\mathbf{x} \in \mathbb{R}^m$, $\sup_{i=1, \dots, N} |\langle \phi_i, \mathbf{x} \rangle| \geq \beta_0 \|\mathbf{x}\|_{\ell^2(\mathbb{R}^m)}$. Así, $\tilde{\mathcal{J}}_n(\tilde{\alpha})$ es un conjunto no vacío si escogemos $\tilde{\alpha} \leq \beta_0$.

La regla de selección (2.51) es similar a las consideradas en [37] y [23]. En [37] la regla es

$$\mathcal{J}_n(t) := \left\{ i : |g_i^n| \geq \frac{t}{\sqrt{m}} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \right\},$$

y el algoritmo desarrollado se llama StOMP. Así, $\tilde{\mathcal{J}}_n(\tilde{\alpha})$ es $\mathcal{J}_n(t)$ con $\tilde{\alpha} = t/\sqrt{m}$. En [23] la regla es

$$\mathcal{D}_n(\delta) := \left\{ i : |g_i^n| \geq \frac{t}{\sqrt{k}} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \right\},$$

y el algoritmo se llama DTresh (y su primo STresh). Así, $\tilde{\mathcal{J}}_n(\tilde{\alpha})$ es $\mathcal{D}_n(t)$ con $\tilde{\alpha} = t/\sqrt{k}$. En ambos algoritmos la actualización de la aproximación se hace usando la misma actualización que en algoritmo OMP.

En la regla (2.51) comparamos con la energía del residuo $\|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}$ en vez de comparar con el máximo de las correlaciones del residuo con los “átomos” de Φ , *i.e.* $\|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}$. Escogemos el término “relajado” ya que, como veremos en el Teorema 2.3.12, la condición sobre δ_{k+1} y δ_k es más débil que la requerida por el Teorema 2.3.5. Por lo tanto, hemos decidido llamar a los algoritmos ambiciosos con esta regla de selección algoritmos WMP, WOMP o WGP *Relajados*, escribiendo RWMP, RWOMP y RWGP, respectivamente.

Consideramos aún que \mathbf{x} es un vector k -disperso con soporte en Γ^\sharp .

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Teorema 2.3.12 *Sea Φ una matriz que satisface RIP con δ_{k+1} . Una condición suficiente para que los algoritmos RWGP, RWMP y RWOMP identifiquen elementos en el soporte Γ^\sharp de \mathbf{x} con la regla de selección **relajada débil** $\tilde{\mathcal{J}}_n(\tilde{\alpha})$ dada por (2.51) es que $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ y*

$$\tilde{\alpha} > \frac{\delta_{k+1}}{(1 - \delta_k)^{1/2}}. \quad (2.52)$$

Demostración. La prueba es similar a la del Teorema 2.3.5. El Lema 2.3.3 es cierto si se reemplaza $\mathcal{J}_k(\alpha)$ por $\tilde{\mathcal{J}}_k(\tilde{\alpha})$, $k = 1, \dots, n$, y se consideran los algoritmos RWMP, RWOMP y RWGP, puesto que su demostración depende de los algoritmos y no de la regla de selección.

El Lema 2.3.4 es el que requiere alguna modificación. Hemos de probar que $\tilde{\mathcal{J}}(\tilde{\alpha}) \subset \Gamma^\sharp = \text{sop}(\mathbf{x})$ cuando $\delta_{k+1} < \tilde{\alpha}(1 - \delta_k)^{1/2}$. Tomando complementarios basta probar que

$$\|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})} < \tilde{\alpha} \|\mathbf{r}\|_{\ell^2(\mathbb{R}^m)}.$$

Puesto que $\mathbf{r} \in \text{span}(\Phi_{\Gamma^\sharp})$ podemos escribir $\mathbf{r} = \Phi_{\Gamma^\sharp} \mathbf{z}$ con $\text{sop} \mathbf{z} \subset \Gamma^\sharp$. Sea λ un índice de $\Gamma^{\sharp c}$ para el que

$$|\langle \phi_\lambda, \mathbf{r} \rangle| = \|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})}.$$

Observar que $|\langle \phi_\lambda, \mathbf{r} \rangle| = |\Phi_\lambda^* \mathbf{r}| = \|\Phi_\lambda^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^\lambda)} = \|\Phi_\lambda^* \mathbf{r}\|_{\ell^2(\mathbb{R}^\lambda)}$ donde Φ_λ es la matriz de tamaño $m \times 1$ formada por la columna λ de Φ . Entonces,

$$\begin{aligned} \frac{\|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})}}{\|\mathbf{r}\|_{\ell^2(\mathbb{R}^m)}} &= \frac{\|\Phi_\lambda^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^\lambda)}}{\|\Phi_{\Gamma^\sharp} \mathbf{z}\|_{\ell^2(\mathbb{R}^m)}} = \frac{\|\Phi_\lambda^* \Phi_{\Gamma^\sharp} \mathbf{z}\|_{\ell^2(\mathbb{R}^\lambda)}^2}{\|\Phi_{\Gamma^\sharp} \mathbf{z}\|_{\ell^2(\mathbb{R}^m)}} \\ &\leq \frac{\delta_{k+1} \|\mathbf{z}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}}{(1 - \delta_k)^{1/2} \|\mathbf{z}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}} \leq \frac{\delta_{k+1}}{(1 - \delta_k)^{1/2}} < \tilde{\alpha}, \end{aligned}$$

donde la desigualdad se debe a (2.45) en el numerador y la parte izquierda de la condición RIP (2.23) en el denominador.

La demostración del Teorema se termina por inducción con el mismo argumento que se dio para finalizar la demostración del Teorema 2.3.5. \blacksquare

Nota 2.3.13 *Supóngase $\mathbf{r} \in \text{span}(\Phi_{\Gamma^\sharp})$ donde Φ es una matriz RIP; entonces*

$$\|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})} \geq \frac{1}{\sqrt{k}} \|\Phi_{\Gamma^\sharp}^* \mathbf{r}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})} \geq \frac{(1 - \delta_k)^{1/2}}{\sqrt{k}} \|\mathbf{r}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})},$$

2.3 Algoritmos ambiciosos en la percepción compresiva

por la desigualdad usual entre normas y el Lema 2.3.2. Así, tomando

$$\tilde{\alpha} \leq \frac{(1 - \delta_k)^{1/2}}{\sqrt{k}} \quad (2.53)$$

siempre se tiene $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$ en cada iteración de los algoritmos con regla de selección Relajada. Nótese que (2.52) y (2.53) podrían verificarse al mismo tiempo para algún valor de $\tilde{\alpha}$ sólo si

$$\frac{\delta_{k+1}}{(1 - \delta_k)^{1/2}} \leq \frac{(1 - \delta_k)^{1/2}}{\sqrt{k}},$$

lo cual da la siguiente restricción: $\sqrt{k} \leq \frac{1 - \delta_k}{\delta_{k+1}}$.

Nota 2.3.14 El Teorema 2.3.12 podría considerarse como una Condición de Reconstrucción Exacta para el algoritmo de Búsqueda de Correspondencia Ortogonal **Relajado Débil** (RWOMP) con matrices RIP y con parámetro δ_{k+1} (ver el Corolario 2.3.6) siempre y cuando (2.53) sea satisfecha. La probabilidad de éxito depende exclusivamente de la probabilidad de que la matriz aleatoria verifique RIP. Se ha probado que las matrices gaussianas, de Bernoulli y parcial aleatoria de Fourier verifican RIP con muy alta probabilidad (concentración exponencial) siempre y cuando el número de mediciones $m \geq Ck \log(N/k)$ para las dos primeras matrices (ver [20],[19]) y $m \geq Ck \log^5(N)$ para la matriz parcial aleatoria de Fourier (ver [20], [92]).

Nota 2.3.15 Puesto que RWOMP selecciona siempre nuevos índices debido a la ortogonalidad expresada en (2.37), el Corolario 2.3.6 también se cumple en esta situación. Por tanto, si se cumplen las condiciones de Teorema 2.3.12, el algoritmo RWOMP recupera cualquier señal k -dispersa \mathbf{x} en como mucho k -iteraciones.

Nota 2.3.16 También se cumple el Corolario 2.3.7 si estamos bajo las hipótesis del Teorema 2.3.12, puesto que no es necesario modificar la demostración. Por tanto, si se cumplen las condiciones del Teorema 2.3.12 y suponemos que $\mathbf{r}^n = 0$ para algún n en cualquiera de los algoritmos RWMP, RWOMP o RWGP, el algoritmo recupera una señal k -dispersa \mathbf{x} en como mucho n iteraciones.

2.3.3 Convergencia

En esta sección estudiamos la convergencia de los algoritmos desarrollados en las Secciones 2.3.1 y 2.3.2.2 para matrices que cumplen RIP (2.23). los resultados se dan en términos de la reducción de la energía de los residuos $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$ de la observación $\mathbf{y} = \Phi \mathbf{x}$, en lugar de darlos en términos de la energía de los residuos de la aproximación, $\mathbf{x} - \mathbf{x}^n$, como se hace en [83] para CoSaMP y en [23] para DThresh.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

2.3.3.1 Convergencia de GP, WGP y RWGP

En [11] el análisis de convergencia de GP se basa en un teorema de existencia (ver Teorema 9.10 en [77]) cuando $\Phi \in \mathbb{C}^{N \times N_1}$ es un diccionario (esto significa que Φ contiene al menos una base para \mathbb{R}^N y así $N_1 \geq N$) que verifica la *Condición de Reconstrucción Exacta* $\text{ERC}_\alpha(\Gamma)$ para diccionarios *cuasi-incoherentes* (ver [102]). El análisis en [103] se hace para matrices aleatorias *admisibles*. El resultado en el siguiente teorema se obtiene para matrices que satisfacen RIP. Consideramos aún que \mathbf{x} es un vector k -disperso con soporte en Γ^\sharp .

Teorema 2.3.17 *Considérense los algoritmos GP, WGP y RWGP. Supongamos que en la iteración n tenemos $\Gamma^s \subset \Gamma^\sharp$, $s = 1, 2, \dots, n$. Sea Φ que verifica RIP con parámetro δ_k . Entonces, para todo vector $\mathbf{x} \in \mathbb{R}^n$ que sea k -disperso con $\text{sop}(\mathbf{x}) = \Gamma^\sharp$,*

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq C_k \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \quad (2.54)$$

con $C_k = (1 - \frac{1-\delta_k}{k(1+\delta_k)})^{1/2} < 1$. En el caso RWGP suponemos que $\tilde{\alpha} \leq \frac{(1-\delta_k)^{1/2}}{\sqrt{k}}$ para que $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$.

Demostración. Para simplificar la notación escribiremos $\mathbf{d}^n = \mathbf{d}_{\Gamma^n}^n$. Tenemos

$$\begin{aligned} \|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)}^2 &= \langle \mathbf{r}^{n-1} - a^n \Phi_{\Gamma^n} \mathbf{d}^n, \mathbf{r}^{n-1} - a^n \Phi_{\Gamma^n} \mathbf{d}^n \rangle \\ &= \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - \overline{a^n} \langle \mathbf{r}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}^n \rangle \\ &\quad - a^n \langle \Phi_{\Gamma^n} \mathbf{d}^n, \mathbf{r}^{n-1} \rangle + \langle a^n \Phi_{\Gamma^n} \mathbf{d}^n, a^n \Phi_{\Gamma^n} \mathbf{d}^n \rangle \\ &= \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - 2 \frac{|\langle \Phi_{\Gamma^n} \mathbf{d}^n, \mathbf{r}^{n-1} \rangle|}{\|\Phi_{\Gamma^n} \mathbf{d}^n\|_2^2} + \frac{|\langle \Phi_{\Gamma^n} \mathbf{d}^n, \mathbf{r}^{n-1} \rangle|^2}{\|\Phi_{\Gamma^n} \mathbf{d}^n\|_2^4} \|\Phi_{\Gamma^n} \mathbf{d}^n\|_{\ell^2(\mathbb{R}^m)}^2 \\ &= \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - \frac{|\langle \mathbf{r}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}^n \rangle|^2}{\|\Phi_{\Gamma^n} \mathbf{d}^n\|_{\ell^2(\mathbb{R}^m)}^2}. \end{aligned} \quad (2.55)$$

Ya que $\mathbf{d}^n = \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}$, el segundo término puede acotarse por abajo con

$$\begin{aligned} \frac{|\langle \mathbf{r}^{n-1}, \Phi_{\Gamma^n} \mathbf{d}^n \rangle|^2}{\|\Phi_{\Gamma^n} \mathbf{d}^n\|_{\ell^2(\mathbb{R}^m)}^2} &= \frac{|\langle \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}, \Phi_{\Gamma^n}^* \mathbf{r}^{n-1} \rangle|^2}{\|\Phi_{\Gamma^n} \Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2} \\ &\geq \frac{\|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^n})}^4}{\|\Phi_{\Gamma^n}\|_{\ell^2(\mathbb{R}^{\Gamma^n}) \rightarrow \ell^2(\mathbb{R}^m)}^2 \|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^n})}^2} \\ &\geq \frac{\|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^n})}^2}{1 + \delta_k}. \end{aligned} \quad (2.56)$$

2.3 Algoritmos ambiciosos en la percepción compresiva

Afirmamos que, debido a la definición de $\mathcal{J}_n(\alpha) = \mathcal{J}_n$ en GP y WGP se tiene

$$\|\Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\mathcal{J}_n})} = \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}. \quad (2.57)$$

En efecto, si $j_0 \in \{1, \dots, N\}$ es un índice tal que

$$|\langle \phi_{j_0}, \mathbf{r}^{n-1} \rangle| = \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}$$

se ha de tener $j_0 \in \mathcal{J}_n(\alpha) = \mathcal{J}_n$ ya que

$$|\langle \phi_{j_0}, \mathbf{r}^{n-1} \rangle| = \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)} \geq \alpha \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}.$$

Por tanto

$$\begin{aligned} \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)} &\geq \|\Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\mathcal{J}_n})} \geq |\langle \phi_{j_0}, \mathbf{r}^{n-1} \rangle| \\ &= \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}, \end{aligned}$$

lo que prueba la afirmación (2.57).

Además, como $\mathcal{J}_n \subset \Gamma^n$ y estamos suponiendo $\Gamma^n \subset \Gamma^\sharp$ es claro que

$$\|\Phi_{\mathcal{J}_n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\mathcal{J}_n})} = \|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^n})} = \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})} = \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}. \quad (2.58)$$

Se tiene que

$$\|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^n})}^2 \geq \|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^n})}^2 = \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}^2 \geq \frac{1}{k} \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})}^2.$$

Puesto que $\mathbf{r}^n \in \text{span}(\Phi_{\Gamma^\sharp})$ (ver Lema 2.3.3), por el Lema 2.3.2 podemos escribir

$$\|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^n})}^2 \geq \frac{1 - \delta_k}{k} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2. \quad (2.59)$$

Sustituyendo este resultado en (2.56) y (2.55) se obtiene

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)}^2 \leq \left(1 - \frac{1 - \delta_k}{k(1 + \delta_k)}\right) \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2$$

lo que demuestra el resultado para los algoritmos GP y WGP.

Para RWGP la regla de selección

$$\tilde{\mathcal{J}}_n(\tilde{\alpha}) = \{i : |\langle \phi_i, \mathbf{r}^{n-1} \rangle| \geq \tilde{\alpha} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}\}$$

no nos permite escribir (2.57). Pero ahora se tiene

$$\|\Phi_{\tilde{\mathcal{J}}_n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\tilde{\mathcal{J}}_n})} = \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}. \quad (2.60)$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

En efecto, como $\mathbf{r} \in \text{span}(\Phi_{\Gamma^\sharp})$ (la prueba es como la del Lema 2.3.3 pero para $\tilde{\mathcal{J}}$) y Φ es RIP con parámetros δ_k , de la Nota 2.3.13

$$\|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})} \geq \frac{(1-\delta_k)^{1/2}}{\sqrt{k}} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}. \quad (2.61)$$

Sea $j_0 \in \Gamma^\sharp$ un índice para el que

$$|\langle \phi_{j_0}, \mathbf{r}^{n-1} \rangle| = \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}.$$

Entonces, $j_0 \in \tilde{\mathcal{J}}_n(\tilde{\alpha})$ porque

$$|\langle \phi_{j_0}, \mathbf{r}^{n-1} \rangle| = \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})} \geq \frac{(1-\delta_k)^{1/2}}{\sqrt{k}} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \geq \tilde{\alpha} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}$$

ya que $\tilde{\alpha} \leq \frac{(1-\delta_k)^{1/2}}{\sqrt{k}}$ para que $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$. Entonces, como $\tilde{\mathcal{J}}_n \subset \Gamma^\sharp$, se tiene

$$\begin{aligned} \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})} &\geq \|\Phi_{\tilde{\mathcal{J}}_n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\tilde{\mathcal{J}}_n})} \geq |\langle \phi_{j_0}, \mathbf{r}^{n-1} \rangle| \\ &= \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}, \end{aligned}$$

lo que prueba (2.60).

Usando (2.60) y (2.61) podemos escribir

$$\begin{aligned} \|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{\Gamma^n})}^2 &\geq \|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^n})}^2 \geq \|\Phi_{\tilde{\mathcal{J}}_n}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\tilde{\mathcal{J}}_n})}^2 \\ &= \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})}^2 \geq \frac{1-\delta_k}{k} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 \end{aligned} \quad (2.62)$$

lo que prueba (2.59) para RWGP y por tanto el resultado. \blacksquare

Nota 2.3.18 En [11, Teorema 3] se prueba que $\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq c_k \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}$ con $c_k^2 = (1 - \frac{\omega}{\|\Phi\|_2^2})$ donde ω es un número positivo real tal que $\|\Phi \mathbf{x}\|_{\ell^\infty(\mathbb{R}^m)}^2 \geq \omega \|\mathbf{x}\|_{\ell^2(\mathbb{R}^m)}^2$ para todo $\mathbf{x} \in \mathbb{R}^m$. Nuestro Teorema 2.3.17 da un valor de c_k que depende de δ_k y la dispersión de \mathbf{x} . Este valor de c_k es un número menor que 1, pero muy cercano a 1 cuando k aumenta. Es fácil comprobar que nunca puede tenerse $c_k \leq 1/2$ cuando $k \geq 2$.

Usando los Teoremas 2.3.5, 2.3.12 y 2.3.17 se deduce lo siguiente:

Corolario 2.3.19 Suponga que Φ satisface RIP con constantes δ_k y δ_{k+1} .

i) Suponga que $\frac{\sqrt{k}\delta_{k+1}}{1-\delta_k} < 1$; entonces, el algoritmo GP satisface (2.54).

2.3 Algoritmos ambiciosos en la percepción compresiva

ii) Sea $0 < \alpha \leq 1$ y suponga $\frac{\sqrt{k}\delta_{k+1}}{1-\delta_k} < \alpha$; entonces, el algoritmo WGP con regla de selección $J_n(\alpha)$ satisface (2.54).

iii) Suponga que $\frac{\delta_{k+1}}{1-\delta_k} < \frac{1}{\sqrt{k}}$ y que $\tilde{\alpha}$ se escoge tal que $\frac{\delta_{k+1}}{(1-\delta_k)^{1/2}} < \tilde{\alpha} < \frac{(1-\delta_k)^{1/2}}{\sqrt{k}}$; entonces, el algoritmo RWGP con regla de selección $\tilde{J}_n(\tilde{\alpha})$ satisface (2.54).

Nota 2.3.20 La convergencia de los algoritmos del tipo Gradient Pursuit (GP, WGP y RWGP) está dada en el Teorema 2.3.17 en términos de la convergencia de la energía de los residuos $\mathbf{r}^n = \mathbf{y} - \mathbf{y}^n$ de la observación. Si Φ satisface RIP la convergencia de los residuos de la aproximación $\mathbf{x} - \mathbf{x}^n$ puede obtenerse a partir del resultado anterior. En efecto, para algoritmos del tipo GP sabemos que $\mathbf{y}^n = \Phi \mathbf{x}^n$ (ver 2.40). Entonces,

$$\begin{aligned} \|\mathbf{y} - \mathbf{y}^n\|_{\ell^2(\mathbb{R}^m)} &= \|\Phi \mathbf{x} - \Phi \mathbf{x}^n\|_{\ell^2(\mathbb{R}^m)} = \|\Phi_{\Gamma^\sharp}(\mathbf{x} - \mathbf{x}^n)\|_{\ell^2(\mathbb{R}^m)} \\ &\geq (1 - \delta_k)^{1/2} \|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)} \end{aligned} \quad (2.63)$$

por la desigualdad izquierda de RIP siempre que $\Gamma^n \subset \Gamma^\sharp$ (que se cumple bajo las condiciones de los Teoremas 2.3.5 y 2.3.12). Análogamente, si también suponemos $\Gamma^{n-1} \subset \Gamma^\sharp$ tenemos

$$\begin{aligned} \|\mathbf{y} - \mathbf{y}^{n-1}\|_{\ell^2(\mathbb{R}^m)} &= \|\Phi \mathbf{x} - \Phi \mathbf{x}^{n-1}\|_{\ell^2(\mathbb{R}^m)} = \|\Phi_{\Gamma^\sharp}(\mathbf{x} - \mathbf{x}^{n-1})\|_{\ell^2(\mathbb{R}^m)} \\ &\leq (1 + \delta_k)^{1/2} \|\mathbf{x} - \mathbf{x}^{n-1}\|_{\ell^2(\mathbb{R}^N)} \end{aligned} \quad (2.64)$$

por la desigualdad derecha de RIP. De (2.63) y (2.64) se deduce

$$\|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)} \leq C_k \left(\frac{1 + \delta_k}{1 - \delta_k}\right)^{1/2} \|\mathbf{x} - \mathbf{x}^{n-1}\|_{\ell^2(\mathbb{R}^N)}, \quad (2.65)$$

donde C_k es la constante del Teorema 2.3.17. Observar que, además de las condiciones del Corolario 2.3.19, si queremos asegurar la convergencia en $\ell^2(\mathbb{R}^N)$ de \mathbf{x}^n a \mathbf{x} necesitamos

$$\left(\frac{1 + \delta_k}{1 - \delta_k}\right)^{1/2} \left(1 - \frac{1 - \delta_k}{k(1 + \delta_k)}\right)^{1/2} < 1,$$

lo que requiere $\delta_k < \frac{1}{2k+1}$.

Si a partir de alguna iteración se tuviera $\Gamma^n = \Gamma^\sharp = \text{sop}(\mathbf{x})$ la reducción de la energía de los residuos sería mayor que la dada en (2.54), como muestra el siguiente resultado.

Teorema 2.3.21 Considerar los algoritmos GP, WGP y RWGP y supongamos que en la iteración n_0 tenemos $\Gamma^{n_0} = \Gamma^\sharp = \text{sop}(\mathbf{x})$. Supongamos que $\alpha, \tilde{\alpha} > \frac{\delta_{k+1}}{(1-\delta_k)^{1/2}}$ y que Φ verifica RIP con parámetro δ_{k+1} . Entonces, para todo $n \geq n_0$,

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq D_k \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \quad (2.66)$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

con $D_k = (1 - \frac{1-\delta_k}{1+\delta_k})^{1/2} = (\frac{2\delta_k}{1+\delta_k})^{1/2} < 1$.

Demostración. La igualdad (2.55) y la desigualdad (2.56) de la demostración del Teorema 2.3.17 siguen siendo válidas en nuestro contexto. Puesto que $\Gamma^{n_0} = \Gamma^\sharp$ se tiene $\Gamma^n = \Gamma^\sharp$ para $n \geq n_0$ ya que $\alpha, \tilde{\alpha} > \frac{\delta_{k+1}}{(1-\delta_k)^{1/2}}$ nos permite concluir que $\tilde{\mathcal{J}}_{n_0}, \tilde{\mathcal{J}}_n \subset \Gamma^\sharp$.

Por tanto, podemos remplazar Γ^n por Γ^\sharp en (2.56) y como $\mathbf{r}^{n-1} \in \text{span}(\Phi_{\Gamma^\sharp})$, por el Lema 2.3.2 podemos escribir

$$\begin{aligned} \|\Phi_{\Gamma^n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 &= \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 \\ &\geq (1 - \delta_k) \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2. \end{aligned}$$

Sustituyendo esta desigualdad en (2.55) y (2.56) se obtiene el resultado para los tres algoritmos. ■

Nota 2.3.22 *La constante D_k del Teorema 2.3.21 puede hacerse tan cercana a 0 como se desee haciendo δ_k pequeño. En particular, basta tomar $\delta_k \leq 1/7 \approx 0.143$ para reducir la energía del residuo a la mitad en una sola iteración, si las condiciones del Teorema 2.3.17 se satisfacen.*

Razonando como en la Nota 2.3.20 se tiene

$$\|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)} \leq \left(\frac{1+\delta_k}{1-\delta_k}\right)^{1/2} \left(1 - \frac{1-\delta_k}{1+\delta_k}\right)^{1/2} \|\mathbf{x} - \mathbf{x}^{n-1}\|_{\ell^2(\mathbb{R}^N)}$$

por lo que basta tener $\delta_k < 1/3$ para asegurar la convergencia de \mathbf{x}^n a \mathbf{x} en $\ell^2(\mathbb{R}^N)$ una vez que $\Gamma^{n_0} = \Gamma^\sharp$.

2.3.3.2 Convergencia de WMP y RWMP

En esta sección estudiamos la convergencia de los algoritmos WMP y RWMP descritos en las Secciones 2.3.1 y 2.3.2.2, respectivamente. Los resultados son similares a los obtenidos en [50] para diccionarios cuasi-incoherentes.

Teorema 2.3.23 *Considerar los algoritmos WMP y RWMP y supongamos que se cumplen las condiciones de los Teoremas 2.3.5 (para WMP) y 2.3.12 (para RWMP) para tener $\Gamma^s \subset \Gamma^\sharp$, $s = 1, 2, \dots$. Entonces, para todo $\mathbf{x} \in \mathbb{R}^N$ que sea k -disperso con $\text{sop}(\mathbf{x}) = \Gamma^\sharp$,*

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq C'_k \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \quad (2.67)$$

con $C'_k = (1 - \frac{(1-\delta_k)^2}{k})^{1/2} < 1$.

Demostración. Tanto para WMP como para RWMP tenemos

$$\begin{aligned}
 \|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)}^2 &= \langle \mathbf{r}^{n-1} - \Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1}, \mathbf{r}^{n-1} - \Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1} \rangle \\
 &= \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - \langle \mathbf{r}^{n-1}, \Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1} \rangle \\
 &\quad - \langle \Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1}, \mathbf{r}^{n-1} \rangle + \|\Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 \\
 &= \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - 2 \|\Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{J_n})}^2 + \|\Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2. \quad (2.68)
 \end{aligned}$$

Por (2.42) tenemos $\|\Phi_{J_n} \Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 \leq (1 + \delta_k) \|\Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{J_n})}^2$ ya que $J_n \subset \Gamma^\sharp$ y $|\Gamma^\sharp| = k$. Por tanto, de (2.68) obtenemos

$$\begin{aligned}
 \|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)}^2 &\leq \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - 2 \|\Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{J_n})}^2 + (1 + \delta_k) \|\Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{J_n})}^2 \\
 &= \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2 - (1 - \delta_k) \|\Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{J_n})}^2. \quad (2.69)
 \end{aligned}$$

Tanto para WMP como para RWMP se tiene

$$\|\Phi_{J_n}^* \mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^{J_n})}^2 \geq \frac{1 - \delta_k}{k} \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2. \quad (2.70)$$

Para WMP la demostración es como en (2.59) del Teorema 2.3.17 y para RWMP sustituyendo Γ^n por $\tilde{J}_n \subset \Gamma^\sharp$ la demostración es como en (2.62).

Sustituyendo (2.70) en (2.69) se obtiene

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)}^2 \leq \left(1 - \frac{(1 - \delta_k)^2}{k}\right) \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}^2$$

que es el resultado deseado. ■

Corolario 2.3.24 *Supongamos que Φ satisface RIP con constantes δ_k y δ_{k+1} .*

a) *Sea $0 < \alpha \leq 1$ y supongamos que $\frac{\sqrt{k}\delta_{k+1}}{1-\delta_k} < \alpha$; entonces el algoritmo WMP con regla de selección $J_n(\alpha)$ satisface (2.67).*

b) *Supongamos que $\frac{\delta_{k+1}}{1-\delta_k} < \frac{1}{\sqrt{k}}$ y que $\tilde{\alpha}$ se elige de manera que*

$$\frac{\delta_{k+1}}{(1 - \delta_k)^{1/2}} < \tilde{\alpha} < \frac{(1 - \delta_k)^2}{\sqrt{k}};$$

entonces, el algoritmo RWMP con regla de selección $\tilde{J}_n(\tilde{\alpha})$ satisface (2.67).

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Nota 2.3.25 La constante C'_k del Teorema 2.3.23 es un número menor que 1, por lo que siempre hay disminución de la energía del residuo. Sin embargo, es cercana a 1 cuando k aumenta. Es fácil comprobar que nunca puede tenerse $C'_k \leq 1/2$ cuando $k \geq 2$.

Nota 2.3.26 Como en la Nota 2.3.20 la disminución de la energía de los residuos dada en (2.67) para WMP y RWMP puede traducirse en convergencia de las estimaciones. Así,

$$\|\mathbf{x} - \mathbf{x}^n\|_{\ell^2(\mathbb{R}^N)} \leq \left(\frac{1 + \delta_k}{1 - \delta_k}\right)^{1/2} C'_k \|\mathbf{x} - \mathbf{x}^{n-1}\|_{\ell^2(\mathbb{R}^N)}.$$

Para que $\left(\frac{1 + \delta_k}{1 - \delta_k}\right)^{1/2} \left(1 - \frac{(1 - \delta_k)^2}{k}\right)^{1/2}$ sea menor que 1 se ha de tener

$$1 - \frac{(1 - \delta_k)^2}{k} < \frac{1 - \delta_k}{1 + \delta_k} \Leftrightarrow \frac{2\delta_k}{1 + \delta_k} < \frac{(1 - \delta_k)^2}{k}.$$

Como $(1 + \delta_k)(1 - \delta_k)^2 = (1 - \delta_k)(1 - \delta_k^2) = 1 - \delta_k - \delta_k^2 + \delta_k^3$ la desigualdad anterior es equivalente a

$$2k\delta_k < 1 - \delta_k - \delta_k^2 + \delta_k^3 \Leftrightarrow \delta_k^3 - \delta_k^2 - (1 + 2k)\delta_k + 1 > 0.$$

Si $\delta_k < \frac{1}{2k+2}$, se tiene que

$$1 > \delta_k + (2k + 1)\delta_k > \delta_k^2 + (2k + 1)\delta_k > \delta_k^2 + (2k + 1)\delta_k - \delta_k^3,$$

por lo que

$$\delta_k < \frac{1}{2k + 2}$$

es suficiente para tener convergencia de las aproximaciones en WMP y RWMP.

2.3.3.3 Convergencia de WOMP y RWOMP

Tanto en WOMP como en RWOMP el residuo \mathbf{r}^n es el vector que realiza la distancia de \mathbf{y} al subespacio $V_n = \{\Phi_{\Gamma^n} \mathbf{x} : \mathbf{x} \in \mathbb{R}^n \text{ con } \text{sop}(\mathbf{x}) \subset \Gamma^n\}$. Como $\Gamma^{n-1} \subset \Gamma^n$, es claro que

$$\|\mathbf{r}^n\|_{\ell^2(\mathbb{R}^m)} \leq \|\mathbf{r}^{n-1}\|_{\ell^2(\mathbb{R}^m)}$$

siempre se cumple. Puede obtenerse una desigualdad estricta observando que el residuo en WOMP o RWOMP, denominado temporalmente \mathbf{r}_{OMP}^n tiene una energía que no supera a la de los residuos en WGP, RWGP, WMP, o RWMP, denominados temporalmente \mathbf{r}_{GP}^n y \mathbf{r}_{MP}^n , ya que tanto \mathbf{y}_{GP}^n como \mathbf{y}_{MP}^n son elementos de V_n .

2.3 Algoritmos ambiciosos en la percepción compresiva

Por tanto, si se cumplen las condiciones del Teorema 2.3.5 (para WGP) o 2.3.12 (para RWGP) podemos aplicar el Teorema 2.3.17 para obtener

$$\begin{aligned} \|\mathbf{r}_{OMP}^n\|_{\ell^2(\mathbb{R}^m)} &\leq \|\mathbf{r}_{GP}^n\|_{\ell^2(\mathbb{R}^m)} \leq C_k \|\mathbf{r}_{GP}^{n-1}\|_{\ell^2(\mathbb{R}^m)} \\ &\leq C_k^n \|\mathbf{r}_{GP}^0\|_{\ell^2(\mathbb{R}^m)} \leq C_k^n \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}, \end{aligned} \quad (2.71)$$

con $C_k < 1$, la constante del Teorema 2.3.17.

Análogamente, pero usando el Teorema 2.3.23 se obtiene

$$\|\mathbf{r}_{OMP}^n\|_{\ell^2(\mathbb{R}^m)} \leq C'_k{}^n \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}, \quad (2.72)$$

con $C'_k < 1$ la constante del Teorema 2.3.23. Entre ambas constantes C_k y C'_k la relación es $C_k < C'_k$ ya que

$$\begin{aligned} C_k < C'_k &\Leftrightarrow \frac{1 - \delta_k}{k(1 + \delta_k)} > \frac{(1 - \delta_k)^2}{k} \Leftrightarrow \frac{1}{1 + \delta_k} > 1 - \delta_k \\ &\Leftrightarrow 1 > 1 - \delta_k^2, \end{aligned}$$

y la última desigualdad es cierta porque $0 < \delta_k < 1$. Por tanto (2.71) da mayor convergencia que (2.72) y prueba que los algoritmos GP convergen más rápidamente que los algoritmos del tipo MP.

Puesto que \mathbf{r}_{OMP}^n es un vector perpendicular a V_n , debería poderse mejorar la constante en (2.71). Como

$$\mathbf{r}_{OMP}^n = \mathbf{y} - \mathbf{y}^n = \mathbf{y} - \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y}$$

donde $\Phi_{\Gamma^n}^\dagger$ es la pseudo-inversa de Φ_{Γ^n} , tenemos que estudiar

$$\begin{aligned} \|\mathbf{r}_{OMP}^n\|_{\ell^2(\mathbb{R}^m)}^2 &= \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}^2 - \langle \mathbf{y}, \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y} \rangle - \langle \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y}, \mathbf{y} \rangle \\ &\quad + \left\| \Phi_{\Gamma^n} \Phi_{\Gamma^n}^\dagger \mathbf{y} \right\|_{\ell^2(\mathbb{R}^m)}^2. \end{aligned} \quad (2.73)$$

Aún no somos capaces de hallar una acotación para (2.73) de la forma

$$\|\mathbf{r}_{OMP}^n\|_{\ell^2(\mathbb{R}^m)} \leq B_k^n \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)} \quad (2.74)$$

con $B_k < C_k$. En el caso de que se cumplan las condiciones de los Teoremas 2.3.5 y 2.3.12 (por ejemplo, si $\delta_{k+1} < \alpha \frac{1-\delta_k}{\sqrt{k}}$ para WOMP y $\frac{\delta_{k+1}}{(1-\delta_k)^{1/2}} < \tilde{\alpha} < \frac{(1-\delta_k)^{1/2}}{\sqrt{k}}$ para RWOMP) las desigualdades (2.71) y (2.72) son triviales si $n \geq k$ ya que los

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

algoritmos del tipo OMP identifican el soporte de una señal k -dispersa \mathbf{x} en como mucho k iteraciones, y entonces $\mathbf{r}^n = 0$.

Por tanto, sólo resulta interesante hallar acotaciones de la forma (2.74) a partir de la expresión (2.73) si conseguimos que se cumpla con valores de δ_{k+1} y δ_k menos restrictivos que los de los Teoremas 2.3.5 y 2.3.12, para los que los algoritmos no identifican el soporte de \mathbf{x} .

2.3.4 Comportamiento de las reglas de selección para algunas matrices aleatorias

El lector puede encontrar en [30] la forma de construir matrices que satisfacen RIP. Estas matrices son de orden $m \times N$ con $m = p^2$ (p un número primo) y $N = p^{r+1}$, $0 < r < p$, y satisfacen RIP con $k < \frac{p}{r} + 1$ y $\delta_k = (k-1)r/p$. Por tanto, $m = p^2 > (k-1)^2 r^2 > (k-1)^2$; que da un valor de m mucho mayor que el necesario para recuperar una señal k -dispersa con el problema de optimización (P_1) que es $m \geq Ck \log(N/k)$.

Es conocido (ver [6]) que existen matrices aleatorias que satisfacen RIP con parámetro δ_k para cualquier $m \geq Ck \log(N/k)$ con probabilidad mayor a $1 - 2e^{-c'm}$. Entre estas se encuentran las que satisfacen una desigualdad de **concentración de la medida**, a saber,

$$\mathbb{P} \left\{ \left| \|\Phi(\omega)\mathbf{x}\|_{\ell^2(\mathbb{R}^m)} - \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)} \right| \geq \varepsilon \|\mathbf{x}\|_{\ell^2(\mathbb{R}^N)}^2 \right\} \leq 2e^{-m c_0(\varepsilon)}, \quad 0 < \varepsilon < 1, \quad (2.75)$$

donde la probabilidad se toma sobre todas las matrices aleatorias $\Phi(\omega)$ de tamaño $m \times N$ y $c_0(\varepsilon) > 0$ es una constante que depende sólo de ε .

Un ejemplo de tales matrices son aquellas $\Phi = (\phi_{i,j})$ tales que $\phi_{i,j}$ es una realización independiente de una variable aleatoria gaussiana $N(0, 1/\sqrt{m})$, es decir, de media 0 y desviación normal $1/\sqrt{m}$. En este caso se tiene $c_0(\varepsilon) = \frac{\varepsilon^2}{4} - \frac{\varepsilon^3}{6}$ (ver [6]).

Otro ejemplo son las matrices cuyas entradas son variables aleatorias independientes de Bernoulli con valores $\{-1/\sqrt{m}, 1/\sqrt{m}\}$, con probabilidad $1/2$ en cada caso. En este caso también se tiene $c_0(\varepsilon) = \frac{\varepsilon^2}{4} - \frac{\varepsilon^3}{6}$ (ver [1]).

En esta sección estudiaremos el comportamiento de las reglas de selección y los algoritmos dados en las secciones 2.3.1 y 2.3.2 con respecto a matrices aleatorias Φ como las descritas en los dos párrafos anteriores. El objeto es probar directamente que este tipo de matrices selecciona elementos del soporte de una señal k -dispersa con

2.3 Algoritmos ambiciosos en la percepción compresiva

gran probabilidad, siguiendo los razonamientos dados en [103] para *Orthogonal Matching Pursuit* (OMP).

Comenzamos con un resultado sobre procesos aleatorios (ver [103] y las referencias citadas en este artículo), cuya demostración se da por completitud:

Lema 2.3.27 a) Sea \mathbf{z} un vector de dimensión m cuyas componentes son v.a. gaussianas $N(0, 1/\sqrt{m})$ i.i.d. Independientemente se elige un vector \mathbf{u} unitario en $\ell^2(\mathbb{R}^m)$. Se tiene, para $0 < \varepsilon \leq 1$,

$$\mathbb{P}\{|\langle \mathbf{u}, \mathbf{z} \rangle| \geq \varepsilon\} \leq e^{-\frac{\varepsilon^2}{2}m}. \quad (2.76)$$

b) Sea \mathbf{w} un vector de dimensión m cuyas componentes son v.a. simétricas de Bernoulli $\{-1/\sqrt{m}, 1/\sqrt{m}\}$ i.i.d. Independientemente se elige un vector \mathbf{u} unitario en $\ell^2(\mathbb{R}^m)$. Se tiene, para $0 < \varepsilon \leq 1$,

$$\mathbb{P}\{|\langle \mathbf{u}, \mathbf{w} \rangle| \geq \varepsilon\} \leq 2e^{-\frac{\varepsilon^2}{2}m}. \quad (2.77)$$

Demostración. El producto interno $\langle \mathbf{u}, \mathbf{z} \rangle = \sum_{i=1}^m u_i z_i$ es una v.a. gaussiana (la suma de v.a. gaussianas es una v.a. gaussiana) con valor esperado $\mathbb{E}\{\langle \mathbf{u}, \mathbf{z} \rangle\} = \sum_{i=1}^m u_i \mathbb{E}\{z_i\} = 0$ y desviación normal

$$\left(\mathbb{E}\left\{|\langle \mathbf{u}, \mathbf{z} \rangle|^2\right\}\right)^{1/2} = \left(\sum_{i=1}^m u_i^2 \mathbb{E}\{z_i^2\} + \sum_{i=1}^m \sum_{j \neq i}^m u_i u_j \mathbb{E}\{z_i z_j\}\right)^{1/2} = \frac{1}{\sqrt{m}}$$

debido a que las v.a. z_i son independientes y que \mathbf{u} , unitario, es independiente de \mathbf{z} . Por tanto, como $|\langle \mathbf{u}, \mathbf{z} \rangle|$ es simétrica

$$\mathbb{P}\{|\langle \mathbf{u}, \mathbf{z} \rangle| \geq \varepsilon\} = 2\mathbb{P}\{\langle \mathbf{u}, \mathbf{z} \rangle \geq \varepsilon\} = 2 \frac{\sqrt{m}}{\sqrt{2\pi}} \int_{\varepsilon}^{\infty} e^{-\frac{1}{2}mx^2} dx = \sqrt{\frac{2}{\pi}} \int_{\varepsilon\sqrt{m}}^{\infty} e^{-\frac{y^2}{2}} dy \quad (2.78)$$

haciendo el cambio de variable $\sqrt{m}x = y$. Sea

$$I = \int_{\varepsilon\sqrt{m}}^{\infty} e^{-\frac{y^2}{2}} dy.$$

Se tiene

$$\begin{aligned} I^2 &= \left(\int_{\varepsilon\sqrt{m}}^{\infty} e^{-\frac{y^2}{2}} dy\right) \left(\int_{\varepsilon\sqrt{m}}^{\infty} e^{-\frac{x^2}{2}} dx\right) = \int_{\varepsilon\sqrt{m}}^{\infty} \left(\int_{\varepsilon\sqrt{m}}^{\infty} e^{-\frac{y^2+x^2}{2}} dy\right) dx \\ &\leq \int_{R_{\varepsilon,m}} \int e^{-\frac{y^2+x^2}{2}} dx dy, \end{aligned}$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

donde

$$R_{\varepsilon, m} = \{(x, y) = w \in \mathbb{R}^2 : x, y \geq 0, \|w\|_2 \geq \varepsilon\sqrt{2m}\}.$$

Pasando a coordenadas polares

$$I^2 \leq \int_0^{\pi/2} \int_{\varepsilon\sqrt{2m}}^{\infty} e^{-\frac{r^2}{2}} r dr d\theta = \frac{\pi}{2} \left[-e^{-\frac{r^2}{2}} \right]_{\varepsilon\sqrt{2m}}^{\infty} = \frac{\pi}{2} e^{-\frac{\varepsilon^2 m}{2}}.$$

Por tanto, $I \leq \sqrt{\frac{\pi}{2}} e^{-\frac{\varepsilon^2 m}{2}}$, que sustituyendo en (2.78) da (2.76).

b) En este caso puede aplicarse la desigualdad de Hoeffding (ver Teorema 4 en [75]) para obtener

$$\begin{aligned} \mathbb{P}\{|\langle \mathbf{u}, \mathbf{w} \rangle| \geq \varepsilon\} &= 2\mathbb{P}\left\{\sum_{i=1}^m u_i w_i \geq \varepsilon\right\} = 2\mathbb{P}\left\{\sum_{i=1}^m u_i w_i - \mathbb{E}\left\{\sum_{i=1}^m u_i w_i\right\} \geq \varepsilon\right\} \\ &= 2e^{-2\varepsilon^2 / \sum_{i=1}^m (u_i/\sqrt{m} - (-u_i/\sqrt{m}))^2} = 2e^{-\frac{\varepsilon^2 m}{2}}, \end{aligned}$$

ya que $\mathbb{E}\{\sum_{i=1}^m u_i w_i\} = \sum_{i=1}^m u_i \mathbb{E}\{w_i\} = 0$ por ser \mathbf{u} independiente de las v.a. w_i y $\|\mathbf{u}\|_2 = 1$. ■

En esta sección consideraremos matrices aleatorias $\Phi(\omega) \in \mathbb{R}^{m \times N}$ que satisfacen las siguientes condiciones, similares a las condiciones para matrices admisibles en [103]:

(M1) Las columnas de $\Phi(\omega)$ son estadísticamente independientes.

(M2) Para cada columna $\phi_j(\omega)$, $j = 1, \dots, N$, de $\Phi(\omega)$ se tiene $\mathbb{E}\left\{\|\phi_j(\omega)\|_{\ell^2(\mathbb{R}^m)}^2\right\} = 1$.

(M3) Sea $\mathbf{u} \in \mathbb{R}^m$ un vector con $\|\mathbf{u}\|_{\ell^2(\mathbb{R}^m)} \leq 1$. Si $\phi(\omega)$ es una columna de $\Phi(\omega)$ independiente de \mathbf{u} ,

$$\mathbb{P}\{|\langle \phi(\omega), \mathbf{u} \rangle| \geq \varepsilon\} \leq q_1 e^{-c_1 \varepsilon^2 m},$$

con q_1, c_1 constantes, $q_1 \geq 1$.

(M4) Para todo conjunto $\Gamma \subset \{1, \dots, N\}$ con $|\Gamma| \leq k < N$ y para todo $\mathbf{r} \in \text{span}(\Phi_\Gamma(\omega))$ se tiene

$$\mathbb{P}\left\{\|\Phi_\Gamma^*(\omega)\mathbf{r}\|_{\ell^2(\mathbb{R}^\Gamma)} \geq \frac{1}{2}\|\mathbf{r}\|_{\ell^2(\mathbb{R}^m)}\right\} \geq 1 - q_2 D^k e^{-c_2 m},$$

con q_2, D y c_2 constantes, $q_2, D > 1$.

2.3 Algoritmos ambiciosos en la percepción compresiva

Estas propiedades son satisfechas por las matrices aleatorias gaussianas y por las de Bernoulli descritas al comienzo de esta sección. La propiedad (M3) es el contenido del Lema 2.3.27 (observar que si $\|\mathbf{u}\|_2 \leq 1$ también se cumple el Lema 2.3.27 ya que si sustituimos \mathbf{u} por un vector unitario en su dirección la probabilidad aumenta) y la (M4) está probada en el Lema 5.1 de [6], en donde se da una demostración a partir de una desigualdad de concentración de la medida, como la dada en (2.75).

2.3.4.1 Identificación probabilística del soporte para algoritmos relajados

En esta sección mostramos que las matrices que satisfacen (M1), (M2), (M3) y (M4) permiten identificar índices del soporte de una señal k -dispersa con gran probabilidad. El resultado dado sigue la línea de argumentación expuesta en [103], con las modificaciones necesarias para acomodarla a la regla de selección relajada $\tilde{I}(\tilde{\alpha})$ dada en (2.51).

Teorema 2.3.28 *Sea $\Phi(\omega) \in \mathbb{R}^{m \times N}$ una matriz aleatoria que satisface las propiedades (M1), (M2), (M3) y (M4). Sea $\mathbf{x} \in \mathbb{R}^N$ con $\text{sop}(\mathbf{x}) = \Gamma^\sharp$ y $|\Gamma^\sharp| \leq k < N$. Sea $\mathbf{y} = \Phi\mathbf{x}$. Suponer que $\tilde{\alpha} \leq 1/2\sqrt{k}$ y dado l , $1 \leq l < N$,*

$$m \geq \max\left\{\frac{1}{c_1\tilde{\alpha}^2} \ln q_3 l(N-k), \frac{2k}{c_2} \ln D\right\}, \quad (2.79)$$

con $q_3 = q_1 + q_2$. los algoritmos RWMP, RWOMP y RWGP, con regla de selección $\tilde{I}_n(\tilde{\alpha})$ dada por (2.51), identifican elementos de Γ^\sharp en las l primeras iteraciones con probabilidad mayor o igual a

$$1 - q_3 l(N-k)e^{-c_1\tilde{\alpha}^2 m}.$$

Demostración. Para cualquiera de los algoritmos dados con la regla de selección $\tilde{I}_n(\tilde{\alpha})$, sea E_n el conjunto de matrices que identifican elementos de $\Gamma^\sharp = \text{sop}(\mathbf{x})$ en la iteración n , $n = 1, 2, 3, \dots$

Comencemos acotando $\mathbb{P}\{E_1\}$, es decir, la probabilidad de que una matriz identifique elementos de Γ^\sharp en la primera iteración. Para que una matriz aleatoria $\Phi(\omega)$ pertenezca a E_1 es suficiente que se cumplan las dos condiciones siguientes:

(A₁) Que $\tilde{I}_1(\tilde{\alpha}) \neq \emptyset$.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

(B_1) Que para $\mathbf{u}_0 = \mathbf{r}^0 / \|\mathbf{r}^0\|_{\ell^2(\mathbb{R}^m)}$ (recordar que $\mathbf{r}^0 = \mathbf{y}$) se cumpla

$$\frac{\|\Phi_{\Gamma^{\#c}}^*(\omega)\mathbf{r}^0\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\#c}})}}{\|\mathbf{r}^0\|_{\ell^2(\mathbb{R}^m)}} = \|\Phi_{\Gamma^{\#c}}^*(\omega)\mathbf{u}_0\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\#c}})} = \max_{j \in \Gamma^{\#c}} |\langle \mathbf{u}_0, \phi_j(\omega) \rangle| < \tilde{\alpha}.$$

Con abuso de notación, llamamos A_1 y B_1 los conjuntos de matrices que satisfacen (A_1) y (B_1) respectivamente. Se tiene

$$\mathbb{P}\{E_1\} \geq \mathbb{P}\{A_1 \cap B_1\} = \mathbb{P}\{A_1|B_1\} \mathbb{P}\{B_1\}. \quad (2.80)$$

Para estimar $\mathbb{P}\{B_1\}$ escribimos

$$\begin{aligned} \mathbb{P}\{B_1\} &= \mathbb{P}\left\{\max_{j \in \Gamma^{\#c}} |\langle \mathbf{u}_0, \phi_j(\omega) \rangle| < \tilde{\alpha}\right\} = \mathbb{P}\left\{\bigcap_{j \in \Gamma^{\#c}} \{|\langle \mathbf{u}_0, \phi_j(\omega) \rangle| < \tilde{\alpha}\}\right\} \\ &= \prod_{j \in \Gamma^{\#c}} \mathbb{P}\{|\langle \mathbf{u}_0, \phi_j(\omega) \rangle| < \tilde{\alpha}\} \end{aligned}$$

debido a la independencia de las columnas ϕ_j de $\Phi(\omega)$ expresada en (M1). Como $\mathbf{u}_0 = \mathbf{y} / \|\mathbf{y}\|_2$ es unitario e $\mathbf{y} = \Phi\mathbf{x} = \Phi_{\Gamma^{\#}}\mathbf{x}_{\Gamma^{\#}}$ sólo depende de las columnas ϕ_j con $j \in \Gamma^{\#}$, \mathbf{u}_0 es independiente de las columnas ϕ_j con $j \in \Gamma^{\#c}$, y se puede usar (M3) para obtener

$$\mathbb{P}\{B_1\} \geq \prod_{j \in \Gamma^{\#c}} (1 - q_1 e^{-c_1 \tilde{\alpha}^2 m}) = (1 - q_1 e^{-c_1 \tilde{\alpha}^2 m})^{N-k} \geq 1 - q_1 (N - k) e^{-c_1 \tilde{\alpha}^2 m}, \quad (2.81)$$

ya que $(1 - x)^n \geq 1 - nx$ si $n \geq 1$ y $x \leq 1$.

Por otro lado, condicionado a B_1 se tiene que $\Phi \in A_1 \Leftrightarrow \|\Phi_{\Gamma^{\#}}\mathbf{y}\|_\infty \geq \tilde{\alpha} \|\mathbf{y}\|_2$. Como $\tilde{\alpha} \leq \frac{1}{2\sqrt{k}}$,

$$\begin{aligned} \mathbb{P}\{A_1|B_1\} &= \mathbb{P}\left\{\|\Phi_{\Gamma^{\#}}^*(\omega)\mathbf{y}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\#}})} \geq \tilde{\alpha} \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}\right\} \\ &\geq \mathbb{P}\left\{\|\Phi_{\Gamma^{\#}}^*(\omega)\mathbf{y}\|_{\ell^2(\mathbb{R}^{\Gamma^{\#}})} \geq \tilde{\alpha} \sqrt{k} \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}\right\} \\ &\geq \mathbb{P}\left\{\|\Phi_{\Gamma^{\#}}^*(\omega)\mathbf{y}\|_{\ell^2(\mathbb{R}^{\Gamma^{\#}})} \geq \frac{1}{2} \|\mathbf{y}\|_{\ell^2(\mathbb{R}^m)}\right\}. \end{aligned}$$

Usando ahora (M4) deducimos (observar que $\mathbf{y} = \Phi\mathbf{x}$, por lo que $\mathbf{y} \in \text{span}(\Phi_{\Gamma^{\#}})$):

$$\mathbb{P}\{A_1|B_1\} \geq 1 - q_2 D^k e^{-c_2 m} \geq 1 - q_2 e^{-c_2 m/2} \quad (2.82)$$

tomando $-c_2 m + k \ln D \leq -c_2 m/2$, es decir, $m \geq \frac{2k}{c_2} \ln D$.

2.3 Algoritmos ambiciosos en la percepción compresiva

Sustituyendo (2.81) y (2.82) en (2.80) se obtiene

$$\begin{aligned}
 \mathbb{P}\{E_1\} &\geq (1 - q_1(N_k)e^{-c_1\tilde{\alpha}^2m})(1 - q_2e^{-c_2m/2}) \\
 &\geq 1 - q_1(N - k)e^{-c_1\tilde{\alpha}^2m} - q_2e^{-c_2m/2} \\
 &\geq 1 - q_3(N - k)e^{-c_1\tilde{\alpha}^2m},
 \end{aligned} \tag{2.83}$$

puesto que $\tilde{\alpha} < 1/2$, con $q_3 = q_1 + q_2$.

Hemos probado el resultado para $l = 1$. Supongamos que el resultado es cierto hasta la iteración $l - 1$, esto es

$$\mathbb{P}\left\{\bigcap_{n=1}^{l-1} E_n\right\} \geq 1 - q_3(l-1)(N-k)e^{-c_1\tilde{\alpha}^2m}. \tag{2.84}$$

Hasta la iteración l tenemos

$$\mathbb{P}\left\{\bigcap_{n=1}^l E_n\right\} = \mathbb{P}\left\{E_l \mid \bigcap_{n=1}^{l-1} E_n\right\} \mathbb{P}\left\{\bigcap_{n=1}^{l-1} E_n\right\}. \tag{2.85}$$

Para que una matriz aleatoria $\Phi(\omega)$ pertenezca a E_l es suficiente que se cumplan las dos condiciones siguientes:

(A_l) Que $\tilde{I}_l(\tilde{\alpha}) \neq \emptyset$.

(B_l) Que para $\mathbf{u}_{l-1} = \mathbf{r}^{l-1} / \|\mathbf{r}^{l-1}\|_2$ se cumpla

$$\frac{\|\Phi_{\Gamma^{\sharp c}}^*(\omega)\mathbf{r}^{l-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})}}{\|\mathbf{r}^{l-1}\|_{\ell^2(\mathbb{R}^m)}} = \|\Phi_{\Gamma^{\sharp c}}^*(\omega)\mathbf{u}_{l-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})} = \max_{j \in \Gamma^{\sharp c}} |\langle \mathbf{u}_{l-1}, \phi_j(\omega) \rangle| < \tilde{\alpha}.$$

Con abuso de notación, llamemos A_l y B_l a los conjuntos de matrices que satisfacen (A_l) y (B_l) respectivamente. Se tiene

$$\mathbb{P}\left\{E_l \mid \bigcap_{n=1}^{l-1} E_n\right\} \geq \mathbb{P}\left\{A_l \mid B_l \cap \left(\bigcap_{n=1}^{l-1} E_n\right)\right\} \mathbb{P}\left\{B_l \mid \bigcap_{n=1}^{l-1} E_n\right\}. \tag{2.86}$$

Con el mismo razonamiento que conduce a (2.81) se obtiene

$$\mathbb{P}\left\{B_l \mid \bigcap_{n=1}^{l-1} E_n\right\} \geq 1 - q_1(N-k)e^{-c_1\tilde{\alpha}^2m} \tag{2.87}$$

ya que como estamos condicionados a que el algoritmo haya elegido índices de Γ^\sharp hasta la iteración $l - 1$, se tiene que \mathbf{r}^{l-1} sólo depende de las columnas ϕ_j con $j \in \Gamma^\sharp$, por lo que \mathbf{u}_{l-1} es un vector unitario independiente de las columnas ϕ_j con $j \in \Gamma^{\sharp c}$.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

Condicionada a B_l , $\Phi \in A_l \Leftrightarrow \|\Phi_{\Gamma^\sharp}(\omega)\mathbf{r}^{l-1}\|_\infty \geq \tilde{\alpha} \|\mathbf{r}^{l-1}\|_2$. Como $\tilde{\alpha} \leq \frac{1}{2\sqrt{k}}$, se tiene

$$\begin{aligned} \mathbb{P} \left\{ A_l | B_l \cap \left(\bigcap_{n=1}^{l-1} E_n \right) \right\} &= \mathbb{P} \left\{ \left\{ \|\Phi_{\Gamma^\sharp}^*(\omega)\mathbf{r}^{l-1}\|_{\ell^\infty(\mathbb{R}^{\Gamma^\sharp})} \geq \tilde{\alpha} \|\mathbf{r}^{l-1}\|_{\ell^2(\mathbb{R}^m)} \right\} \mid \bigcap_{n=1}^{l-1} E_n \right\} \\ &\geq \mathbb{P} \left\{ \left\{ \|\Phi_{\Gamma^\sharp}^*(\omega)\mathbf{y}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})} \geq \tilde{\alpha}\sqrt{k} \|\mathbf{r}^{l-1}\|_{\ell^2(\mathbb{R}^m)} \right\} \mid \bigcap_{n=1}^{l-1} E_n \right\} \\ &\geq \mathbb{P} \left\{ \left\{ \|\Phi_{\Gamma^\sharp}^*(\omega)\mathbf{r}^{l-1}\|_{\ell^2(\mathbb{R}^{\Gamma^\sharp})} \geq \frac{1}{2} \|\mathbf{r}^{l-1}\|_{\ell^2(\mathbb{R}^m)} \right\} \mid \bigcap_{n=1}^{l-1} E_n \right\}. \end{aligned}$$

Ahora es posible usar (M4) ya que $\mathbf{r}^{l-1} \in \text{span}(\Phi_{\Gamma^\sharp})$ debido a la versión para $\tilde{I}(\tilde{\alpha})$ del Lema 2.3.3 (que se prueba fácilmente) y a que estamos calculando la probabilidad condicionada a que el algoritmo haya elegido índices de Γ^\sharp . Por tanto, con el mismo razonamiento que conduce a (2.82) se obtiene

$$\mathbb{P} \left\{ A_l | B_l \cap \left(\bigcap_{n=1}^{l-1} E_n \right) \right\} \geq 1 - q_2 e^{-c_2 m/2}. \quad (2.88)$$

De (2.87) y (2.88) se deduce

$$\mathbb{P} \left\{ E_l | \bigcap_{n=1}^{l-1} E_n \right\} \geq 1 - q_3(N-k)e^{-c_1 \tilde{\alpha}^2 m} \quad (2.89)$$

razonando como en la cadena de desigualdades que conduce a (2.83). Sustituimos (2.89) y (2.84) en (2.85) para obtener

$$\begin{aligned} \mathbb{P} \left\{ \bigcap_{n=1}^l E_n \right\} &\geq (1 - q_3(N-k)e^{-c_1 \tilde{\alpha}^2 m/2})(1 - q_3(l-1)(N-k)e^{-c_1 \tilde{\alpha}^2 m}) \\ &\geq 1 - q_3(N-k)e^{-c_1 \tilde{\alpha}^2 m/2} - q_3(l-1)(N-k)e^{-c_1 \tilde{\alpha}^2 m} \\ &= 1 - q_3 l(N-k)e^{-c_1 \tilde{\alpha}^2 m}, \end{aligned} \quad (2.90)$$

que es la estimación deseada. La condición (2.79) asegura que esta probabilidad es mayor que cero. \blacksquare

Nota 2.3.29 Tomando $\tilde{\alpha} = \frac{1}{2\sqrt{k}}$, la condición (2.79) se escribe

$$m \geq \max \left\{ \frac{4k}{c_1} \ln q_3 l(N-k), \frac{2k}{c_2} \ln D \right\}.$$

Por tanto, si se toma $m \geq Ck \ln l(N-k)$ con C suficientemente grande se tiene garantizada la desigualdad (2.79).

2.3 Algoritmos ambiciosos en la percepción compresiva

Para RWOMP con parámetro $\tilde{\alpha}$ cada iteración añade al menos un nuevo elemento, siempre que $\tilde{\mathcal{J}}_n(\tilde{\alpha}) \neq \emptyset$. En esta situación, RWOMP identifica todos los índices de Γ^\sharp en como mucho k iteraciones. Además, los algoritmos del tipo OMP recuperan cualquier vector k -disperso \mathbf{x} una vez conocido su soporte. Como $k(N - k) \leq N^2/4$ si $k < N$, tenemos el siguiente corolario.

