



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

UNA VERIFICACION EMPIRICA DE LA TEORIA
DE LA PARIDAD DEL PODER ADQUISITIVO
MEDIANTE ONDULETAS

T E S I S

QUE PARA OBTENER EL TITULO DE
A C T U A R I O

P R E S E N T A :

JULIO CESAR CERVANTES PARRA



FACULTAD DE CIENCIAS
UNAM

DIRECTOR DE TESIS:
DR. ALBERTO CONTRERAS CRISTAN

2005



FACULTAD DE CIENCIAS
SECRETARIA DE INVESTIGACIONES CIENTÍFICAS
UNAM

m345365



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Julio César Cervantes Parra

FECHA: 9-Junio-2005

FIRMA: JCP

ACT. MAURICIO AGUILAR GONZÁLEZ

Jefe de la División de Estudios Profesionales de la

Facultad de Ciencias

Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

"Una verificación empírica de la Teoría de la Paridad del Poder Adquisitivo mediante Onduletas"

realizado por Julio César Cervantes Parra

con número de cuenta 09437856-0 , quien cubrió los créditos de la carrera de:

Actuaría

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario Dr. Alberto Contreras Cristán

Propietario Dr. Luis Antonio Rincón Solís

Propietario Mtro. Armando Sánchez Vargas

Suplente M. en C. Agustín Ontiveros Pineda

Suplente Act. Noé Moacyr Vallejo González

Consejo Departamental de Matemáticas

Act. Jaime Vázquez Alamilla



FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMÁTICAS

*Para mis padres,
Maria de la Luz y Francisco*

Índice general

Introducción	2
1. La Transformada en Onduletas	4
1.1. ¿Que es una onduleta?	4
1.2. La Transformada Continua en Onduletas: CWT	7
1.3. La Transformada Discreta en Onduletas: DWT	12
2. Marco Teórico	24
2.1. La Ley del Precio Único y La Paridad de Poder Adquisitivo	24
2.1.1. Desviaciones de la Relación PPA	27
2.2. Evidencia Empírica: Resumen Histórico de las Pruebas de la PPA	29
3. Ejercicio Empírico: Metodología y Resultados	34
3.1. Datos	34
3.2. Metodología	36
3.2.1. Análisis de Regresión y Dependencia	39
3.3. Estimaciones	41
Conclusiones	59
Bibliografía	61
Apéndices	63
1. Aspectos técnicos de la DWT	63
El Filtro de Onduleta	63
El Filtro de Escala	68
El Algoritmo Pirámide	72
2. Regresiones	89

Introducción

Una de las variables que disfruta de mayor atención dentro del ámbito económico es el tipo de cambio entre las divisas de los distintos países. Esta es objeto de numerosos estudios acerca de sus complejas interacciones con otras variables y sus repercusiones en aspectos tan diversos como el crecimiento y el comercio internacional. Sin embargo el aspecto que suscita quizá mayor interés es la determinación de sus propios movimientos (principalmente) en el largo plazo; es por ello que existen numerosos modelos teóricos y métodos empíricos que buscan resolver esta cuestión. De particular importancia es el modelo emanado de la teoría de la paridad del poder adquisitivo (PPA) considerado como el punto de partida de la mayoría de los trabajos en ésta área.

La teoría de paridad del poder adquisitivo (PPA) concebida originalmente por Gustav Cassel y basada en la Ley de un sólo precio establece que frente a mercados competitivos y bajo el supuesto de que no existen costos de transporte ni barreras arancelarias, los productos en los diferentes países deben tener el mismo precio. Esto se supone debido a la existencia de arbitraje¹, lo que generaría una tendencia a llegar a un sólo precio en el mismo bien. Bajo esta premisa el tipo de cambio entre dos divisas dadas estará dado por el cociente de precios (en las respectivas monedas) de un mismo bien. Sin embargo la evidencia de los distintos países muestra que difícilmente se cumple esta proposición. Esta discordancia entre los datos observados y el modelo teórico se justifica principalmente debido a las imperfecciones de los mercados que no les permiten hallarse en equilibrio en todo momento, por tanto se considera que la PPA es una relación de largo plazo. No obstante la noción de largo plazo, dentro de este contexto, es un concepto elusivo pues los métodos econométricos hasta ahora propuestos han sido poco exitosos para avalar esta hipótesis. Dichas metodologías no hacen una distinción clara

¹ Arbitraje: Acto de comprar un artículo en un mercado y venderlo en otro a un precio más elevado con el fin de obtener un beneficio con la diferencia de precios entre los dos mercados.

de los movimientos de largo plazo de las variables en cuestión, es por ello que en este trabajo se propone hacer esta distinción como antesala a otro tipo de análisis, para este fin se utilizará una herramienta relativamente novedosa dentro en el contexto econométrico, a saber la DWT o transformada discreta en onduletas.²

La DWT (discrete wavelet transform) es una transformación ortogonal que expresa a las señales (o series) como combinación lineal de un conjunto de funciones básicas conocidas como onduletas y que representan las diferentes escalas temporales y de frecuencias. De esta manera la DWT distingue los movimientos en diferentes escalas temporales de una señal, razón por la cual puede ser apropiada en el análisis de la teoría de la PPA, al disgregar los componentes de largo plazo de la variables y hacer a un lado los movimientos de corto plazo (ocasionados, por ejemplo en el caso del tipo de cambio por los movimientos diarios de los mercados de divisas y que no reflejan la tendencia permanente de la serie).

Por tanto este trabajo tiene como objeto analizar la hipótesis de la PPA mediante el uso del análisis en onduletas con el fin de distinguir entre los movimientos de largo plazo de los de corto plazo que le ocurren tanto al tipo de cambio (entre el dólar estadounidense y la libra esterlina) como a los índices de precios al consumo (tanto de los Estados Unidos como del Reino Unido)³.

En el capítulo 1 se hace una breve recapitulación acerca de la teoría que sustenta el análisis en onduletas, particularmente de su versión discreta, la DWT (discrete wavelet transform), de manera que sea comprensible el análisis empírico subsecuente. Así mismo se busca que este resumen dé cuenta de las principales propiedades que análisis en onduletas puede aportar a la realización de este ejercicio (econométrico). En el capítulo 2 se recapitulan los aspectos económicos de la teoría que subyace a la PPA y se hace una breve síntesis de la pruebas empíricas que en torno a esta hipótesis se han realizado. En el capítulo 3 se establece propiamente el ejercicio empírico de este trabajo y se muestran sus resultados. Finalmente se presenta un capítulo en donde resumen las principales conclusiones, además un par de apéndices donde se tratan los aspectos más técnicos de la DWT y se reportan algunas de las regresiones realizadas en este trabajo.

²Se asume en este trabajo el término *onduleta* como traducción de los términos *ondelette* (original en francés) y *wavelet* (inglés). Es necesario advertir que existe controversia acerca de dicha traducción; algunos autores españoles utilizan como traducción el término *ondícula*.

³El ejercicio empírico se ha reducido a estos dos países por cuestiones de disponibilidad de datos.

Capítulo 1

La Transformada en Onduletas

1.1. ¿Que es una onduleta?

