

JUAN PABLO HOYOS SANCHEZ



VIABILIDAD DE LA APLICACIÓN DEL PROCESAMIENTO CUÁNTICO DE SEÑALES
EN LA TEORÍA DE WAVELETS

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Maestría en Electrónica y Telecomunicaciones

Popayán
2016

JUAN PABLO HOYOS SANCHEZ

VIABILIDAD DE LA APLICACIÓN DEL PROCESAMIENTO CUÁNTICO DE SEÑALES
EN LA TEORÍA DE WAVELETS

Trabajo de grado presentado a la Facultad de Ingeniería
Electrónica y Telecomunicaciones de la
Universidad del Cauca para la obtención del
Titulo de

Magíster en
Electrónica y Telecomunicaciones

Director:
Ph.D. Pablo Emilio Jojoa Gomez

Popayán
2016

*A mi amada Natalia, y mis adorables hijos,
Pablo Samuel y Luciana.*

*Más vale adquirir sabiduría que oro; más
vale adquirir inteligencia que plata.*
- Proverbios 16:16

Resumen

En el presente trabajo de grado se exploró la aplicación del Procesamiento Cuántico de Señales (*Quantum Signal Processing* - QSP) en la teoría wavelet, el cual es un marco de referencia que hace uso de principios de la mecánica cuántica para aportar nuevos conceptos que sirvan para el diseño o modificación de algoritmos de procesamiento de señales existentes. En especial, se trabajó en tres aspectos claves, como son: cuantización, compresión y optimización.

Para el cuantizador probabilístico presentado en el QSP, se demostró que la función de autocorrelación de las funciones wavelet ortogonales sirve como un mapeo probabilístico, que además brinda grandes ventajas respecto a su implementación. También se estableció una conexión entre los frames de Gabor y los frames Wavelet, que pueden dar origen a un nuevo conjunto de funciones denominadas Wavelet Uniformes Geométricamente. Por último se demostró que bajo ciertas características en la matriz de covarianza de un proceso aleatorio y la matriz wavelet, la transformación wavelet discreta no realiza ningún proceso de decorrelación de las variables del vector aleatorio. Por tanto, se usó una transformación de blanqueamiento diseñada en el QSP que es óptima desde un punto de vista del error cuadrático medio (*Mean Square Error* - MSE), encontrándose que en el dominio wavelet (tiempo-frecuencia) se puede lograr una disminución en la complejidad computacional de la transformación de blanqueamiento, debido a la reducción en el número de entradas distintas de cero que generalmente se puede obtener al transformar al dominio wavelet los vectores aleatorios.

Palabras clave: Procesamiento cuántico de señales, wavelet, cuantización, covarianza, optimización, *Mean Square Error* (MSE).

Abstract

This project explores the *Quantum Signal Processing* (QSP) application in wavelet theory. QSP is a framework that make use of principles of quantum mechanics to brings new concepts for the design or modification of existing signal processing algorithms. Particularly, this work is focused on three key issues: quantization, compression and optimization.

For the probabilistic quantizer presented in QSP, it is shown that the autocorrelation function of orthogonal wavelets serves as a probabilistic mapping, which also provides great advantages to its implementation. It also establishes a connection between the Gabor and Wavelet frames, which can give rise to a new set of functions called Geometrically Uniform Wavelets. Finally, it is demonstrated that under certain characteristics in the covariance matrix of a random process and wavelet matrix, the discrete wavelet transform does not perform any process of decorrelation for the variables of the random process. Then it was used a whitening transformation designed in QSP, optimal from a *Mean Square Error* (MSE) viewpoint, finding that in the wavelet (time-frequency) domain, computational complexity in the whitening transform can be reduced due to the reduction in the number of nonzero inputs that can usually be obtained by transform random vectors to the wavelet domain.

Keywords: Quantum signal processing, wavelet, quantization, covariance, optimization, MSE.

Contenido

Resumen	VII
Lista de Tablas	XII
Lista de Figuras	XV
Lista de Abreviaturas	XVII
Lista de Símbolos	XIX
1. Introducción	1
2. Teoría Wavelet	4
2.1. Espacios vectoriales	4
2.1.1. Espacios de Hilbert	5
2.1.2. Operadores Lineales	6
2.1.3. Transformaciones	7
2.1.4. Funciones átomos	8
2.2. Átomos Wavelet	8
2.2.1. Momentos de desvanecimiento	13
2.3. Frames - Transformada Wavelet Discreta	13
2.4. Análisis Multiresolución	15
2.4.1. Implementación del Análisis Multiresolución	22
2.5. Wavelet Packets	24
3. Procesamiento Cuántico de Señales y su aplicación en la teoría wavelet	26
3.1. Procesamiento Cuántico de Señales	26
3.2. Medición QSP rango uno	27
3.2.1. Diseño de algoritmos mediante el uso de ROM	28
3.3. Medición de Subespacio (SM) y SSM	33
3.4. Combinación de Medidas QSP	36
3.4.1. Clases de Mediciones Combinadas	37
3.5. Observaciones finales	40

4. Aplicación del cuantizador QSP a la teoría Wavelet	42
4.1. Cuantizador probabilístico QSP	42
4.1.1. Cuantizador Probabilístico sin Memoria	45
4.1.2. Implementación del Cuantizador Probabilístico	46
4.1.3. Relación con el Cuantizador <i>dithered</i>	46
4.2. Diseño del mapeo probabilístico $g(x)$	50
4.2.1. Autocorrelación de funciones wavelet	50
4.3. Implementación de la función de autocorrelación	52
4.4. Observaciones finales	54
5. Aplicación de la Transformación QSP en la teoría Wavelet	55
5.1. Variables correlacionadas	56
5.2. Modelo wavelet mediante operadores	56
5.3. Matriz de covarianza en el dominio Wavelet	58
5.3.1. Matriz wavelet	59
5.3.2. Propiedades de la matriz wavelet para una no-decorrelación	62
5.4. Transformación QSP	64
5.5. Transformación blanca en el dominio wavelet	66
5.5.1. Validación de la aplicación mediante simulación	68
5.6. Otras aplicaciones de la transformación QSP	70
5.6.1. Aplicación a cuantización	70
5.6.2. Aplicación a compresión	70
5.7. Observaciones finales	71
6. Conclusiones y Trabajos Futuros	72
6.1. Conclusiones	72
6.2. Trabajos futuros	73
Referencias	74

Lista de Tablas

5-1. Resultados de MSE y <i>sparsity</i>	70
--	----

Lista de Figuras

2-1.	Función wavelet coiflet de orden 4.	9
2-2.	Función scaling coiflet de orden 4.	11
2-3.	Espacios multiresolución.	17
2-4.	Descomposición multiresolución.	20
2-5.	Banco de filtros del proceso de análisis.	23
2-6.	Banco de filtros del proceso de síntesis.	24
2-7.	Árbol binario wavelet packet.	24
2-8.	Árboles binarios wavelet packet.	25
3-1.	Esquema general de la Medición QSP (Eldar, 2001).	27
3-2.	Diagrama en bloques del esquema propuesto MDPWD.	35
3-3.	Comportamiento de la BER contra Eb/No de los esquemas en un canal con desvanecimiento selectivo en frecuencia con una $f_d = 60\text{Hz}$ (Hoyos y Jojoa, 2015).	36
3-4.	Clases de familias wavelets.	38
3-5.	Ejemplo de funciones base B-spline y wavelet con su correspondientes aproximación Gabor (línea punteada) (Unser et al., 1992).	40
4-1.	Función característica de transferencia de un cuantizador uniforme (S P Lipschitz, 1992).	44
4-2.	Descripción del cuantizador mediante QSP (Eldar, 2001).	44
4-3.	Tipos de sistemas de cuantización. a. Cuantización determinístico o no <i>dithered</i> . b. Cuantización <i>dithered</i> sin memoria NSD. c. Cuantización <i>dithered</i> con memoria SD. (S P Lipshitz, 1992).	47
5-1.	a. Correlación entre las variables z_1 y z_2 . b. Correlación entre las variables x_1 y x_2 . c. Diagrama de dispersión de las variables del vector z . d. Diagrama de dispersión de las variables del vector x	57
5-2.	Efecto de la función umbral sobre una matriz de covarianza.	68
5-3.	Efecto de la transformación QSP sobre una matriz Toeplitz.	69

Lista de Abreviaturas

LS	<i>Least Square</i>
MSE	<i>Mean Square Error</i>
MMSE	<i>Minimum Mean Square Error</i>
TIC	Tecnologías de la información y las Comunicaciones
QSP	<i>Quantum Signal Processing</i>
AWGN	<i>Additive White Gaussian Noise</i>
BER	<i>Bit Error Rate</i>
UGC	Uniformes Geométricamente Compuestos
UG	Uniformes Geométricamente
IID	Independientes e Idénticamente Distribuidas
LSB	<i>Least Significant Bit</i>
NSD	<i>Non-Subtractive Dither</i>
SD	<i>Subtractive Dither</i>
ISI	<i>Intersymbol Interference</i>
MF	<i>Matched Filter</i>
f_{dp}	función de densidad de probabilidad
f_c	función característica
ROM	<i>Rank-one Measurement</i>
CSLS	<i>Covariance Shaping Least Square</i>
SM	<i>Subspace Measurements</i>
SSM	<i>Simple Subspace Measurements</i>
POVM	<i>Positive-Operator Valued Measure</i>
SVD	<i>Singular Value Decomposition</i>

Lista de Símbolos

$\langle u, v \rangle$	Producto interno
$\ x\ $	Norma Euclidiana
$\ x\ _F$	Norma de Frobenius
\sup	Valor supremo
$d_x(x, y)$	Métrica o función distancia
\oplus	Suma directa de dos espacios vectoriales
\mathbf{X}	Matriz
\mathbf{X}^*	Operador adjunto
\mathbf{X}^T	Transpuesta de la matriz
\mathbf{X}^{-1}	Inversa de la matriz
\mathbf{I}	Matriz identidad
$\text{traza}(\cdot)$	Traza de una matriz
x^*	Conjugado complejo
\hat{x}	Transformada de Fourier
$x(t)$	Señal continua en el tiempo
$x[n]$	Señal discreta
$\lfloor x \rfloor$	Entero más grande
$\delta(t)$	Distribución Dirac
$\delta[n]$	Dirac discreta
$x \star y(t)$	Convolución en tiempo continuo
$x \star y[n]$	Convolución discreta
$E\{x\}$	Valor esperado de una variable aleatoria x
\mathbb{C}^N	Espacio de vectores de dimensión finita
$\ell^2(\mathbb{Z})$	Espacio de secuencias de energía finita
$\mathcal{L}^2(\mathbb{R})$	Espacio de funciones de energía finita

Capítulo 1

Introducción

La teoría de señales digitales ha sido un área de constante investigación desde sus primeros usos en la teoría de la información (Shannon, 2001). Sus avances han permitido el desarrollo de novedosos sistemas que han impactado notablemente la manera de vivir, llevando a la sociedad a la era de las Tecnologías de la información y las Comunicaciones (TIC).

Un factor importante en el procesamiento digital de señales como en todas las áreas del conocimiento es el desarrollo de nuevas herramientas matemáticas que permitan replantear u optimizar los procedimientos existentes. Una de ellas es la teoría wavelet, que ha mostrado ser una herramienta importante para el análisis y codificación de señales no determinísticas (Akansu y Haddad, 2001; Frisch y Messer, 1992).

En el campo del diseño de sistemas de comunicaciones, varias propuestas de moduladores y ecualizadores que hacen uso de la teoría wavelet han sido presentadas con el fin de reducir el ruido existente en canales *Additive White Gaussian Noise* (AWGN) recurriendo a su adaptabilidad a las características del medio (Newlin, 1998). A su vez, han sido propuestos sistemas de multiacceso que explotan las características de flexibilidad frente a las condiciones de propagación, reducción de interferencia y distorsión, haciéndolos deseables para la nueva generación de sistemas de comunicación (Akansu et al., 1998; Jamin y Mähönen, 2005; Lakshmanan y Nikookar, 2006; You y Ilow, 2004).

Por otra parte, desde que la mecánica cuántica fue presentada en la primera mitad del siglo XX (von Neumann, 1996), se han realizado grandes esfuerzos por mostrar sus capacidades en diferentes campos. Uno de estos campos de acción es la investigación de la computación cuántica que tiene entre sus finalidades, optimizar las redes de computadores actuales (Feynman y Shor, 1982). En el campo de la teoría de la información, la mecánica cuántica ha realizado grandes aportes permitiendo el desarrollo de la criptografía cuántica (Bennett et al., 1992) y la tele portación cuántica, abriendo la posibilidad a la transmisión de datos a altas velocidades y alta fidelidad en la transmisión por medio del cifrado cuántico de llaves (Lee et al., 2011).

El procesamiento de señales no se ha quedado atrás en la incorporación de los axiomas y restricciones presentes en la teoría cuántica con el fin de crear o modificar los algoritmos y procesos existentes. Es por esto que un área reciente de investigación e innovación es el procesamiento cuántico de señales (*Quantum Signal Processing - QSP*) (Eldar, 2001).

Teniendo en cuenta lo anterior y dado que la teoría wavelet recurre al procesamiento clásico de señales, en este documento se presenta el resultado de aplicar los conceptos del QSP a la teoría wavelet, y eso se logró mediante el análisis de estructuras de ambos conceptos que permitió analizar las relaciones entre ambos y así definir aplicaciones del QSP en la teoría wavelet. Por último, se evaluó mediante simulación el desempeño de las aplicaciones obtenidas. De forma resumida, los principales resultados alcanzados en el trabajo de grado fueron los siguientes:

- Se diseñaron aplicaciones del procesamiento cuántico de señales en la teoría wavelet, entre las cuales se destacan: una nueva forma de descomposición wavelet, uso del cuantizador QSP en los coeficientes wavelet, árbol wavelet packet adaptativo en sistemas de comunicación, transformación *Minimum Mean Square Error* (MMSE) en los coeficientes wavelet, y el planteamiento de un nuevo *frame* denominado wavelets uniformes geoméricamente.
- Se demostró que la función wavelet de autocorrelación de las funciones wavelet ortogonales sirven como mapeo para el cuantizador probabilístico QSP, y se determinó una forma recursiva de obtenerla, tal que facilite su implementación.
- Se determinó la propiedad que debe satisfacer la matriz wavelet para no decorrelacionar a un vector aleatorio.
- Se demostró que la transformación QSP de decorrelación puede ser implementada con una menor complejidad computacional en el dominio wavelet (tiempo-frecuencia).

El presente documento está organizado de la siguiente manera:

En el inicio del Capítulo 2 se establece la notación a ser utilizada a lo largo de todo el documento, luego se presentan los aspectos y fundamentos más importantes de la teoría wavelet, buscando una mejor comprensión que permita encontrar nuevas modificaciones de dicha teoría.

En el Capítulo 3 se da una breve introducción al procesamiento cuántico de señales, y también se tratan conceptos que permitan modificar aspectos de la teoría wavelet. Además, se esboza un aporte del trabajo de grado relacionado con aplicaciones del QSP en la teoría wavelet, de las cuales algunas son analizadas en mayor profundidad en los capítulos siguientes. El artículo derivado de este Capítulo fue *Design of dynamic wavelet packet division multiplexing* (Hoyos y Jojoa, 2015).

En el Capítulo 4 se da un análisis del cuantizador probabilístico QSP, el cuantizador *dithered*, y la relación entre ambos. Presentándose a continuación, como aportes del trabajo de grado la descripción la aplicación del cuantizador probabilístico QSP, la demostración que la función de autocorrelación de las funciones wavelet ortogonales sirven como mapeo para el cuantizador probabilístico y una forma iterativa de obtenerlas, similar al método en que son obtenidas las funciones wavelet en la teoría del análisis multiresolución. El artículo derivado de este Capítulo fue *Wavelet autocorrelation function as mapping for a dithered quantizer* (Hoyos y Jojoa, 2016).

En el Capítulo 5 se analiza la teoría wavelet desde una perspectiva matricial, explorándose inicialmente las características de la matriz, y presentándose como aporte del trabajo de grado la identificación de las propiedades que la matriz debe satisfacer para que no decorrelacione las variables de un vector aleatorio. También se desarrolla la aplicación en el dominio wavelet de una transformación QSP que decorrelaciona un vector aleatorio desde una perspectiva del MMSE, y se demuestra que el dominio wavelet trae ventajas en la computación de la transformada QSP.

Por último, en el Capítulo 6 se exponen las conclusiones y trabajos futuros.

Capítulo 2

Teoría Wavelet

En este capítulo se presentan los aspectos más importantes de espacios vectoriales (Kreyszig, 1978), y la teoría wavelet, destacándose: átomos wavelet (Mallat, 2008), frames wavelet (Mallat, 2008; Daubechies, 1992) y análisis multiresolución (Mallat, 2008; Daubechies, 1992; Cohen y Kovačević, 1996).

2.1. Espacios vectoriales

Dada la importancia de éste tema para el desarrollo del trabajo de grado, se recurre a Kreyszig (1978) por su amplio trabajo en el área, siendo complementado con los aportes de otros autores que serán debidamente citados.

Tanto la teoría wavelet como el procesamiento cuántico de señales se soportan sobre espacios vectoriales, de allí su importancia de recordar sus principales aspectos matemáticos, los cuales son tratados a continuación.

Un espacio vectorial X no es más que un conjunto de vectores, los cuales tienen la característica que al sumar dos vectores o al multiplicarlos por un escalar (real o complejo) el resultado es un vector de ese conjunto, es decir para cualquier x y y en X se tiene:

$$x + y \in X \tag{2-1}$$

$$\alpha x \in X, \tag{2-2}$$

donde se observa que todas las combinaciones lineales de $\alpha x + \beta y$ deben estar en X . Ahora es importante recordar algunas nociones de estos conjuntos:

Punto de adherencia: Un vector a de un subespacio A , es un punto de adherencia de A si y solo si existe una sucesión de vectores (x_n) tal que $(x_n) \rightarrow a$. La *Clausura* \bar{A} de un subespacio A es el conjunto formado por todos los vectores a que son puntos de adherencia de A .

Conjunto cerrado: Un conjunto A es cerrado sí y solo sí $\bar{A} = A$.

Base: El conjunto $B = \{v_1, v_2, \dots, v_n\}$ es una base del espacio vectorial V sí y solo sí cada vector $u \in V$ puede ser escrito de una única manera como $u = c_1v_1 + c_2v_2 + \dots + c_nv_n$, es decir los vectores v son linealmente independientes y generan todos los vectores del espacio V .

Dimensión de un espacio: La dimensión de un espacio vectorial de dimensión finita V , es igual al número de vectores de cualquier base de V .

Independencia de espacios: Dos subespacios W_1, W_2 se dicen independientes sí y solo sí el único elemento común a ellos es el vector nulo, $W_1 \cap W_2 = \{0\}$.

Suma directa: Un espacio vectorial V es una *suma directa* de los subespacios W y T , escrito como $V = W \oplus T$, cuando cualquier vector $v \in V$ puede ser escrito unívocamente como $v = v_W + v_T$, donde $v_W \in S$ y $v_T \in T$. Los subespacios W y T forman una descomposición de V , y sí además son ortogonales, entonces la descomposición es llamada una *descomposición ortogonal*.

Proyección: La proyección de un vector u sobre un vector v está dada por

$$P_v u = \frac{\langle u, v \rangle}{\|v\|^2} v, \quad (2-3)$$

donde $\langle \cdot, \cdot \rangle$ es la función producto interno, y $\| \cdot \|$ la función norma, las cuales son definidas a continuación.

2.1.1. Espacios de Hilbert

Sea X un espacio vectorial, un producto interno es un mapeo de $X \times X$ en el campo escalar \mathbb{K} (\mathbb{R} o \mathbb{C}), es decir, cada par de vectores x y y tienen un escalar asociado por $\langle x, y \rangle$. Es llamado un *producto interno* si cumple para los vectores $x, y, z \in X$:

- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- $\langle x, x \rangle \geq 0$
 $\langle x, x \rangle = 0 \iff x = 0$

Si un espacio vectorial tiene definido en él un producto interno, entonces se le llama *espacio de producto interno*. Además, el producto interno sobre X define una norma inducida igual a:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Así, los espacios con norma son llamados *Espacios Normados*. La norma también puede definir una *métrica inducida* igual a $d(x, y) = \|x - y\|$. Si en un espacio normado X , una secuencia de vectores x_0, x_1, \dots cumple que, para cualquier $\epsilon > 0$ existe un entero k_ϵ tal que

$$\|x_n - x_m\| \leq \epsilon \quad \text{para todo } n, m > k_\epsilon \quad (2-4)$$

entonces se dice que la secuencia de vectores es una *secuencia Cauchy*. El espacio X se dice que es *completo* si cada secuencia Cauchy en X converge a un vector x que es elemento del espacio X . Ahora ya se puede dar la definición de un espacio de Hilbert.

Espacios de Hilbert: Un espacio de Hilbert H es un espacio vectorial completo equipado con algún producto interno. Algunos de los espacios de Hilbert más conocidos son:

\mathbb{C}^N - *Espacio de vectores de dimensión finita:* es el conjunto de vectores $x = \mathbf{x}$ de dimensión finita N y valor complejo, con producto interno dado por $\langle x, y \rangle = \sum_{n=0}^{N-1} x_n y_n^* = \mathbf{y}^* \mathbf{x}$.

$\ell^2(\mathbb{Z})$ - *Espacio de secuencias de energía finita:* es el conjunto de las secuencias $x = \{x_n\}$ de valor complejo que cumplen $\sum_{n \in \mathbb{Z}} |x_n|^2 < \infty$, y el producto interno viene dado por $\langle x, y \rangle = \sum_{n \in \mathbb{Z}} x_n y_n^*$.

$\mathcal{L}^2(\mathbb{R})$ - *Espacio de funciones de energía finita:* es el conjunto de todas las funciones $x = x(t)$ que cumplen $\int_{-\infty}^{+\infty} |x(t)|^2 dt < \infty$, y el producto interno viene definido como $\langle x, y \rangle = \int_{-\infty}^{+\infty} x(t) y^*(t) dt$.

También es de interés el *espacio de la señal* (Jacobs y Wozencraft, 1965), el cual es un espacio de Hilbert en donde sus elementos son señales, por tanto una señal $x(t)$ puede ser denotada como un vector, es decir $x = x(t)$, facilitando representar su pertenencia a un espacio vectorial. Por tanto permite referirse a vectores o señales indistintamente.

Base Riesz: (Mallat, 2008) Un conjunto de vectores $\{\psi_j\}_{j \in J} \subset H$, donde J es finito o infinito contable, es llamado una base Riesz para un espacio de Hilbert H si:

1. Es una base para H , y
2. Existen unas constantes de estabilidad $0 < \lambda_{\min} \leq \lambda_{\max} \leq \infty$ tal que cualquier $x \in H$, puede ser expresado como $x = \sum_{j \in J} \alpha_j \psi_j$ y

$$\lambda_{\min} \|x\|^2 \leq \sum_{j \in J} |\alpha_j|^2 \leq \lambda_{\max} \|x\|^2.$$

2.1.2. Operadores Lineales

Los operadores lineales generalizan el concepto de matrices y por tanto permiten explorar las ventajas presentes en los espacios de Hilbert, el cual es el espacio vectorial

donde se desea investigar.

Operador Lineal: Una función $A : H_0 \rightarrow H_1$ se denomina un operador lineal desde H_0 hasta H_1 cuando, para todo $x, y \in H_0$ y $\alpha \in \mathbb{C}$ se cumple con:

- Adición: $A(x + y) = Ax + Ay$
- Escalabilidad $A(\alpha x) = \alpha(Ax)$

Norma de un Operador: La norma de un operador lineal, denotada por $\|A\|$ se define como

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

donde $\sup \|\cdot\|$ representa la mínima cota superior o valor supremo del vector. El operador se dice ser *acotado* si la norma es finita. Los operadores acotados son continuos, ya que: si $x_k \rightarrow x$, entonces $Ax_k \rightarrow Ax$.

Inverso: Un operador lineal acotado $A : H_0 \rightarrow H_1$ se dice ser invertible sí y solo sí existe un operador lineal acotado $B : H_1 \rightarrow H_0$ tal que

$$\begin{aligned} BAx &= x \text{ para cada } x \in H_0 \\ ABx &= x \text{ para cada } x \in H_1 \end{aligned}$$

cuando tal operador existe; es único y entonces el operador inverso de A es denotado por A^{-1} .

Operador Adjunto y Auto-Operador: El operador lineal $A^* : H_1 \rightarrow H_0$ se llama *adjunto* del operador lineal $A : H_0 \rightarrow H_1$ si para cada $x \in H_0$ y $y \in H_1$ se tiene

$$\langle Ax, y \rangle_{H_1} = \langle x, A^*y \rangle_{H_0}$$

El operador A se llama *auto adjunto* o *Hermitiano* si $A = A^*$.

Operador Unitario: Un operador lineal acotado $A : H_0 \rightarrow H_1$ es unitario cuando

- es invertible
- y preserva el producto interno para todo $x, y \in H_0$ $\langle Ax, Ay \rangle_{H_1} = \langle x, y \rangle_{H_0}$

y además $A^* = A^{-1}$.

2.1.3. Transformaciones

Sean dos conjuntos X y Y , una *transformación* T es un mapeo desde X a Y , el cual es obtenido asociando a cada x en X un y de Y , denotado como $y = Tx$, y es llamado la *imagen* de x respecto a T .

Isometría: Una transformación $T : X \rightarrow Y$ se dice ser una isometría si T preserva la distancia, esto es, si para todo x y w en X

$$d_Y(T_x, T_w) = d_X(x, w) \quad (2-5)$$

donde T_x y T_w son las imágenes de x y w , respectivamente.

2.1.4. Funciones átomos

Para las definiciones de esta sección se tomó como referencia (Mallat, 2008) y (Daubechies, 1992).

Átomo tiempo-frecuencia: Son formas de onda elementales que tiene una mínima dispersión en un plano tiempo-frecuencia.

Soporte de una función: Sea f una función con dominio D , el soporte es igual a la clausura del conjunto de puntos donde la función es distinta de cero, $\text{supp}(f) = \overline{\{p \in D | f(p) \neq 0\}}$. El soporte de una función con dominio finito es compacto sí y solo sí el soporte es acotado, entonces se dice que la función tiene *soporte compacto* en D .

Suma de Poisson: En el sentido de igualdad de distribución, se cumple para la función delta de dirack $\delta[n]$ la siguiente relación:

$$\frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} e^{-jnT\omega} = \frac{1}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right), \quad (2-6)$$

así, en particular se tiene

$$\sum_{n=-\infty}^{+\infty} \delta(t - nT) \xleftrightarrow{F} \frac{1}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right) \quad (2-7)$$

Ortogonalidad de Poisson: Sean dos funciones ortogonales de la forma $\langle f(t), f(t + n) \rangle = \delta[n]$ y recurriendo a la suma de Poisson, su transformada de Fourier viene dada por

$$\langle f(t), f(t + n) \rangle = \delta[n] \xleftrightarrow{F} \sum_{k=-\infty}^{+\infty} |\hat{f}(\omega + 2\pi k)|^2 = 1 \quad (2-8)$$

donde \hat{f} corresponde a la transformada de Fourier de f .