Corolario 2.3.30 *Suponer las mismas hipótesis que en el Teorema 2.3.28 sustituyendo (2.79) por*

$$m \geq \max\left\{\frac{1}{c_1 \tilde{\alpha}^2} \ln q_3 \frac{N^2}{4}, \frac{2k}{c_2} \ln D\right\}. \quad (2.91)$$

Entonces, el algoritmo RWOMP recupera el vector k -disperso \mathbf{x} en las primeras k iteraciones con probabilidad mayor o igual a

$$1 - q_3 \frac{N^2}{4} e^{-c_1 \tilde{\alpha}^2 m}.$$

Nota 2.3.31 *Supongamos que queremos obtener el resultado del Teorema 2.3.28 con una probabilidad mayor o igual a $1 - \beta$ para un número $\beta \in (0, 1)$. Bastaría tomar*

$$1 - q_3 l(N - k) e^{-c_1 \tilde{\alpha}^2 m} \geq 1 - \beta.$$

Esto se consigue si

$$\beta e^{c_1 \tilde{\alpha}^2 m} \geq q_3 l(N - k),$$

por lo que basta tomar

$$m \geq \max\left\{\frac{2}{c_1 \tilde{\alpha}^2} \ln \frac{q_3 l(N - k)}{\beta}, \frac{2k}{c_2} \ln D\right\}. \quad (2.92)$$

Para asegurarnos que

$$\ln \frac{q_3 l(N - k)}{\beta} > 1$$

para todo $q_3 \geq 1$, $l < N$ y $k < N$, basta tomar $\beta \in (0, 1/e)$.

Un comentario análogo puede hacerse para el Corolario 2.3.30, y en este caso (2.91) debe ser sustituida por

$$m \geq \max\left\{\frac{1}{c_1 \tilde{\alpha}^2} \ln \frac{q_3 N^2}{4\beta}, \frac{2k}{c_2} \ln D\right\} \quad (2.93)$$

para que RWOMP recupere una señal k -dispersa \mathbf{x} en las k primeras iteraciones con probabilidad mayor o igual a $1 - \beta$.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

2.3.4.2 Identificación probabilística del soporte con la regla de selección $\mathcal{J}(\alpha)$.

En esta sección estudiaremos el comportamiento de las matrices que satisfacen (M1), (M2), (M3) y (M4) con respecto a los algoritmos con regla de selección

$$\mathcal{J}_n(\alpha) = \{i : |\langle \phi_i, \mathbf{r}^{n-1} \rangle| \geq \alpha \|\Phi^* \mathbf{r}^{n-1}\|_{\ell^\infty(\mathbb{R}^N)}\}, \quad (2.94)$$

$0 < \alpha \leq 1$. Una ventaja de $\mathcal{J}_n(\alpha)$ con respecto a la regla de selección $\tilde{\mathcal{J}}_n(\tilde{\alpha})$ dada en (2.51) es que $\mathcal{J}_n(\alpha) \neq \emptyset$ para todo $\alpha \in (0, 1)$ y para todo $n = 1, 2, \dots$

Pero tiene una desventaja. Si queremos probar que $\mathcal{J}_n(\alpha) \subset \Gamma^\sharp = \text{sop}(\mathbf{x})$ necesitamos que se cumpla

$$\frac{\|\Phi_{\Gamma^{\sharp c}}^* \mathbf{r}^{n-1}\|_{\infty}}{\|\Phi^* \mathbf{r}^{n-1}\|_{\infty}} < \alpha.$$

Si escribiéramos $\mathbf{u}_{n-1} = \mathbf{r}^{n-1} / \|\Phi^* \mathbf{r}^{n-1}\|_{\infty}$ como en la demostración del Teorema 2.3.28 no tendríamos $\|\mathbf{u}_{n-1}\|_2 \leq 1$, y no podríamos usar (M3). La propiedad (M4) salva esta situación como veremos en la demostración del siguiente resultado.

Teorema 2.3.32 *Elegir $\Phi(\omega) \in \mathbb{R}^{m \times N}$ una matriz aleatoria que satisface las propiedades (M1), (M2), (M3) y (M4). Sea $\mathbf{x} \in \mathbb{R}^N$ con $\text{sop}(\mathbf{x}) = \Gamma^\sharp$ y $|\Gamma^\sharp| \leq k < N$. Sea $\mathbf{y} = \Phi \mathbf{x}$. Suponer que*

$$m \geq \max\left\{\frac{4k}{c_1 \alpha^2} \ln q_3 l(N - k), \frac{2k}{c_2} \ln D\right\}, \quad (2.95)$$

con $q_3 = q_1 + q_2$. Los algoritmos WMP, WOMP y WGP, con regla de selección $\mathcal{J}_n(\alpha)$ dada en 2.94, identifican elementos de Γ^\sharp en las l primeras iteraciones con probabilidad mayor o igual a

$$1 - q_3 l(N - k) e^{-c_1 \frac{\alpha^2}{4k} m}.$$

Demostración. Para cualquiera de los algoritmos dados con la regla de selección $\mathcal{J}_n(\alpha)$, sea E_n el conjunto de matrices que identifican elementos de $\Gamma^\sharp = \text{sop}(\mathbf{x})$ en la iteración n , $n = 1, 2, \dots$

Comencemos acotando $\mathbb{P}\{E_1\}$, es decir, la probabilidad de que una matriz identifique elementos de Γ^\sharp en la primera iteración. Para que una matriz aleatoria $\Phi(\omega)$ pertenezca a E_1 es suficiente que se cumpla

$$\frac{\|\Phi_{\Gamma^{\sharp c}}^*(\omega) \mathbf{r}^0\|_{\ell^\infty(\mathbb{R}^{\Gamma^{\sharp c}})}}{\|\Phi^*(\omega) \mathbf{r}^0\|_{\ell^\infty(\mathbb{R}^N)}} < \alpha, \quad (\mathbf{r}^0 = \mathbf{y}). \quad (2.96)$$

2.3 Algoritmos ambiciosos en la percepción compresiva

Sea $\mathbf{u}_0 = \mathbf{y}/2 \|\Phi_{\Gamma^\#}^* \mathbf{y}\|_2$; se tiene

$$\begin{aligned} \frac{\|\Phi_{\Gamma^\#}^*(\omega) \mathbf{y}\|_\infty}{\|\Phi_{\Gamma^\#}^*(\omega) \mathbf{y}\|_2} &\leq \frac{\sqrt{k} \|\Phi_{\Gamma^\#}^*(\omega) \mathbf{y}\|_\infty}{\|\Phi_{\Gamma^\#}^*(\omega) \mathbf{y}\|_2} = 2\sqrt{k} \|\Phi_{\Gamma^\#}^*(\omega) \mathbf{u}_0\|_\infty \\ &= 2\sqrt{k} \sup_{j \in \Gamma^\#} |\langle \phi_j(\omega), \mathbf{u}_0 \rangle|. \end{aligned} \quad (2.97)$$

Sea A_1 el conjunto de matrices que satisfacen (M1), (M2), (M3) y (M4) tal que

$$\sup_{j \in \Gamma^\#} |\langle \phi_j(\omega), \mathbf{u}_0 \rangle| < \frac{\alpha}{2\sqrt{k}}. \quad (2.98)$$

De (2.97) deducimos que si $\Phi(\omega) \in A_1$, entonces $\Phi(\omega)$ satisface (2.96) y se tiene $\Phi(\omega) \in E_1$. Por tanto,

$$\mathbb{P}\{E_1\} \geq \mathbb{P}\{A_1\}. \quad (2.99)$$

Sea B_1 el conjunto de matrices $\Phi(\omega)$ que satisfacen (M1), (M2), (M3) y (M4) tal que

$$\|\Phi_{\Gamma^\#}^*(\omega) \mathbf{y}\|_2 \geq \frac{1}{2} \|\mathbf{y}\|_2. \quad (2.100)$$

De (2.99) deducimos

$$\mathbb{P}\{E_1\} \geq \mathbb{P}\{A_1\} \geq \mathbb{P}\{A_1 \cap B_1\} = \mathbb{P}\{A_1|B_1\} \mathbb{P}\{B_1\}. \quad (2.101)$$

Como $\mathbf{y} = \Phi \mathbf{x} = \Phi_{\Gamma^\#} \mathbf{x}_{\Gamma^\#} \in \text{span}(\Phi_{\Gamma^\#})$ de (M4) deducimos

$$\mathbb{P}\{B_1\} \geq 1 - q_2 D^k e^{-c_2 m} \geq 1 - q_2 e^{-c_2 m/2} \quad (2.102)$$

tomando $-c_2 m + k \ln D \leq -c_2 m/2$, es decir, $m \geq \frac{2k}{c_2} \ln D$. Condicionado a B_1 el vector $\mathbf{u}_0 = \mathbf{y}/2 \|\Phi_{\Gamma^\#}^* \mathbf{y}\|_2$ satisface

$$\|\mathbf{u}_0\|_2 = \frac{1}{2} \frac{\|\mathbf{y}\|_2}{\|\Phi_{\Gamma^\#}^* \mathbf{y}\|_2} \leq 1$$

por lo que para acotar

$$\mathbb{P}\{A_1|B_1\} = \mathbb{P}\left\{ \sup_{j \in \Gamma^\#} |\langle \phi_j(\omega), \mathbf{u}_0 \rangle| < \frac{\alpha}{2\sqrt{k}} | B_1 \right\}$$

usamos (M3) y obtenemos

$$\begin{aligned} \mathbb{P}\{A_1|B_1\} &= \mathbb{P}\left\{ \bigcap_{j \in \Gamma^\#} \{|\langle \phi_j(\omega), \mathbf{u}_0 \rangle| < \frac{\alpha}{2\sqrt{k}}\} | B_1 \right\} \\ &= \prod_{j \in \Gamma^\#} \mathbb{P}\left\{ |\langle \phi_j(\omega), \mathbf{u}_0 \rangle| < \frac{\alpha}{2\sqrt{k}} | B_1 \right\} \\ &\geq (1 - q_1 e^{-c_1 \frac{\alpha^2}{4k} m})^{N-k} \end{aligned}$$

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

debido a la independencia de las columnas ϕ_j de $\Phi(\omega)$ expresada en (M1) y a que $\mathbf{u}_0 = \mathbf{y}/2 \|\Phi_{\Gamma^\sharp}^* \mathbf{y}\|_2$ sólo depende de las columnas ϕ_j con $j \in \Gamma^\sharp$, y es por tanto independiente de las columnas ϕ_j con $j \in \Gamma^{\sharp c}$.

Puesto que $(1-x)^n \geq 1-nx$ si $n \geq 1$, $x \leq 1$, podemos escribir

$$\mathbb{P}\{A_1|B_1\} \geq 1 - q_1(N-k)e^{-c_1 \frac{\alpha^2}{4k} m}. \quad (2.103)$$

Sustituyendo (2.103) y (2.102) en (2.101) se obtiene

$$\begin{aligned} \mathbb{P}\{E_1\} &\geq (1 - q_1(N-k)e^{-c_1 \frac{\alpha^2}{4k} m})(1 - q_2e^{-c_1 \frac{m}{2}}) \\ &\geq 1 - q_1(N-k)e^{-c_1 \frac{\alpha^2}{4k} m} - q_2e^{-c_1 \frac{m}{2}} \\ &\geq 1 - q_3(N-k)e^{-c_1 \frac{\alpha^2}{4k} m} \end{aligned} \quad (2.104)$$

con $q_3 = q_1 + q_2$, ya que $\frac{\alpha^2}{4k} \leq 1/4 < 1/2$. Esto prueba el resultado para $l = 1$.

Supongamos que el resultado es cierto hasta la iteración $l-1$, esto es

$$\mathbb{P}\left\{\bigcap_{n=1}^{l-1} E_n\right\} \geq 1 - q_3(l-1)(N-k)e^{-c_1 \frac{\alpha^2}{4k} m}. \quad (2.105)$$

Hasta la iteración l tenemos

$$\mathbb{P}\left\{\bigcap_{n=1}^l E_n\right\} = \mathbb{P}\left\{E_l | \bigcap_{n=1}^{l-1} E_n\right\} \mathbb{P}\left\{\bigcap_{n=1}^{l-1} E_n\right\}. \quad (2.106)$$

Escribir $\mathbf{u}_{l-1} = \mathbf{r}^{l-1}/2 \|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{l-1}\|_2$. Sea A_l el conjunto de matrices que satisfacen (M1), (M2), (M3) y (M4) tal que

$$\sup_{j \in \Gamma^{\sharp c}} |\langle \phi_j(\omega), \mathbf{u}_{l-1} \rangle| < \frac{\alpha}{2\sqrt{k}}. \quad (2.107)$$

Si $\Phi \in A_l$ se tiene

$$\begin{aligned} \frac{\|\Phi_{\Gamma^{\sharp c}}^*(\omega) \mathbf{r}^{l-1}\|_\infty}{\|\Phi_{\Gamma^\sharp}^*(\omega) \mathbf{r}^{l-1}\|_\infty} &\leq \frac{\sqrt{k} \|\Phi_{\Gamma^{\sharp c}}^*(\omega) \mathbf{r}^{l-1}\|_\infty}{\|\Phi_{\Gamma^\sharp}^*(\omega) \mathbf{r}^{l-1}\|_2} = 2\sqrt{k} \|\Phi_{\Gamma^{\sharp c}}^*(\omega) \mathbf{u}_{l-1}\|_\infty \\ &= 2\sqrt{k} \sup_{j \in \Gamma^{\sharp c}} |\phi_j(\omega), \mathbf{u}_{l-1}| < \alpha, \end{aligned}$$

y esto es suficiente para asegurar que $\mathcal{J}_n(\alpha) \subset \Gamma^\sharp$, es decir, $\Phi \in E_l$. Por tanto,

$$\mathbb{P}\left\{E_l | \bigcap_{n=1}^{l-1} E_n\right\} \geq \mathbb{P}\left\{A_l | \bigcap_{n=1}^{l-1} E_n\right\}. \quad (2.108)$$

Sea B_l el conjunto de matrices $\Phi(\omega)$ que satisfacen (M1), (M2), (M3) y (M4) tal que

$$\|\Phi_{\Gamma^\sharp}^*(\omega) \mathbf{r}^{n-1}\|_2 \geq \frac{1}{2} \|\mathbf{r}^{n-1}\|_2. \quad (2.109)$$

2.3 Algoritmos ambiciosos en la percepción compresiva

De (2.108) deducimos

$$\mathbb{P} \left\{ E_l | \cap_{n=1}^{l-1} E_n \right\} \geq \mathbb{P} \left\{ A_l \cap B_l | \cap_{n=1}^{l-1} E_n \right\} = \mathbb{P} \left\{ A_l | B_l \cap (\cap_{n=1}^{l-1} E_n) \right\} \mathbb{P} \left\{ B_l | \cap_{n=1}^{l-1} E_n \right\}. \quad (2.110)$$

Condicionado a B_l ,

$$\|\mathbf{u}_{l-1}\|_2 = \frac{1}{2} \frac{\|\mathbf{r}^{l-1}\|_2}{\|\Phi_{\Gamma^\sharp}^* \mathbf{r}^{l-1}\|_2} \leq 1$$

debido a (2.109). Por (M1) las columnas ϕ_j de $\Phi(\omega)$ son independientes entre sí. Además, condicionado a $\cap_{n=1}^{l-1} E_n$, el vector \mathbf{r}^{n-1} sólo depende de las columnas ϕ_j con $j \in \Gamma^\sharp$ ya que $\mathcal{J}_n(\alpha) \subset \Gamma^\sharp$, $n = 1, 2, \dots, l-1$. Por tanto, \mathbf{u}_{l-1} es independiente de las columnas ϕ_j con $j \in \Gamma^{\sharp c}$ y podemos usar (M3). Se tiene, entonces

$$\begin{aligned} \mathbb{P} \left\{ A_l | B_l \cap (\cap_{n=1}^{l-1} E_n) \right\} &= \mathbb{P} \left\{ \sup_{j \in \Gamma^{\sharp c}} |\langle \phi_j(\omega), \mathbf{u}_{l-1} \rangle| < \frac{\alpha}{2\sqrt{k}} | B_l \cap (\cap_{n=1}^{l-1} E_n) \right\} \\ &= \prod_{j \in \Gamma^{\sharp c}} \mathbb{P} \left\{ |\langle \phi_j(\omega), \mathbf{u}_{l-1} \rangle| < \frac{\alpha}{2\sqrt{k}} | B_l \cap (\cap_{n=1}^{l-1} E_n) \right\} \\ &\geq (1 - q_1 e^{-c_1 \frac{\alpha^2}{4k} m})^{N-k} \geq 1 - q_1 (N-k) e^{-c_1 \frac{\alpha^2}{4k} m}. \end{aligned} \quad (2.111)$$

Condicionado a $\cap_{n=1}^{l-1} E_n$, el vector $\mathbf{r}^{n-1} \in \text{span}(\Phi_{\Gamma^\sharp})$ por el Lema 2.3.3. Podemos usar (M4) para obtener

$$\mathbb{P} \left\{ B_l | \cap_{n=1}^{l-1} E_n \right\} \geq 1 - q_1 D^k e^{-c_2 m} \geq 1 - q_1 e^{-c_2 m/2} \quad (2.112)$$

tomando $-c_2 m + k \ln D \leq -c_2 \frac{m}{2}$, es decir, $m \geq \frac{2k}{c_2} \ln D$.

Sustituimos (2.110) y (2.112) en (2.109) y procedemos como en los cálculos que llevan a (2.104) para obtener

$$\mathbb{P} \left\{ E_l | \cap_{n=1}^{l-1} E_n \right\} \geq 1 - q_3 (N-k) e^{-c_1 \frac{\alpha^2}{4k} m}, \quad (2.113)$$

con $q_3 = q_1 + q_2$. Sustituimos (2.113) y (2.105) en (2.106) para obtener

$$\begin{aligned} \mathbb{P} \left\{ \cap_{n=1}^l E_n \right\} &\geq (1 - q_3 (N-k) e^{-c_1 \frac{\alpha^2}{4k} m}) (1 - q_3 (l-1) (N-k) e^{-c_1 \frac{\alpha^2}{4k} m}) \\ &\geq 1 - q_3 (N-k) e^{-c_1 \frac{\alpha^2}{4k} m} - q_3 (l-1) (N-k) e^{-c_1 \frac{\alpha^2}{4k} m} \\ &= 1 - q_3 l (N-k) e^{-c_1 \frac{\alpha^2}{4k} m}, \end{aligned}$$

que era lo que queríamos demostrar. Para que esta última probabilidad sea mayor a 0 debemos tomar

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

$$m \geq \frac{4k}{c_1 \alpha^2} \ln q_3 l (N - k).$$

■

Se pueden hacer ahora comentarios similares a los hechos al finalizar la demostración del Teorema 2.3.28. Destacamos el siguiente corolario, que en el caso $\alpha = 1$ da el Teorema 6 de [103].

Corolario 2.3.33 *Elegir las mismas condiciones que en el Teorema 2.3.32, sustituyendo (2.95) por*

$$m \geq C \frac{k}{\alpha^2} \ln(q_3 N^2 / 4) \quad (2.114)$$

con C suficientemente grande. Entonces, el algoritmo WOMP recupera el vector k -disperso \mathbf{x} en las k primeras iteraciones con probabilidad mayor o igual a

$$1 - q_3 \frac{N^2}{4} e^{-c_1 \frac{\alpha^2}{4k} m}.$$

2.4 Experimentos

En esta sección presentamos experimentos de los problemas de recuperación dispersa y aproximación no dispersa para las versiones ortogonales y del gradiente de los algoritmos anteriormente estudiados con la regla de selección $\mathcal{J} = \mathcal{J}(\tilde{\alpha})$ dada en (2.51) y los comparamos con los resultados obtenidos con la regla de selección usual $\mathcal{J} = \mathcal{J}(\alpha)$ dada en (2.47). En cada conjunto de experimentos una matriz gaussiana de orden $m \times N$ fue creada. Para algunas señales los resultados se muestran como porcentaje de elementos recuperados y para otras se usa la razón de la energía de la señal original $\mathbf{x} \in \mathbb{R}^N$ y la energía de la diferencia entre la señal y la aproximación dada por

$$SNR = 10 \log_{10} \left(\frac{\|\mathbf{x}\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \right).$$

La Figura 2.2 muestra el el porcentaje de elementos recuperados con el algoritmo RWOMP con parámetro $\alpha = 0.125$. Las matrices gaussianas generadas tienen $N = 256$ y $m = 10\ell$, $\ell = 1, 2, \dots, 25$. Los niveles de dispersión se han escogido como $k = 4, 12, 20, 28, 36$, (cada gráfica corresponde a cada uno de ellos). Por cada par (m, k) , 200 experimentos se realizaron para diferentes señales. Los resultados pueden compararse

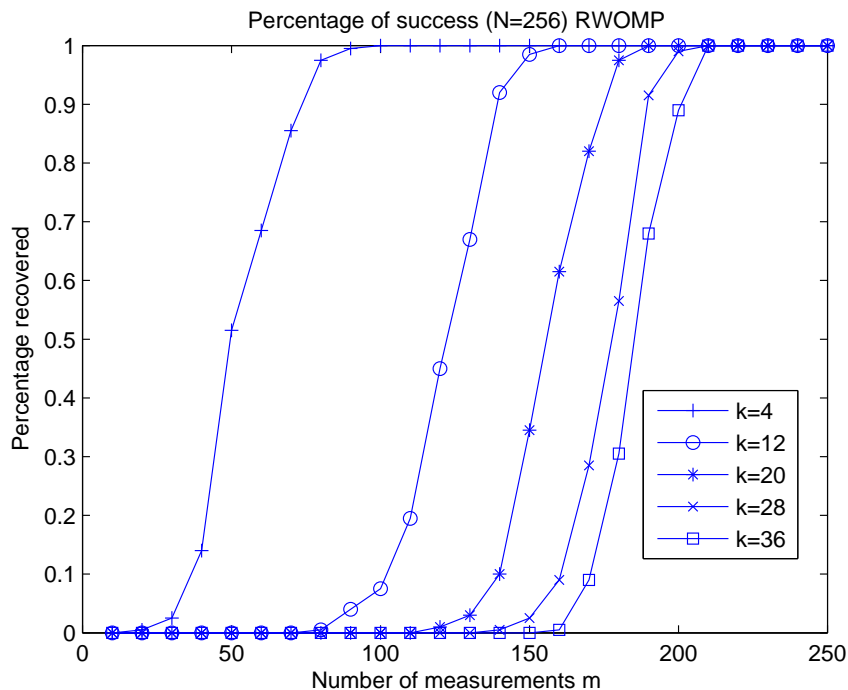


Figure 2.2: Reconstrucción de señales dispersas con RWOMP, $\alpha = 0.125$. Por cada par (m, k) , 200 experimentos se generaron con una matriz gaussiana de orden $m \times 256$. Hasta k iteraciones se permitieron.

con aquellos obtenidos en la Figura 1 de [103] para OMP ($\alpha = 1$). Los parámetros N, k y m toman los mismos valores, pero 1000 experimentos por cada conjunto se realizaron en [103]. Los resultados para RWOMP son mejores que para OMP en [103] para $k = 20, 28, 36$.

A continuación, por propósitos computacionales, introducimos una modificación menor del algoritmo RWOMP, llamada algoritmo k -RWOMP. En cada iteración en RWOMP guardamos los k -mayores elementos de la aproximación ortogonal \mathbf{x}^n . La Figura 2.3 muestra los resultados de aplicar k -RWOMP con los mismos parámetros de los experimentos descritos para la Figura 2.2. Como puede verse, en la Figura 2.3 la reconstrucción exacta se alcanza con menores valores de m .

Los siguientes experimentos se hacen en imágenes. Tomamos una imagen de 64×64 píxeles, que son parte de Lena. Dos descomposiciones de la transformada ondícula se realizan usando las ondículas de *Daubechies 5*. La precepción compresiva se hizo sólo

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

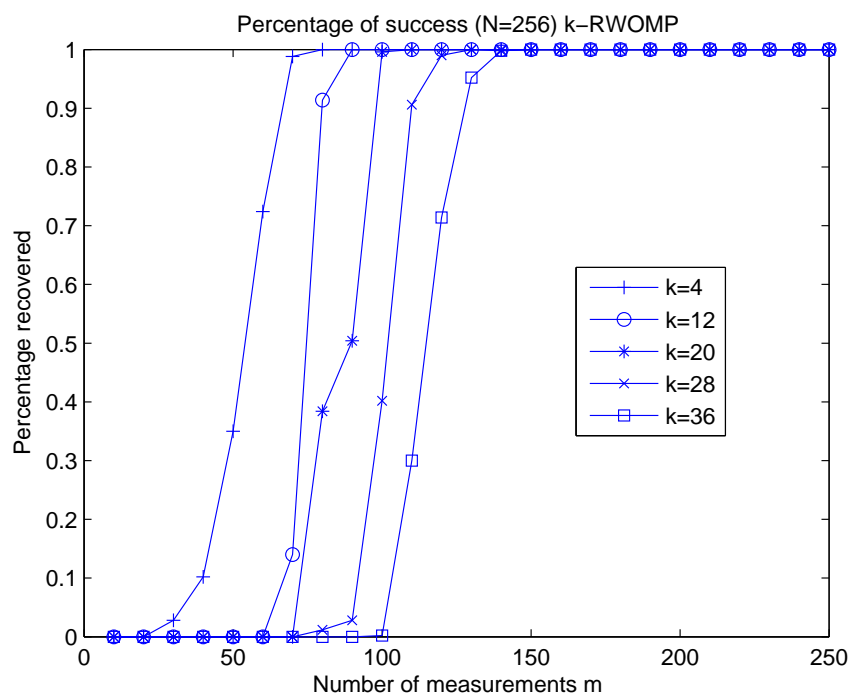


Figure 2.3: Reconstrucción de señales dispersas con k -RWOMP, con los mismos parámetros que para la Figura 2.2. Los resultados son mejores que en la Figura 2.2 para RWOMP. Hasta k iteraciones son permitidas.

en los coeficientes de los detalles. La dispersión k se calcula como la parte entera de 5% de L , donde L es el número de coeficientes verticales, horizontales o diagonales en cada nivel de descomposición. El número de mediciones en cada nivel es la parte entera de $k \log_2(L/k)$ para los coeficientes verticales, horizontales y diagonales. Los algoritmos k-RWOMP y k-RWGP (aquellos con regla de selección $\tilde{\mathcal{J}}(\alpha)$ como se define en (2.51)) se realizaron con $\tilde{\alpha} = 0.125$ y $\tilde{\alpha} = 0.15$, respectivamente. El algoritmo WGP (con regla de selección $\mathcal{J}(\alpha)$ dada en (2.47)) se realizó con $\alpha = 0.8$. Los datos relevantes se dan en la Figura 3. Cuando la dispersión de la aproximación se eleva a 10% de L el algoritmo CoSaMP invierte varios minutos para detenerse, para el resto de algoritmos el tiempo se incrementa a penas perceptiblemente. Por cada algoritmo, los tiempos de realización y los SNR 's difieren para diferentes α y, aún más, pueden diferir con la misma α debido a la aleatoriedad de Φ . También, visualmente hay una mejora para los algoritmos k-Relajados Débiles para los cuales hay menos elementos espurios en las regiones más suaves de la imagen. A juzgar por el SNR, el algoritmo k-RWGP da la mejor aproximación.

2. ALGORITMOS AMBICIOSOS PARA LA PERCEPCIÓN COMPRESIVA

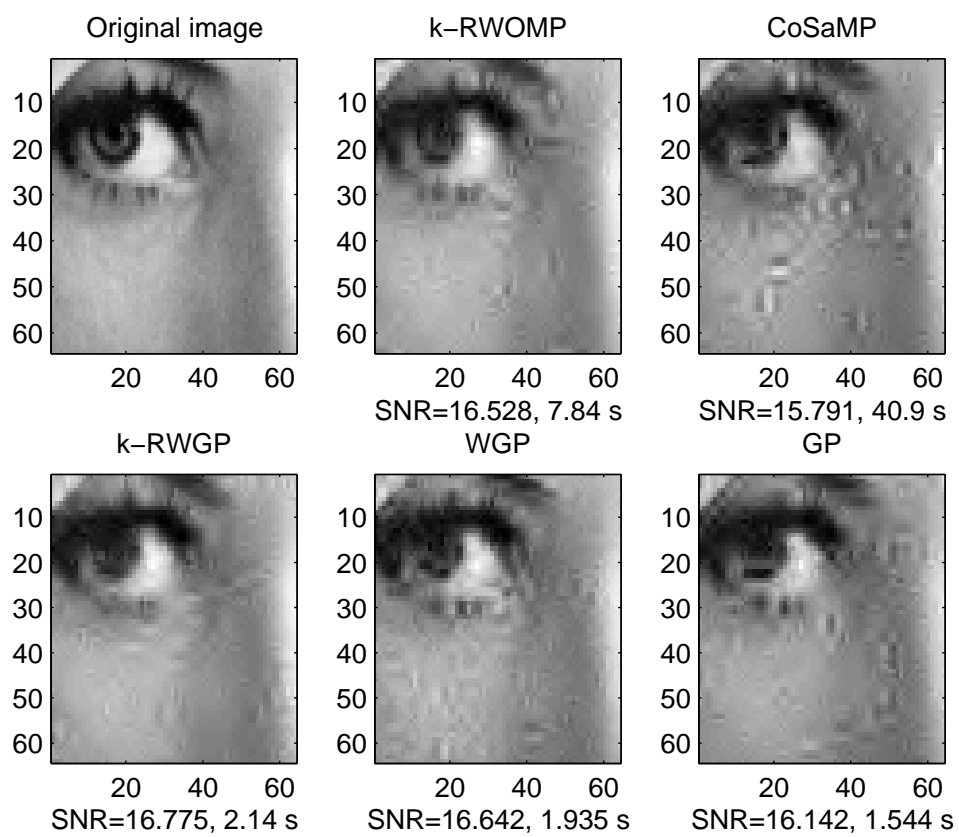


Figure 2.4: k-RWOMP and k-RWGP vs CoSaMP, SWGP and GP.

3

Aproximación no lineal restringida en espacios de sucesiones

*... Al primer soñador le fue deparada en la noche la visión
del palacio y lo construyó;
al segundo, que no supo del sueño del anterior, el poema sobre el palacio...
su primera manifestación fue el palacio;
la segunda el poema.
Quien los hubiera comparado habría visto que eran esencialmente iguales.*

Jorge Luis Borges, *El sueño de Coleridge*.

3.1 Introducción

Dos problemas fundamentales en el procesamiento de señales son el de la compresión y el de la reducción de ruido. El tratamiento de ambos problemas se beneficia de la teoría de la aproximación. En esta teoría se trata de aproximar una función o señal posiblemente complicada, llamada *función objetivo*, por funciones más sencillas y más fáciles de computar, llamadas *aproximantes*.

Aplicaciones de las técnicas aquí usadas son: análisis al procesamiento de señales, estimación estadística y solución numérica de ecuaciones diferenciales.

Uno de los primeros métodos de aproximación usaba como aproximantes polinomios de grado N , tanto algebraicos como trigonométricos. En este caso, la aproximación se hacía a partir de espacios lineales de dimensión finita, es decir, sólo se tomaban en

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

cuenta las primeras N potencias de t , $\{1, t, t^2, \dots, t^{N-1}\}$, o las primeras N potencias de $e^{2\pi it}$, $\{1, e^{2\pi it}, e^{2\pi i2t}, \dots, e^{2\pi i(N-1)t}\}$, por ejemplo. A esto se le llama aproximación lineal. Denotamos L_N al subespacio con N índices “fijos” de los aproximantes. Entonces, una propiedad característica de la **aproximación lineal** es que para dos funciones $f, g \in L_N$ se tiene que $f + g \in L_N$.

Jackson [63] y Bernstein [10] probaron que una función f continua y periódica tiene un *orden de aproximación* $O(N^{-\alpha})$, $0 < \alpha \leq 1$, cuando se aproxima por polinomios trigonométricos de grado menor o igual a N en la norma uniforme, si y sólo si, $f \in \text{Lip } \alpha$, el espacio de Lipschitz de orden α . A continuación hacemos más explícitos estos resultados.

Sea $C([0, 1))$ el espacio de funciones continuas y periódicas. Sea \mathcal{T}_N el conjunto de los polinomios trigonométricos de orden menor o igual a N . Sea $0 < \alpha \leq 1$. El espacio de Lipschitz $\text{Lip } \alpha$ que consiste de todas las funciones $f \in C([0, 1))$ para las que existe $M > 0$ tal que

$$|f(x) - f(y)| \leq M |x - y|^\alpha, \quad \text{para todo } x, y \in \mathbb{T}.$$

Denotamos por

$$\sigma(f, N)_{L^\infty} := \inf_{g \in \mathcal{T}_N} \|f - g\|_{L^\infty}$$

el error de aproximación de $f \in C(\mathbb{T})$ por los elementos de \mathcal{T}_N . Se tienen entonces dos resultados fundamentales para la teoría de la aproximación.

Teorema 3.1.1 (Jackson [63]) *Sea $f \in C(\mathbb{T})$. Entonces, para toda $f \in \text{Lip } \alpha$, $0 < \alpha < 1$, se tiene*

$$\sigma_N(f, \mathcal{T}_N)_{L^\infty} \leq CN^{-\alpha}.$$

Teorema 3.1.2 (Bernstein [10]) *Sean $f \in C(\mathbb{T})$ y $\alpha \in (0, 1)$ tal que*

$$\sigma_N(f, \mathcal{T}_N)_{L^\infty} \leq CN^{-\alpha},$$

para algún $C > 0$ y para todo $N = 1, 2, 3, \dots$. Entonces $f \in \text{Lip } \alpha$.

Este tipo de desigualdades son fundamentales y permean gran parte de la teoría de la aproximación (lineal, no lineal y no lineal restringida).

En el caso de que usemos polinomios a trozos, los aproximantes no deberían restringirse a una partición fija sino que puede depender de la función objetivo. La idea

es simple: se debe usar una partición más fina donde la función objetivo no sea muy suave y una partición menos fina donde sí lo sea. En el caso de funciones en $L^2(\mathbb{T})$ la descomposición en serie de Fourier de una función que pertenece a una clase de funciones que se compone de altas frecuencias no dará mucha información hasta que N sea muy grande no importando que esas altas frecuencias sean pocas, incluso una sola. Así surge la aproximación no lineal. Denotamos Σ_N a la clase de todos los conjuntos de, a lo mucho, N índices no “fijos”. Entonces, una propiedad característica de la **aproximación no lineal** es que para dos funciones $f, g \in \Sigma_N$ se tiene que $f + g \in \Sigma_{K(N)}$, donde K es una función no decreciente tal que $N \leq K(N)$. En este capítulo $K(N) = 2N$ ya que nos enfocaremos en aproximantes que surgen de una base incondicional. Otras funciones K surgen en la aproximación por *splines* o polinomios a trozos cuando los *knots* o (los nudos de) las particiones son libres así como el grado de los polinomios en cada trozo.

3.2 Aproximación no lineal.

Revisamos brevemente algunos resultados clásicos e introducimos algunos algoritmos y conceptos que usaremos más adelante.

3.2.1 Algunos resultados en la ANL

Sea \mathbb{X} un espacio cuasi normado y $\mathcal{B} = \{e_j\}_{j=1}^\infty$ una base de \mathbb{X} . Definimos

$$\Sigma_N = \{g = \sum_{j \in \Gamma} a_j e_j : \Gamma \subset N, |\Gamma| \leq N\}.$$

Para $f \in \mathbb{X}$, el **error de la aproximación** con N términos se define como

$$\sigma(f, N)_{\mathbb{X}} := \inf_{g \in \Sigma_N} \|f - g\|_{\mathbb{X}}. \quad (3.1)$$

Una cuestión importante es encontrar el espacio de funciones para el cual, dado un $\alpha > 0$, se tiene que el error de aproximación tiene una velocidad de decaimiento tal que

$$\sigma(f, N)_{\mathbb{X}} \leq CN^{-\alpha}, \quad N = 1, 2, 3, \dots, \quad (3.2)$$

para algún $C > 0$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

De manera más general, se definen los **espacios de aproximación** con N términos (con aproximantes en la base $\mathcal{B} \subset \mathbb{X}$ y cuyo error es medido en la cuasi-norma \mathbb{X}) como el conjunto de todas las $f \in \mathbb{X}$ tales que

$$\|f\|_{\mathcal{A}_\mu^\xi(\mathbb{X}, \mathcal{B})} := \left(\sum_{N=1}^{\infty} (N^\xi \sigma(f, N)_{\mathbb{X}})^\mu \frac{1}{N} \right)^{1/\mu} < \infty, \quad \xi, \mu \in (0, \infty), \quad (3.3)$$

y $\|f\|_{\mathcal{A}_\infty^\xi(\mathbb{X}, \mathcal{B})} := \sup_{t>0} t^\xi \sigma_\nu(t, \mathbf{s})$ cuando $\xi \in (0, \infty)$ y $\mu = \infty$.

Debe mencionarse que existen otros tipos de aproximaciones como la aproximación racional, destacando el trabajo de Peller en [87] que caracterizó las clases de aproximación racional en la métrica BMO (*bounded mean oscillations*). Además, en el proceso, Peller caracterizó los espacios de interpolación entre BMO y el espacio de Besov $B_1^1(L^1)$. En el caso de los *splines* fue Petrushev en [88] quién caracterizó los espacios de aproximación no lineal con *splines* como espacios de Besov usando el K -funcional.

DeVore y Popov [33] generalizan este último resultado caracterizando los espacios de aproximación (sobre espacios cuasi-normados para un esquema de aproximación no lineal) como espacios de interpolación. Dicho resultado se prueba a partir de las desigualdades tipo Jackson y Bernstein y el K -funcional (usado en el método de la interpolación real). Más específicamente, sea $(\mathbb{X}, \|\cdot\|_{\mathbb{X}})$ un espacio cuasi-normado (de hecho, en [33] se considera más generalmente un grupo Abelian cuasi-normado), para $r > 0$ sea \mathbb{Y}_r un subespacio de \mathbb{X} . Se tiene la desigualdad tipo Jackson de orden r si existe un $C > 0$ tal que para todo $N = 1, 2, 3, \dots$,

$$\sigma(f, N)_{\mathbb{X}} \leq CN^{-r} \|f\|_{\mathbb{Y}_r}, \quad f \in \mathbb{Y}_r. \quad (3.4)$$

Se tiene la desigualdad tipo Bernstein de orden r si existe un $C' > 0$ tal que para todo $N = 1, 2, 3, \dots$,

$$\|g\|_{\mathbb{Y}_r} \leq C' N^r \|g\|_{\mathbb{X}}, \quad g \in \Sigma_N. \quad (3.5)$$

A partir de las desigualdades (3.4) y (3.5) y del K -funcional, DeVore y Popov prueban en [33] que, para $0 < \alpha < r$ y $0 < q \leq \infty$, los espacios de aproximación no lineal se identifican con espacios de interpolación, a saber, $\mathcal{A}_q^\alpha(\mathbb{X}, \Sigma_N) = (\mathbb{X}, \mathbb{Y}_r)_{\alpha/r, q}$.

Uno de los objetivos de la teoría de la aproximación es encontrar la relación que existe entre el *orden de aproximación* (la velocidad con la que el error de aproximación medido en alguna norma decrece) y el espacio de funciones a la que pertenece cuando se usan determinados aproximantes y algún método de aproximación (p.e., los algoritmos

greedy o *thresholding*). En nuestro caso, como ya se dijo anteriormente, usaremos como funciones aproximantes los elementos de una base incondicional. Particularmente, en los ejemplos usaremos las bases de ondículas. Éstas son bases incondicionales para varios espacios de funciones como Lebesgue, Hardy, Sobolev, Besov, Triebel-Lizorkin y todos los espacios de funciones invariantes por reordenamiento con los índices de Boyd entre 0 y 1, (ver [81],[59] y [93]). La suavidad y el decaimiento de las funciones de escala φ y de la ondícula ψ determinan la pertenencia a los espacios de funciones en un cierto rango de parámetros.

Cuando la aproximación no lineal se hace con los elementos de una base de ondículas el siguiente resultado (que también incluye las *splines*) de DeVore-Jawerth-Popov se puede encontrar en [31]. Para los espacios de Besov $B_\tau^\alpha(L^\tau)$ con $1/\tau = (\alpha/d + 1/p)^{-1}$, siguiendo [33] se tiene que, bajo ciertas condiciones de la ondícula, para todo $0 < p < \infty$ y cierto rango de α ,

$$\sum_{N=1}^{\infty} (N^{\alpha/d} \cdot \sigma(f, N)_{L^p})^r \frac{1}{N} < \infty \Leftrightarrow f \in B_\tau^\alpha(L^\tau), \quad (3.6)$$

es decir, se caracteriza cierto espacio de aproximación no lineal (ver (3.3)) con una base de ondículas como cierto espacio de Besov. Como en el caso anterior de DeVore y Popov en [33], para probar (3.6) es suficiente con probar las siguientes desigualdades tipo Jackson

$$\sigma(f, N)_{L^p} \leq Cn^{-\beta/d} \|f\|_{B_\tau^\beta(L^\tau)}, \quad \text{para toda } f \in B_\tau^\beta(L^\tau) \quad (3.7)$$

y tipo Bernstein

$$\|f\|_{B_\tau^\beta(L^\tau)} \leq Cn^{\beta/d} \|f\|_{L^p}, \quad \text{para toda } f \in \Sigma_n. \quad (3.8)$$

para algún $\beta > \alpha$, usando ahora los elementos de la base de ondículas como aproximantes.

La monografía de DeVore [29] da cuenta del desarrollo y algunos de los resultados más importantes de la aproximación no lineal con énfasis en 1) los *splines* y en las ondículas como aproximantes, y 2) los errores medidos en las normas L^p .

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

3.2.2 Un algoritmo, democracia y otras propiedades

Hacemos ahora un paréntesis para introducir un algoritmo y unas propiedades muy importantes para la teoría de la aproximación. El error de aproximación (3.1), conocido en inglés como *best N -term approximation error* o el mejor error de aproximación con N términos, puede significar un cómputo muy costoso ya que se busca entre toda la base (o entre todo el conjunto de aproximantes) aquellos N elementos, *i.e.* $g \in \Sigma_N$, tales que se alcance (de ser posible) en ínfimo de $\|f - g\|_{\mathbb{X}}$. En la práctica, lo más común es hacer lo siguiente. Sea $f = \sum_{j \in \mathbb{N}} s_j e_j$, donde s_j son los coeficientes y e_j son los elementos de la base, y π cualquier biyección de \mathbb{N} tal que

$$\|s_{\pi(1)} e_{\pi(1)}\|_{\mathbb{X}} \geq \|s_{\pi(2)} e_{\pi(2)}\|_{\mathbb{X}} \geq \|s_{\pi(3)} e_{\pi(3)}\|_{\mathbb{X}} \geq \dots$$

Entonces, el **algoritmo ambicioso**¹ (*greedy algorithm*) de etapa N es la correspondencia

$$f = \sum_{j \in \mathbb{N}} s_j e_j \in \mathbb{X} \mapsto G_N^\pi(f) = \sum_{j=1}^N s_{\pi(j)} e_{\pi(j)} \in \Sigma_N.$$

Por la definición de σ , siempre es cierto que $\sigma(f, N)_{\mathbb{X}} \leq \|f - G_N^\pi(f)\|_{\mathbb{X}} := \gamma(f, N)_{\mathbb{X}}$. Si $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ es una base ortonormal para un espacio de Hilbert \mathbb{H} se tiene que $\sigma(f, N)_{\mathbb{H}} = \|f - G_N^\pi(f)\|_{\mathbb{H}}$ para toda $f \in \mathbb{H}$ y todo $N = 1, 2, 3, \dots$. Si existe una constante $C > 0$ tal que, para una base \mathcal{B} , se tiene

$$\|f - G_N^\pi(f)\|_{\mathbb{X}} \leq C \sigma(f, N)_{\mathbb{X}}, \quad f \in \mathbb{X}, N = 1, 2, 3, \dots,$$

se dice que \mathcal{B} es una **base ambiciosa** para \mathbb{X} . La importancia de estas bases es doble: 1) como $\sigma(f, N)_{\mathbb{X}} \approx \gamma(f, N)_{\mathbb{X}}$ las bases ambiciosas producen un casi mejor error de aproximación con N términos y 2) se pueden estudiar y/o analizar los espacios de aproximación $\mathcal{A}_\mu^\xi(\mathbb{X}, \mathcal{B})$ a través del error de aproximación $\gamma(f, N)_{\mathbb{X}}$. Un resultado de Konyagin y Temlyakov en [68] prueba que una base $\mathcal{B} = \{e_k : k \in \mathbb{N}\}$ es ambiciosa en

¹Al *greedy algorithm* se le conoce en castellano generalmente como algoritmo avaricioso. Sin embargo, la traducción que hace el *Concise Oxford Dictionary* de la palabra *greed* es codicia o avaricia. Según el diccionario de la Real Academia Española se tiene que: Avaricia: 1. f. Afán **desordenado** de poseer y adquirir riquezas para atesorarlas. También se tiene que, según la RAE: Codicia: 1. f. Afán excesivo de riquezas. 2. f. Deseo vehemente de algunas cosas buenas. Finalmente, de nuevo según la RAE: Ambicioso, sa.: 1. adj. Que tiene ambición. U. t. c. s. 2. adj. Que tiene ansia o deseo vehemente de algo. U. t. c. s.

\mathbb{X} si, y sólo si, es incondicional y democrática. Una **base democrática** es aquella para la que existe alguna constante $C > 0$ tal que

$$\left\| \sum_{j \in \Gamma} \frac{e_j}{\|e_j\|_{\mathbb{X}}} \right\|_{\mathbb{X}} \leq C \left\| \sum_{j \in \Gamma'} \frac{e_j}{\|e_j\|_{\mathbb{X}}} \right\|_{\mathbb{X}}, \quad (3.9)$$

para todos los conjuntos de índices Γ, Γ' de igual cardinal. Es fácil comprobar que la propiedad de la democracia (3.9) es equivalente a

$$\frac{1}{C} \left\| \sum_{j \in \Gamma} \frac{e_j}{\|e_j\|_{\mathbb{X}}} \right\|_{\mathbb{X}} \leq \left\| \sum_{j \in \Gamma'} \frac{e_j}{\|e_j\|_{\mathbb{X}}} \right\|_{\mathbb{X}} \leq C \left\| \sum_{j \in \Gamma} \frac{e_j}{\|e_j\|_{\mathbb{X}}} \right\|_{\mathbb{X}}, \quad (3.10)$$

para todos los conjuntos finitos Γ, Γ' de igual cardinal. Es fácil ver que para una base ortonormal \mathcal{B} en un espacio de Hilbert \mathbb{H} siempre se tiene que

$$\left\| \sum_{j \in \Gamma} \frac{e_j}{\|e_j\|_{\mathbb{H}}} \right\|_{\mathbb{H}} = (|\Gamma|)^{1/2} = N^{1/2}.$$

La siguiente caracterización para la aproximación en espacios de Hilbert con una base ortogonal fue probada por Stechkin [94] para el caso $\tau = 1$ y por DeVore-Temlyakov [34] para el caso general de τ . Para $\alpha = 1/\tau - 1/2$ se tiene que

$$f \in \mathcal{A}_s^\alpha(\mathbb{H}, \mathcal{B}) \Leftrightarrow \{\langle f, e_j \rangle_{\mathbb{H}}\}_{j \in \mathbb{N}} \in \ell^{\tau, s}, \quad (3.11)$$

donde $\ell^{\tau, s}$ es el espacio de Lorentz discreto clásico (ver la definición en la Sección 3.4.5). Algunas de estas caracterizaciones (o inclusiones) en la aproximación no lineal con N términos en espacios de funciones como los de Triebel-Lizorkin (con pesos), de Orlicz (con pesos), de Lorentz (con pesos) y de Lorentz-Zygmund (con pesos) se encuentran en [47], [48], [49], [56] y [82], y las referencias ahí citadas. Como ejemplos incluimos dos de los primeros resultados sobre estas estimaciones. El primero es de Hsiao, Jawerth, Lucier y Yu (ver la sección 11.3 en [60]) en donde muestran que el espacio de Sobolev (generalizado) H_p^α , $\alpha \in \mathbb{R}$, $1 < p < \infty$, verifica que

$$\frac{1}{C} (|\Gamma|)^{1/p} \leq \left\| \sum_{I \in \Gamma} \frac{\psi_I}{\|\psi_I\|_{H_p^\alpha}} \right\|_{H_p^\alpha} \leq C (|\Gamma|)^{1/p}, \quad (3.12)$$

para todo conjunto finito Γ con cardinal N y en donde ψ es una ondícula con cierta regularidad y cierto número de momentos nulos. En [60] a un espacio con una base

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

que cumple esta propiedad se dice que es un p -espacio. El segundo ejemplo es el de Temlyakov en [95] en donde prueba que la base de Haar en $[0, 1)$ es democrática en los espacios $L^p[0, 1)$, $1 < p < \infty$, y en donde se tiene que

$$\frac{1}{C}(|\Gamma|)^{1/p} \leq \left\| \sum_{I \in \Gamma} \frac{h_I}{\|h_I\|_{L^p}} \right\|_{L^p} \leq C(|\Gamma|)^{1/p}, \quad (3.13)$$

para todo conjunto finito Γ con cardinal N . A esta propiedad (p -espacio) se le ha llegado a conocer como la propiedad p -Temlyakov. A las cotas superiores (resp. inferiores) también se les llama democráticas por la derecha (resp. por la izquierda).

3.2.3 Resultados más generales sobre ANL

Resultados sobre aproximación no lineal en espacios más generales se encuentran en el trabajo de Garrigós y Hernández [46] en donde el error se mide en los espacios de Triebel-Lizorkin. El grueso del trabajo en [46] se hace en espacios de sucesiones y se “transfieren” (ver Subsección 6.2 en [46]) los resultados a los espacios funcionales a través de las caracterizaciones con los coeficientes de las ondículas como se hace en [41]. A continuación se describe someramente la “hoja de ruta” de [46]. Se prueba primeramente que los espacios de sucesiones asociados a los espacios funcionales de Triebel-Lizorkin, $f_{p,r}^s$ (posponemos la definición), cumplen la propiedad de ser p -espacios, es decir,

$$\frac{1}{C}(|\Gamma|)^{1/p} \leq \left\| \sum_{I \in \Gamma} \frac{\psi_I}{\|\psi_I\|_{f_{p,r}^s}} \right\|_{f_{p,r}^s} \leq C(|\Gamma|)^{1/p}, \quad (3.14)$$

para todo conjunto finito Γ . A partir de las dos desigualdades en (3.14), Garrigós y Hernández prueban a continuación las siguientes desigualdades tipo Jackson y Bernstein. Sean $s, \alpha \in \mathbb{R}$, $0 < p, \tau, q < \infty$ y $0 < r \leq \infty$ tales que

$$\max(\tau, q) < p \quad \text{y} \quad \frac{\alpha}{d} - \frac{1}{\tau} = \frac{s}{d} - \frac{1}{p}.$$

Entonces, para toda sucesión $\mathbf{s} \in b_{\tau,q}^\alpha$ (el espacio de sucesiones asociado al espacio de Besov) existe un $C > 0$ tal que (desigualdad tipo Jackson)

$$\sigma(\mathbf{s}, N)_{f_{p,r}^s} \leq C(N+1)^{-(1/(\tau \vee q) - 1/p)} \|\mathbf{s}\|_{b_{\tau,q}^\alpha}, \quad N = 0, 1, 2, \dots \quad (3.15)$$

Sean s, α, p, τ, q y r con el mismo rango que anteriormente y tales que

$$\min(\tau, q) < p \quad \text{y} \quad \frac{\alpha}{d} - \frac{1}{\tau} = \frac{s}{d} - \frac{1}{p}.$$

3.3 Aproximación no lineal restringida

Entonces, existe un $C > 0$ tal que, para todo $N = 1, 2, 3, \dots$, (desigualdad tipo Bernstein)

$$\|\mathbf{s}\|_{b_{\tau,q}^{\alpha}} \leq CN^{1/(\tau \wedge q)-1/p} \|\mathbf{s}\|_{f_{p,r}^s}, \quad \forall \mathbf{s} \in \Sigma_N. \quad (3.16)$$

Aplicando el formalismo general de DeVore-Popov en [33] (las desigualdades tipo Jackson y Bernstein implican que el espacio de aproximación es un espacio de interpolación), Garrigós y Hernández [46] prueban que algunos espacios de aproximación (en p -espacios de sucesiones) se pueden caracterizar como espacios de Lorentz discretos (ver definición en la Sección 3.4.5). Específicamente, si f es un p -espacio de sucesiones para algún $0 < p < \infty$, entonces para todo $\gamma > 0$ y $0 < q \leq \infty$ se tiene

$$\mathcal{A}_q^{\gamma}(f) = \ell^{\tau,q}(f), \quad \frac{1}{\tau} = \gamma + \frac{1}{p}. \quad (3.17)$$

Un par de inclusiones de este tipo (en un contexto ligeramente diferente) aparecen en el trabajo de Gribonval-Nielsen ([49], Teorema 3.1) en donde también se sigue el esquema general de [33]. Kerkyacharian-Picard ([66], Teorema 1) también prueban esta caracterización (en un contexto ligeramente diferente) pero a través de la propiedad p -Temlyakov en vez de las desigualdades tipo Jackson y tipo Bernstein y, por lo tanto, no usan el esquema general (interpolación) de [33].

Por último, como corolarios a (3.17) en [46] se tiene que si $s \in \mathbb{R}$, $0 < p < \infty$, $0 < r \leq \infty$, entonces para todo $\gamma > 0$ y $0 < \theta < 1$,

$$\mathcal{A}_{\tau}^{\gamma/d}(f_{p,r}^s) = \mathcal{A}_{\tau}^{\gamma/d}(b_{p,p}^s) = b_{\tau,\tau}^{s+\gamma}, \quad \text{para } \frac{1}{\tau} = \frac{\gamma}{d} + \frac{1}{p} \quad (3.18)$$

y

$$(f_{p,r}^s, b_{\tau,\tau}^{s+\gamma})_{\theta,\tau\theta} = b_{\tau\theta,\tau\theta}^{s+\theta\gamma} \quad \text{para } \frac{1}{\tau} = \frac{\gamma}{d} + \frac{1}{p} \quad \text{y} \quad \frac{1}{\tau\theta} = \frac{\theta\gamma}{d} + \frac{1}{p}. \quad (3.19)$$

3.3 Aproximación no lineal restringida

Este tipo de aproximación involucra una medida ν sobre el conjunto de índices y se usa para controlar (restringir) los términos (coeficientes de ondículas) en la aproximación. Este tipo de aproximación fue introducido por Cohen-DeVore-Hochmuth en [24]. Además, en la sección 4 de [24] simplifican algunas pruebas de DeVore-Jawerth-Popov [31].

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Como se puede observar de las propiedades p -espacio (3.12), p -Temlyakov (3.13) y de democracia (3.9), los coeficientes están normalizados en la misma norma en la que se “mide” la función. El desarrollo en [24] se hace midiendo el error en la norma de los espacios de Hardy H^p , $0 < p < \infty$ ($H^p = L^p$ si $1 < p < \infty$), pero la normalización se hace en un espacio H^r , $0 < r < \infty$, diferente. Por conveniencia usaremos la notación

$$1_{\Gamma,r} = \sum_{Q \in \Gamma} \frac{\psi_Q}{\|\psi_Q\|_{H^r}}.$$

En [24] se investigan las consecuencias de normalizar los coeficientes en otra norma diferente a la del espacio en el cual se mide el error σ . La “hoja de ruta” es la misma que en el caso [46] de aproximación no lineal para los espacios Triebel-Lizorkin (de hecho [24] es cronológicamente anterior a [46] pero por cuestiones didácticas no hemos introducido la aproximación no lineal restringida hasta ahora). Una de las novedades de [24] es que, aunque el error se mide en H^p , los espacios de aproximación no están necesariamente contenidos en H^p .

En [24] se observa que normalizar los coeficientes en un espacio diferente al que se usa para medir el error es equivalente a realizar un tipo de aproximación no lineal, que llaman **restringida**, en la que una medida ν en el conjunto de cubo diádicos de \mathbb{R}^d se usa para controlar el número de términos de la aproximación. En el caso de [24], $\nu_\alpha(Q) := |Q|^\alpha = |Q|^{1-p/r}$ y para un conjunto Γ se tiene $\nu_\alpha(\Gamma) := \sum_{Q \in \Gamma} |Q|^\alpha$. El primer resultado sobre aproximación no lineal restringida en [24] es que para $0 < p, r < \infty$, $s > 0$ y τ tal que $s = d\beta(1/\tau - 1/p)$, donde $\beta = 1 - \alpha = p/r$, se tiene que los espacios H^p (también desarrollan resultados para los espacios B^p) cumplen un tipo de propiedad como la de p -Temlyakov o de ser un p -espacio. A saber, si $f \in H^p + B_\tau^s$ entonces,

$$\frac{1}{C} \nu_\alpha(\Gamma)^{1/p} \leq \|1_{\Gamma,r}\|_{H^p} \leq C \nu_\alpha(\Gamma)^{1/p}. \quad (3.20)$$

Definimos el error de aproximación no lineal restringida con medida ν_α en el espacio H^p como $\sigma_{\nu_\alpha}(f, t)_{H^p} = \inf_{g \in \Sigma_{t,\nu_\alpha}} \|f - g\|_{H^p}$, en donde $g \in \Sigma_{t,\nu_\alpha}$ significa que el soporte Γ de los coeficientes de la expansión de g tienen una medida ν_α menor o igual a t . A partir de las desigualdades en (3.20) se prueban en [24] las desigualdades tipo Jackson: existe $C > 0$ tal que para todo número real $t > 0$ se tiene

$$\sigma_{\nu_\alpha}(f, t)_{H^p} \leq C t^{-s/\beta d} \|f\|_{B_\tau^s}, \quad (3.21)$$

y tipo Bernstein: para toda $f \in \Sigma_{t, \nu_\alpha}$,

$$\|f\|_{B_\tau^s} \leq C t^{s/\beta d} \|f\|_{H^p}. \quad (3.22)$$

Nuevamente, a partir del formalismo general de DeVore-Popov en [33], se prueba en [24] la siguiente caracterización. Sean $0 < p, r < \infty$, $s > 0$ y $\tau > 0$ tal que $s = d\beta(1/\tau - 1/p)$. Para todo $0 < \gamma < s$ y μ tal que $\gamma = d\beta(1/\mu - 1/p)$, entonces

$$\mathcal{A}_\mu^{\gamma/d\beta}(H^p, \Psi) = B_\mu^\gamma, \quad (3.23)$$

con normas equivalentes y en donde Ψ es una base de ondículas.

3.3.1 ANL restringida en espacios de Banach

En un trabajo posterior a [24], Kerkyacharian-Picard [67] generalizan los resultados anteriores a un esquema de cuasi espacios de Banach con bases incondicionales. Prueban, entre otras cosas, que ciertas desigualdades tipo Jackson y tipo Bernstein son equivalente a la propiedad p -Temlyakov para cuasi espacios de Banach y, a su vez, equivalentes a la caracterización de cierto espacio de aproximación como cierto espacio de Lorentz discreto con pesos (ver definición en la Sección 3.4.5). Específicamente, para un espacio cuasi Banach \mathbb{X} , una base incondicional $\mathcal{B} = \{e_i\}_{i \in \mathbb{N}}$, una medida ν sobre el conjunto de índices, una $f = \sum_{i \in \mathbb{N}} s_i e_i \in \mathbb{X}$ y $0 < p < \infty$, entonces son equivalentes:

1. Para todo $0 < q < p$, los espacios de Lorentz $\ell^q(\mathcal{B}, \nu)$ verifican la desigualdad de Jackson con parámetro $s = 1/q - 1/p$,

$$\sigma_\nu(f, t)_{\mathbb{X}} \leq C t^{-s} \|f\|_{\ell^q(\mathcal{B}, \nu)}, \quad f \in \ell^q(\mathcal{B}, \nu), \quad (3.24)$$

y la desigualdad de Bernstein con parámetro $s = 1/q - 1/p$,

$$\|f\|_{\ell^q(\mathcal{B}, \nu)} \leq C' t^s \|f\|_{\mathbb{X}} \quad f \in \Sigma_{t, \nu}. \quad (3.25)$$

2. El esquema $(\mathbb{X}, \mathcal{B}, \nu)$ verifica la propiedad p -Temlyakov, es decir, para $\nu(\Gamma) < \infty$, se tiene

$$\frac{1}{C} (\inf_{j \in \Gamma} |s_j|) \nu^{1/p}(\Gamma) \leq \|f\|_{\mathbb{X}} \leq C (\sup_{j \in \Gamma} |s_j|) \nu^{1/p}(\Gamma). \quad (3.26)$$

3. Para todo $0 < q < p$ y para $0 < r \leq \infty$,

$$\mathcal{A}_r^s(\mathbb{X}, \mathcal{B}, \nu) = \ell^{q,r}(\mathcal{B}, \nu). \quad (3.27)$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

3.3.2 Thresholding restringido en H^r y B^p

Uno de los métodos numéricos más frecuentes en la aproximación con ondículas es el llamado *thresholding*. Este método consiste en escoger aquellos coeficientes s_Q de la representación de $f = \sum_{Q \in \Gamma} s_Q \psi_Q$ que sobrepasen un umbral (*threshold*), es decir, $|s_Q| > \lambda$. En el marco de la aproximación no lineal restringida y usando una notación distinta a [24], definimos el operador **threshold restringido** (en H^r) como

$$T_\lambda(f, H^r) = \sum_{Q: \|s_Q \psi_Q\|_{H^r} > \lambda} s_Q \psi_Q.$$

En [24] Cohen-DeVore-Hochmuth prueban que el error de aproximación con *thresholding* tiene un decaimiento prescrito si, y sólo si, los coeficientes normalizados en H^r pertenecen a cierto espacio de Lorentz (ver definición en la Sección 3.4.5). A saber, sean $0 < p, r < \infty$, $s > 0$ y τ definido por $s = d\beta(1/\tau - 1/p)$. Para todo $\tau < \mu < p$ una función f satisface

$$\|f - T_\lambda(f, H^r)\|_{H^p} \leq M^{\mu/p} \lambda^{1-\mu/p} \quad (3.28)$$

si, y sólo si,

$$\{\|s_Q \psi_Q\|_{H^r}\}_{Q \in \mathcal{D}} \in \ell^{\mu, \infty}(\nu_\alpha) \quad (3.29)$$

y, además, el menor M que satisface (3.28) es equivalente a $\|\{\|s_Q \psi_Q\|_{H^r}\}_{Q \in \mathcal{D}}\|_{\ell^{\mu, \infty}(\nu_\alpha)}$.