Como su nombre lo sugiere una onduleta es una “onda pequeña”, es decir una función que crece y decae en un periodo de tiempo limitado, (una noción contrastante es a esta es la de “onda grande”, por ejemplo la función *seno* quién se mantiene oscilando sobre todo el eje real). Estos dos conceptos, pueden formularse más formalmente de la siguiente manera; consideremos $\psi(\cdot)$ una función cuyo contradominio pertenece al eje real, entonces esta función es “candidata” para ser onduleta si cumple que:

- La integral de $\psi(\cdot)$ es cero:

$$\int_{-\infty}^{\infty} \psi(u) du = 0 \quad (1.1)$$

- El cuadrado de $\psi(\cdot)$ integra uno:

$$\int_{-\infty}^{\infty} \psi^2(u) du = 1 \quad (1.2)$$

La primera condición es requerida para garantizar un comportamiento ondulatorio de la función (i.e. las desviaciones de la función sobre cero son compensadas por desviaciones bajo de cero), mientras que la segunda asegura que la función es “esencialmente” distinta de cero solo dentro de un

intervalo acotado (condición que, por ejemplo, la función seno no cumple), es decir para cualquier ϵ tal que $0 < \epsilon < 1$ existe un intervalo $[-T, T]$ de longitud finita tal que

$$\int_{-T}^T \psi^2(u) du < 1 - \epsilon$$

Sin embargo es dos condiciones no son suficientes para construir onduletas que tengan propiedades deseables para el análisis de series de tiempo, como la detección de picos y singularidades. Por tanto se añade una más llamada *condición de admisibilidad*; entonces se dice que una ondeleta es admisible si su transformada de Fourier:

$$\Psi(f) \equiv \int_{-\infty}^{\infty} \psi(u) e^{-i2\pi fu} du,$$

es tal que

$$C_\psi \equiv \int_0^\infty \frac{|\Psi(f)|^2}{f} df \quad \text{satisface que} \quad 0 < C_\psi < \infty \quad (1.3)$$

Observemos que la condición (1.1) puede reformularse en términos de la transformada de Fourier como:

$$0 = \Psi(0) = \int_{-\infty}^{\infty} \psi(u) du$$

Así mismo si una función satisface (1.1) y

$$\int_{-\infty}^{\infty} (1 + |u|^\alpha) |\psi(u)| du < \infty \quad \text{para algún} \quad \alpha > 0 \quad (1.4)$$

entonces $C_\psi < \infty$. Es decir (1.1) y (1.4) son suficientes para cumplir la condición de admisibilidad (1.3).

Como ejemplos de estas funciones tenemos en primera instancia, la onduleta de *Haar*, desarrollada en un trabajo pionero por A. Haar, en 1910, y que está definida como:

$$\psi^{(H)}(u) \equiv \begin{cases} -1/\sqrt{2}, & \text{si } -1 < u \leq 0; \\ 1/\sqrt{2}, & \text{si } 0 < u \leq 1; \\ 0, & \text{en otro caso.} \end{cases} \quad (1.5)$$

Para construir otras onduletas consideremos primero la función de densidad de una variable aleatoria normal (gaussiana) con media cero y varianza σ^2 :

$$\phi(u) \equiv \frac{e^{-u^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \quad -\infty < u < \infty$$

su primera derivada es:

$$\frac{d\phi(u)}{du} = -\frac{ue^{-u^2/2\sigma^2}}{\sigma^3\sqrt{2\pi}}$$

renormalizando esta ecuación para satisfacer con la condición (1.2) obtenemos la onduleta

$$\psi^{(fdG)} \equiv \frac{\sqrt{2}ue^{-u^2/2\sigma^2}}{\pi^{1/4}\sigma^{3/2}}$$

Ahora tomando la segunda derivada de la función de densidad gaussiana y realizando la renormalización apropiada se obtiene la onduleta conocida como “Mexican Hat”

$$\psi^{(Mh)} \equiv \frac{2\left(1 - \frac{u^2}{\sigma^2}\right)e^{-u^2/2\sigma^2}}{\pi^{1/4}\sqrt{3}\sigma}$$

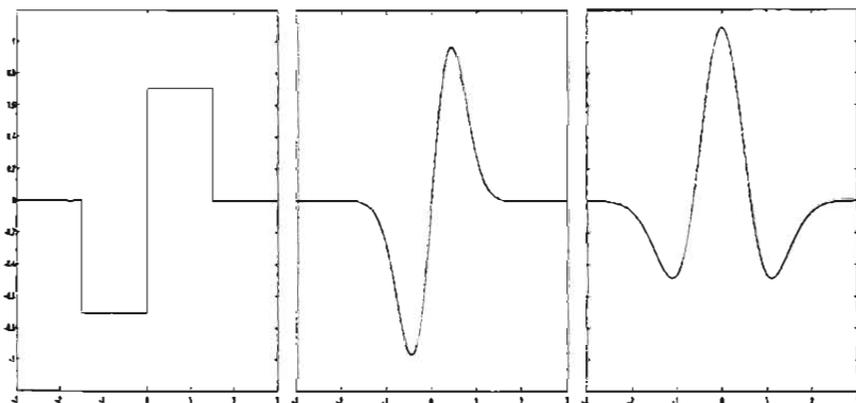


Figura 1.1: De izquierda a derecha: Onduleta de Haar, la primera derivada de la fdp Gaussiana ($\sigma = 0,44311$) y la onduleta Mexican Hat ($\sigma = 0,63628$)

Existen otras onduletas sumamente importantes en las aplicaciones. En primer lugar mencionaremos a las *daubelets*, que fueron el primer tipo de onduletas continuas y ortogonales y cuyo nombre hace honor a una de las pioneras en la investigación en onduletas, Ingrid Daubechies. Las *symmlets* o *least asymmetric wavelets*, son muy similares a las *daubelets* sólo que son tan simétricas como es posible y tienen soporte compacto. Finalmente las *coiflets*, nombradas así en honor de Ronald Coifman, tienen además momentos nulos.

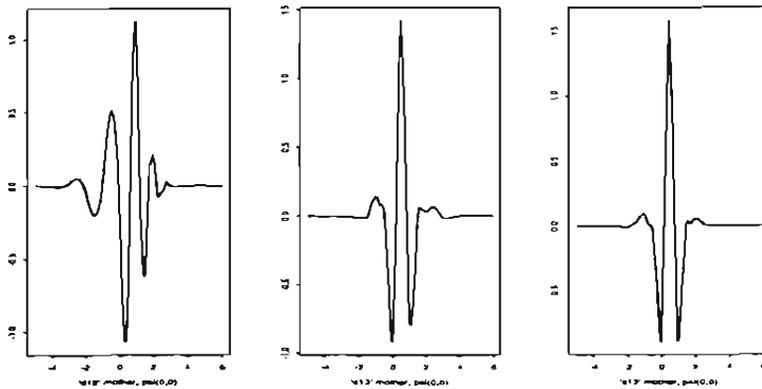


Figura 1.2: De izquierda a derecha: *Daubechies*, *Symmlets* y *Coiflets*

1.2. La Transformada Continua en Onduletas: CWT

La transformada en onduletas, como Daubechies (1992) nos explica, es “una herramienta que corta a las funciones o los datos en pedazos con diferentes componentes de frecuencia y entonces estudia cada componente con una resolución apropiada a su escala¹”, esto se consigue expresando a la función como un promedio ponderado de funciones onduletas con soporte en distintos intervalos de tiempo.