2.2. Átomos Wavelet

Cuando se desea analizar señales que presentan irregularidades o señales noestacionarias, es deseable pensar en utilizar átomos tiempo-frecuencia que tengan diferentes soportes de tiempo (diferentes tamaños) que permitan aislar dichas singularidades. Así, una wavelet $\psi(t) \in \mathcal{L}^2(\mathbb{R})$ es un átomo tiempo-frecuencia que satisface dos propiedades fundamentales (Mallat, 2008; Daubechies, 1992):

1. $\|\psi\| = 1$
2. $\int_{-\infty}^{\infty} \psi(t)dt = 0$

donde la primer condición sugiere normalización y la segunda condición implica que la función debe cambiar de signo en $(-\infty, +\infty)$ y debe desvanecerse en $-\infty$ y $+\infty$, por tanto representa una pequeña onda o wavelet, como se aprecia en la Figura 2-1. Esta función se le conoce como *wavelet madre* (Daubechies, 1992), ya que a partir de ella mediante escalonamientos a y traslaciones (localizaciones) b se pueden obtener un completo diccionario de funciones wavelet $G = \{\psi_{a,b}(t)\}_{b \in \mathbb{R}, a > 0}$ con

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \quad (2-9)$$

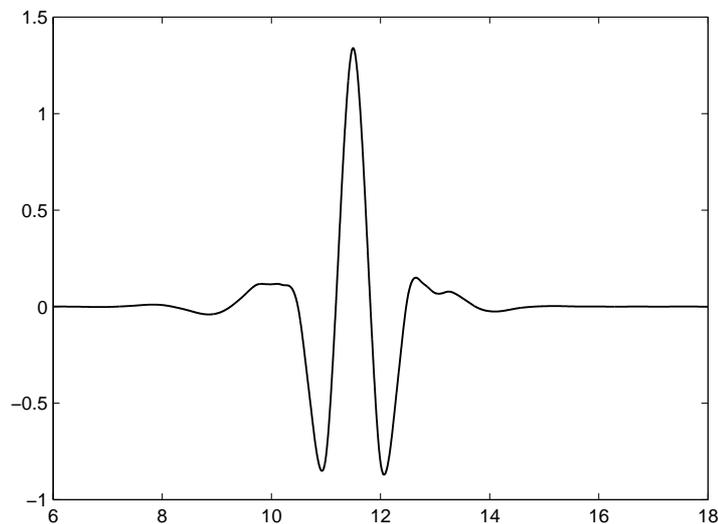


Figura 2-1: Función wavelet coiflet de orden 4.

donde el factor $1/\sqrt{a}$ permite mantener una normalización de la energía a lo largo de las diferentes escalas. Una forma alterna de presentar un diccionario wavelet es recurriendo al uso de operadores lineales, donde el operador de dilatación D sería

$$D_a \psi(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t}{a} \right) \quad (2-10)$$

y el operador de traslación vendría dado por

$$T_b \psi(t) = \psi(t - b) \quad (2-11)$$

así, el diccionario sería igual a $G = \{D_a T_b \psi\}_{b \in \mathbb{R}, a > 0}$. Recordando la relación de la transformada de Fourier de una función escalada $f(\frac{t}{a}) \xrightarrow{F} |a| \hat{f}(aw)$, se puede determinar la transformada de Fourier de las funciones wavelet escaladas dada por $\hat{\psi}_a = \sqrt{a} \hat{\psi}(aw)$,

observándose que la frecuencia es inversa a la escala, $w \propto \frac{1}{a}$, así para una escala $a > 1$ se obtiene una wavelet dilatada en tiempo con un espectro contraído, mientras que si la escala $0 < a < 1$ la wavelet es comprimida en tiempo y dilata en frecuencia. Ahora, sobre este diccionario se puede proyectar una función $f \in \mathcal{L}^2(\mathbb{R})$, mediante el uso del producto interno y debido a que el espacio es de *Hilbert*, la Transformada wavelet continua (Daubechies, 1992) en cualquier tiempo b y escala a sobre el correspondiente átomo wavelet es

$$TW_f(a, b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right) dt \quad (2-12)$$

este mapeo o descomposición es conocido como «Análisis» de la función, y mide el comportamiento de la función f en la vecindad b a una escala a . Si la función presenta irregularidades (picos), se hace necesario el uso de una wavelet con una escala a pequeña, mientras que si la función es regular (suave) en ciertos intervalos, solo será necesario wavelets con grandes escalas a , ésta es la gran ventaja que presenta la transformada wavelet de adaptar su tamaño a las transiciones de la función bajo análisis. Para poder asegurar que tal transformada permita una reconstrucción completa de la función f y mantenga la conservación de energía, la wavelet debe satisfacer una condición más, conocida como condición de admisibilidad y especificada por el teorema 2.1 de Grossmann y Morlet (Grossmann y Morlet, 1984).

Teorema 2.1. (Admisibilidad) Sea $\psi \in \mathcal{L}^2(\mathbb{R})$ una función real tal que

$$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(w)|^2}{w} dw < +\infty \quad (2-13)$$

cuyo valor dependerá de la wavelet escogida, además si se cumple que C_ψ sea finita, entonces cualquier función $f \in \mathcal{L}^2(\mathbb{R})$ puede ser reconstruida y está dada por

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} TW_f(a, b) \psi_{a,b}(t) db \frac{da}{a^2} \quad (2-14)$$

para que la integral (2-13) sea finita, la función wavelet $\psi(t)$ debe de tener una componente $\hat{\psi}(0) = \int_{-\infty}^{\infty} \psi(t) dt = 0$, explicando el porque la función wavelet tiene un promedio igual a cero (condición 2 de las wavelet) y obliga que la función wavelet sea oscilatoria, es decir debe ser una *onda*, además su espectro parece tener un comportamiento de filtro pasa banda. Si es cumplida la condición de admisibilidad, la función f puede ser reconstruida sin pérdida de información y viene dada por (2-14), este proceso es conocido como «Síntesis» de la función o Transformada inversa wavelet (Mallat, 2008; Daubechies, 1992). Pero la ecuación de reconstrucción implica que se deben tener infinitos coeficientes wavelet; es decir cada vez que dilatamos una wavelet ($a > 1$), ésta permite cubrir cada vez menos frecuencias, y si se desea cubrir las frecuencias bajas alrededor de 0, se debe tener átomos wavelet infinitamente dilatados, lo cual en la práctica es imposible. Pero, si se recurre a una función auxiliar definida por

$$|\hat{\phi}(w)|^2 = \int_1^{+\infty} |\hat{\psi}(aw)|^2 \frac{da}{a} \quad (2-15)$$

la función $\phi(t)$ es una agregación de wavelets a escalas mayores que 1 y es llamada *función escala* o *scaling* (Mallat, 2008), un ejemplo se puede apreciar en la Figura 2-2.

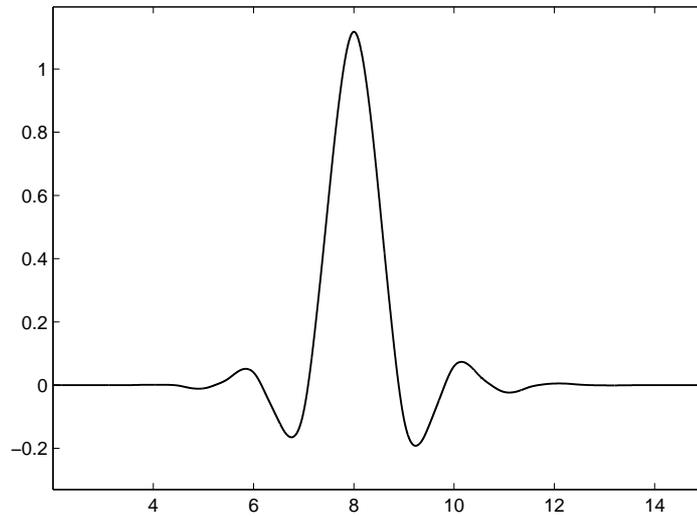


Figura 2-2: Función scaling coiflet de orden 4.

Realizando el cambio de variable $\beta = aw$, y reescribiendo la ecuación (2-15)

$$|\hat{\phi}(w)|^2 = \int_w^{+\infty} \frac{|\hat{\psi}(\beta)|^2}{\beta} d\beta \quad (2-16)$$

se obtiene la condición de admisibilidad cuando $w \rightarrow 0$

$$\lim_{w \rightarrow 0} \int_w^{+\infty} \frac{|\hat{\psi}(\beta)|^2}{\beta} d\beta = \lim_{w \rightarrow 0} |\hat{\phi}(w)|^2 = C_\psi \quad (2-17)$$

entonces, se puede decir que la función escala cumple con dos propiedades:

1. $\|\phi\| = 1$
2. $\int_{-\infty}^{\infty} \phi(t) dt \neq 0$

la propiedad 1 mantiene la normalización de la función escala, mientras que la propiedad 2 indica que la componente dc de la función escala es distinta de cero $\hat{\phi}(0) \neq 0$, por tanto su espectro va a tener una forma de filtro pasa bajas, actuando como complemento a la función wavelet en la proyección de la función f (Mallat, 2008; Daubechies, 1992). A partir de la función escala se puede construir un diccionario de funciones escalas $D_\phi = \{\phi_{a,b}(t)\}_{a>0, b \in \mathbb{R}}$ que sean traslaciones y dilataciones de la función escala madre

$$\phi_{a,b}(t) = \frac{1}{\sqrt{a}} \phi\left(\frac{t-b}{a}\right) \quad (2-18)$$

y la proyección sobre este espacio se puede ver como una aproximación de f en baja frecuencia en la escala a , dada por

$$TE_f(a, b) = \langle f, \phi_{a,b} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \phi^* \left(\frac{t-b}{a} \right) dt \quad (2-19)$$

con esta nueva idea de aproximación de f , la ecuación de reconstrucción de la función f (2-14) se puede empezar a replantear como

$$f(t) = \frac{1}{C_\psi} \int_0^{a_0} \int_{-\infty}^{+\infty} TW_f(a, b) \psi_{a,b}(t) db \frac{da}{a^2} + \frac{1}{C_\psi} \int_{a_0}^{\infty} \int_{-\infty}^{+\infty} TW_f(a, b) \psi_{a,b}(t) db \frac{da}{a^2} \quad (2-20)$$

ahora, expresando la segunda parte de la ecuación como productos de convoluciones se tiene

$$f(t) = \frac{1}{C_\psi} \int_0^{a_0} f * \psi_a^* * \psi_a(t) \frac{da}{a^2} + \frac{1}{C_\psi} \int_{a_0}^{\infty} f * \psi_a^* * \psi_a(t) \frac{da}{a^2} \quad (2-21)$$

con $\psi_a(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t}{a} \right)$. Desarrollando la segunda integral de la igualdad y tomando la transformada de Fourier se obtiene

$$\frac{1}{C_\psi} \int_{a_0}^{\infty} \hat{f}(w) \sqrt{a} \hat{\psi}(aw)^* \sqrt{a} \hat{\psi}(aw) \frac{da}{a^2} = \frac{\hat{f}(w)}{C_\psi} \int_{a_0}^{\infty} |\hat{\psi}(aw)|^2 \frac{da}{a} \quad (2-22)$$

haciendo un cambio de variable $\gamma = a/a_0$ y recordando la ecuación (2-15), se obtiene:

$$\frac{\hat{f}(w)}{C_\psi} \int_1^{\infty} |\hat{\psi}(a_0 \gamma w)|^2 \frac{d\gamma}{\gamma} = \frac{\hat{f}(w)}{C_\psi} |\hat{\phi}_{a_0}(w)|^2 \quad (2-23)$$

y

$$\frac{\hat{f}(w)}{C_\psi} |\hat{\phi}_{a_0}(w)|^2 = \frac{\hat{f}(w)}{C_\psi} \hat{\phi}_{a_0}^*(w) \hat{\phi}_{a_0}(w) \quad (2-24)$$

entonces se puede tomar la transformada de Fourier inversa para obtener la expresión en tiempo

$$\frac{f}{C_\psi} * \frac{\phi_{a_0}^*}{\sqrt{a_0}} * \frac{\phi_{a_0}}{\sqrt{a_0}} = \frac{1}{C_\psi a_0} f * \phi_{a_0}^* * \phi_{a_0} \quad (2-25)$$

y por tanto se obtiene una nueva ecuación de reconstrucción que sería igual a

$$f(t) = \frac{1}{C_\psi} \int_0^{a_0} f * \psi_a^* * \psi_a(t) \frac{da}{a^2} + \frac{1}{C_\psi a_0} f * \phi_{a_0}^* * \phi_{a_0} \quad (2-26)$$

donde la proyección de la función f se realiza para escalas $a < a_0$ en las funciones wavelet, que darían la información de altas frecuencias o detalles, y para las $a > a_0$ en las funciones escala que darían la información de bajas frecuencias o aproximaciones (Mallat, 2008).

2.2.1. Momentos de desvanecimiento

Una característica importante de las funciones wavelets se denominada *Momentos de desvanecimiento* (Mallat, 2008), que vienen definidos como:

$$\int_{-\infty}^{+\infty} t^p \psi_s(t) dt = 0 \quad \forall \quad p = 0, 1, \dots, k-1 \quad (2-27)$$

entonces se dice que la wavelet tiene momentos de desvanecimiento de orden k , o en otras palabras, plantea que la wavelet es ortogonal a polinomios de grado menores o iguales que $k-1$. Son muy importantes cuando la función f es un polinomio o puede ser expresado por uno de grado menor o igual a k , ya que la función al ser proyectada, los coeficientes wavelet van a ser muy pequeños o nulos, y por tanto la proyección se realizará sobre las funciones escala, tendiendo a comprimir funciones no-oscilatorias. Por ejemplo sea $f(t) = t^n; n < k$

$$\begin{aligned} f(t) &= \frac{1}{C_\psi} \int_0^{a_0} f * \psi_a^* * \psi_a(t) \frac{da}{a^2} + \frac{1}{C_\psi} \int_{a_0}^{\infty} f(t) * \psi_a^* * \psi_a(t) \frac{da}{a^2} \\ &= \frac{1}{C_\psi} \int_0^{a_0} t^n * \psi_a^* * \psi_a(t) \frac{da}{a^2} + \frac{1}{C_\psi a_0} TE_f(\cdot, a_0) * \phi_{a_0}(t) \end{aligned}$$

pero

$$\begin{aligned} t^n * \psi_a^* &= \int_{-\infty}^{\infty} t^n \psi_a^*(b-t) dt \\ &= \int_{-\infty}^{\infty} t^n \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right) dt \\ &= 0 \end{aligned}$$

entonces la función es igual a

$$f(t) = \frac{1}{C_\psi a_0} TE_f(\cdot, a_0) * \phi_{a_0}(t)$$

aunque la función de reconstrucción permite recuperar la función f , presenta un alto grado de redundancia e inestabilidad, así para evitar estos problemas se recurre al uso de funciones o vectores denominados *Frames*.

2.3. Frames - Transformada Wavelet Discreta

De acuerdo con Daubechies (1992, 1988, 1993), cuando se desea asegurar una reconstrucción numéricamente estable de una función f , se debe asegurar que los coeficientes tanto de análisis y de síntesis lo sean, y esto se logra mediante el uso de vectores con ciertas características que se les denomina *Frame*, los cuales sin embargo tienen la limitación de que la representación de la función f es redundante, aunque con la ventaja que las familias de vectores pueden no tener una estructura definida. Duffin y Schaeffer (1952) definen los frames mediante el siguiente teorema:

Teorema 2.2. (*Frame*) Una familia de funciones $(\varphi_i)_{i \in I}$ en un espacio de Hilbert H , es llamada un *frame*, si existen $A > 0$ y $B < \infty$, tal que para toda función f en H se cumple que

$$A \|f\|^2 \leq \sum_{i \in I} |\langle f, \varphi_i \rangle|^2 \leq B \|f\|^2$$

donde las constantes A y B son llamadas los límites del *frame*.

Con estos vectores se puede obtener una reconstrucción numéricamente estable de f a partir de sus coeficientes, y cuando las constantes son iguales ($A = B$) se dice que el *frame* es *tight* (estrecho), además si $A = B = 1$ se obtiene una base ortogonal. Si los vectores φ_i son linealmente independientes el *frame* es llamado una *base Riesz* y el *frame* no es redundante. La redundancia viene dada porque el *frame* debe tener una dimensión (número de vectores) mayor o igual a la dimensión del espacio de la función f (Mallat, 2008; Cohen y Kovačević, 1996).

Si los vectores φ_i son linealmente dependientes el rango no va a ser completo y va a ser un subespacio de $\ell^2(\mathbb{Z})$, haciéndolos deseables para aplicaciones como la reducción de ruido. Entonces se tienen dos opciones, 1. eliminar la redundancia y obtener una base ortogonal o 2. tener redundancia para tener libertad en la selección de las bases, las cuales dependerán del tipo de aplicación deseada.

Ahora ya que se ha definido que es un *frame*, se debe plantear la pregunta ¿cómo construir un *frame* de funciones wavelet?, para darle respuesta se opta por la opción 2, es decir se va a dejar cierto grado de redundancia, para tener cierto grado de libertad en la selección de la base (Daubechies, 1992). La forma es discretizando los parámetros a, b , donde se toma la escala como $a = a_0^j, j \in \mathbb{Z}$ con un paso de dilatación fijo $a_0 > 1$. El tiempo de desplazamiento b debe ser seleccionado de tal manera que permita cubrir todo el eje de tiempo, supongase $j = 0$, entonces con $b = nb_0 (b_0 > 0)$ y un adecuado b_0 se logra cubrir todo el eje de tiempo. Pero si $j \neq 0$ el tiempo de desplazamiento debe tener en cuenta la escala de la función debido a que el ancho ha variado y por tanto debe ser proporcional a la escala, entonces parece natural escoger $b = nb_0 a_0^j$. Con estos nuevos parámetros se obtiene la familia de wavelet discretizadas

$$\psi_{j,n}(t) = \frac{1}{\sqrt{a_0^j}} \psi\left(\frac{t - nb_0 a_0^j}{a_0^j}\right) \quad (2-28)$$

que dan origen a la Transformada wavelet discreta (Mallat, 2008; Daubechies, 1992). Ahora, ya que la función $f \in \mathcal{L}^2(\mathbb{R})$, tiene norma finita, los coeficientes de la proyección sobre las funciones wavelet deben ser finitos

$$\sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2 \leq B \|f\|^2 \quad (2-29)$$

entonces la secuencia $(\langle f, \psi_{j,n} \rangle)_{j,n \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$. Por otro lado si $\sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2$ es pequeño, $\|f\|^2$ debería también ser pequeño. Esto nos dice que debería existir un $\alpha < +\infty$

tal que $\sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2 \leq 1$. Tómesese cualquier función $f \in \mathcal{L}^2(\mathbb{R})$ y defínase una nueva función $\tilde{f} = \left[\sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2 \right]^{-1/2} f$, entonces la proyección $\sum_{j,n} |\langle \tilde{f}, \psi_{j,n} \rangle|^2 \leq 1$, implica que $\|\tilde{f}\|^2 < \alpha$ y

$$\|\tilde{f}\|^2 = \frac{\|f\|^2}{\sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2} \leq \alpha$$

despejando

$$A\|f\|^2 \leq \sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2 \quad (2-30)$$

donde $A = 1/\alpha$, y si se reúnen las ecuaciones (2-29) y (2-30) nos dice que una reconstrucción numéricamente estable de f es posible solo si

$$A\|f\|^2 \leq \sum_{j,n} |\langle f, \psi_{j,n} \rangle|^2 \leq B\|f\|^2 \quad (2-31)$$

entonces la familia $\{\psi_{j,n}(t)\}$, $j, n \in \mathbb{Z}$ constituye un *frame*. Ya se ha asegurado la existencia de los dos límites A, B , pero no todos los valores de a_0, b_0 permiten tener un *frame* de wavelets, por tanto el siguiente teorema de Daubechies (1992) establece ciertas condiciones sobre a_0, b_0 para que $\psi_{j,n}(t)$ formen un *frame* para el espacio $\mathcal{L}^2(\mathbb{R})$.

Teorema 2.3. (Condición necesaria) Si $\psi_{j,n}(t)$, $j, n \in \mathbb{Z}$ constituye un *frame* para $\mathcal{L}^2(\mathbb{R})$, entonces los límites del *frame* satisfacen

$$A \leq \frac{C_\psi}{b_0 \ln a_0} \leq B. \quad (2-32)$$

Este teorema nos dice que si la función madre cumple la condición de admisibilidad, existirán un sinnúmero de posibles valores de a, b tal que la familia de funciones wavelet $\psi_{j,n}(t)$ forme un *frame* wavelet. Además, si el *frame* es tight, y $A = B = 1$, que se obtiene con $a_0 = 2, b_0 = 1$, entonces el *frame* es una base ortonormal de acuerdo al siguiente teorema (Mallat, 2008)

Teorema 2.4. (Redundancia) En un espacio de dimensión N , un *frame* de $P \geq N$ vectores normalizados, los límites A y B satisfacen

$$A \leq \frac{P}{N} \leq B \quad (2-33)$$

entonces no habría redundancia ($P = N$).

2.4. Análisis Multiresolución

Este nuevo enfoque es originado por la necesidad de buscar formas eficientes de construcción de bases de wavelets ortogonales, es decir; construir vectores o funciones de la base tal que

$$\langle \psi_i, \psi_k \rangle = \delta(i - k) \quad (2-34)$$

y que a su vez permita lograr tener una transformada wavelet que elimine en algún grado la redundancia presente en la transformación. Mallat (2008) plantea una forma de diseñarlas mediante el Análisis Multiresolución, que además aporta un marco matemático conceptual que facilita el entendimiento de la teoría wavelet y la descomposición en subbandas.

De la fotografía se sabe que la resolución indica el nivel de detalle disponible, por tanto si se desea una mayor resolución se debe reducir el intervalo de muestreo para incrementar el detalle; asociando esta idea a la proyección de una función f , y si cada escala $a = 2^j$ representa un espacio de proyección V_j , entonces la escala debe ser pequeña para poder obtener una proyección $P_{V_j}f$ que se aproxime a la función f , y si además es la mejor entonces decimos que la proyección es ortogonal. A la proyección de la función f sobre el espacio V_j se le llama *aproximación* (Daubechies, 1992).

La idea detrás del análisis multiresolución es lograr una proyección de la función f a diferentes resoluciones o escalas, y para lograrlo el siguiente teorema de Mallat (2008) establece las propiedades que deben tener los espacios de proyección.

Teorema 2.5. (*Definición multiresolución*) Una secuencia $V_j_{\{j \in \mathbb{Z}\}}$ de subespacios cerrados de $\mathcal{L}^2(\mathbb{R})$, es una aproximación multiresolución si satisfacen

1. $f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j, \forall (j, k) \in \mathbb{Z}^2$
2. Los espacios V_j son encajados
 $\dots V_1 \subset V_0 \subset V_{-1} \dots$
3. $f(t) \in V_j \Leftrightarrow f(2^j t) \in V_0, \forall j \in \mathbb{Z}$
4. $\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{\emptyset\}$,
5. $\lim_{j \rightarrow -\infty} V_j = \text{Clausura}(\bigcup_{j=-\infty}^{+\infty} V_j) = \mathcal{L}^2(\mathbb{R})$,

y existe una base Riesz de V_0 .

Analicemos las anteriores propiedades de los espacios de multiresolución, la propiedad 1 indica que cada subespacio V_j es invariante ante traslaciones, que es un requerimiento importante para asegurar la multiresolución. La propiedad 2 implica que con la información presente en el espacio V_j es posible obtener la proyección gruesa en el espacio V_{j+1} , como se aprecia en la Figura 2-3. La propiedad 3 es la que asegura la multiresolución, ya que indica que todos los subespacios son versiones escaladas del espacio central V_0 . La propiedad 4 indica que cuando la escala es demasiado grande el espacio de aproximación es nulo, es decir que no habrá aproximación, mientras que la propiedad 5 indica que si la escala es muy pequeña se alcanza la máxima resolución y por tanto la aproximación sobre este espacio convergerá a la función original f . La existencia de la base Riesz del espacio V_0 y las propiedades 2 y 3 hacen que el trabajo se limite a encontrar esta base, porque a partir de ella se pueden encontrar las demás

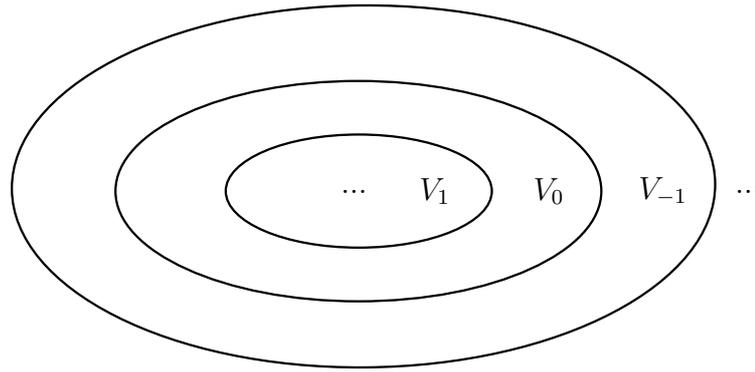


Figura 2-3: Espacios multiresolución.

bases ortogonales de los espacios V_j .

Recurriendo al siguiente teorema (Daubechies, 1992), es posible encontrar una base ortonormal para cada espacio V_j , mediante la ortogonalización de la base Riesz del espacio central V_0 y la construcción de una función ϕ llamada *funcion scaling*.

Teorema 2.6. (*Función Scaling*) Sean $\{V_j\}_{j \in \mathbb{Z}}$ una aproximación multiresolución y ϕ la función scaling con la transformada de fourier

$$\hat{\phi}(w) = \frac{\hat{\theta}(w)}{(\sum_{k=-\infty}^{+\infty} |\hat{\theta}(w + 2k\pi)|^2)^{1/2}} \quad (2-35)$$

entonces para cada espacio $V_j, j \in \mathbb{Z}$ existe una base ortonormal dada por la familia

$$\left\{ \phi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t - 2^j n}{2^j}\right) \right\}_{n \in \mathbb{Z}} \quad (2-36)$$

de acuerdo al anterior teorema, la familia de funciones scaling satisface la propiedad de ortonormalidad de Poisson,

$$\sum_{k=-\infty}^{+\infty} |\hat{\phi}(w + 2\pi k)|^2 = 1 \quad (2-37)$$

una forma alterna para ver el anterior teorema parte de la independencia lineal de los vectores que forma la base Riesz, y sí los límites superior e inferior son iguales, $A = B$, la base Riesz se ha ortogonalizado, pero si además $A = B = 1$, lo cual se logra con las escala diádica $a_0 = 2, b_0 = 1$, la base se ha normalizado, entonces se ha obtenido una base ortonormal para el espacio V_0 que mediante escalonamientos (propiedad 3) sirve para la construcción de bases para todos los espacios V_j .