Nuevamente, en [67], Kerkyacharian-Picard generalizan la caracterización anterior de *thresholding*, i.e. (3.28) y (3.29), a los cuasi espacios de Banach y muestran que es equivalente a la propiedad p -Temlyakov. Específicamente, son equivalentes:

1. El esquema $(\mathbb{X}, \mathcal{B}, \nu)$ verifica la propiedad p -Temlyakov (3.26), es decir, para $\nu(\Gamma) < \infty$, se tiene

$$\frac{1}{C} (\inf_{j \in \Gamma} |s_j|) \nu^{1/p}(\Gamma) \leq \|f\|_{\mathbb{X}} \leq C (\sup_{j \in \Gamma} |s_j|) \nu^{1/p}(\Gamma).$$

2. Para todo $0 < q < p$,

$$\|f\|_{\ell^{q, \infty}(\mathcal{B}, \nu)}^{q/p} \approx \sup_{\lambda} \lambda^{-(1-q/p)} \|f - T_\lambda(f, \mathbb{X})\|_{\mathbb{X}}. \quad (3.30)$$

En este trabajo generalizamos los resultados de Kerkyacharian y Picard para los espacios de sucesiones en donde no se tiene necesariamente que $\eta(t) = t^{1/p}$ sino una función más general.

3.4 ANL restringida en espacios de sucesiones

Presentamos y desarrollamos en esta sección algunos resultados originales sobre la aproximación no lineal restringida.

Aquí mostramos que las inclusiones para los espacios de aproximación no lineal restringida en espacios de sucesiones en términos de espacios de Lorentz con pesos son equivalentes a desigualdades tipo Jackson y Bernstein y, a su vez, equivalentes a una propiedad generalizada del tipo p -Temlyakov por arriba y por abajo.

Trabajar en el espacio de sucesiones no es ninguna limitación. Los espacios de Lebesgue, Sobolev, Hardy y Lipschitz tienen su contraparte en el espacio de sucesiones cuando se usa la transformada φ ([39], [41]) o wavelets ([72], [81], [28],[59], [77], [70]). De manera más general, los espacios de sucesiones Triebel-Lizorkin, $f_{p,r}^s$, y Besov, $b_{p,r}^s$ (ver Sección 3.4.9) permiten representaciones fieles de los espacios de funciones (distribuciones) Triebel-Lizorkin, $F_{p,r}^s$, y Besov, $B_{p,r}^s$, que incluyen todos los mencionados al principio de este párrafo. Cuando los resultados en el espacio de sucesiones se acoplan con el marco de transferencia abstracta diseñada en [46], se recuperan los resultados para los espacios de distribuciones o funciones. Una razón para considerar dicho escenario general, además de las generalizaciones obvias, es que medir el error $\|f - T_\varepsilon(f)\|$ en los espacios de Sobolev, donde las propiedades de suavidad se tienen en cuenta, puede dar una representación visual más fiable de f que cuando el error se mide en L^p . Observe que dos funciones pueden ser visualmente muy diferentes aunque sean cercanas en la norma L^p .

Mostramos también que los espacios de sucesiones de Triebel-Lizorkin, $f_{p,r}^s$, cumplen con la propiedad de Temlyakov (por arriba y por abajo) en el esquema de ANL restringido. Aplicando los resultados de las Subsecciones 3.4.6 y 3.4.8 mostramos que los espacios de aproximación de los espacios de sucesiones de Triebel-Lizorkin se identifican con espacios de Lorentz discretos, que a su vez coinciden con ciertos espacios de Besov para algunos valores particulares de los parámetros. Así, los resultados en [24] y [67] son corolarios de los presentados aquí. Por último, damos un resultado de interpolación entre los espacios de sucesiones de Triebel-Lizorkin y Besov con menos restricciones en los parámetros que aquellos considerados en [24].

En las siguientes subsecciones presentamos el marco de trabajo en el cual desarrollamos nuestros resultados, al mismo tiempo introducimos notación, definiciones y

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

comentarios.

3.4.1 Espacios de sucesiones

Sea \mathcal{D} un conjunto de índices numerable cuyos elementos denotamos por I . El conjunto \mathcal{D} puede ser $\mathbb{N}, \mathbb{Z}, \dots$ o, como en nuestras aplicaciones, el conjunto numerable de cubos diádicos en \mathbb{R}^d .

Denotamos por $S = \mathbb{C}^{\mathcal{D}}$ el conjunto de todas las sucesiones de números complejos $\mathbf{s} = \{s_I\}_{I \in \mathcal{D}}$ definido sobre \mathcal{D} . Por cada $I \in \mathcal{D}$, denotamos por \mathbf{e}_I el elemento de S con entrada 1 en I y 0 en cualquier otro caso. Escribimos $\mathcal{E} = \{\mathbf{e}_I\}_{I \in \mathcal{D}}$ para la base canónica de S . Usaremos la notación $\sum_{I \in \Gamma} s_I \mathbf{e}_I$, $\Gamma \in \mathcal{D}$, para denotar el elemento de S cuyas entradas son s_I cuando $I \in \Gamma$ y 0 en cualquier otro caso. Observe que no hay un significado de convergencia impuesto a la notación anterior, aún cuando Γ no sea finito.

Definición 3.4.1 *Un espacio lineal de sucesiones $f \subset S$ es un **retículo (de sucesiones) cuasi Banach** si existe una cuasi norma $\|\cdot\|_f$ en f con respecto al cual f es completo y satisface:*

- (a) *Monotonicidad: si $\mathbf{t} \in f$ y $|s_I| \leq |t_I|$ para todo $I \in \mathcal{D}$, entonces $\mathbf{s} \in f$ y $\|\{s_I\}\|_f \leq \|\{t_I\}\|_f$.*
- (b) *Si $\mathbf{s} \in f$, entonces $\lim_{n \rightarrow \infty} \|s_{I_n} \mathbf{e}_{I_n}\|_f = 0$, para alguna enumeración $\mathcal{J} = \{I_1, I_2, \dots\}$.*

Diremos que un retículo (de sucesiones) cuasi Banach f está incluido en S , y escribiremos $f \hookrightarrow S$ si

$$\lim_{n \rightarrow \infty} \|\mathbf{s}^n - \mathbf{s}\|_f = 0 \Rightarrow \lim_{n \rightarrow \infty} s_I^n = s_I, \forall I \in \mathcal{D}.$$

Nota 3.4.2 *Cuando $\mathcal{E} = \{\mathbf{e}_I\}_{I \in \mathcal{D}}$ es una base de Schauder para f , la condición (a) en la Definición 3.4.1 implica que \mathcal{E} es una base incondicional para f con constante $C = 1$.*

Nota 3.4.3 *El espacio cuasi Banach f puede ser dotado con una cuasi norma equivalente, la cual puede denotarse de nuevo por $\|\cdot\|_f$, para la cual existe un $\rho \in (0, 1]$ tal que $\|\mathbf{s} + \mathbf{t}\|_f \leq \|\mathbf{s}\|_f^\rho + \|\mathbf{t}\|_f^\rho$ para todo $\mathbf{s}, \mathbf{t} \in f$. En este caso decimos que $\|\cdot\|_f$ satisface la desigualdad triangular de potencia ρ o desigualdad ρ -triangular. Observe que también satisface la desigualdad triangular de potencia μ tal que $0 < \mu \leq \rho$*

3.4.2 Aproximación no lineal restringida en espacios de sucesiones

Una medida positiva ν sobre \mathcal{D} es cualquier función $\nu : \mathcal{D} \rightarrow \mathbb{R}$ tal que $\nu(I) > 0$ para todo $I \in \mathcal{D}$. En la aproximación clásica con N términos ν es la medida de contar (*i.e.*, $\nu(I) = 1$ para todo $I \in \mathcal{D}$), pero medidas más generales se usan en la aproximación no lineal restringida. La medida ν controlará los términos en la aproximación.

Definición 3.4.4 *Decimos que (f, ν) es un **esquema canon** (para la ANL restringida) si*

- i) f es un retículo (de sucesiones) cuasi Banach incluido en S .*
- ii) ν es una medida positiva sobre \mathcal{D} como se explica en el primer párrafo de esta sección.*

Sea (f, ν) un esquema canon. Para $t > 0$, definimos

$$\Sigma_{t,\nu} := \left\{ \mathbf{t} = \sum_{I \in \Gamma} t_I \mathbf{e}_I : \nu(\Gamma) \leq t \right\}.$$

Nótese que $\Sigma_{t,\nu}$ es no lineal, pero se cumple que $\Sigma_{t,\nu} + \Sigma_{t,\nu} \subset \Sigma_{2t,\nu}$.

Dado $\mathbf{s} \in S$, el error- f (o riesgo- f) de la aproximación a \mathbf{s} por elementos de $\Sigma_{t,\nu}$ está dado por

$$\sigma_\nu(t, \mathbf{s})_f := \inf_{\mathbf{t} \in \Sigma_{t,\nu}} \|\mathbf{s} - \mathbf{t}\|_f.$$

Nótese que elementos $\mathbf{s} \in S$ que no estén en f podrían tener un riesgo- f finito ya que los elementos de $\Sigma_{t,\nu}$ podrían tener un número infinito de entradas.

Definición 3.4.5 (Espacios de aproximación restringida) *Sea (f, ν) un esquema canon. Para $0 < \xi < \infty$ y $0 < \mu \leq \infty$, $\mathcal{A}_\mu^\xi(f, \nu)$ se define como todas las sucesiones $\mathbf{s} \in S$ tales que*

$$\|\mathbf{s}\|_{\mathcal{A}_\mu^\xi(f, \nu)} := \left(\int_0^\infty [t^\xi \sigma_\nu(t, \mathbf{s})_f]^\mu \frac{dt}{t} \right)^{1/\mu} < \infty, \quad (3.31)$$

con la modificación usual cuando $\mu = \infty$.

Nótese que los espacios $\mathcal{A}_\mu^\xi(f, \nu)$ dependen de la base canónica \mathcal{E} de S . Cuando f sea claro en el contexto escribiremos $\mathcal{A}_\mu^\xi(\nu)$ en vez de $\mathcal{A}_\mu^\xi(f, \nu)$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Nota 3.4.6 Si $\mathbf{s} \in f$ entonces, por el hecho de que $\sigma_\nu(t, \mathbf{s})_f \leq \|\mathbf{s}\|_f$, es fácil ver que (3.31) puede remplazarse por $\|\mathbf{s}\|_f$ más la misma integral de 1 a ∞ . Necesitamos considerar el rango completo $0 < t < \infty$ ya que no suponemos que $\mathbf{s} \in f$. No obstante, las propiedades de los espacios de aproximación no lineal restringida son las mismos que las de los espacios de aproximación con N -términos (ver [89] o [32]).

Partiendo la integral en pedazos diádicos y usando la monotonicidad del error- f , $\sigma_\nu(t, \mathbf{s})_f$, tenemos la siguiente cuasi norma equivalente para los espacios de aproximación restringida:

$$\|\mathbf{s}\|_{A_{\mu}^{\xi}(\nu)} \approx \left(\sum_{k \in \mathbb{Z}} [2^{k\xi} \sigma_\nu(2^k, \mathbf{s})_f]^\mu \right)^{1/\mu}. \quad (3.32)$$

3.4.3 Dos definiciones importantes

Es bien sabido el papel fundamental de las desigualdades tipo Jackson y tipo Bernstein para la teoría de la aproximación no lineal (ver [29]). Considerando nuestro esquema canon (f, ν) tenemos las siguientes definiciones.

Definición 3.4.7 Dado $r > 0$, un retículo (de sucesiones) cuasi Banach $g \hookrightarrow S$ satisface la desigualdad de Jackson de orden r si existe un $C > 0$ tal que

$$\sigma_\nu(t, \mathbf{s})_f \leq Ct^{-r} \|\mathbf{s}\|_g, \quad \text{para toda } \mathbf{s} \in g \text{ y para todo } t > 0.$$

Definición 3.4.8 Dado $r > 0$, un retículo (de sucesiones) cuasi Banach $g \hookrightarrow S$ satisface la desigualdad de Bernstein de orden r si existe un $C > 0$ tal que

$$\|\mathbf{t}\|_g \leq Ct^r \|\mathbf{t}\|_f, \quad \text{para toda } \mathbf{t} \in \Sigma_{t,\nu} \cap f \text{ y para todo } t > 0.$$

En estas definiciones no suponemos que $g \hookrightarrow f$, pero sí que $\mathbf{t} \in \Sigma_{t,\nu} \cap f$. Así, la Definición 3.4.8 tiene sentido.

3.4.4 Funciones de peso para espacios de Lorentz discretos

La propiedad clásica p -Temlyakov está expresada típicamente en términos de $(N)^{1/p}$, en donde N es el número de aproximantes y p es el parámetro de integración en algún espacio de funciones. En [24] y [67] se considera la función $(\nu(\Gamma))^{1/p}$. Para generalizar estos trabajos se pensó en otras funciones que se describen continuación.

3.4 ANL restringida en espacios de sucesiones

Definición 3.4.9 Denotaremos por \mathbb{W} el conjunto de todas las funciones continuas $\eta : [0, \infty) \mapsto [0, \infty)$ tales que

i) $\eta(0) = 0$ and $\lim_{t \rightarrow \infty} \eta(t) = \infty$.

ii) η es no decreciente.

iii) η tiene la propiedad doblante: existe $C > 0$ tal que $\eta(2t) \leq C\eta(t)$ para todo $t > 0$.

Un elemento típico de la clase \mathbb{W} es $\eta(t) = t^{1/p}$, $0 < p < \infty$. Las funciones en la clase \mathbb{W} serán usadas para definir espacios de Lorentz discretos generales. Ocasionalmente, necesitaremos asumir una condición más fuerte en $\eta \in \mathbb{W}$. Para $\eta \in \mathbb{W}$ definimos la función de dilatación

$$M_\eta(s) = \sup_{t>0} \frac{\eta(st)}{\eta(t)}, \quad s > 0.$$

Ya que η es no decreciente, $M_\eta(s) \leq 1$ para $0 < s \leq 1$.

Definición 3.4.10 Decimos que $\eta \in \mathbb{W}_+$ si $\eta \in \mathbb{W}$ y existe $s_0 \in (0, 1)$ para el cual $M_\eta(s_0) < 1$.

Observe que para $\eta \in \mathbb{W}_+$ y $r > 0$, $\eta^r \in \mathbb{W}_+$. Además, si $\eta \in \mathbb{W}$ y $r > 0$, $t^r \eta(t) \in \mathbb{W}_+$.

Lema 3.4.11 Sea $\eta \in \mathbb{W}_+$ y s_0 como en la definición 3.4.10. Entonces, existe $C > 0$ tal que para todo $t > 0$,

$$\sum_{j=0}^{\infty} \eta(s_0^j t) \leq C\eta(t).$$

Demostración. Sea $\delta := M_\eta(s_0) < 1$. Por definición de M_η se tiene

$$1 > \delta \geq \frac{\eta(s_0^{j+1}t)}{\eta(s_0^j t)} \quad \text{para todo } j = 0, 1, 2, \dots$$

Por lo tanto,

$$\sum_{j=0}^{\infty} \eta(s_0^j t) \leq \sum_{j=0}^{\infty} \delta^j \eta(t) = \eta(t) \frac{1}{1 - \delta}.$$

■

Lema 3.4.12 Dada $\eta \in \mathbb{W}_+$, existe $g \in C^1$, $g \in \mathbb{W}_+$ tal que $g \approx \eta$ y $g'(t)/g(t) \approx 1/t$, $t > 0$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Demostración. Defínase $g(t) = \int_0^t \frac{\eta(s)}{s} ds$. Con s_0 como en la Definición 3.4.10

$$g(t) = \sum_{j=0}^{\infty} \int_{s_0^{j+1}t}^{s_0^j t} \frac{\eta(s)}{s} ds \leq \sum_{j=0}^{\infty} \eta(s_0^j t) \log(s_0^{-1}) \leq C\eta(t) \log(s_0^{-1}),$$

por el Lema 3.4.11 ($C = \frac{1}{1-\delta}$, ver la prueba del Lema 3.4.11). Por otra parte

$$g(t) \geq \int_{t/2}^t \frac{\eta(s)}{s} ds \geq \eta(t/2) \log 2 \geq D\eta(t) \log 2$$

por la propiedad doblante de η . Esto muestra que

$$C_1\eta(t) \leq g(t) \leq C_2\eta(t), \quad t \in (0, \infty), \quad (3.33)$$

con $0 < C_1 \leq C_2 < \infty$. Como prueba alternativa uno puede ver que η satisface la hipótesis del Lema 1.4 en [69] (p. 54) para concluir $g \approx \eta$. La función g es claramente no decreciente y (3.33) muestra que $g \in \mathbb{W}$. Es claro que $g \in C^1$ con $g'(t) = \eta(t)/t$. Así,

$$\frac{g'(t)}{g(t)} \approx \frac{\eta(t)/t}{\eta(t)} \approx \frac{1}{t}.$$

Falta probar que $g \in \mathbb{W}_+$. Para probar esto, observe que una función $\eta \in \mathbb{W}$ es un elemento de \mathbb{W}_+ si y sólo si

$$i_\eta := \lim_{t \rightarrow 0^+} \frac{\log M_\eta(t)}{\log t} > 0$$

(i_η se llama el índice de dilatación (de Boyd) de η - ver [8]). Usando (3.33)

$$i_g = \lim_{t \rightarrow 0^+} \frac{\log M_g(t)}{\log t} \geq \lim_{t \rightarrow 0^+} \frac{\log(\frac{C_1}{C_2} M_\eta(t))}{\log t} = \lim_{t \rightarrow 0^+} \frac{\log(M_\eta(t))}{\log t} = i_\eta$$

y, similarmente

$$i_g \leq \lim_{t \rightarrow 0^+} \frac{\log(\frac{C_2}{C_1} M_\eta(t))}{\log t} = i_\eta.$$

Así, $i_g = i_\eta > 0$ lo que prueba $g \in \mathbb{W}_+$. ■

3.4.5 Espacios de Lorentz discretos generales

Antes de la definición de estos espacios recordamos algunas definiciones clásicas (ver p.e. [32] y [8]). Para una sucesión $\mathbf{s} = \{s_I\}_{I \in \mathcal{D}} \in S$, \mathcal{D} numerable, el reordenamiento no creciente de \mathbf{s} con respecto a la medida ν en \mathcal{D} es

$$\begin{aligned} \mathbf{s}_\nu^*(t) &= \inf\{\lambda > 0 : \nu(\{I \in \mathcal{D} : |s_I| > \lambda\}) \leq t\} \\ &= \sup\{\lambda > 0 : \nu(\{I \in \mathcal{D} : |s_I| > \lambda\}) > t\}. \end{aligned}$$

así, para todo $t, \lambda \geq 0$, se tiene

$$0 \leq \lambda < \mathbf{s}_\nu^*(t) \iff 0 \leq t < \nu(\{I \in \mathcal{D} : |s_I| > \lambda\}). \quad (3.34)$$

Aún más, es bien sabido (ver Sección 1.3 en [9]), que \mathbf{s} y \mathbf{s}_ν^* tienen la misma función de distribución. En particular, para $t > 0$,

$$\nu(\{I \in \mathcal{D} : |s_I| > \mathbf{s}_\nu^*(t)\}) \leq t \leq \nu(\{I \in \mathcal{D} : |s_I| \geq \mathbf{s}_\nu^*(t)\}). \quad (3.35)$$

Para $\eta \in \mathbb{W}$, ν una medida en \mathcal{D} , y $\mu \in (0, \infty]$, el **espacio de Lorentz discreto** $\ell_\eta^\mu(\nu)$ es el conjunto de todas las $\mathbf{s} = \{s_I\}_{I \in \mathcal{D}} \in S$ tal que

$$\|\mathbf{s}\|_{\ell_\eta^\mu(\nu)} := \left(\int_0^\infty [\eta(t)\mathbf{s}_\nu^*(t)]^\mu \frac{dt}{t} \right)^{1/\mu} < \infty, \quad 0 < \mu < \infty \quad (3.36)$$

y

$$\|\mathbf{s}\|_{\ell_\eta^\infty(\nu)} := \sup_{t>0} \eta(t)\mathbf{s}_\nu^*(t) < \infty.$$

Si $\eta(t) = t^{1/p}$, $1 \leq p < \infty$, entonces $\ell_\eta^\mu(\nu) = \ell^{p,\mu}(\nu)$ son los espacios clásicos de Lorentz (discretos). Cuando, además, $p = \mu$, $\ell^{p,p}(\nu) = \ell^p(\nu)$, $0 < p < \infty$, son los espacios de sucesiones $\mathbf{s} \in S$ tales que

$$\|\mathbf{s}\|_{\ell^p(\nu)} = \left(\sum_{I \in \mathcal{D}} |s_I|^p \nu(I) \right)^{1/p}.$$

Para $\xi > 0$ y $\eta \in \mathbb{W}$, escribimos $\tilde{\eta}(t) = t^\xi \eta(t) \in \mathbb{W}_+$ y $\ell_\eta^\mu(\nu)$ sera denotado por $\ell_{\xi,\eta}^\mu(\nu)$.

Proposición 3.4.13 *Sean $\eta \in \mathbb{W}$ y ν una medida en \mathcal{D} . Para una sucesión $\mathbf{s} = \{s_I\}_{I \in \mathcal{D}} \in S$ se tiene*

$$\|\mathbf{s}\|_{\ell_\eta^\infty(\nu)} \approx \sup_{\lambda>0} \lambda \eta(\nu(I \in \mathcal{D} : |s_I| > \lambda)).$$

Además, si $0 < \mu < \infty$ y $\eta \in \mathbb{W}_+$,

$$\|\mathbf{s}\|_{\ell_\eta^\mu(\nu)} \approx \left(\int_0^\infty [\lambda \eta(\nu(I \in \mathcal{D} : |s_I| > \lambda))]^\mu \frac{d\lambda}{\lambda} \right)^{1/\mu}.$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Demostración. El caso $\mu = \infty$ se sigue de la parte iii) de la Proposición 2.2.5 en [21].

Para $0 < \mu < \infty$, sea $w(t) = [\eta(t)]^\mu/t$, $0 < t < \infty$. Escribiendo $\lambda_\nu(t, \mathbf{s}) = \nu(\{I \in \mathcal{D} : |s_I| > t\})$ para la función de distribución de \mathbf{s} con respecto de la medida ν y $W(s) = \int_0^s w(t)dt$, $0 < s < \infty$, la parte ii) de la Proposición 2.2.5 in [21] da

$$\|\mathbf{s}\|_{\ell_\eta^\mu(\nu)} = \left(\int_0^\infty \mu t^\mu W(\lambda_\nu(t, \mathbf{s})) \frac{dt}{t} \right)^{1/\mu}.$$

Ya que $\eta \in \mathbb{W}_+$, entonces η^μ satisface la hipótesis del Lema 1.4 en [69] (p. 54) con lo que concluimos

$$W(s) = \int_0^s \frac{\eta(t)^\mu}{t} dt \approx [\eta(s)]^\mu$$

(ver también la prueba del Lema 3.4.12 y el comentario que sigue la Definición 3.4.10). Esto prueba el resultado. ■

Una sucesión $\mathbf{u} = \{u_I\}_{I \in \mathcal{D}} \in S$ tal que $u_I > 0$ para todo $I \in \mathcal{D}$ será llamada **sucesión de pesos**.

Definición 3.4.14 Sea $\mathbf{u} = \{u_I\}_{I \in \mathcal{D}}$ una sucesión de pesos y ν una medida positiva como en la Subsección 3.4.2. Para $0 < \mu \leq \infty$ y $\eta \in \mathbb{W}$ se define el espacio $\ell_\eta^\mu(\mathbf{u}, \nu)$ como el conjunto de todas las sucesiones $\mathbf{s} = \sum_{I \in \mathcal{D}} s_I \mathbf{e}_I \in S$ tal que

$$\|\mathbf{s}\|_{\ell_\eta^\mu(\mathbf{u}, \nu)} := \|\{u_I s_I\}_{I \in \mathcal{D}}\|_{\ell_\eta^\mu(\nu)} < \infty.$$

Estos espacios se usarán en las Subsecciones 3.4.6 y 3.4.8 para caracterizar las desigualdades tipo Jackson y Bernstein en el escenario de la aproximación no lineal restringida. Para las aplicaciones (ver Subsecciones 3.4.7, 3.4.9-3.4.11) tomaremos $u_I = \|\mathbf{e}_I\|_g$, $I \in \mathcal{D}$, donde g es un retículo (de sucesiones) cuasi Banach.

Lema 3.4.15 Sea \mathbf{u} y ν como en la Definición 3.4.14 y escribáse $1_{\Gamma, \mathbf{u}} = \sum_{I \in \Gamma} u_I^{-1} \mathbf{e}_I$, $\Gamma \subset \mathcal{D}$ y $\nu(\Gamma) < \infty$.

- a) Si $\eta \in \mathbb{W}$, $\|1_{\Gamma, \mathbf{u}}\|_{\ell_\eta^\infty(\mathbf{u}, \nu)} = \eta(\nu(\Gamma))$.
- b) Si $0 < \mu < \infty$ y $\eta \in \mathbb{W}$, $\|1_{\Gamma, \mathbf{u}}\|_{\ell_\eta^\mu(\mathbf{u}, \nu)} \geq \eta(\nu(\Gamma))$, y si $\eta \in \mathbb{W}_+$, $\|1_{\Gamma, \mathbf{u}}\|_{\ell_\eta^\mu(\mathbf{u}, \nu)} \approx \eta(\nu(\Gamma))$.

Demostración. (a) Escribiendo $1_{\Gamma, \mathbf{u}} = \sum_{I \in \mathcal{D}} s_I \mathbf{e}_I$ se tiene $s_I = u_I^{-1}$ para todo $I \in \Gamma$ y $s_I = 0$ si $I \notin \Gamma$. Así, $u_I s_I = 1$ para todo $I \in \Gamma$ y $u_I s_I = 0$ si $I \notin \Gamma$. Esto implica

$$\{u_I s_I\}_\nu^*(t) = \begin{cases} 1, & 0 < t < \nu(\Gamma) \\ 0, & t \geq \nu(\Gamma) \end{cases}. \quad (3.37)$$

Por la Definición 3.4.14

$$\|1_{\Gamma, \mathbf{u}}\|_{\ell_\eta^\infty(\mathbf{u}, \nu)} = \sup_{0 < t < \nu(\Gamma)} \eta(t) = \eta(\nu(\Gamma)).$$

(b) Usando 3.37 se tiene

$$\begin{aligned} \|1_{\Gamma, \mathbf{u}}\|_{\ell_\eta^\mu(\mathbf{u}, \nu)} &= \left(\int_0^{\nu(\Gamma)} [\eta(t)]^\mu \frac{dt}{t} \right)^{1/\mu} \geq \left(\int_{\nu(\Gamma)/2}^{\nu(\Gamma)} [\eta(t)]^\mu \frac{dt}{t} \right)^{1/\mu} \\ &\geq \eta(\nu(\Gamma)/2) \log 2 \geq C \eta(\nu(\Gamma)) \end{aligned}$$

ya que η es doblante. Para la desigualdad contraria, puesto que $\eta \in \mathbb{W}_+$, podemos usar la Proposición 3.4.13 y obtenemos

$$\|1_{\Gamma, \mathbf{u}}\|_{\ell_\eta^\mu(\mathbf{u}, \nu)} \approx \left(\int_0^1 [\lambda \eta(\nu(\Gamma))]^\mu \frac{d\lambda}{\lambda} \right)^{1/\mu} \approx \eta(\nu(\Gamma)).$$

■

3.4.6 Desigualdades tipo Jackson

Damos condiciones equivalentes para que algunas desigualdades tipo Jackson se verifiquen en el escenario de la aproximación no lineal restringida. Los siguientes resultados generalizan aquellos en [24] y [67] para el mismo escenario, así como los de [66] y [48] para el caso $\nu(I) = 1$ para todo $I \in \mathcal{D}$ (la medida de contar).

Empezamos con un resultado previo que es la parte 2 del Lema 2 en [67].

Lema 3.4.16 Sean F y G dos funciones no decrecientes continuas por la izquierda, tales que $F(0) = G(0) = 0$. Entonces, para cualquier sucesión $\mathbf{d} = \{d_I\}$ y $t > 0$,

$$\int_0^{\mathbf{d}_\nu^*(t)} G(\nu(\{I \in \mathcal{D} : |d_I| > \lambda\})) dF(\lambda) = F(\mathbf{d}_\nu^*(t))G(t) + \int_t^\infty F(\mathbf{d}_\nu^*(s))G(s).$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Demostración. Usando el hecho que para todo $s, \lambda \geq 0$, $1_{[0, \nu(\{I \in \mathcal{D}: |d_I| > \lambda\})]}(s) = 1_{[0, \mathbf{d}_\nu^*]}(\lambda)$ (ver 3.34), el teorema de Fubini da

$$\begin{aligned}
& \int_0^{\mathbf{d}_\nu^*(t)} G(\nu(\{I \in \mathcal{D} : |d_I| > \lambda\})) dF(\lambda) \\
&= \int_0^{\mathbf{d}_\nu^*(t)} \int_0^{\nu(\{I \in \mathcal{D}: |d_I| > \lambda\})} dG(s) dF(\lambda) \\
&= \int_0^\infty 1_{[0, \mathbf{d}_\nu^*(t)]}(\lambda) \int_0^\infty 1_{[0, \nu(\{I \in \mathcal{D}: |d_I| > \lambda\})]}(s) dG(s) dF(\lambda) \\
&= \int_0^\infty \left[\int_0^\infty 1_{[0, \mathbf{d}_\nu^*(t)]}(\lambda) 1_{[0, \mathbf{d}_\nu^*(s)]}(\lambda) dF(\lambda) \right] dG(s) \\
&= \int_0^\infty F(\min\{\mathbf{d}_\nu^*(s), \mathbf{d}_\nu^*(t)\}) dG(s) \\
&= F(\mathbf{d}_\nu^*(t))G(t) + \int_t^\infty F(\mathbf{d}_\nu^*(s)) dG(s),
\end{aligned}$$

porque $\mathbf{d}_\nu^*(t)$ es una función no creciente. ■

Teorema 3.4.17 *Sea (f, ν) un esquema canon (ver Definición 3.4.4) y sea $\mathbf{u} = \{u_I\}_{I \in \mathcal{D}}$ una sucesión de pesos. Sean $\xi > 0$ y $\mu \in (0, \infty]$ fijos. Entonces, para cualquier función $\eta \in \mathbb{W}_+$, los siguientes son equivalentes:*

1) *Existe $C > 0$ tal que para todo $\Gamma \subset \mathcal{D}$ con $\nu(\Gamma) < \infty$,*

$$\left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f \leq C\eta(\nu(\Gamma)).$$

2) $\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu) \hookrightarrow \mathcal{A}_\mu^\xi(f, \nu)$.

3) *El espacio $\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)$ satisface la desigualdad de Jackson de orden ξ , esto es, existe $C > 0$ tal que*

$$\sigma_\nu(t, \mathbf{s})_f \leq Ct^{-\xi} \|\mathbf{s}\|_{\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)}, \quad \text{para todo } \mathbf{s} \in \ell_{\xi, \eta}^\mu(\mathbf{u}, \nu) \text{ y todo } t > 0.$$

Demostración. 2) \Rightarrow 3) Esto es inmediato ya que $\mathcal{A}_\mu^\xi(\nu) \hookrightarrow \mathcal{A}_\infty^\xi(\nu)$ y 3) es equivalente a $\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu) \hookrightarrow \mathcal{A}_\infty^\xi(f, \nu)$.

3) \Rightarrow 1) Se prueba en el Lemma 1 en [67] (ver también la prueba del Teorema 2.1 en [46]) que para $\mathbf{s} = \sum_{I \in \mathcal{D}} s_I \mathbf{e}_I$,

$$\sigma_\nu(t, \mathbf{s}) \approx \tilde{\sigma}_\nu(t, \mathbf{s}) := \inf_{\nu(\Gamma) \leq t} \|\mathbf{s} - \sum_{\Gamma} s_I \mathbf{e}_I\|_f, \quad (3.38)$$

3.4 ANL restringida en espacios de sucesiones

esto es, el ínfimo que define $\sigma_\nu(t, \mathbf{s})$, puede tomarse, hasta una equivalencia, sobre los elementos de $\Sigma_{t, \nu}$ cuyos coeficientes son aquellos de \mathbf{s} . Probamos (3.38) por completitud. Claramente $\sigma_\nu(t, \mathbf{s}) \leq \tilde{\sigma}_\nu(t, \mathbf{s})$; más aún, para cualquier $t_I, I \in \mathcal{D}$ tenemos

$$\|\mathbf{s} - \sum_{I \in \Gamma} s_I \mathbf{e}_I\|_f \leq \|\mathbf{s} - \sum_{I \in \Gamma} t_I \mathbf{e}_I\|_f + \|\sum_{I \in \Gamma} s_I \mathbf{e}_I - \sum_{I \in \Gamma} t_I \mathbf{e}_I\|_f \leq 2\|\mathbf{s} - \sum_{I \in \Gamma} s_I \mathbf{e}_I\|_f.$$

por la monotonicidad de $\|\cdot\|_f$ (ver parte a) de la Definición 3.4.1.

Sea $0 < s_0 < 1$ tal que $M_\eta(s_0) < 1$ como en la Definición 3.4.10. Sea $\Gamma \subset \mathcal{D}$ con $\nu(\Gamma) < \infty$ y escribáse $1_\Gamma := 1_{\Gamma, \mathbf{u}} = \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I}$. Por (3.38), para $\Lambda_0 = \Gamma \subset \mathcal{D}$ se puede encontrar un conjunto $\Lambda_1 \subset \Lambda_0$ con $\nu(\Lambda_1) \leq s_0 \nu(\Lambda_0)$ tal que

$$\|1_{\Lambda_0} - 1_{\Lambda_1}\|_f \approx \sigma_\nu(s_0 \nu(\Lambda_0), 1_{\Lambda_0}).$$

Repetimos este argumento para encontrar subconjuntos anidados (posiblemente un número finito en algunos casos)

$$\Gamma = \Lambda_0 \supset \Lambda_1 \supset \dots \supset \Lambda_j \supset \Lambda_{j+1} \supset \dots,$$

tales que $\nu(\Lambda_{j+1}) \leq s_0 \nu(\Lambda_j)$ y

$$\|1_{\Lambda_j} - 1_{\Lambda_{j+1}}\|_f \approx \sigma_\nu(s_0 \nu(\Lambda_j), 1_{\Lambda_j}), \quad j = 0, 1, 2, \dots$$

Por la desigualdad ρ -triangular (ver Nota 3.4.3) para f se tiene

$$\|1_\Gamma\|_f^\rho \leq \sum_{j=0}^{\infty} \|1_{\Lambda_j} - 1_{\Lambda_{j+1}}\|_f^\rho \approx \sum_{j=0}^{\infty} \sigma_\nu^\rho(s_0 \nu(\Lambda_j), 1_{\Lambda_j}).$$

Usando la hipótesis y el Lema 3.4.15 obtenemos

$$\sigma_\nu(s_0 \nu(\Lambda_j), 1_{\Lambda_j}) \leq C[s_0 \nu(\Lambda_j)]^{-\xi} \|1_{\Lambda_j}\|_{\ell_{\xi, \eta}^\mu(\nu)} \approx \eta(\nu(\Lambda_j)).$$

Concatenando estas desigualdades se deduce que

$$\begin{aligned} \|1_\Gamma\|_f &\lesssim \left[\sum_{j=0}^{\infty} \eta^\rho(\nu(\Lambda_j)) \right]^{1/\rho} \lesssim \left[\sum_{j=0}^{\infty} \eta^\rho(s_0^j \nu(\Lambda_0)) \right]^{1/\rho} \\ &\lesssim \eta(\nu(\Lambda_0)) = \eta(\nu(\Gamma)) \end{aligned}$$

por el Lema 3.4.11, ya que η y η^ρ pertenecen a \mathbb{W}_+ .

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

1) \Rightarrow 2) Por el Lema 3.4.12 podemos asumir $\eta \in C^1$ y $\eta'(t)/\eta(t) \approx 1/t$, $t > 0$. Empezamos acotando $\sigma_\nu(t, \mathbf{s})$ para $\mathbf{s} \in \ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)$. Recordese que $\mathbf{d} := \{s_I u_I\}_{I \in \mathcal{D}} \in \ell_{\xi, \eta}^\mu(\nu)$. Ya que $\nu(\{I \in \mathcal{D} : |u_I s_I| > \mathbf{d}_\nu^*(t)\}) \leq t$ (ver (3.35)) se tiene

$$\sigma_\nu(t, \mathbf{s}) = \inf_{\mathbf{t} \in \Sigma_{t, \nu}} \|\mathbf{s} - \mathbf{t}\|_f \leq \left\| \sum_{|u_I s_I| \leq \mathbf{d}_\nu^*(t)} s_I \mathbf{e}_I \right\|_f.$$

Para $j = 0, 1, 2, \dots$, sea $\Lambda_j = \{I \in \mathcal{D} : 2^{-j-1} \mathbf{d}_\nu^*(t) < |s_I u_I| \leq 2^{-j} \mathbf{d}_\nu^*(t)\}$. La desigualdad ρ -triangular y la propiedad de la monotonicidad de f , junto con la hipótesis, implican

$$\begin{aligned} [\sigma_\nu(t, \mathbf{s})]^\rho &\leq \sum_{j=0}^{\infty} \left\| \sum_{I \in \Lambda_j} s_I \mathbf{e}_I \right\|_f^\rho = \sum_{j=0}^{\infty} \left\| \sum_{I \in \Lambda_j} s_I u_I \frac{\mathbf{e}_I}{u_I} \right\|_f^\rho \\ &\leq C \sum_{j=0}^{\infty} [2^{-j} \mathbf{d}_\nu^*(t)]^\rho \eta^\rho(\nu(\Lambda_j)) \\ &\leq C \int_0^{\mathbf{d}_\nu^*(t)} \lambda^\rho [\eta(\nu(\{I \in \mathcal{D} : |s_I u_I| > \lambda\}))]^\rho \frac{d\lambda}{\lambda}. \end{aligned}$$

Aplicando el Lema 3.4.16 con $F(\lambda) = \frac{\lambda^\rho}{\rho}$ y $G(\lambda) = [\eta(\lambda)]^\rho$ da

$$\begin{aligned} [\sigma_\nu(t, \mathbf{s})]^\rho &\leq \frac{1}{\rho} [\mathbf{d}_\nu^*(t) \eta(t)]^\rho + \frac{1}{\rho} \int_t^\infty [\mathbf{d}_\nu^*(s)]^\rho d\eta^\rho(s) \\ &\approx [\mathbf{d}_\nu^*(t) \eta(t)]^\rho + \int_t^\infty [\mathbf{d}_\nu^*(s)]^\rho \eta^\rho(s) \frac{ds}{s} \end{aligned}$$

donde hemos usado que $\eta'(s)/\eta(s) \approx 1/s$. Por lo tanto,

$$\sigma_\nu(t, \mathbf{s}) \lesssim \mathbf{d}_\nu^*(t) \eta(t) + \left(\int_t^\infty [\mathbf{d}_\nu^*(s) \eta(s)]^\rho \frac{ds}{s} \right)^{1/\rho}.$$

Así,

$$\begin{aligned} \|\mathbf{s}\|_{\mathcal{A}_\mu^\xi(\nu)}^\mu &= \int_0^\infty [t^\xi \sigma_\nu(t, \mathbf{s})]^\mu \frac{dt}{t} \\ &\lesssim \int_0^\infty [t^\xi \eta(t) \mathbf{d}_\nu^*(t)]^\mu \frac{dt}{t} + \int_0^\infty \left[t^\xi \left(\int_t^\infty [\mathbf{d}_\nu^*(s) \eta(s)]^\rho \frac{ds}{s} \right)^{1/\rho} \right]^\mu \frac{dt}{t} \\ &:= I + II. \end{aligned}$$

El primer término, I , es precisamente $\|\mathbf{s}\|_{\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)}^\mu$. Para II , usamos la desigualdad de Hardy (ver [8], p.124) con un ρ tal que $\mu/\rho > 1$ (nótese que esto siempre es posible

ya que si f satisface la desigualdad ρ -triangular también satisface la desigualdad ρ' -triangular para cualquier $0 < \rho' \leq \rho$) para obtener

$$\begin{aligned} II^{1/\mu} &= \left[\int_0^\infty t^{\xi\mu} \left(\int_t^\infty [\mathbf{d}_\nu^*(s)\eta(s)]^\rho \frac{ds}{s} \right)^{\mu/\rho} \frac{dt}{t} \right]^{1/\mu} \\ &\leq \left[\frac{1}{\rho\xi} \int_0^\infty [\mathbf{d}_\nu^*(s)\eta(s)]^\mu s^{\xi\mu} \frac{ds}{s} \right]^{1/\mu} \\ &= C \|\mathbf{s}\|_{\ell_{\xi,\eta}^\mu(\mathbf{u},\nu)}. \end{aligned}$$

Esto prueba el resultado. ■

Tomando $\eta(t) = t^{1/p}$, $0 < p < \infty$, y $u_I = \|\mathbf{e}_I\|_f$ en el Teorema 3.4.17, la condición 1) es llamada en [67] la propiedad p -Temlyakov (por arriba) para f . En este caso, $\ell_{\xi,\eta}^\mu(\mathbf{u},\nu) = \ell^{q,\mu}(\mathbf{u},\nu)$ con $\frac{1}{q} = \xi + \frac{1}{p}$.

Tomando ν como la medida de contar en \mathcal{D} recuperamos el Teorema 3.6 en [48] a partir del Teorema 3.4.17.

3.4.7 ANLR e interpolación real

Es bien conocido que la aproximación con N -términos y la interpolación real están interconectados. Si las desigualdades tipo Jackson y Bernstein se verifican para la medida de contar ν , los espacios de aproximación se caracterizan en términos de espacios de interpolación (ver p.e. el Teorema 3.1 en [33] o la Sección 9, Capítulo 7 de [32]).

Como se apunta en [24] la teoría arriba mencionada puede desarrollarse en un contexto más general. En nuestro caso lo desarrollamos en el esquema abstracto que hemos introducido en la Subsección 3.4.2. Las pruebas son modificaciones directas de aquellas dadas en las referencias ya citadas en el primer párrafo de esta subsección.

Teorema 3.4.18 *Sea (f,ν) un esquema canon. Supóngase que el retículo (de sucesiones) cuasi Banach $g \subset S$ satisface las desigualdades de Jackson y Bernstein para algún $r > 0$. Entonces, para $0 < \xi < r$ y $0 < \mu \leq \infty$, se tiene*

$$A_\mu^\xi(f,\nu) = (f,g)_{\xi/r,\mu},$$

donde $(f,g)_{\xi/r,\mu}$ denota el espacio de interpolación real entre f y g .

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

No es difícil mostrar que los espacios $\mathcal{A}_q^r(f, \nu)$, $0 < r < \infty$, $0 < q \leq \infty$, satisfacen las desigualdades de Jackson y Bernstein de orden r , así que por el Teorema 3.4.18,

$$\mathcal{A}_\mu^\xi(f, \nu) = (f, \mathcal{A}_q^r(f, \nu))_{\xi/r, \mu},$$

para $0 < \xi < r$ y $0 < \mu \leq \infty$. A partir de esto y usando el teorema de la reiteración para interpolación real obtenemos el siguiente resultado que se usará en la prueba del Teorema 3.4.21.

Corolario 3.4.19 *Sea $0 < \alpha_0, \alpha_1 < \infty$, $0 < q, q_0, q_1 \leq \infty$ y $0 < \theta < 1$. Entonces,*

$$(\mathcal{A}_{q_0}^{\alpha_0}(f, \nu), \mathcal{A}_{q_1}^{\alpha_1}(f, \nu))_{\theta, q} = \mathcal{A}_q^\alpha(f, \nu), \quad \alpha = (1 - \theta)\alpha_0 + \theta\alpha_1,$$

para un esquema canon (f, ν) .

3.4.8 Desigualdades tipo Bernstein

Damos condiciones equivalentes para que algunas desigualdades tipo Bernstein se verifiquen en el escenario de la aproximación no lineal restringida. Este resultado generaliza aquellos obtenidos en [24] y [67] para el mismo escenario, así como aquellos obtenidos en [66] y [48] para el caso $\nu(I) = 1$ para todo $I \in \mathcal{D}$ (la medida de contar).

Empezamos con un teorema de representación para los espacios $\mathcal{A}_\mu^\xi(f, \nu)$. La demostración se sigue de aquella en [89] reemplazando la medida de contar por una medida positiva ν .

Proposición 3.4.20 *Sea (f, ν) un esquema canon (ver Definición 3.4.4). Sean $\xi > 0$ y $\mu \in (0, \infty]$. Las siguientes son equivalentes:*

i) $\mathbf{s} \in \mathcal{A}_\mu^\xi(f, \nu)$.

ii) Existe $\mathbf{s}_k \in \sum_{2^k, \nu} \cap f$, $k \in \mathbb{Z}$, tal que $\mathbf{s} = \sum_{k=-\infty}^{\infty} \mathbf{s}_k$ y $\{2^{k\xi} \|\mathbf{s}_k\|_f\}_{k \in \mathbb{Z}} \in \ell^\mu(\mathbb{Z})$.

Más aún,

$$\|\mathbf{s}\|_{\mathcal{A}_\mu^\xi(f, \nu)} \approx \inf \left\{ \left[\sum_{k=-\infty}^{\infty} (2^{k\xi} \|\mathbf{s}_k\|_f)^\mu \right]^{1/\mu} \right\},$$

donde el ínfimo se toma sobre todas las representaciones de \mathbf{s} como en ii).

3.4 ANL restringida en espacios de sucesiones

Demostración. i) \Rightarrow ii) Sea $\mathbf{s} \in \mathcal{A}_\mu^\xi(f, \nu)$. Escójase $\varphi_k \in \Sigma_{2^{k-1}, \nu}$ tal que $\|\mathbf{s} - \varphi_k\|_f \leq 2\sigma_\nu(2^{k-1}, \mathbf{s})$. Sea $\mathbf{s}_k = \varphi_k - \varphi_{k-1}$, de manera que $\mathbf{s}_k \in \Sigma_{2^k, \nu}$. Ya que $\mathbf{s} \in \mathcal{A}_\mu^\xi(f, \nu)$, se tiene $\sigma_\nu(2^{k-1}, \mathbf{s}) \rightarrow 0$ cuando $k \rightarrow \infty$; la suposición $f \hookrightarrow S$ implica

$$\lim_{k \rightarrow \infty} \varphi_k = \mathbf{s} \quad \text{in } \mathcal{D} \text{ (término a término).}$$

Por otra parte $\lim_{k \rightarrow -\infty} \varphi_k = 0$ ya que $\nu(\text{supp } \varphi_k) \rightarrow 0$ cuando $k \rightarrow -\infty$. Así,

$$\mathbf{s} = \lim_{k \rightarrow \infty} \varphi_k = \sum_{k=-\infty}^{\infty} \mathbf{s}_k.$$

Ahora,

$$\|\mathbf{s}_k\|_f^\rho \leq \|\mathbf{s} - \varphi_k\|_f^\rho + \|\mathbf{s} - \varphi_{k-1}\|_f^\rho \leq 2 \cdot 2^\rho [\sigma_\nu(2^{k-2}, \mathbf{s})]^\rho.$$

Por lo tanto,

$$\sum_{k \in \mathbb{Z}} [2^{k\xi} \|\mathbf{s}_k\|_f]^\mu \leq C \sum_{k \in \mathbb{Z}} [2^{k\xi} \sigma_\nu(2^{k-2}, \mathbf{s})]^\mu \approx \|\mathbf{s}\|_{\mathcal{A}_\mu^\xi(\nu)}^\mu,$$

por la caracterización discreta de los espacios de aproximación restringidos dados en la Subsección 3.4.2. Es fácil ver que el resultado también se verifica para $\mu = \infty$.

ii) \Rightarrow i) Obsérvese que $\sum_{k=-\infty}^{\ell-1} \mathbf{s}_k \in \Sigma_{2^\ell, \nu}$ ya que cada $\mathbf{s}_k \in \Sigma_{2^k, \nu}$. Tómesese ρ tal que $0 < \rho < \mu$ y $\|\cdot\|_f$ satisfaga la desigualdad ρ -triangular. Se tiene

$$[\sigma_\nu(2^\ell, \mathbf{s})]^\rho \leq \left\| \mathbf{s} - \sum_{k=-\infty}^{\ell-1} \mathbf{s}_k \right\|_f^\rho \leq \sum_{k=\ell}^{\infty} \|\mathbf{s}_k\|_f^\rho.$$

Con $p = \mu/\rho > 1$, (aquí $0 < \mu < \infty$) y $u > 0$ tal que $u < \xi\rho$ se tiene

$$\begin{aligned} \|\mathbf{s}\|_{\mathcal{A}_\mu^\xi(\nu)}^\mu &\approx \sum_{\ell=-\infty}^{\infty} [2^{\ell\xi} \sigma_\nu(2^\ell, \mathbf{s})]^\mu \leq \sum_{\ell=-\infty}^{\infty} [2^{\ell\xi} (\sum_{k=\ell}^{\infty} \|\mathbf{s}_k\|_f^\rho)^{1/\rho}]^\mu \\ &= \sum_{\ell=-\infty}^{\infty} 2^{\ell\xi\mu} \left(\sum_{k=\ell}^{\infty} 2^{-ku} 2^{ku} \|\mathbf{s}_k\|_f^\rho \right)^p \\ &\leq \sum_{\ell=-\infty}^{\infty} 2^{\ell\xi\mu} \left(\sum_{k=\ell}^{\infty} 2^{-kup'} \right)^{p/p'} \left(\sum_{k=\ell}^{\infty} 2^{kup} \|\mathbf{s}_k\|_f^\mu \right) \\ &\approx \sum_{\ell \in \mathbb{Z}} 2^{\ell(\xi\mu - up)} \left(\sum_{k=\ell}^{\infty} 2^{kup} \|\mathbf{s}_k\|_f^\mu \right) \\ &= \sum_{k \in \mathbb{Z}} 2^{kup} \|\mathbf{s}_k\|_f^\mu \left(\sum_{\ell=-\infty}^k 2^{\ell(\xi\mu - up)} \right) \approx \sum_{k \in \mathbb{Z}} 2^{k\xi\mu} \|\mathbf{s}_k\|_f^\mu. \end{aligned}$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Ya que la última expresión es finita por hipótesis, hemos probado $\mathbf{s} \in \mathcal{A}_\mu^\xi(f, \nu)$ para $0 < \mu < \infty$. Para $\mu = \infty$ se tiene

$$\begin{aligned} \|\mathbf{s}\|_{\mathcal{A}_\infty^\xi(\nu)} &\approx \sup_{\ell \in \mathbb{Z}} 2^{\ell\xi} \sigma_\nu(2^\ell, \mathbf{s}) \leq \sup_{\ell \in \mathbb{Z}} 2^{\ell\xi} \left(\sum_{k=\ell}^{\infty} \|\mathbf{s}_k\|_f^\rho \right)^{1/\rho} \\ &= \sup_{\ell \in \mathbb{Z}} 2^{\ell\xi} \left(\sum_{k=\ell}^{\infty} 2^{-k\xi\rho} 2^{k\xi\rho} \|\mathbf{s}_k\|_f^\rho \right)^{1/\rho} \\ &\leq \left(\sup_{k \in \mathbb{Z}} 2^{k\xi} \|\mathbf{s}_k\|_f \right) \sup_{\ell \in \mathbb{Z}} 2^{\ell\xi} \left(\sum_{k=\ell}^{\infty} 2^{-k\xi\rho} \right)^{1/\rho} \approx \sup_{k \in \mathbb{Z}} 2^{k\xi} \|\mathbf{s}_k\|_f. \end{aligned}$$

■

Teorema 3.4.21 *Sea (f, ν) un esquema canon (ver Definición 3.4.4) y sea $\mathbf{u} = \{u_I\}_{I \in \mathcal{D}}$ una sucesión de pesos. Sean $\xi > 0$ y $\mu \in (0, \infty]$ fijos. Entonces, para cualquier función $\eta \in \mathbb{W}$, las siguientes son equivalentes:*

1) *Existe $C > 0$ tal que para todo $\Gamma \subset \mathcal{D}$ con $\nu(\Gamma) < \infty$,*

$$\frac{1}{C} \eta(\nu(\Gamma)) \leq \left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f.$$

2) *El espacio $\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)$ satisface la desigualdad de Bernstein de orden ξ , esto es, existe $C > 0$ tal que*

$$\|\mathbf{s}\|_{\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)} \leq C t^\xi \|\mathbf{s}\|_f, \quad \text{para todo } \mathbf{s} \in \Sigma_{t, \nu} \cap f \text{ y todo } t > 0.$$

3) $\mathcal{A}_\mu^\xi(f, \nu) \leftrightarrow \ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)$.

Demostración. 1) \Rightarrow 2) Sea $\mathbf{s} \in \Sigma_{t, \nu} \cap f$ y escríbase $\mathbf{s} = \{s_I\}_{I \in \mathcal{D}}$. Dado $0 < \tau \leq \nu(\Gamma)$ escójase $\Lambda_\tau = \{I \in \Gamma : |u_I s_I| \geq \mathbf{d}_\nu^*(\tau)\}$ donde $\mathbf{d} = \{u_I s_I\}_{I \in \mathcal{D}}$. Se tiene $\tau \leq \nu(\Lambda_\tau)$ (ver (3.35)). Aplicando la hipótesis 1) y la monotonicidad de f se obtiene

$$\mathbf{d}_\nu^*(\tau) \eta(\tau) \leq \mathbf{d}_\nu^*(\tau) \eta(\nu(\Lambda_\tau)) \lesssim \mathbf{d}_\nu^*(\tau) \left\| \sum_{I \in \Lambda_\tau} \frac{\mathbf{e}_I}{u_I} \right\|_f \leq \left\| \sum_{I \in \Lambda_\tau} s_I \mathbf{e}_I \right\|_f \lesssim \|\mathbf{s}\|_f.$$

Ya que $\nu(\{I \in \mathcal{D} : |s_I u_I| > 0\}) = \nu(\Gamma) \leq t$ se tiene $\mathbf{d}_\nu^*(\tau) = 0$ para $\tau \geq t$. así,

$$\|\mathbf{s}\|_{\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)}^\mu = \int_0^t [\tau^\xi \eta(\tau) \mathbf{d}_\nu^*(\tau)]^\mu \frac{d\tau}{\tau} \lesssim \|\mathbf{s}\|_f^\mu \int_0^t \tau^{\xi\mu} \frac{d\tau}{\tau} \approx t^{\xi\mu} \|\mathbf{s}\|_f^\mu.$$

El caso $\mu = \infty$ es similar.

3.4 ANL restringida en espacios de sucesiones

2) \Rightarrow 1) Sea $1_\Gamma := 1_{\Gamma, \mathbf{u}} = \sum_{I \in \Gamma} \frac{e_I}{u_I}$ y $\nu(\Gamma) = t$, de manera que $1_\Gamma \in \Sigma_{t, \nu}$. Podemos asumir $1_\Gamma \in f$, ya que de otra forma el lado derecho de 1) es ∞ , y el resultado es trivialmente cierto. La hipótesis 2) da

$$\|1_\Gamma\|_f \gtrsim t^{-\xi} \|1_\Gamma\|_{\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)} \gtrsim \eta(\nu(\Gamma)),$$

donde la última desigualdad se debe al Lema 3.4.15.

3) \Rightarrow 2) Para $\mathbf{s} \in \Sigma_{t, \nu} \cap f$, $\sigma_\nu(\tau, \mathbf{s}) = 0$ si $\tau \geq t$. Así, por 3)

$$\begin{aligned} \|\mathbf{s}\|_{\ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)} &\lesssim \|\mathbf{s}\|_{\mathcal{A}_{\tilde{\mu}}^\xi(\nu)} = \left(\int_0^t [\tau^\xi \sigma_\nu(\tau, \mathbf{s})]^\mu \frac{d\tau}{\tau} \right)^{1/\mu} \\ &\lesssim \|\mathbf{s}\|_f \left(\int_0^t \tau^{\xi\mu} \frac{d\tau}{\tau} \right)^{1/\mu} = t^\xi \|\mathbf{s}\|_f, \end{aligned}$$

donde hemos usado $\sigma_\nu(\tau, \mathbf{s}) \leq \|\mathbf{s}\|_f$ para todo $\tau > 0$.

2) \Rightarrow 3) Ya hemos probado que 1) \Leftrightarrow 2); ya que 1) no depende de ξ, μ , entonces 2) se verifica para todo $\tilde{\xi} > 0$ y todo $\mu \in (0, \infty]$. Para cualquier $\tilde{\xi} > 0$ tómesese $\tilde{\rho}$ tal que $\ell_{\tilde{\xi}, \eta}^\mu(\mathbf{u}, \nu)$ satisface la desigualdad $\tilde{\rho}$ -triangular. Por la Proposición 3.4.20 podemos encontrar $\mathbf{s}_k \in \Sigma_{2^k, \nu} \cap f$, $k \in \mathbb{Z}$, tal que $\mathbf{s} = \sum_{k \in \mathbb{Z}} \mathbf{s}_k$ (in \mathcal{D}) y

$$\|\mathbf{s}\|_{\mathcal{A}_{\tilde{\rho}}^{\tilde{\xi}}(\nu)} \approx \left(\sum_{k \in \mathbb{Z}} [2^{k\tilde{\xi}} \|\mathbf{s}_k\|_f]^{\tilde{\rho}} \right)^{1/\tilde{\rho}}.$$

Aplicando la hipótesis 2) a $\ell_{\tilde{\xi}, \eta}^\mu(\mathbf{u}, \nu)$ obtenemos

$$\|\mathbf{s}\|_{\ell_{\tilde{\xi}, \eta}^{\tilde{\rho}}(\mathbf{u}, \nu)} \lesssim \sum_{k \in \mathbb{Z}} [\|\mathbf{s}_k\|_{\ell_{\tilde{\xi}, \eta}^\mu(\mathbf{u}, \nu)}]^{\tilde{\rho}} \lesssim \sum_{k \in \mathbb{Z}} (2^{k\tilde{\xi}} \|\mathbf{s}_k\|_f)^{\tilde{\rho}} \approx \|\mathbf{s}\|_{\mathcal{A}_{\tilde{\rho}}^{\tilde{\xi}}(\nu)}.$$

Esto significa que para $\mu \in (0, \infty]$ y cualquier $\tilde{\xi} > 0$ tenemos la inclusión continua

$$\mathcal{A}_{\tilde{\rho}}^{\tilde{\xi}}(\nu) \hookrightarrow \ell_{\tilde{\xi}, \eta}^\mu(\mathbf{u}, \nu), \quad (3.39)$$

donde $\tilde{\rho}$ es el exponente de la desigualdad $\tilde{\rho}$ -triangular para $\ell_{\tilde{\xi}, \eta}^\mu(\mathbf{u}, \nu)$. Del Corolario 3.4.19, para $\xi = (\xi_0 + \xi_1)/2$ y cualquier $\rho \in (0, 1]$ se tiene

$$(\mathcal{A}_{\rho}^{\xi_0}(\nu), \mathcal{A}_{\rho}^{\xi_1}(\nu))_{1/2, \mu} = \mathcal{A}_{\mu}^{\xi}(\nu).$$

Aplicando primero (3.39) con $\tilde{\xi} = \xi_0$, y después con $\tilde{\xi} = \xi_1$ y $\rho = \min\{\tilde{\rho}_0, \tilde{\rho}_1\}$ se obtiene

$$\mathcal{A}_{\mu}^{\xi}(\nu) = (\mathcal{A}_{\rho}^{\xi_0}(\nu), \mathcal{A}_{\rho}^{\xi_1}(\nu))_{1/2, \mu} \hookrightarrow (\ell_{\xi_0, \eta}^\mu(\mathbf{u}, \nu), \ell_{\xi_1, \eta}^\mu(\mathbf{u}, \nu))_{1/2, \mu} = \ell_{\xi, \eta}^\mu(\mathbf{u}, \nu)$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

donde la última igualdad es un resultado en interpolación real de espacios discretos de Lorentz que puede encontrarse en [80] (Teorema 3). ■

Tomando $\eta(t) = t^{1/p}$, $0 < p < \infty$, y $u_I = \|\mathbf{e}_I\|_f$ en el Teorema 3.4.21, la condición 1) es llamada en [67] la propiedad p -Temlyakov (por abajo) para f . En este caso, $\ell_{\xi,\eta}^\mu(\mathbf{u}, \nu) = \ell^{q,\mu}(\mathbf{u}, \nu)$ con $\frac{1}{q} = \xi + \frac{1}{p}$.

El Teorema 3.4.21 generaliza el Teorema 5 en [67] para un esquema canon. Para la demostración del Teorema 3.4.21 hemos hecho uso de la teoría de interpolación real (ver Subsección 3.4.7) para acortarla. Una prueba alternativa es como la de [67], en la cual no usan interpolación real.

Tomando ν como la medida de contar en \mathcal{D} recuperamos el Teorema 4.2 en [48] a partir del Teorema 3.4.21.

los Teoremas 3.4.17 y 3.4.21 permiten obtener el siguiente resultado que da una caracterización de los espacios de aproximación como espacios de Lorentz discretos con pesos, siempre que se cumpla la propiedad de Temlyakov generalizada.

Teorema 3.4.22 *Sea (f, ν) un esquema canon (ver Definición 3.4.4) y sea $\mathbf{u} = \{u_I\}_{I \in \mathcal{D}}$ una sucesión de pesos. Entonces, para cualquier $\eta \in \mathbb{W}_+$, las siguientes condiciones son equivalentes:*

1. Existe $C > 0$ tal que para todo $\Gamma \subset \mathcal{D}$ con $\nu(\Gamma) < \infty$,

$$\frac{1}{C}\eta(\nu(\Gamma)) \leq \left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f \leq C\eta(\nu(\Gamma)).$$

2. $\mathcal{A}_\mu^\xi(f, \nu) = \ell_{\xi,\eta}^\mu(\mathbf{u}, \nu)$ para todo $\xi > 0$, $\mu \in (0, \infty]$.