Ahora para introducir la definición de la transformada en onduletas consideremos $x(\cdot)$ una función con valores reales de una variable independiente t a la que, en adelante, llamaremos el “tiempo”. Nos referiremos a $x(\cdot)$ como la señal, por simple conveniencia. Consideremos la integral

$$\frac{1}{b-a} \int_a^b x(u) du \equiv \alpha(a, b) \quad (1.6)$$

donde asumimos que $a < b$ y que $x(\cdot)$ es tal que la integral anterior está bien definida. En los libros elementales de cálculo, $\alpha(a, b)$ se llama el valor promedio de $x(\cdot)$ sobre el intervalo $[a, b]$. Lo anterior está relacionado con la noción de una media muestral de una conjunto de N observaciones. Para ver esta

¹La escala es longitud (medida) del intervalo de tiempo sobre el que se concentra el análisis, en este caso es la longitud del soporte de las funciones que caracterizan los distintos componentes frecuenciales.

conexión, supongan por un momento que $x(\cdot)$ es una función escalonada de la forma:

$$x(t) = x_j \text{ para } a + \frac{j}{N}(b-a) < t \leq a + \frac{j+1}{N}(b-a) \text{ y } j = 0, \dots, N-1$$

de la definición de la integral de Riemann se sigue fácilmente que

$$\frac{1}{b-a} \int_a^b x(u) du = \frac{1}{b-a} \sum_{j=0}^{N-1} x_j \frac{b-a}{N} = \frac{1}{N} \sum_{j=0}^{N-1} x_j$$

En lugar de considerar a $\alpha(a, b)$ como una función de los puntos finales de el intervalo $[a, b]$ sobre el que se realiza la integración, podemos considerarlo simplemente como una función de la longitud del intervalo, i.e. $\lambda \equiv b - a$, y el centro del tiempo del intervalo, i.e. $t = (a + b)/2$. Nos referiremos a λ como la escala asociada con el promedio. Usando λ y t podemos definir

$$A(\lambda, t) \equiv \alpha\left(t - \frac{\lambda}{2}, t + \frac{\lambda}{2}\right) = \frac{1}{\lambda} \int_{t-\frac{\lambda}{2}}^{t+\frac{\lambda}{2}} x(u) du$$

Se llama a $A(\lambda, t)$ el valor promedio de la señal $x(\cdot)$ sobre la escala λ centrada alrededor de t . Ahora definimos el cambio del promedio como:

$$D(1, t - \frac{1}{2}) \equiv A(1, t) - A(1, t - 1) = \int_{t-\frac{1}{2}}^{t+\frac{1}{2}} x(u) du - \int_{t-\frac{3}{2}}^{t-\frac{1}{2}} x(u) du$$

o equivalentemente

$$D(1, t) \equiv A\left(1, t + \frac{1}{2}\right) - A\left(1, t - \frac{1}{2}\right) = \int_t^{t+1} x(u) du - \int_{t-1}^t x(u) du$$

La cantidad $|D(1, t)|$ puede interpretarse como el cambio promedio de la señal por unidad de tiempo alrededor de t . En general el cambio promedio en la escala λ se define como:

$$D(\lambda, t) \equiv A\left(\lambda, t + \frac{\lambda}{2}\right) - A\left(\lambda, t - \frac{\lambda}{2}\right) = \frac{1}{\lambda} \int_t^{t+\lambda} x(u) du - \frac{1}{\lambda} \int_{t-\lambda}^t x(u) du \quad (1.7)$$

Debido a que las dos integrales en (1.7) involucran intervalos adyacentes ajenos, es fácil combinarlos en una sola integral sobre todo el eje real para obtener

$$D(\lambda, t) = \int_{-\infty}^{\infty} \tilde{\psi}_{\lambda, t}(u) x(u) du$$

donde

$$\tilde{\psi}_{\lambda,t}(u) \equiv \begin{cases} -1/\lambda, & \text{si } t - \lambda < u \leq t; \\ 1/\lambda, & \text{si } t < u \leq t + \lambda; \\ 0, & \text{en otro caso.} \end{cases}$$

en el caso especial en el que $\lambda = 1$ y $t = 0$ tenemos:

$$\tilde{\psi}_{1,0}(u) \equiv \begin{cases} -1, & \text{si } -1 < u \leq 0; \\ 1, & \text{si } 0 < u \leq 1; \\ 0, & \text{en otro caso.} \end{cases}$$

Si comparamos lo anterior con la onduleta de Haar $\psi^{(H)}(\cdot)$ de la ecuación (1.5), vemos que $\tilde{\psi}_{1,0}(u) = \sqrt{2}\psi^{(H)}(u)$. Entonces el esquema de mirar a las diferencias de promedios en una escala unitaria en el tiempo $t = 0$ es equivalente (dentro de una proporcionalidad constante) a integrar el producto de la señal $x(\cdot)$ y la onduleta de Haar. En efecto

$$\int_{-\infty}^{\infty} \psi^{(H)}(u)x(u)du \equiv W^{(H)}(1,0),$$

la onduleta de Haar extrae la información acerca de que tan grande es la diferencia entre dos escalas de unidades de $x(\cdot)$ alrededor de $t = 0$.

Es sencillo ajustar la onduleta de Haar para extraer la información acerca de los cambios en escalas unitarias para otros valores de t : simplemente necesitamos cambiar el lugar en que tiene soporte la función $\psi^{(H)}(\cdot)^2$. Definimos

$$\psi_{1,t}^{(H)} \equiv \psi^{(H)}(u-t) \quad \text{tal que} \quad \psi_{1,t}^{(H)}(u) = \begin{cases} -1/\sqrt{2}, & t - 1 < u \leq t; \\ 1/\sqrt{2}, & t < u \leq t + 1; \\ 0, & \text{en otro caso.} \end{cases}$$

Debido a que $\psi_{1,t}^{(H)}(\cdot)$ es sólo una versión trasladada de $\psi^{(H)}(\cdot)$ la función $\psi_{1,t}^{(H)}(\cdot)$ satisface las dos propiedades básicas de una onduleta (integra cero y su cuadrado integra la unidad). La integral del producto de $\psi_{1,t}^{(H)}(\cdot)$ con $x(\cdot)$ es:

$$W^{(H)}(1,t) \equiv \int_{-\infty}^{\infty} \psi_{1,t}^{(H)}(u)x(u)du = \int_t^{t+1} x(u)du - \int_{t-1}^t x(u)du = D(1,t).$$

²A la onduleta básica a partir de la cual se hacen traslaciones y reescalamientos se le conoce comúnmente como onduleta madre.

Es posible extraer información similar para la señal acerca de otras escalas λ y tiempos t considerando³:

$$\psi_{\lambda,t}^{(H)}(u) \equiv \frac{1}{\sqrt{\lambda}} \psi^{(H)}\left(\frac{u-t}{\lambda}\right) = \begin{cases} -1/\sqrt{2\lambda}, & \text{si } t - \lambda < u \leq t; \\ 1/\sqrt{2\lambda}, & \text{si } t < u \leq t + \lambda; \\ 0, & \text{en otro caso.} \end{cases}$$

Es fácil ver que $\psi_{\lambda,t}^{(H)}(\cdot)$ satisface las propiedades de onduletas (1.1) y (1.2). Utilizando esta onduleta obtenemos

$$W^{(H)}(\lambda, t) \equiv \int_{-\infty}^{\infty} \psi_{\lambda,t}^{(H)}(u)x(u)du \propto D(\lambda, t) \quad (1.8)$$