Ahora, ya que $V_1 \subset V_0$, la funcion scaling $\phi_1(t)$ se puede expresar como

$$\phi_1(t) = \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} h[n] \phi(t - n) \quad (2-38)$$

con

$$h[n] = \left\langle \frac{1}{\sqrt{2}} \phi \left(\frac{t}{2} \right), \phi(t - n) \right\rangle \quad (2-39)$$

donde $h[n]$ son los coeficientes de la función *scaling*, por tanto $h[n]$ se puede interpretar como un filtro discreto pasa bajo (Daubechies, 1992). Esta es una importante observación porque permite la construcción de funciones de tiempo continuo partiendo de filtros discretos iterados. Si se define la transformada de Fourier del filtro *scaling* como $\hat{h}(w) = \sum_{n=-\infty}^{+\infty} h[n]e^{-jwn}$, entonces la transformada de Fourier de la función *scaling* sería

$$\hat{\phi}(2w) = \frac{\hat{h}(w)}{\sqrt{2}} \hat{\phi}(w) \quad (2-40)$$

por inducción, se tiene que para cualquier $k \geq 0$ la ecuación se convierte en $\hat{\phi}(2^{-k+1}w) = \frac{1}{\sqrt{2}} \hat{h} \left(\frac{w}{2^k} \right) \hat{\phi} \left(\frac{w}{2^k} \right)$, la cual es una ecuación recursiva, expresándola se tendría

$$\hat{\phi}(w) = \left\{ \prod_{m=1}^{+\infty} \frac{1}{\sqrt{2}} \hat{h} \left(\frac{w}{2^m} \right) \right\} \hat{\phi}(0) \quad (2-41)$$

donde $\hat{\phi}(0) = \lim_{m \rightarrow +\infty} \hat{\phi}(w/2^k)$. Para que el producto infinito del filtro $h[n]$ de origen a la transformada de Fourier de la función *scaling* (converga), debe satisfacer las propiedades descritas por el siguiente teorema (Mallat, 2008)

Teorema 2.7. (Condiciones necesarias de $h[n]$) Si $\hat{h}(w)$ es de periodo 2π y continuamente diferenciable en una vecindad de $w = 0$, y si satisface

$$|\hat{h}(w)|^2 + |\hat{h}(w + \pi)|^2 = 2 \quad \hat{h}(0) = \sqrt{2} \quad \inf_{w \in [-\pi/2, \pi/2]} |\hat{h}(w)| > 0, \quad (2-42)$$

entonces

$$\hat{\phi}(w) = \left\{ \prod_{m=1}^{+\infty} \frac{1}{\sqrt{2}} \hat{h} \left(\frac{w}{2^m} \right) \right\} \quad (2-43)$$

es la transformada de Fourier de una función *scaling* $\phi \in \mathcal{L}^2(\mathbb{R})$.

Este teorema de Mallat y Meyer plantea que toda función *scaling* tiene asociada un filtro pasa bajo, pero no todo filtro pasa bajo genera una función *scaling*, y además de establecer las propiedades necesarias, implica con la primera condición la ortogonalidad del filtro, es decir $\sum_{n \in \mathbb{Z}} h[n]h[n - 2k] = \delta[k]$, por otra parte también de forma implícita indica que $|\hat{\phi}(0)| = 1$. Los filtros que cumplen la propiedad (2-42) son conocidos como *Filtros espejos en Cuadratura* o *QMF (Quadrature mirror filter)* (Mallat, 2008). En la práctica la forma de obtener la función *scaling* generalmente no es realizar la transformada de Fourier inversa, sino que se recurre al *algoritmo en cascada* para wavelet de soporte compacto propuesto por Daubechies (1992).

Así, se ha construido la base ortonormal de funciones scaling para cualquier espacio V_j , entonces mediante la proyección ortogonal de la función f en el espacio V_j se obtiene la *Aproximación* de la función f a la escala 2^j , que viene dada por la combinación lineal de las funciones scaling

$$P_{V_j} f = \sum_{n=-\infty}^{\infty} \langle f, \phi_{j,n} \rangle \phi_{j,n} \quad (2-44)$$

Construcción de la Función Wavelet

La teoría de Multiresolución (Mallat, 2008) asegura que tenemos una secuencia de subespacios cerrados encajados, es decir que $V_{j+1} \subset V_j$, entonces es posible expresar el subespacio faltante como el complemento ortogonal de V_{j+1}

$$V_j = V_{j+1} \oplus W_{j+1} \quad (2-45)$$

Así, cada espacio de aproximación V_j es descompuesto en un espacio de aproximación grueso V_{j+1} y un espacio de detalles W_{j+1} , como se observa en la Figura 2-4. Pero, ¿qué representa W_{j+1} ?, si V_j representa la aproximación de la función f en la escala 2^j y además con esta información se puede obtener la aproximación de f a una escala mayor 2^{j+1} , entonces si se desea analizar la función se debería tener la aproximación de f en cada espacio V_j , lo cual haría un proceso muy redundante. Por tanto, sería mejor si se procesara la información que está en la escala 2^j y que desaparece en la escala 2^{j+1} , es decir la información diferente entre las escalas, y esta información no es más que las fluctuaciones o «detalles» de f que aparecen en la escala fina 2^j , y que se pueden obtener mediante la proyección de la función sobre el espacio W_{j+1} .

$$P_{V_j} f = P_{V_{j+1}} f + P_{W_{j+1}} f \iff P_{W_{j+1}} f = P_{V_j} f - P_{V_{j+1}} f \quad (2-46)$$

Para realizar la proyección ortogonal sobre W_{j+1} , se debe encontrar una base ortonormal, y de manera similar a la función scaling, la base se puede lograr escalando y trasladando una nueva función llamada *función Wavelet* (Mallat, 2008).

Entonces, como $W_1 \subset V_0$, es posible expresar la función wavelet $\psi_1(t)$ en función de la ya conocida función scaling

$$\psi_1(t) = \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right) = \sum_{n=-\infty}^{+\infty} g[n] \phi(t-n) \quad (2-47)$$

donde

$$g[n] = \left\langle \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle \quad (2-48)$$

y como los espacios W_1 y V_1 son ortogonales, las funciones base también lo deben ser

$$\langle \psi_1(t), \phi_{1,m}(t) \rangle = \left\langle \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right), \frac{1}{\sqrt{2}} \phi\left(\frac{t-2m}{2}\right) \right\rangle = 0 \quad (2-49)$$

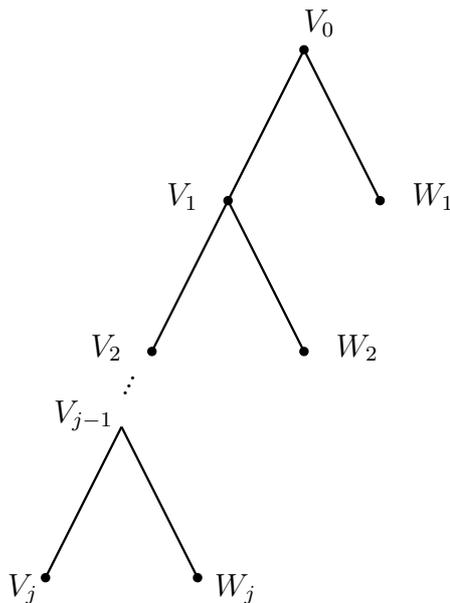


Figura 2-4: Descomposición multiresolución.

reemplazado (2-38) y (2-47) en (2-49), se tiene

$$\left\langle \sum_{n=-\infty}^{+\infty} g[n]\phi(t-n), \sum_{k=-\infty}^{+\infty} h[k-2m]\phi(t-k) \right\rangle = 0 \quad (2-50)$$

y por la propiedad de linealidad del producto interno

$$\begin{aligned} \sum_{n=-\infty}^{+\infty} \langle g[n]\phi(t-n), \sum_{k=-\infty}^{+\infty} h[k-2m]\phi(t-k) \rangle &= 0 \\ \sum_{n=-\infty}^{+\infty} g[n] \langle \phi(t-n), \sum_{k=-\infty}^{+\infty} h[k-2m]\phi(t-k) \rangle &= 0 \end{aligned} \quad (2-51)$$

analizado el producto interno

$$\begin{aligned} I &= \langle \phi(t-n), \sum_{k=-\infty}^{+\infty} h[k-2m]\phi(t-k) \rangle \\ I &= \sum_{k=-\infty}^{+\infty} \langle \phi(t-n), h[k-2m]\phi(t-k) \rangle \\ I &= \sum_{k=-\infty}^{+\infty} h[k-2m] \langle \phi(t-n), \phi(t-k) \rangle \end{aligned}$$

y por ortogonalidad $\langle \phi(t-n), \phi(t-k) \rangle = \delta[n-k]$, entonces la sumatoria es distinta de cero solo cuando $n = k$, por tanto

$$I = h[n-2m]$$

por lo que reemplazando en (2-51) se obtiene la condición de ortogonalidad de los filtros

$$\sum_{n=-\infty}^{+\infty} g[n]h[n-2m] = 0.$$

Existen muchas formas de escoger $g[n]$ de forma que satisfaga la ortogonalidad con $h[n]$, una forma es la dada por Strang y Nguyen (1997) que consiste en invertir a $h[n]$ y alternarlo en signo, o en otras palabras es construir a partir de un filtro pasa bajo $h[n]$ un filtro pasa alto $g[n]$, con sus coeficientes iguales a

$$g[n] = (-1)^{1-n}h[1-n] \quad (2-52)$$

además, la transformada de Fourier de (2-47) es $\hat{\psi}(2w) = \frac{1}{\sqrt{2}}\hat{g}(w)\hat{\phi}(w)$, por tanto al reescalar se obtiene

$$\hat{\psi}(w) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{w}{2}\right)\hat{\phi}\left(\frac{w}{2}\right) \quad (2-53)$$

con lo cual se ha cumplido las dos condiciones del siguiente teorema de Mallat (2008), para que la familia de funciones wavelet formen una base ortonormal.

Teorema 2.8. (Condiciones bases Wavelets) Sea $\psi(t)$ una función con transformada de Fourier

$$\hat{\psi}(w) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{w}{2}\right)\hat{\phi}\left(\frac{w}{2}\right) \quad (2-54)$$

con

$$\hat{g}(w) = e^{-jw}\hat{h}^*(w+\pi) \quad (2-55)$$

entonces la familia

$$\left\{\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-2^jn}{2^j}\right)\right\}_{n \in \mathbb{Z}} \quad (2-56)$$

es una base ortonormal de cualquier espacio $\{W_j\}$. Para todas las escalas, $\{\psi_{j,n}, j, n \in \mathbb{Z}\}$ es una base ortonormal del espacio $\mathcal{L}^2(\mathbb{R})$.

con lo cual ya se tienen las bases ortonormales para cada uno de los espacios de detalles W_j , entonces la proyección de la función sobre el espacio W_j vendría dada por

$$P_{W_j}f = \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \quad (2-57)$$

Ahora, por la ortogonalidad de los subespacios W_j , sabemos que $V_j = V_{j+1} \oplus W_{j+1}$, y a su vez $V_{j+1} = V_{j+2} \oplus W_{j+2}$, y ya que $V_\infty = \{0\}$, entonces de forma general se tiene

$$\mathcal{L}^2(\mathbb{R}) = \bigoplus_{j=-\infty}^{+\infty} W_j \quad (2-58)$$

por tanto la función f puede ser representada de manera única como

$$f = \sum_{j=-\infty}^{+\infty} P_{W_j}f = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \quad (2-59)$$

lo que indica que la descomposición de la señal se puede ver como una agregación de detalles en todas las escalas.

2.4.1. Implementación del Análisis Multiresolución

En la sección anterior se ha obtenido la función scaling de forma recursiva para el nivel particular $j = 1$, ahora se va a obtener la ecuación para cualquier nivel j y tiempo m . De la teoría de multiresolución se conoce que la función scaling del espacio V_j se puede obtener como una combinación lineal de la base ortonormal del espacio V_{j-1} , es decir

$$\phi_{j,m}(t) = \sum_{n=-\infty}^{+\infty} \langle \phi_{j,m}(t), \phi_{j-1,n}(t) \rangle \phi_{j-1,n}(t) \quad (2-60)$$

donde $\phi_{j,m}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t-2^j m}{2^j}\right)$. Si se expresa el producto interno como integral y se realiza el cambio de variable $t' = (t - 2^j m)2^{-j-1}$, se obtiene

$$\langle \phi_{j,m}(t), \phi_{j-1,n}(t) \rangle = h[n - 2m] \quad (2-61)$$

por tanto, la ecuación recursiva de funciones scaling para cualquier nivel j , es dada por

$$\phi_{j,m}(t) = \sum_{n=-\infty}^{+\infty} h[n - 2m] \phi_{j-1,n}(t). \quad (2-62)$$

De igual forma ocurre con las funciones wavelet, ya que el espacio $W_j \subset V_{j-1}$, es posible expresarlas como una combinación lineal de la base ortonormal de V_{j-1}

$$\psi_{j,m}(t) = \sum_{n=-\infty}^{+\infty} \langle \psi_{j,m}(t), \phi_{j-1,n}(t) \rangle \phi_{j-1,n}(t) \quad (2-63)$$

desarrollando el producto interno,

$$\langle \psi_{j,m}(t), \phi_{j-1,n}(t) \rangle = g[n - 2m] \quad (2-64)$$

se obtiene la ecuación recursiva de las funciones wavelet para cualquier nivel j

$$\psi_{j,m}(t) = \sum_{n=-\infty}^{+\infty} g[n - 2m] \phi_{j-1,n}(t) \quad (2-65)$$

ahora, si se toma el producto interno en ambas ecuaciones (2-62) y (2-65), se tendría

$$\langle f, \phi_{j,m}(t) \rangle = \sum_{n=-\infty}^{+\infty} h[n - 2m] \langle f, \phi_{j-1,n}(t) \rangle \quad (2-66)$$

$$\langle f, \psi_{j,m}(t) \rangle = \sum_{n=-\infty}^{+\infty} g[n - 2m] \langle f, \phi_{j-1,n}(t) \rangle \quad (2-67)$$

y definamos $a_j[m] = \langle f, \phi_{j,n}(t) \rangle$ el coeficiente de aproximación en el nivel j , y $d_j[m] = \langle f, \psi_{j,m}(t) \rangle$ el coeficiente wavelet del nivel j , entonces se tendrían los *Coefficientes de Descomposición* (Mallat, 2008):

$$a_j[m] = \sum_{n=-\infty}^{+\infty} h[n - 2m] a_{j-1}[n] \quad (2-68)$$

$$d_j[m] = \sum_{n=-\infty}^{+\infty} g[n - 2m] a_{j-1}[n] \quad (2-69)$$

si se parte del nivel $j = 0$, lo cual es válido dado que cualquier espacio V_j mediante escalonamientos puede ser obtenido de V_0 , entonces a partir del coeficiente scaling a_0 podemos obtener tanto a_1 y d_1 , luego a partir de a_1 se obtienen a_2 y d_2 y así sucesivamente hasta el nivel finito j deseado, es decir las aproximaciones se subdividen sucesivamente en aproximaciones y detalles, que corresponde al proceso de *análisis* (Mallat, 2008), además el coeficiente a_0 puede ser reescrito como la sucesión de detalles d_1, \dots, d_j y una aproximación a_j , tal como se puede apreciar en la Figura 2-5.

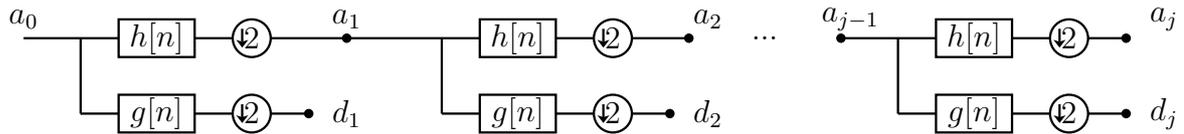


Figura 2-5: Banco de filtros del proceso de análisis.

En la Figura 2-5 aparece un factor de submuestreo por 2, este se debe a que el coeficiente a_0 es convolucionado con el filtro $h[n]$ y luego se mantienen solo las muestras par o impar (dependiendo de m), lo cual corresponde a un proceso de submuestreo. Ahora se debe obtener una expresión que permita a partir de los coeficientes a_j y d_j obtener el coeficiente a_{j-1} , que no es otra cosa que el proceso inverso o de *síntesis* (Mallat, 2008).

Se conoce que $V_{j-1} = V_j \oplus W_j$, entonces las funciones $\phi_{j-1,m}(t)$ se puede obtener como una combinación lineal de las bases

$$\phi_{j-1,m}(t) = \sum_{n=-\infty}^{+\infty} \langle \phi_{j-1,m}(t), \phi_{j,n}(t) \rangle \phi_{j,n}(t) + \sum_{n=-\infty}^{+\infty} \langle \phi_{j-1,m}(t), \psi_{j,n}(t) \rangle \psi_{j,n}(t) \quad (2-70)$$

recurriendo a (2-61) y (2-64) se obtiene

$$\phi_{j-1,m}(t) = \sum_{n=-\infty}^{+\infty} h[m-2n] \phi_{j,n}(t) + \sum_{n=-\infty}^{+\infty} g[m-2n] \psi_{j,n}(t) \quad (2-71)$$

y lo que queda es tomar el producto interno con la función f

$$\langle f, \phi_{j-1,m}(t) \rangle = \sum_{n=-\infty}^{+\infty} h[m-2n] \langle f, \phi_{j,n}(t) \rangle + \sum_{n=-\infty}^{+\infty} g[m-2n] \langle f, \psi_{j,n}(t) \rangle \quad (2-72)$$

reemplazando por coeficientes, se tiene la ecuación de reconstrucción

$$a_{j-1}[m] = \sum_{n=-\infty}^{+\infty} h[m-2n] a_j[n] + \sum_{n=-\infty}^{+\infty} g[m-2n] d_j[n] \quad (2-73)$$

entonces el coeficiente de aproximación a_{j-1} es obtenido mediante un sobremuestreo por 2, de los coeficientes a_j y d_j como se aprecia en la Figura 2-6

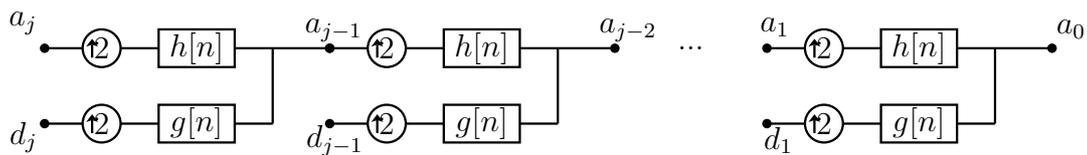


Figura 2-6: Banco de filtros del proceso de síntesis.

2.5. Wavelet Packets

Para aplicaciones en las cuales las señales tienen características que varían con el tiempo, como por ejemplo la música, o cuando se desea poder realizar clasificaciones, es necesario desarrollar una transformación que fuera de alguna manera adaptativa. Así surgió la teoría de funciones *Wavelet Packet* (Coifman et al., 1992), que es una generalización del análisis multiresolución, donde no se limita a la descomposición de los espacios de aproximaciones V_j , sino que permite obtener también una descomposición de los espacios de detalles W_j , por tanto se debe cambiar la nomenclatura utilizada para representar los espacios, supóngase que la escala de aproximación inicial es k , entonces en la raíz del árbol se asocia el espacio $V_k = W_{k,0}$ por tanto $\phi_k = \psi_{k,0}$, y es posible asegurar por el teorema de Coifman et al. (1992) que cualquier nodo hijo (j, l) donde j representa el nivel de profundidad y l el número de nodos a la izquierda, se puede descomponer como

$$\psi_{j+1,2l}(t) = \sum_{n=-\infty}^{+\infty} h[n]\psi_{j,l}(t - 2^j n) \quad \psi_{j+1,2l+1}(t) = \sum_{n=-\infty}^{+\infty} g[n]\psi_{j,l}(t - 2^j n) \quad (2-74)$$

donde las familias $\{\psi_{j+1,2l}\}$ y $\{\psi_{j+1,2l+1}\}$ constituyen bases ortonormales para los espacios $W_{j+1,2l}$ y $W_{j+1,2l+1}$ respectivamente. La subdivisión recursiva permite construir el árbol binario de la Figura 2-7.

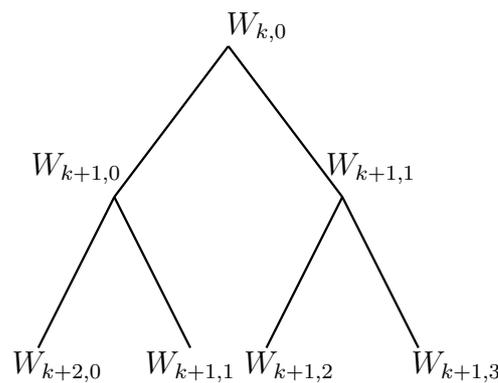


Figura 2-7: Árbol binario wavelet packet.

La teoría wavelet packet permite construir una librería de funciones de la cual varias bases pueden ser extraídas, y la adaptabilidad se obtiene al poder encontrar una base

wavelet packet que ofrezca una máxima separación de una característica basada en un criterio, el cual se conoce como *Función Costo*. Esta nueva descomposición se puede resumir en un árbol binario, el cual debe cumplir que cada nodo solo puede tener dos o ningún hijo nodo, Figura 2-8, y es de gran importancia estas funciones porque permiten construir diferentes árboles, que en el dominio de frecuencia se ven como diferentes descomposiciones en frecuencia, tal como se aprecia en la Figura 2-8

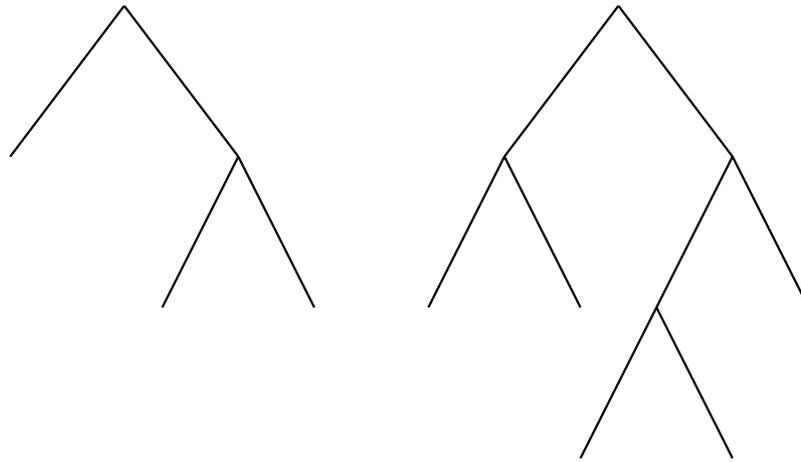


Figura 2-8: Árboles binarios wavelet packet.

Capítulo 3

Procesamiento Cuántico de Señales y su aplicación en la teoría wavelet

En el presente capítulo se da una introducción a los temas más relevantes del procesamiento cuántico de señales (*Quantum Signal Processing - QSP*), y se presentan los aportes del proyecto de grado que corresponden a aplicaciones del QSP en la teoría wavelet, que demuestran la utilidad del marco QSP para derivar nuevos métodos de procesamiento al expresar la teoría wavelet en términos de una o varias mediciones QSP. Entre estas aplicaciones desarrolladas se tiene: una nueva forma de descomposición wavelet al recurrir a nuevos mapeos QSP, uso del cuantizador QSP en los coeficientes wavelet, árbol wavelet packet adaptativo en sistemas de comunicación (Hoyos y Jojoa, 2015), transformación MMSE en los coeficientes wavelet, y el planteamiento de un nuevo *frame* denominado wavelets uniformes geoméricamente. Las aplicaciones de cuantización y transformación MMSE serán exploradas en los Capítulos 4 y 5.

3.1. Procesamiento Cuántico de Señales

El procesamiento cuántico de señales es un marco de referencia desarrollado por Eldar (2001), el cual recurre al uso de algunos principios y formalismos de la mecánica cuántica para replantear el procesamiento clásico de señales y así poder crear o modificar diferentes algoritmos. En la mecánica cuántica existen tres principios fundamentales que son *Medición*, *Consistencia de la medición* y *Cuantización de la medición*, por tanto el QSP debe realizar la definición de ellos.

Medición QSP: Consiste en aplicar un algoritmo a una señal (Eldar, 2001). Sea la señal de entrada x que pertenece a algún espacio de Hilbert H , a la cual se le desea aplicar una medición QSP M para producir una señal de salida $y = M(x) \in H$. Si la señal x no es adecuada para M entonces es posible construir una transformación que realice un mapeo del espacio de la señal X al espacio de Hilbert H , $T_x : X \rightarrow H$ de tal manera que la nueva señal $\check{x} = T_x(x)$ sea adecuada para la medición M . Finalmente a la salida se le aplica el mapeo inverso para obtener la señal en el mismo espacio

$T_y : H \rightarrow Y$, tal como se puede apreciar en la Figura 3-1.



Figura 3-1: Esquema general de la Medición QSP (Eldar, 2001).

Consistencia de la Medición: La señal y obtenida después de una medición debe ser tal que al realizar de nuevo el proceso, es decir, realizar una nueva medición sobre y , se obtendrá el resultado inicial, y (Eldar, 2001). Si M representa el operador de medición, y x la señal el objeto a medir, se tiene:

$$M(y) = y \quad (3-1)$$

donde $y = M(x)$.

Cuantización de la Medición: Se exige, para todas las mediciones M , que la señal resultante de una medición debe pertenecer al conjunto de señales determinadas de M (Eldar, 2001).

En el análisis realizado en el marco de referencia QSP, se desarrollaron los siguientes 3 tipos de medición para el diseño y/o modificación de algoritmos:

1. Medición de rango uno (Rank-one Measurement - ROM) se definen por vectores de medición.
2. Medición de subespacios (Subspace Measurements - SM) definidas por proyecciones.
3. Medición simple de subespacios (Simple Subspace Measurements - SSM) son proyecciones lineales.

3.2. Medición QSP rango uno

Una medición de rango uno o *Rank-one Measurement* (ROM) (Eldar, 2001), es en general un mapeo no-lineal (probabilístico) entre H y el conjunto de señales determinadas de M , descrito por un conjunto de vectores de medición $\{q_i \in H, i \in I\}$, donde I representa un conjunto índice. En la mecánica cuántica estos vectores de medición q_i deben ser ortonormales y por tanto linealmente independientes, pero parte de la gran utilidad que presenta este nuevo marco es la no limitación física presente en los sistemas cuánticos, es decir; las mediciones no están limitadas a las condiciones físicas presentes en los sistemas cuánticos, por lo que los vectores q_i se pueden asumir que

no son múltiplos de otro, ni que sean ortonormales, entonces los espacios unidimensional S_i generados por los vectores de medición q_i son disjuntos.

Los *subespacios de medición* S_i (Eldar, 2001), son los subespacios unidimensional que contienen todos los vectores que son múltiplos de q_i , así las *señales determinadas* son las señales que caen completamente en cada subespacio ($x \in S_i$), y por tanto son todas las señales de la forma aq_i , que se denotan como X_M . P_i representa la proyección sobre el espacio S_i , entonces la medición M se define como

$$M(x) = \begin{cases} P_i x = x & \text{si } x \in X_M \\ P_i x, \text{ con } i = f_M(\{\langle x, q_k \rangle\}) & \text{si } x \notin X_M \end{cases} \quad (3-2)$$

donde f_M es un mapeo entre la señal de entrada y el conjunto de índices, que solo depende del producto interno de la señal x con los vectores q_i , y que permite que la señal no determinada x sea cuantizada en una de las señales determinadas $P_i x$. Como ejemplos se tendrían $f_M = \arg \max_{k \in I} \langle x, q_k \rangle$, o $f_M = \arg \max_{k \in I} (\langle x, q_k \rangle - \langle q_k, q_k \rangle)$.

Se puede observar que la medición QSP genera algoritmos los cuales son no-lineales cuando las señales no son determinadas, pero que son caracterizados por operadores de proyección lineales, es decir; en última instancia la medición QSP se puede ver como un operador P_i (una proyección sobre el espacio unidimensional S_i), que actúa sobre una señal de entrada x , en donde el operador P_i es escogido de un conjunto de operadores $\{P_i, i \in I\}$ mediante alguna regla dada por el mapeo f_M . Por tanto, en la medición QSP siempre se va a tener que la señal de salida es una señal determinada de M .