3.4.9 Espacios de sucesiones asociados a espacios de suavidad

Un gran número de espacios usados en análisis son casos particulares de los espacios de Triebel-Lizorkin y de Besov.

Dados $s \in \mathbb{R}$, $0 < p < \infty$, y $0 < r \leq \infty$, los espacios de Triebel-Lizorkin en \mathbb{R}^d se denotan por $F_{p,r}^s := F_{p,r}^s(\mathbb{R}^d)$ donde s es un parámetro de suavidad, p mide la integrabilidad y r mide un refinamiento de la suavidad. Las definiciones de estos

3.4 ANL restringida en espacios de sucesiones

espacios se pueden encontrar en [41] y [42]. Los espacios de Lebesgue $L^p(\mathbb{R}^d) = F_{p,2}^0$, $1 < p < \infty$, los espacios de Hardy $H^p(\mathbb{R}^d) = F_{p,2}^0$, $0 < p \leq 1$, y los espacios de Sobolev $W_p^s(\mathbb{R}^d) = F_{p,2}^s$, $s > 0$, $1 < p < \infty$, están incluidos en esta colección.

Dados $s \in \mathbb{R}^d$, $0 < p, r \leq \infty$, los espacios de Besov en \mathbb{R}^d se denotan $B_{p,r}^s := B_{p,r}^s(\mathbb{R}^d)$, con una interpretación de los parámetros similar a la de los espacios de Triebel-Lizorkin. Estos espacios incluyen las clases de Lipschitz (ver [42]).

Existen caracterizaciones de $F_{p,r}^s$ y $B_{p,r}^s$ en términos de espacios de sucesiones. Tales caracterizaciones fueron dadas primeramente en [41] usando la transformada φ . Las bases de ondículas con regularidad y condiciones de momentos apropiados también proveen dichas caracterizaciones.

A continuación se da una breve descripción de las bases de ondículas en \mathbb{R}^d . Sea \mathcal{D} el conjunto de cubos diádicos en \mathbb{R}^d definidos por

$$Q_{j,k} = 2^{-j}([0, 1]^d + k), \quad j \in \mathbb{Z}, k \in \mathbb{Z}^d.$$

Una colección finita de funciones $\Psi = \{\psi^{(1)}, \dots, \psi^{(L)}\} \subset L^2(\mathbb{R}^d)$ con $L = 2^d - 1$ es una familia de ondículas (ortonormales) si el conjunto

$$\mathcal{W} := \{\psi_{Q_{j,k}}^{(\ell)}(x) := 2^{\frac{j d}{2}} \psi^{(\ell)}(2^j x - k) : Q_{j,k} \in \mathcal{D}, \ell = 1, 2, \dots, L\} \quad (3.40)$$

es una base ortonormal de $L^2(\mathbb{R}^d)$. Esta es la definición que aparece en [72]. Se pueden consultar las propiedades de las ondículas en [81], [28], [59], [77].

Definición 3.4.23 *Dados $s \in \mathbb{R}$, $0 < p < \infty$ y $0 < r \leq \infty$, sea $f_{p,r}^s$ el espacio de sucesiones $\mathbf{s} = \{s_Q\}_{Q \in \mathcal{D}}$ tales que*

$$\|\mathbf{s}\|_{f_{p,r}^s} := \left\| \left[\sum_{Q \in \mathcal{D}} (|Q|^{-s/d+1/p-1/2} |s_Q| \chi_Q^{(p)}(\cdot))^r \right]^{1/r} \right\|_{L^p(\mathbb{R}^d)} < \infty,$$

donde $\chi_Q^{(p)}(\cdot) = \chi_Q(\cdot) |Q|^{-1/p}$ y $\chi_Q(\cdot)$ denota la función característica de Q .

Definición 3.4.24 *Dados $s \in \mathbb{R}$, $0 < p, r \leq \infty$, sea $b_{p,r}^s$ el espacio de sucesiones $\mathbf{s} = \{s_Q\}_{Q \in \mathcal{D}}$ tal que*

$$\|\mathbf{s}\|_{b_{p,r}^s} := \left[\sum_{j \in \mathbb{Z}} \left(\sum_{|Q|=2^{-j d}} (|Q|^{-s/d+1/p-1/2} |s_Q|)^p \right)^{r/p} \right]^{1/r} < \infty,$$

con las modificaciones obvias cuando $p, r = \infty$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Con las condiciones apropiadas en los elementos de la familia de ondículas $\Psi = \{\psi^{(1)}, \dots, \psi^{(L)}\}$, ($L = 2^d - 1$), puede mostrarse que \mathcal{W} es una base incondicional de $F_{p,r}^s$ o $B_{p,r}^s$ y si $f = \sum_{\ell=1}^L \sum_{Q \in \mathcal{D}} s_Q^\ell \psi_Q^{(\ell)}$, entonces

$$\|f\|_{F_{p,r}^s} \approx \sum_{\ell=1}^L \left\| \{s_Q^\ell\}_{Q \in \mathcal{D}} \right\|_{f_{p,r}^s} \quad \text{y} \quad \|f\|_{B_{p,r}^s} \approx \sum_{\ell=1}^L \left\| \{s_Q^\ell\}_{Q \in \mathcal{D}} \right\|_{b_{p,r}^s}. \quad (3.41)$$

Condiciones sobre Ψ para que estas equivalencias se verifiquen pueden encontrarse en [81], [59], [2], [70], [72]. Cuando una familia de ondículas Ψ provee una base incondicional para $F_{p,r}^s$ o $B_{p,r}^s$, con equivalencias como en (3.41), diremos que Ψ es **admisibile** para $F_{p,r}^s$ o $B_{p,r}^s$, respectivamente.

Las equivalencias (3.41) nos permiten trabajar al nivel de las sucesiones. Obviaremos la suma sobre ℓ ya que sólo cambia las constantes en los cálculos posteriores. Los resultados probados para los espacios de sucesiones $f_{p,r}^s$ o $b_{p,r}^s$, pueden transferirse al los espacios $F_{p,q}^s$ o $B_{p,r}^s$, a través del marco de transferencia abstracto desarrollado en [46].

Hacemos notar que los espacios de Triebel-Lizorkin y de Besov caracterizados como en (3.41) son llamados **homogéneos** y se denotan frecuentemente por $\dot{F}_{p,r}^s$ and $\dot{B}_{p,r}^s$. El caso inhomogéneo requiere pequeñas modificaciones. También el caso de espacios anisotrópicos como aquellos considerados en [46] y [70] requieren modificaciones menores, incluso aquellos espacios definidos en dominios acotados. Nos restringiremos a los casos caracterizados por (3.41).

3.4.10 ANLR para espacios de sucesiones de Triebel-Lizorkin

Como consecuencias de los teoremas desarrollados en las Subsecciones 3.4.6 y 3.4.8 obtendremos resultados para la aproximación restringida en espacios de sucesiones de Triebel-Lizorkin. Cuando son acoplados con el marco de transferencia abstracto desarrollado en [46], nuestros resultados generalizan aquellos en [24] y, con modificaciones menores, aquellos en [61].

Lema 3.4.25 *Sea $\Gamma \subset \mathcal{D}$ (no necesariamente finito), $x \in \cup_{Q \in \Gamma} Q$, y $\gamma \neq 0$. Defínase*

$$S_\Gamma^\gamma(x) = \sum_{Q \in \Gamma} |Q|^\gamma \chi_Q(x).$$

3.4 ANL restringida en espacios de sucesiones

i) Si $\gamma > 0$ y existe Q^x , el mayor cubo en Γ que contiene a x , entonces $S_\Gamma^\gamma(x) \approx |Q^x|^\gamma \chi_{Q^x}(x) = |Q^x|^\gamma$.

ii) Si $\gamma < 0$ y existe Q_x , el menor cubo en Γ que contiene a x , entonces $S_\Gamma^\gamma(x) \approx |Q_x|^\gamma \chi_{Q_x}(x) = |Q_x|^\gamma$.

Demostración. i) Es claro que $|Q^x|^\gamma \chi_{Q^x}(x) \leq S_\Gamma^\gamma(x)$ ya que el lado derecho de esta desigualdad contiene al menos el cubo Q^x (y posiblemente más). Para la desigualdad contraria, agrandamos la suma que define $S_\Gamma^\gamma(x)$ para incluir todos los cubos diádicos contenidos en Q^x . Por lo tanto,

$$S_\Gamma^\gamma(x) \leq \sum_{Q \subset Q^x: Q \in \mathcal{D}} |Q|^\gamma = \sum_{j=0}^{\infty} (2^{-jd} |Q^x|)^\gamma = |Q^x|^\gamma \sum_{j=0}^{\infty} 2^{-jd\gamma} \approx |Q^x|^\gamma$$

ya que $\gamma > 0$.

ii) Es claro que $|Q_x|^\gamma \chi_{Q_x}(x) \leq S_\Gamma^\gamma(x)$ ya que el lado derecho de esta desigualdad contiene al menos el cubo Q_x (y posiblemente más). Para la desigualdad contraria, agrandamos la suma definiendo $S_\Gamma^\gamma(x)$ para incluir todos los cubos que contienen a Q_x . Así,

$$S_\Gamma^\gamma(x) \leq \sum_{Q \supset Q_x: Q \in \mathcal{D}} |Q|^\gamma = \sum_{j=0}^{\infty} (2^{jd} |Q_x|)^\gamma = |Q_x|^\gamma \sum_{j=0}^{\infty} 2^{jd\gamma} \approx |Q_x|^\gamma,$$

ya que $\gamma < 0$. ■

El menor cubo Q_x de Γ que contiene a $x \in \cup_{Q \in \Gamma} Q$ ha sido usado por otros autores en el contexto de la aproximación no lineal con bases de ondículas (ver [60], [24], [46], [47]). Hasta donde sabemos, el mayor cubo Q^x de Γ que contiene $x \in \cup_{Q \in \Gamma} Q$ no ha sido usado antes.

Teorema 3.4.26 Sea $s_1, s_2 \in \mathbb{R}$, $0 < p_1, p_2 < \infty$, $0 < q_1, q_2 \leq \infty$. Para $\alpha \in \mathbb{R}$ y $\Gamma \subset \mathcal{D}$ defínase $\nu_\alpha(\Gamma) = \sum_{Q \in \Gamma} |Q|^\alpha$. Además, sea $\nu_\alpha(\Gamma) < \infty$. Entonces,

$$\left\| \sum_{Q \in \Gamma} \frac{\mathbf{e}_Q}{\|\mathbf{e}_Q\|_{f_{p_2, q_2}^{s_2}}} \right\|_{f_{p_1, q_1}^{s_1}} \approx [\nu_\alpha(\Gamma)]^{1/p_1}, \quad (3.42)$$

si y sólo si $\alpha \neq 1$ y $\alpha = p_1 \left(\frac{s_2 - s_1}{d} - \frac{1}{p_2} \right) + 1$ o $\alpha = 1$, $\frac{s_2 - s_1}{d} = \frac{1}{p_2}$ y $p_1 = q_1$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Demostración. Empezamos por probar (3.42). Escribimos $f_1 := f_{p_1, q_1}^{s_1}$ y $f_2 := f_{p_2, q_2}^{s_2}$ para simplificar la notación en esta prueba. Por la Definición 3.4.23 tenemos $\|e_Q\|_{f_2} = |Q|^{-s_2/d+1/p_2-1/2}$ y

$$\begin{aligned} \left\| \sum_{Q \in \Gamma} \frac{e_Q}{\|e_Q\|_{f_2}} \right\|_{f_1} &= \left(\int_{\mathbb{R}^d} \left[\sum_{Q \in \Gamma} |Q|^{\frac{s_2-s_1}{d} q_1} |Q|^{-q_1/p_2} \chi_Q(x) \right]^{p_1/q_1} dx \right)^{1/p_1} \\ &= \left(\int_{\mathbb{R}^d} \left[\sum_{Q \in \Gamma} (|Q|^{\frac{\alpha-1}{p_1}} \chi_Q(x))^{q_1} \right]^{p_1/q_1} dx \right)^{1/p_1}. \end{aligned} \quad (3.43)$$

Considérese primero el caso $\alpha > 1$. Ya que $\nu_\alpha(\Gamma) < \infty$, el mayor Q^x contenido en Γ existe para todo $x \in \cup_{Q \in \Gamma} Q$. Aplicando el Lema 3.4.25, parte i), primero con $\gamma = \frac{\alpha-1}{p_1} q_1 > 0$ y después con $\gamma = \alpha - 1 > 0$, obtenemos

$$\left[\sum_{Q \in \Gamma} |Q|^{\frac{\alpha-1}{p_1} q_1} \chi_Q(x) \right]^{p_1/q_1} \approx |Q^x|^{\alpha-1} \chi_{Q^x}(x) \approx \sum_{Q \in \Gamma} |Q|^{\alpha-1} \chi_Q(x),$$

para todo $x \in \cup_{Q \in \Gamma} Q$. A partir de (3.43) deducimos

$$\begin{aligned} \left\| \sum_{Q \in \Gamma} \frac{e_Q}{\|e_Q\|_{f_2}} \right\|_{f_1} &\approx \left(\int_{\mathbb{R}^d} \sum_{Q \in \Gamma} |Q|^{\alpha-1} \chi_Q(x) dx \right)^{1/p_1} \\ &= \left(\sum_{Q \in \Gamma} |Q|^\alpha \right)^{1/p_1} = [\nu_\alpha(\Gamma)]^{1/p_1}. \end{aligned}$$

Considérese ahora $\alpha < 1$. Si $\alpha \leq 0$, ya que $\nu_\alpha(\Gamma) < \infty$, el cubo menor Q_x contenido en Γ existe para todo $x \in \cup_{Q \in \Gamma} Q$ (nótese que $\alpha = 0$ es el caso clásico de la medida de contar). Si $0 < \alpha < 1$, se puede mostrar que el conjunto E_α de todos los $x \in \cup_{Q \in \Gamma} Q$ para los cuales Q_x no existe tiene medida cero. Para ver esto, escribáse $\mathcal{D}_k = \{Q \in \mathcal{D} : |Q| = 2^{-kd}, k \in \mathbb{Z}\}$. Entonces, para todo $m \geq 0$, $E_\alpha \subset \cup_{k \geq m} \cup_{Q \in \Gamma \cap \mathcal{D}_k} Q$; por lo tanto

$$\begin{aligned} |E_\alpha| &\leq \sum_{k \geq m} \sum_{Q \in \Gamma \cap \mathcal{D}_k} |Q| = \sum_{k \geq m} \sum_{Q \in \Gamma \cap \mathcal{D}_k} |Q|^\alpha |Q|^{1-\alpha} \\ &\leq \nu_\alpha(\Gamma) \sum_{k \geq m} 2^{-kd(1-\alpha)} \approx \nu_\alpha(\Gamma) 2^{-md(1-\alpha)} \end{aligned}$$

ya que $1 - \alpha > 0$. Dejando que $m \rightarrow \infty$ se deduce $|E_\alpha| = 0$.

3.4 ANL restringida en espacios de sucesiones

Aplíquese el Lema 3.4.25, parte ii), primero con $\gamma = \frac{\alpha-1}{p_1}q_1 < 0$ y después con $\gamma = \alpha - 1 < 0$, para obtener

$$\left[\sum_{Q \in \Gamma} |Q|^{\frac{\alpha-1}{p_1}q_1} \chi_Q(x) \right]^{p_1/q_1} \approx |Q_x|^{\alpha-1} \chi_{Q_x}(x) \approx \sum_{Q \in \Gamma} |Q|^{\alpha-1} \chi_Q(x)$$

para todo $x \in \cup_{Q \in \Gamma} Q$ si $\alpha \leq 0$ y todo $x \in \cup_{Q \in \Gamma} Q \setminus E_\alpha$ si $0 < \alpha < 1$. En cualquier caso, de (3.43) se deduce

$$\left\| \sum_{Q \in \Gamma} \frac{\mathbf{e}_Q}{\|\mathbf{e}_Q\|_{f_2}} \right\|_{f_1} \approx \left(\int_{\mathbb{R}^d} \sum_{Q \in \Gamma} |Q|^{\alpha-1} \chi_Q(x) dx \right)^{1/p_1} = \left(\sum_{Q \in \Gamma} |Q|^\alpha \right)^{1/p_1} = [\nu_\alpha(\Gamma)]^{1/p_1}.$$

Para $\alpha = 1$ el conjunto E_1 de todos los $x \in \cup_{Q \in \Gamma} Q$ para los cuales Q_x existe tiene también medida cero. De hecho

$$|E_1| \leq \sum_{k \geq m} \sum_{Q \in \Gamma \cap \mathcal{D}_k} |Q| = \sum_{k \geq m} \nu_1(\Gamma \cap \mathcal{D}_k)$$

y la última suma tiende a 0 cuando $m \rightarrow \infty$ ya que es la cola de la suma convergente $\sum_{k \geq 1} \nu_1(\Gamma \cap \mathcal{D}_k) \leq \nu_1(\Gamma) < \infty$.

Supóngase ahora que (3.42) se verifica. Para $N \in \mathbb{N}$ y $L = 2^l$ considere el conjunto $\Gamma_{N,L} = \{[0, L]^d + Lj : j \in \mathbb{N}^d, 0 \leq |j| < N\}$ de N^d cubos diádicos disjuntos de lado L . Para esta colección tenemos

$$\nu_\alpha(\Gamma_{N,L}) = \sum_{\Gamma_{N,L}} |Q|^\alpha = (L^\alpha N)^d. \quad (3.44)$$

También,

$$\left\| \sum_{Q \in \Gamma_{N,L}} \frac{\mathbf{e}_Q}{\|\mathbf{e}_Q\|_{f_2}} \right\|_{f_1} = \left(\int_{\mathbb{R}^d} [S_{\Gamma_{N,L}}^\gamma(x)]^{p_1/q_1} dx \right)^{1/p_1},$$

con $\gamma = (\frac{s_2-s_1}{d} - \frac{1}{p_2})q_1$. Ya que $S_{\Gamma_{N,L}}^\gamma(x) = L^{d\gamma} \sum_{\Gamma_{N,L}} \chi_Q(x) = L^{d\gamma} \chi_{[0, NL]^d}(x)$, se obtiene

$$\left\| \sum_{Q \in \Gamma_{N,L}} \frac{\mathbf{e}_Q}{\|\mathbf{e}_Q\|_{f_2}} \right\|_{f_1} = L^{d\gamma/q_1} (LN)^{d/p_1} = L^{d(\frac{\gamma}{q_1} + \frac{1}{p_1})} N^{d/p_1}. \quad (3.45)$$

Escójase $N, N' \in \mathbb{N}$, $L = 2^l, L' = 2^{l'}$ tal que $L^\alpha N = (L')^\alpha N'$; de esta forma (3.44) implica $\nu_\alpha(\Gamma_{N,L}) = \nu_\alpha(\Gamma_{N',L'})$. Por (3.42) y (3.45) se deduce

$$L^{d(\frac{\gamma}{q_1} + \frac{1}{p_1})} N^{d/p_1} \approx (L')^{d(\frac{\gamma}{q_1} + \frac{1}{p_1})} (N')^{d/p_1} \Leftrightarrow \left(\frac{L}{L'}\right)^{d(\frac{\gamma}{q_1} + \frac{1-p_1}{p_1})} \approx 1.$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Esto fuerza que $\frac{\gamma}{q_1} = \frac{\alpha-1}{p_1}$, o equivalentemente $\frac{s_2-s_1}{d} - \frac{1}{p_2} = \frac{\alpha-1}{p_1}$ como se deseaba probar.

Para $\alpha = 1$ todavía debemos probar que $p_1 = q_1$. Sea $N \in \mathbb{N}$ y $\Gamma_N = \{Q \subset [0, 1]^d : 2^{-Nd} < |Q| \leq 1\}$. Se tiene $\nu_1(\Gamma_N) = N$ y

$$\left\| \sum_{Q \in \Gamma_N} \frac{\mathbf{e}_Q}{\|\mathbf{e}_Q\|_{f_2}} \right\|_{f_1} = \left(\int_{\mathbb{R}^d} \left(\sum_{\Gamma_N} \chi_Q(x) \right)^{p_1/q_1} dx \right)^{1/p_1} = N^{1/q_1}. \quad (3.46)$$

Para la misma $N \in \mathbb{N}$, tómesese $\tilde{\Gamma}_N = \{[0, 1]^d + j\mathbf{e}_1^d : 0 \leq j < N\}$ para que $\nu(\tilde{\Gamma}_N) = N$ y

$$\left\| \sum_{Q \in \tilde{\Gamma}_N} \frac{\mathbf{e}_Q}{\|\mathbf{e}_Q\|_{f_2}} \right\|_{f_1} = \left(\int_{\mathbb{R}^d} \chi_{[0, N] \times [0, 1]^{d-1}}(x) dx \right)^{1/p_1} = N^{1/p_1}. \quad (3.47)$$

Por (3.42) aplicado a Γ_N y $\tilde{\Gamma}_N$ junto con (3.46) and (3.47) se obtiene $N^{1/q_1} \approx N^{1/p_1}$. Esto fuerza $p_1 = q_1$, como se deseaba probar. ■

Los Teoremas 3.4.17 y 3.4.21 con $\eta(t) = t^{1/p_1}$ y $u_Q = \|\mathbf{e}_Q\|_{f_{p_2, q_2}^{s_2}}$ junto con el Teorema 3.4.26 muestran que la aproximación no lineal con el error medido en $f_{p_1, q_1}^{s_1}$ cuando la base está normalizada en $f_{p_2, q_2}^{s_2}$ está relacionado al uso de la medida $\nu_\alpha(Q) = |Q|^\alpha$, $Q \in \mathcal{D}$ ($\alpha = p_1(\frac{s_2-s_1}{d} - \frac{1}{p_2}) + 1$, cuando $\alpha \neq 1$; y para $\alpha = 1$ se tienen $\frac{s_2-s_1}{d} = \frac{1}{p_2}$ y $p_1 = q_1$) para controlar el número de términos en la aproximación. Nótese que los segundos parámetros de suavidad q_1 y q_2 no juegan ningún papel en α .

Los Teoremas 3.4.17 y 3.4.21 junto con el Teorema 3.4.26 también nos permiten identificar los espacios de aproximación en el escenario de los espacios de Triebel-Lizorkin como espacios de Lorentz discretos. Más específicamente se tiene:

Corolario 3.4.27 Sean $s_1, s_2 \in \mathbb{R}$, $0 < p_1, p_2 < \infty$, $0 < q_1, q_2 \leq \infty$ y defínase $\alpha = p_1(\frac{s_2-s_1}{d} - \frac{1}{p_2}) + 1$. Para $\Gamma \subset \mathcal{D}$ sea $\nu_\alpha(\Gamma) = \sum_{Q \in \Gamma} |Q|^\alpha$. Sean $\xi > 0$ y $\mu \in (0, \infty]$. Si $\alpha \neq 1$,

$$\mathcal{A}_\mu^\xi(f_{p_1, q_1}^{s_1}, \nu_\alpha) = \ell^{\tau, \mu}(\mathbf{u}, \nu_\alpha),$$

donde $\frac{1}{\tau} = \xi + \frac{1}{p_1}$ y $\mathbf{u} = \{\|\mathbf{e}_Q\|_{f_{p_2, q_2}^{s_2}}\}_{Q \in \mathcal{D}}$. Si $\alpha = 1$ el resultado se verifica con $p_1 = q_1$.

Demostración. Es consecuencia directa al aplicar los Teoremas 3.4.17 y 3.4.21 a $f = f_{p_1, q_1}^{s_1}$, \mathbf{u} y ν_α , como en el enunciado del corolario, y con $\eta(t) = t^{1/p_1}$.

■

Para valores particulares de los parámetros, los espacios de Lorentz discretos que aparecen en el Corolario 3.4.27 pueden identificarse con ciertos espacios de Besov.

Lema 3.4.28 Sean $s_1, s_2 \in \mathbb{R}$, $0 < p_1, p_2 < \infty$, $0 < q_2 \leq \infty$ y defínase $\alpha = p_1(\frac{s_2 - s_1}{d} - \frac{1}{p_2}) + 1$. Para $\Gamma \subset \mathcal{D}$ sea $\nu_\alpha(\Gamma) = \sum_{Q \in \Gamma} |Q|^\alpha$. Dado $\tau \in (0, \infty)$ se tiene

$$\ell^{\tau, \tau}(\mathbf{u}, \nu_\alpha) = b_{\tau, \tau}^\gamma,$$

(con cuasi normas iguales) donde $\mathbf{u} = \{\|\mathbf{e}_Q\|_{f_{p_2, q_2}^{s_2}}\}_{Q \in \mathcal{D}}$ y $\gamma = s_1 + d(\frac{1}{\tau} - \frac{1}{p_1})(1 - \alpha)$.

Demostración. Sea $f_2 := f_{p_2, q_2}^{s_2}$ para simplificar notación. Ya que

$$\|\mathbf{e}_Q\|_{f_2} = |Q|^{-s_2/d + 1/p_2 - 1/2} = |Q|^{-\gamma/d + (1-\alpha)/\tau - 1/2},$$

para $\mathbf{s} = \sum_{Q \in \mathcal{D}} s_Q \mathbf{e}_Q$ se tiene

$$\begin{aligned} \|\mathbf{s}\|_{\ell^{\tau, \tau}(\mathbf{u}, \nu_\alpha)}^\tau &= \left\| \{ \|s_Q \mathbf{e}_Q\|_{f_2} \}_{Q \in \mathcal{D}} \right\|_{\ell^{\tau, \tau}(\nu_\alpha)}^\tau = \sum_{Q \in \mathcal{D}} \|s_Q \mathbf{e}_Q\|_{f_2}^\tau |Q|^\alpha \\ &= \sum_{Q \in \mathcal{D}} (|s_Q| |Q|^{-\gamma/d + (1-\alpha)/\tau - 1/2} |Q|^{\alpha/\tau})^\tau \\ &= \sum_{Q \in \mathcal{D}} (|s_Q| |Q|^{-\gamma/d + 1/\tau - 1/2})^\tau = \|\mathbf{s}\|_{b_{\tau, \tau}^\gamma}^\tau. \end{aligned}$$

■

Nota 3.4.29 Si consideramos el punto de vista de [67], entonces solo podemos probar $\ell^{\tau, \tau}(\mathbf{u}, \nu_\alpha) = b_{\tau, \tau}^\gamma \cap f_{p_1, q_1}^{s_1}$ y la equivalencia de cuasi normas se verifica si tomamos $\|\cdot\|_{b_{\tau, \tau}^\gamma}$ en el lado derecho

Cuando $\tau < p_1$ y $\frac{\gamma}{d} - \frac{1}{\tau} = \frac{s_1}{d} - \frac{1}{p_1}$ se sabe que $b_{\tau, \tau}^\gamma \hookrightarrow f_{p_1, q_1}^{s_1}$ (ver [40] o [14]). Esta situación ocurre cuando $\alpha = 0$ (la medida de contar) pero no es verdad en nuestra situación más general.

El siguiente resultado identifica ciertos espacios de aproximación no lineal restringidos, cuando el error es medido en los espacios de Triebel-Lizorkin, como espacios de Besov. Es un corolario directo del Lema 3.4.28 y el Corolario 3.4.27.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Corolario 3.4.30 Sean $s_1, s_2 \in \mathbb{R}, 0 < p_1, p_2 < \infty, 0 < q_1 \leq \infty$ y defínase $\alpha = p_1(\frac{s_2 - s_1}{d} - \frac{1}{p_2}) + 1$. Para $\Gamma \subset \mathcal{D}$ sea $\nu_\alpha(\Gamma) = \sum_{Q \in \Gamma} |Q|^\alpha$. Dado $\xi > 0$ defínase τ por $\frac{1}{\tau} = \xi + \frac{1}{p_1}$. Si $\alpha \neq 1$,

$$\mathcal{A}_\tau^\xi(f_{p_1, q_1}^{s_1}, \nu_\alpha) = b_{\tau, \tau}^\gamma \quad (\text{cuasi normas equivalentes}),$$

donde $\gamma = s_1 + d\xi(1 - \alpha)$. Si $\alpha = 1$ el resultado se verifica con $\gamma = s_1$ y $p_1 = q_1$.

Nota 3.4.31 Si aplicáramos el Teorema 1 en [67] obtendríamos $\mathcal{A}_\tau^\xi(f_{p_1, q_1}^{s_1}, \mathbf{u}, \nu_\alpha) = b_{\tau, \tau}^\gamma \cap f_{p_1, q_1}^{s_1}$ con equivalencia de cuasi normas, como en la Nota 3.4.29

Los resultados obtenidos en [24] para aproximación no lineal restringida con ondículas en el espacio de Hardy $H^p, 0 < p < \infty$, cuando los coeficientes de ondículas están restringidos a $H^r, 0 < r < \infty$, son simples consecuencias de los resultados anteriores y del marco de transferencia abstracta desarrollado en [46]. Para ver esto, note que los espacios de sucesiones asociados a H^p y H^r (r, p como antes) son $f_{p, 2}^0$ y $f_{r, 2}^0$, respectivamente.

Así, para una base de ondículas $\mathcal{W} = \{\psi_Q^{(\ell)} : Q \in \mathcal{D}, \ell = 1, \dots, L\}, (L = 2^d - 1)$, admisible para H^p y $B_{\tau, \tau}^\gamma$,

$$\mathcal{A}_\tau^\xi(H^p, \mathcal{W}, \nu_\alpha) = B_{\tau, \tau}^\gamma, \quad (3.48)$$

donde $\gamma = \frac{dp}{r}\xi$, τ definido por $\frac{1}{\tau} = \xi + \frac{1}{p}$, y $\alpha = 1 - p/r$ ($\neq 1$). Este es el Corolario 6.3 en [24]. Note que $\mathcal{A}_\tau^\xi(H^p, \mathcal{W}, \nu_\alpha)$ corresponde al espacio de aproximación donde los coeficientes de ondículas están normalizados en H^r . En la notación de arriba enfatizamos que los espacios de aproximación están definidos usando una base de ondículas.

En esta situación, las desigualdades de Jackson y Bernstein (Teoremas 5.1 y 5.2 en [24]) se siguen de (3.48) y el hecho de que los espacios de aproximación siempre satisfacen las desigualdades de Jackson y Bernstein.

La otra situación considerada en [24] es $B_p := B_{p, p}^0, 0 < p < \infty$, cuando los coeficientes de ondículas están normalizados en $H_r, 0 < r < \infty$. En este caso, los espacios de sucesiones asociados a $B_{p, p}^0 = F_{p, p}^0$ y H^r son $f_{p, p}^0$ y $f_{r, 2}^0$, respectivamente. El Corolario 3.4.30 produce entonces

$$\mathcal{A}_\tau^\xi(B_p, \mathcal{W}, \nu_\alpha) = B_{\tau, \tau}^\gamma, \quad (3.49)$$

donde $\gamma = \frac{dp}{r}\xi$, τ definido por $\frac{1}{\tau} = \xi + \frac{1}{p}$, con ν_α y \mathcal{W} como antes. Esto es más general que el Corolario 6.1 en [24] y una comparación con (3.48) prueba inmediatamente una

versión más general del Teorema 6.3 en [24]. Por supuesto, las desigualdades de Jackson y Bernstein de los Teoremas 5.4 y 5.5 en [24] también se siguen de nuestros resultados.

Para mostrar un ejemplo no tratado en [24] considere la base ortonormal de ondículas $\mathcal{W} = \{\psi_Q^{(\ell)} : Q \in \mathcal{D}, \ell = 1, \dots, L\}$ ($L = 2^d - 1$), admisible para el espacio de Sobolev $W_2^s, s > 0$. Queremos medir el error en W_2^s pero restringimos los coeficientes de ondículas a L^2 . ya que los espacios de sucesiones asociados a W_2^s y L^2 son $f_{2,2}^s$ y $f_{2,2}^0$, respectivamente, el Corolario 3.4.30 y el marco de transferencia abstracto de [46] da

$$\mathcal{A}_\tau^\xi(W_2^s, \mathcal{W}, \nu_\alpha) = b_{\tau,\tau}^\gamma$$

donde $\gamma = s + d\xi(1 - \alpha)$, $\alpha = -2s/d$ and τ definido por $\frac{1}{\tau} = \xi + \frac{1}{2}$.

Hacemos notar que, definiendo espacios de sucesiones apropiadamente, un poco de más trabajo muestra los resultado probados en [61], el caso anisotrópico.

3.4.11 Interpolación real para espacios de Triebel-Lizorkin

Una vez que los espacios de aproximación restringida para espacios de sucesiones de Triebel-Lizorkin han sido identificados (ver Corolario 3.4.30) usamos el Teorema 3.4.18 para obtener resultados acerca de interpolación real. Este método ya ha sido usado anteriormente (ver [33] o [46]). Pero en el caso clásico, los parámetros de los espacios interpolados están limitados. con la teoría de la aproximación restringida probaremos resultados de interpolación para un conjunto mucho más grande de parámetros.

Teorema 3.4.32 Sean $s \in \mathbb{R}$, $0 < p < \infty$, $0 < q \leq \infty$. Para $0 < \tau < p$, $0 < \theta < 1$ y $\gamma \neq s$ ($\gamma \in \mathbb{R}$) se tiene

$$(f_{p,q}^s, b_{\tau,\tau}^\gamma)_{\theta,\tau_\theta} = b_{\tau_\theta,\tau_\theta}^{(1-\theta)s+\theta\gamma} \quad \text{con} \quad \frac{1}{\tau_\theta} = \frac{1-\theta}{p} + \frac{\theta}{\tau}.$$

Demostración. Escribase $\xi = 1/\tau - 1/p > 0$ y escójase $\alpha \neq 1$ tal que $\gamma = s + (1 - \alpha)\xi d$ (i.e. $\alpha = 1 - \frac{\gamma-s}{\xi d}$), lo cual es posible ya que $\gamma \neq s$. Un vez que se escoge α , tómesese $s_2 \in \mathbb{R}$ de tal forma que $\alpha = p(\frac{s_2-s}{d})$ (i.e. $s_2 = s + \frac{\alpha d}{p}$). El Teorema 3.4.26 muestra que $f_{p,q}^s$ satisface 1) de los Teoremas 3.4.17 y 3.4.21 con $\eta(t) = t^{1/p}$, para el espacio de “normalización” $f_2 := f_{p,p}^{s_2}$ y ν_α (nótese que $\alpha = p(\frac{s_2-s}{d})$ es la condición requerida en el Lema 3.4.28). Así, el espacio $\ell^{\tau,\mu}(\mathbf{u}, \nu_\alpha)$ satisface las desigualdades de Jackson y Bernstein de orden $\xi = 1/\tau - 1/p > 0$, donde $\mathbf{u} = \{\|e_Q\|_{f_2}\}_{Q \in \mathcal{D}}$. Tomando $\mu = \tau$,

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

el Lema 3.4.28 muestra que $b_{\tau,\tau}^\gamma$ satisface las desigualdades de Jackson y Bernstein de orden $\xi = 1/\tau - 1/p > 0$, ya que $\gamma = s + (1 - \alpha)\xi d$ (que es la condición requerida).

Por el Teorema 3.4.18, para $0 < \theta < 1$,

$$(f_{p,q}^s, b_{\tau,\tau}^\gamma)_{\theta,\tau_\theta} = \mathcal{A}_{\tau_\theta}^{\theta\xi}(f_{p,q}^s, \nu_\alpha).$$

Ya que $\frac{1}{\tau_\theta} = \frac{(1-\theta)}{p} + \frac{\theta}{\tau} = \theta(\frac{1}{\tau} - \frac{1}{p}) + \frac{1}{p} = \theta\xi + \frac{1}{p}$, el Corolario 3.4.30 da

$$\mathcal{A}_{\tau_\theta}^{\theta\xi}(f_{p,q}^s, \nu_\alpha) = b_{\tau_\theta,\tau_\theta}^{\tilde{\gamma}}$$

con $\tilde{\gamma} = s + d\theta\xi(1 - \alpha) = s + d\theta\xi\frac{(\gamma-s)}{\xi d} = s + \theta(\gamma - s) = (1 - \theta)s + \theta\gamma$, lo que prueba el resultado. ■

Nota 3.4.33 Aunque el teorema es presentado como un resultado de interpolación entre espacios de sucesiones de Triebel-Lizorkin y de Besov, es en realidad un resultado de interpolación entre los espacios de sucesiones de Triebel-Lizorkin, ya que $b_{\tau,\tau}^\gamma = f_{\tau,\tau}^\gamma$. Así, el resultado puede establecerse como

$$(f_{p,q}^s, f_{\tau,\tau}^\gamma)_{\theta,\tau_\theta} = f_{\tau_\theta,\tau_\theta}^{(1-\theta)s+\theta\gamma} \quad \text{con} \quad \frac{1}{\tau_\theta} = \frac{1-\theta}{p} + \frac{\theta}{\tau}. \quad (3.50)$$

Nota 3.4.34 Nótese que no necesitamos la restricción $\frac{\gamma}{d} - \frac{1}{\tau} = \frac{s}{d} - \frac{1}{p}$ característica de este tipo de resultados cuando se usa la aproximación no lineal clásica (ver [33] o [46]).

Nota 3.4.35 Por el marco de transferencia abstracto diseñado en [46], el resultado del Teorema 3.4.32 puede ser traducido a un resultado para espacios (homogeneos) de Triebel-Lizorkin. Los casos no homogéneo y el de dominios acotados pueden obtenerse con modificaciones menores en la prueba.

Nota 3.4.36 El mérito del Teorema 3.4.32 es que prueba resultados para un gran conjunto de parámetros usando la teoría de aproximación no lineal restringida. No obstante, muchos (pero no todos, hasta donde sabemos) han sido probados. Uno puede leer del Teorema 3.5 en [14] el siguiente resultado:

$$(F_{p_0,q_0}^{s_0}, F_{p_1,q_1}^{s_1})_{\theta,p} = F_{p,p}^s = B_{p,p}^s \quad (3.51)$$

cuando $p_i < q_i$, $i = 0, 1$, $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, $s = (1 - \theta)s_0 + \theta s_1$ y $s_0 \neq s_1$. Comparándolo con (3.50) vemos que (3.51) tiene un conjunto más grande de parámetros en el segundo espacio, mientras que (3.50) no tiene la restricción $p < q$ que se requiere en (3.51). Ambas deficiencias se deben a los métodos de las pruebas. Por otra parte, el Teorema 2.42/1 (página 184) de [99] muestra (3.51) sin $p_i < q_i$ pero suponiendo $1 < p_i, q_i < \infty$.

3.5 Espacios de Orlicz

En los ejemplos con ondículas de la aproximación no lineal restringida hemos encontrado espacios que cumplen la propiedad de Temlyakov con función $\eta(t) = t^{1/p}$. Como veremos en esta sección, los espacios de Orlicz son ejemplos en los que la función $\eta(t)$ es diferente a $t^{1/p}$. Aquí, generalizamos la propiedad de Temlyakov (3.20), de L^p con $1 < p < \infty$, y de H^p con $0 < p \leq 1$, de [24] para un espacio de Orlicz $L^\Phi(\mathbb{R}^d)$. La dificultad radica en encontrar las expresiones para los lados izquierdo y derecho de (3.20). Esto requiere el uso de las funciones de dilatación de la función fundamental de un espacio de Orlicz. A continuación damos un resumen de los espacios de Orlicz y, al mismo tiempo, establecemos notación.

Una función de Young es una función convexa no-decreciente $\Phi : [0, \infty) \rightarrow [0, \infty)$ tal que $\lim_{t \rightarrow 0^+} \Phi(t) = 0$ y $\lim_{t \rightarrow \infty} \Phi(t) = \infty$. Asumimos que Φ es estrictamente creciente y finita, de esta forma es una biyección continua de $[0, \infty)$. Dada una función de Young Φ , el espacio de Orlicz $L^\Phi(\mathbb{R}^d)$ es el conjunto de todas las funciones medibles $f : \mathbb{R}^d \rightarrow \mathbb{C}$ tal que $\Phi(|f(x)|/\lambda) \in L^1(\mathbb{R}^d)$ para algún $\lambda > 0$. Remitimos al lector a [8] y a [91] para un tratamiento completo de estos espacios. Se sabe que L^Φ se convierte en un espacio de funciones de Banach invariante al reordenamiento con la norma de Luxemburg

$$\|f\|_{L^\Phi(\mathbb{R}^d)} := \inf\{\lambda > 0 : \int_{\mathbb{R}^d} \Phi\left(\frac{|f(x)|}{\lambda}\right) dx \leq 1\}, \quad (3.52)$$

(ver p.e. [8], p. 269). La función fundamental del espacio de Orlicz $L^\Phi(\mathbb{R}^d)$ puede computarse explícitamente por medio de

$$\varphi(t) = \frac{1}{\Phi^{-1}(1/t)}, \quad t > 0, \quad (3.53)$$

(ver [8], p. 276). Así, φ es una biyección continua estrictamente creciente de $[0, \infty)$. Definiendo

$$h_\varphi^+(t) = \sup_{s>0} \frac{\varphi(st)}{\varphi(s)}, \quad t > 0,$$

los índices de Boyd inferior y superior de $L^\Phi(\mathbb{R}^d)$ están dados por

$$\begin{aligned} i_\varphi &:= \lim_{t \rightarrow 0^+} \frac{\log h_\varphi^+(t)}{\log t} = \sup_{0 < t < 1} \frac{\log h_\varphi^+(t)}{\log t} \\ I_\varphi &:= \lim_{t \rightarrow \infty} \frac{\log h_\varphi^+(t)}{\log t} = \inf_{1 < t < \infty} \frac{\log h_\varphi^+(t)}{\log t}, \end{aligned} \quad (3.54)$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

(ver [8], p. 277). Se puede probar que $0 \leq i_\varphi \leq I_\varphi \leq 1$.

Además de $\Phi(t) = t^{1/p}$, $1 \leq p < \infty$, otro ejemplo típico de una función de Young es $\Phi(t) = t^p[\log(e+t)]^\alpha$, $1 \leq p < \infty$, $\alpha \in \mathbb{R}$. Cuando $\alpha = 0$, $L^\Phi(\mathbb{R}^d)$ coincide con los espacios de Lebesgue, y si $\alpha > 0$, $L^\Phi(\mathbb{R}^d)$ es el espacio de Zygmund $L^p(\log L)^\alpha$. Véase [47] para más ejemplos.

La función $\Phi : [0, \infty) \rightarrow [0, \infty)$ es **doblante** si existe $D > 0$ tal que $\Phi(2t) \leq D\Phi(t)$ para todo $t > 0$. No es difícil ver que para una función de Young Φ , $i_\varphi > 0$ es equivalente a Φ doblante. Aún más, si $\tilde{\Phi}$ es la **función complementaria de Young** a Φ (ver p.e. [91], p. 6, para la definición precisa) entonces Φ y $\tilde{\Phi}$ doblante equivale a decir que $L^\Phi(\mathbb{R}^d)$ y $L^{\tilde{\Phi}}(\mathbb{R}^d)$ es un par de espacios de Orlicz reflexivos con $0 < i_\varphi \leq I_\varphi < 1$. La dualidad está dada por

$$\|f\|_{L^\Phi(\mathbb{R}^d)} \approx \sup \left\{ \left| \int_{\mathbb{R}^d} f(x)g(x)dx \right| : \|g\|_{L^{\tilde{\Phi}}(\mathbb{R}^d)} \leq 1 \right\}. \quad (3.55)$$

Cuando $L^\Phi(\mathbb{R}^d)$ y $L^{\tilde{\Phi}}(\mathbb{R}^d)$ es un par reflexivo de espacios de Orlicz podemos considerar las funciones fundamentales $\varphi(t)$ y $\tilde{\varphi}(t)$ de cada uno de ellos. En [8](p. 276) se prueba que $t \leq \Phi^{-1}(t)\tilde{\Phi}^{-1}(t) \leq 2t$, y así

$$\varphi(t)\tilde{\varphi}(t) \approx t, \quad t > 0. \quad (3.56)$$

Aún más, se sigue de (3.56) que

$$i_{\tilde{\varphi}} = 1 - I_\varphi \quad \text{y} \quad I_{\tilde{\varphi}} = 1 - i_\varphi \quad (3.57)$$

(ver [8], p. 280).

Hacemos ahora un breve repaso de las caracterizaciones de algunos espacios funcionales usando ondículas, que serán necesarias en los resultados que siguen.

Definición 3.5.1 Para $0 < p < \infty$ diremos que la base de ondículas \mathcal{W} como en (3.40) es p -admisibles si \mathcal{W} es una base incondicional de $H^p(\mathbb{R}^d)$ y si $f = \sum_{\ell=1}^L \sum_{Q \in \mathcal{D}} \langle f, \psi_Q^\ell \rangle \psi_Q^\ell$, entonces

$$\|f\|_{H^p(\mathbb{R}^d)} \approx \sum_{\ell=1}^L \left\| \left[\sum_{Q \in \mathcal{D}} (|Q|^{-1/2} |\langle f, \psi_Q^\ell \rangle| \chi_Q(\cdot))^2 \right]^{1/2} \right\|_{L^p(\mathbb{R}^d)} \quad (3.58)$$

donde χ_Q denota la función característica del cubo diádico Q . Recordamos que $H^p(\mathbb{R}^d)$ coincide con el espacio de Lebesgue $L^p(\mathbb{R}^d)$ si $1 < p < \infty$.

Para $1 < p < \infty$, el sistema d -dimensional de Haar es p -admisibles. También son admisibles las ondículas que surgen de un análisis multirresolución r regular ([81], p. 22), y las ondículas que pertenecen a la clase de regularidad $\mathcal{R}^{0,M}$ ([59], p. 264 para $d = 1$ o [70]). Para $0 < p \leq 1$ se necesita un poco más de regularidad (ver [43] o [70] para los detalles).

De igual forma que para $H^p(\mathbb{R}^d)$, tenemos la siguiente definición para los espacios de Orlicz.

Definición 3.5.2 Para una función de Young Φ diremos que una base de ondículas \mathcal{W} como en (3.40) es Φ -admisibles si para todo $f \in L^\Phi(\mathbb{R}^d)$ se tiene $f = \sum_{\ell=1}^L \sum_{Q \in \mathcal{D}} \langle f, \psi_Q^\ell \rangle \psi_Q^\ell$ con convergencia incondicional en $L^\Phi(\mathbb{R}^d)$ y

$$\|f\|_{L^\Phi(\mathbb{R}^d)} \approx \sum_{\ell=1}^L \left\| \left[\sum_{Q \in \mathcal{D}} (|Q|^{-1/2} |\langle f, \psi_Q^\ell \rangle| \chi_Q(\cdot))^2 \right]^{1/2} \right\|_{L^\Phi(\mathbb{R}^d)}, \quad (3.59)$$

donde χ_Q denota la función característica del cubo Q .

En [93] se prueba que si \mathcal{W} es p -admisibles para todo $1 < p < \infty$, entonces \mathcal{W} es Φ -admisibles siempre y cuando los índices de Boyd (3.54) de $L^\Phi(\mathbb{R}^d)$ estén en $(0, 1)$.

Empezamos probando el lado derecho de la propiedad de Temlyakov.

Proposición 3.5.3 Sea $L^\Phi(\mathbb{R}^d)$ un espacio de Orlicz con índices de Boyd en $(0, 1)$ y sea \mathcal{W} una base de ondículas r -admisibles y Φ -admisibles. Sea $\Gamma = \cup_{\ell=1}^L \Gamma_\ell$ con $\Gamma_\ell \subset \mathcal{D}$ y tal que $\nu_\Phi(\Gamma) := \sum_{\ell=1}^L \sum_{Q \in \Gamma_\ell} |Q| \Phi(|Q|^{-1/r}) < \infty$. Existe $C > 0$ tal que

$$\sum_{\ell=1}^L \left\| \sum_{Q \in \Gamma_\ell} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r(\mathbb{R}^d)}} \right\|_{L^\Phi(\mathbb{R}^d)} \leq CLh_\Phi^+(\nu_\Phi(\Gamma)). \quad (3.60)$$

Demostración. Comenzamos con un solo ℓ fijo y escribimos $\Gamma_\ell = \Gamma$ para simplificar la notación. Ya que \mathcal{W} es r -admisibles, entonces $\|\psi_Q^\ell\|_{H^r} \approx |Q|^{\frac{1}{r}-\frac{1}{2}}$ por (3.58). Entonces,

$$\begin{aligned} \left\| \sum_{Q \in \Gamma} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r(\mathbb{R}^d)}} \right\|_{L^\Phi(\mathbb{R}^d)} &\approx \left\| \sum_{Q \in \Gamma} |Q|^{\frac{1}{2}-\frac{1}{r}} \psi_Q^\ell \right\|_{L^\Phi(\mathbb{R}^d)} \\ &\approx \left\| \left[\sum_{Q \in \Gamma} |Q|^{-\frac{2}{r}} \chi_Q(\cdot) \right]^{1/2} \right\|_{L^\Phi(\mathbb{R}^d)}, \end{aligned} \quad (3.61)$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

por (3.59). Sea $S_{\Gamma,r}(x) = \sum_{Q \in \Gamma} |Q|^{-2/r} \chi_Q(x)$, $x \in \cup_{Q \in \Gamma} Q$. Para un $x \in \cup_{Q \in \Gamma} Q$, sea Q_x el cubo más pequeño de Γ que contenga a x , si existe. Mostraremos que $\nu_\Phi(\Gamma) < \infty$ implica que Q_x existe (y es único) para c.t. $x \in \cup_{Q \in \Gamma} Q$. De hecho, si $E_\Gamma = \{x \in \cup_{Q \in \Gamma} Q : Q_x \text{ no existe}\}$ y escribimos $\mathcal{D}_k = \{Q \in \mathcal{D} : |Q| = 2^{-kd}\}$, $k \in \mathbb{Z}$, tenemos que para todo $m \geq 0$, $E_\Gamma \subset \cup_{k \geq m} \cup_{Q \in \Gamma \cap \mathcal{D}_k} Q$. Entonces,

$$\begin{aligned} |E_\Gamma| &\leq \sum_{k \geq m} \sum_{Q \in \Gamma \cap \mathcal{D}_k} |Q| \frac{\Phi(|Q|^{-1/r})}{\Phi(|Q|^{-1/r})} \leq \sum_{k \geq m} \frac{1}{\Phi(2^{kd/r})} \sum_{Q \in \Gamma \cap \mathcal{D}_k} |Q| \Phi(|Q|^{-1/r}) \\ &\leq \nu_\Phi(\Gamma) \sum_{k \geq m} \frac{1}{\Phi(2^{kd/r})}. \end{aligned}$$

Como Φ es convexa y $\Phi(0) = 0$, $\Phi(1) = \Phi(2^{-kd/r} 2^{kd/r}) \leq 2^{-kd/r} \Phi(2^{kd/r})$. Por lo tanto,

$$|E_\Gamma| \leq \nu_\Phi(\Gamma) \sum_{k \geq m} \frac{1}{\Phi(1)} 2^{-kd/r} \lesssim \nu_\Phi(\Gamma) 2^{-md/r}.$$

Dejando que $m \rightarrow \infty$ se deduce $|E_\Gamma| = 0$.

Para todo $x \in \cup_{Q \in \Gamma} Q \cap E_\Gamma^c$ claramente tenemos $S_\Gamma(x) \geq |Q_x|^{-2/r} \chi_{Q_x}(x)$. Probamos ahora la desigualdad inversa (excepto por una constante). Extendiendo la suma para incluir todos los cubos diádicos que contengan a Q_x , para $x \in E_\Gamma$, se obtiene

$$\begin{aligned} S_\Gamma(x) &\leq \sum_{Q \supset Q_x : Q \in \Gamma} |Q|^{-2/r} \chi_{Q_x}(x) = \sum_{j=0}^{\infty} 2^{-2jd/r} |Q_x|^{-2/r} \\ &= C |Q_x|^{-2/r} \chi_{Q_x}(x). \end{aligned}$$

Por lo tanto,

$$S_\Gamma(x) \approx |Q_x|^{-2/r} \chi_{Q_x}(x) \quad \text{para todo } x \in \cup_{Q \in \Gamma} Q \setminus E_\Gamma. \quad (3.62)$$

Si $\Gamma_{\min} := \{Q_x : x \in \cup_{Q \in \Gamma} Q \setminus E_\Gamma\}$ es la familia de cubos **mínimos** de Γ , (3.62) muestra que $S_\Gamma(x) \approx S_{\Gamma_{\min}}(x)$. Para cualquier $Q \in \Gamma$, defínase la **sombra** de Q como la unión de todos los cubos de Γ estrictamente contenidos en Q , esto es, $\text{sombra}(Q) = \cup\{R : R \in \Gamma, R \subsetneq Q\}$. Defínase la **luz** de Q como $\text{luz}(Q) = Q - \text{sombra}(Q)$ (estos objetos han sido usados en [47]). Se sigue que $Q \in \Gamma_{\min}$ si y sólo si $\text{luz}(Q) \neq \emptyset$ y

$$\cup_{Q \in \Gamma} Q = \cup_{Q \in \Gamma_{\min}} \text{luz}(Q),$$

donde los conjuntos en la última unión son disjuntos a pares. Por lo tanto, por (3.62) podemos escribir, para $x \in \cup_{Q \in \Gamma} Q \setminus E_\Gamma$,

$$S_\Gamma(x) \approx \sum_{Q \in \Gamma_{\min}} |Q|^{-2/r} \chi_{\text{luz}(Q)}(x), \quad (3.63)$$

donde en la última suma hay a lo sumo un término diferente de cero. para cada x . De (3.61) podemos escribir

$$\left\| \sum_{Q \in \Gamma} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r(\mathbb{R}^d)}} \right\|_{L^\Phi(\mathbb{R}^d)} \approx \left\| \sum_{Q \in \Gamma} |Q|^{-1/r} \chi_{\text{luz}(Q)}(\cdot) \right\|_{L^\Phi(\mathbb{R}^d)}. \quad (3.64)$$

Para $Q \in \Gamma_{\min}$, escójase $u_Q \in \mathbb{R}^+$ tal que $\varphi(u_Q) = |Q|^{1/r}$ donde φ es la función fundamental de $L^\Phi(\mathbb{R}^d)$ (ver (3.53) para la definición). Entonces, para todo $t \geq 0$

$$h_\varphi^+(t) = \sup_{s>0} \frac{\varphi(st)}{\varphi(s)} \geq \frac{\varphi(u_Q t)}{\varphi(u_Q)} = \frac{\varphi(u_Q t)}{|Q|^{1/r}}.$$

Sea $\lambda := h_\varphi^+(\nu_\Phi(\Gamma))$, de manera que $\lambda \geq \varphi(u_Q \nu_\Phi(\Gamma)) / |Q|^{1/r}$. Así,

$$\frac{1}{\lambda |Q|^{1/r}} \leq \frac{1}{\varphi(u_Q \nu_\Phi(\Gamma))}. \quad (3.65)$$

Se tiene

$$\begin{aligned} \int_{\mathbb{R}^d} \Phi \left(\frac{\sum_{Q \in \Gamma_{\min}} |Q|^{-1/r} \chi_{\text{luz}(Q)}(x)}{\lambda} \right) dx &= \sum_{Q \in \Gamma_{\min}} \Phi \left(\frac{1}{\lambda |Q|^{1/r}} \right) |\text{luz}(Q)| \\ &\leq \sum_{Q \in \Gamma_{\min}} \Phi \left(\frac{1}{\varphi(u_Q \nu_\Phi(\Gamma))} \right) |Q| \\ &= \sum_{Q \in \Gamma_{\min}} \Phi \left(\Phi^{-1} \left(\frac{1}{u_Q \nu_\Phi(\Gamma)} \right) \right) |Q| \\ &= \sum_{Q \in \Gamma_{\min}} \frac{1}{u_Q \nu_\Phi(\Gamma)} |Q|, \end{aligned} \quad (3.66)$$

donde hemos usado que las luces de los cubos son conjuntos disjuntos, (3.65), Φ es creciente y (3.53).

Ahora,

$$\begin{aligned} \sum_{Q \in \Gamma_{\min}} \frac{1}{u_Q \nu_\Phi(\Gamma)} |Q| &= \frac{1}{\nu_\Phi(\Gamma)} \sum_{Q \in \Gamma_{\min}} \frac{|Q|}{\varphi^{-1}(|Q|^{1/r})} \\ &= \frac{1}{\nu_\Phi(\Gamma)} \sum_{Q \in \Gamma_{\min}} |Q| \Phi(|Q|^{-1/r}) \leq 1. \end{aligned} \quad (3.67)$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Así, de la definición de la norma de Luxemburg (3.52), junto con (3.64), (3.66) y (3.67) se deduce el resultado para un ℓ fijo.

Recordando que

$$h_{\varphi}^{+}(t) = \sup_{s>0} \frac{\varphi(st)}{\varphi(s)}, \quad t > 0,$$

definimos

$$h_{\varphi}^{-}(t) = \inf_{s>0} \frac{\varphi(st)}{\varphi(s)}, \quad t > 0.$$

Para el caso en que se tenga diferentes ψ^{ℓ} se necesita el siguiente lema:

Lema 3.5.4 *Sea $0 < t_j < \infty$, $j = 1, 2, \dots, L$ and $T = \sum_{j=1}^L t_j$. Se tiene*

$$h_{\varphi}^{\pm}(T) \leq \sum_{j=1}^L h_{\varphi}^{\pm}(t_j) \leq Lh_{\varphi}^{\pm}(T).$$

Demostración. Sea $0 < a < b < \infty$; usando que $\varphi(t)/t$ es no creciente ([8], p. 67) se deduce $h_{\varphi}^{\pm}(a)/a \geq h_{\varphi}^{\pm}(b)/b$. De hecho, para h_{φ}^{+}

$$\begin{aligned} \frac{h_{\varphi}^{+}(a)}{a} &= \sup_{s>0} \frac{\varphi(sa)}{a\varphi(s)} = \sup_{s>0} \frac{\varphi(sa)/sa}{\varphi(s)/s} \\ &\geq \sup_{s>0} \frac{\varphi(sb)/sb}{\varphi(s)/s} = \frac{h_{\varphi}^{+}(b)}{b}. \end{aligned}$$

De manera similar para h^{-} . Así,

$$\sum_{j=1}^L h_{\varphi}^{\pm}(t_j) = \sum_{j=1}^L t_j \frac{h_{\varphi}^{\pm}(t_j)}{t_j} \geq \sum_{j=1}^L t_j \frac{h_{\varphi}^{\pm}(T)}{T} = h_{\varphi}^{\pm}(T).$$

Esto prueba el lado izquierdo de la desigualdad. Para probar el lado derecho de la desigualdad se usa que φ es creciente (ya que Φ es creciente) para obtener $h_{\varphi}^{\pm}(a) \leq h_{\varphi}^{\pm}(b)$ cuando $0 < a < b < \infty$. Así,

$$\sum_{j=1}^L h_{\varphi}^{\pm}(t_j) \leq h_{\varphi}^{\pm}(T) \sum_{j=1}^L 1 = Lh_{\varphi}^{\pm}(T).$$

■

Se puede concluir la demostración de la Proposición 3.5.3 usando el Lema 3.5.4:

$$\begin{aligned}
 \sum_{\ell=1}^L \left\| \sum_{Q \in \Gamma_\ell} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r}} \right\|_{L^\Phi(\mathbb{R}^d)} &\leq C \sum_{\ell=1}^L h_\varphi^+(\nu_\Phi(\Gamma_\ell)) \leq CLh_Q^+(\sum_{\ell=1}^L \nu_\Phi(\Gamma_\ell)) \\
 &= CLh_Q^+(\nu_\Phi(\Gamma)). \tag{3.68}
 \end{aligned}$$

■

Ahora queremos hallar una acotación por abajo de la expresión de la izquierda de (3.60). Para ello usaremos la dualidad de los espacios de Orlicz. Empezamos con una serie de resultados técnicos.

En el siguiente lema hacemos explícito un resultado que hemos probado en la prueba de la Proposición 3.5.3.

Lema 3.5.5 *Sea Φ una función de Young y $\Gamma \subset \mathcal{D}$ tal que, para $0 < r < \infty$, $\nu_\Phi(\Gamma) = \sum_{Q \in \Gamma} |Q| \Phi(|Q|^{-1/r}) < \infty$. Entonces, para c.t. $x \in \cup_{Q \in \Gamma} Q$ el cubo más pequeño de Γ que contiene a x existe.*

Lema 3.5.6 *Sea Φ una función de Young con función complementaria $\tilde{\Phi}$ y $0 < r < \infty$. Supóngase que $\varphi(t) := 1/\Phi^{-1}(1/t)$ satisface $i_\varphi > 0$. Si $\Gamma \subset \mathcal{D}$ es tal que $\nu_\Phi(\Gamma) := \sum_{Q \in \Gamma} |Q| \Phi(|Q|^{-1/r}) < \infty$, para c.t. $x \in \cup_{Q \in \Gamma} Q$ se tiene*

$$S_\Gamma(x) := \sum_{Q \in \Gamma} \frac{\chi_Q(x)}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2} \approx \frac{\chi_{Q_x}(x)}{[\varphi(1/\tilde{\Phi}(|Q_x|^{-1/r}))]^2}, \tag{3.69}$$

donde Q_x es el cubo más pequeño de Γ que contiene a x .