Variando λ podemos obtener “fotografías” de como los promedios de $x(\cdot)$ cambian sobre diferentes escalas de un periodo de longitud λ al siguiente alrededor de t . Cuando λ es grande el soporte de la onduleta es más amplio y por tanto nos ofrece información sobre variaciones de la señal a gran escala, mientras que para λ pequeño el soporte de la onduleta es menos amplio y nos informa sobre variaciones (de los promedios) de la señal en escalas pequeñas. Una manera de visualizar el proceso es en términos de fotografías en las que en las escalas mayores la foto nos da una imagen amplia sin detalles y una foto en las escalas menores completa los detalles (mediante “acercamientos”). De esta manera las señales que oscilen muy rápidamente arrojaran valores mayores de $W^{(H)}(\lambda, t)$ para λ pequeña, mientras que las señales que oscilan más lentamente tendrán mayores valores de $W^{(H)}(\lambda, t)$ para valores grandes de λ .⁴

A la colección de variables $\{W^{(H)}(\lambda, t) : \lambda > 0, -\infty < t < \infty\}$ se conoce como la transformada continua en onduletas de Haar de $x(\cdot)$. La interpretación de esta transformada es que $W^{(H)}(\lambda, t)$ es proporcional a la diferencia entre dos promedios adyacentes de escala λ , con el primer promedio comenzando en t y el segundo promedio terminando en t . De manera similar podemos tomar cualquier otra onduleta $\psi(\cdot)$ y construir la CWT basada en ella mediante:

$$W(\lambda, t) \equiv \int_{-\infty}^{\infty} \psi_{\lambda,t}(u)x(u)du \quad \text{donde} \quad \psi_{t,\lambda}(u) \equiv \frac{1}{\sqrt{|\lambda|}} \psi\left(\frac{u-t}{\lambda}\right) \quad (1.9)$$

³La variable λ modula la longitud del soporte de la onduleta, por ejemplo si el soporte de $\psi_{1,t}(\cdot)$ es $[t-d, t+d]$, entonces el soporte de $\psi_{\lambda,t}(\cdot)$ es $[\lambda(t-d), \lambda(t+d)]$. Para mantener la energía unitaria se añade el término $\frac{1}{\sqrt{\lambda}}$.

⁴Para una ilustración empírica de estas afirmaciones remítase a: Bruce y Gao (1996) cap.2 así como Percival y Walden (2000) cap 4.10

o bien en términos del producto interno

$$W(\lambda, t) = \langle x, \psi_{t,\lambda} \rangle = \int_{-\infty}^{\infty} x(u) \psi_{t,\lambda}(u) du \quad (1.10)$$

la ecuación anterior nos permite interpretar a $\{W(\lambda, t) : \lambda > 0, -\infty < t < \infty\}$ como la proyección de la señal $x(\cdot)$ sobre el conjunto de todas las traslaciones y reescalamientos de la onduleta madre: $\{\psi_{\lambda,t} : \lambda > 0, -\infty < t < \infty\}$ ⁵.

Para las onduletas $\psi^{(fdG)}(\cdot)$ y $\psi^{(Mh)}(\cdot)$ podemos hacer una interpretación de lo que $W(\lambda, t)$ está diciendonos a cerca de $x(\cdot)$ en una forma similar a la onduleta de Haar. Considerando primero a $\psi^{(fdG)}(\cdot)$, se puede argumentar que esta da la diferencia entre promedios ponderados adyacentes; i.e. mientras que la onduleta de Haar se enfoca a diferencias de promedios que son análogos a una media muestral como:

$$\frac{1}{N} \sum_{j=0}^{N-1} x_j$$

la onduleta $\psi^{(fdG)}(\cdot)$ involucra diferencias de promedios ponderados análogos a:⁶

$$\frac{\sum_{j=0}^{N-1} w_j x_j}{\sum_{j=0}^{N-1} w_j}$$

(las ponderaciones de los dos promedios son diferenciadas e invertidas con respecto a cada una debido a las simetrías de $\psi^{(fdG)}(\cdot)$). Por otro lado la onduleta “Mexican hat” $\psi^{(Mh)}(\cdot)$ nos da una diferencia entre un promedio ponderado sobre la escala unitaria y los dos promedios ponderados que están alrededor de él.⁷

Ahora establecemos un hecho fundamental acerca de la CWT; esto es, que preserva toda la información de $x(\cdot)$. Si $\psi(\cdot)$ satisface la condición de admisibilidad (1.3) y si la señal satisface que

$$\int_{-\infty}^{\infty} x^2(t) dt < \infty$$

entonces podemos recuperar a $x(\cdot)$ a partir de su CWT via:

$$x(t) = \frac{1}{C_\psi} \int_0^\infty \left[\int_{-\infty}^\infty W(\lambda, t) \frac{1}{\sqrt{\lambda}} \psi\left(\frac{t-u}{\lambda}\right) du \right] \frac{d\lambda}{\lambda^2} \quad (1.11)$$

⁵Este conjunto haría las veces de una “base” para L^2 . Ogden (1997) muestra que el conjunto ortonormal $\{\psi_{\lambda,t} : \lambda, t \in \mathbb{Z}\}$ es suficiente para aproximar “arbitrariamente bien” cualquier función en L^2 mediante una combinación lineal de elementos en este conjunto.

⁶Percival y Walden (2000) p. 10-11

⁷Idem

por lo tanto una función y su CWT son dos representaciones de la misma entidad matemática; más aún:

$$\int_{-\infty}^{\infty} x^2(t)dt = \frac{1}{C_\psi} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W^2(\lambda, t)dt \right] \frac{d\lambda}{\lambda^2} \quad (1.12)$$

lo anterior dice que $(W(\lambda, t)/\lambda)^2$ define esencialmente una función de densidad de la energía de la señal, que la descompone a través de las diferentes escalas y tiempos. Esto nos permite apreciar la utilidad fundamental de la CWT, a saber, la capacidad de presentar cierta información acerca de $x(\cdot)$ que no es fácilmente disponible, (por ejemplo, en una gráfica de la función $x(\cdot)$ versus t) pues expresa a la señal a partir de sus proyecciones sobre las distintas localizaciones de las onduletas (variable t) y las distintas escalas (variable λ), de este modo quedan de relieve las características fundamentales de la señal, es decir se enfatizan las escalas en las que se presentan las mayores variaciones además de que se localizan en el tiempo.

1.3. La Transformada Discreta en Onduletas: DWT

Como se ha establecido la transformada continua en onduletas para una función de una variable es una función de dos variables, sin embargo dicha transformación es redundante pues cambios muy pequeños en la escala tienen efectos de magnitud similar en la transformada, por tanto la información adicional de dicha frecuencia es muy poca. Para “minimizar” la transformación uno puede seleccionar valores discretos de λ y t (que sirvan como representantes) y tener aún una transformación que es invertible. Sin embargo el muestreo que preserva toda la información acerca de la función descompuesta no puede ser menos fino que cierto valor crítico de muestreo. De acuerdo con Vidakovic (1999) el valor crítico que produce la mínima base está definido por:

$$\lambda = 2^{-j}, \quad t = k2^{-j}, \quad \text{con } j, k \in \mathbb{Z} \quad (1.13)$$

Un muestreo de escalas con estas características es conocido como *diádico*. Cualquier muestreo menos fino (con menos puntos) no permite obtener una única transformación inversa; es decir, la función original no será completamente recuperable. Más aún bajo algunas condiciones más sobre la función onduleta ψ , dicho muestreo produce una base ortogonal para L^2 , a saber, $\{\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k) : j, k \in \mathbb{Z}\}$

Hay otras elecciones de discretización; por ejemplo un muestreo más general está dado por:

$$\lambda = \lambda_0^{-j}, \quad t = kt_0\lambda_0^{-j}, \quad \text{con } j, k \in \mathbb{Z}, \quad \lambda_0 > 1, t_0 > 0$$

para dicho muestreo también son posibles reconstrucciones numéricamente estables si el sistema constituye un “frame”⁸. En este caso

$$\psi_{j,k}(x) = \lambda_0^{j/2} \psi \left(\frac{x - kt_0\lambda_0^{-j}}{\lambda_0^{-j}} \right) = \lambda_0^{j/2} \psi \left(\lambda_0^j x - kt_0 \right)$$

En lo sucesivo nos ocuparemos de escalas diádicas, lo cual no constituye una restricción grave, pues como ya se ha establecido, la DWT de una señal discreta y ella misma son dos representaciones de la misma entidad matemática, para este tipo de escalas, es decir, es posible reconstruir la señal a partir de su DWT.