Dada que una ROM en realidad depende de la elección del mapeo probabilístico f_M , de los vectores de medición q_i y de la proyección P_i , el marco QSP provee un método sistemático que permite modificar o crear nuevos e interesantes algoritmos de procesamiento de señales o imágenes, ya que al no estar sujeto a las limitantes presentes en la física cuántica, es posible utilizar nuevas restricciones que llevan al desarrollo de los algoritmos. Por ejemplo, el mapeo f_M puede utilizar mapeos probabilísticos o determinísticos, abriendo un abanico mas amplio de posibilidades. También los vectores de medida q_i ya no están como en la mecánica cuántica limitados a ser ortonormales y las proyecciones P_i tampoco deben ser solo ortogonales.

3.2.1. Diseño de algoritmos mediante el uso de ROM

Al usar una medición ROM se presentan dos posibilidades (Eldar, 2001):

1. Aplicar la medición ROM sobre la señal.
2. Usar algunos parámetros de la medición ROM, pero sin aplicar directamente la medición ROM sobre la señal.

a continuación se entrará en mas detalle sobre estas posibilidades.

Diseño de algoritmos con aplicación de medición ROM

El paso clave en el diseño de un nuevo algoritmo es la definición o identificación de los vectores de medición q_i , ya que estos a su vez van a definir las posibles salidas del algoritmo, y luego se debe asociar un espacio de Hilbert H a estos vectores. Si la señal de entrada x no pertenece al espacio se puede utilizar una transformación T_x para obtener una representación $\check{x} \in H$, entonces ya es posible usar la medición ROM sobre la representación de la señal para obtener la salida del algoritmo $\check{y} = M(\check{x})$, la cual si la señal de entrada es una señal determinada será igual a la representación $\check{y} = M(\check{x}) = \check{x}$ y si por el contrario no lo es, la salida será una señal determinada dada por $\check{y} = P_i x$ que dependerá del mapeo f_M . Por último si es necesario se utilizaría una transformación T_y para que la señal de salida y esté en el mismo espacio vectorial que x , lo cual se puede apreciar en la Figura 3-1. En conclusión, el marco de referencia QSP provee las transformaciones, los vectores de medición y los operadores de proyección para el diseño de nuevos algoritmos.

De manera interesante se logra apreciar como la teoría Wavelet encaja perfectamente en el marco QSP, gracias en parte a que la transformación wavelet se puede ver como una proyección de la señal en el dominio wavelet, y por tanto actúa como un operador sobre la señal x , que lleva a aplicaciones en diferentes campos de la ingeniería como por ejemplo la transmisión, codificación y análisis de la señal. El marco QSP provee nuevos caminos para modificar no solo la teoría wavelet sino cualquier algoritmo que pueda ser expresado como en la Figura 3-1. Por ejemplo, además de utilizar mapeos determinísticos o probabilísticos, también es posible recurrir a mediciones combinadas o se puede definir nuevas limitantes a los vectores de medición, una de ellas puede ser en el producto interno utilizado.

En este orden de ideas, en este trabajo se diseñaron las siguientes aplicaciones del marco QSP en la teoría wavelet.

En la siguiente aplicación (Ejemplo 1), se muestra como puede ser expresado mediante el concepto de la medición ROM, los coeficientes resultantes del proceso de análisis de una señal al utilizar la descomposición wavelet packet.

Ejemplo 1 Descomposición Wavelet: En la descomposición de una señal x se utiliza el proceso de análisis; en el cual se utiliza la transformada wavelet para obtener los coeficientes a_j , y cuando la señal ha sido contaminada con ruido blanco aditivo es común utilizar una función costo para su eliminación. En el proceso de análisis o Transformada Wavelet discreta se utilizan las funciones wavelet $\psi(t)$, que no hacen más que realizar un proceso de correlación con la función x a través del producto interno. Para implementar la medición M , primero se debe definir los vectores de medición

$\{q_i(t) = \psi_i(t)\}$, y luego se define el mapeo f_M como

$$f_M(x(t)) = \arg \max_{j \in J} \{ \langle x(t), \psi_j(t) \rangle \} = i \quad (3-3)$$

el cual se encarga de seleccionar la i -ésima wavelet de un conjunto J que mejor se aproxima a la señal. Iterando este proceso hasta cumplir algún criterio, por ejemplo, un umbral de porcentaje de la energía, permitiría descartar ciertas wavelet que no den mayor información. Se observa como este mapeo probabilístico se postula a ser una muy buena aplicación en la Teoría Wavelet, sometiendo los coeficientes de análisis a este mapeo f_M se obtendría una representación reducida similar a un proceso de compresión:

$$\begin{aligned} x &= f_M \left\{ \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle x, \psi_{j,n} \rangle \psi_{j,n} \right\} \\ \Rightarrow x &= \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} f_M \{ \langle x, \psi_{j,n} \rangle \} \psi_{j,n}. \end{aligned} \quad (3-4)$$

De la selección del mapeo se pueden desarrollar nuevos e interesantes aportes a la teoría wavelet. El mapeo f_M puede ser completamente aleatorio; seleccionando solo ciertos coeficientes de acuerdo a cierta distribución, a partir de los cuales se buscaría poder recuperar la señal original. Este tipo de descomposición abre un camino muy interesante ya que permite explorar que bases wavelet optimizan el proceso de recuperación respecto alguna norma, y mas aún, determinar la probabilidad con que se puede recuperar la señal original. Por tanto, se tendría nuevos caminos de investigación.

Además, otros ejemplos que el mapeo f_M podría representar son un nuevo cuantizador o un estimador de parámetros, que podrían aportar en el proceso de análisis de una función.

Ejemplo 2 Modulación Wavelet: En la transmisión mediante señales wavelet, una parte esencial es la recepción; la cual se realiza mediante filtros sintonizados o *Matched Filter* (MF). En el marco QSP se desarrolló un nuevo MF llamado filtro sintonizado ortogonal (*Orthogonal matched filter* - OMF) (Eldar y Oppenheim, 2001) al imponer una restricción en el producto interno de los vectores de medición, que da origen a un receptor compuesto por un banco de correladores con señales que están sintonizadas y con una estructura definida, cercanas en un sentido de mínimos cuadrados o *Least Square* (LS) a las señales transmitidas, presentando un comportamiento mejor cuando el ruido no es gaussiano. De manera interesante, en (Eldar y Oppenheim, 2001) se demostró que este receptor puede ser implementado como un MF seguido por una transformación de covarianza, que optimiza

la correlación de las salidas del MF antes de la detección, por tanto, podría ser aplicado en un sistema de transmisión que utiliza las funciones wavelet, buscando mejorar su desempeño ante ruidos aditivos distintos al gaussiano.

Ejemplo 3 Compresión: En la conversión de una señal analógica f , esta es discretizada en una señal $f[n]$ de tamaño N , que luego se descompone en una base ortonormal wavelet:

$$f = \sum_{k \in T} \langle f, \psi_k \rangle \psi_k \quad (3-5)$$

Los coeficientes $\langle f, \psi_k \rangle$ son aproximados por valores cuantizados $Q(\langle f, \psi_k \rangle)$. Al mapear a la medición QSP se puede recurrir al uso del cuantizador QSP (Eldar, 2001), que se describe a continuación.

Ejemplo 4 Cuantizador Probabilístico: Se puede decir que cuantización es un proceso que transforma la entrada en una salida digital, la cual se toma de un conjunto finito de posibles salidas. Se puede hacer elemento a elemento (cuantización escalar) o se puede hacer a cada vector (cuantización vectorial) (Gersho y Gray, 1991). Un cuantizador realiza una partición en n niveles $\{y_i, 1 \leq i \leq n\}$, y cada entrada de x es mapeada a un nivel mediante una función $y_i = Q(x)$ donde $i = \arg \min |x - y_j|$, es decir; se está escogiendo el nivel más cercano al valor de entrada, lo cual es conocido como la regla de la vecindad más cercana.

Para realizar la descripción del cuantizador como una medición QSP es necesario la definición de los vectores de medición $q_i, 1 \leq i \leq n$, que en este caso corresponderían a los vectores ortonormales del mapeo $T : R \rightarrow R^n$, que asocia a cada uno de los niveles y_i un vector en $q_i \in R^n$. Si la entrada del mapeo corresponde a un nivel y_i entonces se tiene $T(y_i) = q_i$, pero si la entrada es otro valor $x \neq y_i$, entonces $T(x) = z$, donde $z = \sum_{k=1}^N q_k / (x - y_k)$. Recurriendo a la ecuación (3-4) de descomposición de la función y reemplazando el mapeo por Q , se tiene

$$\hat{f} = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} Q(\langle f, \psi_{j,n} \rangle) \psi_{j,n} \quad (3-6)$$

donde \hat{f} representa una aproximación de la señal f , y los coeficientes son cuantizados por el cuantificador probabilístico Q desarrollado por Eldar (2001). Esta aplicación de QSP en la teoría wavelet será explorada más en profundidad en el Capítulo 4.

Uso de parámetros para el diseño de Algoritmos

El marco QSP ofrece otra posibilidad en la creación de nuevos algoritmos, que consiste en recurrir al uso de los parámetros de la medición ROM para procesar los datos de

interés. Debido a que cualquier algoritmo lineal puede ser expresado como un proceso realizado sobre los datos por un conjunto de vectores de medición, esta perspectiva facilita la adopción de limitantes de la mecánica cuántica que aporten en la búsqueda de algoritmos más óptimos (Eldar, 2001).

Si recurrimos al uso de operadores lineales, la transformada Wavelet la podemos expresar como un operador lineal de análisis $\Phi^* : H \rightarrow \ell^2(\mathbb{Z})$, el cual es posible expresarlo de forma matricial, donde las filas de la matriz son las funciones ψ_i , y el vector resultante sería igual a $y = \Phi^*x$. Este nuevo punto de vista facilita el uso de los parámetros aportados por el QSP para su modificación, por ejemplo, establecer nuevas condiciones sobre la matriz de covarianza y derivar algoritmos óptimos sujetos a tal condición. De esta manera el QSP extiende el concepto de mínimos cuadrados o LS de la mecánica cuántica para desarrollar algoritmos que minimicen el criterio MSE sujeto a condiciones sobre la matriz de covarianza.

La siguiente aplicación hace uso de este nuevo punto de vista.

Ejemplo 5 Transformación MMSE: Cuando se tiene un vector aleatorio $x \in \mathbb{C}^n$ y se desea utilizar alguna transformación W que permita obtener un vector $y = Wx$ de-correlacionado, es decir; la matriz de covarianza original C_x se convierta en una matriz identidad de la forma $C_y = a^2 I_n$, para alguna constante a . La transformación debe ser tal que

$$\begin{aligned} C_y &= \mathbb{E}\{yy^*\} = \mathbb{E}\{Wx(Wx)^*\} \\ C_y &= W\mathbb{E}\{xx^*\}W^* \\ C_y &= WC_xW^* = a^2 I_n \end{aligned}$$

Para la selección de la transformación blanca se presentan diferentes opciones, pero generalmente el vector y no es sensible a los cambios del vector de entrada x , lo cual no es muy deseable en estimadores. Teniendo en cuenta lo anterior y recurriendo a ideas de la detección cuántica, el marco QSP propuso una transformación blanca óptima donde el vector resultante y es tan cercano a x sujeto el criterio MSE, llamada *Transformación blanca MMSE* (Eldar, 2003). En el Capítulo 5 se explorará la aplicación en la teoría wavelet.

En la siguiente aplicación el estimador desarrollado por el marco QSP al recurrir a ideas de la detección cuántica e ideas de LS, es utilizado en la teoría wavelet con el objetivo de obtener un algoritmo lineal sujeto a restricciones estocásticas en el producto interno.

Ejemplo 6 Estimador CSLS: En los dispositivos de detección digital, como cámaras o micrófonos, se presenta el problema que sus mediciones están contaminadas por ruido, las cuales pueden ser modeladas como:

$$y[n] = a[n] + w[n] \quad (3-7)$$

donde $a[n]$ por ejemplo podría representar los coeficientes de análisis en una descomposición wavelet $a[n] = \Phi^* x$, y donde $w[n]$ describe el ruido intrínseco a la señal con matriz de covarianza C_w , por ejemplo en situaciones donde la intensidad de la luz es baja y se toma una fotografía, además también el ruido generado por el propio dispositivo. A menudo el ruido w es modelado como un vector aleatorio con una distribución de probabilidad Gaussiana. La idea es lograr obtener un estimador D tal que el vector estimado $\hat{x}_{LS} = Dy$ de x , permita obtener un vector estimado $\hat{y} = \Phi^* \hat{x}_{LS}$ el cual este lo más cercano posible al vector original de medidas y . Entonces se busca minimizar

$$\epsilon_{LS} = (y - \Phi^* \hat{x}_{LS})^* C_w^{-1} (y - \Phi^* \hat{x}_{LS}) \quad (3-8)$$

pero en muchos casos se vuelve a presentar el problema que el vector y no varía al ritmo que lo hace x , lo que originaría un bajo desempeño del estimador LS.

Para mejorar el desempeño del estimador, el marco QSP propuso una modificación basada en ideas desarrolladas en la detección cuántica, donde se tiene un control en la forma espectral y el rango dinámico de la covarianza del error, mediante la selección de un estimado de x que minimice el error de la varianza sujeto a restricciones de la covarianza del error. Al estimador se le llama estimador LS con conformación de covarianza o *Covariance Shaping Least Square* (CSLS) (Eldar y Oppenheim, 2003). Así, $\hat{x}_{CSLS} = Dy$ es escogido para minimizar

$$\epsilon = \mathbb{E}((y' - \Phi^* D \hat{y}')^* C_w^{-1} (y' - \Phi^* D \hat{y}')) \quad (3-9)$$

donde $y' = y - \mathbb{E}\{y\}$, y ya que la covarianza del error en el estimado \hat{x}_{CSLS} es igual a la covarianza del estimado \hat{x}_{CSLS} , se le impuso la restricción que debe ser proporcional a una matriz de covarianza R . Por tanto, el estimador D debe satisfacer

$$DC_w D^* = a^2 R \quad (3-10)$$

donde $a > 0$ es una constante que puede ser especificada o escogida tal que minimice el error. La aplicación del estimador en la teoría wavelet queda propuesta para ser desarrollada en trabajos futuros.

3.3. Medición de Subespacio (SM) y SSM

Una medición de subespacio o *Subspace Measurements* (SM) (Eldar, 2001) pretende determinar cual subespacio esta más adaptado de acuerdo a algún criterio a la señal que se esta midiendo. Si M^S es un SM definido sobre el espacio H , entonces M^S es mapeo no lineal descrito por un conjunto de operadores de proyección $\{P_i, i \in I\}$,

donde P_i es una proyección sobre un subespacio $S_i \subseteq H$, y los subespacios solo se interceptan en \emptyset . Pero a diferencia de la mecánica cuántica, tales operadores de proyección no están limitados a ser ortogonales, permitiendo mayor grado de libertad en su diseño.

Las señales determinadas $x \in H$ son similares a las definidas para una ROM, es decir; si $x \in S_i$ entonces la señal x es una señal determinada de M^S , y el conjunto de todas las señales determinadas es denotado por X_M^S . Un SM es un mapeo probabilístico entre H y X_M^S con las siguientes características:

- Es un mapeo identidad si la señal es determinística, esto es, si $x \in S_i$, entonces $M^S(x) = P_i x = x$
- Es un mapeo probabilístico si $x \notin S_i$, entonces $M^S(x) = P_i x, i \in I$ donde $i = f_M^S(\{\langle P_k x, P_k x \rangle, k \in I\})^1$

y $f_M^S : H \rightarrow I$ al igual que las ROM, define un mapeo probabilístico entre las señales y los índices, y que depende solo del producto interno. Por ejemplo, se podría escoger un mapeo que seleccione el índice con mayor norma euclidiana, es decir

$$f_M^S(x) = \arg \max_{k \in I} \|P_k x\|^2 \quad (3-11)$$

Se observa como una medición SM proyecta la señal sobre un subespacio de acuerdo a algún criterio (por ejemplo una norma), y se puede ver como selecciona el subespacio que es más parecido es la señal de entrada. De manera resumida se tiene

$$M^S(x) = \begin{cases} x & \text{si } x \in X_M^S \\ P_i x, \text{ con } i = f_M^S(\{\langle P_k x, P_k x \rangle\}) & \text{si } x \notin X_M^S \end{cases} \quad (3-12)$$

Un caso especial parece cuando la medición M^S es definida por un único operador de proyección P , entonces $M^S(x) = Px$, y es llamada *Simple Subspace Measurements* (SSM) (Eldar, 2001).

A continuación se da una aplicación de SM en la teoría wavelet.

Ejemplo 7 Árbol adaptativo Wavelet: En la representación de una señal x mediante los arboles wavelet, la selección de las bandas pasa bajas y pasa altas se realiza de acuerdo a un criterio, el más conocido ha sido propuesto por Coifman y Wickerhauser (1992), donde el criterio utilizado fue la entropía. Pero, el uso del QSP abre un nuevo camino de utilizar nuevos criterios asociados a un mapeo que facilite la extracción de mayor información de la señal de interés, o que permita una descomposición más óptima. Con esta idea se llegó al diseño de un sistema de multiplexación por división de paquetes dinámica (MDPWD) (Hoyos y Jojoa, 2015), el cual parte

¹es importante notar que $\langle P_k x, P_k x \rangle = \|P_k x\|^2$

de los resultados obtenidos en (Hoyos y Jojoa, 2014), donde se demostró la mejoría al utilizar arboles asimétricos, para desarrollar un nuevo esquema que busca realizar una asignación dinámica de las wavelet packet que funcionan como ondas portadoras, con el objetivo de reducir la localización en frecuencia y por tanto reducir el nivel de interferencia presente en la comunicación.

El esquema propuesto MDPWD (Figura 3-2) consta de un selector de funciones wavelet packet, que realiza una asignación dinámica de las subportadoras o funciones wavelet packet del espacio finito equiprobable de árboles binarios S a cada usuario en un instante de tiempo, sin tener a priori información del canal de comunicación, buscando disminuir la localización en frecuencia de las subportadoras.

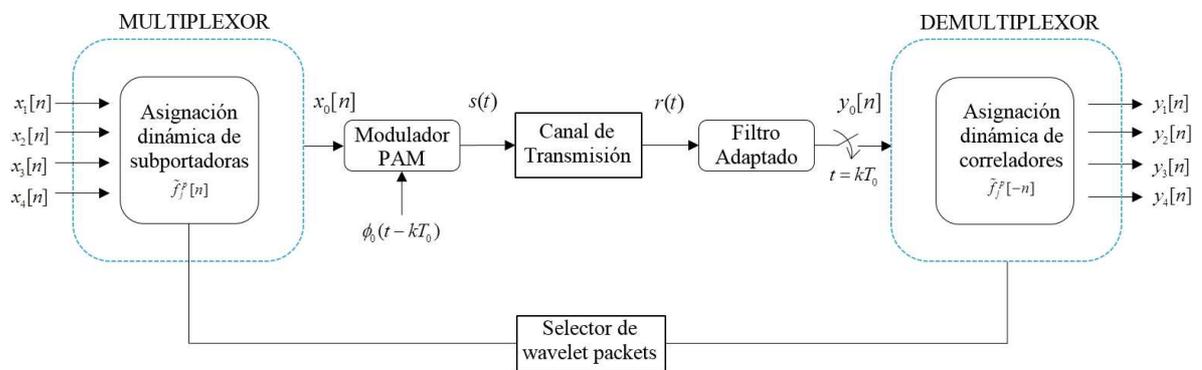


Figura 3-2: Diagrama en bloques del esquema propuesto MDPWD.

Para observar como este esquema disminuía el nivel de interferencia, se evaluó su desempeño en un canal con desvanecimiento selectivo en frecuencia dado por

$$y(t) = r_1(t)s(t) + r_2(t)s(t - \tau) + n(t) \quad (3-13)$$

donde $s(t)$ es la señal de entrada al canal, τ corresponde al retraso, $n(t)$ representa el ruido blanco o gaussiano y $r(t)$ son procesos gaussianos complejos con densidad espectral de potencia dad por

$$R(f) = \begin{cases} \frac{1}{\pi f_d \sqrt{1-(f/f_d)^2}} & |f| < f_d \\ 0, & \text{p.d.v} \end{cases} \quad (3-14)$$

con una frecuencia doppler máxima de $f_d = 60\text{Hz}$. Debido a que el desvanecimiento selectivo en frecuencia afecta en un instante de tiempo cierto rango de frecuencias, el esquema propuesto al variar la descomposición de subbandas mediante el cambio del árbol binario permite combatir el desvanecimiento selectivo al evitar que los usuarios utilicen las mismas frecuencias durante toda la transmisión, y por tanto sean menos afectados por las

efectos adversos de un canal de comunicaciones variante en el tiempo, como se observa en la Figura 3-3; donde se aprecia que el nuevo esquema supera el comportamiento respecto a la tasa de error binario o *Bit Error Rate* (BER), tanto del sistema de Multiplexación por División de Paquetes Wavelet (*Wavelet Packet Division Multiplexing* - WPDM), como del sistema de acceso múltiple-multiplexación por división wavelet ortogonal MA-OWDM (Liew et al., 2005).

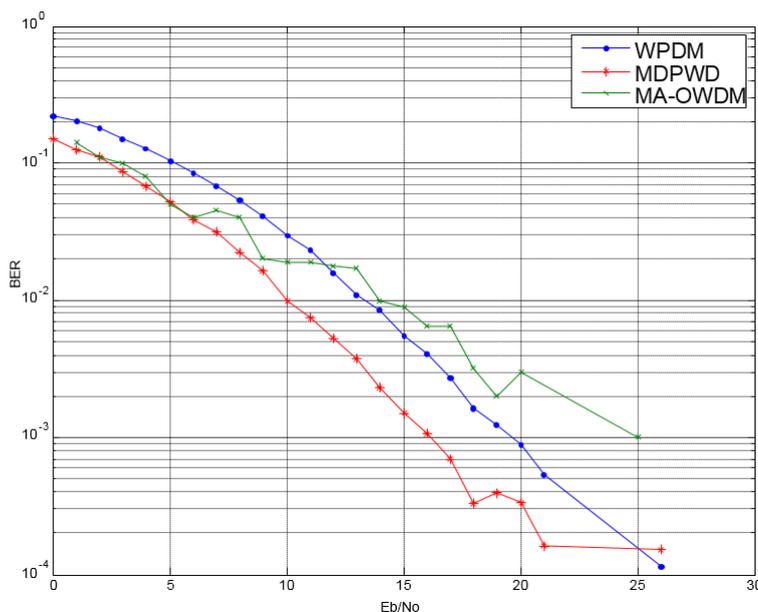


Figura 3-3: Comportamiento de la BER contra E_b/N_0 de los esquemas en un canal con desvanecimiento selectivo en frecuencia con una $f_d = 60\text{Hz}$ (Hoyos y Jojoa, 2015).

3.4. Combinación de Medidas QSP

En algunos casos se conoce de ante mano información sobre los sistemas cuánticos, como por ejemplo que el sistema cae completamente en uno o varios subespacios, lo cual llevaría a que los vectores de medición se diseñen para que estén limitados a tales subespacios. Esta idea es la que lleva a las nociones de Medidas con valores de operadores positivos o *Positive-Operator Valued Measure* (POVM) (Eldar, 2001) que buscan aplicar mediciones sucesivamente. Esta característica los hace muy deseables en aplicaciones donde se desea extraer más información de los datos. Una forma alternativa de entender que es un POVM, es recurrir al QSP y verlo como un ROM seguido de una proyección o simple subespacio de medida (SSM).

Recurriendo a la ventaja de no tener limitaciones físicas contrario a la mecánica cuántica, es posible incluir nuevas formas de combinar las mediciones que permiten ampliar

el abanico de aplicaciones. A continuación se describen las combinaciones de medidas.

3.4.1. Clases de Mediciones Combinadas

El marco QSP definió la combinación utilizando solo ROM y SSM. A la ROM se le asocia la medición M_1 y la SSM se le asocia la medición M_2 , al combinarlas; es decir al tener por ejemplo $M_{12} = M_1(M_2(x))$, se pueden tener cuatro combinaciones (Eldar, 2001):

1. ROM seguido de una SSM, M_{21}
2. SSM seguido de una ROM, M_{12}
3. ROM seguido de una ROM, $M_{11'}$
4. SSM seguido de una SSM, $M_{22'}$

ROM seguido de una SSM

En esta combinación de medidas se tiene que M_2 al ser una SSM realiza una proyección del resultado de la primer medida M_1 sobre el rango del operador E , entonces los posibles resultados de la medición serán proporcionales a la proyección de los vectores de medición utilizados en M_1 , y se llamaran *vectores de medición efectivos* (Eldar, 2001). El marco QSP mostró que tales vectores conforman un *frame* para el rango de E ($R(E)$). Dada la flexibilidad que aporta el marco QSP se pueden establecer restricciones sobre los vectores de M_1 . Si los vectores de medición de M_1 fueran las funciones wavelet ψ_i mediante las restricciones nombradas, que serían en últimas establecidas por el tipo de proyección E utilizadas, por ejemplo si fuese una proyección oblicua daría conexión a las wavelet oblicuas, observándose como el marco QSP ofrece una interesante manera de crear o conectar con las funciones existentes, las cuales se puede apreciar en la Figura 3-4.

Es de nuestro interés explorar en más detalle la combinación 1, ya que lleva al diseño de una estructura de vectores conocida como *Frames Uniformes Geométricamente* (Eldar y Bolcskei, 2003).

Frames Uniformes Geométricamente (UG): Un *frame* es una generalización de las bases (Duffin y Schaeffer, 1952), que permiten una representación redundante ya que los vectores tiene más libertad en su estructura y pueden ser linealmente dependientes, además pueden dar origen a representaciones escasas (Mallat, 2008). El número de vectores n del *frame* es mayor que la dimensión del espacio m , y el factor de redundancia viene dado por $f = n/m$. Un conjunto de vectores X es geoméricamente uniforme si, dados dos puntos x y x' en X , existe una isometría T que transforma x a

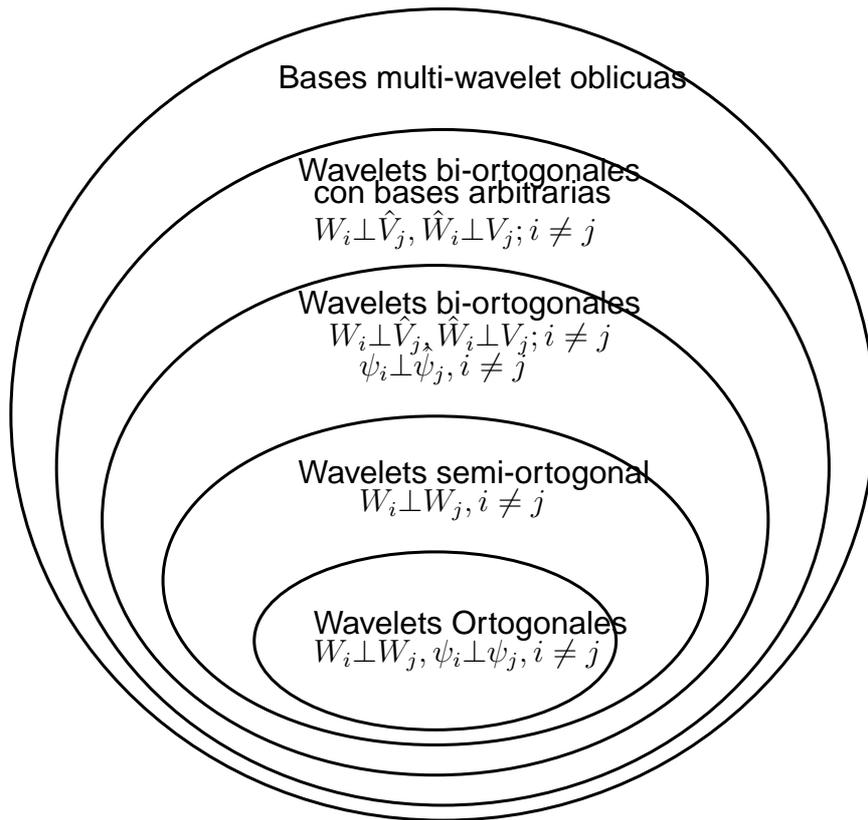


Figura 3-4: Clases de familias wavelets.

x' mientras deja X invariante:

$$\mathbf{T}(x) = x' \quad (3-15)$$

$$\mathbf{T}(X) = X \quad (3-16)$$

Recurriendo a los anteriores conceptos, el QSP logra desarrollar una nueva área de acción denominada Frames Uniformes Geométricamente (UG), que son vectores con propiedades simétricas muy fuertes (Eldar y Bolcskei, 2003). El marco QSP además de desarrollar su definición, logra la creación de un método para la construcción de frames uniformes geoméricamente. Se puede decir que la característica principal de estos vectores es ser generados a partir de una función ψ conocida como vector «generador», mediante una transformación $\psi_i = U_i\psi$, donde las transformaciones U_i son unitarias y forman un grupo abeliano Q^2 . Así, se observa que esta clase de *frame* tienen otra forma de construcción que es muy distinta a la construcción de wavelet, la cual se realiza mediante traslaciones (T) y dilataciones (D).