Demostración. Para c.t. $x \in \cup_{Q \in \Gamma} Q$, el cubo más pequeño existe, por el Lema 3.5.5. Para tales x el lado izquierdo de (3.69) es claramente mayor o igual que el lado derecho. Para probar la desigualdad inversa extendemos la suma en el lado izquierdo de (3.69) para obtener

$$\begin{aligned}
 S_\Gamma(x) &= \sum_{Q \in \Gamma} \frac{\chi_Q(x)}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2} \leq \sum_{Q \in \mathcal{D}: Q \supset Q_x} \frac{1}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2} \\
 &= \sum_{j=0}^{\infty} \frac{1}{[\varphi(1/\tilde{\Phi}(2^{-jd/r} |Q_x|^{-1/r}))]^2} \\
 &= \sum_{j=0}^{\infty} [\Phi^{-1}(\tilde{\Phi}(2^{-jd/r} |Q_x|^{-1/r}))]^2.
 \end{aligned}$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Como $\tilde{\Phi}$ es convexa y $\tilde{\Phi}(0) = 0$ se tiene $\tilde{\Phi}(2^{-jd/r} |Q_x|^{-1/r}) \leq 2^{-jd/r} \tilde{\Phi}(|Q_x|^{-1/r})$. Ya que Φ^{-1} es creciente se tiene

$$S_\Gamma(x) \leq \sum_{j=0}^{\infty} [\Phi^{-1}(2^{-jd/r} \tilde{\Phi}(|Q_x|^{-1/r}))]^2. \quad (3.70)$$

Se puede probar que $i_{\Phi^{-1}} = i_\varphi > 0$ y que $\Phi^{-1} \in \mathbb{W}_+$ y, por lo tanto, $[\Phi^{-1}]^2 \in \mathbb{W}_+$ (ver la Definición 3.4.10). Por el Lema 3.4.11 se deduce

$$S_\Gamma(x) \leq C[\Phi^{-1}(\tilde{\Phi}(|Q_x|^{-1/r}))]^2 = C \frac{1}{[\varphi(1/\tilde{\Phi}(|Q_x|^{-1/r}))]^2},$$

como se deseaba probar. ■

Nota 3.5.7 Aunque el Lema 3.4.11 da $\sum_{j=0}^{\infty} \eta(s_0^j t) \leq C\eta(t)$, $t > 0$, con s_0 como en la definición de $\eta \in \mathbb{W}_+$, no es difícil ver que para cualquier $s \in (0, 1)$,

$$\sum_{j=0}^{\infty} \eta(s^j t) \leq C_s \eta(t), \quad t > 0,$$

donde ahora la constante depende de s . Hemos usado este resultado en la prueba recién dada para $s = 2^{-d/r} < 1$ y $t = \tilde{\Phi}(|Q_x|^{-1/r})$.

Del Lema 3.4.11 se tiene

$$\sum_{j=0}^{\infty} \eta(s_0^j t) \leq C_{s_0} \eta(t), \quad t > 0.$$

Queremos probar que la desigualdad previa se verifica para todo $s \in (0, 1)$, i.e.

$$\sum_{j=0}^{\infty} \eta(s^j t) \leq C_s \eta(t), \quad t > 0.$$

Si $0 < s \leq s_0$, ya que η es creciente se tiene

$$\sum_{j=0}^{\infty} \eta(s^j t) \leq \sum_{j=0}^{\infty} \eta(s_0^j t) \leq C_{s_0} \eta(t), \quad t > 0.$$

Para $0 < s_0 < s < 1$ escójase $k_0 \in \mathbb{Z}_+$ tal que $s^{k_0} < s_0 \leq s^{k_0-1}$ (nótese que $s < 1$). Entonces,

$$\begin{aligned} \sum_{j=0}^{\infty} \eta(s^j t) &= \sum_{\ell=0}^{\infty} \sum_{m=0}^{k_0-1} \eta(s^{k_0\ell+m} t) = \sum_{\ell=0}^{\infty} \sum_{m=0}^{k_0-1} \eta(s^{k_0\ell} s^m t) \\ &\leq \sum_{\ell=0}^{\infty} \sum_{m=0}^{k_0-1} \eta(s^{k_0\ell} t) \leq \sum_{\ell=0}^{\infty} k_0 \eta(s_0^\ell t) \leq k_0 C_{s_0} \eta(t) \approx \frac{\log s}{\log s_0} C_{s_0} \eta(t). \end{aligned}$$

Lema 3.5.8 Sea $L^\Phi(\mathbb{R}^d)$ un espacio de Orlicz con índice de Boyd en $(0, 1)$, sea $\tilde{\Phi}$ la función complementaria de Φ y $0 < r < \infty$. Supóngase que la base de ondículas \mathcal{W} como en (3.40) es Φ -admisibile. Existe $C_1 > 0$ tal que para todo $\ell = 1, \dots, L$ y todos los conjuntos $\Gamma \subset \mathcal{D}$ con $\nu_{\tilde{\Phi}}(\Gamma) := \sum_{Q \in \Gamma} |Q| \tilde{\Phi}(|Q|^{-1/r}) < \infty$ se tiene

$$\left\| \sum_{Q \in \Gamma} \frac{|Q|^{1/2} \psi_Q^\ell(\cdot)}{\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))} \right\|_{L^\Phi(\mathbb{R}^d)} \leq C_1 h_\varphi^+(\nu_{\tilde{\Phi}}(\Gamma)), \quad (3.71)$$

donde φ es la función fundamental de $L^\Phi(\mathbb{R}^d)$.

Demostración. Como \mathcal{W} es Φ -admisibile por (3.59) el lado izquierdo de (3.71) es equivalente a

$$\left\| \left(\sum_{Q \in \Gamma} \frac{\chi_Q(x)}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2} \right)^{1/2} \right\|_{L^\Phi(\mathbb{R}^d)} = \left\| (S_\Gamma(x))^{1/2} \right\|_{L^\Phi(\mathbb{R}^d)}.$$

Sea $\Gamma_{\min} = \{Q \in \Gamma : Q = Q_x \text{ for some } x \in \cup_{Q \in \Gamma} Q\}$. Del Lema 3.5.6

$$\sum_{Q \in \Gamma} \frac{\chi_Q(x)}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2} \approx \sum_{Q \in \Gamma_{\min}} \frac{\chi_{\text{luz}(Q)}(x)}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2},$$

donde $\text{luz}(Q)$ se definió después de (3.62). Compárese la desigualdad anterior con (3.63). Como $\{\text{luz}(Q) : Q \in \Gamma\}$ es una familia disjunta de conjuntos, entonces

$$\begin{aligned} \left\| (S_\Gamma(x))^{1/2} \right\|_{L^\Phi(\mathbb{R}^d)} &\approx \left\| \left(\sum_{Q \in \Gamma_{\min}} \frac{\chi_{\text{luz}(Q)}(x)}{[\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))]^2} \right)^{1/2} \right\|_{L^\Phi(\mathbb{R}^d)} \\ &= \left\| \sum_{Q \in \Gamma_{\min}} \frac{\chi_{\text{luz}(Q)}(x)}{\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))} \right\|_{L^\Phi(\mathbb{R}^d)}. \end{aligned} \quad (3.72)$$

Sea $\lambda = h_\varphi^+(\nu_{\tilde{\Phi}}(\Gamma))$ y para cada $Q \in \Gamma$ escójase $u_Q = 1/\tilde{\Phi}(|Q|^{-1/r})$. Entonces, por cada $Q \in \Gamma$,

$$\lambda \geq \frac{\varphi(u_Q \nu_{\tilde{\Phi}}(\Gamma))}{\varphi(u_Q)} \Leftrightarrow \frac{1}{\lambda \varphi(u_Q)} \leq \frac{1}{\varphi(u_Q \nu_{\tilde{\Phi}}(\Gamma))}.$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Ya que Φ es creciente y $\{\text{luz}(Q) : Q \in \Gamma\}$ es una familia disjunta de conjuntos,

$$\begin{aligned}
 \int_{\mathbb{R}^d} \Phi \left(\frac{1}{\lambda} \sum_{Q \in \Gamma_{\min}} \frac{\chi_{\text{luz}(Q)}(x)}{\varphi(1/\tilde{\Phi}(|Q|^{-1/r}))} \right) dx &= \sum_{Q \in \Gamma_{\min}} \Phi \left(\frac{1}{\lambda \varphi(u_Q)} \right) |\text{luz}(Q)| \\
 &\leq \sum_{Q \in \Gamma_{\min}} \Phi \left(\frac{1}{\varphi(u_Q \nu_{\tilde{\Phi}}(\Gamma))} \right) |Q| \\
 &= \sum_{Q \in \Gamma_{\min}} \frac{1}{u_Q \nu_{\tilde{\Phi}}(\Gamma)} |Q| \\
 &= \frac{1}{\nu_{\tilde{\Phi}}(\Gamma)} \sum_{Q \in \Gamma_{\min}} |Q| \tilde{\Phi}(|Q|^{-1/r}) \leq 1.
 \end{aligned}$$

Esto muestra (3.71) por la definición de la norma de Luxemburg de $L^\Phi(\mathbb{R}^d)$. ■

Proposición 3.5.9 *Sea $L^\Phi(\mathbb{R}^d)$ un espacio de Orlicz con índices de Boyd en $(0, 1)$ y $0 < r < \infty$. Sea \mathcal{W} una base de ondículas r -, Φ -, $\tilde{\Phi}$ -admisibles. Sea $\Gamma = \cup_{\ell=1}^L \Gamma^\ell$ con $\Gamma_\ell \subset \mathcal{D}$ y tal que $\nu_\Phi(\Gamma) = \sum_{\ell=1}^L \sum_{Q \in \Gamma_\ell} |Q| \Phi(|Q|^{-1/r}) < \infty$. Existe $C > 0$ tal que*

$$\frac{1}{C} h_{\tilde{\Phi}}^-(\nu_\Phi(\Gamma)) \leq \sum_{\ell=1}^L \left\| \sum_{Q \in \Gamma_\ell} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r}} \right\|_{L^\Phi(\mathbb{R}^d)}. \quad (3.73)$$

Demostración. Comenzamos con un solo ℓ fijo y escribimos $\Gamma_\ell = \Gamma$ para simplificar la notación. Como \mathcal{W} es r -admissible, $\|\psi_Q^\ell\|_{H^r} \approx |Q|^{\frac{1}{r}-\frac{1}{2}}$ por (3.58). Ya que $0 < i_{\tilde{\Phi}} \leq I_{\tilde{\Phi}} < 1$, $L^{\tilde{\Phi}}(\mathbb{R}^d)$ es el dual de $L^\Phi(\mathbb{R}^d)$. Así, por la dualidad expresada en (3.55) se tiene

$$\begin{aligned}
 \left\| \sum_{Q \in \Gamma} \frac{\psi_Q^\ell(\cdot)}{\|\psi_Q^\ell\|_{H^r}} \right\|_{L^\Phi(\mathbb{R}^d)} &= \left\| \sum_{Q \in \Gamma} |Q|^{-1/r+1/2} \psi_Q^\ell(\cdot) \right\|_{L^\Phi(\mathbb{R}^d)} \\
 &= \sup_{\|g\|_{L^{\tilde{\Phi}}(\mathbb{R}^d)} \leq 1} \int_{\mathbb{R}^d} \left(\sum_{Q \in \Gamma} |Q|^{-1/r+1/2} \psi_Q^\ell(x) \right) g(x) dx \quad (3.74)
 \end{aligned}$$

Considérese

$$g(x) = \frac{1}{C_1 h_{\tilde{\Phi}}^+(\nu_\Phi(\Gamma))} \sum_{Q \in \Gamma} \frac{|Q|^{1/2} \psi_Q^\ell(x)}{\tilde{\varphi}(1/\Phi(|Q|^{-1/r}))},$$

donde C_1 es la constante en el Lema 3.5.8 cuando se aplica a $\tilde{\varphi}$, la función fundamental de $L^{\tilde{\Phi}}(\mathbb{R}^d)$. Por (3.71)

$$\|g\|_{L^{\tilde{\Phi}}(\mathbb{R}^d)} \leq 1,$$

así que por (3.74) se puede escribir

$$\begin{aligned}
 \left\| \sum_{Q \in \Gamma} \frac{\psi_Q^\ell(\cdot)}{\|\psi_Q^\ell\|_{H^r}} \right\|_{L^\Phi(\mathbb{R}^d)} &\geq \frac{1}{C_1 h_{\tilde{\varphi}}^+(\nu_\Phi(\Gamma))} \int_{\mathbb{R}^d} \left(\sum_{Q \in \Gamma} |Q|^{-1/r+1/2} \psi_Q^\ell(x) \right) \\
 &\quad \left(\sum_{Q \in \Gamma} \frac{|Q|^{1/2} \psi_Q^\ell(x)}{\tilde{\varphi}(1/\Phi(|Q|^{-1/r}))} \right) dx \\
 &= \frac{1}{C_1 h_{\tilde{\varphi}}^+(\nu_\Phi(\Gamma))} \sum_{Q \in \Gamma} \frac{|Q|^{-1/r+1}}{\tilde{\varphi}(1/\Phi(|Q|^{-1/r}))}. \tag{3.75}
 \end{aligned}$$

Usando $\tilde{\varphi}(t)\varphi(t) \approx t$, $0 < t < \infty$ (ver (3.56)), se deduce

$$\tilde{\varphi} \left(\frac{1}{\Phi(|Q|^{-1/r})} \right) \approx \frac{1/\Phi(|Q|^{-1/r})}{\varphi(1/\Phi(|Q|^{-1/r}))} = \frac{1}{|Q|^{1/r} \Phi(|Q|^{-1/r})}.$$

De (3.75) se obtiene

$$\begin{aligned}
 \left\| \sum_{Q \in \Gamma} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r}} \right\|_{L^\Phi(\mathbb{R}^d)} &\geq \frac{1}{C_1 h_{\tilde{\varphi}}^+(\nu_\Phi(\Gamma))} \sum_{Q \in \Gamma} |Q| \Phi(|Q|^{-1/r}) \\
 &= \frac{\nu_\Phi(\Gamma)}{C_1 h_{\tilde{\varphi}}^+(\nu_\Phi(\Gamma))}. \tag{3.76}
 \end{aligned}$$

Pero, $\tilde{\varphi}(t)\varphi(t) \approx t$, $0 < t < \infty$, implica

$$h_{\tilde{\varphi}}^+(t) = \sup_{s>0} \frac{\tilde{\varphi}(st)}{\tilde{\varphi}(s)} \approx \sup_{s>0} \frac{st\varphi(s)}{\varphi(st)s} = t \sup_{s>0} \frac{\varphi(s)}{\varphi(st)} = t \left(\inf_{s>0} \frac{\varphi(st)}{\varphi(s)} \right)^{-1} = \frac{t}{h_{\tilde{\varphi}}^-(t)},$$

y (3.76) da

$$\left\| \sum_{Q \in \Gamma} \frac{\psi_Q^\ell}{\|\psi_Q^\ell\|_{H^r}} \right\|_{L^\Phi(\mathbb{R}^d)} \gtrsim \frac{1}{C_1} h_{\tilde{\varphi}}^-(\nu_\Phi(\Gamma)) \tag{3.77}$$

como queríamos probar.

Para probar (3.73) se usa el Lema 3.5.4 para escribir

$$\frac{1}{C} h_{\tilde{\varphi}}^-(\nu_\Phi(\Gamma)) = \frac{1}{C} h_{\tilde{\varphi}}^-\left(\sum_{\ell=1}^L \nu_\Phi(\Gamma_\ell)\right) \leq \frac{1}{C} \sum_{\ell=1}^L h_{\tilde{\varphi}}^-(\nu_\Phi(\Gamma_\ell))$$

y se finaliza aplicando (3.77) a cada ℓ . ■

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

3.6 *Thresholding* restringido en espacios de sucesiones

Thresholding es una técnica muy utilizada en el procesamiento de señales para comprimirlas o para reducir su ruido. El algoritmo *thresholding* más sencillo para una señal f , denotado por $T_\varepsilon(f)$, se define, para $\varepsilon > 0$, como

$$f = \sum_{j=1}^{\infty} x_j \mathbf{e}_j \implies T_\varepsilon(f) = \sum_{j:|x_j|>\varepsilon} x_j \mathbf{e}_j. \quad (3.78)$$

suponiendo que los elementos $\{\mathbf{e}_j\}_{j=1}^{\infty}$ que se usan para la representación de la señal estén normalizados.

Es tradicional medir el error que se produce al aproximar f por $T_\varepsilon(f)$ en la misma norma en la que se tiene la representación de la izquierda de (3.78), generalmente en un espacio L^2 . Sin embargo, se argumenta en [66] que si el error es pequeño en L^p o, como se dice en la comunidad estadística, la aproximación tiene un riesgo- L^p pequeño, $T_\varepsilon(f)$ puede ser visualmente una mejor aproximación de f que si tiene un riesgo- L^2 pequeño.

Una situación más general se considera en [24], donde se tiene una representación de una señal f en una base de ondículas en un espacio de Hardy H^r , $0 < r < \infty$ (observar que $H^r = L^r$ si $1 < r < \infty$), pero el error se mide en un espacio de Hardy H^p , $0 < p < \infty$. Se muestra en [24] que esta situación es equivalente a realizar una aproximación no lineal restringida (ver Sección 3.4.2 para la definición).

El objetivo de esta sección es extender este resultado de [24] a espacios cuasi-Banach más generales que los de Hardy, pero que tengan una caracterización en términos de una cuasi-norma en un espacio de sucesiones. Sabemos que esto es cierto para muchos espacios clásicos del Análisis.

Para ello consideramos $S = \mathbb{C}^{\mathcal{D}}$ el conjunto de todas las sucesiones de números complejos $\mathbf{s} = \{s_I\}_{I \in \mathcal{D}}$ definidas sobre un conjunto numerable \mathcal{D} y $\mathcal{E} = \{\mathbf{e}_I\}_{I \in \mathcal{D}}$ la base canónica (ver Sección 3.4.1). Sea $\mathbf{u} = \{u_I\}_{I \in \mathcal{D}} \in S$ con $u_I > 0$ para todo $I \in \mathcal{D}$ una **sucesión de pesos**.

Definición 3.6.1 Dado $\mathbf{x} = \sum_{I \in \mathcal{D}} x_I \mathbf{e}_I \in S$, $\lambda > 0$ y \mathbf{u} una sucesión de pesos, el **operador de thresholding restringido** $T_\lambda(\mathbf{x}, \mathbf{u})$ se define como

$$T_\lambda(\mathbf{x}, \mathbf{u}) = \sum_{I:|x_I|>\lambda} x_I \mathbf{e}_I. \quad (3.79)$$

3.6 Thresholding restringido en espacios de sucesiones

En esta sección se mostrará que la velocidad de decaimiento de $\|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f$ (f es un retículo cuasi-Banach de sucesiones) permite caracterizar los espacios de aproximación restringida $\mathcal{A}_\mu^\xi(f, \nu)$ (ν es una medida positiva sobre \mathcal{D}). Mostraremos que para esto es esencial que se cumpla la propiedad generalizada de Temlyakov, a saber,

$$\frac{1}{C}\eta(\nu(\Gamma)) \leq \left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f \leq C\eta(\nu(\Gamma)). \quad (3.80)$$

Antes de enunciar con detalle estos resultados y dar su demostración, necesitamos exponer una nueva propiedad de las funciones de la clase \mathbb{W}_+ (ver Definiciones 3.4.9 y 3.4.10).

Lema 3.6.2 *Para cualquier $\eta \in \mathbb{W}_+$ existen $\rho > 0$ suficientemente pequeña y $C > 0$ tales que*

$$\eta^\rho \left(\sum_{j=0}^{\infty} x_j \right) \leq C \sum_{j=0}^{\infty} \eta^\rho(x_j)$$

para cualesquiera sucesiones $\{x_j\}_{j=0}^{\infty} \subset [0, \infty)$ con $\sum_{j=0}^{\infty} x_j < \infty$.

Demostración. Escribábase $M_\eta(t) = \sup_{0 < t < \infty} \frac{\eta(st)}{\eta(s)}$ para la función de dilatación de η y sean

$$i_\eta = \lim_{t \rightarrow 0} \frac{\log M_\eta(t)}{\log t} = \sup_{0 < t < 1} \frac{\log M_\eta(t)}{\log t} \geq 0$$

y

$$I_\eta = \lim_{t \rightarrow \infty} \frac{\log M_\eta(t)}{\log t} = \inf_{t > 1} \frac{\log M_\eta(t)}{\log t} \geq 0$$

los índices de dilatación de η (ver p. 53 de [69]).

Como $\eta \in \mathbb{W}_+$, $i_\eta > 0$. De hecho, si $i_\eta = 0$, $\frac{\log M_\eta(t)}{\log t} = 0$ para todo $t \in (0, 1)$; así pues $\log M_\eta(t) = 0$ para todo $t \in (0, 1)$ y se tiene $M_\eta(t) = 1$ para todo $t \in (0, 1)$, lo que contradice la definición de \mathbb{W}_+ . Además, η doblante implica que $I_\eta < \infty$. De hecho, $M_\eta(2) = \sup_{0 < s < \infty} \frac{\eta(s2)}{\eta(s)} \leq C$, así que $I_\eta \leq \frac{\log M_\eta(2)}{\log 2} \leq \frac{C}{\log 2} < \infty$.

Ahora, tómesese $\rho > 0$ tal que $\rho < 1/I_\eta$. Ya que $M_{\eta^\rho}(t) = [M_\eta]^\rho$ se deduce $i_{\eta^\rho} = \rho i_\eta > 0$ y $I_{\eta^\rho} = \rho I_\eta < 1$. Por el Corolario 2 en la p. 55 de [69], η^ρ es equivalente a su menor mayorante cóncava. Así pues, existe φ cóncava en $[0, \infty)$ tal que

$$A\varphi(t) \leq \eta^\rho(t) \leq B\varphi(t), \quad 0 < t < \infty \quad (3.81)$$

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

con $0 < A \leq B < \infty$ independiente de t . Como $\eta(0) = 0$ se tiene que $\varphi(0) = 0$. También, ya que φ es cóncava, φ/t es decreciente (ver p. 47 en [69]). Pero entonces, φ es sub-aditiva, pues para $t_1, t_2 \in [0, \infty)$:

$$\begin{aligned} \varphi(t_1 + t_2) &= t_1 \frac{\varphi(t_1 + t_2)}{t_1 + t_2} + t_2 \frac{\varphi(t_1 + t_2)}{t_1 + t_2} \\ &\leq t_1 \frac{\varphi(t_1)}{t_1} + t_2 \frac{\varphi(t_2)}{t_2} = \varphi(t_1) + \varphi(t_2), \end{aligned} \quad (3.82)$$

(este argumento se toma de la p. 55 de [69]). La prueba se termina usando (3.81) y (3.82):

$$\begin{aligned} \left[\eta \left(\sum_{j=0}^{\infty} x_j \right) \right]^{\rho} &\leq B \varphi \left(\sum_{j=0}^{\infty} x_j \right) \leq B \sum_{j=0}^{\infty} \varphi(x_j) \\ &\leq B \sum_{j=0}^{\infty} \frac{1}{A} [\eta(x_j)]^{\rho}, \end{aligned}$$

donde hemos usado la continuidad de φ en la segunda desigualdad (cualquier función creciente cóncava es continua). ■

Nota 3.6.3 *Este lema muestra que, excepto por una constante C , la potencia ρ de la función $\eta \in \mathbb{W}_+$ es subaditiva (lo cual denotaremos como $\eta \in \mathcal{PS}$) para alguna $\rho > 0$ pequeña.*

3.6.1 La caracterización por *thresholding*

Damos a continuación la caracterización de los espacios de aproximación restringida mediante *thresholding*, basándonos en la propiedad de Temlyakov.

Los siguientes resultados son generalizaciones de los contenidos en [24] y [67]. Sea $\mathcal{C}_+ := \{\eta : [0, \infty) \mapsto [0, \infty] : \eta(0) = 0, \lim_{t \rightarrow \infty} \eta(t) = \infty, \eta \text{ continua, finita en todas partes y creciente}\}$.

Proposición 3.6.4 *Sea (f, ν) un esquema canon (ver Definición 3.4.4). Para $\eta \in \mathcal{C}_+$ las siguientes condiciones son equivalentes:*

1. Existe $C > 0$ tal que para todo $\Gamma \subset \mathcal{D}$,

$$\left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f \leq C \eta(\nu(\Gamma)). \quad (3.83)$$

3.6 Thresholding restringido en espacios de sucesiones

2. Para todo $\varepsilon \in (0, 1)$ y escribiendo $\eta_1(t) = \eta^{1/\varepsilon}(t)$, se tiene

$$\sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f \leq C \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon. \quad (3.84)$$

Demostración. Comenzamos probando que $1 \Rightarrow 2$. Para $\lambda > 0$, sean $\Lambda_j = \{I \in \mathcal{D} : 2^{-j-1}\lambda < |x_I| u_I \leq 2^{-j}\lambda\}$ y $\Gamma_j = \{I \in \mathcal{D} : 2^{-j-1}\lambda < |x_I| u_I\}$. Si f satisface la desigualdad ρ -triangular, entonces usando la propiedad 1,

$$\|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f^\rho \leq C^\rho \sum_{j=0}^{\infty} (2^{-j}\lambda)^\rho \eta^\rho(\nu(\Lambda_j)).$$

Como η es creciente y $\Lambda_j \subset \Gamma_j$, las definiciones de η_1 y $\ell_{\eta_1}^\infty(\nu, \mathbf{u})$ dan

$$\eta(\nu(\Lambda_j)) \leq \eta(\nu(\Gamma_j)) \lesssim (2^{-j-1}\lambda)^{-\varepsilon} \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon.$$

Entonces,

$$\begin{aligned} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f^\rho &\leq C^\rho \sum_{j=0}^{\infty} (2^{-j}\lambda)^\rho (2^{-j-1}\lambda)^{-\rho\varepsilon} \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^{\rho\varepsilon} \\ &= C_{\rho, \varepsilon} \lambda^{\rho(1-\varepsilon)} \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^{\rho\varepsilon}, \end{aligned}$$

ya que $\varepsilon < 1$. Esto prueba $1. \Rightarrow 2$.

Para probar que $2 \Rightarrow 1$, considérese $\mathbf{x} = \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I}$. Entonces,

$$\|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f = 0, \quad \lambda \leq 1 \quad \text{y} \quad \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f = \|\mathbf{x}\|_f, \quad \lambda > 1.$$

Por lo tanto, se tiene por una parte que

$$\sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f = \|\mathbf{x}\|_f.$$

Por otra parte

$$\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})} \approx \sup_{\lambda > 0} \lambda \eta_1(\nu(I : |x_I| u_I > \lambda)) = \eta_1(\nu(\Gamma)).$$

Por la definición de η_1 ,

$$\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon = \eta(\nu(\Gamma)),$$

y $2. \Rightarrow 1$. se prueba de inmediato. ■

Proposición 3.6.5 *Sea (f, ν) un esquema canon y \mathbf{u} una sucesión de pesos. Sea $\eta \in \mathcal{C}_+$ tal que $\eta \in \mathcal{PS}$. Las siguientes condiciones son equivalentes:*

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

1. Existe $C > 0$ tal que para todo $\Gamma \subset \mathcal{D}$,

$$\frac{1}{C}\eta(\nu(\Gamma)) \leq \left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f. \quad (3.85)$$

2. Para todo $\varepsilon \in (0, 1)$ y escribiendo $\eta_1(t) = \eta^{1/\varepsilon}(t)$, se tiene

$$\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon \leq C \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f. \quad (3.86)$$

Demostración. Comenzamos demostrando que 1. \Rightarrow 2. Sea $M = \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f$.

Para $\lambda > 0$, let $\Lambda_j = \{I \in \mathcal{D} : 2^j \lambda < |x_I| u_I \leq 2^{j+1} \lambda\}$ y $\Gamma_j = \{I \in \mathcal{D} : |x_I| u_I \leq 2^{j+1} \lambda\}$.

Tomando ρ como en la definición de \mathcal{PS} , se obtiene

$$\eta^\rho(\nu(I : |x_I| u_I > \lambda)) = \eta^\rho \left(\sum_{j=0}^{\infty} \nu(\Lambda_j) \right) \leq C \sum_{j=0}^{\infty} \eta^\rho(\nu(\Lambda_j)).$$

Como η es creciente y $\Lambda_j \subset \Gamma_j$, (3.85) da

$$\begin{aligned} \eta^\rho(\nu(I : |x_I| u_I > \lambda)) &\leq C \sum_{j=0}^{\infty} \left\| \sum_{I \in \Lambda_j} \frac{\mathbf{e}_I}{u_I} \right\|_f^\rho \\ &\leq C \sum_{j=0}^{\infty} \frac{1}{(2^j \lambda)^\rho} \left\| \sum_{I \in \Lambda_j} x_I \mathbf{e}_I \right\|_f^\rho \\ &\leq C \sum_{j=0}^{\infty} \frac{1}{(2^j \lambda)^\rho} \left\| \sum_{I \in \Gamma_j} x_I \mathbf{e}_I \right\|_f^\rho \\ &= C \sum_{j=0}^{\infty} \frac{1}{(2^j \lambda)^\rho} \|\mathbf{x} - T_{2^{j+1}\lambda}(\mathbf{x}, \mathbf{u})\|_f^\rho, \end{aligned}$$

donde las segunda y tercera desigualdades son debidas a que f es un retículo. Por la definición de M se deduce

$$\begin{aligned} \eta^\rho(\nu(I : |x_I| u_I > \lambda)) &\leq C \sum_{j=0}^{\infty} \frac{1}{(2^j \lambda)^\rho} (2^{j+1} \lambda)^{\rho(1-\varepsilon)} M^\rho \\ &= C M^\rho \lambda^{-\rho\varepsilon} \sum_{j=0}^{\infty} 2^{-j\rho\varepsilon} = C_{\rho, \varepsilon} M^\rho \lambda^{-\rho\varepsilon}. \end{aligned}$$

De ahí que

$$\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon = \sup_{\lambda > 0} \eta(\nu(I : |x_I| u_I \geq \lambda)) \lambda^\varepsilon \leq C_{\rho, \varepsilon} M.$$

Esto prueba 1. \Rightarrow 2. La implicación 2. \Rightarrow 1. se prueba de manera similar como en la Proposición 3.6.4. ■

Como corolario tenemos la siguiente caracterización.

Teorema 3.6.6 *Sea (f, ν) un esquema canon y \mathbf{u} una sucesión de pesos. Sea $\eta \in \mathcal{C}_+$ tal que $\eta \in \mathcal{PS}$. Las siguientes condiciones son equivalentes:*

1. *Existe $C > 0$ tal que para todo $\Gamma \subset \mathcal{D}$ (no necesariamente finito)*

$$\left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_{\mathbb{X}} \approx \eta(\nu\{\Gamma\}). \quad (3.87)$$

2. *Para todo $\varepsilon \in (0, 1)$ y escribiendo $\eta_1(t) = \eta^{1/\varepsilon}(t)$, se tiene*

$$\sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f \approx \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon. \quad (3.88)$$

Nota 3.6.7 *Puesto que $\mathbb{W}_+ \subset \mathcal{PS}$ debido al Lema 3.6.2, el Teorema 3.6.6 es cierto con la hipótesis $\eta \in \mathbb{W}_+$. Esta condición es también suficiente para que se cumplan los Teoremas 3.4.17 y 3.4.21, con lo que se tiene la caracterización de espacios de aproximación usando thresholding.*

Nota 3.6.8 *Por el Lema 3.6.2 sabemos que $\mathbb{W}_+ \subset \mathcal{PS}$. La inclusión es estricta ya que $\eta(t) = \log(t+1) \notin \mathbb{W}_+$ (ya que $i_\eta = 0$) y $\eta \in \mathcal{PS}$. Esto último es claro ya que η es subaditiva: para $t_1, t_2 \geq 0$*

$$t_1 + t_2 + 1 \leq (t_1 + 1)(t_2 + 1) \Rightarrow \eta(t_1 + t_2) \leq \eta(t_1) + \eta(t_2).$$

Nota 3.6.9 *Sea $p \in (0, \infty)$ y tomar $\eta(t) = t^p \in \mathbb{W}_+$. El Teorema 3.6.6 con $\varepsilon = q/p < 1$ produce la equivalencia 1. \Leftrightarrow 2. de la Proposición 2 de [67] para retículos de sucesiones. Tomando el espacio de sucesiones $f_p^{0,2}$ asociado a H^p , $0 < p < \infty$, se obtiene el Teorema 7.1 de [24] con $\varepsilon = \mu/p < 1$*

3.7 Cuantización restringida

Puede obtenerse otra caracterización de los espacios de aproximación por medio de la cuantización. Este proceso mapea valores en \mathbb{R} a ciertos valores predeterminados con

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

el fin de almacenarlos o transmitirlos de forma digital. Definimos el **operador de cuantización restringida** como

$$Q_\lambda(\mathbf{x}, \mathbf{u}) = \sum_{I \in \mathcal{D}} \operatorname{sgn} x_I \left[\frac{|x_I|}{\lambda/u_I} \right] \frac{\lambda}{u_I} \mathbf{e}_I, \quad (3.89)$$

donde, para todo $x \geq 0$, $[x]$ es el mayor entero menor o igual que x , es decir, $[x] \leq x < [x] + 1$.

Proposición 3.7.1 *Sea (f, ν) un esquema canon (ver Definición 3.4.4). Para $\eta \in \mathbb{C}_+$, cada una de las condiciones (3.83) y (3.84) es equivalente a:*

(3) Para $\varepsilon \in (0, 1)$ y escribiendo $\eta_1(t) = \eta^{1/\varepsilon}(t)$, se tiene

$$\sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \leq C \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon. \quad (3.90)$$

Demostración. (3.90) \Rightarrow (3.83). Considérese $\mathbf{x} = \sum_{I \in \Gamma} \mathbf{e}_I/u_I$. Observar que

$$Q_\lambda(\mathbf{x}, \mathbf{u}) = \sum_{I \in \Gamma} \left[\frac{1}{\lambda} \right] \frac{\lambda}{u_I} \mathbf{e}_I.$$

Si $\lambda > 1$, $Q_\lambda(\mathbf{x}, \mathbf{u}) = 0$ y $\|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f = \|\mathbf{x}\|_f$. Para $\lambda \leq 1$,

$$\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u}) = \sum_{I \in \Gamma} \left(1 - \left[\frac{1}{\lambda} \right] \lambda\right) \frac{\mathbf{e}_I}{u_I}.$$

Puesto que $[\frac{1}{\lambda}] \leq \frac{1}{\lambda} < [\frac{1}{\lambda}] + 1$, se deduce que

$$\left| 1 - \left[\frac{1}{\lambda} \right] \lambda \right| = 1 - \left[\frac{1}{\lambda} \right] \lambda < 1 + \lambda - 1 = \lambda.$$

Por tanto, para $\lambda \leq 1$,

$$\|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \leq \lambda \left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f = \lambda \|\mathbf{x}\|_f,$$

ya que f es un retículo. Esto implica

$$\sup_{\lambda > 1} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f = \sup_{\lambda > 1} \lambda^{-(1-\varepsilon)} \|\mathbf{x}\|_f = \|\mathbf{x}\|_f$$

y

$$\sup_{\lambda \leq 1} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \leq \sup_{\lambda \leq 1} \lambda^{-(1-\varepsilon)} \lambda \|\mathbf{x}\|_f = \|\mathbf{x}\|_f.$$

Por tanto

$$\sup_{\lambda>0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f = \|\mathbf{x}\|_f. \quad (3.91)$$

Razonando como en la demostración del Lema 3.4.15 a) se tiene

$$\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})} = \eta_1(\nu(\Gamma)) = [\eta(\nu(\Gamma))]^{1/\varepsilon}. \quad (3.92)$$

Usando (3.90) da

$$\left\| \sum_{I \in \Gamma} \frac{\mathbf{e}_I}{u_I} \right\|_f = \|\mathbf{x}\|_f \leq C \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon = C \eta(\nu(\Gamma)).$$

(3.83) y (3.84) \Rightarrow (3.90). De la desigualdad ρ -triangular

$$\|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \leq C_\rho (\|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f + \|T_\lambda(\mathbf{x}, \mathbf{u}) - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f).$$

Si $|x_I| u_I < \lambda$, $\lceil \frac{|x_I| u_I}{\lambda} \rceil = 0$. Por tanto

$$\begin{aligned} & Q_\lambda(\mathbf{x}, \mathbf{u}) - T_\lambda(\mathbf{x}, \mathbf{u}) \\ &= \sum_{|x_I| u_I \geq \lambda} \operatorname{sgn} x_I \lceil \frac{|x_I| u_I}{\lambda} \rceil \frac{\lambda}{u_I} \mathbf{e}_I - \sum_{|x_I| u_I > \lambda} x_I \mathbf{e}_I \\ &= \lambda \left[\sum_{|x_I| u_I > \lambda} \left(\operatorname{sgn} x_I \lceil \frac{|x_I| u_I}{\lambda} \rceil - \frac{|x_I| u_I}{\lambda} \right) \frac{\mathbf{e}_I}{u_I} + \sum_{|x_I| u_I = \lambda} \operatorname{sgn} x_I \frac{\mathbf{e}_I}{u_I} \right]. \end{aligned}$$

Por la definición de $[x]$, el valor absoluto de los coeficientes de \mathbf{e}_I/u_I en la expresión anterior es menor o igual que 1. Por la propiedad de monotonía de f podemos escribir

$$\|T_\lambda(\mathbf{x}, \mathbf{u}) - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \leq \lambda \left\| \sum_{|x_I| u_I > \lambda} \frac{\mathbf{e}_I}{u_I} \right\|_f.$$

Por lo tanto, usando (3.84) y (3.83)

$$\begin{aligned} & \sup_{\lambda>0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \\ & \leq C_\rho \left(\sup_{\lambda>0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f + \sup_{\lambda>0} \lambda^\varepsilon \left\| \sum_{I: |x_I| u_I > \lambda} \frac{\mathbf{e}_I}{u_I} \right\|_f \right) \\ & \leq C_\rho \left(C \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon + C' \sup_{\lambda>0} \lambda^\varepsilon \eta(\{I : |x_I| u_I > \lambda\}) \right) \\ & = C'_\rho \left(\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon + \left(\sup_{\lambda>0} \lambda \eta_1(\nu(\{I : |x_I| u_I > \lambda\})) \right)^\varepsilon \right) \\ & = C_\rho \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon, \end{aligned}$$

porque $\eta_1 = \eta^{1/\varepsilon}$. ■

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

Proposición 3.7.2 Sea (f, ν) un esquema canon y \mathbf{u} una sucesión de pesos. Sea $\eta \in \mathcal{C}_+$ tal que $\eta \in \mathcal{PS}$. Entonces, cada una de las condiciones (3.85) y (3.86) es equivalente a:

(3) Para $\varepsilon \in (0, 1)$ y escribiendo $\eta_1(t) = \eta^{1/\varepsilon}(t)$, se tiene

$$\|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon \leq C \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f. \quad (3.93)$$

Demostración. Observar que (3.85) y (3.86) son equivalentes por la Proposición 3.6.5.

(3.93) \Rightarrow (3.85). Como en ocasiones anteriores, elíjase $\mathbf{x} = \sum_{I \in \Gamma} \frac{e_I}{u_I}$. Usando (3.91) y (3.93) se obtiene

$$\begin{aligned} \|\mathbf{x}\|_f &= \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f \geq C \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon \\ &= C\eta(\nu(\Gamma)), \end{aligned}$$

donde la última igualdad es como en (3.92).

(3.86) \Rightarrow (3.93). Para $\mathbf{x} = \sum_{I \in \Gamma} x_I \mathbf{e}_I$ se tiene

$$\mathbf{x} - T_{\lambda/2}(\mathbf{x}, \mathbf{u}) = \sum_{|x_I|u_I \leq \lambda/2} x_I \mathbf{e}_I$$

y

$$\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u}) = \sum_{|x_I|u_I < \lambda} x_I \mathbf{e}_I + \sum_{|x_I|u_I > \lambda} (x_I - \operatorname{sgn} x_I \lfloor \frac{|x_I|u_I}{\lambda} \rfloor \frac{\lambda}{u_I}) \mathbf{e}_I.$$

Por la propiedad de monotonía de f ,

$$\|\mathbf{x} - T_{\lambda/2}(\mathbf{x}, \mathbf{u})\|_f \leq \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f.$$

De (3.86) deducimos

$$\begin{aligned} \|\mathbf{x}\|_{\ell_{\eta_1}^\infty(\nu, \mathbf{u})}^\varepsilon &\leq C \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - T_\lambda(\mathbf{x}, \mathbf{u})\|_f \\ &\leq C \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_{2\lambda}(\mathbf{x}, \mathbf{u})\|_f \\ &= C 2^{1-\varepsilon} \sup_{\lambda > 0} (2\lambda)^{-(1-\varepsilon)} \|\mathbf{x} - Q_{2\lambda}(\mathbf{x}, \mathbf{u})\|_f \\ &= C' \sup_{\lambda > 0} \lambda^{-(1-\varepsilon)} \|\mathbf{x} - Q_\lambda(\mathbf{x}, \mathbf{u})\|_f, \end{aligned}$$

que es (3.93). ■

3.8 Experimentos

En esta sección mostramos algunos experimentos con imágenes en los que se comparan los resultados de hacer *thresholding* de los coeficientes en $L^2 \approx \ell^2 \approx f_{2,2}^0$ con los de hacer *thresholding* en $f_{2,2}^s$ con parámetros de suavidad $s = -1, -2$. Comparamos el uso de los *thresholding* usual y restringido tanto para el problema de la compresión (aproximación a la señal original con el menor número de elementos de la base de ondículas) como para el de reducción de ruido en la señal (estimación de la señal original).

Nuestra señal original es la típica de Lena lena512.bmp (262144 bytes) a la cual denotaremos con \mathbf{x} y que puede verse en la Figura 3.1. A \mathbf{x} se le añade ruido en forma de una matriz de 512×512 con entradas gaussianas $r \times N(0, 1)$, donde r es un número que regula la energía del ruido. En todos los ejemplos a continuación se ha tomado $r = 15$. Esta señal se muestra en la Figura 3.2. A la señal contaminada se le denotará con $\hat{\mathbf{x}}$. La diferencia en ℓ^2 entre \mathbf{x} y $\hat{\mathbf{x}}$ es

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = \left(\sum_{i=1}^{512} \sum_{j=1}^{512} (x_{i,j} - \hat{x}_{i,j})^2 \right)^{1/2}.$$

Otra medida habitual de comparación entre \mathbf{x} y $\hat{\mathbf{x}}$ es la de la razón de la energía de la señal a la energía del ruido SNR definido como:

$$SNR = 10 \log_{10} \left(\frac{\|\mathbf{x}\|_2}{\|\mathbf{x} - \hat{\mathbf{x}}\|_2} \right).$$

El SNR entre \mathbf{x} y $\hat{\mathbf{x}}$ es 12.87.

Generalmente, el objetivo consiste en comprimir lo más posible la señal (aproximarla con el menor número posible de coeficientes de ondículas) sin que se degrade visualmente o estimar la señal \mathbf{x} a partir de $\hat{\mathbf{x}}$ de manera que el SNR sea lo más grande posible. Veremos que el hecho de hacer *thresholding* restringido nos da compresiones y estimaciones que pueden tener menor SNR que las estimaciones con *thresholding* usual, pero que visualmente son mejores. En algunos casos los resultados con *thresholding* restringido son con menos coeficientes que con *thresholding* usual. En muchos casos el *thresholding* restringido mejora tanto el SNR como la percepción visual. Incluso cuando visualmente no hay una mejora clara el SNR del *thresholding* restringido es mayor que el del *thresholding* usual.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

La señal original \mathbf{x} se representa con una descomposición de ondículas de 5 niveles como

$$\mathbf{x} = \sum_{|Q|=1} \alpha_Q \varphi_Q + \sum_{j=0}^4 \sum_{|Q|=2^{-2j}} \beta_Q \psi_Q, \quad (3.94)$$

donde φ_Q son las funciones de escala (los coeficientes α_Q dan la tendencia de la imagen) y ψ_Q son las ondículas (los coeficientes β_Q dan los detalles de la imagen). En los experimentos usamos las ondículas de Haar denotadas con *db1* y las de Daubechies 3 denotadas por *db3*. La señal contaminada por ruido gaussiano se representa por

$$\hat{\mathbf{x}} = \sum_{|Q|=1} \hat{\alpha}_Q \varphi_Q + \sum_{j=0}^4 \sum_{|Q|=2^{-2j}} \hat{\beta}_Q \psi_Q. \quad (3.95)$$

La aproximación de \mathbf{x} haciendo *thresholding* sobre \mathbf{x} se representa como

$$T_\lambda(\mathbf{x}, f_{2,2}^s) = \sum_{|Q|=1} \alpha_Q \varphi_Q + \sum_{j=0}^4 \sum_{\substack{|Q|=2^{-2j} \\ \|\beta_Q \psi_Q\|_{f_{2,2}^s} > \lambda}} \beta_Q \psi_Q. \quad (3.96)$$

Cuando $s = 0$ (*thresholding* usual) se tiene $\|\beta_Q \psi_Q\|_{\ell_2} = |\beta_Q|$ porque las ψ_Q son ortonormales. Cuando $s = -1, -2$ (*thresholding* restringido) se tiene $\|\beta_Q \psi_Q\|_{f_{2,2}^s} = |\beta_Q| |Q|^{-s/2} = |\beta_Q| 2^{sj}$.

La estimación de \mathbf{x} haciendo *thresholding* sobre $\hat{\mathbf{x}}$ se representa como

$$T_\lambda(\hat{\mathbf{x}}, f_{2,2}^s) = \sum_{|Q|=1} \hat{\alpha}_Q \varphi_Q + \sum_{j=0}^4 \sum_{\substack{|Q|=2^{-2j} \\ \|\hat{\beta}_Q \psi_Q\|_{f_{2,2}^s} > \lambda}} \hat{\beta}_Q \psi_Q. \quad (3.97)$$

Así pues, el SNR al hacer *thresholding* es

$$SNR = 10 \log_{10} \left(\frac{\|\mathbf{x}\|_2}{\|\mathbf{x} - T_\lambda(\mathbf{y}, f_{2,2}^s)\|_2} \right),$$

donde $\mathbf{y} = \mathbf{x}$ en el caso sin ruido e $\mathbf{y} = \hat{\mathbf{x}}$ en el caso con ruido.

En las Figuras 3.3 y 3.4 se realizan un *thresholding* usual y uno restringido, respectivamente, sobre la señal sin ruido para el problema de la aproximación/compresión. La descomposición de ondículas se hace con la base de Haar en ambos casos. El *thresholding* usual ($s = 0$) se hace con el nivel del umbral (*threshold*) $\lambda = 30$. El *thresholding*

restringido se hace con $s = -1$ y el nivel del umbral $\lambda = 3$. Para el *thresholding* usual el error en norma ℓ^2 es $E2 \sim 649$ y para el restringido es $E2 = 681$. El número de coeficientes ($NoC \simeq 12266$) es casi el mismo para ambos y representa el 4.68% de los datos originales. Aunque el $SNR = 13.02$ es ligeramente mayor para la Figura 3.3 que el $SNR = 12.8$ para la Figura 3.4, visualmente la Figura 3.4 está menos “píxeleada”.

En las Figuras 3.5, 3.6 y 3.7 se realizan un *thresholding* usual y dos restringidos, respectivamente, para el problema de la estimación de la señal con ruido. La descomposición se hace con ondículas de Haar. El primer *thresholding* restringido se hace con $s = -1$ y $\lambda = 3$, y tiene un error en norma $\ell^2 = 567$ que da un $SNR = 13.61$ mayor que el del *thresholding* usual ($s = 0$) con umbral $\lambda = 30$ cuyo error en norma ℓ^2 es $E2 = 578$ que da un $SNR = 13.53$. Esto se consigue a pesar de que el número de coeficientes para el primer *thresholding* restringido, $NoC = 24380$, es menor que el número de coeficientes del *thresholding* usual, $NoC = 27581$, que representa el 9.5% del total de datos. Visualmente, hay una gran mejora de la Figura 3.5 a la Figura 3.6, ya que la primera está muy “píxeleada”. De la Figura 3.6 a la Figura 3.7 (con parámetros $s = -2$ y $\lambda = 0.35$) la mejora no es obvia pero se recuperan algunos detalles. Para la Figura 3.7 se tiene un $SNR = 13.15$ con un número de coeficientes $NoC = 26591$.

En las Figuras 3.8 y 3.9 se realizan un *thresholding* usual y uno restringido, respectivamente, para el problema de la estimación de la señal con ruido. En ambos casos se usan ondículas de Daubechies 3, *db3*. Para el caso restringido ($s = -1$, $\lambda = 3$) se tiene un $SNR = 14.41$ que es mayor que para el usual ($s = 0$, $\lambda = 30$) con un $SNR = 14.12$. Esta mejora en el SNR se obtiene a pesar de que para el *thresholding* restringido el número de coeficientes, $NoC = 24460$, es menor que el del *thresholding* usual, $NoC = 25325$. Visualmente, también hay una mejora ya que hay elementos espurios (*spikes*) en las zonas suaves de la Figura 3.8 que no están tan acentuados en la Figura 3.9. Además, se recuperan detalles en la parte superior del sombrero en el caso restringido.

En las Figuras 3.10 y 3.11 se realizan un *thresholding* usual ($s = 0$, $\lambda = 20$) y uno restringido ($s = -1$, $\lambda = 2$), respectivamente, para el problema de la estimación de la señal con ruido. Se usan ondículas de Daubechies 3, *db3*. El número de coeficientes para el *thresholding* restringido, $NoC = 43292$, es sólo $\sim 2/3$ que para el *thresholding* usual, $NoC = 64707$. Sin embargo, el SNR es mayor para el caso restringido ($SNR = 14.13$)

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES



Figure 3.1: Imagen de Lena original.

que para el usual ($SNR = 13.5$). Visualmente, son muy similares ya que estamos tomando un gran número de coeficientes.

En las Figuras 3.12 y 3.13 se realizan un *thresholding* usual ($s = 0$, $\lambda = 40$) y uno restringido ($s = -1$, $\lambda = 5.9$), respectivamente, para el problema de la estimación de la señal con ruido. Se usan ondículas de Daubechies 3, *db3*. El SNR es menor para el caso restringido, $SNR = 13.95$, que para el usual, $SNR = 14.19$. Sin embargo, visualmente, con el *thresholding* restringido se recuperan detalles en la parte superior del sombrero.

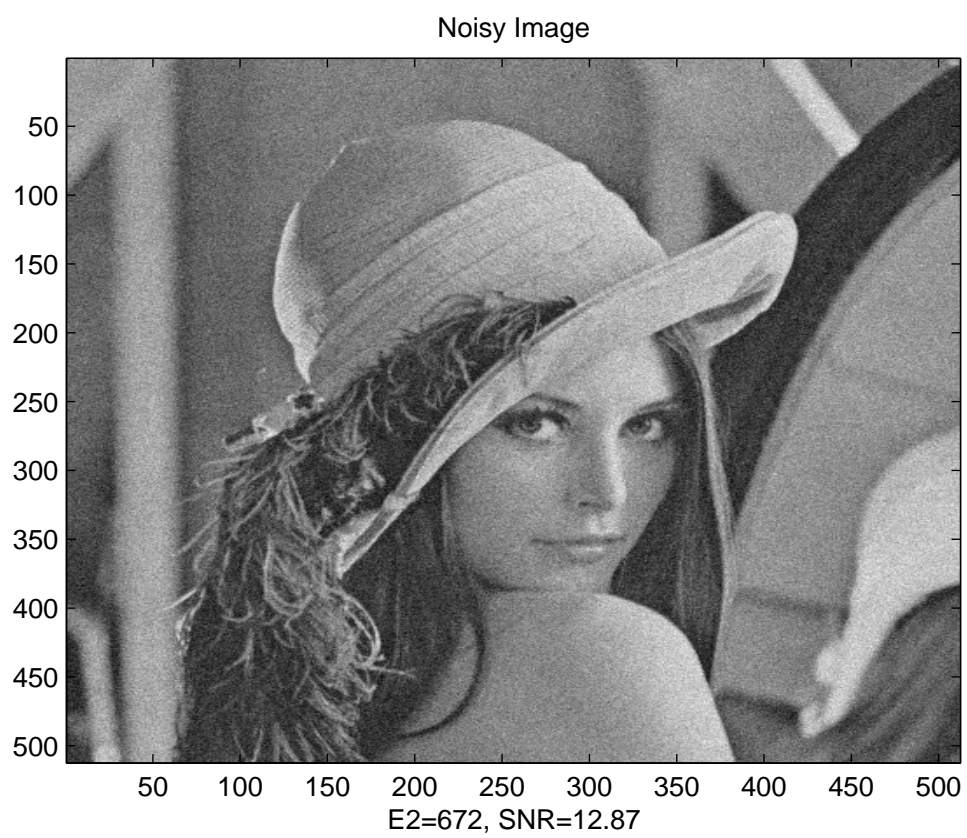


Figure 3.2: Imagen de Lena con ruido.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

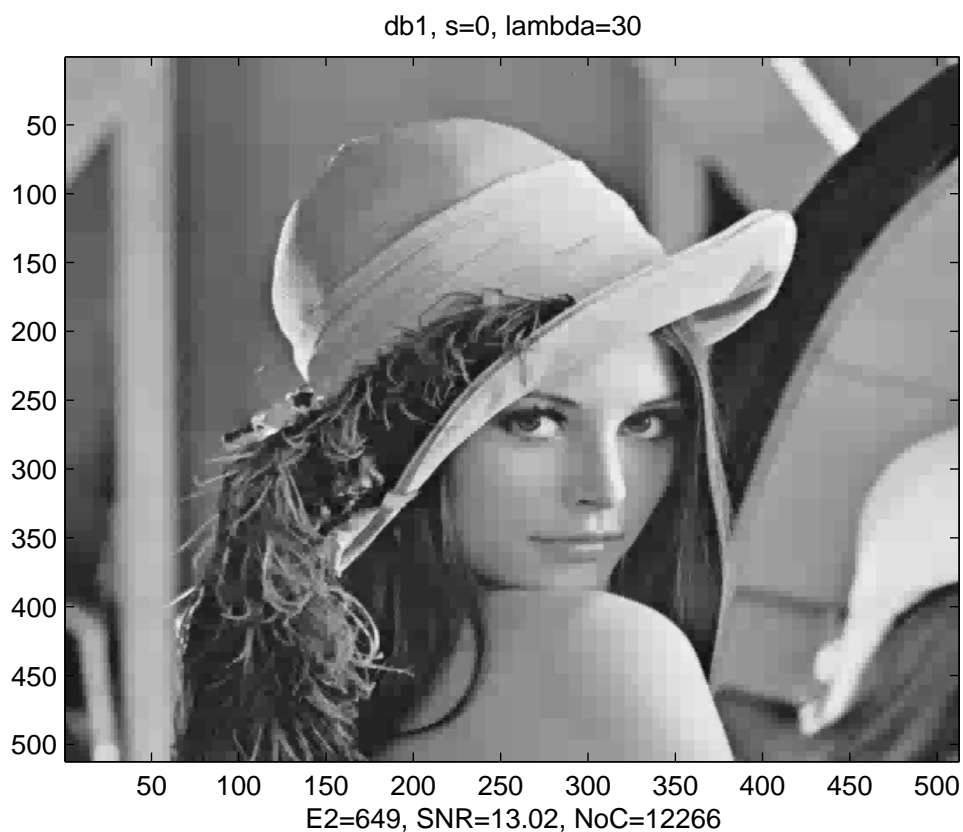


Figure 3.3: Aproximación de Lena a partir de la señal original haciendo *thresholding* usual con parámetros $s = 0$, $\lambda = 30$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 649$ y da un $SNR = 13.02$ con un número de coeficientes $NoC = 12266$.

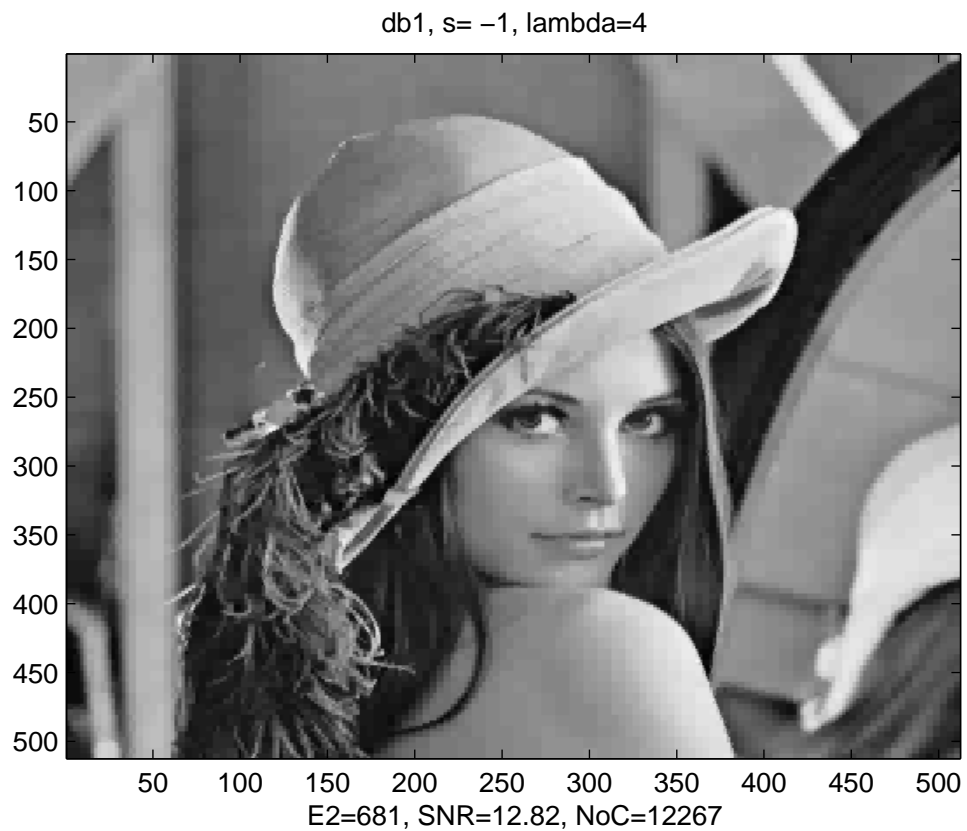


Figure 3.4: Aproximación de Lena a partir de la señal original haciendo *thresholding* restringido parámetros $s = 0$, $\lambda = 30$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 681$ y da un $SNR = 12.8$ con un número de coeficientes $NoC = 12267$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

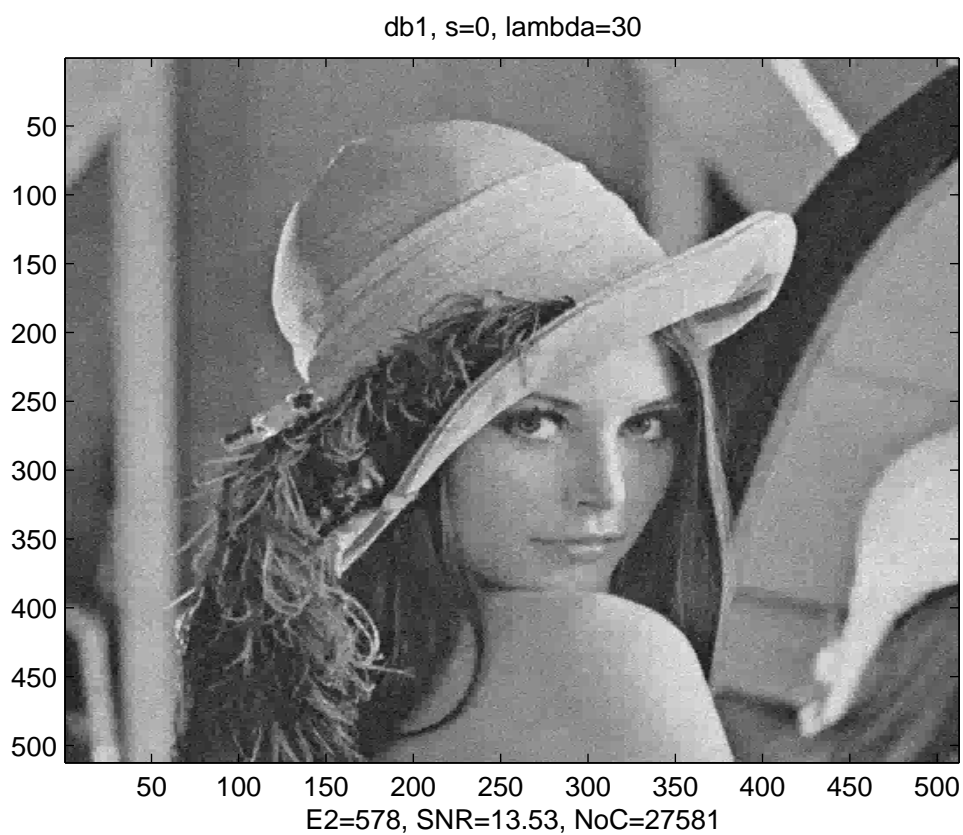


Figure 3.5: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* usual ($s=0$) con umbral $\lambda = 30$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 578$ que da un $SNR = 13.53$ con $NoC = 27581$.