Aunque hemos introducido la DWT como un submuestreo de la CWT, esta se puede justificar independientemente como una poderosa herramienta para el análisis de las series de tiempo debido a sus propiedades prácticas. En primera instancia recordemos que en las aplicaciones se tienen datos discretos, los cuales no permiten una extensión obvia para la señal en tiempo continuo y por ende de su CWT, se recurre entonces a la DWT que esta diseñada para series discretas y arroja interpretaciones similares a las de CWT. Algunas otras características de la DWT que Percival y Walden (2000) enfatizan son la capacidad de la transformada para remover la correlación en algunas series⁹ y la eficiencia de su cálculo a través del algoritmo pirámide o piramidal, el cual requiere menos operaciones que el algoritmo de la transformada rápida de Fourier.¹⁰ Además la partición de la energía de la serie de tiempo que resulta de la DWT está directamente asociada

⁸Vidakovic (1999). Un conjunto numerable $\{f_n | n \in \mathbb{N}\}$ de un espacio de Hilbert separable \mathcal{H} constituye una frame si y sólo si existen constantes A y B ($0 < A \leq B < \infty$) tales que para toda $x \in \mathcal{H}$

$$A \cdot \|x\|^2 \leq \sum_{n \in \mathbb{Z}} |\langle x, f_n \rangle|^2 \leq B \cdot \|x\|^2$$

⁹No es el caso de las series estudiadas en este trabajo.

¹⁰Este requiere solamente $O(N)$ multiplicaciones para ser calculado, mientras que calcular directamente el producto de la matriz \mathcal{W} de $N \times N$ y el vector \mathbf{X} requiere N^2 multiplicaciones (la notación $O(a_N)$ significa que existe una constante C tal que el verdadero número de multiplicaciones es menor o igual a Ca_N para todo N). Este algoritmo es, inclusive, más rápido que el algoritmo de la transformada rápida de Fourier que requiere $O(N \log_2 N)$ multiplicaciones.

con las diferentes escalas y tiempos, así esta descomposición de energía es muy parecida a un análisis de varianza (ANOVA), basado en la escala que es análogo al análisis de varianza basado en el poder del espectro (dado por la transformada de Fourier).

Ahora estudiaremos más a detalle las propiedades de la DWT, primero como transformada ortonormal y luego desde la perspectiva de filtros. Estas dos alternativas nos permitirán interpretar los coeficientes de la DWT como proyecciones de la señal sobre subespacios con distintos componentes frecuenciales. Comenzamos definiendo nuestro objeto de estudio: la señal. Sea X_0, X_1, \dots, X_{N-1} una serie de tiempo de N variables con valores reales, en lo sucesivo denotaremos a la serie como $\{X_t : t = 0, \dots, N-1\}$ o solamente $\{X_t\}$ si es claro cuales son los valores del índice. Sea \mathbf{X} el vector columna de dimensión N cuyo t -ésimo elemento es X_t para $t = 0, \dots, N-1$. Asumiremos también que $N = 2^J$ para algún entero positivo J .

Al igual que la transformada de Fourier discreta ortonormal, la transformada en onduletas discreta (DWT) de $\{X_t\}$ es una transformada ortonormal. A continuación asumiremos que existe una matriz \mathcal{W} de valores reales y dimensiones $N \times N$ que define a la DWT y que satisface que $\mathcal{W}^T \mathcal{W} = I_N$. Más adelante discutiremos respecto a como definir sus renglones. Sea $\mathbf{W} = \mathcal{W}\mathbf{X}$, donde \mathcal{W} es la matriz antes mencionada, entonces denotamos por $\{W_n : n = 0, \dots, N-1\}$ a los elementos del vector (columna) \mathbf{W} de dimensión $N = 2^J$, cuyo n -ésimo elemento es el n -ésimo coeficiente W_n de la DWT.

La ortonormalidad de \mathcal{W} implica que $\mathbf{X} = \mathcal{W}^T \mathbf{W}$ y $\|\mathbf{W}\|^2 = \|\mathbf{X}\|^2$. Entonces W_n^2 representa la contribución de la energía atribuible al coeficiente de la DWT cuyo índice es n .

Ahora con el objetivo de mostrar que el n -ésimo coeficiente W_n está asociado con una escala particular y un conjunto particular de tiempos, consideremos la matriz \mathcal{W} asociada a la onduleta de Haar para $N = 16$. Las filas de esta matriz para $n = 0, 8, 12, 14$ y 15 son:¹¹

$$\begin{aligned} \mathcal{W}_{0\bullet}^T &= \left[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \underbrace{0, \dots, 0}_{14 \text{ ceros}}\right], & \mathcal{W}_{8\bullet}^T &= \left[-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \underbrace{0, \dots, 0}_{12 \text{ ceros}}\right] \\ \mathcal{W}_{12\bullet}^T &= \left[-\frac{1}{\sqrt{8}}, \dots, -\frac{1}{\sqrt{8}}, \frac{1}{\sqrt{8}}, \dots, \frac{1}{\sqrt{8}}, \underbrace{0, \dots, 0}_{8 \text{ ceros}}\right] \\ &\quad \underbrace{\hspace{2cm}}_{4 \text{ de estos}} \quad \underbrace{\hspace{2cm}}_{4 \text{ de estos}} \\ \mathcal{W}_{14\bullet}^T &= \left[-\frac{1}{4}, \dots, -\frac{1}{4}, \frac{1}{4}, \dots, \frac{1}{4}\right] & \mathcal{W}_{15\bullet}^T &= \left[\frac{1}{4}, \dots, \frac{1}{4}\right] \\ &\quad \underbrace{\hspace{2cm}}_{8 \text{ de estos}} \quad \underbrace{\hspace{2cm}}_{8 \text{ de estos}} \quad \underbrace{\hspace{2cm}}_{16 \text{ de estos}} \end{aligned}$$

las once filas restantes de la matriz son versiones trasladadas de las anterior-

¹¹En adelante se hará uso de la notación $A_{j\bullet}$ y $A_{\bullet k}$ para referirnos a la j -ésima fila y la k -ésima columna respectivamente de una matriz \mathbf{A} .