Además, un *frame* UG tiene la interesante propiedad que al verlos lucen iguales geoméricamente desde cualquier punto del conjunto, la cual es una característica muy importante para aplicaciones como en la codificación de canal (Forney, 1991) y codificación de descripción múltiple y optimización numérica de los mismos (Goyal et al., 2001),

²Es decir; Q contiene la transformación identidad, el inverso de cada U_i^{-1} , el producto $U_i U_j \in Q$ y $U_i U_j = U_j U_i$.

también permite reducir la complejidad computacional requerida en su ejecución. La importancia radica cuando por ejemplo, en una transmisión se han perdido paquetes, y se desea determinar la calidad de la reconstrucción; una forma es saber como se comporta el *frame* cuando uno o más de los vectores del *frame* se han removido, por tanto es importante ser capaz de estimar las cotas del *frame* reducido. Los *frame* UG tienen la propiedad que si los paquetes son perdidos con igual probabilidad, las cotas del *frame* se deterioran uniformemente sin importar que vector del *frame* ha sido removido.

Por otra parte se puede encontrar los frames Uniformes Geométricamente Compuestos (UGC) que es una extensión de frames UG, los cuales son generados por un grupo Abeliano finito usando *múltiples* vectores generadores (Eldar y Bolcskei, 2003). Es importante resaltar que estos vectores no son necesariamente UG, aunque si consisten de un subconjunto de vectores que si son UG. Ahora, si los vectores generadores del *frame* UGC son UG, y generados por un grupo Abeliano O , en (Eldar y Bolcskei, 2003) se demostró que este tipo de UGC son una generalización de los frames de Gabor los cuales son generados a partir de una función madre $g(t)$, la cual es sometida a modulaciones mediante un operador $Mg(t) = e^{2\pi i b t} g(t)$ y traslaciones $T_b g(t) = g(t - b)$, y donde Q es el grupo de traslaciones y O es el grupo de modulaciones.

Dado que en (Christensen y Goh, 2014) se ha logrado demostrar que es posible basados en ciertos dual frames de Gabor construir un dual frames Wavelet y viceversa, y aunque la redundancia se preserva en casos especiales; como por ejemplo wavelet ortonormal o shannon, esta conexión abre un nuevo camino ya que los UGC son una generalización de los frames Gabor, entonces se podría pensar en utilizar la relación para la construcción de una generalización de los *frame* wavelet. Esta relación entre tales *frames* se logra mediante el uso de una transformación que diseñaron y que tiene una característica especial, que consiste en la preservación de la distancia de los vectores del *frame*. La transformación consiste en asociar a una función $g \in \mathcal{L}^2(\mathbb{R})$ que presenta la propiedad $g(\log_\theta |\cdot|) \in \mathcal{L}^2(\mathbb{R})$, una función $\psi \in \mathcal{L}^2(\mathbb{R})$ tal que su transformada de Fourier venga dada por

$$\hat{\psi}(\alpha) = \begin{cases} g(\log_\theta |\alpha|) & \text{si } \alpha \neq 0 \\ 0, & \text{si } x = 0. \end{cases} \quad (3-17)$$

donde $\theta > 1$, ya que para los valores entre 0 y 1 la transformación no aporta mayor información. De forma generalizada tenemos que para cualquier valor $a > 0, j \in \mathbb{Z}$ y $\alpha \in \mathbb{R} - \{0\}$

$$\hat{\psi}(a^j \alpha) = g(j \log_\theta(a) + \log_\theta(|\alpha|)). \quad (3-18)$$

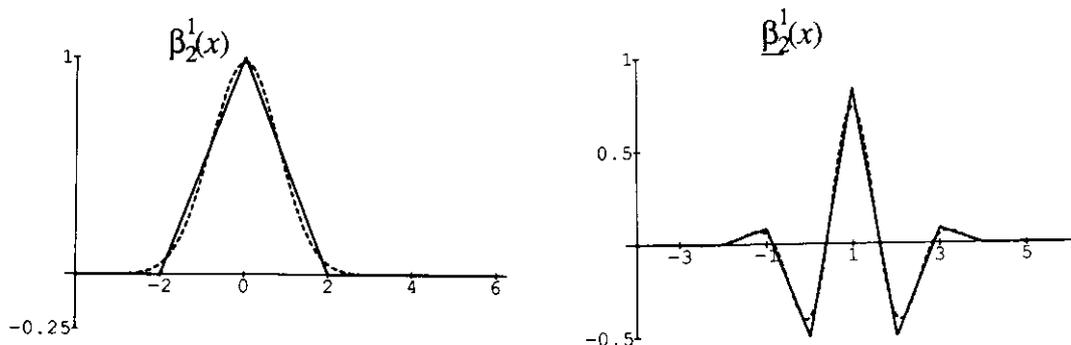
Para que la transformación pueda ser aplicada se debe satisfacer las condiciones presentes en el siguiente teorema (Christensen y Goh, 2014)

Teorema 3.1. (Gabor-Wavelet) Sean $b > 0, \alpha > 0$ y $\theta > 1$. Se asume que $g, \tilde{g} \in \mathcal{L}^2(\mathbb{R})$ son funciones acotadas con soporte en el intervalo $[C, D]$ para $C, D \in \mathbb{R}$ y

$\{E_{mb}T_{n\alpha}g\}_{m,n\in\mathbb{Z}}$ y $\{E_{mb}T_{n\alpha}\tilde{g}\}_{m,n\in\mathbb{Z}}$ forman un dual frame para $\mathcal{L}^2(\mathbb{R})$. Con $a := \theta^\alpha$, si $b \leq 1/(2\theta^N)$, entonces $\{D_{aj}T_{kb}\psi\}_{j,k\in\mathbb{Z}}$ y $\{D_{aj}T_{kb}\tilde{\psi}\}_{j,k\in\mathbb{Z}}$ son un dual frame para $\mathcal{L}^2(\mathbb{R})$.

y adicionalmente en (Unser et al., 1992) se prueba que ciertos B-spline wavelet convergen a funciones Gabor punto a punto y en todas las normas L_p con $p \in [1, +\infty)$, ratificando aún más la conexión mencionada. Un ejemplo de tales funciones se puede apreciar en la Figura 3-5. Ya que se tiene del marco QSP que bajo ciertas condiciones un *frame* UGC se puede ver como una generalización de los frames Gabor, y si se recurre a la anterior transformación se abre una muy interesante aplicación, la construcción de un *frame Wavelet UGC*.

n=1 (piecewise linear)



n=3 (cubic)

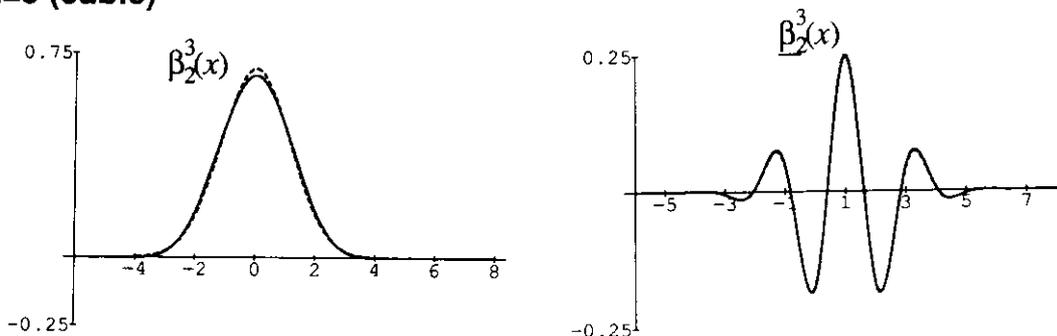


Figura 3-5: Ejemplo de funciones base B-spline y wavelet con su correspondientes aproximación Gabor (línea punteada) (Unser et al., 1992).

3.5. Observaciones finales

En este capítulo se ha sugerido ciertas aplicaciones del QSP en la teoría wavelet, que resultan de una exploración preliminar y quedan otras muchas por explorar. Se observó como el concepto de medición ha permitido llegar a nuevas ideas en cuantización, co-

rrelación y frames, que serán exploradas en los capítulos siguientes.

También se determinó de forma general la conexión entre los frames Gabor y Wavelet que permita la creación de un nuevo *frame* wavelet, denominado *Wavelets Uniformes Geométricamente* que podría ser muy útil en proceso de codificación de canal y codificación de descripción múltiple.

Capítulo 4

Aplicación del cuantizador QSP a la teoría Wavelet

Es de importancia resaltar que debido a la gran cantidad de aplicaciones del QSP en la teoría Wavelet, se optó por seleccionar aplicaciones que den un mayor aporte al estado del arte y que permitan en un futuro no muy lejano abrir nuevos temas de interés en la comunidad científica de la teoría wavelet. Por tanto, en este Capítulo se discutirá la aplicación del cuantizador QSP (Eldar, 2001), y se presenta como aporte del proyecto de grado la demostración del uso de la función de autocorrelación de las funciones wavelets ortogonales como mapeo para el cuantizador *dithered*, y una nueva forma de implementarla (Hoyos y Jojoa, 2016).

En el inicio del presente Capítulo se da una introducción al cuantizador probabilístico y cuantizador *dithered*, presentándose a continuación la relación directa entre ambos cuantizadores, en especial la relación con el cuantizador *dithered* sin memoria. Posteriormente se presenta como aporte del proyecto de grado la demostración del uso de la función de autocorrelación de las funciones wavelets ortogonales como mapeo para el cuantizador probabilístico, y una forma de implementar la función de autocorrelación recursivamente, muy similar a la forma de implementar las funciones wavelet, que aporta en la reducción de la complejidad computacional de las señales dither. Esta nueva forma de implementar un cuantizador *dithered* podría ser aplicado a los coeficientes de descomposición del proceso de análisis multiresolución wavelet, de forma similar a lo desarrollado por Kozaitis y Goswami (2000, 2001) que utilizaron el cuantizador *dithered* en imágenes.

4.1. Cuantizador probabilístico QSP

El proceso de conversión Análogo a Digital es integrado por dos importantes procesos: el primero consiste en el *Muestreo* que permite la adquisición de muestras de la señal en el tiempo, y el segundo la *Cuantización* de la amplitud de los valores, que asocia un código binario a cada muestra de la señal (Proakis, 2007).

En muchas aplicaciones se ha observado como el proceso de cuantización juega un rol muy importante, llevando al desarrollo de sistemas de cuantización cada vez más complejos (Goyal et al., 2001; Proakis, 2007; Mallat, 2008). Uno de ellos es el cuantizador *dithered*, el cual tiene propiedades muy interesantes como producir ruido de cuantización independiente de la fuente y reconstrucción continua en el lado del decodificador (S P Lipshitz, 1992). En (West y Scheets, 2012) se demostró como al utilizar una señal dither determinística se incrementa la resolución de un cuantizador uniforme, y en (Pamarti et al., 2007) se ha recurrido al uso de señales dither en un modulador digital sigma delta. Por otra parte Akyol y Rose (2009) propuso un cuantizador *dithered* no-uniforme el cual supera al cuantizador *dithered* uniforme y al cuantizador determinístico. A pesar de las bondades del cuantizador *dithered* una de sus desventajas es la complejidad computacional en la inyección de las señales dither, por eso Sanyal y Sun (2012) propone un nuevo método para añadir señales dither en cuantizadores vectoriales, que alcanza un mejor desempeño en la forma del desfase a un costo hardware bajo comparado con las existentes técnicas dither, y Anabtawi et al. (2014) propone una nueva técnica de inyección dither para los moduladores sigma delta. Además, la utilidad de este esquema de cuantización se ve limitada por la complejidad computacional en la generación de procesos aleatorios sujetos a una distribución de probabilidad conjunta (Wannamaker et al., 2000).

Teniendo en cuenta estas limitantes, bajo el marco QSP, Eldar (2001) desarrolló el cuantizador probabilístico QSP, el cual tiene una relación directa con el cuantizador *dithered*, y permite reducir su complejidad computacional.

Sea una señal x que toma valores de amplitud continuos (\mathbb{R}) en un instante de tiempo, entonces el proceso de cuantización se encarga de realizar un mapeo de este mundo analógico al digital, asociando al valor de amplitud un código binario. Es importante resaltar que una característica del proceso de cuantización es la degradación inevitable que sufre la señal (Proakis, 2007). De manera más precisa, la entrada al cuantizador es una secuencia que debe pertenecer al espacio $\ell^2(\mathbb{Z})$ (con el objetivo de brindar estabilidad), así, a cada valor de entrada x_n le corresponderá un valor cuantizado dado por $y_n = Q(x_n)$, donde Q selecciona uno de los niveles de cuantización de acuerdo a una regla definida, y la diferencia entre cada nivel es igual a Δ . Para el desarrollo del cuantizador QSP, se tomará como base el cuantizador escalar definido por:

$$Q(x) = \Delta \left\lfloor \frac{x}{\Delta} + \frac{1}{2} \right\rfloor \quad (4-1)$$

donde el operador $\lfloor \cdot \rfloor$ regresa el entero más grande que es menor o igual a su argumento. El paso Δ es comúnmente referido como el bit menos significativo o *Least Significant Bit* (LSB), ya que un cambio en la señal de entrada de máximo un paso generará en la salida un cambio en el LSB del código binario (S P Lipshitz, 1992); la función de transferencia del cuantizador escalar se observa en la Figura 4-1.

Para desarrollar la descripción del cuantizador tal que permita la construcción de una

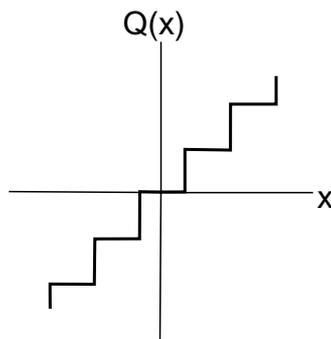


Figura 4-1: Función característica de transferencia de un cuantizador uniforme (S P Lipshitz, 1992).

medición QSP, primero es necesario definir los vectores $q_i, 1 \leq i \leq m$ que se encargan de realizar un mapeo $T_x : \mathbb{R} \rightarrow l_2$, tal que si la entrada x es igual a un nivel de cuantización, es decir $x = i\Delta$ el mapeo le asociará un vector $T_x(x) = q_i$, y si por el contrario la entrada tiene otro valor $x \neq k\Delta$, el mapeo le asociará un vector $T_x(x) = u$, donde $u = \sum_{k=1}^N q_k / (x - k\Delta)$. El mapeo se puede resumir como

$$T_x(x) = u = \begin{cases} q_i & \text{si } x = i\Delta \\ \sum_{k=1}^N q_k / (x - k\Delta) & \text{si } x \neq i\Delta. \end{cases} \quad (4-2)$$

La medición QSP será construida tomando los vectores de representación q_i como los vectores de medición, y la medición se desarrolla sobre u . Si u es igual a un vector de medición q_i , se tiene que $M(u) = u$, pero si no es igual entonces $M(u) = bq_i, b \in \mathbb{C}$, donde $i = f_M(\langle u, q_k \rangle)$. De forma general la medición M se puede expresar como

$$M(u) = \hat{y} = \begin{cases} z & \text{si } z = q_i \\ bq_i, \text{ con } i = f_M(\{\langle z, q_k \rangle\}) & \text{si } z \neq q_i. \end{cases} \quad (4-3)$$

y de la elección del mapeo probabilístico f_M dependerá el desempeño del cuantizador QSP. Por último, para la obtención del valor cuantizado se define un mapeo $T_y : l_2 \rightarrow \mathbb{R}$, tal que $T_y(\hat{y}) = i\Delta$, si $\hat{y} = bq_i$ (lo cual siempre se cumple por la definición de M), el mapeo le asociará un nivel $T_y(\hat{y}) = i\Delta$. La descripción del cuantizador utilizando el marco QSP se observa en la Figura 4-2.



Figura 4-2: Descripción del cuantizador mediante QSP (Eldar, 2001).

En el espacio vectorial \mathbb{R} el cuantizador probabilístico QSP presenta una gran ventaja ya que puede expresarse de forma abreviada como $Q_M(x) = i\Delta$, con $i = f(x)$ y donde f es el mapeo probabilístico, pero en este caso solo depende de los números $\{x - i\Delta\}$, aún más, el cuantizador en este caso solo depende de la entrada x . Este

cuantizador es llamado Cuantizador QSP con mapeo f (Eldar, 2001). Al representar el cuantizador como una medición QSP se obtiene una gran libertad en el diseño, ya que con la elección del mapeo f se da origen a nuevos tipos de cuantizadores, por ejemplo el cuantizador uniforme es obtenido si se escoge $f = \arg \min |x - i\Delta|$.

El cuantizador QSP surge al escoger un mapeo probabilístico, dando origen a cuantizadores probabilísticos (Eldar, 2001). De forma general se pueden clasificar los cuantizadores en dos tipos: Cuantizadores probabilísticos sin memoria que son aquellos cuantizadores que solo dependen de la entrada actual, y Cuantizador probabilísticos con memoria que depende de entradas anteriores de la señal (S P Lipshitz, 1992). En este trabajo se presenta únicamente el análisis correspondiente para el caso del cuantizador sin memoria.

4.1.1. Cuantizador Probabilístico sin Memoria

Para el desarrollo del cuantizador probabilístico se debe escoger un mapeo probabilístico $f : R \times V \rightarrow \mathbb{Z}$ donde $\{V = \mathbb{Z}\}$ representa el espacio muestral de una variable discreta auxiliar a . Así, la variable a solo puede tomar valores $a_i = i \in \mathbb{Z}$ con probabilidad $p_i = g(x - i\Delta)$, para alguna función $g(x)$. El mapeo se encargaría de escoger un índice i a cada entrada mediante $f(x, i) = i$, y la salida del cuantizador vendría siendo

$$y = i\Delta, \text{ con probabilidad igual a } p_i = g(x - i\Delta). \quad (4-4)$$

Este cuantizador se conoce como cuantizador probabilístico con mapeo $g(x)$. Como se está trabajando con variables aleatorias, y se desea que p_i represente probabilidades, $g(x)$ debe satisfacer:

$$g(x - i\Delta) \geq 0 \quad \text{para todo } i \in \mathbb{Z} \quad (4-5)$$

por lo que de forma general se tiene que $g(x) \geq 0$. También debe satisfacer la condición de normalización

$$\sum_{i=-\infty}^{+\infty} g(x - i\Delta) = 1. \quad (4-6)$$

Las anteriores son las condiciones que debe satisfacer el mapeo $g(x)$ en el dominio del tiempo, pero para caracterizarla completamente es necesario analizar lo que debe satisfacer en el dominio de la frecuencia, para eso se debe buscar la forma de re-expresar la anterior ecuación, una opción que se tiene al recordar las propiedades de la convolución de una función con el impulso es expresarla como una suma de funciones Diracs retrazadas:

$$g(x) \star \sum_{i=-\infty}^{+\infty} \delta(x - i\Delta) = 1 \quad (4-7)$$

y obtener su transformada de Fourier

$$\begin{aligned} G(w) \frac{2\pi}{\Delta} \sum_{i=-\infty}^{+\infty} \delta\left(w - \frac{2\pi}{\Delta}i\right) &= 2\pi\delta(w) \\ \Rightarrow \frac{2\pi}{\Delta} \sum_{i=-\infty}^{+\infty} G\left(\frac{2\pi}{\Delta}i\right) \delta\left(w - \frac{2\pi}{\Delta}i\right) &= 2\pi\delta(w) \end{aligned} \quad (4-8)$$

donde $G(w)$ es la transformada de Fourier de la función $g(x)$, la cual debe que satisfacer la siguiente condición en el dominio de la frecuencia (Eldar, 2001):

$$G\left(\frac{2\pi}{\Delta}i\right) = \Delta\delta[i], i \in \mathbb{Z}. \quad (4-9)$$

En conclusión, la función $g(x)$ debe cumplir dos condiciones: 1. ser una función no negativa, y 2. su transformada de Fourier debe satisfacer (4-9). Aún más, ya que $g(x)$ representa la función de densidad de probabilidad (fdp) de la variable auxiliar a , y se sabe que la función característica (fc) M viene dada por

$$M = \mathcal{F}\{g(x)\} \quad (4-10)$$

entonces $G(w)$ representa la función característica de la variable auxiliar a .

4.1.2. Implementación del Cuantizador Probabilístico

La implementación del cuantizador probabilístico es sencilla, primero se deben generar las probabilidades de acuerdo a (4-4), para luego generar la variable auxiliar a con alfabeto \mathbb{Z} , el cual se puede generar mediante un generador de números aleatorios, y a cada elemento $a = i$ se le asocia una probabilidad p_i , entonces la salida cuantizada será $y = i\Delta$. La dificultad se encuentra en determinar el valor i , así que una forma es generar la variable aleatoria uniformemente distribuida a sobre el intervalo $[0, 1]$, y tomar la probabilidades acumuladas $s_i = \sum_{k=1}^i p_k$, entonces la $p(s_i \leq a \leq s_{i+1}) = s_{i+1} - s_i = p_i$, así para obtener la salida basta con generar una realización de a , y entonces la salida cuantizada será $y = i\Delta$ donde el valor i es el que satisface $s_i \leq a \leq s_{i+1}$ (Eldar, 2001).

4.1.3. Relación con el Cuantizador *dithered*

En los sistemas de cuantización generalmente se asume que el error es una función determinística de la entrada, es decir el error se trata como un proceso aleatorio aditivo, independiente de la señal de entrada x y uniformemente distribuido (Schuchman, 1964; Gersho y Gray, 1991). Este modelo es muy útil para procesar señales complejas (casi-aleatorias) con amplitudes grandes respecto al paso definido Δ , pero si la entrada presenta valores comparable en magnitud al paso del cuantizador, la señal del error total ϵ es fuertemente dependiente de la entrada y se hace necesario la búsqueda de nuevos esquemas de cuantización que permitan reducir tal dependencia. Uno de ellos es el cuantizador *dithered*, cuyo fin es controlar las propiedades estadísticas

del error total y su relación a la señal de entrada del sistema (Wannamaker et al., 2000).

De forma general se tiene en los sistemas de cuantización determinísticos los sistemas *dithered* sin memoria o *Non-Subtractive Dither* (NSD) y los sistemas *dithered* con memoria o *Subtractive Dither* (SD), los cuales se pueden apreciar en la Figura 4-3.

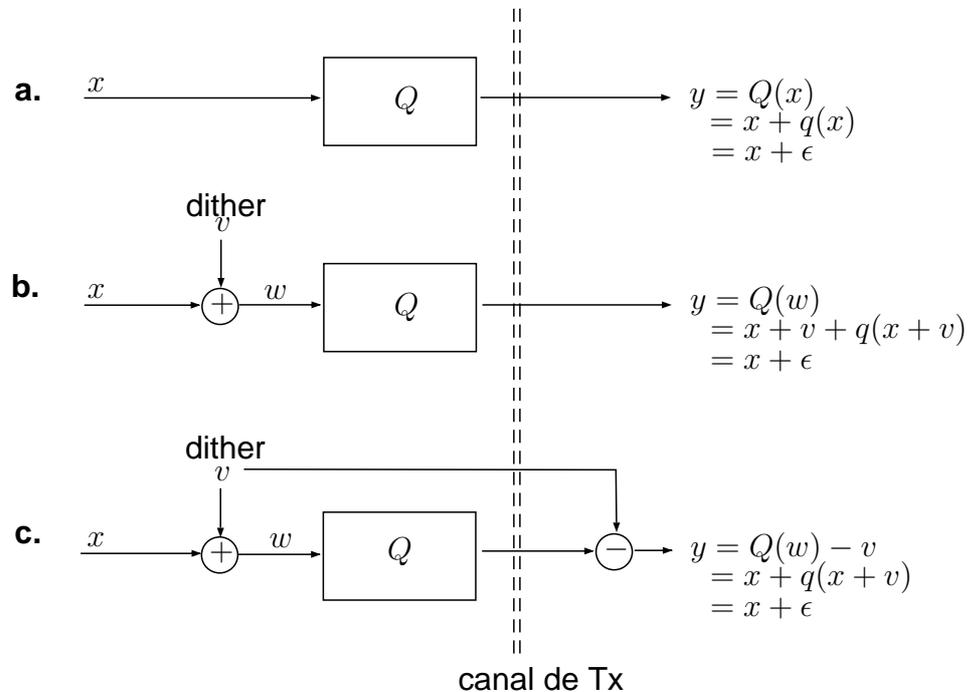


Figura 4-3: Tipos de sistemas de cuantización. a. Cuantización determinístico o no *dithered*. b. Cuantización *dithered* sin memoria NSD. c. Cuantización *dithered* con memoria SD. (S P Lipshitz, 1992).

Los cuantizadores *dither* reciben su nombre debido a la adición a la señal de entrada x una señal aleatoria v con fdp f_v llamada señal *Dither*, la cual se asume estacionaria ¹ y estadísticamente independiente de x . Esta señal origina que la entrada al cuantizador $w = x + v$ deje de ser una función determinística de x y a su vez también deje de serlo el error total ϵ . El SD representa el proceso ideal de cuantización, ya que permite que el error total sea un proceso ruidoso aditivo independiente de la entrada bajo ciertas condiciones en la función característica de la señal *dither* (Wannamaker et al., 2000), pero debido a que la señal *dither* debe ser sustraída en la salida del sistema, en sistemas de codificación (transmisión y/o almacenamiento) la señal *dither* debe viajar junto con la señal cuantizada y en recepción debe ser sincronizada, aumentando la complejidad al hacer necesario la incorporación de nuevos elementos, lo cual hace menos práctica su aplicación. Por otra parte, el NSD presenta una estructura mucho más simple con la limitante que solo ciertos momentos estadísticos del error total pueden ser independientes de la entrada, pero permite regular los momentos conjuntos

¹Un proceso aleatorio estacionario es aquel cuyas propiedades estadísticas son invariantes en el tiempo.

del error total, aunque no llega a lograr que el error y la entrada sean estadísticamente independientes.

La salida de un cuantizador NSD viene dada por:

$$y = Q(w) = Q(x + v) \quad (4-11)$$

con el error total igual a

$$\epsilon = y - x \quad (4-12)$$

y el error de cuantización $q(w)$ es función de la entrada al cuantizador w mediante la expresión

$$q(w) = Q(w) - w. \quad (4-13)$$

La diferencia entre el *error de cuantización* y el *error total* se hace necesaria plasmar, ya que en un sistema de cuantización sin señales dihter, la entrada al cuantizador w es la misma que al sistema x , y los errores son idénticos, pero en sistemas dihter las señales son distintas y para una descripción adecuada del sistema y del cuantizador se hace necesario la adopción de los dos tipos de error (S P Lipshitz, 1992).