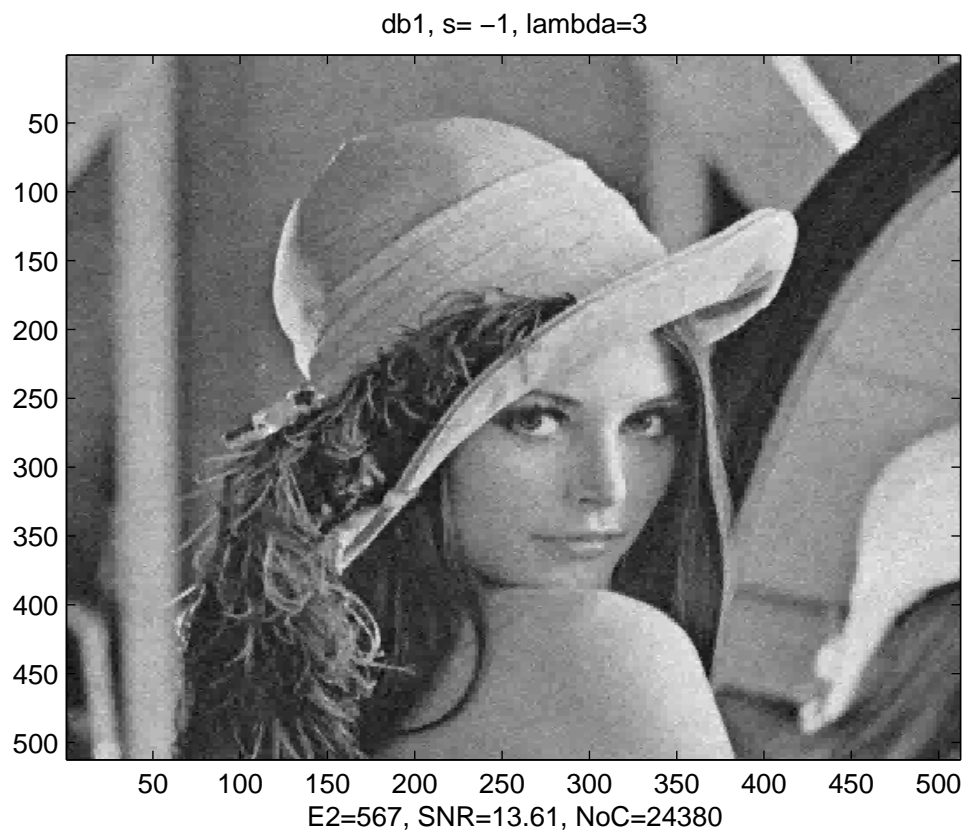


Figure 3.6: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* restringido ($s=-1$) con umbral $\lambda = 3$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 567$ que da un $SNR = 13.61$ con $NoC = 24380$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

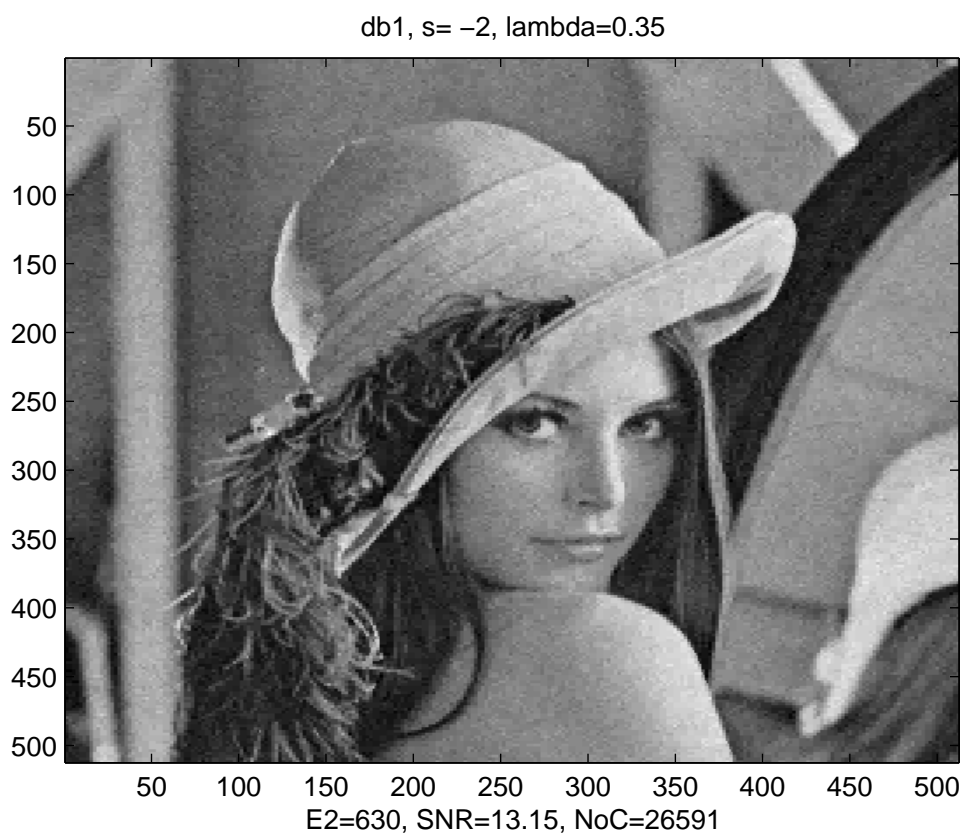


Figure 3.7: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* restringido ($s=-2$) con umbral $\lambda = 0.35$ sobre una descomposición de ondículas de Haar. El error en norma ℓ^2 es $E2 = 630$ que da un $SNR = 13.15$ con $NoC = 26591$.

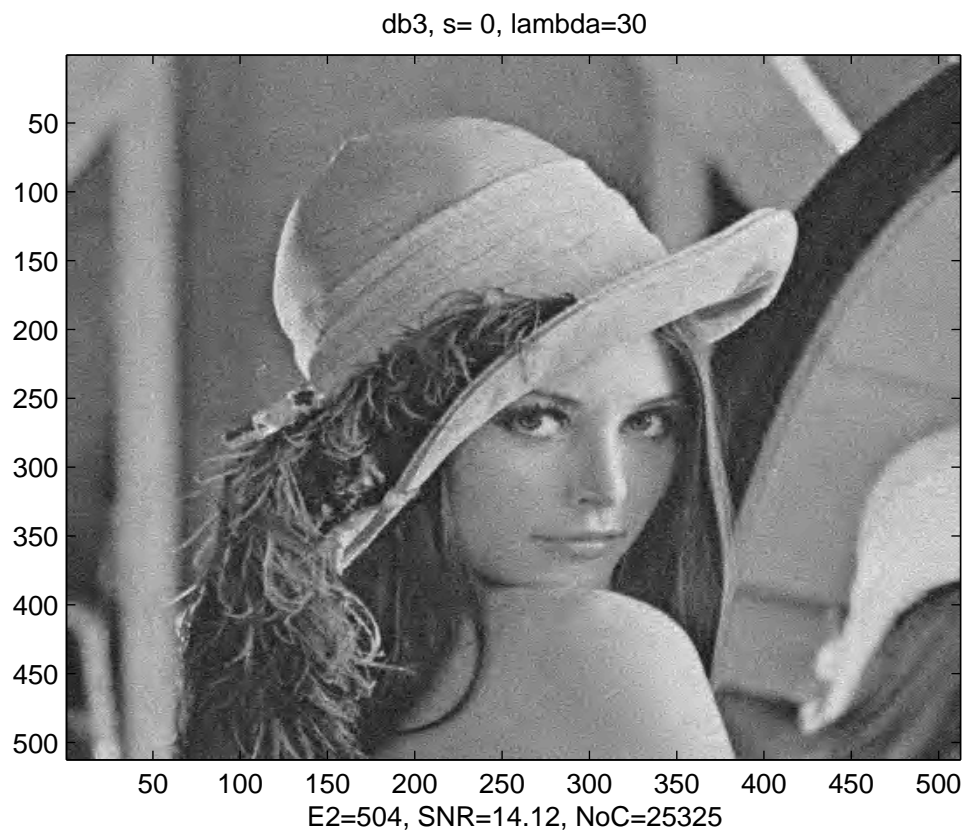


Figure 3.8: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* usual ($s = 0$) con $\lambda = 30$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 504$ que da un $SNR = 14.12$ con $NoC = 25325$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

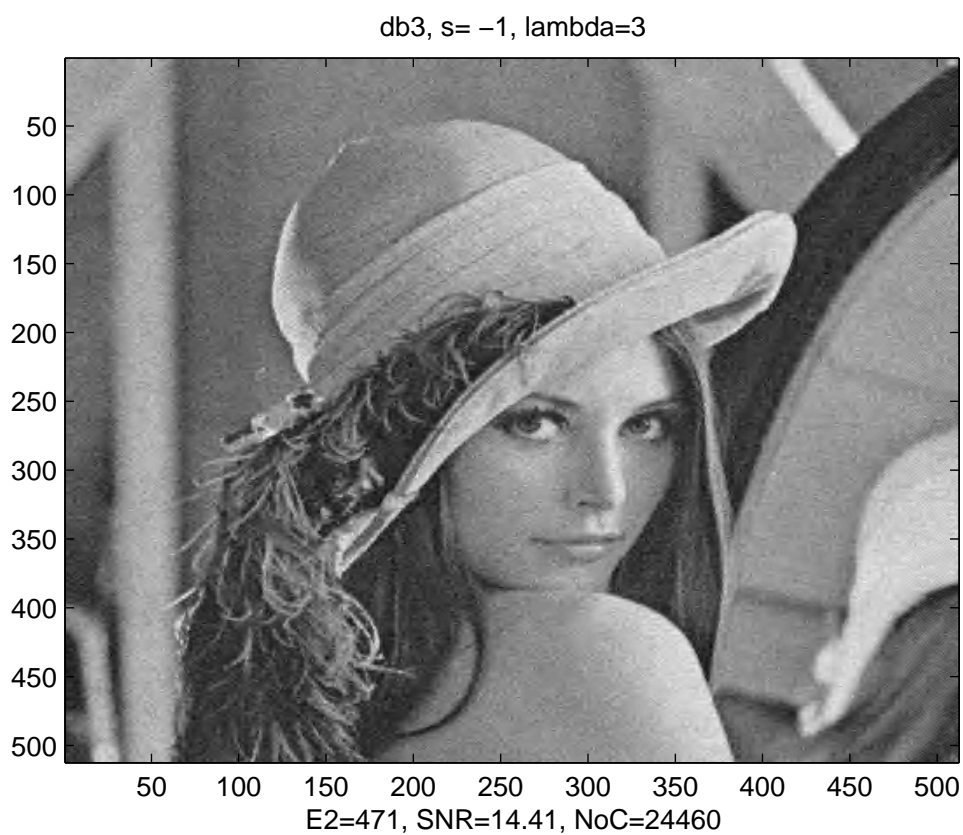


Figure 3.9: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* restringido ($s = -1$) con $\lambda = 3$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 471$ que da un $SNR = 14.41$ con $NoC = 24460$.

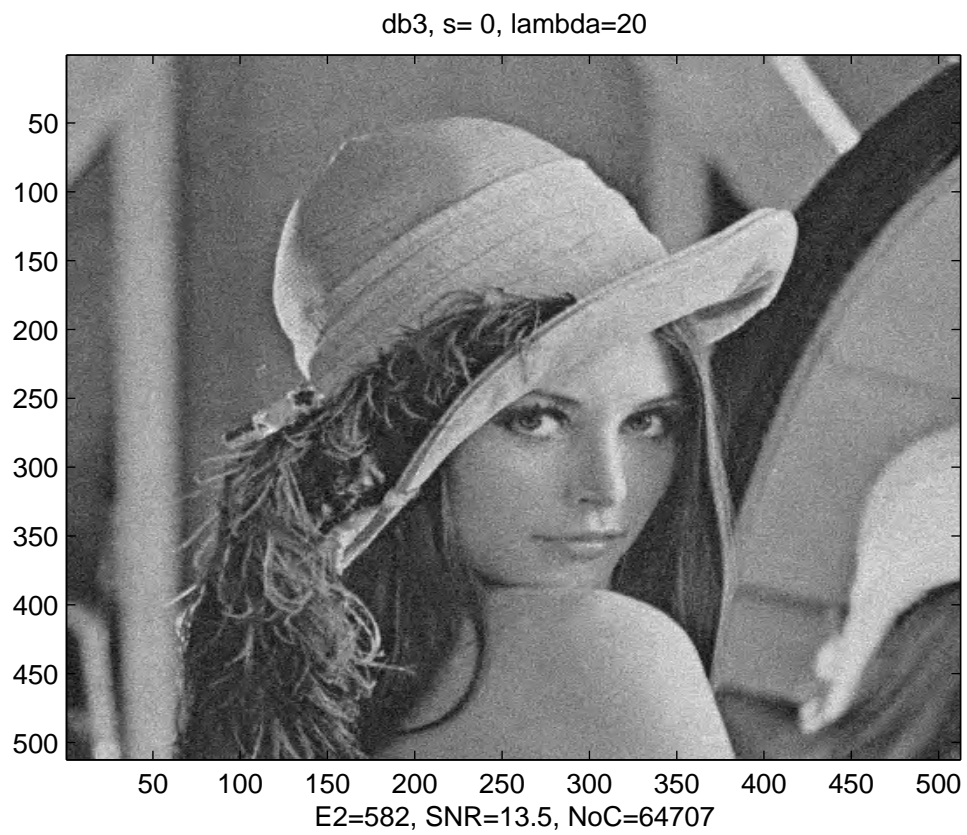


Figure 3.10: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* usual ($s = 0$) con $\lambda = 20$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 582$ que da un $SNR = 13.5$ con $NoC = 64707$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

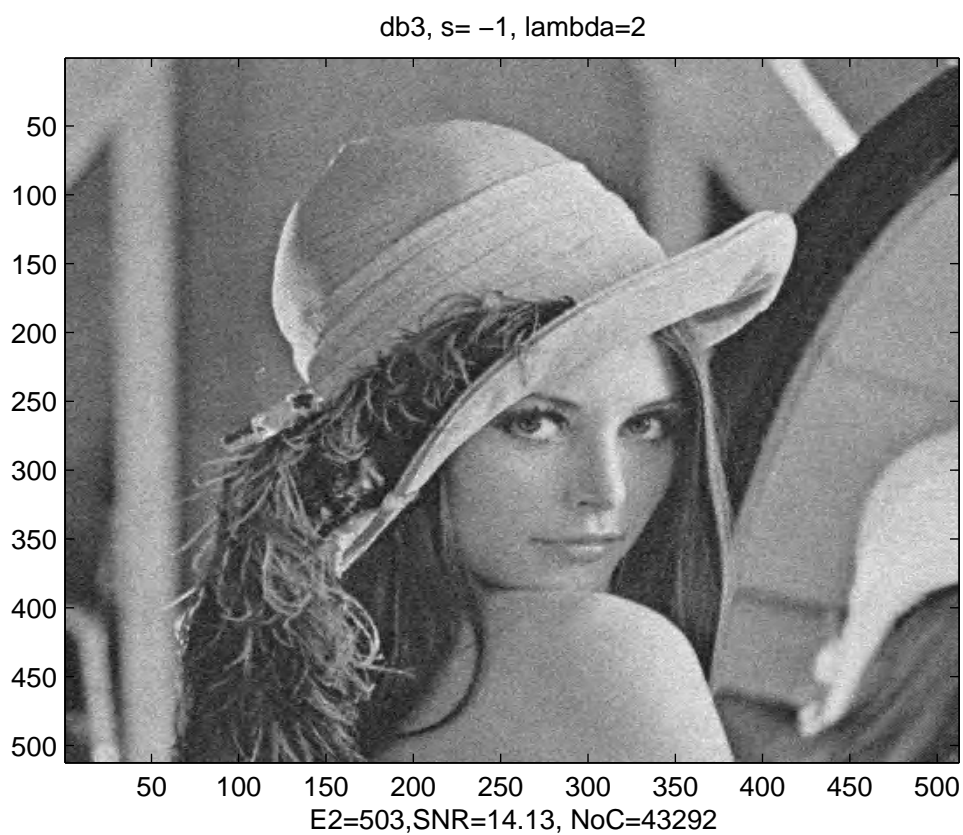


Figure 3.11: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* restringido ($s = -1$) con $\lambda = 2$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 503$ que da un $SNR = 14.13$ con $NoC = 43292$.

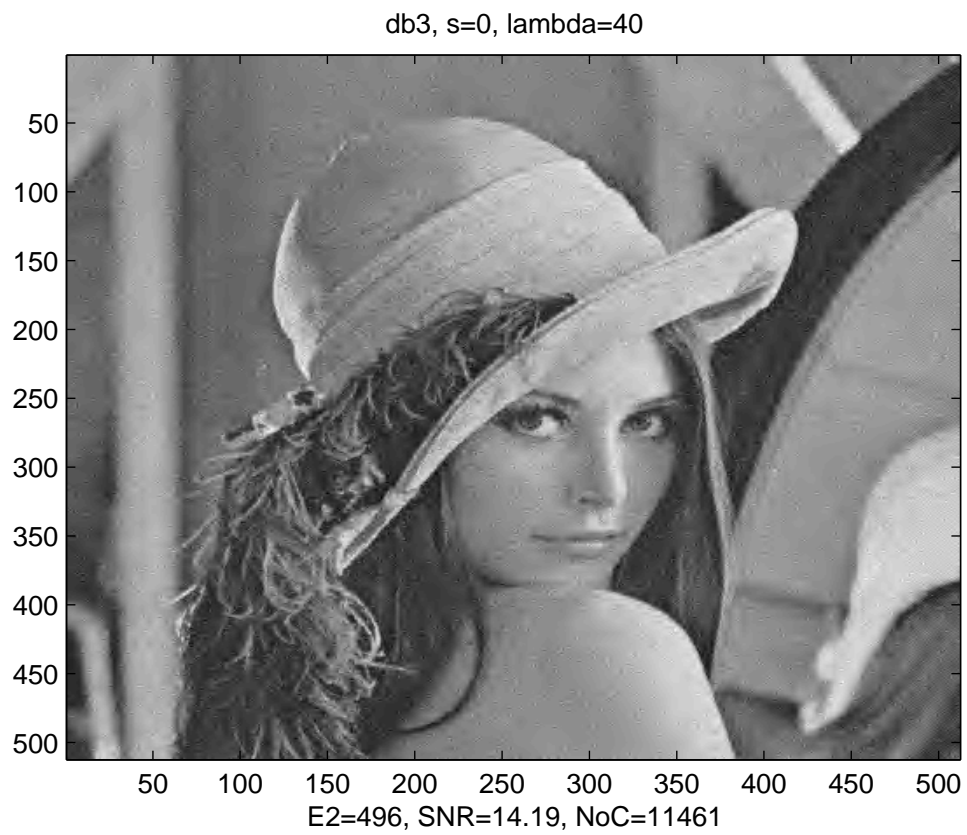


Figure 3.12: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* usual ($s = 0$) con $\lambda = 40$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 496$ que da un $SNR = 14.19$ con $NoC = 11461$.

3. APROXIMACIÓN NO LINEAL RESTRINGIDA EN ESPACIOS DE SUCESIONES

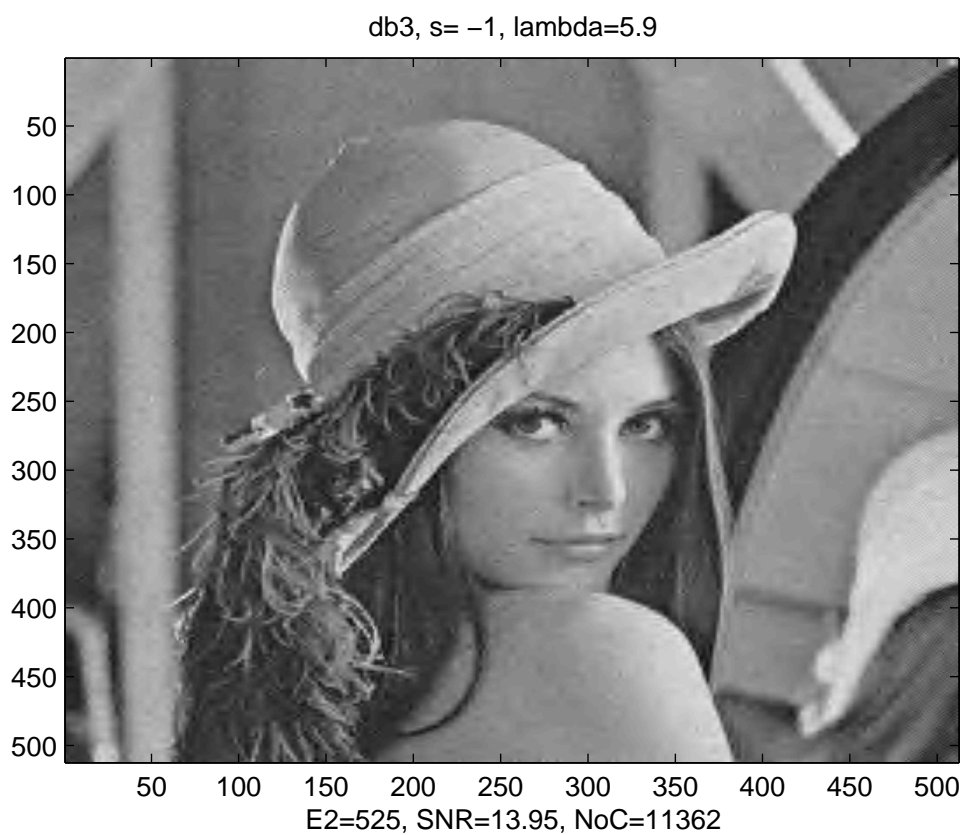


Figure 3.13: Estimación de Lena a partir de la señal con ruido haciendo *thresholding* restringido ($s = -1$) con $\lambda = 5.9$ sobre una descomposición de ondículas de Daubechies 3. El error en norma ℓ^2 es $E2 = 525$ que da un $SNR = 13.95$ con $NoC = 11362$.

4

Shearlets y nuevos espacios de distribuciones

*Así como del fondo de la música
brota una nota
que mientras vibra crece y se adelgaza
hasta que en otra música enmudece,
brota del fondo del silencio
otro silencio, aguda torre, espada,
y sube y crece y nos suspende...
... desembocamos al silencio
en donde los silencios enmudecen.*

Octavio Paz, *Silencio*.

4.1 Introducción y resultados.

Las ondículas multidimensionales tradicionales (separables) son construidas a partir de productos tensoriales de ondículas de una dimensión. Estas ondículas no permiten percibir la geometría de las discontinuidades en bajas dimensiones. Para tratar de solucionar este problema se han desarrollado recientemente las ondículas direccionales y bancos de filtros ([4], [5]), las *curvelets* ([16]), las *contourlets* ([35]) y las *shearlets*.

Las *shearlets* pueden considerarse un caso particular de las ondículas con dilatación compuesta desarrolladas por Guo, Lim, Labate, Weiss y Wilson en [55]. Producen marcos de Parseval para $L^2(\mathbb{R}^2)$ o subespacios de éste (dependiendo del muestreo discreto de los parámetros) por lo que son sistemas de reproducción. Se ha demostrado que las *shearlets* son sistemas que permiten una representación óptima de funciones que están

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

en $C^2(\mathbb{R}^2)$ excepto en discontinuidades a lo largo de curvas C^2 (*cartoon-like images*), ver [53], [52]. Algunas publicaciones pueden encontrarse en <http://www.shearlet.org>.

Comenzamos repasando la definición de ondícula ortonormal asociada a una matriz de dilatación A , es decir, una matriz A de tamaño $d \times d$ con coeficientes reales y cuyos autovalores tienen todos módulo mayor que 1. Una función $\psi \in L^2(\mathbb{R}^d)$ se dice que es una **ondícula ortonormal** si la colección

$$\{\psi_{\nu,k}(x) = |\det A|^{\nu/2} \psi(A^\nu x - k) : \nu \in \mathbb{Z}, k \in \mathbb{Z}^d\} \quad (4.1)$$

es una base ortonormal de $L^2(\mathbb{R}^d)$. Escribimos $\mathcal{Q} = \{Q_{\nu,k} = A^{-\nu}([0, 1]^d + k) : \nu \in \mathbb{Z}, k \in \mathbb{Z}^d\}$ para denotar el conjunto de todos los “cubos” (con dilatación asociada a A) de \mathbb{R}^d . Para ν fijo $\mathcal{Q}^\nu = \{Q_{\nu,k} : k \in \mathbb{Z}^d\}$ es una partición de \mathbb{R}^d por “cubos” de tamaño $|\det A|^{-\nu}$. Se tiene la siguiente fórmula de reproducción para toda $f \in L^2(\mathbb{R}^d)$:

$$f = \sum_{Q \in \mathcal{Q}} \langle f, \psi_Q \rangle \psi_Q. \quad (4.2)$$

Las *shearlets* que usaremos en este capítulo están definidas en \mathbb{R}^2 y se construyen en el lado de la transformada de Fourier en $\hat{\mathbb{R}}^2$. Partimos el plano $\hat{\mathbb{R}}^2$ en tres regiones: un **cono horizontal**, un **cono vertical** y un cuadrado centrado en el origen. El **cono horizontal** es

$$\mathcal{D}^h = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| \geq \frac{1}{8}, \left| \frac{\xi_2}{\xi_1} \right| \leq 1.\}$$

El **cono vertical** es

$$\mathcal{D}^v = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \geq \frac{1}{8}, \left| \frac{\xi_1}{\xi_2} \right| \leq 1.\}$$

El **cuadrado central** es

$$\mathcal{R} = \{(\xi_1, \xi_2) \in \hat{\mathbb{R}}^2 : |\xi_1|, |\xi_2| \leq \frac{1}{8}\}.$$

Sean $\psi_1, \psi_2 \in L^2(\mathbb{R})$ definidas de la siguiente manera: $\hat{\psi}_1, \hat{\psi}_2 \in C^\infty(\mathbb{R})$ con $\text{supp } \hat{\psi}_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ y $\text{supp } \hat{\psi}_2 \subset [-1, 1]$ tales que

$$\sum_{j \geq 0} \left| \hat{\psi}_1(2^{-2j}\omega) \right|^2 = 1, \quad \text{para } |\omega| \geq \frac{1}{8} \quad (4.3)$$

y

$$\left| \hat{\psi}_2(\omega - 1) \right|^2 + \left| \hat{\psi}_2(\omega) \right|^2 + \left| \hat{\psi}_2(\omega + 1) \right|^2 = 1, \quad \text{para } |\omega| \leq 1. \quad (4.4)$$

La construcción de funciones con estas propiedades puede verse en [55].

Para definir las *shearlets* en el cono horizontal \mathcal{D}^h consideramos

$$A_h = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad B_h = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

y definimos $\hat{\psi}^h(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\frac{\xi_2}{\xi_1})$. Puede demostrarse ([55, §5.2.1] y Sección 4.2.3 en este trabajo) que la colección

$$\{\psi_{j,\ell,k}^h(x) = 2^{\frac{3j}{2}} \psi^h(B_h^\ell A_h^j x - k) : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\} \quad (4.5)$$

es un marco de Parseval para $L^2((\mathcal{D}^h)^\vee) = \{f \in L^2(\mathbb{R}^2) : \text{sop } \hat{f} \subset \mathcal{D}^h\}$.

Puesto que

$$(\psi_{j,\ell,k}^h)^\wedge(\xi) = |\det A_h|^{-j/2} \hat{\psi}(\xi A_h^{-j} B_h^{-\ell}) e^{-2\pi i \xi A_h^{-j} B_h^{-\ell} k} \quad (4.6)$$

tenemos que

$$\text{sop } (\psi_{j,\ell,k}^h)^\wedge \subset \{\xi \in \hat{\mathbb{R}}^2 : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], \left| \frac{\xi_2}{\xi_1} - \ell 2^{-j} \right| \leq 2^{-j}\}. \quad (4.7)$$

Por tanto, el recubrimiento del cono horizontal \mathcal{D}^h en el plano de frecuencias $\hat{\mathbb{R}}^2$ es el que se muestra en la Figura 4.1 (se esquematizan sólo las escalas $j = 0, 1, 2$).

Las definiciones de las *shearlets* en el cono vertical se hacen de manera similar considerando las matrices

$$A_v = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, \quad B_v = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

y definiendo $\hat{\psi}^v(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\hat{\psi}_2(\frac{\xi_1}{\xi_2})$. Se tiene que la colección

$$\{\psi_{j,\ell,k}^v(x) = 2^{\frac{3j}{2}} \psi^v(B_v^\ell A_v^j x - k) : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\} \quad (4.8)$$

es un marco de Parseval para $L^2((\mathcal{D}^v)^\vee) = \{f \in L^2(\mathbb{R}^2) : \text{supp } \hat{f} \subset \mathcal{D}^v\}$. Si $\varphi \in L^2(\mathbb{R}^2)$ tal que $\varphi \in C_c^\infty(\mathbb{R}^2)$, $\text{supp } \hat{\varphi} \subset [-\frac{1}{4}, \frac{1}{4}]^2$, $|\hat{\varphi}| = 1$ para $\xi \in [-\frac{1}{8}, \frac{1}{8}]^2 = \mathcal{R}$, se cumple

$$\begin{aligned} P(\xi) &= |\hat{\varphi}(\xi)|^2 \chi_{\mathcal{R}}(\xi) + \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left| \hat{\psi}^h(\xi A_h^{-j} B_h^{-\ell}) \right|^2 \chi_{\mathcal{D}^h}(\xi) \\ &\quad + \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left| \hat{\psi}^v(\xi A_v^{-j} B_v^{-\ell}) \right|^2 \chi_{\mathcal{D}^v}(\xi) = 1, \quad \text{para todo } \xi \in \hat{\mathbb{R}}^2, \quad (4.9) \end{aligned}$$

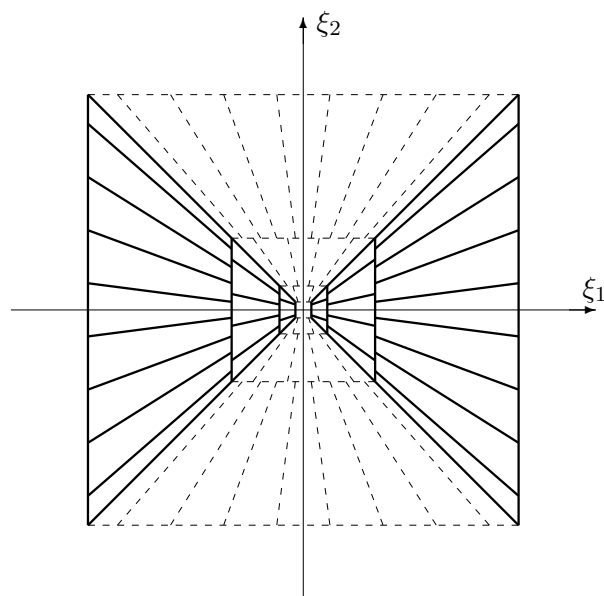


Figure 4.1: Esquema del recubrimiento del cono horizontal \mathcal{D}^h en el plano de frecuencias $\hat{\mathbb{R}}^2$

entonces se tiene un marco de Parseval para $L^2((\mathcal{D}^h)^\vee) \cup L^2((\mathcal{D}^v)^\vee) \cup L^2(\mathcal{R}^\vee)$, ver [71, Teorema 9].

La gran desventaja de esta construcción reside en que los operadores de las proyecciones ortogonales sobre \mathcal{D}^h , \mathcal{D}^v y \mathcal{R} tiene el efecto de destruir las propiedades de localización espacial de las “*shearlets* en la frontera”, es decir, aquellas que tienen el parámetro $\ell = \pm 2^j$, ver p. 3 en [54]. Para remediar esta desventaja, Guo y Labate en [54], modifican convenientemente las definiciones de las *shearlets*, ver Sección 4.5.1. En la Sección 4.5 demostramos que la fórmula de representación con *shearlets* suaves es válida para cualquier distribución $f \in \mathcal{S}'$ con convergencia en \mathcal{S}' (ver Teorema 4.5.3). La demostración de este hecho sigue la línea argumental de [13] y [39].

La Sección 4.3 está destinada a probar relaciones de casi-ortogonalidad entre elementos de las colecciones (4.5) y (4.8), las cuales serán usadas en las demostraciones de los resultados de la Sección 4.4, cuyo contenido describimos a continuación.

Los espacios de Triebel-Lizorkin son una familia de espacios que incluyen a los espacios L^p , $1 < p < \infty$, los espacios de Hardy H^p , $0 < p \leq 1$, y los espacios de Sobolev $W^{s,p}$, $1 < p < \infty$, $s \in \mathbb{N}$. El lector puede consultar los libros/artículos de Triebel ([100] y [101]), Frazier-Jawerth ([41]), Frazier-Jawerth-Weiss ([42]), y su contraparte con pesos en los artículos de Bui ([14] y [15]), para conocer sus propiedades. Seguimos [41] para dar su definición de los espacios de Triebel-Lizorkin inhomogéneos.

Sean $\varphi, \Phi \in \mathcal{S}(\mathbb{R}^n)$ tales que

1. $\text{sop } \hat{\varphi} \subset \{\xi \in \hat{\mathbb{R}}^d : \frac{1}{2} < |\xi| < 2\}$, $\text{sop } \hat{\Phi} \subset \{\xi \in \hat{\mathbb{R}}^d : |\xi| \leq 2\}$,
2. $|\hat{\varphi}(\xi)| \geq c > 0$ si $\frac{3}{5} \leq |\xi| \leq \frac{5}{3}$, $|\hat{\Phi}(\xi)| \geq c > 0$ si $|\xi| \leq \frac{5}{3}$.

Para $\alpha \in \mathbb{R}$, $0 < q \leq \infty$, $0 < p < \infty$, el espacio inhomogéneo de Triebel-Lizorkin $\mathbf{F}_p^{\alpha,q}$ es la colección de todas las $f \in \mathcal{S}'$ tales que

$$\|f\|_{\mathbf{F}_p^{\alpha,q}} = \|\Phi * f\|_{L^p} + \left\| \left(\sum_{\nu=1}^{\infty} (2^{\nu\alpha} |\varphi_\nu * f|)^q \right)^{1/q} \right\|_{L^p} < \infty. \quad (4.10)$$

donde $\varphi_\nu(x) = 2^{\nu d} \varphi(2^\nu x)$. Puesto que $\hat{\varphi}_\nu(\xi) = \hat{\varphi}(\xi 2^{-\nu})$ las propiedades 1) y 2) nos indican que la definición (4.10) está asociada a la descomposición de $\hat{\mathbb{R}}^d$ en anillos diádicos concéntricos.

Se podrían también definir los espacios de Triebel-Lizorkin $\mathbf{F}_p^{\alpha,q}(A, w)$ asociados a una matriz de dilatación A y a un peso $w \in A_\infty$ sustituyendo L^p por $L^p(w)$, $2^{\nu\alpha}$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

por $|\det A|^{\nu\alpha}$ y φ_ν por $\varphi_\nu(x) = |\det A|^j \varphi(A^j x)$, $j = 1, 2, 3, \dots$, para funciones Φ y φ adecuadas, tal como lo hacen Bownik-Ho en [13], en donde el lector puede encontrar un estudio detallado de sus propiedades.

En la Sección 4.4 definimos los espacios $\mathbf{F}_p^{\alpha,q}(AB)$ que llamamos espacios de **Triebel-Lizorkin AB -anisotrópicos inhomogéneos** (o también **Triebel-Lizorkin anisotrópicos inhomogéneos asociados a las shearlets**) usando la descomposición del plano de frecuencias dado en la Figura 4.1 al que se ha añadido la descomposición en frecuencias del cono vertical y del cuadrado central.

De forma precisa, sean ψ^h y ψ^v las funciones descritas con anterioridad a (4.5) y (4.8) y $\hat{\varphi} \in C_c^\infty(\hat{\mathbb{R}}^2)$ tal que $\text{supp } \hat{\varphi} \subset [-\frac{1}{4}, \frac{1}{4}]^2$ y $|\hat{\varphi}| = 1$ para $\xi \in [-\frac{1}{8}, \frac{1}{8}]^2$. Definimos $\mathbf{F}_p^{\alpha,q}(AB)$, $\alpha \in \mathbb{R}$, $0 < q \leq \infty$, $0 < p < \infty$, como la colección de todas las $f \in \mathcal{S}'$ tales que

$$\begin{aligned} \|f\|_{\mathbf{F}_p^{\alpha,q}(AB)} &= \|f * \varphi\|_{L^p} \\ &+ \left\| \left(\sum_{\mathfrak{d}=\{h,v\}} \left\{ \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} [|Q_j|^{-\alpha} |\tilde{\psi}_{A_\mathfrak{d}^{-j} B_\mathfrak{d}^{-\ell}} * f|^q] \right\} \right)^{1/q} \right\|_{L^p}, \end{aligned} \quad (4.11)$$

donde $\tilde{\psi}(x) = \overline{\psi(-x)}$, $\psi_M(x) = |\det A|^{-1} \psi(M^{-1}x)$ y $|Q_j|^{-\alpha} = |\det A|^{j\alpha} = 2^{3j\alpha}$.

Es conocido (ver [41]) que los espacios $\mathbf{F}_p^{\alpha,q}$ definidos en (4.10) pueden caracterizarse en términos de los coeficientes $\langle f, \Phi_Q \rangle$ y $\langle f, \varphi_Q \rangle$ donde $Q = Q_{\nu,k} = 2^{-\nu}([0, 1]^d + k)$ de forma que $f \in \mathbf{F}_p^{\alpha,q}$ si y sólo si

$$\left\| \left(\sum_{|Q|=1} (|\langle f, \Phi_Q \rangle| \tilde{\chi}_Q)^q \right)^{1/q} \right\|_{L^p} + \left\| \left(\sum_{|Q|<1} (|Q|^{-\frac{\alpha}{d}} |\langle f, \varphi_Q \rangle| \tilde{\chi}_Q)^q \right)^{1/q} \right\|_{L^p} < \infty \quad (4.12)$$

donde $\varphi_Q(x) = \varphi_{Q_{\nu,k}}(x) = 2^{\frac{\nu d}{2}} \varphi(2^\nu x - k)$, $\tilde{\chi}_Q(x) = |Q|^{-\frac{1}{2}} \chi_Q(x)$ y $|Q| = |Q_{\nu,k}| = 2^{-\nu d}$.

En la Sección 4.4 damos una caracterización de los espacios $\mathbf{F}_p^{\alpha,q}(AB)$ definidos en (4.11) usando los coeficientes de shearlets en los conos (ver Teorema 4.4.11). Para entender esta descripción sea $Q_0 = [0, 1]^2$ y

$$\begin{aligned} Q_{j,\ell,k}^h &= A_h^{-j} B_h^{-\ell}(Q_0 + k), \quad k \in \mathbb{Z}^2, -2^j \leq \ell \leq 2^j, j \geq 0 \\ Q_{j,\ell,k}^v &= A_v^{-j} B_v^{-\ell}(Q_0 + k), \quad k \in \mathbb{Z}^2, -2^j \leq \ell \leq 2^j, j \geq 0. \end{aligned} \quad (4.13)$$

Se tiene que $f \in \mathbf{F}_p^{\alpha,q}(AB)$ si y sólo si

$$\left\| \left(\sum_{|Q|=1} (|\langle f, \varphi_Q \rangle| \tilde{\chi}_Q)^q \right)^{1/q} \right\|_{L^p} + \left\| \left(\sum_{\mathfrak{d}=\{h,v\}} \sum_{Q \in \mathcal{Q}_{AB}} (|Q|^{-\alpha} |\langle f, \psi_{Q^{\mathfrak{d}}} \rangle| \tilde{\chi}_Q)^q \right)^{1/q} \right\|_{L^p} < \infty \quad (4.14)$$

donde \mathcal{Q}_{AB} denota el conjunto de todos los $Q_{j,\ell,k}^{\mathfrak{d}}$ definidos en (4.13) con $\mathfrak{d} = h$ y $\mathfrak{d} = v$ según consideremos el cono horizontal o el vertical. Recordar que $\psi_{Q^{\mathfrak{d}}}$ han sido definidos en (4.5) y (4.8) según sea $\mathfrak{d} = h$ o $\mathfrak{d} = v$. La demostración de este resultado sigue la línea argumental desarrollada en [41] y [13], con las adaptaciones que requiere la nueva situación.

En la Sección 4.6 damos algunas inclusiones continuas entre los espacios $\mathbf{F}_p^{\alpha,q}(AB)$ y los espacios $\mathbf{F}_p^{\alpha,q}$ (ver Teoremas 4.6.2 y 4.6.3). En concreto

$$F_p^{\alpha_1,q} \hookrightarrow F_p^{\alpha_2,q}(AB) \quad (4.15)$$

si $3\alpha_2 + \frac{1}{q} + \lambda > 2\alpha_1$ siempre que $\lambda > 2 \max(1, \frac{1}{q}, \frac{1}{p})$, y

$$F_p^{\alpha_2,q}(AB) \hookrightarrow F_p^{\alpha_1,q} \quad (4.16)$$

si $2\alpha_1 + 2 - \min(1, q) \leq 3\alpha_2$. Las regiones en las que se cumplen (4.15) y (4.16) se dibujan en la Figura 4.2.

Finalmente, mostramos que, dependiendo de los valores de los parámetros, hay sucesiones de funciones $\{f^{(j)}\}_{j \geq 0}$ tal que $\|f^{(j)}\|_{\mathbf{F}_{p_2}^{\alpha_2,q_2}(AB)} \approx 1$, pero $\lim_{j \rightarrow \infty} \|f^{(j)}\|_{\mathbf{F}_{p_1}^{\alpha_1,q_1}} = 0$, y sucesiones $\{g^{(\nu)}\}_{\nu \geq 0}$ tales que $\|g^{(\nu)}\|_{\mathbf{F}_{p_1}^{\alpha_1,q_1}} \approx 1$, pero $\lim_{\nu \rightarrow \infty} \|g^{(\nu)}\|_{\mathbf{F}_{p_2}^{\alpha_2,q_2}(AB)} = 0$ (ver Teoremas 4.6.4 y 4.6.5).

4.2 Shearlets

Las *shearlets* son una generalización de las ondículas que capturan mejor las propiedades geométricas de las funciones. Se basan en las llamadas ondículas con dilatación compuesta (ver [55]). Damos una introducción básica a la construcción de los diferentes tipos de *shearlets*: continuas, discretas y discretas en el cono, en las siguientes tres subsecciones. Un punto $x \in \mathbb{R}^d$ es un vector columna $x = (x_1, \dots, x_d)^t$ y un punto ξ en el dual $\hat{\mathbb{R}}^d$ es un vector fila $\xi = (\xi_1, \dots, \xi_d)$.

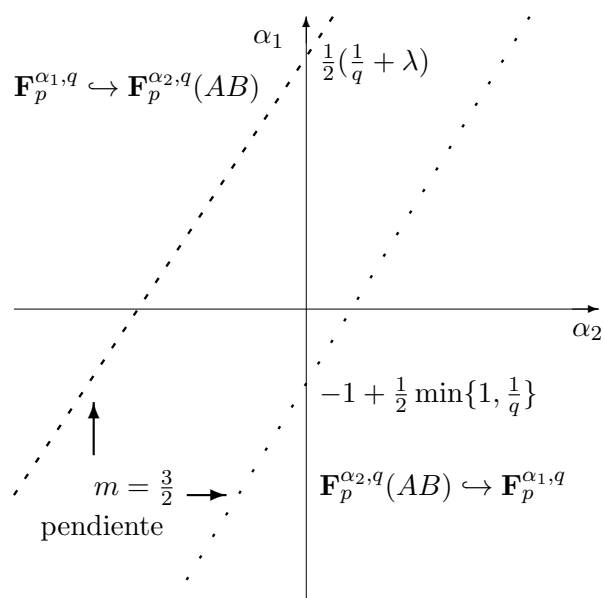


Figure 4.2: Relación de inclusión entre $\mathbf{F}_p^{\alpha_2, q}(AB)$ y $\mathbf{F}_p^{\alpha_1, q}$

4.2.1 Shearlets continuas

Un **sistema afín continuo** en $L^2(\mathbb{R}^d)$ es una colección de funciones de la forma

$$\{T_t D_M \psi : t \in \mathbb{R}^d, M \in G\},$$

donde $\psi \in L^2(\mathbb{R}^d)$, T_t es el operador de traslación $T_t f(x) = f(x-t)$, D_M es el operador de dilatación $D_M f(x) = |\det M|^{-1/2} f(M^{-1}x)$ (normalizado en $L^2(\mathbb{R}^d)$), y G es un subconjunto de $GL_d(\mathbb{R})$. Considérese $d = 2$ y G el grupo de dilatación de 2 parámetros (ver [55] para una definición aún más general)

$$G = \{M_{as} = \begin{pmatrix} a & \sqrt{as} \\ 0 & \sqrt{a} \end{pmatrix} : (a, s) \in \mathbb{R}_+ \times \mathbb{R}\}.$$

La matriz M_{as} es el producto $S_s A_a$ de la matriz de **transformación shear** que preserva el área $S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ y de la matriz de **dilatación anisotrópica** $A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}$. Así mismo, además, que ψ es dada por

$$\hat{\psi}(\xi) = \hat{\psi}_1(\xi_1) \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right), \quad (4.17)$$

para cualquier $\xi = (\xi_1, \xi_2) \in \hat{\mathbb{R}}^2$, $\xi_1 \neq 0$, y

$$\int_0^\infty \left| \hat{\psi}_1(a\omega) \right|^2 \frac{da}{a} = 1, \quad \text{for a.e. } \omega \in \mathbb{R},$$

y $\|\psi_2\|_{L^2(\mathbb{R})} = 1$. Entonces, se dice que una ψ es una *shearlet* continua si

$$\{\psi_{ast}(x) = a^{-3/4} \psi(M_{as}^{-1}(x-t)) : a \in \mathbb{R}_+, s \in \mathbb{R}, t \in \mathbb{R}^2\},$$

es un sistema de reproducción para $L^2(\mathbb{R}^2)$, esto es,

$$\|f\|_{L^2(\mathbb{R}^2)}^2 = \int_{\mathbb{R}^2} \int_{\mathbb{R}} \int_0^\infty |\langle f, \psi_{ast} \rangle|^2 \frac{da}{a^3} ds dt,$$

para toda $f \in L^2(\mathbb{R}^2)$ (ver [105]).

4.2.2 Shearlets discretas

Ya que $L^2(\mathbb{R}^2)$ es un espacio de Hilbert separable, sucede que, por medio de un “muestreo” de los parámetros de las *shearlets* continuas, existe una construcción de un sistema parecido a una base para $L^2(\mathbb{R}^2)$ con un sistema “discreto” de *shearlets*.

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Una familia numerable $\{e_j : j \in \mathcal{J}\}$ de elementos en un espacio de Hilbert separable \mathbb{H} se llama un **marco** si existen constantes $0 < A \leq B < \infty$, tales que $A \|f\|_{\mathbb{H}}^2 \leq \sum_{j \in \mathcal{J}} |\langle f, e_j \rangle|^2 \leq B \|f\|_{\mathbb{H}}^2$, para todas las funciones $f \in \mathbb{H}$. Un marco se llama **ajustado** si $A = B$, y se llama un **marco de Parseval** si $A = B = 1$. Así, si $\{e_j : j \in \mathcal{J}\}$ es un marco de Parseval para \mathbb{H} , entonces $\|f\|_{\mathbb{H}}^2 = \sum_{j \in \mathcal{J}} |\langle f, e_j \rangle|^2$, para toda $f \in \mathbb{H}$, lo cual es equivalente a la fórmula de reproducción $f = \sum_{j \in \mathcal{J}} \langle f, e_j \rangle e_j$, con convergencia en \mathbb{H} .

Con un muestreo diádico del parámetro continuo, establecemos $a_j = 2^j, j \in \mathbb{Z}$, como el parámetro de escala discreto. Para cubrir el plano de frecuencias un mayor número de “direcciones” se necesitan al crecer el parámetro de escala. Así, el parámetro *shearing* se establece como $s_{j,k} = k\sqrt{a_j} = k2^{j/2}, k \in \mathbb{Z}$. Finalmente, el enrejado ajustado a las dilataciones anisotrópicas es $t_{jkm} = S_{s_{jk}} A_{a_j} m, m \in \mathbb{Z}^2$. Observando que $T_{S_{k2^{j/2}} A_{2^j} m} D_{S_{k2^{j/2}} A_{2^j} m} = D_{A_{2^j} S_k} T_m$, se obtiene el sistema discreto

$$\{\psi_{j,k,m} = D_{A_{2^j} S_k} T_m \psi : j, k \in \mathbb{Z}, m \in \mathbb{Z}^2\}. \quad (4.18)$$

Sea $\psi \in L^2(\mathbb{R}^2)$ como en (4.17) donde $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ satisface $\text{supp } \hat{\psi}_1 \subset [-1/2, -1/4] \cup [1/4, 1/2]$ y la fórmula discreta de Calderón

$$\sum_{j \in \mathbb{Z}} \left| \hat{\psi}_1(2^j \omega) \right|^2 = 1, \quad \text{para } \omega \in \mathbb{R},$$

y donde $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ satisface $\text{supp } \hat{\psi}_2 \subset [-1, 1]$ y

$$\sum_{k \in \mathbb{Z}} \left| \hat{\psi}_2(k + \omega) \right|^2 = 1, \quad \text{para } \omega \in \mathbb{R}.$$

Entonces, el sistema de *shearlets* discretas (4.18) es un marco de Parseval para $L^2(\mathbb{R}^2)$ (ver [55]). Así, el plano $\hat{\mathbb{R}}^2$ se cubre por bandas diádicas verticales (asociadas con $\hat{\psi}_1$) y cada una de éstas cubierta por $\left| \hat{\psi}_2(k + \omega) \right|^2$.

4.2.3 Shearlets discretas en el cono

A pesar de que el sistema discreto de *shearlets* tiene la propiedad de ser un marco de Parseval para $L^2(\mathbb{R}^2)$, el cubrimiento de $\hat{\mathbb{R}}^2$ no es homogénea. Con pequeñas modificaciones en el sistema discreto de *shearlets* arriba mencionado, uno puede obtener un marco de Parseval para funciones en $L^2(\mathbb{R}^2)$ cuya transformada de Fourier tiene soporte en el **cono horizontal**

$$\mathcal{D}^h = \{(\xi_1, \xi_2) \in \hat{\mathbb{R}}^2 : |\xi_1| \geq \frac{1}{8}, \left| \frac{\xi_2}{\xi_1} \right| \leq 1\}. \quad (4.19)$$

Considérese $\hat{\psi}_1, \hat{\psi}_2 \in C^\infty(\mathbb{R})$ con $\text{supp } \hat{\psi}_1 \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$ y $\text{supp } \hat{\psi}_2 \subset [-1, 1]$ tales que

$$\sum_{j \geq 0} \left| \hat{\psi}_1(2^{-2j}\omega) \right|^2 = 1, \quad \text{para } |\omega| \geq \frac{1}{8} \quad (4.20)$$

y

$$\left| \hat{\psi}_2(\omega - 1) \right|^2 + \left| \hat{\psi}_2(\omega) \right|^2 + \left| \hat{\psi}_2(\omega + 1) \right|^2 = 1, \quad \text{para } |\omega| \leq 1. \quad (4.21)$$

Se sigue de (4.21) que, para $j \geq 0$,

$$\sum_{\ell=-2^j}^{2^j} \left| \hat{\psi}_2(\xi - \ell) \right|^2 = 1, \quad \text{para } |\xi| \leq 2^j,$$

y haciendo el cambio de variable $\xi = 2^j\omega$,

$$\sum_{\ell=-2^j}^{2^j} \left| \hat{\psi}_2(2^j\omega - \ell) \right|^2 = 1, \quad \text{para } |\omega| \leq 1. \quad (4.22)$$

Sean

$$A_h = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad B_h = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

y $\hat{\psi}^h(\xi) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\frac{\xi_2}{\xi_1})$. De (4.20) y (4.22) se sigue que

$$\begin{aligned} \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} \left| \hat{\psi}^h(\xi A_h^{-j} B_h^{-\ell}) \right|^2 &= \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} \left| \hat{\psi}_1(2^{-2j}\xi_1) \right|^2 \left| \hat{\psi}_2(2^j \frac{\xi_2}{\xi_1} - \ell) \right|^2 \\ &= \sum_{j \geq 0} \left| \hat{\psi}_1(2^{-2j}\xi_1) \right|^2 \sum_{\ell=-2^j}^{2^j} \left| \hat{\psi}_2(2^j \frac{\xi_2}{\xi_1} - \ell) \right|^2 = 1, \end{aligned} \quad (4.23)$$

para $\xi = (\xi_1, \xi_2) \in \mathcal{D}^h \subset \hat{\mathbb{R}}^2$ y al cual llamaremos la **condición de marco de Parseval** (para el cono horizontal). Ya que $\text{supp } \hat{\psi}^h \subset [-\frac{1}{2}, \frac{1}{2}]^2$, (4.5) implica que el sistema de *shearlets*

$$\{\psi_{j,\ell,k}^h(x) = 2^{3j/2}\psi^h(B_h^\ell A_h^j x - k) : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\}, \quad (4.24)$$

es un marco de Parseval para $L^2((\mathcal{D}^h)^\vee) = \{f \in L^2(\mathbb{R}^2) : \text{supp } \hat{f} \subset \mathcal{D}^h\}$ (ver [55], Sección 5.2.1). Esto significa que

$$\sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} \sum_{k \in \mathbb{Z}^2} \left| \langle f, \psi_{j,\ell,k}^h \rangle \right|^2 = \|f\|_{L^2(\mathbb{R}^2)}^2,$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

para toda $f \in L^2(\mathbb{R}^2)$ tal que $\text{supp } \hat{f} \subset \mathcal{D}^h$. Esto se ve fácilmente aplicando cambios de variable, el teorema de Plancherel, la identidad de Parseval y (4.5)

$$\begin{aligned}
\sum_{j,k,\ell} \left| \langle f, \psi_{j,\ell,k}^h \rangle \right|^2 &= \sum_{j,k,\ell} \left| \int_{\hat{\mathbb{R}}^2} \hat{f}(\xi) \overline{(\psi_{j,\ell,k}^h)^\wedge(\xi)} d\xi \right|^2 \\
&= \sum_{j,k,\ell} \left| \int_{\hat{\mathbb{R}}^2} \hat{f}(\xi) |\det A_h|^{-j/2} \overline{\hat{\psi}^h(\xi A_h^{-j} B_h^{-\ell})} e^{2\pi i \xi A_h^{-j} B_h^{-\ell} k} d\xi \right|^2 \\
&= \sum_{j,k,\ell} \left| \int_Q \hat{f}(\omega B_h^\ell A_h^j) |\det A_h|^{j/2} \overline{\hat{\psi}^h(\omega)} e^{2\pi i \omega k} d\omega \right|^2 \\
&= \sum_{j,\ell} \int_Q \left| \hat{f}(\omega B_h^\ell A_h^j) \right|^2 |\det A_h|^j \left| \hat{\psi}^h(\omega) \right|^2 d\omega \\
&= \sum_{j,\ell} \int_{\hat{\mathbb{R}}^2} \left| \hat{f}(\xi) \right|^2 \left| \hat{\psi}^h(\xi A_h^{-j} B_h^{-\ell}) \right|^2 d\xi \\
&= \left\| \hat{f} \right\|_{L^2(\mathcal{D}^h)}^2 = \|f\|_{L^2((\mathcal{D}^h)^\vee)}^2.
\end{aligned}$$

Existen varios ejemplos de funciones ψ_1, ψ_2 que satisfacen las propiedades arriba descritas (ver [53]). Ya que $\hat{\psi}^h \in C_c^\infty(\hat{\mathbb{R}}^2)$, existen C_N tales que $|\psi^h(x)| \leq C_N(1+|x|)^{-N}$ para todo $N \in \mathbb{N}$. La transformada de Fourier de (4.24) es

$$(\psi_{j,\ell,k}^h)^\wedge(\xi) = |\det A_h|^{-j/2} \hat{\psi}(\xi A_h^{-j} B_h^{-\ell}) e^{-2\pi i \xi A_h^{-j} B_h^{-\ell} k}.$$

Así, las propiedades del soporte de las *shearlets* horizontales son más evidentes al observar que

$$\text{supp } (\psi_{j,\ell,k}^h)^\wedge \subset \left\{ \xi \in \hat{\mathbb{R}}^2 : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], \left| \frac{\xi_2}{\xi_1} - \ell 2^{-j} \right| \leq 2^{-j} \right\},$$

y que pueden verse en la Figura 4.1.

Se puede construir un marco de Parseval para el cono vertical

$$\mathcal{D}^v = \left\{ (\xi_1, \xi_2) \in \hat{\mathbb{R}}^2 : |\xi_2| \geq \frac{1}{8}, \left| \frac{\xi_1}{\xi_2} \right| \leq 1 \right\},$$

definiendo $\hat{\psi}^v(\xi) = \hat{\psi}_1(\xi_2) \hat{\psi}_2(\frac{\xi_1}{\xi_2})$ y con matrices anisotrópica y *shear*

$$A_v = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, \quad B_v = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Si $\varphi \in L^2(\mathbb{R}^2)$ tal que $\varphi \in C_c^\infty(\mathbb{R}^2)$, $\text{supp } \hat{\varphi} \subset [-\frac{1}{4}, \frac{1}{4}]^2$, $|\hat{\varphi}| = 1$ para $\xi \in [-\frac{1}{8}, \frac{1}{8}]^2 = \mathcal{R}$, se cumple

$$\begin{aligned} P(\xi) &= |\hat{\varphi}(\xi)|^2 \chi_{\mathcal{R}}(\xi) + \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left| \hat{\psi}^h(\xi A_h^{-j} B_h^{-\ell}) \right|^2 \chi_{\mathcal{D}^h}(\xi) \\ &\quad + \sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left| \hat{\psi}^v(\xi A_v^{-j} B_v^{-\ell}) \right|^2 \chi_{\mathcal{D}^v}(\xi) = 1, \quad \text{para todo } \xi \in \hat{\mathbb{R}}^2, \end{aligned} \quad (4.25)$$

entonces se tiene un marco de Parseval para $L^2((\mathcal{D}^h)^\vee) \cup L^2((\mathcal{D}^v)^\vee) \cup L^2(\mathcal{R}^\vee)$, ver [71, Teorema 9].

4.3 Notación y casi ortogonalidad

Ya que todos los resultados en el cono horizontal \mathcal{D}^h pueden ser enunciados para el cono vertical \mathcal{D}^v , con las modificaciones obvias, dejamos el super índice h y hacemos el desarrollo para el cono horizontal. Consideremos

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

como las matrices anisotrópica y *shear*, respectivamente. Para seguir los argumentos de [41] se escogen $\varphi, \psi \in \mathcal{S}$ como en la Sección 4.1 y en particular satisfacen las ecuaciones (4.20), (4.21) y (4.22). Siguiendo la notación para la dilatación isotrópica usual, $\varphi_t(x) := \frac{1}{t} \varphi(\frac{x}{t})$, denotamos para una matriz $M \in GL_2(\mathbb{R})$ la dilatación anisotrópica $\varphi_M(x) = |\det M|^{-1} \varphi(M^{-1}x)$ (no confundir con el operador de dilatación normalizado en L^2 como en la Subsección 4.2.1). También escribimos $\tilde{\varphi}(x) = \overline{\varphi(-x)}$. Con $Q_0 = [0, 1)^2$, definimos el paralelogramo

$$Q_{j,\ell,k} = A^{-j} B^{-\ell} (Q_0 + k); \quad (4.26)$$

por lo tanto, $\int \chi_{Q_{j,\ell,k}} = |Q_{j,\ell,k}| = |Q_{j,\ell}| = |Q_j| = 2^{-3j} = |\det A|^{-j}$. Escribimos $\tilde{\chi}_Q(x) = |Q|^{-1/2} \chi_Q(x)$. Si $\mathcal{Q}_{AB} := \{Q_{j,\ell,k} : j \geq 0, \ell = -2^j, \dots, 2^j, k \in \mathbb{Z}^2\}$ y $\mathcal{Q}^{j,\ell} := \{Q_{j,\ell,k} : k \in \mathbb{Z}^2\}$, entonces $\mathcal{Q}^{j,\ell}$ es una partición de \mathbb{R}^2 como se muestra en la Figura 4.3. Para acortar la notación y aclarar la exposición, identificaremos los multi índices (j, ℓ, k) y (i, m, n) con P y Q , respectivamente. De esta forma escribimos $\psi_P = \psi_{j,\ell,k}$ o $\psi_Q = \psi_{i,m,n}$. También, sean x_P y x_Q las esquinas izquierdas inferiores $A^{-j} B^{-\ell} k$ y

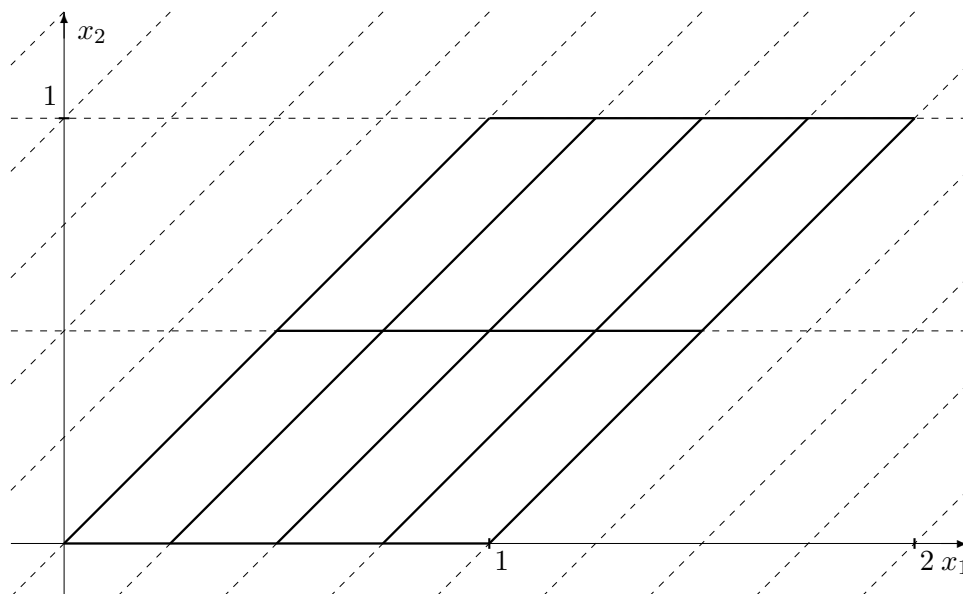


Figure 4.3: Esquema del recubrimiento del plano \mathbb{R}^2 con paralelogramos en $\mathcal{Q}^{1,2}$. Los paralelogramos con líneas sólidas cubren el paralelogramo $Q_{j,\ell,k} = Q_{0,1,0}$.

$A^{-i}B^{-m}n$ de los “cubos” $P = Q_{j,\ell,k}$ y $Q = Q_{i,m,n}$. Sea $B_r(x)$ la bola euclídea centrada en x con radio r .