res:

$$\begin{aligned} W_{1\bullet} &= T^2 W_{0\bullet} & W_{2\bullet} &= T^4 W_{0\bullet} & \dots & W_{7\bullet} &= T^{14} W_{0\bullet} \\ W_{9\bullet} &= T^4 W_{8\bullet} & W_{10\bullet} &= T^8 W_{8\bullet} & W_{11\bullet} &= T^{12} W_{8\bullet} \\ W_{13\bullet} &= T^8 W_{12\bullet} \end{aligned}$$

donde T es una matriz de traslaciones ¹² y cuya forma específica es

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix}$$

Retomemos ahora la noción de escala para el caso de una señal discreta. Para un entero positivo λ sea

$$\bar{X}_t(\lambda) \equiv \frac{1}{\lambda} \sum_{l=0}^{\lambda-1} X_{t-l} \quad (1.14)$$

que representa el promedio de los λ valores (datos) contiguos con índices desde $t - \lambda + 1$ hasta t (notemos que $\bar{X}_t(1) = X_t$, a quién podemos ver simplemente como “un promedio de un sólo punto” y que $\bar{X}_{N-1}(N) = \bar{X}$ que es el promedio muestral de todos los N valores). Nos referiremos a $\bar{X}_t(\lambda)$ como el promedio muestral para la escala λ sobre el conjunto de tiempos $t - \lambda + 1$ a t . Como $\mathbf{W} = \mathbf{W}\mathbf{X}$ es posible escribir:

$$\mathbf{W} = \begin{bmatrix} W_0 \\ \vdots \\ W_7 \\ W_8 \\ \vdots \\ W_{11} \\ W_{12} \\ W_{13} \\ W_{14} \\ W_{15} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}(X_1 - X_0) \\ \vdots \\ \frac{1}{\sqrt{2}}(X_{15} - X_{14}) \\ \frac{1}{2}(X_3 + X_2 - X_1 - X_0) \\ \vdots \\ \frac{1}{2}(X_{15} + X_{14} - X_{13} - X_{12}) \\ \frac{1}{\sqrt{8}}(X_7 + \dots + X_4 - X_3 - \dots - X_0) \\ \frac{1}{\sqrt{8}}(X_{15} + \dots + X_{12} - X_{11} - \dots - X_8) \\ \frac{1}{4}(X_{15} + \dots + X_8 - X_7 - \dots - X_0) \\ \frac{1}{4}(X_{15} + \dots + X_0) \end{bmatrix}$$

Usando la definición para $\bar{X}_t(\lambda)$ podemos escribir a las W_n 's como:

¹²La matriz de traslaciones es una transformación (matriz) ortogonal ya que $T^{-1} = T^T$

$$\begin{aligned}
W_0 &= \frac{1}{\sqrt{2}}[\bar{X}_1(1) - \bar{X}_0(1)], \dots, W_7 = \frac{1}{\sqrt{2}}[\bar{X}_{15}(1) - \bar{X}_{14}(1)] \\
W_8 &= \bar{X}_3(2) - \bar{X}_1(2), \dots, W_{11} = \bar{X}_{15}(2) - \bar{X}_{13}(2) \\
W_{12} &= \sqrt{2}[\bar{X}_7(4) - \bar{X}_3(4)], \quad W_{13} = \sqrt{2}[\bar{X}_{15}(4) - \bar{X}_{11}(4)] \\
W_{14} &= 2[\bar{X}_{15}(8) - \bar{X}_7(8)] \\
W_{15} &= \bar{X}_{15}(16)
\end{aligned}$$

A partir de las ecuaciones anteriores notemos que los primeros ocho coeficientes de la DWT, W_0, \dots, W_7 , son proporcionales a diferencias (cambios) en promedios adyacentes de $\{X_t\}$ en la escala unitaria; los siguientes cuatro coeficientes, i.e. W_8, \dots, W_{11} son diferencias de promedios adyacentes en la escala de dos; W_{12} y W_{13} son proporcionales a diferencias en la escala de cuatro, W_{14} es proporcional a una diferencia en la escala de ocho y finalmente el coeficiente W_{15} es proporcional al promedio de todos los datos. Observemos que las ecuaciones anteriores tienen como análogo en el caso continuo a la ecuación (1.8), de esta manera la DWT preserva la interpretación de la CWT en cuanto a que cada coeficiente de onduletas es proporcional a la diferencia entre dos promedios adyacentes en su escala correspondiente.

Generalizando para $N = 2^J$ y para la onduleta de Haar, los elementos de \mathbf{W} pueden organizarse de tal manera que los primeros $N/2$ coeficientes están asociados con cambios en la escala unitaria; los siguientes $N/4$ coeficientes, con cambios en la escala de dos; y así sucesivamente hasta que llegamos a los coeficientes W_{N-4} y W_{N-3} , que están asociados con cambios en la escala de $N/4$; luego el coeficiente W_{N-2} está asociado con un cambio en la escala de $N/2$; y finalmente W_{N-1} es proporcional al promedio de todos los datos (i.e. en la escala N).¹³ Entonces hay exactamente $N/(2\tau_j)$ coeficientes de la DWT asociados con cambios en la escala de τ_j , donde $\tau_j \equiv 2^{j-1}$ para $j = 1, \dots, J$, (notemos que $\tau_1 = 1$ y $\tau_J = N/2$); adicionalmente hay un sólo coeficiente W_{N-1} asociado con un promedio en la escala de N . Los $N - 1$ coeficientes que están asociados son cambios en varias escalas son llamados *coeficientes de onduletas*, mientras que W_{N-1} es llamado el *coeficiente de escala*. El patrón observado en la DWT de Haar es similar para todas las onduletas que pertenecen a la clase de Daubechies¹⁴, en particular para la

¹³Percival y Walden (2000) p. 59

¹⁴Las onduletas de Daubechies son precisamente aquellas que preservan la interpretación descrita anteriormente para los coeficientes de su DWT, es decir que el i -ésimo coeficiente de onduletas está asociado a la escala $\tau_j = 2^{j-1}$ para j tal que $\frac{2^{j-1}-1}{2^{j-1}} < \frac{i}{N} \leq \frac{2^j-1}{2^j}$ o como veremos más adelante está asociado a frecuencias en el intervalo $[1/2^{j+1}, 1/2^j]$; es necesario aclarar que esta interpretación no es una consecuencia automática de la definición de una onduleta. Estas onduletas están asociadas a filtros obtenidos mediante la *factorización espectral* del módulo al cuadrado de su transformada de Fourier. Refiérase a Percival y Walden (2000) capítulo 4.8

onduleta que se utiliza en el ejercicio empírico de este trabajo, la *symmlet* o *least asymmetric wavelet*.

Es importante mencionar que τ_j es una escala estandarizada sin unidades; sin embargo en las aplicaciones prácticas, es necesario tener en cuenta el intervalo Δt entre las observaciones, en cuyo caso τ_j corresponde a la escala $\tau_j \Delta t$, quién si tiene interpretación práctica.