Como la salida y del cuantizador solo puede tomar valores que son múltiplos enteros del paso de cuantización $y = k\Delta, k \in \mathbb{Z}$, la probabilidad será igual a si la entrada al cuantizador w cae entre $(2k - 1)\Delta/2$ y $(2k + 1)\Delta/2$, entonces:

$$\begin{aligned} p_k(y) &= P[(2k - 1)\Delta/2 \leq w \leq (2k + 1)\Delta/2] \\ p_k(y) &= P[k\Delta - \Delta/2 \leq x + v \leq k\Delta + \Delta/2] \\ p_k(y) &= P[k\Delta - \Delta/2 - x \leq v \leq k\Delta + \Delta/2 - x] \\ p_k(y) &= \int_{k\Delta - \Delta/2 - x}^{k\Delta + \Delta/2 - x} f_v(v) dv \end{aligned}$$

si se define la función rectangular

$$\mathbf{1}_\Delta = \begin{cases} 1 & -\Delta/2 \leq x \leq \Delta/2 \\ 0 & \text{para los demás valores} \end{cases} \quad (4-14)$$

se puede expresar la probabilidad de la salida del cuantizador como la convolución de f_v con la función rectangular $\mathbf{1}_\Delta$, reduciéndose la expresión a

$$p_k(y) = \mathbf{1}_\Delta(x - k\Delta) \star f_v(-x). \quad (4-15)$$

Es interesante observar que en la probabilidad de la salida del cuantizador no se describe específicamente la distribución de la variable de entrada x , facilitando su análisis.

Además, si se compara con la función definida para el cuantizador probabilístico, es decir

$$p_i = g(x - i\Delta)$$

e igualándolas $g(x - i\Delta) = \mathbf{1}_\Delta(x - k\Delta) \star f_v(-x)$, se observa que la forma del pulso $g(x)$ en el dominio del tiempo sería igual a

$$g(x) = \mathbf{1}_\Delta(x) \star f_v(-x), \quad (4-16)$$

indicando la relación existente entre f_v y $g(x)$, estableciendo que un cuantizador NSD con una señal dither con una fdp f_v , es equivalente a un cuantizador probabilístico QSP con mapeo $g(x)$. Además, si se obtiene su transformada de Fourier, se tiene que la función característica de la variable auxiliar a del cuantizador probabilístico sería:

$$G(w) = \Delta \text{sinc}\left(\frac{\Delta w}{2}\right) F_v(-w) \quad (4-17)$$

donde F_v representa la función característica de la variable aleatoria o señal dither v , que es igual a la transformada de Fourier de la función de densidad de probabilidad f_v de la señal dither (Eldar, 2001). La anterior ecuación muestra la condición en frecuencia que se debe satisfacer para que el cuantizador *dithered* describa a un cuantizador probabilístico con mapeo $g(x)$. También, se observa que satisface la ecuación (4-9), ya que si $w = \frac{2\pi}{\Delta}k$ se tendría $\text{sinc}(\pi k) = \delta[k]$ y como $F_v(0) = 1$, entonces sin importar la distribución de la variable aleatoria f_v , si el mapeo $G(w)$ es generado teniendo en cuenta la ecuación (4-17) se satisface la condición en frecuencia del cuantizador probabilístico (4-9). En el dominio del tiempo ya que tanto la fdp f_v y la función rectangular $\mathbf{1}_\Delta$ son no negativas, también lo será el mapeo $g(x)$. Entonces si $g(x)$ se diseña bajo la ecuación (4-16), se tendrá también que las condiciones (4-5) y (4-9) son satisfechas.

Por otra parte, partiendo del cuantizador probabilístico con mapeo $g(x)$ es posible el diseño de un cuantizador *dithered* siempre y cuando la función característica de la señal dither venga dada por

$$F_v(w) = \begin{cases} \frac{G(-w)}{\Delta \text{sinc}(\frac{\Delta w}{2})} & w \neq \frac{2\pi}{\Delta}k, k \in \mathbb{Z} \setminus 0 \\ M(w) & w = \frac{2\pi}{\Delta}k, k \in \mathbb{Z} \end{cases} \quad (4-18)$$

donde $M(w)$ representa los valores que la función F_v toma en los instantes $w = \frac{2\pi}{\Delta}k$ y deben ser tal que la función $f_v(x) \geq 0$ para que represente una fdp (Eldar, 2001).

Se ha observado que un cuantizador probabilístico bajo ciertas condiciones es equivalente a un cuantizador *dithered*, claro está desde un punto de vista de la salida de ambos cuantizadores. Pero la característica interesante o ventaja que presenta el cuantizador probabilístico es una implementación mucho más simple respecto al *dithered*, y se debe principalmente a que es necesario solo un generador de números uniformemente aleatorios. Esta separación lleva a generar cualquier tipo de distribución

deseada al escoger una adecuada función $g(x)$. En comparación, en típicas aplicaciones (Wannamaker et al., 2000; S P Lipshitz, 1992; Schuchman, 1964; Akyol y Rose, 2009) el cuantizador *dithered* es mucho más complejo, ya que este necesita generar una variable aleatoria con distribución f_v , la cual generalmente viene dada como una suma de variables aleatorias Independientes e Idénticamente Distribuidas (IID), distribuidas uniformemente sobre $[-\Delta/2, \Delta/2]$, y además puede llegar tener un grado de dificultad mucho mayor en el momento de la implementación.

Al igual que el cuantizador *dithered*, el cuantizador probabilístico permite la independencia de i momentos estadísticos del error total y la entrada del cuantizador $E(\epsilon^i|x)$, siempre y cuando el mapeo cumpla la siguiente condición:

$$G(w = \frac{2\pi}{\Delta}k)^{(i)} = 0, \forall k = \pm 1, \pm 2, \dots \quad (4-19)$$

donde el superíndice (i) denota la i -ésima derivada. Para $i = 1$ el error y la entrada son no correlacionadas. Si la condición (4-19) es satisfecha se tendría

$$E(\epsilon^i|x) = E(\epsilon^i) = \frac{(-j)^i}{\Delta} G(w = 0)^{(i)} = \frac{1}{\Delta} \int_x x^i g(x) dx \quad (4-20)$$

de acuerdo con la propiedad de la transformada de Fourier $\mathcal{F}[(-j)^k x^k g(x)] = G(w)^{(k)}$ (Eldar, 2001).

4.2. Diseño del mapeo probabilístico $g(x)$

Establecida la relación entre el mapeo $g(x)$ y la función característica de la señal dither, en este trabajo se exploraron diferentes posibilidades de funciones que satisfagan la condición de no negatividad y que su transformada de Fourier cumpla con la condición en frecuencia del cuantizador probabilístico (4-9), una de ellas es la función de autocorrelación de las funciones wavelet ortogonales, la cual es analizada a continuación.

4.2.1. Autocorrelación de funciones wavelet

Del análisis multiresolución se obtuvo las funciones wavelet ortogonales $\{\psi_j, j \in \mathbb{Z}\}$, que satisfacen

$$\langle \psi_i, \psi_k \rangle = \delta(i - k) \quad (4-21)$$

y para analizar en el dominio de la frecuencia, primero es posible expresar la anterior ecuación como

$$\int_{-\infty}^{+\infty} \psi(t - i) \psi^*(t - k) dt = \psi \star \psi^*(i - k) = \delta[i - k] \quad (4-22)$$

y como $\{\psi(i - k) = \psi(n), n \in \mathbb{Z}\}$ se tiene

$$\int_{-\infty}^{+\infty} \psi(t) \psi^*(t - n) dt = \delta_n \quad (4-23)$$

que es igual a la función de autocorrelación $\gamma(n)$, así se tiene

$$\gamma(n) = \int_{-\infty}^{+\infty} \psi(t)\psi^*(t-n)dt = \delta_n \begin{cases} 1 & n = 0 \\ 0 & n \neq 0. \end{cases} \quad (4-24)$$

que no es más que el criterio de Nyquist en el dominio del tiempo (Proakis, 2007). Para obtener la transformada de Fourier de la función de autocorrelación se puede ver que (4-23) representa un proceso de muestreo de la función $\psi(t)$, es decir, se puede re-escribir como

$$\psi \star \psi^* \sum_{n=-\infty}^{+\infty} \delta(t-n) = \delta[n] \quad (4-25)$$

ahora, aplicando la transformada de Fourier $\mathcal{F}\{\psi \star \psi^* \sum_{n=-\infty}^{+\infty} \delta(t-n)\}$ y recurriendo a la ortogonalidad de Poisson (2-8) se tiene la condición de ortogonalidad de la funciones wavelet en el dominio de la frecuencia

$$\sum_{k=-\infty}^{+\infty} |\Psi(w - 2\pi k)|^2 = 1 \quad (4-26)$$

donde $|\Psi(w)|^2 = \mathcal{F}\{\psi(t) \star \psi^*(-t)\}$ es una función no-negativa. Para una familia de wavelets ortonormales $\{\psi_j(t - nT); n \in \mathbb{Z}\}$, con un intervalo de símbolo $T = 2^j T_0$, y con una función de autocorrelación $\gamma_j(t) = \psi_j(t) \star \psi_j^*(-t)$, la ecuación (4-26) se puede generalizar como:

$$\sum_{k=-\infty}^{+\infty} \Gamma_j(w - \frac{2\pi k}{T}) = T \quad (4-27)$$

donde Γ_j es la transformada de Fourier de la función de autocorrelación γ_j , que satisface la condición en el ámbito de la comunicaciones sobre un canal de banda limitada, conocida como el criterio de Nyquist para la conformación del pulso, y que evita la presencia de interferencia intersímbolo o *Intersymbol Interference* (ISI) cuando $\psi_j(t)$ es usada como un filtro de conformación de pulso (*pulse-shaping filter*) y $\psi_j^*(-t)$ es usada como el filtro adaptado (*matched filter*). En conclusión, las funciones escala y wavelet pueden no ser pulsos Nyquist, pero si lo son sus funciones de autocorrelación.

De (4-27) se observa que la transformada de Fourier de la función de autocorrelación satisface las condiciones en tiempo de un mapeo probabilístico (4-5) ya que por definición es nonegativa, y sí se toma un paso $\Delta = 2\pi/T$, entonces

$$\sum_{i=-\infty}^{+\infty} \frac{1}{T} \Gamma(x - \Delta i) = 1 \quad (4-28)$$

satisface la condición de normalización (4-6). Queda por verificar la condición en frecuencia (4-9), pero si se toma $g(x) = \frac{1}{T} \Gamma(x)$, se tiene que

$$\sum_{i=-\infty}^{+\infty} g(x - \Delta i) = 1 \quad (4-29)$$

es igual a la misma ecuación de normalización (4-6), por tanto su transformada de Fourier satisface la condición en frecuencia (4-9). Lo anterior se puede resumir en el siguiente teorema:

Teorema 4.1. *(Autocorrelación wavelet como mapeo) Sea $\psi_j(t)$ una función wavelet ortonormal, y su función de autocorrelación $\gamma_j(t)$ definida como $\gamma_j(t) = \psi_j(t) \star \psi_j^*(-t)$. Se dice que $\mathcal{F}\{\gamma_j(t)\}$ satisface las condiciones de un mapeo probabilístico $g(x)$, cuando se toma un paso de cuantización igual a $\Delta = 2\pi/T$.*

Se ha demostrado que la función de autocorrelación de wavelets ortonormales cumple todas las condiciones para servir como un mapeo probabilístico, pero queda ver si ¿existen wavelets que lo sean? o ¿es posible su construcción?. Para ambas preguntas la respuesta es «afirmativa». En primer lugar, entre las wavelets que satisfacen el criterio de Nyquist se tiene la función Haar y las Oliveira definidas en (Oliveira et al., 2002) que dan origen a una nueva familia de funciones wavelet a partir del filtro de coseno alzado, permitiendo que estas nuevas wavelet satisfagan el criterio de Nyquist; y se pensaría que podrían ser utilizadas como un mapeo probabilístico $g(x)$, pero recordando que toda función wavelet debe mantener un nivel dc cero, y por tanto debe cambiar en signo, no podrían ser por definición una función no-negativa. Entonces, una wavelet vista en el dominio del tiempo solo puede satisfacer el criterio Nyquist, ahora queda ver su comportamiento en frecuencias; si su transformada es no-negativa como en el caso de las wavelet oliveira, entonces su transformada puede ser utilizada como mapeo probabilístico. En conclusión, se puede establecer que algunas funciones wavelet satisfacen las condiciones de un mapeo probabilístico.

Ya establecida que la función de autocorrelación de las funciones wavelet ortonormales sirven como mapeos para un cuantizador probabilístico, es de gran interés determinar una forma eficiente de implementarla, por tal motivo en el presente trabajo de grado se desarrolló una nueva perspectiva que permita llegar a tal fin, la cual es descrita a continuación.

4.3. Implementación de la función de autocorrelación

Las familias de funciones wavelet se pueden clasificar de manera general como (Mallat, 2008; Daubechies, 1992; Cohen y Kovačević, 1996):

1. Wavelets con soporte finito (tiempo limitado) y por tanto un ancho de banda ilimitado,
2. Wavelets con ancho de banda limitado y soporte infinito en tiempo.

que dan origen a dos definiciones de las funciones de autocorrelación: para el primer caso se asociaría a una definición basada en potencia y para el segundo caso a energía, y más aún, la transformada de Fourier de la función de autocorrelación es definida como

1. Densidad espectral de potencia para wavelet de soporte finito, y
2. Densidad espectral de energía para wavelet de soporte infinito.

y ya que la función de autocorrelación es real, ambas densidades no contienen información de fase. El objetivo es la búsqueda de una manera eficiente de obtener la función de autocorrelación. Ya que las funciones wavelet de soporte compacto pueden ser implementadas fácilmente por el algoritmo de mallat y a trozos (Mallat, 2008), y como son de tamaño finito se logran implementar con exactitud. Además, se demostró en Saito y Beylkin (1993) que la función de autocorrelación γ es la «función fundamental» en el esquema de interpolación iterativa de Lagrange, por tanto puede ser obtenida de manera recursiva, similar a la forma en que son obtenidas las funciones wavelet en la teoría del análisis multiresolución. Retomando la expresión de las funciones wavelet y *scaling*

$$\phi_{j,m}(t) = \sum_{n=-\infty}^{+\infty} h[n-2m]\phi_{j-1,n}(t) = \phi_{j-1,n} \star \bar{h}[2m] \quad (4-30)$$

$$\psi_{j,m}(t) = \sum_{n=-\infty}^{+\infty} g[n-2m]\phi_{j-1,n}(t) = \psi_{j-1,n} \star \bar{g}[2m] \quad (4-31)$$

donde $\bar{f} = f(-t)$. La función de autocorrelación escala viene dada por $\rho(n) = \phi_j(t) \star \phi_j^*(-t)$, y si se recurre a las anteriores ecuaciones, se puede expresar la autocorrelación escala de manera recursiva como

$$\rho_j(n) = \phi_{j,n} \star \bar{\phi}_{j,n}^* \quad (4-32)$$

$$\rho_j(n) = \phi_{j-1,n} \star \bar{h}[2m] \star (\bar{\phi}_{j-1,n} \star \bar{h}[2m])^* \quad (4-33)$$

re-ordenando se tiene

$$\rho_j(n) = h[2m] \star \bar{h}[2m] \star \phi_{j-1,n} \star \bar{\phi}_{j-1,n}^* \quad (4-34)$$

y como la convolución $\phi_{j-1,n} \star \bar{\phi}_{j-1,n}^*$ representa la autocorrelación a las escala $j-1$, es decir $\rho_{j-1}(n)$, entonces es posible expresar la autocorrelación escala de manera recursiva como

$$\rho_j(n) = h[2m] \star \bar{h}[2m] \star \rho_{j-1}(n). \quad (4-35)$$

De manera similar la autocorrelación wavelet vendría dada por

$$\gamma_j(n) = g[2m] \star \bar{g}[2m] \star \gamma_{j-1}(n). \quad (4-36)$$

Las anteriores expresiones se pueden plasmar en el siguiente teorema:

Teorema 4.2. (Autocorrelación wavelet recursiva) *La función de autocorrelación escala ρ y la autocorrelación wavelet γ vienen definidas por*

$$\rho_j(n) = H \star \rho_{j-1}(n). \quad \gamma_j(n) = G \star \gamma_{j-1}(n) \quad (4-37)$$

donde los filtros viene dados por $H = h[2m] \star \bar{h}[2m]$ y $G = g[2m] \star \bar{g}[2m]$.

En conclusión, de manera similar al caso del análisis multiresolución se tiene la ventaja de asociarles unos filtros a las funciones de autocorrelación, lo que facilita su implementación, aunque no sean de cuadratura. Como se sabe N representa el número de coeficientes del filtro h y $N_j = (2^j)(N - 1)$ define la longitud de la j -ésima wavelet Daubechies discreta (Daubechies, 1992). Las anteriores formas de obtener de forma recursiva la función de autocorrelación trae consigo grandes ventajas, por ejemplo se pasa de un número de operaciones de N^2 a ser igual a la longitud N .

Por último, en Saito y Beylkin (1993) se demostró que las funciones de autocorrelación son posibles dilatarlas y trasladarlas, de forma similar a las funciones wavelet. Lo cual permite realizar una descomposición de una función f , con gran impacto especialmente imágenes; en parte por la simetría que presenta la función de autocorrelación. Así, para la autocorrelación de la función wavelet se tiene

$$\gamma(n)_{j,k} = \frac{1}{\sqrt{2^j}} \gamma\left(\frac{n-k}{2^j}\right) \quad (4-38)$$

y de igual manera se puede definir la función de autocorrelación para las funciones escala, que vendría dada por

$$\rho(n)_{j,k} = \frac{1}{\sqrt{2^j}} \rho\left(\frac{n-k}{2^j}\right). \quad (4-39)$$

4.4. Observaciones finales

En este capítulo se ha desarrollado un nuevo mapeo probabilístico recurriendo a la función de autocorrelación de las funciones wavelet ortogonales, la cual se demostró cumple con todas las condiciones necesarias para un cuantizador NSD. Además, se logró obtener una expresión recursiva, que permite mostrar que la función de autocorrelación es similar a las funciones wavelet, es decir, pueden ser obtenida de forma recursiva, facilitando su implementación.

Capítulo 5

Aplicación de la Transformación QSP en la teoría Wavelet

Para una función f que ha sido contaminada por ruido blanco w , al ser representada en el dominio de la transformada wavelet, el ruido presente en los coeficientes continúa siendo blanco, y aunque la energía de la función se concentre en ciertos coeficientes, la energía del ruido no. De manera general siempre se ha supuesto que la transformada wavelet realiza un blanqueamiento, donde los coeficientes han sido decorrelacionados o marginalmente Gaussianos, es decir; que todos los coeficientes son independientes (Mallat, 2008). Pero en Crouse et al. (1998), se ha empezado a reconocer que los coeficientes wavelet realmente no son marginalmente Gaussianos, y no son independientes respecto a los coeficientes en la misma y diferentes escalas, por tanto se hace necesario determinar las relaciones estadísticas conjuntas. Por ejemplo, en Atto y Berthoumieu (2012) se ha visto que la transformada discreta wavelet aproximadamente decorrelaciona ciertos procesos estocásticos, pero no lo hace en la gran mayoría, por tanto se hace pertinente la búsqueda de un proceso que permita optimizar la decorrelación de las variables aleatorias.

Es de resaltar que se sigue utilizando el modelado Gaussiano debido en gran parte a la reducción en la complejidad del modelado al asumir esta condición, además permite que los resultados sigan siendo válidos y cercanos al ideal. Así, se han venido desarrollando nuevos modelos que permitan estudiar la estadística de la transformada wavelet. Aunque no es de nuestro interés en este trabajo la evaluación de estos modelos, es importante tenerlos presente para la búsqueda de una mejora en el desempeño.

En este capítulo se presenta como aporte del trabajo de investigación la caracterización de las condiciones bajo las cuales la matriz de la transformada wavelet discreta no tiene efecto de decorrelación sobre las variables aleatorias de un vector aleatorio. También se presenta la aplicación de una transformación QSP que permita mejorar la decorrelación de un vector aleatorio al expresarlo en las funciones base wavelet, y se demuestra que en el dominio wavelet la transformada QSP puede ser calculada y aplicada con una menor complejidad computacional.

5.1. Variables correlacionadas

Se desea explorar si al tener un vector aleatorio $x \in \mathbb{R}^n$, cuyas variables aleatorias están correlacionadas, la transformada wavelet tiene algún grado de decorrelación y en caso contrario buscar un método que permita realizarlo. Para la generación de un vector correlacionado existen muchos métodos, y en el presente trabajo se optó por la descomposición *Cholesky* (Golub y Van Loan, 1996). Se debe primero generar un vector aleatorio de variables independientes z con distribución Normal o Gaussiana $N(0, \sigma^2)$, luego se debe encontrar la matriz C que representa la matriz de covarianza deseada, la cual debe tener la característica de ser semi-definida positiva, y así poder hacer uso de la descomposición Cholesky para obtener la siguiente descomposición

$$\Sigma = LL^T \quad (5-1)$$

donde L es llamado el factor Cholesky, y es una matriz triangular superior con todas la entradas de la diagonal principal positivas. El vector con variables correlacionadas vendría dado por

$$x = Lz \quad (5-2)$$

y su matriz de covarianza sería igual a

$$\begin{aligned} C &= E\{Lz(Lz)^T\} \\ &= LE\{zz^T\}L^T \end{aligned} \quad (5-3)$$

como el vector z esta compuesto por variables independientes, entonces $E\{zz^T\} = I$. Reemplazando se tiene la igualdad deseada

$$C = LL^T = \Sigma. \quad (5-4)$$

En la Figura 5-1 se observa tanto la correlación de las variables del vector aleatorio, como su matriz de covarianza mediante un diagrama de dispersión, para el caso del vector no-correlacionado (z) y correlacionado (x), para una dimensión de $n = 2$.

Para poder analizar como actúa la transformación wavelet sobre este vector, es necesario primero expresar la transformación en una forma que permita una mejor manipulación matemática.

5.2. Modelo wavelet mediante operadores

Se tiene que el proceso tanto de síntesis como de análisis es posible expresarlo mediante operadores lineales, y para el caso del análisis multiresolución se sabe que las funciones $\{\psi_j\}_{j \in J}$ forman una base Riesz para un espacio de Hilbert H , entonces se puede definir el operador de síntesis asociado como $\Phi : \ell^2(\mathbb{Z}) \rightarrow H$, donde $\ell^2(\mathbb{Z})$ representa el conjunto de secuencias con norma finita. El operador es representado por una

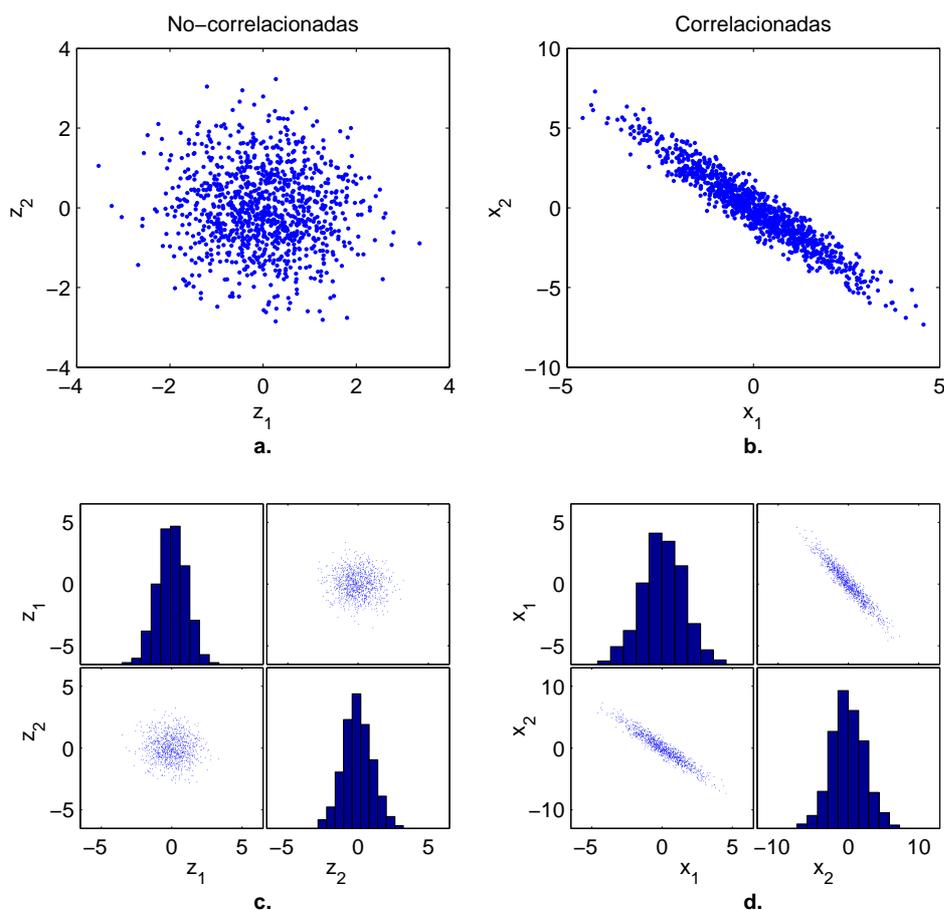


Figura 5-1: **a.** Correlación entre las variables z_1 y z_2 . **b.** Correlación entre las variables x_1 y x_2 . **c.** Diagrama de dispersión de las variables del vector z . **d.** Diagrama de dispersión de las variables del vector x .

matriz cuyas columnas son los vectores de la base wavelet. Si se recurre al operador Φ es posible recuperar la función original mediante

$$x = \Phi\alpha = \sum_{j \in J} \alpha_j \psi_j, \tag{5-5}$$

El adjunto del operador Φ es llamado el operador síntesis, denotado por $\Phi^* : H \rightarrow \ell^2(\mathbb{Z})$, que permite decomponer la función así

$$(\Phi^*x)_j = \langle x, \psi_j \rangle = \alpha_j, j \in J. \tag{5-6}$$

Dado que en el análisis multiresolución las funciones wavelet son ortonormales, los operadores por tanto son también ortogonales y unitarios $\|\Phi\| = 1$. Así, se tiene

$$\Phi\Phi^* = I. \tag{5-7}$$

La expansión de la función x al recurrir a los operadores lineales de análisis y síntesis

viene dada por

$$\mathbf{x} = \Phi \Phi^* \mathbf{x} \quad (5-8)$$

$$= \sum_{j \in J} \langle \mathbf{x}, \psi_j \rangle \psi_j. \quad (5-9)$$

5.3. Matriz de covarianza en el dominio Wavelet

Sea $\mathbf{x} \in \mathbb{R}^n$ un vector aleatorio con valor esperado cero, al expresarlo en el dominio tiempo-frecuencia, es decir, al aplicar el operador Φ^* o lo que es lo mismo, la transformada wavelet, los coeficientes vendrían dados por $\alpha = \Phi^* \mathbf{x}$, y más aún, cada coeficiente viene dado por

$$\alpha_j = \langle \mathbf{x}, \psi_j \rangle = \sum_{n=0}^{n=+\infty} x[n] \psi_j^*[n] \quad j \in J. \quad (5-10)$$

El vector conformado por todos los coeficientes wavelet $\alpha = [\alpha_1, \dots, \alpha_J]^T$ obtenidos del vector \mathbf{x} , será a su vez también un vector aleatorio, donde J representa la profundidad obtenida en la descomposición de la señal, y para el caso de la transformada wavelet discreta sería igual a $J = \log_2(n)$, obteniéndose una dimensión del vector α igual a la del vector \mathbf{x} .