Los elementos de la colección afín

$$\mathcal{A}_{AB} := \{\psi_{j,\ell,k}(x) = |\det A|^{j/2} \psi(B^\ell A^j x - k) : j \geq 0, -2^j \leq \ell \leq 2^j, k \in \mathbb{Z}^2\},$$

tienen transformada de Fourier

$$(\psi_{j,\ell,k})^\wedge(\xi) = |\det A|^{-j/2} \hat{\psi}(\xi A^{-j} B^{-\ell}) \mathbf{e}^{-2\pi i \xi A^{-j} B^{-\ell} k}.$$

Usando la dilatación anisotrópica es fácil verificar que

$$\psi_{A^{-j}B^{-\ell}}(x - A^{-j}B^{-\ell}k) = |\det A|^{j/2} \psi_{j,\ell,k}(x) = |P|^{-1/2} \psi_P(x)$$

y así

$$\left(\psi_{A^{-j}B^{-\ell}}(\cdot - A^{-j}B^{-\ell}k)\right)^\wedge(\xi) = \hat{\psi}(\xi A^{-j} B^{-\ell}) \mathbf{e}^{-2\pi i \xi A^{-j} B^{-\ell} k}.$$

También se tiene

$$\begin{aligned}
 |\langle f, \psi_P \rangle| &= |\langle f, \psi_{j,\ell,k} \rangle| \\
 &= \left| \int_{\mathbb{R}^2} f(x) |\det A|^{-j/2} \psi_{A^{-j}B^{-\ell}}(x - A^{-j}B^{-\ell}k) dx \right| \\
 &= |P|^{1/2} \left| (f * \tilde{\psi}_{A^{-j}B^{-\ell}})(x_P) \right|. \tag{4.27}
 \end{aligned}$$

4.3.1 Casi ortogonalidad

De la condición del soporte en $\hat{\psi}_1$, la definición de la matriz A y (4.20), el conjunto de todas las *shearlets* en la escala j (para todos los parámetros *shear* y de traslación) interactúa con el conjunto de todas las *shearlets* solamente en las escalas $j - 1$, j y $j + 1$ (para todos los parámetros *shear* y de traslación). El siguiente resultado es para funciones en \mathcal{S} no necesariamente *shearlets*.

Lema 4.3.1 Sean $g, h \in \mathcal{S}$. Para $i = j - 1, j, j + 1 \geq 0$, identifíquese Q con (i, m, n) . Entonces, para todo $N > 2$, existe $C_N > 0$ tal que

$$|g_{A^{-j}B^{-\ell}} * h_Q(x)| \leq \frac{C_N |Q|^{-\frac{1}{2}}}{(1 + 2^i |x - x_Q|)^N}, \quad |\ell| \leq 2^j$$

para todo $x \in \mathbb{R}^2$.

Para probar el Lema 4.3.1 necesitamos dos resultados previos. El primero relaciona las distancias euclídeas bajo los operadores de dilatación 2^{j-1} y $B^\ell A^j$. El segundo es un resultado básico de “casi ortogonalidad” del cual derivamos el Lema 4.3.1.

Lema 4.3.2 Sean A y B como en la Sección 4.3. Entonces ,

$$2^{j-2} |x| \leq \left| B^\ell A^j x \right|,$$

para todo $j \geq 0$, $\ell = -2^j, \dots, 2^j$ y todo $x \in \mathbb{R}^2$.

Demostración. Es suficiente con probar el lema para $|x| = 1$. Considérese $|x_1| \geq \sqrt{3/4}$. Entonces, como $|\ell| \leq 2^j$,

$$\begin{aligned}
 \left| B^\ell A^j x \right| &\geq 2^j |2^j x_1 + \ell x_2| \geq 2^j |2^j |x_1| - |\ell| |x_2| | \geq 2^{2j} \left| |x_1| - (1 - |x_1|^2)^{1/2} \right| \\
 &\geq 2^{2j} \left| \frac{\sqrt{3}}{2} - \frac{1}{2} \right| \geq 2^{2j} \left| \frac{3}{4} - \frac{2}{4} \right| = 2^{2(j-1)} \geq 2^{(j-2)}.
 \end{aligned}$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Para $|x_1| \leq \sqrt{3/4}$, $|x_2| \geq 1/2$ y $|B^\ell A^j x| \geq 2^j |x_2| \geq 2^{j-1}$. ■

Obsérvese que este resultado no puede aplicarse al caso de las “*shearlets* discretas”, ya que el parámetro *shear* en este caso recorre todo \mathbb{Z} haciendo los elipsoides $B^\ell A^j \partial\mathbb{U}$ cada vez más delgados cuando $\ell \rightarrow \pm\infty$ y, consecuentemente, interseca al círculo unitario $\partial\mathbb{U}$.

El siguiente resultado es fundamental para probar nuestro primer resultado de “casi ortogonalidad” en forma de una convolución en el dominio del espacio (ver (4.27)).

Lema 4.3.3 Sean $g, h \in \mathcal{S}$. Entonces, para cada $N > 2$, $i = j - 1, j, j + 1 \geq 0$, $|m| \leq 2^j$ y $|\ell| \leq 2^j$ existe $C_N > 0$ tal que

$$|g_{j,\ell,k} * h_{i,m,n}(x)| \leq \frac{C_N}{(1 + 2^i |x - A^{-i} B^{-m} n - A^{-j} B^{-\ell} k|)^N},$$

para todo $x \in \mathbb{R}^2$.

Demostración. Como $g, h \in \mathcal{S}$ entonces $|g(x)|, |h(x)| \leq \frac{C_N}{(1+|x|)^N}$ para todo $N \in \mathbb{N}$. Entonces,

$$|g_{j,\ell,k} * h_{i,m,n}(x)| \leq \int_{\mathbb{R}^2} \frac{C_N |\det A|^{(j+i)/2}}{(1 + |B^\ell A^j y|)^N} \cdot \frac{C_N}{(1 + |B^m A^i(x' - y)|)^N} dy$$

donde $x' = x - A^{-j} B^{-\ell} k - A^{-i} B^{-m} n$. observe que, ya que $i = j - 1, j, j + 1$, $|\det A|^{(j+i)/2} \simeq |\det A|^j \simeq |\det A|^i$. Siguiendo [59, §6], defínanse

$$\begin{aligned} E_1 &= \{y \in \mathbb{R}^d : |B^m A^i(x' - y)| \leq 3\} \\ E_2 &= \{y \in \mathbb{R}^d : |B^m A^i(x' - y)| > 3, |y| \leq |x'|/2\} \\ E_3 &= \{y \in \mathbb{R}^d : |B^m A^i(x' - y)| > 3, |y| > |x'|/2\}. \end{aligned}$$

El Lema 4.3.2 permite las siguientes tres cotas. Para $y \in E_1$ se tiene $1 + 2^i |x'| \leq 1 + 2 |B^m A^i(x' - y)| + 2^{i+1} |y| \leq 7 + 2^{j+2} |y| \leq 8(1 + |B^\ell A^j y|)$. Si $y \in E_3$, $1 + 2^i |x'| \leq 1 + 2^{i+1} |y| \leq 8(1 + |B^\ell A^j y|)$. Cuando $y \in E_2$, $2^{i-1} |x'| \leq 2^i |x'| - 2^i |y| < 2^i |x' - y|$, lo que implica $4 |B^m A^i(x' - y)| = |B^m A^i(x' - y)| + 3 |B^m A^i(x' - y)| \geq 3 + \frac{3}{2} 2^i |x' - y| \geq 1 + 2^i |x' - y| \geq 1 + 2^{i-1} |x'| \geq (1 + 2^i |x'|)/2$. Así, como $|j - i| \leq 1$,

$$\begin{aligned} |g_{j,\ell,k} * h_{i,m,n}(x)| &\lesssim \frac{C_N |\det A|^i}{(1 + 2^i |x'|)^N} \int_{E_1 \cup E_3} \frac{C_N}{(1 + |B^m A^i(x' - y)|)^N} dy \\ &\quad + \frac{C_N |\det A|^i}{(1 + 2^i |x'|)^N} \int_{E_2} \frac{C_N}{(1 + |B^\ell A^j y|)^N} dy \\ &\lesssim \frac{C_N}{(1 + 2^i |x'|)^N} \end{aligned}$$

para algún $C_N > 0$ y todo $N > 2$, haciendo un cambio de variables para acotar las integrales con una constante independiente de i, j, l y m . El resultado se sigue reemplazando de vuelta $x' = x - A^{-j}B^{-\ell}k - A^{-i}B^{-m}n$ en la estimación anterior. ■

Como corolario al Lema 4.3.3 se tiene la primera propiedad de casi ortogonalidad para las operaciones anisotrópicas y shear para funciones en \mathbb{S} .

Demostración del Lema 4.3.1. Identifíquese $(j, \ell, 0)$ con P y (i, m, n) con Q . Escribáse $|g_{A^{-j}B^{-\ell}} * h_{i,m,n}(x)| = \left| |P|^{-1/2} g_P * h_Q \right|$. Como $|i - j| \leq 1$, $|P|^{-1/2} \sim |Q|^{-1/2}$. Entonces, el Lema 4.3.3 da

$$\left| |P|^{-1/2} g_P * h_Q \right| \leq \frac{C_N |P|^{-1/2}}{(1 + 2^j |x - x_Q|)^N} \lesssim \frac{C_N |Q|^{-1/2}}{(1 + 2^j |x - x_Q|)^N}.$$

■

Por construcción, para el caso específico de las “*shearlets* en el cono” tenemos la siguiente propiedad.

Lema 4.3.4 *Sea $\hat{\psi}$ como en la Subsección 4.2.3. Entonces, el soporte de una $(\psi_{j,\ell,k})^\wedge$ horizontal se solapa con el soporte de a lo mucho otras 10 $(\psi_{i,m,n})^\wedge$ horizontales para $(j, \ell) \neq (i, m)$ y todos los $k, n \in \mathbb{Z}^2$.*

Demostración. Esta es una consecuencia directa de la construcción y dilatación de las *shearlets*. Ya que k y n son parámetros de traslación no se ven reflejados en el soporte de $(\psi_{j,\ell,k})^\wedge$ o $(\psi_{i,m,n})^\wedge$. Sabemos que

$$\text{supp } (\psi_{j,\ell,k})^\wedge \subset \left\{ \xi \in \hat{\mathbb{R}}^2 : \xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], \left| \frac{\xi_2}{\xi_1} - \ell 2^{-j} \right| \leq 2^{-j} \right\}.$$

Por lo que respecta a ξ_1 para que $\text{supp } (\psi_{i,m,n})^\wedge$ tenga intersección no vacía con $\text{supp } (\psi_{j,\ell,k})^\wedge$ se ha de tener $2^{2j-4} < 2^{2i-1}$ y $2^{2j-1} > 2^{2i-4}$. De esto se deduce $j - \frac{3}{2} < i < j + \frac{3}{2}$, lo que ocurre sólo cuando $i = j - 1, j, j + 1$. Luego, sólo las escalas $j - 1$ y $j + 1$ tienen intersección no vacía con $\text{supp } (\psi_{j,\ell,k})^\wedge$.

1) Fijemos (j, ℓ) . Buscamos los valores de m para los que $\text{supp } (\psi_{j,m,n})^\wedge$ tiene intersección no vacía con $\text{supp } (\psi_{j,\ell,k})^\wedge$. Escribimos $|\xi_2/\xi_1 - \ell 2^{-j}| \leq 2^{-j}$ como

$$\xi_1(\ell - 1)2^{-j} \leq \xi_2 \leq \xi_1(\ell + 1)2^{-j}.$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Hemos de tener

$$(\ell + 1)2^{-j} > (m - 1)2^{-j} \text{ y } (\ell - 1)2^{-j} < (m + 1)2^{-j}.$$

De aquí se deduce $\ell - 2 < m < \ell + 2$, lo que sólo ocurre cuando $m = \ell - 1, \ell, \ell + 1$. Excluyendo $m = \ell$, se tienen dos solapamientos como máximo.

2) Miremos ahora a la escala $j - 1$. Una $(\psi_{j,\ell,k})^\wedge$ fija tiene un soporte delimitado por

$$|\xi_1| = 2^{2j-4}, |\xi_1| = 2^{2j-1}, \xi_2 = \xi_1(\ell - 1)2^{-j}, \xi_2 = \xi_1(\ell + 1)2^{-j}.$$

Una $(\psi_{j-1,m,n})^\wedge$ tiene soporte delimitado por

$$|\xi_1| = 2^{2j-6}, |\xi_1| = 2^{2j-3}, \xi_2 = \xi_1(m - 1)2^{-(j-1)}, \xi_2 = \xi_1(m + 1)2^{-(j-1)}.$$

Para que haya intersección no vacía se ha de tener

$$(\ell + 1)2^{-j} > (m - 1)2^{-(j-1)} \text{ y } (\ell - 1)2^{-j} < (m + 1)2^{-(j-1)}.$$

De aquí se deduce $\ell/2 - 3/2 < m < \ell/2 + 3/2$. Si $\ell = 2s$ es par, $s - 3/2 < m < s + 3/2$ da $m = s - 1, s, s + 1$, que son tres valores. Si $\ell = 2s + 1$ es impar $s - 1 < m < s + 2$, lo que implica $m = s, s + 1$. Por tanto, se tienen como máximo tres solapamientos.

3) Miremos ahora a la escala $j + 1$. Una $(\psi_{j+1,m,n})^\wedge$ tiene un soporte delimitado por

$$|\xi_1| = 2^{2j-2}, |\xi_1| = 2^{2j+1}, \xi_2 = \xi_1(m - 1)2^{-(j+1)}, \xi_2 = \xi_1(m + 1)2^{-(j+1)}.$$

Para que haya intersección no vacía se ha de tener

$$(\ell + 1)2^{-j} > (m - 1)2^{-(j+1)}, (\ell - 1)2^{-j} < (m + 1)2^{-(j+1)}.$$

De aquí se deduce $2\ell - 3 < m < 2\ell + 3$. Por tanto, $m = 2\ell - 2, 2\ell - 1, 2\ell, 2\ell + 1, 2\ell + 2$. Se tienen, por tanto, cinco solapamientos como máximo.

Sumando el número máximo de solapamientos en cada escala da el resultado. ■

Nota 4.3.5 Ya que los parámetros de traslación k y n no afectan el soporte en el dominio de la frecuencia, entonces para

$$f = T_\psi \mathbf{s} = \sum_{Q \in \mathcal{Q}_{AB}} s_Q \psi_Q = \sum_{i \geq 0} \sum_{m = -2^i, \dots, 2^i} \sum_{n \in \mathbb{Z}^2} s_{i,m,n} \psi_{i,m,n},$$

formalmente se tiene

$$(\tilde{\psi}_{A^{-j}B^{-\ell}} * f)(x) = \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} \sum_{Q \in \mathbb{Q}^{i,m}} s_Q(\tilde{\psi}_{A^{-j}B^{-\ell}} * \psi_Q)(x),$$

donde la suma $\sum_{i=j-1}^{j+1} \sum_{m(\ell,i)}$ tiene a lo mucho $10+1$ términos por el Lema 4.3.4 y $m(\ell, i)$ es el índice de las shearlets que rodean a $(\psi_{j,\ell,k})^\wedge$.

Nota 4.3.6 Del Lema 4.3.4 el número de solapamientos de shearlets horizontales/verticales en el dominio de Fourier está acotado para todas las escalas, ya que el sistema vertical para \mathcal{D}^v es una rotación ortonormal del sistema horizontal para \mathcal{D}^h , dejando todas las distancias y ángulos de los soportes inalterados.

4.4 La caracterización

Comenzaremos definiendo los espacios de Triebel-Lizorkin asociados a la descomposición característica de las shearets (ver Figura 4.1), pero teniendo en cuenta tanto las “direcciones” horizontales como las verticales y las bajas frecuencias representadas por φ . Daremos después la definición de los espacios de sucesiones, también teniendo en cuenta la “partición” de \mathbb{R}^2 inducida por las *shearlets*. Sin embargo, todas las demostraciones se harán para las shearlets horizontales ya que las demostraciones son similares para las verticales y las bajas frecuencias que induce φ se tratan igual que en la literatura existente (ver p.e. la Sección 12 de [41]).

4.4.1 Espacios Triebel-Lizorkin AB -anisotrópicos inhomogéneos

Sean $\varphi, \psi^h, \psi^v \in \mathcal{S}$ como en la Sección 4.2.3.

Definición 4.4.1 Sean $\alpha \in \mathbb{R}$, $0 < p < \infty$ y $0 < q \leq \infty$. Los **espacios de distribuciones Triebel-Lizorkin AB -anisotrópicos inhomogéneos** $\mathbf{F}_p^{\alpha,q}(AB)$ se definen como la colección de todas las $f \in \mathcal{S}'$ tales que

$$\begin{aligned} \|f\|_{\mathbf{F}_p^{\alpha,q}(AB)} &= \|f * \varphi\|_{L^p} \\ &+ \left\| \left(\sum_{\mathfrak{d}=\{h,v\}} \left\{ \sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} [|Q_j|^{-\alpha} \left| \tilde{\psi}_{A_{(\mathfrak{d})}^{-j}B_{(\mathfrak{d})}^{-\ell}} * f \right|^q] \right\} \right)^{1/q} \right\|_{L^p}, \end{aligned} \quad (4.28)$$

donde $\tilde{\psi}(x) = \overline{\psi(-x)}$ y $\psi_M(x) = |\det M|^{-1} \psi(M^{-1}x)$.

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Recuérdese que \mathcal{Q}_{AB} es el conjunto de los elementos $Q_{j,\ell,k}^{(\mathfrak{d})} = A_{(\mathfrak{d})}^{-j} B_{(\mathfrak{d})}^{-\ell} (Q_0 + k)$, $\mathfrak{d} = \{h, v\}$ definidos en (4.13). Usaremos temporalmente la notación \mathcal{Q}_{AB}^h para las descomposiciones de \mathbb{R}^2 asociadas al cono horizontal. Análogamente para \mathcal{Q}_{AB}^v . Para trabajar en el nivel de las sucesiones con los coeficientes de las *shearlets* también damos la siguiente definición.

Definición 4.4.2 Sean $\alpha \in \mathbb{R}$, $0 < p < \infty$ y $0 < q \leq \infty$. Los **espacios de sucesiones Triebel-Lizorkin AB-anisotrópicos inhomogéneos** $\mathbf{f}_p^{\alpha,q}(AB)$ se definen como la colección de todas las sucesiones con valores complejos $\mathbf{s} = \{s_Q : Q \in \mathcal{R}, |Q| = 1\} \cup \{s_Q : Q \in \mathcal{Q}_{AB}^{(\mathfrak{d})}, \mathfrak{d} = 1, 2\}$, tales que

$$\begin{aligned} \|\mathbf{s}\|_{\mathbf{f}_p^{\alpha,q}(AB)} &= \left\| \left(\sum_{Q \in \mathcal{R}: |Q|=1} (|s_Q| \tilde{\chi}_Q(\cdot))^q \right)^{1/q} \right\|_{L^p} \\ &+ \left\| \left(\sum_{\mathfrak{d}=\{h,v\}} \sum_{Q \in \mathcal{Q}_{AB}^{(\mathfrak{d})}} (|Q|^{-\alpha} |s_Q| \tilde{\chi}_Q(\cdot))^q \right)^{1/q} \right\|_{L^p} < \infty, \end{aligned} \quad (4.29)$$

donde $\tilde{\chi}_Q(x) = |Q|^{1/2} \chi_Q(x)$.

Como ya hemos mencionado, las demostraciones las haremos sólo para las *shearlets* “horizontales” en \mathcal{D}^h . Por abuso de notación seguiremos usando $\mathbf{F}_p^{\alpha,q}(AB)$ y $\mathbf{f}_p^{\alpha,q}(AB)$ aunque sólo usaremos el valor $\mathfrak{d} = h$ de las expresiones (4.28) y (4.29).

4.4.2 Dos resultados básicos

Definimos formalmente los **operadores de análisis** y **síntesis** como

$$S_\psi f = \{\langle f, \psi_Q \rangle\}_{Q \in \mathcal{Q}_{AB}} \quad \text{and} \quad T_\psi \mathbf{s} = \sum_{Q \in \mathcal{Q}_{AB}} s_Q \psi_Q, \quad (4.30)$$

respectivamente.

Para la prueba de nuestro resultado principal (el Teorema 4.4.11) seguimos los argumentos de [41]. La prueba está basada en una especie de desigualdad de Peetre para acotar $S_\psi : \mathbf{F}_p^{\alpha,q}(AB) \rightarrow \mathbf{f}_p^{\alpha,q}(AB)$, y una caracterización de $\mathbf{f}_p^{\alpha,q}(AB)$ para acotar $T_\psi : \mathbf{f}_p^{\alpha,q}(AB) \rightarrow \mathbf{F}_p^{\alpha,q}(AB)$. Empezamos con una definición de un resultado bien conocido.

Definición 4.4.3 La *función maximal de Hardy-Littlewood*, $\mathcal{M}f(x)$, se define como

$$\mathcal{M}f(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy,$$

para una función localmente integrable f en \mathbb{R}^2 y donde $B_r(x)$ es la bola con centro en x y radio r .

Es bien sabido que \mathcal{M} es acotado en L^p , $1 < p \leq \infty$. También es cierto que la siguiente desigualdad vectorial se cumple (ver [38]).

Teorema 4.4.4 [Fefferman-Stein] Para $1 < p < \infty$ y $1 < q \leq \infty$, existe una constante $C_{p,q}$ tal que

$$\left\| \left\{ \sum_{i=1}^{\infty} (\mathcal{M}f_i)^q \right\}^{1/q} \right\|_{L^p} \leq C_{p,q} \left\| \left\{ \sum_{i=1}^{\infty} f_i^q \right\}^{1/q} \right\|_{L^p},$$

para cualquier sucesión $\{f_i : i = 1, 2, \dots\}$ de funciones localmente integrables.

Definimos ahora la **función maximal shear anisotrópica de Peetre**. Para todo $\lambda > 0$,

$$(\psi_{j,\ell,\lambda}^{**}f)(x) := \sup_{y \in \mathbb{R}^2} \frac{|(\psi_{A^{-j}B^{-\ell}} * f)(x-y)|}{(1 + |B^\ell A^j y|)^{2\lambda}}. \quad (4.31)$$

Entonces, se tiene una **desigualdad de Peetre shear anisotrópica**.

Lema 4.4.5 Sea ψ de banda limitada y $f \in S'$. Entonces, para cualquier $\lambda > 0$, existe una constante C_λ tal que

$$(\psi_{j,\ell,\lambda}^{**}f)(x) \leq C_\lambda \left\{ \mathcal{M}(|\psi_{A^{-j}B^{-\ell}} * f|^{1/\lambda})(x) \right\}^\lambda, \quad x \in \mathbb{R}^2.$$

Para la demostración de este lema algunas definiciones y resultados previos son necesarios.

Definición 4.4.6 Para una función g definida en \mathbb{R}^2 y para todo número real $\lambda > 0$ la **función maximal de Peetre** (ver Lema 2.1 de [86]) es

$$g_\lambda^*(x) = \sup_{y \in \mathbb{R}^2} \frac{|g(x-y)|}{(1 + |y|)^{2\lambda}}, \quad x \in \mathbb{R}^2.$$

Lema 4.4.7 Sea $g \in S'(\mathbb{R}^2)$ tal que $\text{supp}(\hat{g}) \subseteq \{\xi \in \hat{\mathbb{R}}^2 : |\xi| \leq R\}$ para algún $R > 0$. Entonces, para cada $\lambda > 0$ existe un $C_\lambda > 0$ tal que, para $|\alpha| = 1$,

$$(\partial^\alpha g)_\lambda^*(x) \leq C_\lambda g_\lambda^*(x), \quad x \in \mathbb{R}^2.$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Demostración. Ya que $g \in \mathcal{S}'$ tiene soporte compacto en el dominio de Fourier, g es regular. Más precisamente, por el teorema de Paley-Wiener-Schwartz g es lentamente creciente (a lo mucho polinomial) e infinitamente diferenciable (Theorem 7.3.1 en [62]). Sea γ una función en la clase de Schwartz tal que $\hat{\gamma}(\xi) = 1$ si $|\xi| \leq R$. Entonces, $\hat{\gamma}(\xi)\hat{g}(\xi) = \hat{g}(\xi)$ para todo $\xi \in \hat{\mathbb{R}}^2$. Así pues, $\gamma * g = g$ y $\partial^\alpha g = \partial^\alpha \gamma * g$. Más aún,

$$\begin{aligned} |\partial^\alpha g(x-y)| &= \left| \int_{\mathbb{R}^2} \partial^\alpha \gamma(x-y-z)g(z)dz \right| = \left| \int_{\mathbb{R}^2} \partial^\alpha \gamma(w-y)g(x-w)dw \right| \\ &\leq \int_{\mathbb{R}^2} |\partial^\alpha \gamma(w-y)| (1+|w-y|)^{2\lambda} (1+|y|)^{2\lambda} \frac{|g(x-w)|}{(1+|w|)^{2\lambda}} dw, \end{aligned}$$

por la desigualdad triangular. Por lo tanto,

$$|\partial^\alpha g(x-y)| \leq g_\lambda^*(x) (1+|y|)^{2\lambda} \int_{\mathbb{R}^2} |\partial^\alpha \gamma(w-y)| (1+|w-y|)^{2\lambda} dw.$$

Ya que $\gamma \in \mathcal{S}$, la última integral es igual a una constante finita c_λ , independiente de y , y se obtiene

$$|\partial^\alpha g(x-y)| \leq c_\lambda g_\lambda^*(x) (1+|y|)^{2\lambda},$$

lo que muestra el resultado deseado. ■

Se tiene una relación entre la función maximal de Hardy-Littlewood y g_λ^* .

Lema 4.4.8 Sean $\lambda > 0$ y $g \in \mathcal{S}'$ tales que $\text{supp}(\hat{g}) \subseteq \{\xi \in \hat{\mathbb{R}}^2 : |\xi| \leq R\}$ para algún $R > 0$. Entonces, existe una constante $C_\lambda > 0$ tal que

$$g_\lambda^*(x) \leq C_\lambda \left(\mathcal{M}(|g|^{1/\lambda})(x) \right)^\lambda, \quad x \in \mathbb{R}^2.$$

Demostración. Como g es de banda limitada, g es diferenciable en \mathbb{R}^2 (por el teorema de Paley-Wiener-Schwartz), así que podemos considerar valores puntuales de g . Sean $x, y \in \mathbb{R}^2$ y $0 < \delta < 1$. Escójase $z \in \mathbb{R}^2$ tal que $z \in B_\delta(x-y)$. Aplicamos el teorema de valor medio a g y los extremos $x-y$ y z para obtener

$$|g(x-y)| \leq |g(z)| + \delta \sup_{w: w \in B_\delta(x-y)} (|\nabla g(w)|).$$

Tomando la $(1/\lambda)^a$ potencia e integrando con respecto a la variable z sobre $B_\delta(x-y)$, se obtiene

$$\begin{aligned} |g(x-y)|^{1/\lambda} &\leq \frac{c_\lambda}{|B_\delta(x-y)|} \int_{B_\delta(x-y)} |g(z)|^{1/\lambda} dz \\ &\quad + c_\lambda \delta^{1/\lambda} \sup_{w: w \in B_\delta(x-y)} (|\nabla g(w)|)^{1/\lambda}. \end{aligned} \quad (4.32)$$

Como $B_\delta(x-y) \subset B_{\delta+|y|}(x)$,

$$\int_{B_\delta(x-y)} |g(z)|^{1/\lambda} dz \leq \int_{B_{\delta+|y|}(x)} |g(z)|^{1/\lambda} dz \leq |B_{\delta+|y|}(x)| \mathcal{M}(|g|^{1/\lambda})(x),$$

y el término sup en el lado derecho de (4.32) está acotado por

$$\begin{aligned} \sup_{w:w \in B_{\delta+|y|}(x)} (|\nabla g(w)|)^{1/\lambda} &= \sup_{t:|t| < \delta+|y|} (|\nabla g(x-t)|)^{1/\lambda} \\ &\leq (1 + \delta + |y|)^2 [(\nabla g)_\lambda^*(x)]^{1/\lambda}. \end{aligned}$$

Substituyendo estas dos últimas desigualdades en (4.32) da

$$\begin{aligned} |g(x-y)|^{1/\lambda} &\leq c_\lambda \frac{|B_{\delta+|y|}(x)|}{|B_\delta(x-y)|} \mathcal{M}(|g|^{1/\lambda})(x) \\ &\quad + c_\lambda \delta^{1/\lambda} (1 + \delta + |y|)^2 [(\nabla g)_\lambda^*(x)]^{1/\lambda}, \end{aligned}$$

y como $|B_{\delta+|y|}(x)| / |B_\delta(x-y)| = (\delta + |y|)^2 / \delta^2$, se obtiene

$$\begin{aligned} |g(x-y)|^{1/\lambda} &\leq c_\lambda \frac{(\delta + |y|)^2}{\delta^2} \mathcal{M}(|g|^{1/\lambda})(x) \\ &\quad + c_\lambda \delta^{1/\lambda} (1 + \delta + |y|)^2 [(\nabla g)_\lambda^*(x)]^{1/\lambda}. \end{aligned}$$

Tomando la λ^a potencia da

$$\frac{|g(x-y)|}{(1+|y|)^{2\lambda}} \leq c'_\lambda \left\{ \frac{1}{\delta^{2\lambda}} [\mathcal{M}(|g|^{1/\lambda})(x)]^\lambda + \delta [(\nabla g)_\lambda^*(x)] \right\},$$

ya que $\delta < 1$ implica $(1 + \delta + |y|) \leq 2(1 + |y|)$. Tomando δ suficientemente pequeña para que $c'_\lambda C_\lambda \delta < 1/4$ (donde C_λ es la constante en el Lema 4.4.7) se obtiene

$$g_\lambda^*(x) \leq c_\lambda [\mathcal{M}(|g|^{1/\lambda})(x)]^\lambda + \frac{1}{2} g_\lambda^*(x).$$

Supóngase por el momento que $g \in \mathcal{S}$, así pues $g_\lambda^*(x) < \infty$. Entonces, podemos sustraer el segundo término en el lado derecho de la desigualdad anterior del lado izquierdo y la demostración está completa para $g \in \mathcal{S}$. Para quitar la suposición $g \in \mathcal{S}$, se aplica un argumento usual de regularización a una $g \in \mathcal{S}'$ como se hace en p. 22 de [100] o en el Lema A.4 de [41]. Sea $\gamma \in \mathcal{S}$ que satisfaga $\text{supp } \hat{\gamma} \subset B(0,1)$, $\hat{\gamma}(\xi) \geq 0$ y $\gamma(0) = 1$ ($\gamma(0) = \int \hat{\gamma}(\xi)$). Por la inversión de la transformada de Fourier $|\gamma(x)| \leq 1$ para todo $x \in \mathbb{R}^2$. Para $0 < \delta < 1$, sea $g_\delta(x) = g(x)\gamma(\delta x)$. Entonces, $\text{supp } \hat{g}_\delta$ también es compacto, $g_\delta \in \mathcal{S}$, $|g_\delta| \leq |g|$ para todo $x \in \mathbb{R}^2$ y $g_\delta \rightarrow g$ uniformemente en conjuntos

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

compactos cuando $\delta \rightarrow 0$. Aplicando el resultado previo a g_δ y dejando que $\delta \rightarrow 0$ se obtiene el resultado para \mathcal{S}' en general. ■

Demostración del Lema 4.4.5. Sea $g(x) = (\psi_{A^{-j}B^{-\ell}} * f)(x)$. Ya que ψ es de banda limitada, también lo es g . Por una parte, como $j \geq 0$ y $2^{j-1}|y| \leq |B^\ell A^j y|$,

$$\begin{aligned} g_\lambda^*(t) &= \sup_{y \in \mathbb{R}^2} \frac{|g(t-y)|}{(1+|y|)^\lambda} \geq \sup_{y \in \mathbb{R}^2} \frac{|(\psi_{A^{-j}B^{-\ell}} * f)(t-y)|}{(1+2^j|y|)^\lambda} \\ &= \sup_{y \in \mathbb{R}^2} \frac{|(\psi_{A^{-j}B^{-\ell}} * f)(t-y)|}{2^\lambda(2^{-1}+2^{j-1}|y|)^\lambda} \\ &\geq 2^{-\lambda} \sup_{y \in \mathbb{R}^2} \frac{|(\psi_{A^{-j}B^{-\ell}} * f)(t-y)|}{(1+|B^\ell A^j y|)^\lambda} = 2^{-\lambda} |(\psi_{j,\ell,\lambda}^{**})(t)|. \end{aligned}$$

Por otra parte,

$$\begin{aligned} \mathcal{M}(|g|^{1/\lambda})(t) &= \sup_{r>0} \frac{1}{|B_r(t)|} \int_{B_r(t)} |(\psi_{A^{-j}B^{-\ell}} * f)(y)|^{1/\lambda} dy \\ &= \mathcal{M}(|\psi_{A^{-j}B^{-\ell}} * f|^{1/\lambda})(t). \end{aligned}$$

El resultado se sigue del Lema 4.4.8 con $t = x$. ■

Sea $Q \in \mathcal{Q}^{j,\ell}$, $j \geq 0$, $\ell = 2^{-j}, \dots, 2^j$. Para todo $r > 0$ y $N > 0$, defínase

$$(s_{r,N}^*)_Q := \left(\sum_{P \in \mathcal{Q}^{j,\ell}} \frac{|s_P|^r}{(1+2^j|x_Q - x_P|)^N} \right)^{1/r}.$$

para $\{s_P : P \in \mathcal{Q}^{j,\ell}\}$ una colección de números complejos. Escribimos $\mathbf{s}_{r,N}^*$ para denotar la sucesión $\{(s_{r,N}^*)_Q\}_{Q \in \mathcal{Q}_{AB}}$.

Tenemos la siguiente caracterización de los espacios $\mathbf{f}_p^{\alpha,q}(AB)$ que se usará en la prueba de la acotabilidad de $T_\psi : \mathbf{f}_p^{\alpha,q}(AB) \rightarrow \mathbf{F}_p^{\alpha,q}(AB)$.

Lema 4.4.9 Sean $\alpha \in \mathbb{R}$, $0 < p < \infty$ and $0 < q \leq \infty$. Entonces, para todo $r > 0$ y $N > 3 \max(1, r/q, r/p)$ existe $C > 0$ tal que para todo $\mathbf{s} = \{s_Q\}_{Q \in \mathcal{Q}_{AB}}$ se tiene

$$\|\mathbf{s}\|_{\mathbf{f}_p^{\alpha,q}(AB)} \leq \|\mathbf{s}_{r,N}^*\|_{\mathbf{f}_p^{\alpha,q}(AB)} \leq C \|\mathbf{s}\|_{\mathbf{f}_p^{\alpha,q}(AB)}.$$

Para probar el Lema 4.4.9 usaremos el siguiente resultado.

Lema 4.4.10 Sean $0 < a \leq r$ y $Q = Q_{j,m,n}$, $j \geq 0$, $m = -2^j, \dots, 2^j$. Para todo $N > 0$ y cualquier sucesión de números complejos $\{s_P\}_{P \in Q^{j,\ell}}$ ($j \geq 0$, $\ell = -2^j, \dots, 2^j$) y para cualquier $x \in Q$ se tiene

$$(s_{r,N}^*)_Q := \left(\sum_{P \in Q^{j,\ell}} \frac{|s_P|^r}{(1 + 2^j |x_Q - x_P|)^N} \right)^{1/r} \leq C_{a,r} 2^{j(\frac{r}{a} - \frac{N}{3})} \left[\mathcal{M} \left(\sum_{P \in Q^{j,\ell}} |s_P|^a \chi_P \right) (x) \right]^{r/a}.$$

Si $N > \frac{3r}{a}$ y $0 < q \leq \infty$ se tiene

$$\sum_{P \in Q^{j,\ell}} [(s_{r,N}^*)_P \tilde{\chi}_P(x)]^q \leq C_{a,r} \left[\mathcal{M} \left(\sum_{P \in Q^{j,\ell}} (|s_P| \tilde{\chi}_P)^a \right) (x) \right]^{q/a}.$$

Demostración. Si $Q = Q_{j,m,n}$ y $P = P_{j,\ell,k}$ escribimos $x_Q = A^{-j}B^{-m}n$ y $x_P = A^{-j}B^{-\ell}k$ para denotar las “esquinas inferiores izquierdas” de Q y P , respectivamente. Escribáse $d_{Q,P} = |x_Q - x_P|$. Definimos

$$D_0 = \{P \in Q^{j,\ell} : 2^{3j/2} d_{Q,P} \leq 1\}$$

y

$$D_\nu = \{P \in Q^{j,\ell} : 2^{\nu-1} < 2^{3j/2} d_{Q,P} \leq 2^\nu\}, \quad \nu = 1, 2, 3, \dots,$$

que forman una partición de conjuntos disjuntos de $Q^{j,\ell}$. Por tanto

$$\sum_{P \in Q^{j,\ell}} \frac{|s_P|^r}{(1 + 2^j d_{Q,P})^N} = \left(\sum_{P \in D_0} + \sum_{\nu=1}^{\infty} \sum_{P \in D_\nu} \right) \frac{|s_P|^r}{(1 + 2^j d_{Q,P})^N}.$$

La expresión anterior puede acotarse por

$$\begin{aligned} \sum_{P \in D_0} |s_P|^r + \sum_{\nu=1}^{\infty} \sum_{P \in D_\nu} \frac{|s_P|^r}{(2^{3j/2} d_{Q,P})^{2N/3}} \\ \leq \sum_{P \in D_0} |s_P|^r + \sum_{\nu=1}^{\infty} 2^{2N/3} \sum_{P \in D_\nu} \frac{|s_P|^r}{2^{\nu 2N/3}} \\ \leq 2^{2r/a} \sum_{\nu=0}^{\infty} 2^{-\nu 2N/3} \left(\sum_{P \in D_\nu} |s_P|^r \right). \end{aligned}$$

Como $a \leq r$ se tiene $\ell^a \hookrightarrow \ell^r$, por lo que la expresión anterior se acota por

$$2^{2r/a} \sum_{\nu=0}^{\infty} 2^{-\nu 2N/3} \left(\sum_{P \in D_\nu} |s_P|^a \right)^{r/a}.$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

El mayor diámetro de $Q_{j,m,n}$ se alcanza cuando $m = \pm 2^j$ y su valor es $\text{diam } Q_{j,m,n} = \sqrt{(4^{-j} + 2^j 4^{-j})^2 + (2^{-j})^2} < 2^{-j} \sqrt{5}$. Análogamente, el mayor diámetro de $P = P_{j,\ell,k}$ se alcanza cuando $\ell = \pm 2^j$ y su valor es $\text{diam } P_{j,\ell,k} = \sqrt{(4^{-j} + 2^j 4^{-j})^2 + (2^{-j})^2} < 2^{-j} \sqrt{5}$. Por tanto, si $x \in Q_{j,m,n}$ y $P \in \mathcal{D}_\nu$,

$$P \subset B_{2^{-j}\sqrt{5}+2^\nu-\frac{3j}{2}+2^{-j}\sqrt{5}}(x) = B_{2\sqrt{5}\cdot 2^{-j}+2^\nu-\frac{3j}{2}}(x) \subset B_{2\sqrt{5}(2^{-j}+2^\nu-\frac{3j}{2})}(x).$$

Sea $\tilde{B}(x) = B_{2\sqrt{5}(2^{-j}+2^\nu-\frac{3j}{2})}(x)$. Se tiene que

$$\sum_{P \in \mathcal{D}^\nu} |s_P|^a = 2^{3j} \int_{\tilde{B}(x)} \left(\sum_{P \in \mathcal{D}^\nu} |s_P|^a \right) \chi_P(y) dy.$$

Por tanto, para $x \in Q_{j,m,n}$ podemos continuar la acotación anterior para obtener

$$\begin{aligned} \sum_{P \in \mathcal{Q}^{j,\ell}} \frac{|s_P|^r}{(1 + 2^j d_{Q,P})^N} &\leq 2^{2ra} \sum_{\nu=0}^{\infty} 2^{-\nu 2N/3} \left(2^{3j} \frac{|\tilde{B}(x)|}{|\tilde{B}(x)|} \int_{\tilde{B}(x)} \sum_{P \in \mathcal{D}^\nu} |s_P|^a \chi_P(y) dy \right)^{r/a} \\ &\leq 2^{ra} \sum_{\nu=0}^{\infty} 2^{-\nu 2N/3} 2^{3jr/a} |\tilde{B}(x)|^{r/a} \left(\mathcal{M} \left(\sum_{P \in \mathcal{D}^\nu} |s_P|^a \chi_P \right) (x) \right)^{r/a} \end{aligned}$$

Si $0 \leq \nu \leq j/2$ se tiene que $-j \geq \nu - 3j/2$ y por tanto $|\tilde{B}(x)| \leq C 2^{-2j}$. Por el contrario, si $\nu > j/2$ se tiene que $-j < \nu - 3j/2$ y entonces $|\tilde{B}(x)| \leq 2^{2\nu-3j}$. En consecuencia, como $\mathcal{D}_\nu \subset \mathcal{Q}^{j,\ell}$, $\nu = 0, 1, \dots$, se tiene

$$\sum_{P \in \mathcal{Q}^{j,\ell}} \frac{|s_P|^r}{(1 + 2^j d_{Q,P})^N} \leq 2^{ra} \left[\mathcal{M} \left(\sum_{P \in \mathcal{Q}^{j,\ell}} |s_P|^a \chi_P \right) (x) \right]^{r/a} \left(\sum_{\nu=0}^{\infty} 2^{-\nu 2N/3} 2^{3jr/a} |\tilde{B}(x)|^{r/a} \right),$$

y

$$\begin{aligned} \sum_{\nu=0}^{\infty} 2^{-\nu 2N/3} 2^{3jr/a} |\tilde{B}(x)|^{r/a} &\leq C \sum_{\nu=0}^{j/2} 2^{-\nu 2N/3} 2^{3jr/a} 2^{-2jr/a} \\ &\quad + C \sum_{\nu > j/2} 2^{-\nu 2N/3} 2^{3jr/a} 2^{(2\nu-3j)r/a} \\ &\leq C 2^{jr/a} 2^{-jN/3} + C 2^{-j(N/3-r/a)} = C 2^{j(r/a-N/3)}. \end{aligned}$$

Cuando $N > 3r/a$, se tiene que $2^{j(r/a-N/3)} \leq 1$ para todo $j = 0, 1, \dots$. Para probar la segunda desigualdad, multiplíquense ambos lados por $\tilde{\chi}_Q(x)$, elévese a la potencia q

y súpese sobre $Q \in \mathcal{Q}^{j,\ell}$ para obtener

$$\begin{aligned}
 \sum_{Q \in \mathcal{Q}^{j,\ell}} [(s_{r,N}^*)_Q \tilde{\chi}_Q(x)]^q &\leq C \sum_{Q \in \mathcal{Q}^{j,\ell}} \left[\mathcal{M} \left(\sum_{P \in \mathcal{Q}^{j,\ell}} |s_P|^a \chi_P \right) (x) \right]^{q/a} \tilde{\chi}_Q^q(x) \\
 &= C \sum_{Q \in \mathcal{Q}^{j,\ell}} \left[\mathcal{M} \left(\sum_{P \in \mathcal{Q}^{j,\ell}} (|s_P| \tilde{\chi}_P)^a \right) (x) \right]^{q/a} \chi_Q(x) \\
 &= C \left[\mathcal{M} \left(\sum_{P \in \mathcal{Q}^{j,\ell}} (|s_P| \tilde{\chi}_P)^a \right) (x) \right]^{q/a},
 \end{aligned}$$

ya que $\mathcal{Q}^{j,\ell}$ es una partición de \mathbb{R}^2 . ■

Demostración del Lema 4.4.9. Estamos suponiendo que $N > 3 \max(1, r/q, r/p)$.

Si $r < \min(q, p)$, escójase $a = r$. De otra forma, si $r \geq \min(q, p)$, escójase a tal que

$3r/N < a < \min(r, q, p)$. Siempre es posible elegir un a de esta forma puesto que

$N > 3 \max(1, r/q, r/p)$ implica $3r/N < \min(r, q, p)$. En ambos casos se tiene que

$$0 < a \leq r < \infty, \quad N > 3r/a, \quad q/a > 1, \quad p/a > 1.$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

El argumento previo es similar al de [13]. Entonces,

$$\begin{aligned}
\|s_{r,N}^*\|_{\mathbf{f}_p^{\alpha,q}(AB)} &= \left\| \left(\sum_{P \in Q_{AB}} (|P|^{-\alpha} (s_{r,N}^*)_P \tilde{\chi}_P)^q \right)^{1/q} \right\|_{L^p} \\
&= \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} 2^{3j\alpha q} \sum_{P \in Q^{j,\ell}} (s_{r,N}^*)_P^q \tilde{\chi}_P^q \right)^{1/q} \right\|_{L^p} \\
&\leq C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} 2^{3j\alpha q} \left[\mathcal{M} \left(\sum_{P \in Q^{j,\ell}} (|s_P| \tilde{\chi}_P)^a \right) \right]^{q/a} \right)^{1/q} \right\|_{L^p} \\
&= C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left[\mathcal{M} \left(\sum_{P \in Q^{j,\ell}} (|P|^{-\alpha} |s_P| \tilde{\chi}_P)^a \right) \right]^{q/a} \right)^{1/q} \right\|_{L^p} \\
&= C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left[\mathcal{M} \left(\sum_{P \in Q^{j,\ell}} (|P|^{-\alpha} |s_P| \tilde{\chi}_P)^a \right) \right]^{q/a} \right)^{a/q} \right\|_{L^{p/a}}^{1/a} \\
&\leq C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left(\sum_{P \in Q^{j,\ell}} (|P|^{-\alpha} |s_P| \tilde{\chi}_P)^a \right)^{q/a} \right)^{a/q} \right\|_{L^{p/a}}^{1/a} \\
&= C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{P \in Q^{j,\ell}} (|P|^{-\alpha} |s_P| \tilde{\chi}_P)^q \right)^{a/q} \right\|_{L^{p/a}}^{1/a} \\
&= C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{P \in Q^{j,\ell}} (|P|^{-\alpha} |s_P| \tilde{\chi}_P)^q \right)^{1/q} \right\|_{L^p} \\
&= C \|s\|_{\mathbf{f}_p^{\alpha,q}(AB)},
\end{aligned}$$

la primera desigualdad es por la segunda parte del Lema 4.4.10 y la segunda por el Teorema 4.4.4.

La desigualdad inversa es trivial ya que $|s_Q| \leq (s_{r,N}^*)_Q$ siempre se verifica. ■

4.4.3 Acotabilidad de S_ψ y T_ψ

Estamos en condiciones de demostrar el resultado principal de este capítulo en el que se muestra la acotabilidad de los operadores de análisis S_ψ y de síntesis T_ψ entre los espacios de Triebel-Lizorkin definidos en (4.28) y sus correspondientes espacios de sucesiones (4.29). Recordar que

$$S_\psi f = \{\langle f, \psi_Q \rangle\}_{Q \in \mathcal{Q}_{AB}} \quad \text{and} \quad T_\psi \mathbf{s} = \sum_{Q \in \mathcal{Q}_{AB}} s_Q \psi_Q,$$

como se definieron en (4.30).

Teorema 4.4.11 Sean $\alpha \in \mathbb{R}$, $0 < p < \infty$ y $0 < q \leq \infty$. Entonces, los operadores $S_\psi : \mathbf{F}_p^{\alpha,q}(AB) \rightarrow \mathbf{f}_p^{\alpha,q}(AB)$ y $T_\psi : \mathbf{f}_p^{\alpha,q}(AB) \rightarrow \mathbf{F}_p^{\alpha,q}(AB)$ están bien definidos y son acotados.

Demostración. Probamos sólo el caso $q < \infty$. Para probar la acotabilidad de S_ψ supóngase $f \in \mathbf{F}_p^{\alpha,q}(AB)$. Identifíquese P con (j, ℓ, k) . Entonces, $\left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(x_P) \right| \chi_P = |\langle f, \psi_P \rangle| \tilde{\chi}_P$, como en (4.27). Ya que $\mathcal{Q}^{j,\ell}$ es una partición de \mathbb{R}^2 , se tiene que para $x \in P'$ y $P' \in \mathcal{Q}^{j,\ell}$,

$$\begin{aligned} & \sum_{P \in \mathcal{Q}^{j,\ell}} [|P|^{-\alpha} |(S_\psi f)_P| \tilde{\chi}_P(x)]^q \\ &= |\det A|^{j\alpha q} \sum_{P \in \mathcal{Q}^{j,\ell}} \left[\left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(x_P) \right| \chi_P(x) \right]^q \\ &\leq |\det A|^{j\alpha q} \sum_{P \in \mathcal{Q}^{j,\ell}} \sup_{y \in P} \left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(y) \right|^q \chi_P(x) \\ &\leq |\det A|^{j\alpha q} \sup_{|z| < \text{Diam}(P')} \left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(x-z) \right|^q \\ &= |\det A|^{j\alpha q} \sup_{z \in \mathcal{Q}_{j,\ell,0}} \left[\frac{\left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(x-z) \right|}{(1 + |B^\ell A^j z|)^{2/\lambda}} \right]^q (1 + |B^\ell A^j z|)^{q2/\lambda} \\ &\leq |\det A|^{j\alpha q} \left[\sup_{z \in \mathbb{R}^2} \frac{\left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(x-z) \right|}{(1 + |B^\ell A^j z|)^{2/\lambda}} \right]^q (1 + \text{Diam}(Q_{0,0,0}))^{2q/\lambda} \\ &= C_{q,\lambda} |\det A|^{j\alpha q} (\tilde{\psi}_{j,\ell,1/\lambda}^{**} f)^q(x) \\ &\leq C_{q,\lambda} |\det A|^{j\alpha q} \left\{ \mathcal{M} \left(\left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f \right|^\lambda \right) (x) \right\}^{q/\lambda}, \end{aligned}$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

por el Lema 4.4.5 (con $1/\lambda$ en vez de λ en la última desigualdad). Ahora, tómesese $0 < \lambda < \min(p, q)$. Entonces, la estimación previa y el Teorema 4.4.4 dan

$$\begin{aligned}
\|S_\psi f\|_{\mathbf{f}_p^{\alpha, q}(AB)} &= \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \sum_{P \in Q^{j, \ell}} [|P|^{-\alpha} |(S_\psi f)_P| \tilde{\chi}_P]^q \right)^{1/q} \right\|_{L^p} \\
&\leq C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left\{ \mathcal{M} \left(|\det A|^{j\alpha\lambda} |\tilde{\psi}_{A^{-j}B^{-\ell}} * f|^\lambda \right) \right\}^{q/\lambda} \right)^{1/q} \right\|_{L^p} \\
&= C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} \left\{ \mathcal{M} \left(|\det A|^{j\alpha\lambda} |\tilde{\psi}_{A^{-j}B^{-\ell}} * f|^\lambda \right) \right\}^{q/\lambda} \right)^{\lambda/q} \right\|_{L^{p/\lambda}}^{1/\lambda} \\
&\leq C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} |\det A|^{j\alpha q} |\tilde{\psi}_{A^{-j}B^{-\ell}} * f|^q \right)^{\lambda/q} \right\|_{L^{p/\lambda}}^{1/\lambda} \\
&= C \left\| \left(\sum_{j \geq 0} \sum_{\ell = -2^j}^{2^j} [|\det A|^{j\alpha} |\tilde{\psi}_{A^{-j}B^{-\ell}} * f|]^q \right)^{1/q} \right\|_{L^p} \\
&= C \|f\|_{\mathbf{f}_p^{\alpha, q}(AB)}.
\end{aligned}$$

Para probar la acotabilidad de T_ψ supóngase que $\mathbf{s} = \{s_Q\}_Q \in \mathbf{f}_p^{\alpha, q}$ y $f = T_\psi \mathbf{s} = \sum_{Q \in \mathcal{Q}_{AB}} s_Q \psi_Q$. Entonces,

$$(\tilde{\psi}_{A^{-j}B^{-\ell}} * f)(x) = \sum_{i=j-1}^{j+1} \sum_{m(\ell, i)} \sum_{Q \in \mathcal{Q}^{i, m}} s_Q (\tilde{\psi}_{A^{-j}B^{-\ell}} * \psi_Q)(x),$$

Donde la suma $\sum_{i=j-1}^{j+1} \sum_{m(\ell,i)}$ tiene a lo mucho 11 términos por el Lema 4.3.4 (ver también la Nota 4.3.5). Para $x \in Q$ y $Q \in \mathcal{Q}^{i,m}$, el Lema 4.3.1 da

$$\begin{aligned}
 \left| \tilde{\psi}_{A^{-j}B^{-\ell}} * f(x) \right| &\leq \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} \sum_{Q \in \mathcal{Q}^{i,m}} |s_Q| \left| \tilde{\psi}_{A^{-j}B^{-\ell}} * \psi_Q(x) \right| \\
 &\leq C \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} \sum_{Q \in \mathcal{Q}^{i,m}} |s_Q| \frac{|Q|^{-1/2}}{(1+2^i|x-x_Q|)^N} \\
 &\leq C' \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} \sum_{Q \in \mathcal{Q}^{i,m}} |s_Q| \frac{|Q|^{-1/2}}{(1+2^i|x_{Q'}-x_Q|)^N} \\
 &= C' \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} |Q|^{-1/2} (s_{1,N}^*)_{Q'} \chi_{Q'}(x) \\
 &= C' \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} \sum_{Q \in \mathcal{Q}^{i,m}} (s_{1,N}^*)_Q \tilde{\chi}_Q(x),
 \end{aligned}$$

para todo $N > 2$ y porque $\mathcal{Q}^{i,m}$ es una partición de \mathbb{R}^2 . Con $r = 1$, sea $N > 3 \max(1, 1/q, 1/p)$. Entonces, la estimación previa y el Lema 4.4.9 da

$$\begin{aligned}
 &\|T_\psi \mathbf{s}\|_{\mathbf{F}_p^{\alpha,q}(AB)} \\
 &= \left\| \left(\sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} [|Q_j|^{-\alpha} |\tilde{\psi}_{A^{-j}B^{-\ell}} * f|^q] \right)^{1/q} \right\|_{L^p} \\
 &\leq C \left\| \left(\sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} \left[2^{3j\alpha} \sum_{i=j-1}^{j+1} \sum_{m(\ell,i)} \sum_{Q \in \mathcal{Q}^{i,m}} (s_{1,N}^*)_Q \tilde{\chi}_Q \right]^q \right)^{1/q} \right\|_{L^p} \\
 &= C \left\| \left(\sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} \left[\sum_{\nu=-1}^{+1} \sum_{m(\nu)} \sum_{Q \in \mathcal{Q}^{(j+\nu), \ell(j+\nu)}} 2^{3\alpha(j+\nu-\nu)} (s_{1,N}^*)_Q \tilde{\chi}_Q \right]^q \right)^{1/q} \right\|_{L^p} \\
 &\leq C_{p,q} \sum_{\nu=-1}^{+1} \sum_{m(\nu)} 2^{-3\nu\alpha} \left\| \left(\sum_{j \geq 0} \sum_{\ell=-2^j}^{2^j} \sum_{Q \in \mathcal{Q}^{j,\ell}} [|Q_{j,\ell}|^{-\alpha} (s_{1,N}^*)_Q \tilde{\chi}_Q]^q \right)^{1/q} \right\|_{L^p} \\
 &= C_{\alpha,p,q} \|\mathbf{s}_{1,N}^*\|_{\mathbf{F}_p^{\alpha,q}(AB)} \leq C_{\alpha,p,q} \|\mathbf{s}\|_{\mathbf{F}_p^{\alpha,q}(AB)},
 \end{aligned}$$

donde en la penúltima desigualdad usamos que $\sum_{\nu=-1}^{+1} \sum_{\bar{m}(\nu)}$ tiene un número finito de términos y que $\mathcal{Q}^{j,\ell}$ es una partición de \mathbb{R}^2 . ■

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

La siguiente es una reformulación de la Nota 2.6 en [41] en el ámbito de las “*shearlets* en el cono”.

Nota 4.4.12 Sean $\psi^{(1)}, \psi^{(2)} \in \mathcal{S}(\mathbb{R}^2)$ tales que $\hat{\psi}_1^{(1)}, \hat{\psi}_2^{(1)}$ y $\hat{\psi}_1^{(2)}, \hat{\psi}_2^{(2)}$ (y sus soportes respectivos) estén definidos como en la Subsección 4.2.3 y que verifiquen (4.20) y (4.21). Defínase $\mathbf{F}_p^{\alpha,q}(AB, \psi^{(1)})$ y $\mathbf{F}_p^{\alpha,q}(AB, \psi^{(2)})$ como en (4.28) con $\psi^{(1)}$ y $\psi^{(2)}$ en lugar de ψ . Note que la prueba de la acotabilidad de $T_\psi : \mathbf{f}_p^{\alpha,q} \rightarrow \mathbf{F}_p^{\alpha,q}$ de arriba requiere sólo que $\psi \in \mathcal{S}$ sea tal que $\hat{\psi}$ verifique (4.20) y (4.21) con $\text{supp } \hat{\psi}$ como en la Subsección 4.2.3. De esta forma podemos aplicar el Lema 4.3.1 a $\psi^{(1)}$ y a $\psi^{(2)}$ y el Lema 4.3.4 al $\text{supp } \hat{\psi}_1^{(1)}$ y al $\text{supp } \hat{\psi}_2^{(2)}$. Por lo tanto,

$$\begin{aligned} \|f\|_{\mathbf{F}_p^{\alpha,q}(AB, \psi^{(1)})} &= \left\| \sum_{Q \in \Omega_{AB}} (S_{\psi^{(2)}} f)_Q \psi_Q^{(2)} \right\|_{\mathbf{F}_p^{\alpha,q}(AB, \psi^{(1)})} \\ &\leq \left\| (S_{\psi^{(2)}} f)_Q \right\|_{\mathbf{f}_p^{\alpha,q}(AB)} \leq C' \|f\|_{\mathbf{F}_p^{\alpha,q}(AB, \psi^{(2)})}. \end{aligned}$$

Esto muestra que la definición de $\mathbf{F}_p^{\alpha,q}(AB)$ es independiente de la elección de $\psi \in \mathcal{S}$ siempre y cuando se satisfagan los requerimientos de la Subsección 4.2.3.

4.5 La identidad con los marcos de Parseval suaves

Recientemente, Guo and Labate en [54] encontraron una forma de evitar el uso de las funciones características en el dominio de Fourier (ver (4.25)) para restringir las *shearlets* horizontales/verticales a sus respectivos conos. El uso de estas funciones características afectan la suavidad de las *shearlets en la frontera* (aquellas con $\ell = \pm 2^j$). Modifican ligeramente la definición de estas *shearlets* en la frontera en vez de proyectarlas en el cono. Este nuevo sistemas de *shearlets* no es afín. Sin embargo, sí producen la misma teselación de frecuencia que el de la Subsección 4.2.3.

4.5.1 El nuevo sistema de *shearlets* suaves

Esta subsección es un breve resumen de algunos de los resultados en [54] y la intención es mostrar la construcción de los marcos de Parseval suaves. Sea ϕ en C^∞ una función univariable tal que $0 \leq \hat{\phi} \leq 1$, con $\hat{\phi} = 1$ en $[-1/16, 1/16]$ y $\hat{\phi} = 0$ afuera de $[-1/8, 1/8]$

4.5 La identidad con los marcos de Parseval suaves

(i.e., ϕ es una función de escala de una ondícula de Meyer re-escalada). Para $\xi \in \hat{\mathbb{R}}^2$, sea $\hat{\Phi}(\xi) = \hat{\phi}(\xi_1)\hat{\phi}(\xi_2)$ y $W^2(\xi) = \hat{\Phi}^2(2^{-2}\xi) - \hat{\Phi}^2(\xi)$. Se sigue que

$$\hat{\Phi}(\xi) + \sum_{j \geq 0} W^2(2^{-2j}\xi) = 1, \quad \text{for all } \xi \in \hat{\mathbb{R}}^2.$$

Sea ahora $v \in C^\infty(\mathbb{R})$ tal que $v(0) = 1$, $v^{(n)}(0) = 0$ para todo $n \geq 1$, $\text{supp } v \subset [-1, 1]$ y

$$|v(u-1)|^2 + |v(u)|^2 + |v(u+1)|^2 = 1, \quad |u| \leq 1.$$

Se sigue que, para cualquier $j \geq 0$,

$$\sum_{m=-2^j}^{2^j} |v(2^j u - m)|^2 = 1, \quad |u| \leq 1.$$

Estas funciones tienen propiedades similares a aquellas de la Sección 4.2.3, ver [54].

Con $V_h(\xi_1, \xi_2) = v(\frac{\xi_2}{\xi_1})$, el sistema de las *shearlets* horizontales para $L^2(\mathbb{R}^2)$ se define como la colección numerable de funciones

$$\{\psi_{j,\ell,k}^h : j \geq 0, |\ell| < 2^j, k \in \mathbb{Z}^2\},$$

cuyos elementos se definen por su transformada de Fourier

$$(\psi_{j,\ell,k}^h)^\wedge(\xi) = |\det A_h|^{-j/2} W(2^{-2j}\xi) V_h(\xi A_h^{-j} B_h^{-\ell}) e^{-2\pi i \xi A_h^{-j} B_h^{-\ell} k}, \quad (4.33)$$

donde A_h y B_h son como en la Subsección 4.2.3. Similarmente, uno puede construir el sistema de *shearlets* verticales como en la Subsección 4.2.3. Para las *shearlets en la frontera* sean $j \geq 1$, $\ell = \pm 2^j$ and $k \in \mathbb{Z}^2$, se define

$$(\psi_{j,\ell,k}^h)^\wedge(\xi) = 2^{-\frac{3}{2}j - \frac{1}{2}} W(2^{-2j}\xi) v(2^j \frac{\xi_2}{\xi_1} - \ell) e^{-2\pi i \xi 2^{-1} A_h^{-j} B_h^{-\ell} k}$$

para $\xi \in \mathcal{D}^h$, y

$$(\psi_{j,\ell,k}^v)^\wedge(\xi) = 2^{-\frac{3}{2}j - \frac{1}{2}} W(2^{-2j}\xi) v(2^j \frac{\xi_1}{\xi_2} - \ell) e^{-2\pi i \xi 2^{-1} A_v^{-j} B_v^{-\ell} k}$$

para $\xi \in \mathcal{D}^v$. Para $j = 0$, $\ell = \pm 1$ y $k \in \mathbb{Z}^2$ defínase

$$(\psi_{0,\ell,k}^h)^\wedge(\xi) = W(\xi) v(\frac{\xi_2}{\xi_1} - \ell) e^{-2\pi i \xi k}$$

para $\xi \in \mathcal{D}^h$, y

$$(\psi_{0,\ell,k}^v)^\wedge(\xi) = W(\xi) v(\frac{\xi_1}{\xi_2} - \ell) e^{-2\pi i \xi k}$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

para $\xi \in \mathcal{D}^v$. Estas *shearlets en la frontera* son también $C^\infty(\hat{\mathbb{R}}^2)$ (ver [54]). Este nuevo sistema no es afín ya que la función W no es invariante a la operación *shear*. Sin embargo, como ya se mencionó, generan la misma teselación de la frecuencia.