Descompongamos los elementos del vector \mathbf{W} en $J + 1$ subvectores. Los primeros J subvectores se denotan por \mathbf{W}_j , $j = 1, \dots, J$, y el j -ésimo vector contiene todos los coeficientes de la DWT para la escala τ_j . Note que \mathbf{W}_j es un vector columna con $N/2^j$ elementos. El subvector final está denotado como \mathbf{V}_J y contiene sólo al coeficiente de escala W_{N-1} . Es posible escribir

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \dots \\ \mathbf{W}_J \\ \mathbf{V}_J \end{bmatrix}$$

cuando $N = 2^J = 16$ de tal manera que $J = 4$ entonces tenemos

$$\mathbf{W}_1^T = [W_0, W_1, W_2, W_3, W_4, W_5, W_6, W_7]$$

$$\mathbf{W}_2^T = [W_8, W_9, W_{10}, W_{11}]$$

$$\mathbf{W}_3^T = [W_{12}, W_{13}]$$

$$\mathbf{W}_4^T = [W_{14}]$$

$$\mathbf{V}_4^T = [W_{15}]$$

Análogamente definamos a \mathcal{W}_j y \mathcal{V}_J como las matrices que particionando las filas de \mathcal{W} conmesuran con las particiones de \mathbf{W} en $\mathbf{W}_1, \dots, \mathbf{W}_j$ y \mathbf{V}_J . Entonces la matriz \mathcal{W}_1 de $\frac{N}{2} \times N$ está formada por las filas desde $n = 0$ hasta $n = \frac{N}{2} - 1$ de \mathcal{W} ; la matriz \mathcal{W}_2 de $\frac{N}{4} \times N$ está formada por las filas desde $n = \frac{N}{2}$ hasta $n = \frac{3N}{4} - 1$; y así sucesivamente hasta que llegamos a las matrices \mathcal{W}_J y \mathcal{V}_J , que son las últimas dos filas de \mathcal{W} . Entonces tenemos que:

$$\mathcal{W} = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{W}_2 \\ \vdots \\ \mathcal{W}_J \\ \mathcal{V}_J \end{bmatrix} \quad (1.15)$$

donde \mathcal{W}_j es una matriz de $\frac{N}{2^j} \times N$ para $j = 1, \dots, J$ y \mathcal{V}_J es un vector fila de N elementos (y que de hechos todos son iguales a $\frac{1}{\sqrt{N}}$). En el ejemplo

de la DWT con $N = 16$, \mathcal{W}_1 es un matriz de 8×16 cuyas filas son las ocho primeras filas de \mathcal{W} ; i.e.

$$\mathcal{W}_1 = [\mathcal{W}_{0\bullet}, \mathcal{W}_{1\bullet}, \mathcal{W}_{2\bullet}, \mathcal{W}_{3\bullet}, \mathcal{W}_{4\bullet}, \mathcal{W}_{5\bullet}, \mathcal{W}_{6\bullet}, \mathcal{W}_{7\bullet}]^T;$$

de la misma manera \mathcal{W}_2 es la matriz de 4×16 dada por:

$$\mathcal{W}_2 = [\mathcal{W}_{8\bullet}, \mathcal{W}_{9\bullet}, \mathcal{W}_{10\bullet}, \mathcal{W}_{11\bullet}]^T$$

mientras que \mathcal{W}_3 , \mathcal{W}_4 y \mathcal{V}_4 son las matrices de 2×16 , 1×16 y 1×16 respectivamente dadas por:

$$\mathcal{W}_3 = [\mathcal{W}_{12\bullet}, \mathcal{W}_{13\bullet}]^T, \quad \mathcal{W}_4 = \mathcal{W}_{14\bullet}^T \quad \text{y} \quad \mathcal{V}_4 = \mathcal{W}_{15\bullet}^T.$$

La condición de conservación de la energía la podemos escribir como:

$$\|\mathbf{X}\|^2 = \|\mathbf{W}\|^2 = \sum_{j=1}^J \|\mathbf{W}_j\|^2 + \|\mathbf{V}_J\|^2$$

de tal manera que $\|\mathbf{W}_j\|^2$ representa la contribución de la energía de $\{X_t\}$ debida a cambios en la escala τ_j . Para todas las DWTs formadas utilizando alguna ondeleta de la clase de Daubechies se cumple que $\mathcal{W}_{N-1}/\sqrt{N} = \bar{X}$; ¹⁵ esto implica que $\|\mathbf{V}_J\| = N\bar{X}^2$ y es posible descomponer la varianza como:

$$\hat{\sigma}_X^2 = \frac{1}{N} \|\mathbf{X}\|^2 - \bar{X}^2 = \frac{1}{N} \|\mathbf{W}\|^2 - \bar{X}^2 = \frac{1}{N} \sum_{j=1}^J \|\mathbf{W}_j\|^2 \quad (1.16)$$

de tal manera que $\|\mathbf{W}_j\|^2/N$ representa la contribución de la varianza muestral de $\{X_t\}$ debida a cambios en la escala τ_j . Esta descomposición de la varianza muestral puede ser utilizada para poder definir un poder del espectro empírico de la ondeleta $\{P_{\mathcal{W}}(\tau_j) : \tau_j = 1, 2, 4, \dots, N/2\}$ para $\{X_t\}$ como

$$P_{\mathcal{W}}(\tau_j) \equiv \frac{1}{N} \|\mathbf{W}_j\|^2, \quad \text{para el cual tenemos} \quad \sum_{j=1}^J P_{\mathcal{W}}(\tau_j) = \hat{\sigma}_X^2$$

Consideremos ahora la síntesis en onduletas de \mathbf{X}

$$\mathbf{X} = \mathcal{W}^T \mathbf{W} = \sum_{n=0}^{N-1} W_n \mathcal{W}_{n\bullet} = \sum_{j=1}^J \mathcal{W}_j^T \mathbf{W}_j + \mathcal{V}_J^T \mathbf{V}_J \quad (1.17)$$

¹⁵Para una demostración de este hecho remitase a Percival y Walden (2000) p.517

Ahora definamos $\mathcal{D}_j \equiv \mathcal{W}_j^T \mathbf{W}_j$ para $j = 1, \dots, J$, el cual es un vector columna de dimensión N cuyos elementos están asociados con cambios en \mathbf{X} en la escala τ_j ,¹⁶ i.e. $\mathbf{W}_j = \mathcal{W}_j \mathbf{X}$ representa la porción del análisis $\mathbf{W} = \mathcal{W} \mathbf{X}$ atribuible a la escala τ_j , mientras que $\mathcal{W}_j^T \mathbf{W}_j$ es la porción de la síntesis $\mathbf{X} = \mathcal{W}^T \mathbf{W}$ atribuible a la escala τ_j . Sea $\mathcal{S}_J \equiv \mathcal{V}_J^T \mathbf{V}_J$, el cual tiene todos sus elementos iguales a la media muestral \bar{X} .¹⁷ Podemos escribir

$$\mathbf{X} = \sum_{j=1}^J \mathcal{D}_j + \mathcal{S}_J \quad (1.18)$$

la cual define un *análisis multiresolución* (MRA) de \mathbf{X} ; i.e. expresa a la serie \mathbf{X} como la suma de un vector constante \mathcal{S}_J y otros J vectores \mathcal{D}_j , $j = 1, \dots, J$, cada uno de los cuales contiene una serie de tiempo relacionada con variaciones in \mathbf{X} a cierta escala.¹⁸ Nos referiremos \mathcal{D}_j como el detalle al j -ésimo nivel de onduletas.