Como el vector de coeficientes es aleatorio, el valor esperado vendría siendo

$$E\{\Phi \mathbf{x}\} = \Phi E\{\mathbf{x}\} = 0 \quad (5-11)$$

y también se podría pensar en obtener la matriz de covarianza

$$\begin{aligned} C_\alpha &= E\{\alpha \alpha^T\} \\ &= E\{\Phi^* \mathbf{x} (\Phi^* \mathbf{x})^T\} \\ &= E\{\Phi^* \mathbf{x} \mathbf{x}^T \Phi^{*T}\} \\ &= \Phi^* E\{\mathbf{x} \mathbf{x}^T\} \Phi^{*T} \end{aligned}$$

obteniéndose la relación

$$C_\alpha = \Phi^* C_x \Phi^{*T}, \quad (5-12)$$

y donde se aprecia que el vector de coeficientes será no correlacionado solo en el caso que la matriz de covarianza C_x sea diagonal, que se obtiene en casos como por ejemplo cuando las variables son IID.

Esta relación de similitud es interesante de estudiar porque muestra que es posible utilizar la matriz de covarianza en el dominio wavelet como referente para procesos que buscan extraer información, o decorrelacionar las variables del vector original. Con ventajas como, por ejemplo, que la matriz C_α tiene generalmente muchos elementos ilegibles que se pueden asumir cero dando origen a matrices escasas o *sparse*, sin

afectar notablemente en la aproximación a la matriz original, y con grandes ventajas en su computación y almacenamiento (Mallat, 2008; Dasarathy et al., 2015).

Tomando la norma de *Frobenius* (Meyer, 2000), que viene definida por

$$\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} = (\text{traza}(A^T A))^{1/2} \quad (5-13)$$

reexpresando la ecuación (5-12) se tiene

$$\Phi^{*-1} C_\alpha = C_x \Phi^{*T} \quad (5-14)$$

para el caso de una base wavelet ortonormal $\Phi^{*-1} = \Phi^{*T}$,

$$\Phi^{*T} C_\alpha = C_x \Phi^{*T} \quad (5-15)$$

tomando la norma de Frobenius a ambos lados

$$\begin{aligned} \|\Phi^{*T} C_\alpha\|_F^2 &= \|C_x \Phi^{*T}\|_F^2 \\ \|\Phi^{*T} C_\alpha\|_F^2 &= \|\Phi^{*T} C_x\|_F^2 \\ \text{traza}((\Phi^{*T} C_\alpha)^T \Phi^{*T} C_\alpha) &= \text{traza}((\Phi^{*T} C_x)^T \Phi^{*T} C_x) \\ \text{traza}(C_\alpha^T \Phi^* \Phi^{*T} C_\alpha) &= \text{traza}(C_x^T \Phi^* \Phi^{*T} C_x) \end{aligned}$$

y como se ha supuesto una base ortogonal, entonces $\Phi^* \Phi^{*T} = I$. Reemplazando se tiene

$$\text{traza}(C_\alpha^T C_\alpha) = \text{traza}(C_x^T C_x)$$

obteniéndose la importante relación

$$\|C_\alpha\|_F = \|C_x\|_F \quad (5-16)$$

indicando que las matrices de covarianza son equivalentes desde un punto de vista de sus autovalores y de su polinomio característico, más no una igualdad. Así, es natural pensar bajo que condiciones es posible que se de una igualdad. A continuación se aborda esta condición, para lo cual primero se va a explorar la estructura de la matriz wavelet (Strang, 1989; Strang y Strela, 1995).

5.3.1. Matriz wavelet

Se sabe que las bases escala y wavelet son la respuesta al impulso de los filtros h y g , por tanto es necesario analizar la estructura dada por el banco de filtros del árbol wavelet desde una perspectiva discreta, y una forma es abstraer la información mediante el álgebra lineal. Si a un vector x de tamaño N , le es aplicado un filtro h con coeficientes $(h_0, h_1, \dots, h_{N-1})$, se sabe que la salida viene dada por la convolución $h \star x$, que es igual a

$$y[n] = \sum_{i=1}^N x[i] h[n-i]. \quad (5-17)$$

El filtro h no es más que una combinación de desplazamientos (adelantos y retrasos), donde un retraso viene dado por

$$\mathbf{d}_k = \delta[n - k] \quad (5-18)$$

y un adelanto por

$$\mathbf{d}_{-k} = \delta[n + k] \quad (5-19)$$

entonces el filtro puede ser expresado como

$$\begin{aligned} h[n] &= \dots + h[-1]\mathbf{d}_{-1} + h[0]\mathbf{I} + h[1]\mathbf{d} \\ h[n] &= \sum h[i]\mathbf{d}_i \end{aligned} \quad (5-20)$$

El operador \mathbf{d} es representado por una matriz con una diagonal de unos, donde el desplazamiento k define la posición de la diagonal; es decir si $k = 0$ se tendría la matriz identidad \mathbf{I} , si $k = 1$ la diagonal de unos se desplaza hacia abajo dando origen a una matriz triangular inferior

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & 0 & 0 & 0 & \cdot \\ \cdot & 1 & 0 & 0 & 0 & \cdot \\ \cdot & 0 & 1 & 0 & 0 & \cdot \\ \cdot & 0 & 0 & 1 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot \\ x[0] \\ x[1] \\ x[2] \\ x[3] \\ \cdot \end{bmatrix} = \begin{bmatrix} \cdot \\ x[-1] \\ x[0] \\ x[1] \\ x[2] \\ \cdot \end{bmatrix} \quad (5-21)$$

si $k = -1$ daría origen a una matriz triangular superior. Como los filtro wavelet son lineales, entonces pueden ser expresado en forma *matricial* (Strang y Nguyen, 1997). Para el caso de los filtros wavelet se tiene que son invariantes en el tiempo, dando la propiedad de diagonales constantes en la matriz, además cada columna del filtro es un retraso de la previa columna y cada i -ésima diagonal del filtro es constante con un valor igual al i -ésimo coeficiente. Para un filtro con 4 coeficientes, la matriz sería algo como

$$\mathbf{H} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & h[0] & 0 & 0 & 0 & \cdot \\ \cdot & h[1] & h[0] & 0 & 0 & \cdot \\ \cdot & h[2] & h[1] & h[0] & 0 & \cdot \\ \cdot & h[3] & h[2] & h[1] & h[0] & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad \begin{array}{l} \vdots \\ \text{Diag. -1} \\ \text{Diag. 0} \\ \dots \text{Diag. 1} \end{array} \quad (5-22)$$

y la forma de la matriz es conocida como *Toeplitz* (Gray, 2005). Así, la convolución es representada por la matriz

$$h \star x = \mathbf{H}x. \quad (5-23)$$

Como en el proceso de la transformación wavelet discreta se hace un proceso de submuestro o downsampling, el cual elimina la mitad de información al reducir la entradas impares y mantener solo las entradas pares, que en forma matricial sería equivalente a

$$(\downarrow 2) = \begin{bmatrix} \cdot & 1 & & & & & & \\ & 0 & 0 & 1 & & & & \\ & & & & 0 & 0 & 1 & \\ & & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & \cdot \\ & & & & & & & \cdot \end{bmatrix} \quad (5-24)$$

es de resaltar que la transpuesta de la matriz de downsampling es upsampling $(\downarrow 2)^T = (\uparrow 2)$. Multiplicando la matriz de downsampling con la matriz del filtro, que equivale a eliminar las filas impares de la matriz del filtro, se obtiene

$$\begin{bmatrix} \cdot & \cdot \\ \cdot & h[1] & h[0] & 0 & 0 & 0 & 0 & \cdot \\ \cdot & h[3] & h[2] & h[1] & h[0] & 0 & 0 & \cdot \\ \cdot & 0 & 0 & h[3] & h[2] & h[1] & h[0] & \cdot \\ \cdot & \cdot \end{bmatrix} \quad (5-25)$$

la cual es una matriz *block-toeplitz* (Strang, 1989), con bloques de tamaño 1×2 . Como en la transformada wavelet discreta actúan los filtros h y g que forman un banco de filtros y sirven para obtener los coeficientes de detalles y de aproximación, al combinar las matrices se tendría la **matriz wavelet** (Strang, 1989; Strang y Nguyen, 1997; Strang, 2012), que viene dada por

$$\Phi^* = \begin{bmatrix} (\downarrow 2)\mathbf{H} \\ (\downarrow 2)\mathbf{G} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot \\ \cdot & h[1] & h[0] & 0 & 0 & 0 & 0 & \cdot \\ \cdot & g[1] & g[0] & 0 & 0 & 0 & 0 & \cdot \\ \cdot & h[3] & h[2] & h[1] & h[0] & 0 & 0 & \cdot \\ \cdot & g[3] & g[2] & g[1] & g[0] & 0 & 0 & \cdot \\ \cdot & 0 & 0 & h[3] & h[2] & h[1] & h[0] & \cdot \\ \cdot & 0 & 0 & g[3] & g[2] & g[1] & g[0] & \cdot \\ \cdot & \cdot \end{bmatrix} \quad (5-26)$$

y esta estructura es conocida como *block-Toeplitz*, con bloques 2×2 . Cuando se manejan datos de dimensión finita, que es el caso práctico, generalmente se usan bases wavelet periódicas como en el caso de Fourier, las cuales dan origen a una matriz en bloques circular o block-circulant ¹. Primero reexpresemos la matriz wavelet como

$$\begin{bmatrix} \cdot & A_0 & & & \\ \cdot & A_1 & A_0 & & \\ \cdot & & A_1 & A_0 & \\ \cdot & & & A_1 & \cdot \end{bmatrix} \quad (5-27)$$

¹Una matriz circulante también es una matriz toeplitz (el recíproco no es cierto), generalmente se omite la inclusión del termino toeplitz. Para el propósito de este trabajo se tomara indistintamente el término matriz circular o matriz circulante

donde A_j representa un bloque 2×2 . Entonces la matriz circular para este caso vendría dada por

$$\begin{bmatrix} A_0 & A_1 \\ A_1 & A_0 \end{bmatrix} \quad (5-28)$$

y de forma general se tendría

$$\Phi_c^* = \begin{bmatrix} A_0 & A_1 & A_2 & \cdot & \cdot & A_{k-1} \\ A_{k-1} & A_0 & A_1 & \cdot & \cdot & A_{k-2} \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ A_1 & A_2 & A_3 & \cdot & \cdot & A_0 \end{bmatrix} \quad (5-29)$$

también en (Kautsky y Turcajová, 1995) se muestran que las bases biortogonales también dan origen a matrices en bloques circulares.

Se ha determinado la estructura de la matriz wavelet, queda por determinar que propiedades adicionales se deben tener para obtener la igualdad de la matriz de covarianza en tiempo y en el dominio wavelet. La anterior interrogante nos motivo para presentar en la siguiente sección como aporte del trabajo de grado, un análisis de las características a cumplir por la matriz wavelet para no decorrelacionar a un vector aleatorio.

5.3.2. Propiedades de la matriz wavelet para una no-decorrelación

El gran auge que ha tenido la teoría wavelet se debe en gran parte a su fácil implementación, ya que cada función wavelet o escala es implementada sobre un filtro digital, así las propiedades deseadas en la función son transmitidas en el diseño del filtro. Así, se han creado multitud de filtros que dan origen a bases con ciertas propiedades, cada una asociada a ciertas aplicaciones (Vaidyanathan, 1993; Strang y Nguyen, 1997). Una de ellas es la propiedad de *simetría*, muy adecuada para los problemas de borde de señales de dimensión finita. Otra es la condición de fase lineal deseada porque se opone a la distorsión de fase no lineal y mantiene la forma de la señal, de gran importancia en el procesamiento de señales e imágenes.

La propiedad de ortogonalidad ha sido ampliamente discutida, pero cabe recordar su importancia ya que permite una descomposición única sin redundancia. Las condiciones de ortogonalidad y simetría no es posible obtenerlas en wavelet reales, pero si en wavelet complejas con una fase aproximadamente lineal (Gao et al., 2002).

Resumiendo las propiedades, se tendría

$$\begin{aligned} \Phi^* &= (\Phi^*)^T \text{ Simetría,} \\ \Phi^T &= (\Phi^*)^{-1} \text{ Ortogonalidad.} \end{aligned}$$

Recordando la expresión de la matriz de covarianza del vector de coeficientes (5-12) y como se sabe que la matriz de covarianza debe ser semi-definida positiva y simétrica, se establece

$$\begin{aligned} C_\alpha &= C_\alpha^T \\ C_\alpha &= \Phi^* C_x \Phi^{*T} = AB = C_\alpha^T \end{aligned} \quad (5-30)$$

donde $A = \Phi^* C_x$ y $B = \Phi^{*T}$. Como $C_\alpha = AB = C_\alpha^T$, se tiene que

$$AB = B^T A^T = BA^T \quad (5-31)$$

resta determinar si A es simétrica, ya B lo es. La matriz A está compuesta por la multiplicación de la matriz wavelet y la matriz de covarianza $\Phi^* C_x$, en la sección anterior se observó que la matriz wavelet es block-toeplitz, y debido a las múltiples características de las wavelets, la matriz puede tomar otras estructuras, como por ejemplo centrosimétricas, block-toeplitz con bloques circulantes, etc (Kautsky, 1995; Strang, 2012; Kautsky y Turcajová, 1995; Deng-feng et al., 2003; Kautsky y Turcajova, 1994). Además, la matriz de covarianza para algunos procesos, por ejemplo, en (Hartley y Naik, 2001) donde se estudia la tasa de decesos de sectores del centro de una ciudad podrían ser circularmente correlacionados, o en (Khattree y Naik, 1994) donde se observó que los mensajes de receptores satelitales ubicados en los n lados de un polígono, presentan una covarianza con estructura simétrica y circular, con bloques circulares. Para el caso de las matrices por bloques que conmutan, en Lara (2001) se define el siguiente teorema

Teorema 5.1. *(Matrices por bloques conmutativas) Sean $A = \text{circ}(A_1, A_2, \dots, A_m)$ y $B = \text{circ}(B_1, B_2, \dots, B_m)$ matrices por bloques circulantes, si $A_j B_k = B_k A_j$ para $k, j = 1, \dots, m$ entonces $AB = BA$.*

El cual plantea que para llegar a la conmutación se debe analizar los bloques de las matrices, y como en Gray (2005) se demuestra que las matrices circulantes conmutan, entonces se llega al siguiente teorema

Teorema 5.2. *(Matrices BCBC) Todas las matrices por bloques circulantes con bloques circulantes (BCBC) del mismo tamaño conmutan. (Combescure, 2009)*

pero además, se sabe que las matrices simétricas también conmutan, es decir

Teorema 5.3. *(Matrices SC-toeplitz) Cualquier par de matrices simétricas y circulares (SC) toeplitz del mismo tamaño conmutan. (Liang et al., 2011)*

por tanto, una directa aplicación del teorema da origen al siguiente teorema

Teorema 5.4. *(Matrices BCSC) Todas las matrices por bloques circulantes con bloques simétricos y circulantes (BCSC) del mismo tamaño conmutan.*

dadas las flexibilidad en la construcción de las wavelet y por tanto en la matriz wavelet, es posible construir una matriz wavelet con tales características. Si el proceso presenta una matriz BCBC o BCSC, se tiene

$$A = \Phi^* C_x = C_x \Phi^* = A^T \quad (5-32)$$

demostrando que $A^T = A$, entonces A y B son conmutativas. En resumen se tiene que

$$\begin{aligned} A = A^T &\rightarrow \Phi^* C_x = C_x^T \Phi^{*T} = C_x \Phi^* \\ B = B^T &\rightarrow \Phi^* = \Phi^{*T} = \Phi^* \end{aligned} \quad (5-33)$$

reemplazando en (5-12) se tiene

$$C_\alpha = C_x \Phi^* \Phi^{*T} \quad (5-34)$$

y como las wavelet además poseen la propiedad de ortogonalidad, $\Phi^* \Phi^{*T} = I$, la matriz de covarianza del vector de coeficientes wavelet será igual a

$$C_\alpha = C_x. \quad (5-35)$$

Se observa que para este tipo de wavelet la matriz de covarianza de los coeficientes wavelet es igual a la matriz de covarianza del vector, demostrando que no se ha realizado ningún proceso de decorrelación de las variables aleatorias del vector x . Lo anterior se puede resumir en el siguiente teorema:

Teorema 5.5. *(No-decorrelacion de la transformada wavelet discreta) Si la matriz wavelet Φ y la covarianza C_x del vector aleatorio x son matrices BCBC o BSBC, la transformada wavelet discreta no decorrelaciona las variables aleatorias, y la matriz de covarianza de los coeficientes wavelet C_α y C_x son iguales.*

Para este tipo de casos en donde la transformada wavelet no tiene efecto en la estructura de la matriz de covarianza del proceso, y en los casos en los cuales la transformada wavelet no alcanza a decorrelacionar completamente las variables aleatorias, se hace necesario buscar una transformación que decorrelacione las variables, una de ellas es la diseñada en el QSP que se explica a continuación.

5.4. Transformación QSP

En el análisis estadístico y en procesos de regresión se exploran diferentes formas de obtener la información que queda oculta debido a diferentes fenómenos. En el procesamiento cuántico de señales partiendo de ideas de la detección cuántica se desarrolló una transformación lineal de la forma de covarianza que minimizan el MSE entre el dato transformado y el original (Eldar, 2001).

Supongase que se tiene un vector aleatorio $x \in \mathbb{R}^n$ con matriz de covarianza C_x y media cero. Se desea mediante una transformación lineal T obtener un nuevo vector

$\mathbf{y} = \mathbf{T}\mathbf{x}$ con cierta matriz de covarianza C_y , con la restricción que el vector \mathbf{y} debe estar en un sentido de MSE lo más cercano al vector \mathbf{x} , es decir

$$\epsilon_{mse} = \sum_{i=1}^n E\{(\mathbf{y}_i - \mathbf{x}_i)^2\} \rightarrow 0. \quad (5-36)$$

Partiendo que la matriz de covarianza viene dada por

$$C_y = E\{(\mathbf{y} - E\{\mathbf{y}\})(\mathbf{y} - E\{\mathbf{y}\})^*\} \quad (5-37)$$

y como $\mathbf{y} = \mathbf{T}\mathbf{x}$ y $E\{\mathbf{y}\} = 0$, reemplazando en (5-37) se tiene

$$C_y = \mathbf{T}C_x\mathbf{T}^* \quad (5-38)$$

así, la matriz de covarianza del vector resultante puede ser controlada mediante la transformación \mathbf{T} , es decir, siendo \mathbf{Z} la forma de la matriz deseada, se tiene

$$C_y = \mathbf{T}C_x\mathbf{T}^* = \alpha\mathbf{Z}. \quad (5-39)$$

Este tipo de transformaciones es muy útil, por ejemplo cuando se envía una señal y por el efecto de múltiples caminos en el receptor se obtienen múltiples señales, las cuales estarían correlacionadas y se desea buscar un proceso que las descorrelacione, es decir $\mathbf{Z} = \mathbf{I}$, para luego ser utilizadas en la optimización del receptor, o cuando se tiene un banco de correladores y se desea optimizar la identificación de las diferentes señales de los usuarios. Teniendo en cuenta que en ambos casos el objetivo es lograr un proceso de decorrelación que no introduzca distorsión, en el QSP se obtuvo la transformación óptima desde un punto de vista del MSE, y es resumida por el siguiente teorema (Eldar, 2001, 2003):

Teorema 5.6. (Transformación QSP) Sea $\mathbf{x} \in \mathbb{C}^n$ un vector aleatorio con matriz de covarianza $C_x = \mathbf{V}\mathbf{D}\mathbf{V}^*$ de rango- m . Sea $C_y = \alpha^2\mathbf{Z}$ la matriz de covarianza de rango- l del vector $\mathbf{y} = \mathbf{T}\mathbf{x}$, con $\mathbf{Z} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*$, y $\mathbf{D}^{1/2}\mathbf{V}^*\mathbf{Q}\mathbf{\Lambda}^{1/2} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^*$. La transformación de covarianza óptima que minimiza el MSE entre el vector \mathbf{x} y \mathbf{y} viene dada por

$$\mathbf{T} = \alpha\mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{W}\mathbf{U}^*(\mathbf{D}^{1/2})^\dagger\mathbf{V}^* \quad (5-40)$$

donde α es escogido para minimizar el MSE y viene dado por $\alpha = \text{Traza}((C_x\mathbf{Z})^{1/2}) / \text{Traza}(\mathbf{Z})$. Además, si las matrices son de rango completo, $l = n = m$ se tiene $\mathbf{T} = \alpha(\mathbf{Z}C_x)^{(-1/2)}\mathbf{Z}$.

Donde las matrices C_x y \mathbf{Z} tienen la eige-descomposición con \mathbf{V} y \mathbf{Q} matrices unitarias, y \mathbf{D} y $\mathbf{\Lambda}$ siendo las matrices diagonales de la respectiva descomposición. Además $\mathbf{D}^{1/2}\mathbf{V}^*\mathbf{Q}\mathbf{\Lambda}^{1/2} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^*$ es la Singular Value Decomposition (SVD) de $\mathbf{D}^{1/2}\mathbf{V}^*\mathbf{Q}\mathbf{\Lambda}^{1/2}$. Cuando la matriz $\mathbf{Z} = \mathbf{I}$, la transformación QSP recibe el nombre de transformación blanca. A pesar que la transformación es óptima desde una mirada del MSE, surgen interrogantes asociados a la complejidad y eficiencia de la transformación, debido principalmente en la necesidad de conocer la matriz de covarianza, ya que por ejemplo

en procesos de grandes dimensiones, la determinación de la covarianza podría incrementar aún más la complejidad, pero en Dasarathy et al. (2015); Chen et al. (2015) se presentan novedosos métodos que permiten obtener una aproximación a la matriz de covarianza sin la necesidad de almacenar todo el proceso aleatorio, dando una disminución en la complejidad del proceso de transformación.

5.5. Transformación blanca en el dominio wavelet

La transformación blanca desarrollada en (5-40) es completamente aplicable tanto en el dominio del tiempo como en el dominio wavelet, pero con la ventaja que las bases wavelet pueden ser seleccionadas de acuerdo a cierto conocimiento previo de la señal, tal que pueden aproximadamente representar la señal con pocos coeficientes (Mallat, 2008), dando origen a una matriz de covarianza *sparse*, lo cual además de mejorar su computación es muy adecuada para su almacenamiento o para su recuperación mediante algoritmos de aproximación (Dasarathy et al., 2015; Chen et al., 2015). Por ejemplo, cuando la matriz C_x es una matriz Calderon-Zygmund, la transformación $\Phi^* C_x \Phi^{*T}$ puede dar origen a una matriz muy *sparse*, en el sentido que muchas entradas de la matriz serán muy pequeñas y pueden ser descartadas (Beylkin et al., 1991).

Teniendo en cuenta lo anterior, a continuación como aporte del trabajo de grado se demuestra que la transformación QSP en el dominio wavelet puede ser implementada con una menor complejidad computacional, cuando se recurre a funciones umbral que den origen a matrices *sparse*.

Otra ventaja viene de una de las principales aplicaciones de la transformación wavelet, la eliminación de ruido, realizada mediante el uso de una función umbral o *thresholding* en los coeficientes wavelet. La idea es tomar los coeficientes obtenidos luego de aplicar la transformación wavelet y dejar solo aquellos que son mayores que cierto valor ya que se asume que estos representan la señal, y a los coeficientes de pequeño valor que representa el ruido se le asigna el valor de cero. Supongase que la señal ruidosa viene expresada como

$$\mathbf{y} = \mathbf{x} + \mathbf{e}, \mathbf{e} \sim N(0, \sigma^2) \quad (5-41)$$

donde \mathbf{x} es la señal de interés y \mathbf{e} representa el ruido. Obteniendo los coeficientes de la representación de la señal (5-6), se tiene:

$$\boldsymbol{\alpha} = \Phi^* \mathbf{y} = \Phi^* \mathbf{x} + \Phi^* \mathbf{e} \quad (5-42)$$

aplicando una función umbral dura o *hard thresholding* v_T^h (Mallat, 2008)

$$v_T^h(x) = x \mathbf{1}(|x| > T) \begin{cases} x & \text{si } |x| > T \\ 0 & \text{si } |x| \leq T \end{cases} \quad (5-43)$$

donde T representa el valor umbral y $\mathbf{1}$ es la función indicador. Otra opción muy conocida es la función umbral suave o *soft thresholding* v_T^s (Mallat, 2008) dada por

$$v_T^s(x) = \text{sign}(x)(|x| - T)_+ = \begin{cases} x + T & \text{si } x \leq -T \\ 0 & \text{si } |x| < T \\ x - T & \text{si } x \geq T \end{cases} \quad (5-44)$$

donde

$$(x)_+ = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \geq 0 \end{cases} \quad (5-45)$$

Aplicando el umbral duro a los coeficientes wavelet se tendría

$$\tilde{\alpha} = v_T^h(\alpha) \quad (5-46)$$

y de forma interesante además de la eliminación del ruido, al observar la matriz de covarianza se tiene

$$\begin{aligned} C &= E\{v_T^h(\Phi^* \mathbf{x})(v_T^h(\Phi^* \mathbf{x}))^T\} \\ &= \Phi^* \mathbf{x} \mathbf{1}(|\Phi^* \mathbf{x}| > T)_+ (\Phi^* \mathbf{x} \mathbf{1}(|\Phi^* \mathbf{x}| > T)_+)^T \end{aligned} \quad (5-47)$$

expresando el producto como producto matricial, y recordando que $\tilde{\alpha} \in \mathbb{C}^n$, se tiene

$$C_{ij} = E\left\{\sum_{k=1}^n (\Phi^* \mathbf{x} \mathbf{1}(|\Phi^* \mathbf{x}| > T)_+)_i (\Phi^* \mathbf{x} \mathbf{1}(|\Phi^* \mathbf{x}| > T)_+)_j^T\right\} \quad (5-48)$$

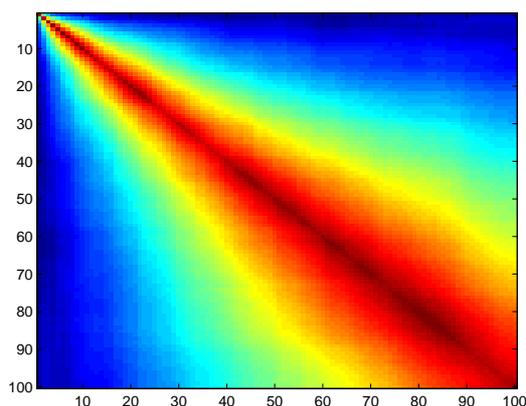
analizando su estructura de forma general

$$C = \begin{cases} (\Phi^* x)_i^2 \mathbf{1}(|(\Phi^* x)_i| > T)_+ & \text{si } i = j \\ ((\Phi^* x)_i (\Phi^* x)_j \mathbf{1}(|(\Phi^* x)_i| > T)_+ (\mathbf{1}(|(\Phi^* x)_j| > T)_+)^T & \text{si } i \neq j \end{cases} \quad (5-49)$$

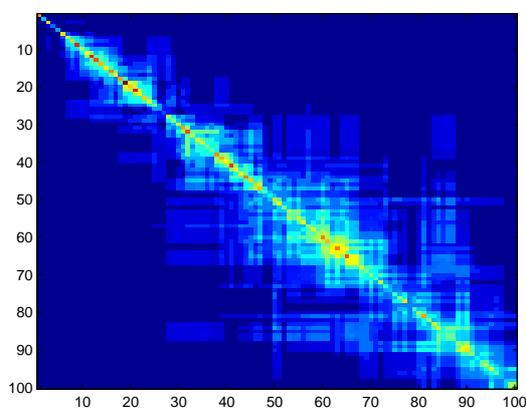
y para el umbral suave se tiene algo similar

$$C = \begin{cases} (|(\Phi^* x)_i| - T)_+^2 & \text{si } i = j \\ \text{sign}((\Phi^* x)_i) \text{sign}((\Phi^* x)_j) (|(\Phi^* x)_i| - T)_+ (|(\Phi^* x)_j| - T)_+ & \text{si } i \neq j \end{cases} \quad (5-50)$$

en ambos casos la diagonal de la matriz es siempre positiva, y en las demás posiciones de la matriz se tiene una doble condición de umbral, lo cual en algunos casos puede llevar a tener una matriz con pocos valores diferentes de cero o matriz *sparse*. También se observa que el umbral da origen a un estimador que garantiza las condiciones de matriz semi-definida positiva y simétrica, características esenciales de cualquier matriz de covarianza. En la Figura 5-2(a) se observa la matriz de covarianza de los coeficientes wavelets, y como cambia su estructura a una matriz *sparse* (Figura 5-2(b)) cuando los coeficientes wavelet son sometidos a un umbral duro.