La nueva condición de marco de Parseval es ahora escrita como (véase la prueba del Teorema 2.3 en [54])

$$\left| \hat{\Phi}(\xi) \right|^2 + \sum_{\mathfrak{d}=1}^2 \sum_{j \geq 0} \sum_{|\ell| < 2^j} \left| \hat{\psi}^{\mathfrak{d}}(\xi A_{\mathfrak{d}}^{-j} B_{\mathfrak{d}}^{-\ell}) \right|^2 + \sum_{j \geq 0} \sum_{\ell = \pm 2^j} \left| \hat{\psi}(\xi A^{-j} B^{-\ell}) \right|^2 = 1, \quad (4.34)$$

para todo $\xi \in \hat{\mathbb{R}}^2$ y donde $\mathfrak{d} = 1, 2$ representa las direcciones horizontal y vertical, y donde omitimos el subíndice para las matrices de las *shearlets en la frontera*. Nótese que ahora no existen las funciones características como en (4.25).

4.5.2 La identidad de reconstrucción en \mathcal{S}'

El objetivo es mostrar que, con los marcos de Parseval suaves del sistema de *shearlets* de Guo y Labate en [54], $T_\psi \circ S_\psi$ es la identidad en \mathcal{S}' y, por tanto, en $\mathbf{F}_p^{\alpha,q}(AB)$. Primero mostramos que cualquier $f \in \mathcal{S}'$ admite una descomposición tipo Littlewood-Paley con dilataciones *shear* anisotrópicas, para lo cual seguimos [13]. Luego, mostramos la identidad de reconstrucción en \mathcal{S}' siguiendo [42]. Denotamos $\hat{\Phi} = (\psi_{-1})^\wedge$.

Lema 4.5.1 *Sea $\{\psi_{j,\ell,k} : j \geq 0, \ell = -2^j, \dots, 2^j, k \in \mathbb{Z}^2\}$ el sistema de *shearlets* suaves que verifica (4.34). Entonces, para cualquier $f \in \mathcal{S}'$,*

$$f = f * \tilde{\psi}_{-1} * \psi_{-1} + \sum_{\mathfrak{d}=1}^2 \sum_{j \geq 0} \sum_{|\ell| < 2^j} f * \tilde{\psi}_{A_{\mathfrak{d}}^{-j} B_{\mathfrak{d}}^{-\ell}}^{\mathfrak{d}} * \psi_{A_{\mathfrak{d}}^{-j} B_{\mathfrak{d}}^{-\ell}}^{\mathfrak{d}} + \sum_{j \geq 0} \sum_{\ell = \pm 2^j} f * \tilde{\psi}_{A^{-j} B^{-\ell}} * \psi_{A^{-j} B^{-\ell}},$$

con convergencia en \mathcal{S}' .

Demostración. Como la transformada de Fourier \mathcal{F} es un isomorfismo de \mathcal{S}' , es suficiente mostrar que

$$\begin{aligned} \hat{f}(\xi) &= \hat{f}(\xi) |(\psi_{-1})^\wedge(\xi)|^2 + \sum_{\mathfrak{d}=1}^2 \sum_{j \geq 0} \sum_{|\ell| < 2^j} \hat{f}(\xi) \left| \hat{\psi}^{\mathfrak{d}}(\xi A_{\mathfrak{d}}^{-j} B_{\mathfrak{d}}^{-\ell}) \right|^2 \\ &+ \sum_{j \geq 0} \sum_{\ell = \pm 2^j} \hat{f}(\xi) \left| \hat{\psi}(\xi A^{-j} B^{-\ell}) \right|^2, \end{aligned}$$

4.5 La identidad con los marcos de Parseval suaves

converge en \mathcal{S}' . Mostraremos esto sólo para aquellas *shearlets* con $j \geq 0$ (ψ_{-1} es de hecho una función de escala de una ondícula de Meyer). Supóngase que \hat{f} tiene orden $\leq m$. Esto es, existen un entero $n \geq 0$ y una constante $C > 0$ tal que

$$\left| \langle \hat{f}, g \rangle \right| \leq C \sup_{|\alpha| \leq n, |\beta| \leq m} \|g\|_{\alpha, \beta}, \quad \text{for all } g \in \mathcal{S},$$

donde $\|g\|_{\alpha, \beta} = \sup_{\xi \in \hat{\mathbb{R}}^2} |\xi^\alpha| |\partial^\beta g(\xi)|$ denota la semi-norma usual en \mathcal{S} para multi-índices α y β . Entonces,

$$\left| \langle \hat{f} |(\psi_{A^{-j}B^{-\ell}})^\wedge|^2, g \rangle \right| = \left| \langle \hat{f}, |(\psi_{A^{-j}B^{-\ell}})^\wedge|^2 g \rangle \right| \leq C \sup_{|\alpha| \leq n, |\beta| \leq m} \left\| |(\psi_{A^{-j}B^{-\ell}})^\wedge|^2 g \right\|_{\alpha, \beta}.$$

Como en el Lema 2.5 en [53], uno puede probar que

$$\sup_{|\beta|=m} \left\| \partial^\beta |(\psi_{A^{-j}B^{-\ell}})^\wedge|^2 \right\|_\infty \leq C 2^{-jm}.$$

Así pues, por las condiciones sobre el soporte de $(\psi_{A^{-j}B^{-\ell}})^\wedge(\xi)$ (ver Subsección 4.2.3)

$$\begin{aligned} & \sup_{|\alpha| \leq n, |\beta| \leq m} \left\| |(\psi_{A^{-j}B^{-\ell}})^\wedge|^2 g \right\|_{\alpha, \beta} \\ & \leq C \sup_{\xi \in \hat{\mathbb{R}}^2} \left[(1 + |\xi|)^n \sup_{|\beta| \leq m} \left| \partial^\beta |(\psi_{A^{-j}B^{-\ell}})^\wedge(\xi)|^2 \right| \sup_{|\beta| \leq m} \left| \partial^\beta g(\xi) \right| \right] \\ & \leq C \sup_{\xi \in \text{supp}(\psi_{A^{-j}B^{-\ell}})^\wedge(\xi)} (1 + |\xi|)^n \sup_{|\beta| \leq m} \left| \partial^\beta g(\xi) \right| \\ & \leq C \sup_{|\alpha| \leq n+1, |\beta| \leq m} \|g\|_{\alpha, \beta} \sup_{\xi \in \text{supp}(\psi_{A^{-j}B^{-\ell}})^\wedge(\xi)} (1 + |\xi|)^{-1} \\ & \leq C \sup_{|\alpha| \leq n+1, |\beta| \leq m} \|g\|_{\alpha, \beta} (1 + 2^{2j-4})^{-1} \leq C 2^{-2j}, \end{aligned}$$

lo que prueba la convergencia en \mathcal{S}' . La igualdad es una consecuencia directa de (4.34). ■

Lema 4.5.2 Sean $g \in \mathcal{S}'$ y $h \in \mathcal{S}$ tales que

$$\text{supp } \hat{g}, \text{supp } \hat{h} \subset [-1/2, 1/2]^2 B^\ell A^j = QB^\ell A^j.$$

Entonces,

$$g * h = \sum_{k \in \mathbb{Z}^2} |\det A|^{-j} g(A^{-j}B^{-\ell}k)h(x - A^{-j}B^{-\ell}k),$$

con convergencia en \mathcal{S}' .

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Demostración. Supóngase primero que $g \in \mathcal{S}$. podemos expresar \hat{g} por su serie de Fourier como

$$\hat{g}(\xi) = \sum_{k \in \mathbb{Z}^2} |\det A|^{-j/2} e^{-2\pi i \xi A^{-j} B^{-\ell} k} \cdot \left(\int_{QB^\ell A^j} \hat{g}(\omega) \cdot |\det A|^{-j/2} e^{2\pi i \omega A^{-j} B^{-\ell} k} d\omega \right).$$

Por la fórmula de inversión de Fourier en $\hat{\mathbb{R}}^2$ se tiene

$$\hat{g}(\xi) = \sum_{k \in \mathbb{Z}^2} |\det A|^{-j} e^{-2\pi i \xi A^{-j} B^{-\ell} k} \cdot g(A^{-j} B^{-\ell} k), \quad \xi \in QB^\ell A^j.$$

ya que \hat{g} tiene soporte compacto, $g(A^{-j} B^{-\ell} k)$ tiene sentido (por el teorema de Paley-Wiener). Como $\text{supp } \hat{h} \subset QB^\ell A^j$, $g * h = (\hat{g} \hat{h})^\vee$,

$$\begin{aligned} g * h &= \sum_{k \in \mathbb{Z}^2} |\det A|^{-j} g(A^{-j} B^{-\ell} k) [e^{-2\pi i \xi A^{-j} B^{-\ell} k} \hat{h}(\cdot)]^\vee \\ &= \sum_{k \in \mathbb{Z}^2} |\det A|^{-j} g(A^{-j} B^{-\ell} k) h(x - A^{-j} B^{-\ell} k), \end{aligned}$$

lo que prueba la convergencia para $g \in \mathcal{S}$. Para quitar esta suposición se usa el mismo argumento de regularización usual como en la prueba del Lema 4.4.8. ■

Teorema 4.5.3 *Sea el sistema de shearlets $\{\psi_{j,\ell,k}\}$ construido como en la Subsección 4.5.1 tal que sea un marco de Parseval suave que verifica (4.34). La composición de los operadores de análisis y síntesis $T_\psi \circ S_\psi$ (ver (4.30) para las definiciones) es la identidad*

$$f = \sum_{Q \in \mathcal{Q}_{AB}} \langle f, \psi_Q \rangle \psi_Q,$$

en \mathcal{S}' .

Demostración. Identifíquese P con (j, ℓ, k) y escríbase $f * \tilde{\psi}_{A^{-j} B^{-\ell}}(A^{-j} B^{-\ell} k) = f * \tilde{\psi}_{A^{-j} B^{-\ell}}(x_P) = |\det A|^{j/2} \langle f, \psi_P \rangle$, como en (4.27). También, como en la Sección 4.3.1, $\psi_{A^{-j} B^{-\ell}}(x - A^{-j} B^{-\ell} k) = |\det A|^{j/2} \psi_P(x)$. Sean $g = f * \tilde{\psi}_{A^{-j} B^{-\ell}}$ y $h = \psi_{A^{-j} B^{-\ell}}$. Por construcción, el soporte de $(\psi_{j,0,k})^\wedge(\xi)$ es un subconjunto de QA^j , esto implica $\text{supp}(\psi_{j,\ell,k})^\wedge(\xi) \subset QB^\ell A^j$. Por tanto, el Lema 4.5.2 da

$$\begin{aligned} f * \tilde{\psi}_{A^{-j} B^{-\ell}} * \psi_{A^{-j} B^{-\ell}} &= \sum_{k \in \mathbb{Z}^2} \langle f, \psi_{j,\ell,k} \rangle \psi_{j,\ell,k} \\ &= \sum_{P \in \mathcal{Q}^{j,\ell}} \langle f, \psi_P \rangle \psi_P. \end{aligned}$$

Sumando apropiadamente sobre $\mathfrak{d} = 1, 2$, $j \geq 0$ y $\ell = -2^j, \dots, 2^j$, el Lema 4.5.1 da el resultado. ■

4.6 $\mathbf{F}_{p_1}^{\alpha_1, q_1}$ y $\mathbf{F}_{p_2}^{\alpha_2, q_2}(AB)$

En esta sección relacionamos los espacios clásicos diádicos y los altamente anisotrópicos al probar inclusiones de los espacios Triebel-Lizorkin clásicos (isotrópicos diádicos) inhomogéneos en los recién definidos espacios Triebel-Lizorkin *shear* anisotrópicos inhomogéneos (y viceversa) para ciertos parámetros de suavidad. También mostramos que hay sucesiones de funciones cuya norma en los espacios *shear* anisotrópicos (resp. diádicos) está acotada por arriba y por abajo pero cuya norma en los espacios diádicos (resp. *shear* anisotrópicos) se desvanece. Sean A y B como en la Subsección 4.2.3.

Empezamos con algunas definiciones concernientes a los espacios clásicos diádicos (ver Sección 12 en [41]). Sean $\varphi, \theta, \Phi, \Theta$ las funciones de análisis y síntesis de la transformada φ de Frazier y Jawerth. Entonces, φ, θ, Φ y Θ satisfacen: 1) $\varphi, \theta, \Phi, \Theta \in \mathcal{S}$, 2) $\text{supp } \hat{\varphi}, \text{supp } \hat{\theta} \subset \{\xi \in \hat{\mathbb{R}}^2 : \frac{1}{2} \leq |\xi| \leq 2\}$ y $\text{supp } \hat{\Phi}, \text{supp } \hat{\Theta} \subset \{\xi \in \hat{\mathbb{R}}^2 : |\xi| \leq 2\}$, 3) $|\hat{\varphi}(\xi)|, |\hat{\theta}(\xi)| \geq c > 0$ si $\frac{3}{5} \leq |\xi| \leq \frac{5}{3}$ y $|\hat{\Phi}(\xi)|, |\hat{\Theta}(\xi)| \geq c > 0$ if $|\xi| \leq \frac{5}{3}$, y 4) $\hat{\Phi}(\xi)\hat{\Theta}(\xi) + \sum_{\nu \in \mathbb{Z}_+} \overline{\hat{\varphi}(2^{-\nu}\xi)}\hat{\theta}(2^{-\nu}\xi) = 1$. Sea \mathcal{D}_+ el conjunto de cubos diádicos con $l(Q) \leq 1$ donde $l(Q)$ es la longitud del lado del cubo Q .

Para $\alpha \in \mathbb{R}$, $0 < q \leq \infty$, $0 < p < \infty$, el espacio (diádico) inhomogéneo de Triebel-Lizorkin $\mathbf{F}_p^{\alpha, q}$ es la colección de todas las $f \in \mathcal{S}'$ tales que (ver el Lema 12.1 en [41] para una discusión sobre los índices de dilatación)

$$\|f\|_{\mathbf{F}_p^{\alpha, q}} = \|\Phi * f\|_{L^p} + \left\| \left(\sum_{\nu=1}^{\infty} (2^{\nu\alpha} |\varphi_\nu * f|)^q \right)^{1/q} \right\|_{L^p} < \infty,$$

donde $\varphi_\nu(x) = 2^{2\nu}\varphi(2^\nu x)$.

Para $\alpha \in \mathbb{R}$, $0 < q \leq \infty$, $0 < p < \infty$, el espacio de sucesiones inhomogéneo de Triebel-Lizorkin $\mathbf{f}_p^{\alpha, q}$ es la colección de todas las sucesiones con valores complejos \mathbf{s}

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

tales que

$$\|\mathbf{s}\|_{\mathbf{F}_p^{\alpha,q}} = \left\| \left(\sum_{Q \in \mathcal{D}_+} (2^{\nu\alpha} |s_Q| \tilde{\chi}_Q)^q \right)^{1/q} \right\|_{L^p} < \infty,$$

donde $\tilde{\chi}_Q(x) = |Q|^{-\frac{1}{2}} \chi_Q(x)$ es la función característica normalizada en L^2 de $Q \in \mathcal{D}_+$.

Sea $\mathbf{s} = \{s_Q\}_Q$, en donde identificamos Q con el par $(\nu, k) \in \mathbb{Z}_+ \times \mathbb{Z}^2$. Para $0 < r \leq \infty$ y $\lambda > 0$, defínase la sucesión $\mathbf{s}_{r,\lambda}^* = \{(s_{r,\lambda}^*)_Q\}_{Q \in \mathcal{D}_+}$ por

$$(s_{r,\lambda}^*)_{Q'} = \left(\sum_{Q: l(Q)=l(Q')} \frac{|s_Q|^r}{(1 + l(Q')^{-1} |x_Q - x_{Q'}|)^\lambda} \right)^{1/r},$$

ver p. 48 de [41] y comparar con la sucesión definida de manera similar en la Subsección 4.4.2.

4.6.1 Las inclusiones

Empezamos con un resultado sobre casi ortogonalidad de funciones bajo las dilataciones *shear* anisotrópicas y diádicas.

Lema 4.6.1 Sean $\psi, \varphi \in \mathcal{S}$. Para $j \geq 0$ y $|\ell| \leq 2^j$,

$$\int_{\mathbb{R}^2} \left| \psi(B^\ell A^j(x-y)) \right| |\varphi(2^{2j}y)| dy \leq \frac{C_N 2^{-3j}}{(1 + 2^j |x|)^N},$$

para cada $N > 2$.

Demostraciones. Como $\psi, \varphi \in \mathcal{S}$,

$$\begin{aligned} & \int_{\mathbb{R}^2} \left| \psi(B^\ell A^j(x-y)) \right| |\varphi(2^{2j}y)| dy \\ & \lesssim \int_{\mathbb{R}^2} \frac{1}{(1 + |B^\ell A^j(x-y)|)^N} \frac{1}{(1 + |2^{2j}y|)^N} dy. \end{aligned}$$

Defínanse los conjuntos

$$E_1 = \{y \in \mathbb{R}^2 : |y| > |x|/2\}$$

$$E_2 = \{y \in \mathbb{R}^2 : |y| \leq |x|/2\}.$$

Si $y \in E_1$, $1 + 2^j |x| \leq 1 + 2^{j+1} |y| \leq 2(1 + 2^{2j} |y|)$. Cuando $y \in E_2$, $\frac{1}{2} |x| \leq |x| - |y| \leq |x - y|$. Entonces, por el Lemma 4.3.2

$$4(1 + |B^\ell A^j(x-y)|) \geq 1 + 4 |B^\ell A^j(x-y)| \geq 1 + 4 \cdot 2^{j-1} |x-y| \geq 1 + 2^j |x|.$$

Así,

$$\begin{aligned} & \int_{\mathbb{R}^2} \left| \psi(B^\ell A^j(x-y)) \right| |\varphi(2^{2j}y)| dy \\ & \lesssim \frac{1}{(1+2^j|x|)^N} \int_{E_1} \frac{1}{(1+|B^\ell A^j(x-y)|)^N} dy \\ & \quad + \frac{1}{(1+2^j|x|)^N} \int_{E_2} \frac{1}{(1+2^{2j}|y|)^N} dy \\ & \lesssim \left[\frac{2^{-3j}}{(1+2^j|x|)^N} + \frac{2^{-4j}}{(1+2^j|x|)^N} \right] \lesssim \frac{2^{-3j}}{(1+2^j|x|)^N}, \end{aligned}$$

para cada $N > 2$. ■

Teorema 4.6.2 *Sea $\alpha_1, \alpha_2 \in \mathbb{R}$, $0 < q \leq \infty$, $0 < p < \infty$ y $\lambda > 2 \max(1, 1/q, 1/p)$. Si $3\alpha_2 + \frac{1}{q} + \lambda < 2\alpha_1$,*

$$\mathbf{F}_p^{\alpha_1, q} \hookrightarrow \mathbf{F}_p^{\alpha_2, q}(AB).$$

Demostración. Para acortar la notación escribimos $\mathbf{F}_1 = \mathbf{F}_p^{\alpha_1, q}$, $\mathbf{f}_1 = \mathbf{f}_p^{\alpha_1, q}$ y $\mathbf{F}_2 = \mathbf{F}_p^{\alpha_2, q}(AB)$. De hecho probaremos que, para $f = \sum_{Q \in \mathcal{D}_+} s_Q \varphi_Q \in \mathbf{F}_1$,

$$\|f\|_{\mathbf{F}_2} \lesssim \|\mathbf{s}_{1, \lambda}^*\|_{\mathbf{f}_1} \lesssim \|\mathbf{s}\|_{\mathbf{f}_1} \lesssim \|f\|_{\mathbf{F}_1},$$

donde, por supuesto, la desigualdad que nos interesa probar es la primera y las últimas dos se prueban en [41], en donde la segunda desigualdad requiere la condición $\lambda > 2 \max\{1, 1/q, 1/p\}$. De las condiciones del soporte compacto de $(\varphi_{\nu, k})^\wedge$ y $(\tilde{\psi}_{A^{-j}B^{-\ell}})^\wedge$ y sus expansiones diádicas y *shear* anisotrópicas, respectivamente, se tiene formalmente

$$\tilde{\psi}_{A^{-j}B^{-\ell}} * f = \sum_{\nu=2j-5}^{2j} \sum_{k \in \mathbb{Z}^2} s_{\nu, k} \tilde{\psi}_{A^{-j}B^{-\ell}} * \varphi_{\nu, k}.$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

Por lo tanto, escribiendo $\tilde{\psi}_{A^{-j}B^{-\ell}}(x) = |\det A|^j \tilde{\psi}(B^\ell A^j x)$ y $\varphi_{\nu,k}(x) = 2^\nu \varphi(2^\nu x - k)$, el Lema 4.6.1 da

$$\begin{aligned} \|f\|_{\mathbf{F}_2} &= \left\| \left(\sum_{j \geq 0} 2^{3j\alpha_2 q} \sum_{|\ell| \leq 2^j} \left| \sum_{\nu=2^j-5}^{2^j} \sum_{k \in \mathbb{Z}^2} s_{\nu,k} \tilde{\psi}_{A^{-j}B^{-\ell}} * \varphi_{\nu,k}(\cdot) \right|^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{j \geq 0} 2^{3j\alpha_2 q} \sum_{|\ell| \leq 2^j} \left| \sum_{k \in \mathbb{Z}^2} s_{2^j,k} \tilde{\psi}_{A^{-j}B^{-\ell}} * \varphi_{2^j,k}(\cdot) \right|^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{j \geq 0} 2^{3j\alpha_2 q} \sum_{|\ell| \leq 2^j} \left[\sum_{k \in \mathbb{Z}^2} |s_{2^j,k}| \frac{2^{2j}}{(1+2^j|\cdot+2^{-2j}k|)^N} \right]^q \right)^{1/q} \right\|_{L^p}, \end{aligned}$$

para cada $N > 2$. Sea $\lambda > 2 \max(1, 1/q, 1/p)$. Siguiendo la prueba de la segunda parte del Teorema 4.4.11, si $x \in Q$ y $Q \in \mathcal{D}^{2^j}$,

$$\begin{aligned} &\sum_{k \in \mathbb{Z}^2} \frac{|s_{2^j,k}| 2^{2j}}{(1+2^j|x-2^{-2j}k|)^\lambda} \\ &= \sum_{k \in \mathbb{Z}^2} \frac{|s_{2^j,k}| 2^{2j} \cdot 2^{j\lambda}}{2^{j\lambda}(1+2^j|x-2^{-2j}k|)^\lambda} \leq 2^{j\lambda} \sum_{k \in \mathbb{Z}^2} \frac{|s_{2^j,k}| 2^{2j}}{(1+2^{2j}|x-2^{-2j}k|)^\lambda} \\ &\lesssim 2^{j\lambda} \sum_{Q \in \mathcal{D}^{2^j}} |Q|^{-\frac{1}{2}} |(s_{1,\lambda}^* Q)| \chi_Q(x) = 2^{j\lambda} \sum_{Q \in \mathcal{D}^{2^j}} |(s_{1,\lambda}^* Q)| \tilde{\chi}_Q(x), \end{aligned}$$

ya que \mathcal{D}^{2^j} es una partici3n de \mathbb{R}^2 . As3,

$$\begin{aligned} \|f\|_{\mathbf{F}_2} &\lesssim \left\| \left(\sum_{j \geq 0} 2^{3j\alpha_2 q} (2^{j+1} + 1) [2^{j\lambda} \sum_{Q \in \mathcal{D}^{2^j}} |(s_{1,\lambda}^* Q)| \tilde{\chi}_Q(\cdot)]^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{j \geq 0} \sum_{Q \in \mathcal{D}^{2^j}} [2^{3j\alpha_2 + \frac{j}{q} + j\lambda} |(s_{1,\lambda}^* Q)| \tilde{\chi}_Q(\cdot)]^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{j \geq 0} \sum_{Q \in \mathcal{D}^j} [2^{3j\alpha_2/2 + \frac{j}{2q} + j\lambda/2} |(s_{1,\lambda}^* Q)| \tilde{\chi}_Q(\cdot)]^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{Q \in \mathcal{D}_+} [|Q|^{-\frac{\alpha_1}{2}} |(s_{1,\lambda}^* Q)| \tilde{\chi}_Q(\cdot)]^q \right)^{1/q} \right\|_{L^p} = \|s_{1,\lambda}^*\|_{\mathbf{f}_1}, \end{aligned}$$

ya que $2^{3j\alpha_2/2 + \frac{j}{2q} + j\lambda/2} \leq 2^{j\alpha_1} = |Q|^{-\frac{\alpha_1}{2}}$ para un $Q \in \mathcal{D}^j$. Siguiendo la prueba del Lema 2.3 en [41] se obtiene $\|s_{1,\lambda}^*\|_{\mathbf{f}_1} \lesssim \|s\|_{\mathbf{f}_1}$ (con la restricci3n sobre λ), y del Teorema

2.2 en [41], $\|s\|_{\mathbf{f}_1} \lesssim \|f\|_{\mathbf{F}_1}$, lo que finaliza la prueba. \blacksquare

Teorema 4.6.3 Sean $\alpha_1, \alpha_2 \in \mathbb{R}$, $0 < q \leq \infty$, y $0 < p < \infty$. Si $2\alpha_1 + 2 - \min(1, 1/q) \leq 3\alpha_2$,

$$\mathbf{F}_p^{\alpha_2, q}(AB) \hookrightarrow \mathbf{F}_p^{\alpha_1, q}.$$

Demostración. Para acortar la notación escribáse $\mathbf{F}_1 = \mathbf{F}_p^{\alpha_1, q}$, $\mathbf{F}_2 = \mathbf{F}_p^{\alpha_2, q}(AB)$ y $\mathbf{f}_2 = \mathbf{f}_p^{\alpha_2, q}$. Supóngase $f = \sum_{P \in Q_{AB}} s_P \psi_P \in \mathbf{F}_2$ y sea $\lambda > 3 \max(1, 1/q, 1/p)$. De las condiciones de soporte compacto de $(\varphi_{2^\nu})^\wedge$ y $(\psi_{j, \ell, k})^\wedge$, $j \sim \lfloor \nu/2 \rfloor$. Así, escribiendo $\psi_{\nu, \ell, k}(x) = |\det A|^{\nu/2} \psi(B^\ell A^\nu x - k)$ y $\varphi_{2^{2\nu}}(x) = 2^{4\nu} \varphi(2^{2\nu}x)$,

$$\begin{aligned} \|f\|_{\mathbf{F}_1} &= \left\| \left(\sum_{\nu \geq 0} (|Q_\nu|^{-\alpha_1/2} |\varphi_{2^\nu} * f(\cdot)|)^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{\nu \geq 0} 2^{\nu \alpha_1 q} \left| \sum_{|\ell| \leq 2^{\lfloor \nu/2 \rfloor}} \sum_{k \in \mathbb{Z}^2} s_{\lfloor \nu/2 \rfloor, \ell, k} \cdot \varphi_{2^\nu} * \psi_{\lfloor \nu/2 \rfloor, \ell, k}(\cdot) \right|^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{\nu \geq 0} 2^{2\nu \alpha_1 q} \left| \sum_{|\ell| \leq 2^\nu} \sum_{k \in \mathbb{Z}^2} s_{\nu, \ell, k} \cdot \varphi_{2^{2\nu}} * \psi_{\nu, \ell, k}(\cdot) \right|^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{\nu \geq 0} 2^{2\nu \alpha_1 q} \left(\sum_{|\ell| \leq 2^\nu} \sum_{k \in \mathbb{Z}^2} |s_{\nu, \ell, k}| \cdot \frac{2^{-3\nu} \cdot 2^{4\nu} \cdot 2^{3\nu/2}}{(1 + 2^\nu |\cdot - A^{-\nu} B^{-\ell} k|)^N} \right)^q \right)^{1/q} \right\|_{L^p}, \end{aligned}$$

para cada $N > 2$, por el Lema 4.6.1. Continuando como en la segunda parte de la demostración de Teorema 4.4.11,

$$\begin{aligned} \|f\|_{\mathbf{F}_1} &\lesssim \left\| \left(\sum_{\nu \geq 0} 2^{2\nu \alpha_1 q + \nu q} \left(\sum_{|\ell| \leq 2^\nu} \sum_{P \in Q^{\nu, \ell}} |s_P| \cdot \frac{|P|^{-1/2}}{(1 + 2^\nu |\cdot - x_P|)^N} \right)^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{\nu \geq 0} 2^{\nu q(2\alpha_1 + 1)} \left[\sum_{|\ell| \leq 2^\nu} \sum_{P \in Q^{\nu, \ell}} (s_{1, N}^*)_P \tilde{\chi}_P(\cdot) \right]^q \right)^{1/q} \right\|_{L^p}. \end{aligned}$$

En este punto no se puede usar el argumento de “partición de \mathbb{R}^2 ” ya que $\sum_{|\ell| \leq 2^\nu} \sum_{P \in Q^{\nu, \ell}} \tilde{\chi}_P(x)$ no es una partición de \mathbb{R}^2 . Así pues, si $0 < q \leq 1$ usamos la desigualdad

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

q -triangular $|a + b|^q \leq |a|^q + |b|^q$ o la desigualdad de Hölder si $1 < q$ para obtener

$$\begin{aligned} \|f\|_{\mathbf{F}_1} &\lesssim \left\| \left(\sum_{P \in \mathcal{Q}_{AB}} [2^{\nu(2\alpha_1 + 2 - \min(1, 1/q))} |(s_{1,\lambda}^*)_{P}| \tilde{\chi}_P(\cdot)]^q \right)^{1/q} \right\|_{L^p} \\ &\lesssim \left\| \left(\sum_{P \in \mathcal{Q}_{AB}} [|P|^{-\alpha_2} |(s_{1,\lambda}^*)_{P}| \tilde{\chi}_P(\cdot)]^q \right)^{1/q} \right\|_{L^p} = \|s_{1,\lambda}^*\|_{\mathbf{f}_2}. \end{aligned}$$

Por el Lema 4.4.9 y el Teorema 4.4.11 la prueba está completa. ■

4.6.2 Más relaciones

Un cubo diádico en la escala ν se denotará por Q_ν y un cuadrilátero anisotrópico en la escala j se denotará por P_j .

Teorema 4.6.4 Sean $\alpha_1, \alpha_2 \in \mathbb{R}$, $0 < q_1, q_2 \leq \infty$ y $0 < p_1, p_2 < \infty$. Entonces, existen sucesiones de funciones $\{f^{(j)}\}_{j \geq 0}$ tales que $\|f^{(j)}\|_{\mathbf{F}_{p_2}^{\alpha_2, q_2}(AB)} \approx 1$, pero que $\|f^{(j)}\|_{\mathbf{F}_{p_1}^{\alpha_1, q_1}} \rightarrow 0$, $j \rightarrow \infty$, cuando $3(\alpha_2 - 1/p_2) > 2(\alpha_1 - 1/p_1) + 1$.

Demostración. Para $j \geq 0$, sea $P_j = P_{j,0,0}$ y definir $\mathbf{s}^{(j)} = \{s_Q^{(j)}\}_{Q \in \mathcal{Q}_{AB}}$ de manera que

$$s_Q^{(j)} = \begin{cases} 0 & \text{si } Q \neq P_j \\ |P_j|^{\alpha_2 - \frac{1}{p_2} + \frac{1}{2}} & \text{si } Q = P_j \end{cases}.$$

Entonces, $\|\mathbf{s}^{(j)}\|_{\mathbf{F}_{p_2}^{\alpha_2, q_2}} = 1$, para todo $j \geq 0$. Así, $f^{(j)}(x) = \sum_{Q \in \mathcal{Q}_{AB}} s_Q^{(j)} \psi_Q(x) = |P_j|^{\alpha_2 - \frac{1}{p_2} + \frac{1}{2}} \psi_{j,0,0}(x) \in \mathbf{F}_{p_2}^{\alpha_2, q_2}(AB)$ con $\|f^{(j)}\|_{\mathbf{F}_{p_2}^{\alpha_2, q_2}} \approx 1$. De las condiciones sobre el soporte compacto de $\hat{\varphi}$ y $\hat{\psi}$ se tiene que el soporte de $(\psi_{j,0,0})^\wedge$ se solapa con el soporte de $(\varphi_{\nu,0})^\wedge$ sólo cuando $2j - 5 \leq \nu < 2j$. Así, como $|f^{(j)} * \varphi_{2^{2j}}| = 2^{2j} |f^{(j)} * \varphi_{\nu,0}|$ y suponiendo $\nu = 2j$, el Lema 4.6.1 da

$$\begin{aligned} |f^{(j)} * \varphi_{2^\nu}(x)| &= \left| \int_{\mathbb{R}^2} |P_j|^{\alpha_2 - \frac{1}{p_2} + \frac{1}{2}} |\det A|^{j/2} \psi(A^j(x-y)) 2^{4j} \varphi(2^{2j}y) dy \right| \\ &\leq 2^{-3j(\alpha_2 - \frac{1}{p_2} + \frac{1}{2}) + \frac{3j}{2} + 4j} \int_{\mathbb{R}^2} |\psi(A^j(x-y))| |\varphi(2^{2j}y)| dy \\ &\lesssim \frac{2^{-3j(\alpha_2 - \frac{1}{p_2}) + j}}{(1 + 2^j|x|)^N}, \end{aligned}$$

para todo $N > 2$. Entonces, para N tal que $Np_1 > 2$, se tiene

$$\begin{aligned} \|f^{(j)}\|_{\mathbf{F}_{p_1}^{\alpha_1, q_1}} &= \left\| \left(\sum_{\nu=2j-5}^{2j-1} [2^{\nu\alpha_1} |f^{(j)} * \varphi_{2^\nu}|]^{q_1} \right)^{1/q_1} \right\|_{L^{p_1}} \\ &\leq C_{N, q_1} \left(\int_{\mathbb{R}^2} 2^{2j\alpha_1 p_1} \cdot \frac{[2^{-3j(\alpha_2 - \frac{1}{p_2}) + j}]^{p_1}}{(1 + 2^j |x|)^{Np_1}} dx \right)^{1/p_1} \\ &= C_{N, q_1} 2^{2j\alpha_1 - 3j(\alpha_2 - \frac{1}{p_2}) + j - \frac{2j}{p_1}}, \end{aligned}$$

que tiende a 0 conforme $j \rightarrow \infty$ si $2(\alpha_1 - \frac{1}{p_1}) + 1 < 3(\alpha_2 - \frac{1}{p_2})$. ■

Teorema 4.6.5 Sean $\alpha_1, \alpha_2 \in \mathbb{R}$, $0 < q_1, q_2 \leq \infty$ y $0 < p_1, p_2 < \infty$. Entonces, existen sucesiones de funciones $\{f^{(\nu)}\}_{\nu \geq 0}$ tales que $\|f^{(\nu)}\|_{\mathbf{F}_{p_1}^{\alpha_1, q_1}} \approx 1$, pero que $\|f^{(\nu)}\|_{\mathbf{F}_{p_2}^{\alpha_2, q_2}(AB)} \rightarrow 0$, $\nu \rightarrow \infty$, cuando $2\alpha_1 - 4/p_1 > 3\alpha_2 + 1/q_2 - 2/p_2$.

Demostración. Para $\nu \geq 0$, sea $Q_\nu = Q_{\nu,0} \in \mathcal{D}_+$ y definir $\mathbf{s}^{(\nu)} = \{s_Q^{(\nu)}\}_{Q \in \mathcal{D}_+}$ de manera que

$$s_Q^{(\nu)} = \begin{cases} 0 & \text{si } Q \neq Q_\nu \\ |Q_\nu|^{\frac{\alpha_1}{2} - \frac{1}{p_1} + \frac{1}{2}} & \text{si } Q = Q_\nu \end{cases}.$$

Entonces $\|\mathbf{s}^{(\nu)}\|_{\mathbf{F}_{p_1}^{\alpha_1, q_1}} = 1$, para todo $\nu \geq 0$. Así, $f^{(\nu)}(x) = \sum_{Q \in \mathcal{D}_+} s_Q^{(\nu)} \varphi_Q(x) = |Q_\nu|^{\frac{\alpha_1}{2} - \frac{1}{p_1} + \frac{1}{2}} \varphi_{\nu,0}(x) \in \mathbf{F}_{p_1}^{\alpha_1, q_1}$ con $\|f^{(\nu)}\|_{\mathbf{F}_{p_1}^{\alpha_1, q_1}} \approx 1$, para todo $\nu \geq 0$. Con los mismos argumentos sobre las condiciones del soporte de $(\varphi_{\nu,0})^\wedge(\xi)$ y $(\psi_{j,\ell,0})^\wedge(\xi)$ podemos asumir $\nu = 2j$ y se obtiene

$$\begin{aligned} \left| f^{(2j)} * \psi_{A^{-j}B^{-\ell}}(x) \right| &= \left| \int_{\mathbb{R}^2} s_{2j,0} 2^{2j} \varphi(2^{2j}y) |\det A|^j \psi(B^\ell A^j(x-y)) dy \right| \\ &\leq 2^{-4j(\frac{\alpha_1}{2} - \frac{1}{p_1} + \frac{1}{2}) + 2j + 3j} \int_{\mathbb{R}^2} \left| \psi(B^\ell A^j(x-y)) \right| |\varphi(2^{2j}y)| dy \\ &\lesssim \frac{2^{-4j(\frac{\alpha_1}{2} - \frac{1}{p_1})}}{(1 + 2^j |x|)^N}, \end{aligned}$$

4. SHEARLETS Y NUEVOS ESPACIOS DE DISTRIBUCIONES

para cada $N > 2$, por el Lema 4.6.1. Entonces, para $N > 2$ tal que $Np_2 > 2$,

$$\begin{aligned}
\|f^{(2j)}\|_{\mathbf{F}_2} &\lesssim \left\| \left(\sum_{|\ell| \leq 2^j} [|P_j|^{-\alpha_2} |f^{2j} * \psi_{A^{-j}B^{-\ell}}|]^{q_2} \right)^{1/q_2} \right\|_{L^{p_2}} \\
&\lesssim \left(\int_{\mathbb{R}^2} \left(\sum_{|\ell| \leq 2^j} [2^{3j\alpha_2} \frac{2^{-4j(\frac{\alpha_1}{2} - \frac{1}{p_1})}}{(1+2^j|x|)^N}]^{q_2} \right)^{p_2/q_2} dx \right)^{1/p_2} \\
&= \left(\int_{\mathbb{R}^2} ((2^{j+1} + 1) [\frac{2^{j(3\alpha_2 - 4(\frac{\alpha_1}{2} - \frac{1}{p_1}))}}{(1+2^j|x|)^N}]^{q_2})^{p_2/q_2} dx \right)^{1/p_2} \\
&\lesssim 2^{j(3\alpha_2 - 4(\frac{\alpha_1}{2} - \frac{1}{p_1}) + \frac{1}{q_2})} \left(\int_{\mathbb{R}^2} \frac{1}{(1+2^j|x|)^{Np_2}} dx \right)^{1/p_2} \\
&\lesssim 2^{j(3\alpha_2 - 4(\frac{\alpha_1}{2} - \frac{1}{p_1}) + \frac{1}{q_2})} \cdot 2^{-\frac{2j}{p_2}},
\end{aligned}$$

que tiende a 0, cuando $j \rightarrow \infty$, si $3\alpha_2 - 4(\frac{\alpha_1}{2} - \frac{1}{p_1}) + \frac{1}{q_2} - \frac{2}{p_2} < 0$. ■

4.7 Pesos

Para extender este trabajo al caso de pesos $w \in A_\infty = \cup_{p>1} A_p$, se pueden seguir [14] y [13]. Los espacios $\mathbf{F}_p^{\alpha,q}(AB)$ y $\mathbf{f}_p^{\alpha,q}(AB)$ se definen con las cuasi normas $L^p(w)$. Para la versión con pesos de la desigualdad Fefferman-Stein referimos al lector a [3] o a la Nota 6.5 del Capítulo V en [44]. Se añade $w \in A_{p_0}$ al enunciado del Lema 4.4.9 y N se escoge tal que $N > 3 \max(1, r/q, rp_0/p)$. Para el Teorema 4.4.11 uno añade $w \in A_\infty$. Para su prueba, λ debe escogerse tal que $0 < \lambda < \min(p/p_0, q)$ y N tal que $N > 3 \max(1, 1/q, p_0/p)$.

Referencias

- [1] D. Achlioptas, *Database-friendly Random Projections: Johnson-Lindenstrauss with Binary Coins*, Journal of Comp. & Sys. Sci., 66 (4), (2003), p.671-687, special issue of invited papers from PODS'01. [9](#), [15](#), [24](#), [50](#)
- [2] A. Almeida, *Wavelet bases in generalized Besov spaces*, J. Math. Anal. Appl., 304:198-211, 2005. [98](#)
- [3] K. F. Andersen, R. T. John, *Weighted inequalities for vector-valued maximal functions and singular integrals*, Studia Math. **69**, 19-31, 1980. [186](#)
- [4] J. P. Antoine, R. Murenzi and P. Vandergheynst, *Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns*, Appl. Comput. harmon. Anal., 6, pp.314-345, 1999. [143](#)
- [5] R. H. Bamberger and M. J. T. Smith, *A filter bank for directional decomposition of images: Theory and design*, IEEE Trans. Signal Process., 40, pp.882-893, 1992. [143](#)
- [6] R. Baraniuk, M. Davenport, R. DeVore and M. Wakin, *A Simple Proof of the Restricted Isometry Property for Random Matrices*, Constructive Approximation, vol. 28, no. 3, pp. 253-263, 2008. [9](#), [24](#), [50](#), [53](#)
- [7] A.R. Barron, *Universal approximation bounds for superposition of n sigmoidal functions*, IEEE Trans. Inform. Theory, vol. 39, pp. 930-945, 1993. [9](#)
- [8] C. Bennett and R. C. Sharpley, *Interpolation of Operators*, Academic Press, 1988. [84](#), [85](#), [90](#), [107](#), [108](#), [112](#)

REFERENCIAS

- [9] J. Bergh and J. Löfström, *Interpolation Spaces: an Introduction*, Springer-Verlag, 1976. [85](#)
- [10] S. N. Bernstein, *Sur l'ordre de meilleure approximation des fonctions continues par des polynomes*, Académie Royale de Belgique, Classe des Sciences, Mémoires Colletion in 4°, ser. II, 13, 49-194, 1912. [68](#)
- [11] T. Blumensath, M. E. Davies, *Gradient Pursuits*, IEEE Transactions on Signal Processing, vol. 56, no. 6, pp. 2370-2382, 2008. [30](#), [31](#), [32](#), [42](#), [44](#)
- [12] T. Blumensath, M. E. Davies, *Stagewise Weak Gradient Pursuits*, IEEE Transactions on Signal Processing, vol. 57, no. 11, pp. 4333-4346, 2009. [30](#), [31](#), [32](#)
- [13] M. Bownik and K-P. Ho *Atomic and molecular decompositions of anisotropic Triebel-Lizorkin spaces*, Trans. Am. Math. Soc., 358(4):1469-1510, 2005. [147](#), [148](#), [149](#), [170](#), [176](#), [186](#)
- [14] H.-Q. Bui, *Weighted Besov and Triebel-Lizorkin spaces: Interpolation by the real method*, Hiroshima Math. J., 12, 581-605, 1982. [103](#), [106](#), [147](#), [186](#)
- [15] H.-Q. Bui, Characterization of weighted Besov and Triebel-Lizorkin spaces via temperatures, *J. Funct. Anal.*, **55** (1984), 39-62. [147](#)
- [16] E. J. Candès and D. L. Donoho, *Curvelets - A surprising effective nonadaptive representation for objects with edges*, in Curves and Surfaces, C. Rabut, A. Cohen, and L. L. Schumaker, eds., Vanderbilt University Press, Nashville, TN, 2000. [143](#)
- [17] E. Candès, J. Romberg and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. on Information Theory, 52(2) pp. 489-509, 2006. [8](#), [19](#)
- [18] E. Candès, J. Romberg and T. Tao, *Stable signal recovery from incomplete and Inaccurate measurements*, Comm. Pure Appl. Math. 59, 1207-1223, 2006. [19](#), [21](#)
- [19] E. Candès and T. Tao, *Near optimal signal recovery from random projections: Universal encoding strategies?*, IEEE Trans. on Information Theory, 52(12), pp. 5406 - 5425, 2006. [8](#), [18](#), [19](#), [21](#), [41](#)

- [20] E. Candès and T. Tao, *Decoding by linear programming*, IEEE Trans. on Information Theory, 51(12), pp. 4203 - 4215, 2005. [9](#), [19](#), [20](#), [21](#), [26](#), [32](#), [41](#)
- [21] M. Carro, J. Raposo and J. Soria, *Recent developments in the theory of Lorentz spaces and weighted inequalities*, Memoirs Amer. Math. Soc., 107, 2007. [86](#)
- [22] A. Cohen, W. Dahmen and R. DeVore, *Compressed sensing and best k-term approximation*, Journal of the AMS, 22, 211-231, 2009. [16](#)
- [23] A. Cohen, W. Dahmen and R. DeVore, *Instance Optimal Decoding by Thresholding in Compressed Sensing*, Proceedings of the 8th International Conference in Harmonic Analysis and Partial Differential Equations, El Escorial, Contemporary Mathematics, AMS, 2009. [32](#), [39](#), [41](#)
- [24] A. Cohen, R. A. DeVore and R. Hochmuth, *Restricted Nonlinear Approximation*, Constr. Approx., 16:85-113, 2000. [75](#), [76](#), [77](#), [78](#), [79](#), [82](#), [87](#), [91](#), [92](#), [98](#), [99](#), [104](#), [105](#), [107](#), [118](#), [120](#), [123](#)
- [25] M. A. Davenport and M. B. Wakin, *Analysis of Orthogonal Matching Pursuit using the Restricted Isometry Property*, IEEE Transactions on Information Theory, vol. 56, no. 9, pp. 4395-4401, 2010. [37](#)
- [26] W. Dai, O. Milenkovic, *Subspace pursuit for compressive sensing: closing the gap between performance and complexity*, IEEE Trans, Inform, Theory, vol. 55, no 5, pp. 2230-2249, 2009. [37](#)
- [27] S. Dasgupta and A. Gupta, *An elementary proof of the Johnson-Lindenstrauss lemma*, ICSI Technical Report TR-99-006, U.C. Berkeley, 1999. [9](#), [15](#), [24](#)
- [28] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Maths., vol. 61. Philadelphia: SIAM, 1992. [79](#), [97](#)
- [29] R. A. DeVore, *Nonlinear approximation*, Acta Numerica, 51-150, 1998. [71](#), [82](#)
- [30] R. A. DeVore, *Deterministic construction of compressed sensing matrices*, Journal of Complexity, 23:918-925, 2007. [9](#), [50](#)
- [31] R. DeVore, B. Jawerth and V.A. Popov, *Compression of wavelet decompositions*, Amer. J. Math., 114:737-785, 1992. [71](#), [75](#)

REFERENCIAS

- [32] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, 1993. [82](#), [85](#), [91](#)
- [33] R. DeVore, and V.A. Popov, *Interpolation spaces and nonlinear approximation*, en Function Spaces and Applications (M Cwikel, J. Peetre, Y. Sagher, H. Wallin, eds.), Vol. 1302 de *Lecture Notes in Mathematics*, Springer, Berlin, 191-205, 1988. [70](#), [71](#), [75](#), [77](#), [91](#), [105](#), [106](#)
- [34] R. A. DeVore and V. N. Temlyakov, *Some remarks on greedy algorithms*, Adv. Comput. Math., 5, (2-3), 173-187, 1996. [9](#), [28](#), [73](#)
- [35] M. N. Do and M. Vetterli, , *The contourlet transform: An efficient directional multiresolution image representation*, IEEE Trans. Image Process., 14 , pp. 2091-2106, 2005. [143](#)
- [36] D.L. Donoho, *Compressed sensing*, IEEE Trans. on Information Theory, 52(4), pp. 1289 - 1306, April 2006. [8](#), [9](#), [16](#), [18](#)
- [37] D.L. Donoho, Y. Tsaig, I. Drori and J.-L. Starck, *Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit*, Reporte técnico, 2006. [39](#)
- [38] C. Fefferman and E. M. Stein, *Some maximal inequalities*, Amer. J. Math., 93, 107-115, 1971. [163](#)
- [39] M. Frazier, B. Jawerth, *Decomposition of Besov spaces*, Indiana Univ. Math. J., 34 (1985) 777-799. [79](#), [147](#)
- [40] M. Frazier and B. Jawerth, *The φ transform and applications to distribution spaces*, in: M. Cwikel, J. Peetre, Y. Sagher and H. E. Wallin, (Eds.), Function Spaces and Applications, Lecture Notes in Math., vol. 1302, Springer, Berlin, 1988, 223-246. [103](#)
- [41] M. Frazier and B. Jawerth, *A discrete transform and decomposition of distribution spaces*, J. Func. Anal., 93:34-170, 1990. [4](#), [74](#), [79](#), [97](#), [147](#), [148](#), [149](#), [155](#), [161](#), [162](#), [165](#), [174](#), [179](#), [180](#), [181](#), [182](#), [183](#)

- [42] M. Frazier, B. Jawerth and G. Weiss, *Littlewood-Paley theory and the study of function spaces*, CBMS Regional Conference Series in Mathematics, 79, 1991. [97](#), [147](#), [176](#)
- [43] J. García-Cuerva y J. M. Martell, *Wavelet characterization of weighted spaces*, J. Geom. Anal., 11 No. 2 (2001), 241-262. [109](#)
- [44] J. García-Cuerva, J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland (1985). [186](#)
- [45] A. Y. Garnaev, E. D. Gluskin, *On widths of the Euclidean ball*, Soviet Mathematics - Doklady, 30, 200-203, 1984. [18](#)
- [46] G. Garrigós and E. Hernández, *Sharp Jackson and Bernstein Inequalities for N -term Approximation in Sequence Spaces with Applications*, Ind. Univ. Math. Journal, 53, 6:1741-1764, 2004. [74](#), [75](#), [76](#), [79](#), [88](#), [98](#), [99](#), [104](#), [105](#), [106](#)
- [47] G. Garrigós, E. Hernández and J. M. Martell, *Wavelets, Orlicz spaces, and greedy bases*, Appl. Comput. Harmon. Anal., 24, 70-93, 2008. [73](#), [99](#), [108](#), [110](#)
- [48] G. Garrigós, E. Hernández and M. Natividade, *Democracy Functions and Optimal Embeddings for Approximations Spaces*, to appear in Adv. Comput. Math., 2011. [73](#), [87](#), [91](#), [92](#), [96](#)
- [49] R. Gribonval and M. Nielsen, *Some remarks on non-linear approximation with Schauder bases*, East. J. Approximation, 7(2), 1-19, 2001. [73](#), [75](#)
- [50] R. Gribonval and P. Vandergheynst, *On the exponential convergence of Matching Pursuits in Quasi-Incoherent Dictionaries*, IEEE Trans. Information Theory, vol. 52, no.1, pp 255-261, 2006. [37](#), [46](#)
- [51] G. H. Golub, F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 3rd. ed., 1996. [31](#)
- [52] K. Guo, G. Kutyniok and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators*, in: *Wavelets and Splines*, G. Chen and M. Lai (eds.), Nashboro Press, Nashville, TN (2006), pp. 189-201. [144](#)

REFERENCIAS

- [53] K. Guo, and D. Labate, *Optimally sparse multidimensional representation using shearlets*, SIAM J. Math. Anal. (2007), 39:298-318. [4](#), [144](#), [154](#), [177](#)
- [54] K. Guo, and D. Labate, *The construction of smooth Parseval frames of shearlets*, preprint, 2011. [147](#), [174](#), [175](#), [176](#)
- [55] Guo, K., W.-Q. Lim, D. Labate, G. Weiss and E. Wilson, *Wavelets with composite dilation and their MRA properties*, Appl. Comput. Harmon. Anal. (2006), 20:202-236. [4](#), [143](#), [145](#), [149](#), [151](#), [152](#), [153](#)
- [56] E. Hernández, J.M. Martell and M. Natividade, *Quantifying Democracy of Wavelet bases in Lorentz Spaces*, Constr. Approx., Vol. 33, 1:1-14, 2011. [73](#)
- [57] E. Hernández, and D. Vera, *Greedy type algorithms for RIP matrices. A study of two selection rules*, submitted, 2012. [2](#)
- [58] E. Hernández, and D. Vera, *Restricted non-linear approximation in sequence spaces and applications to wavelet bases and interpolation*, Monatsh. Math., 2012, (DOI) 10.1007/s00605-012-0425-6. [3](#)
- [59] E. Hernández, and G. Weiss, *A first course on wavelets*, CRC Press, Boca Raton FL, 1996. [71](#), [79](#), [97](#), [98](#), [109](#), [158](#)
- [60] C. Hsiao, B. Jawerth, B. J. Lucier, and X. M. Yu, *Near optimal compression of orthonormal wavelet expansions*, en Wavelets: Mathematics and Applications, J. J. Benedetto, M. W. Frazier (eds.), Stud. Adv. Math., CRC, Boca Raton, FL, 425-446, 1994. [73](#), [99](#)
- [61] R. Hochmuth, *Anisotropic wavelet bases and thresholding*, Math. Nachr., 280, No. 5-6, 523-533, 2007. [98](#), [105](#)
- [62] L. Hörmander, *The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis*, Second Ed., Springer-Verlag, Berlin, 1990. [164](#)
- [63] D. Jackson, *On the approximation by trigonometric sums and polynomials*, TAMS, 13:491-515, 1912. [68](#)
- [64] L. Jones, *On a conjecture of Huber concerning the convergence of projection pursuit regression*, The Annals of Statistics, vol. 15, pp. 880-882, 1987. [9](#), [28](#)

- [65] B. S. Kashin, *Diameters of certain finite-dimensional sets in classes of smooth functions*, Izv. Akad. Nauk SSSR, Ser. Mat. 41 (2) 334-351, 1977. [18](#)
- [66] G. Kerkyacharian, and D. Picard, *Entropy, universal coding, approximation and bases properties*, Constr. Approx., 20, 1-37, 2004. [75](#), [87](#), [92](#), [118](#)
- [67] G. Kerkyacharian and D. Picard, *Nonlinear Approximation and Muckenhoupt Weights*, Constr. Approx., 24:123-156, 2006. [77](#), [78](#), [79](#), [82](#), [87](#), [88](#), [91](#), [92](#), [96](#), [103](#), [104](#), [120](#), [123](#)
- [68] S. V. Konyagin, and V. N. Temlyakov, *A remark on greedy approximation in Banach spaces*, East J. Approx., 5:365-379, 1999. [72](#)
- [69] S. Krein, J. Petunin and E. Semenov , *Interpolation of Linear Operators*, Translations Math. Monographs, vol. 55, Amer. Math. Soc., Providence, RI, 1992. [84](#), [86](#), [119](#), [120](#)
- [70] G. Kyriazis, *Multilevel characterization of anisotropic function spaces*, SIAM J. Math. Anal. 36, (2004), 441-462. [79](#), [98](#), [109](#)
- [71] D. Labate and G. Weiss, *Continuous and discrete reproducing systems that arise from translations. Theory and applications of composite wavelets*, pp. 48, in: Four Short Courses on Harmonic Analysis, Birkhauser (2009). [147](#), [155](#)
- [72] P. G. Lemarié and Y. Meyer, *Ondelettes et bases Hilbertiennes*, Rev. Mat. Iberoamericana, 2:1-18, 1986. [79](#), [97](#), [98](#)
- [73] E. Liu, V. N. Temlyakov, *Orthogonal supergreedy algorithm and applications in Compressed Sensing*, Preprint. [37](#)
- [74] G. G. Lorentz, M. von Golitschek and Y. Makovoz, *Constructive approximation: Advanced problems*, vol. 304, Springer-Verlag, Berlin, 1996. [25](#)
- [75] G. Lugosi, *Concentration-of-measure inequalities*, Lecture Notes, december 2004. Available on <http://www.econ.upf.edu/lugosi/surveys.html> [52](#)
- [76] R. Maleh, *Improved RIP Analysis of Orthogonal Matching Pursuit*, Preprint. [37](#)

REFERENCIAS

- [77] S. Mallat, *A Wavelet tour on signal processing*, Academic Press, 2nd. ed., 1999. [39](#), [42](#), [79](#), [97](#)
- [78] Q. Mo, Y. Shen, *Remarks on the Restricted Isometry Property in Orthogonal Matching Pursuit algorithm*, Preprint, 2012. [37](#)
- [79] S. Mallat, Z. Zhang, “Matching Pursuit with time-frequency dictionaries”, *IEEE Trans. Signal Processing*, vol. 41, num. 12, pp. 3397-3415, 1993. [28](#)
- [80] C. Merucci, *Applications of interpolation with a function parameter to Lorentz, Sobolev and Besov spaces. Interpolation spaces and allied topics in analysis*, Lecture Notes in Math., 1070, Springer, Berlin, 183-201, 1984. [96](#)
- [81] Y. Meyer, *Ondelettes et Opérateurs*, Vol. 1, Herman, Paris, 1990. [Traducción en inglés: *Wavelet and operators*, Cambridge University Press, 1992.] [71](#), [79](#), [97](#), [98](#), [109](#)
- [82] M. Natividade, *Best approximation with wavelets in weighted Orlicz spaces*, Monatsh. Math., Vol. 164, Number 1, 87-114, 2011. [73](#)
- [83] D. Needell and J. A. Tropp, *CoSaMP: Iterative signal recovery from incomplete and inaccurate samples*, Appl. Comput. Harmon. Anal. 26, no. 3, pp. 301-321, 2008. [32](#), [41](#)
- [84] D. Needell and R. Vershynin, *Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit*, Found. Comput. Math., 9, 317-334, 2009. [26](#), [32](#)
- [85] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer, 2nd ed., 2006. [30](#)
- [86] J. Peetre, On spaces of Triebel-Lizorkin type, Ark. Mat., 13:123-130, 1975. [163](#)
- [87] V. Peller, *Hankel operators of the class S_p , investigations of the rate of rational approximation and other applications*, Mat. Sbornik, 122, 481-510, 1980. [70](#)
- [88] P.P. Petrushev, *Direct and converse theorems for spline and rational approximation and Besov spaces*, in Function Spaces and Applications, (M. Cwikel et al., eds), Vol. 1302 of *Lecture Notes in Mathematics*, Springer, Berlin, 363-377, 1988. [70](#)

- [89] A. Pietsch, *Approximation Spaces*, J. Approx. Theory, 32, 115-134, 1981. [82](#), [92](#)
- [90] A. Pinkus, *n-widths in Approximation Theory*, Springer-Verlag, 1985. [18](#)
- [91] M. M. Rao and Z. D. Ren, *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, 146, Marcel Dekker, (1991). [107](#), [108](#)
- [92] M. Rudelson and R. Vershynin, *Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements*, CISS, 40th Annual Conference on Information Sciences and Systems, 2006. [9](#), [21](#), [41](#)
- [93] P. M. Soardi, *Wavelet Bases in Rearrangement Invariant Function Spaces*, Proc. Amer. Math. Soc., no. 12, 125:3669-3673, 1997. [71](#), [109](#)
- [94] S. B. Stechkin, *On absolute convergence of orthogonal series*, Dokl. Akad. Nauk SSSR, 102:37-40, 1955. [73](#)
- [95] V. N. Temlyakov, *The best m-term approximation and greedy algorithms*, Adv. Comput. Math., 8:249-265, 1998. [74](#)
- [96] Temlyakov, *Weak greedy algorithms*, Adv. in Comput. Math., vol. 12, pp. 213-227, 2000. [9](#), [28](#)
- [97] V.N. Temlyakov, *Nonlinear methods of approximation*, Found. Comput. Math., vol. 17, pp. 269-280, 2003. [9](#), [28](#)
- [98] V.N. Temlyakov, *Greedy approximation*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2011. [9](#), [28](#)
- [99] H. Triebel, *Interpolation theory, function spaces, differential operators*, North-Holland, 1978. [106](#)
- [100] H. Triebel, *Theory of function spaces*, Monographs in Math., #78, Birkhäuser (1983). [147](#), [165](#)
- [101] H. Triebel, *Theory of function spaces II*, Monographs in Math., #84, Birkhäuser Verlag, Basel (1992). [147](#)
- [102] J. A. Tropp, *Greedy is good: Algorithmic results for sparse approximation*, IEEE Trans. Info. Theory, vol. 50, num. 10, pp. 2231-2242, 2004. [37](#), [42](#)

REFERENCIAS

- [103] J. A. Tropp and A. Gilbert, *Signal recovery from random measurements via Orthogonal Matching Pursuit*, IEEE Trans. Info. Theory, vol. 53, num. 12, pp. 4655-4666, 2007. [42](#), [51](#), [52](#), [53](#), [62](#), [63](#)
- [104] D. Vera, Triebel-Lizorkin spaces and shearlets on the cone in \mathbb{R}^2 , Appl. Comput. Harmon. Anal. (2012), <http://dx.doi.org/10.1016/j.acha.2012.08.006> . [4](#)
- [105] G. Weiss and E. Wilson, *The mathematical theory of wavelets*. Proceeding of the NATO-ASI meeting. Harmonic analysis 2000. A celebration. Kluwer Publisher, 2001. [151](#)