(a) Covarianza coeficientes wavelet.



(b) Covarianza coeficientes wavelet + umbral.

Figura 5-2: Efecto de la función umbral sobre una matriz de covarianza.

En conclusión se observa que el uso de la transformación de blanqueamiento en el dominio wavelet puede traer beneficios adicionales respecto a los brindados en el dominio del tiempo, ya que muchos vectores tienen representaciones escasas en el dominio wavelet, lo cual lleva a su vez que su matriz de covarianza sea a escasa. Y como en Chen et al. (2015); Dasarathy et al. (2015) se demostró las reducciones en cuanto a computación cuando se hace uso de matrices *sparse*, y recordando que la transformación de blanqueamiento requiere para su cálculo de la estimación de la matriz de covarianza del proceso, entonces se puede concluir que la transformada QSP mejoraría su computación y su estimación en el dominio wavelet.

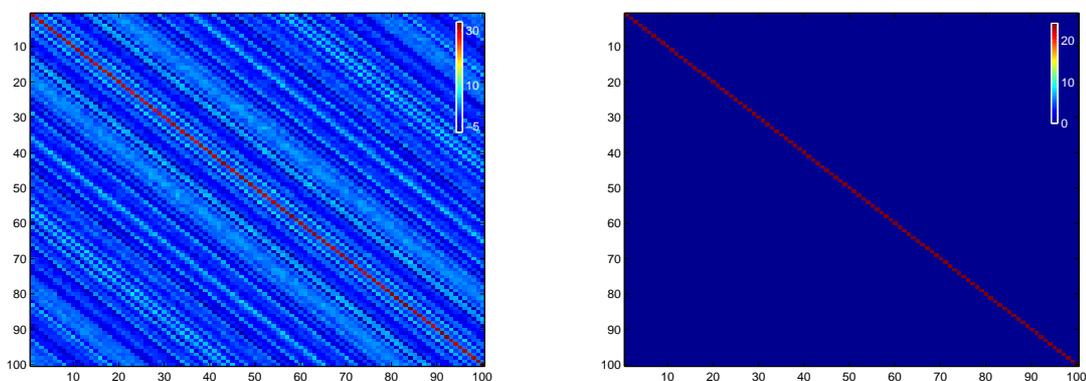
5.5.1. Validación de la aplicación mediante simulación

En la Sección 5.3.2 se determinó bajo que circunstancias la transformada wavelet no aporta en la decorrelación de las variables aleatorias, por tanto, como se había mencionado anteriormente se hace necesario recurrir a la transformación QSP para deco-

relacionar el vector aleatorio. Así, en esta sección se proveen algunos resultados de simulación que sugieren el comportamiento y desempeño de la transformación QSP en la transformada wavelet discreta.

Con el fin de mostrar que la transformada QSP es óptima desde un punto de vista del MSE, y que al ser aplicada en el dominio wavelet trae ventajas respecto a su computación, se diseñó una simulación que parte de un proceso gaussiano correlacionado x_C , obtenido mediante la descomposición Cholesky, con una dimensión $n = 100$. Se generaron 2^{11} realizaciones con el fin de que los resultados fueran estadísticamente correctos, por tanto el número total de datos generados sería de $d = 2^{11}n = 204800$, y el número de entradas distintas de cero o *sparsity* del proceso coincide con el número de datos d . La transformación wavelet discreta fue implementada mediante la caja de herramientas WaveLab (Donoho et al., 1995), con el objetivo que los resultados obtenidos sean fácilmente reproducibles. Los resultados presentados a continuación son obtenidos a partir de esta simulación.

En la Figura 5-3(a) se aprecia la matriz de covarianza del proceso correlacionado x_C , que para este ejemplo es una matriz Toeplitz. Al aplicar la transformación QSP en el dominio wavelet, se observa que el proceso es decorrelacionado, cambiando la forma de su matriz de covarianza a una estructura en diagonal, como se ve en la Figura 5-3(b).



(a) Matriz de covarianza del proceso aleatorio. (b) Matriz de covarianza del proceso aleatorio decorrelacionada.

Figura 5-3: Efecto de la transformación QSP sobre una matriz Toeplitz.

En la Tabla 5-1 se muestran los errores MSE, observando que la transformación QSP es óptima, con errores menores al 1 por ciento, y además se observa que a pesar que la transformada es óptima desde un punto de vista del MSE, su error es muy alto respecto al vector original no correlacionado x .

Por otra parte, de la Tabla 5-1 se observa como el dominio wavelet permite que la re-

Tabla 5-1: Resultados de MSE y *sparsity*.

MSE	x_C	x	<i>sparsity</i>
Φ^*x_C	0.042542	0.44766	94405
$T\Phi^*x_C$	0.08779	0.43509	87648
x_C			204800

presentación del proceso se pueda realizar con muy pocas entradas distintas de cero o *sparsity*, ya que se pasa de tener un número de entradas de x_C de 204,800, a 94405 cuando es transformado al dominio wavelet Φ^*x_C , y sí además se utiliza la transformada QSP, el vector $T\Phi^*x_C$ tiene un nivel de *sparsity* de 87648, por tanto se logra obtener una disminución alrededor de más del 50 por ciento respecto al vector original x_C , y una reducción del 7,15 respecto al vector Φ^*x_C . Ratificando así, las ventajas que se obtienen al aplicar la transformación QSP en el dominio wavelet, ya que el alto nivel de *sparsity* de los coeficientes wavelet permite que la aplicación de la transformación se realice con una menor complejidad computacional.

5.6. Otras aplicaciones de la transformación QSP

5.6.1. Aplicación a cuantización

La transformación QSP permite que la matriz de covarianza tome la estructura que se desee, y es óptima desde el punto de vista del MSE, por tanto es muy útil en estimadores diagonales, los cuales dan origen a representaciones escasas ya que analizan coeficiente por coeficiente, pero también en el uso de los no-diagonales que analizan en bloque a los coeficientes, es decir; toman en cuenta las relaciones entre éstos, por tanto para este caso es interesante utilizar la transformación QSP para correlacionar los coeficientes y luego aplicar el estimador en bloque, mejorando de esta forma el desempeño del estimador.

5.6.2. Aplicación a compresión

La transformación permite que la covarianza de cualquier vector aleatorio x sea modificada por la matriz R , y en aplicaciones como compresión de imágenes, se hace importante explorar los tipos de correlación, tales como correlación espacial, espectral o temporal. Por ejemplo, sea x nuestro dato a comprimir, si se aplica la transformación T_c , las variables aleatorias del vector x estarían correlacionadas, permitiendo que al llevarlo al dominio wavelet mediante la transformación wavelet discreta Φ^* , está pueda extraer la mayor cantidad de información, y sea posible lograr niveles óptimos de *sparsity*. Para recuperar el vector original x a partir de los coeficientes wavelet $\alpha = \Phi^*T_c x$, existirían dos procedimientos: el primero consistiría en aplicar una transformación T_d que se encargaría de decorrelacionar los coeficientes previamente obtenidos y luego

la transformada inversa Φ , el segundo llevaría los coeficientes al dominio del tiempo mediante la transformada inversa Φ para luego aplicar la transformación de decorrelación T_d . Aunque ambos procesos son validos, desde un punto de vista computacional, o equivalentemente, analizando que proceso requiere menos operaciones, de acuerdo a la Tabla 5-1 es mejor el primer procedimiento, ya que cuando se esta en el dominio wavelet, el nivel de *sparsity* es mucho mayor y la transformada T_d se computaría mucho mas rápido. Adicionalmente se podría obtener mas reducciones en la computación si se recurre a optimizaciones lineales o convexas (Boyd y Vandenberghe, 2004). De manera general el proceso de compresión vendría dado por

$$y = \Phi T_d \Phi^* T_c x. \quad (5-51)$$

Es importante resaltar que para tal logro tendría que haber un sacrificio en el valor del MSE, y a su vez tener en cuenta la carga computacional adicional generada al utilizar las transformadas T_d y T_c , además, por supuesto, la de la transformada wavelet. Aunque con las constantes mejoras en la reducción del número de operaciones y por tanto el tiempo de ejecución de la transformada wavelet, posiblemente no afectaría notablemente en el desempeño total.

5.7. Observaciones finales

En este capítulo se exploró y se demostró cuales son las propiedades matriciales que deben compartir la matriz de covarianza del proceso aleatorio y la matriz wavelet, para que la transformación wavelet discreta no tenga un efecto de decorrelación en las variables del vector aleatorio.

Ante estas condiciones se usó la transformación de blanqueamiento propuesta por el QSP, mostrando que al ser utilizada en el dominio wavelet (tiempo-frecuencia) puede llegar a ser más eficiente que al ser utilizada en el dominio del tiempo, dada la característica de *sparsity* que se obtiene en las matrices de covarianza del vector aleatorio, cuando se es analizado en el dominio wavelet. Además se enumeraron posibles aplicaciones de la transformación QSP, que quedan por ser desarrolladas en trabajos futuros.

Capítulo 6

Conclusiones y Trabajos Futuros

6.1. Conclusiones

- El procesamiento cuántico de señales (QSP) al utilizar los principios y problemas de la física cuántica aporta nuevas formas de analizar y concebir la teoría wavelet, lo cual abre a interesantes temas de investigación, tales como cuantización, compresión, detección y optimización que aportan al crecimiento de conocimiento en la teoría wavelet.
- Como hipótesis principal se definió la creación de una nueva forma de representación de la señal, expresada en el ejemplo 1 del Capítulo 3, la cual se ha podido validar en parte al uso de los mapeos determinísticos y probabilísticos, quedando aún muchas otras opciones sin explorar.
- Se demostró que la función de autocorrelación de las funciones wavelet ortogonales cumple con las condiciones para ser utilizada como mapeo en el cuantizador probabilístico desarrollado en el QSP, permitiendo generar señales dither con una distribución dada por la función de autocorrelación, que tiene la ventaja de ser obtenida de forma recursiva a partir de un filtro digital, tal como lo son las funciones wavelet.
- Se observa un gran abanico de posibles aplicaciones del QSP en la teoría wavelet, algunas planteadas en el capítulo 3, y en especial la creación de wavelet uniformes geoméricamente.
- Se exploró la estructura de la matriz wavelet, y se logró determinar un par de condiciones plasmadas en el teorema 5.5, que de ser satisfechas la transformada wavelet discreta no realiza ningún proceso de decorrelación sobre el proceso aleatorio.
- Tanto matemáticamente como por simulación, se demostró las ventajas en el cálculo y la aplicación de la transformación QSP, cuando se es desarrollada en el dominio wavelet.

6.2. Trabajos futuros

- La aplicación del QSP en la teoría wavelet abre un camino en el uso de mapeos probabilísticos en la descomposición de funciones y se enlaza con líneas como *compressed sensing* o *sparse representation*, siendo un muy interesante tópico para investigaciones futuras.
- El marco del QSP brinda una transformada de covarianza óptima desde un punto de vista del MSE, muy útil para decorrelacionar como también para correlacionar los coeficientes wavelet, por ejemplo; un tema interesante de investigación es la optimización de procesos de cuantificación en la cual se utilizan las correlaciones entre los coeficientes.
- Se propone para una siguiente investigación la creación de Wavelets Uniformes Geométricamente, que pueden aportar a los sistemas de codificación.
- Se alcanzó a esbozar la aplicación del estimador desarrollado por el QSP en la teoría wavelet, por su importancia se proyecta como un interesante tema de investigaciones futuras.
- Queda por explorar la aplicación del MF desarrollado en el QSP en un multiplexor que utilice los filtros wavelet como conformadores del pulso y evaluar su desempeño en un canal con ruido aditivo gaussiano y no-gaussiano.

Referencias

- Akansu, A.N. y Haddad, R.A. (2001). *Multiresolution Signal Decomposition (Second Edition)*. Academic Press, San Diego, second edition edition. doi:<http://dx.doi.org/10.1016/B978-012047141-6/50000-8>. URL <http://www.sciencedirect.com/science/article/pii/B9780120471416500008>.
- Akansu, A., Duhamel, P., Lin, X., y de Courville, M. (1998). Orthogonal transmultiplexers in communication: a review. *Signal Processing, IEEE Transactions on*, 46(4), 979–995. doi:10.1109/78.668551.
- Akyol, E. y Rose, K. (2009). Nonuniform dithered quantization. In *Data Compression Conference, 2009. DCC '09.*, 435–435. doi:10.1109/DCC.2009.78.
- Anabtawi, N., Ferzli, R., y Harmanani, H. (2014). Efficient shaped quantizer dithering implementation for sigma delta modulators. In *Electronics, Circuits and Systems (ICECS), 2014 21st IEEE International Conference on*, 766–769. doi:10.1109/ICECS.2014.7050098.
- Atto, A. y Berthoumieu, Y. (2012). Wavelet packets of nonstationary random processes: Contributing factors for stationarity and decorrelation. *Information Theory, IEEE Transactions on*, 58(1), 317–330. doi:10.1109/TIT.2011.2167496.
- Bennett, C., Bessette, F., Brassard, G., Salvail, L., y Smolin, J. (1992). Experimental quantum cryptography. *Journal of Cryptology*, 5, 3–28.
- Beylkin, G., Coifman, R., y Rokhlin, V. (1991). Fast wavelet transforms and numerical algorithms I. *Comm. Pure Appl. Math.*, 44(2), 141–183. doi:10.1002/cpa.3160440202. URL <http://dx.doi.org/10.1002/cpa.3160440202>.
- Boyd, S. y Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, New York, NY, USA.
- Chen, Y., Chi, Y., y Goldsmith, A.J. (2015). Exact and stable covariance estimation from quadratic sampling via convex programming. *Information Theory, IEEE Transactions on To appear*.
- Christensen, O. y Goh, S. (2014). From dual pairs of gabor frames to dual pairs of wavelet frames and vice versa. *Applied and Computational Harmonic Analysis*, 36(2), 198–214. doi:10.1016/j.acha.2013.03.005.

- Cohen, A. y Kovačević, J. (1996). Wavelets: A mathematical background. 84(4), 514–522.
- Coifman, R.R., Meyer, Y., y Wickerhauser, M.V. (1992). Wavelet analysis and signal processing. In M.B. Ruskai, G. Beylkin, R.R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, y L. Raphael (eds.), *Wavelets and Their Applications*, 153–178. Jones and Bartlett, Boston. URL <http://www.math.wustl.edu/victor/papers/wasp.pdf>.
- Coifman, R. y Wickerhauser, M. (1992). Entropy-based algorithms for best basis selection. *Information Theory, IEEE Transactions on*, 38(2), 713–718. doi:10.1109/18.119732.
- Combescure, M. (2009). Block-circulant matrices with circulant blocks, Weil sums, and mutually unbiased bases. II. The prime power case. *Journal of Mathematical Physics*, 50(3), 032104. doi:10.1063/1.3078420.
- Crouse, M., Nowak, R., y Baraniuk, R. (1998). Wavelet-based statistical signal processing using hidden markov models. *Signal Processing, IEEE Transactions on*, 46(4), 886–902. doi:10.1109/78.668544.
- Dasarathy, G., Shah, P., Bhaskar, B., y Nowak, R. (2015). Sketching sparse matrices, covariances, and graphs via tensor products. *Information Theory, IEEE Transactions on*, 61(3), 1373–1388. doi:10.1109/TIT.2015.2391251.
- Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. In *Commun Pure Appl Math* 41:906–966.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- Daubechies, I. (1993). Orthonormal bases of compactly supported wavelets ii: Variations on a theme. *SIAM J. Math. Anal.*, 24(2), 499–519. doi:10.1137/0524031. URL <http://dx.doi.org/10.1137/0524031>.
- Deng-feng, L., Han-lin, C., y Qiu-hui, S. (2003). Matrix approach in wavelet analysis. *Journal of Mathematical Research and Exposition*, 23, 427–436. doi:10.3770/j.issn:1000-341X.2003.03.008.
- Donoho, D., Maleki, A., y Shahram, M. (1995). Wavelab and reproducible research. http://www-stat.stanford.edu/~wavelab/Wavelab_850/index_wavelab850.html. URL http://www-stat.stanford.edu/~wavelab/Wavelab_850/index_wavelab850.html.
- Duffin, R.J. y Schaeffer, A.C. (1952). A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72, 341–366.
- Eldar, Y.C. (2001). *Quantum Signal Processing*. Ph.D. thesis, Massachusetts Institute of Technology.

- Eldar, Y. (2003). Minimum mean-squared error covariance shaping. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, volume 6, VI-713-16 vol.6. doi: 10.1109/ICASSP.2003.1201781.
- Eldar, Y. y Bolcskei, H. (2003). Geometrically uniform frames. *Information Theory, IEEE Transactions on*, 49(4), 993-1006. doi:10.1109/TIT.2003.809602.
- Eldar, Y. y Oppenheim, A. (2001). Orthogonal matched filter detection. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, volume 5, 2837-2840 vol.5. doi: 10.1109/ICASSP.2001.940237.
- Eldar, Y. y Oppenheim, A. (2003). Covariance shaping least-squares estimation. *Signal Processing, IEEE Transactions on*, 51(3), 686-697. doi:10.1109/TSP.2002.808125.
- Feynman, R. y Shor, P.W. (1982). Simulating physics with computers. *SIAM Journal on Computing*, 26, 1484-1509.
- Forney, G. (1991). Geometrically uniform codes. *Information Theory, IEEE Transactions on*, 37(5), 1241-1260. doi:10.1109/18.133243.
- Frisch, M. y Messer, H. (1992). The use of the wavelet transform in the detection of an unknown transient signal. *Information Theory, IEEE Transactions on*, 38(2), 892-897. doi:10.1109/18.119748.
- Gao, X., Nguyen, T., y Strang, G. (2002). A study of two-channel complex-valued filter-banks and wavelets with orthogonality and symmetry properties. *Signal Processing, IEEE Transactions on*, 50(4), 824-833. doi:10.1109/78.992127.
- Gersho, A. y Gray, R.M. (1991). *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Norwell, MA, USA.
- Golub, G.H. y Van Loan, C.F. (1996). *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA.
- Goyal, V.K., Kovačević, J., y Kelner, J.A. (2001). Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10(3), 203 - 233. doi:http://dx.doi.org/10.1006/acha.2000.0340. URL <http://www.sciencedirect.com/science/article/pii/S1063520300903403>.
- Gray, R.M. (2005). Toeplitz and circulant matrices: A review. *Commun. Inf. Theory*, 2(3), 155-239. doi:10.1561/0100000006. URL <http://dx.doi.org/10.1561/0100000006>.
- Grossmann, A. y Morlet, J. (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis*, 15(4), 723-736. doi:10.1137/0515056. URL <http://dx.doi.org/10.1137/0515056>.

- Hartley, A.M. y Naik, D.N. (2001). Estimation of familial correlations under autoregressive circular covariance. *Communications in Statistics - Theory and Methods*, 30(8-9), 1811–1828. doi:10.1081/STA-100105699.
- Hoyos, J.P. y Jojoa, P. (2014). Evaluación de la multiplexación por división de paquetes wavelet. *Ingeniería y Universidad, Pontificia Universidad Javeriana*. doi:10.11144/Javeriana.IYU18-1.emdp.
- Hoyos, J.P. y Jojoa, P. (2015). Design of dynamic wavelet packet division multiplexing. *enviado a IEEE América Latina*.
- Hoyos, J.P. y Jojoa, P. (2016). Wavelet autocorrelation function as mapping for a dithered quantizer. *enviado a la revista Ingeniería e Investigación, Universidad Nacional de Colombia*.
- Jacobs, I.M. y Wozencraft, J.M. (1965). *Principles of communication engineering*. Wiley, New York, NY. URL <https://cds.cern.ch/record/105069>.
- Jamin, A. y Mähönen, P. (2005). Wavelet packet modulation for wireless communications: Research articles. *Wirel. Commun. Mob. Comput.*, 5(2), 123–137. doi:10.1002/wcm.v5:2. URL <http://dx.doi.org/10.1002/wcm.v5:2>.
- Kautsky, J. (1995). An algebraic construction of discrete wavelet transforms. *Applications of Mathematics*, 38, 169–193.
- Kautsky, J. y Turcajová, R. (1995). Discrete biorthogonal wavelet transforms as block circulant matrices. *Linear Algebra and its Applications*, 223–224, 393 – 413. doi:[http://dx.doi.org/10.1016/0024-3795\(94\)00298-R](http://dx.doi.org/10.1016/0024-3795(94)00298-R). URL <http://www.sciencedirect.com/science/article/pii/002437959400298R>. Honoring Miroslav Fiedler and Vlastimil Ptak.
- Kautsky, J. y Turcajova, R. (1994). An matrix approach to discrete wavelets. In L.P. Charles K. Chui Laura Montefusco (ed.), *Wavelets: Theory, Algorithms, and Applications*, 117–135. Academic Press.
- Khattree, R. y Naik, D.N. (1994). Estimation of interclass correlation under circular covariance. *Biometrika*, 81, 612–17.
- Kozaitis, S.P. y Goswami, H. (2000). Wavelet-based image compression using randomized quantization. doi:10.1117/12.390486. URL <http://dx.doi.org/10.1117/12.390486>.
- Kozaitis, S.P. y Goswami, H. (2001). Adaptive threshold selection technique for denoising in dithered quantizers. doi:10.1117/12.421195. URL <http://dx.doi.org/10.1117/12.421195>.
- Kreyszig, E. (1978). *Introductory Functional Analysis with Application*. Wiley.

- Lakshmanan, M. y Nikookar, H. (2006). A review of wavelets for digital wireless communication. *Wireless Personal Communications*, 37(3-4), 387–420. doi:10.1007/s11277-006-9077-y. URL <http://dx.doi.org/10.1007/s11277-006-9077-y>.
- Lara, T. (2001). Matrices circulantes. *Divulgaciones Matemáticas*, 9(1), 85–102.
- Lee, N., Benichi, H., Takeno, Y., Takeda, S., Webb, J., Huntington, E., y Furusawa, A. (2011). Teleportation of nonclassical wave packets of light. *Science*, 332(6027), 330–333. doi:10.1126/science.1201034. URL <http://www.sciencemag.org/content/332/6027/330.abstract>.
- Liang, Y., von Rosen, T., y von Rosen, D. (2011). Block circular symmetry in multilevel models. *Research Report / Department of Statistics, Stockholm University*, 612–17.
- Liew, B., Berber, S., y Sandhu, G. (2005). Performance of a multiple access orthogonal wavelet division multiplexing system. In *Information Technology and Applications, 2005. ICITA 2005. Third International Conference on*, volume 2, 350–353. doi:10.1109/ICITA.2005.213.
- Mallat, S. (2008). *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd edition.
- Meyer, C.D. (ed.) (2000). *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- Newlin, H. (1998). Developments in the use of wavelets in communication systems. In *Military Communications Conference, 1998. MILCOM 98. Proceedings., IEEE*, volume 1, 343–349 vol.1. doi:10.1109/MILCOM.1998.722601.
- Oliveira, H., Soares, L., y Falk, T. (2002). A family of wavelets and a new orthogonal multiresolution analysis based on the nyquist criterion. *Proc. IEEE/SBrT Int. Telecomm. Symp.*, 48(2), 384–389.
- Pamarti, S., Welz, J., y Galton, I. (2007). Statistics of the quantization noise in 1-bit dithered single-quantizer digital delta ndash;sigma modulators. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 54(3), 492–503. doi:10.1109/TCSI.2006.887616.
- Proakis (2007). *Digital Communications 5th Edition*. McGraw Hill.
- S P Lipshitz, R A Wannamker, J.V. (1992). Quantization and dither: A theoretical survey. *Journal of Audio Eng. Soc*, 40, 355–375.
- Saito, N. y Beylkin, G. (1993). Multiresolution representations using the autocorrelation functions of compactly supported wavelets. *Signal Processing, IEEE Transactions on*, 41(12), 3584–3590. doi:10.1109/78.258102.

- Sanyal, A. y Sun, N. (2012). A simple and efficient dithering method for vector quantizer based mismatch-shaped dacs. In *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, 528–531. doi:10.1109/ISCAS.2012.6272082.
- Schuchman, L. (1964). Dither signals and their effect on quantization noise. *Communication Technology, IEEE Transactions on*, 12(4), 162–165. doi:10.1109/TCOM.1964.1088973.
- Shannon, C.E. (2001). A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1), 3–55. doi:10.1145/584091.584093. URL <http://doi.acm.org/10.1145/584091.584093>.
- Strang, G. y Strela, V. (1995). Short wavelets and matrix dilation equations. *Signal Processing, IEEE Transactions on*, 43(1), 108–115. doi:10.1109/78.365291.
- Strang, G. (1989). Wavelets and dilation equations: A brief introduction. *SIAM Review*, 31(4), 614–627. doi:10.1137/1031128. URL <http://dx.doi.org/10.1137/1031128>.
- Strang, G. (2012). Banded matrices with banded inverses and $a = Ipu$. *Proc. Fifth Intl. Congress of Chinese Mathematicians : ICCM2010, International Press and Amer. Math. Soc.*, 771–784.
- Strang, G. y Nguyen, T.Q. (1997). *Wavelets and filter banks*. Wellesley-Cambridge Press.
- Unser, M., Aldroubi, A., y Eden, M. (1992). On the asymptotic convergence of b-spline wavelets to gabor functions. *Information Theory, IEEE Transactions on*, 38(2), 864–872. doi:10.1109/18.119742.
- Vaidyanathan, P.P. (1993). *Multirate Systems and Filter Banks*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- von Neumann, J. (1996). *Mathematical Foundations of Quantum Mechanics*. Princeton University Press.
- Wannamaker, R., Lipshitz, S., Vanderkooy, J., y Wright, J. (2000). A theory of non-subtractive dither. *Signal Processing, IEEE Transactions on*, 48(2), 499–516. doi:10.1109/78.823976.
- West, N. y Scheets, G. (2012). Increasing the resolution of a uniform quantizer using a deterministic dithering signal. In *AUTOTESTCON, 2012 IEEE*, 54–57. doi:10.1109/AUTEST.2012.6334521.
- You, M. y Ilow, J. (2004). A multi-wavelet packet modulation in wireless communications. In *Electrical and Computer Engineering, 2004. Canadian Conference on*, volume 4, 2367–2370 Vol.4. doi:10.1109/CCECE.2004.1347722.