La ortonormalidad de \mathcal{W} implica que para $1 \leq j, k \leq J$,

$$\mathcal{D}_j^T \mathcal{D}_k = \mathbf{W}_j^T \mathcal{W}_j \mathcal{W}_k^T \mathbf{W}_k = \begin{cases} \mathbf{W}_j^T \mathbf{W}_j, & k = j \\ 0, & \text{en otro caso} \end{cases}$$

entonces $\|\mathcal{D}_j\|^2 = \|\mathbf{W}_j\|^2$ y usando la ecuación (1.16), tenemos que

$$\hat{\sigma}_{\mathbf{X}}^2 = \frac{1}{N} \sum_{j=1}^J \|\mathcal{D}_j\|^2, \quad j = 1, \dots, J,$$

donde $\|\mathcal{D}_j\|^2 / N$ puede ser interpretada como la varianza muestral de N elementos de \mathcal{D}_j . Entonces el poder del espectro discreto de onduletas $P_{\mathcal{W}}(\tau_j) \equiv \|\mathbf{W}_j\|^2 / N$ puede ser expresado en términos de los detalles como $P_{\mathcal{W}}(\tau_j) = \|\mathcal{D}_j\|^2 / N$

Recordemos que hemos definido $\mathcal{S}_J = \mathcal{V}_J^T \mathbf{V}_J$ como un vector cuyos elementos son todos \bar{X} . Para $0 \leq j \leq J - 1$ sea

$$\mathcal{S}_j \equiv \sum_{k=j+1}^J \mathcal{D}_k + \mathcal{S}_J \quad (1.19)$$

¹⁶Retomando la analogía con el caso continuo Vidakovic(1999) y Ogden(1997) muestran como cada \mathcal{D}_j es la proyección de la señal \mathbf{X} sobre el subespacio de funciones $W_j = \text{span} \{\psi_{j,k} : k \in \mathbb{Z}\}$

¹⁷Ibid. En cambio \mathcal{S}_J es la proyección de la señal en el subespacio $V_0 = \{f \in L^2(\mathbb{R}) : f \text{ es constante por tramos en } [k, (k+1)), k \in \mathbb{Z}\}$.

¹⁸Ibid. En este caso la señal se ha proyectado sobre el espacio $V_J = \{f \in L^2(\mathbb{R}) : f \text{ es constante por tramos en } [k2^{-J}, (k+1)2^{-J}), k \in \mathbb{Z}\}$. Muestran también que $V_J = \bigoplus_{j=0}^{J-1} W_j \oplus V_0$ de tal suerte que V_J es una aproximación al espacio $L^2(\mathbb{R})$ al que se le puede ver como $\bigoplus_{j \in \mathbb{N}} W_j \oplus V_0 = L^2(\mathbb{R})$. (Ogden, 1996, p.18)

y para $j \geq 1$

$$\mathbf{X} - \mathcal{S}_j = \sum_{k=1}^j \mathcal{D}_k$$

podemos argumentar que \mathcal{S}_j es una versión suavizada de \mathbf{X} por que la diferencia entre los dos vectores involucra sólo detalles en la escala $\tau_j = 2^{j-1}$ y menores, entonces mientras en índice j crece, \mathcal{S}_j debe ser más suave (como en el caso cuando $j = J$ donde todos los elementos son el mismo). No referiremos a \mathcal{S}_j como el j -ésimo *wavelet smooth* para \mathbf{X} . Para ilustrar los conceptos hasta ahora definidos tomemos como ejemplo a la función doppler a la que se le ha añadido un ruido blanco.¹⁹ El grafico siguiente nos muestra en la parte superior a los datos generados, mientras que en la columna izquierda se muestran los detalles hasta el cuarto nivel, en la derecha se muestran los *smooths* también hasta el cuarto nivel, donde es posible apreciar como se tienen “versiones” cada vez más suavizadas de la señal.

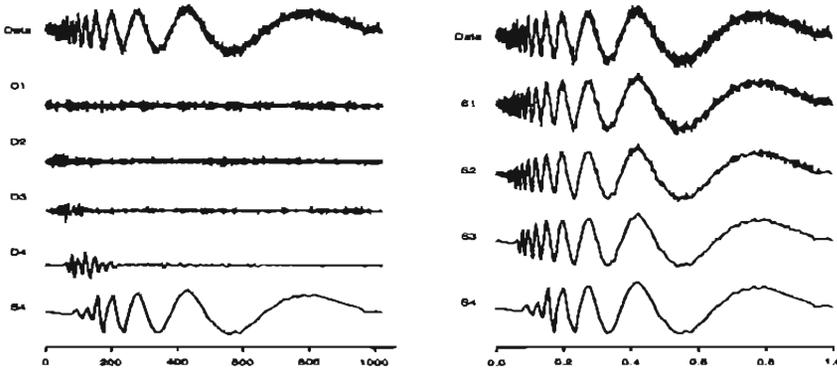


Figura 1.3: Ejemplo

Similarmente definimos el j -ésimo *wavelet rough* para \mathbf{X} como:

$$\mathcal{R}_j \equiv \begin{cases} 0, & j = 0; \\ \sum_{k=1}^j \mathcal{D}_k, & 1 \leq j \leq J \end{cases} \quad (1.20)$$

entonces tenemos que $\mathbf{X} = \mathcal{S}_j + \mathcal{R}_j$ para toda j , donde \mathcal{S}_j está asociado con las escalas τ_{j+1} y mayores, mientras que \mathcal{R}_j está asociado a la escala

¹⁹Ejemplo tomado de Bruce y Gao (1996) pp. 27-30. La función doppler está definida como $\text{doppler}(t) = \sqrt{t(1-t)} \sin\left(\frac{2.1\pi}{t+0.05}\right)$.

τ_j y menores. Notemos que $S_j - S_{j+1} = D_{j+1}$ y $R_j - R_{j+1} = D_{j+1}$; i.e. los detalles son las diferencias entre los *smooths* y los *roughs* adyacentes.

Antes de concluir este capítulo, describiremos muy brevemente las características más relevantes, para este trabajo, del algoritmo que permite el cálculo de la DWT de una señal. El algoritmo pirámide tiene como componentes fundamentales un par de filtros, a saber, el filtro de onduletas $\{h_l\}$ y el filtro de escala $\{g_l\}$. El filtro de onduletas es una sucesión finita de números reales, que está caracterizada por tres condiciones, dos de ellas tienen su símil en las ecuaciones (1.1) y (1.2), es decir la suma de todos los elementos del filtro es igual a cero y la suma de sus cuadrados es igual a uno; se añade además otra condición más bien técnica que asegura que los renglones de la matriz \mathcal{W} , que serán formadas mediante este filtro, sean ortogonales entre sí.²⁰ Apartir del filtro de onduleta se construye el llamado filtro de escala, mediante su “cuadratura espejo”, esto es, $g_l \equiv (-1)^{l+1}h_{L-1-l}$. Dada su construcción este último filtro comparte algunas características con el filtro de onduletas y en otros casos tiene propiedades “duales” a él. Por ejemplo aunque también cumple la condición de ortonormalidad y la suma de sus elementos al cuadrado es igual a uno, la suma de ellos es $\sqrt{2}$ o $-\sqrt{2}$.²¹

El algoritmo pirámide, en su primera etapa, filtra la señal mediante $\{h_l\}$ y $\{g_l\}$, para obtener los $N/2$ coeficientes de onduletas y los $N/2$ coeficientes de escala de primer nivel respectivamente (se realiza un filtrado circular y un procedimiento llamado *submuestreo*, que consiste en “descartar” todos los coeficientes de subíndice par resultantes del filtrado²²). Se forman entonces dos vectores; con los coeficientes de onduletas se obtiene \mathbf{W}_1 (que a su vez constituye la primera mitad del vector \mathbf{W} y permite calcular el primer detalle D_1) y con los coeficientes de escala se forma el vector \mathbf{V}_1 (que servirá de insumo para la segunda etapa del algoritmo al ser tratado en manera similar a la que en la primera se trató a \mathbf{X} ; este vector también permite calcular el *smooth* de primer nivel S_1). Estas dos series de coeficientes aportan información de la señal en distintos intervalos de frecuencia. Así mientras que los coeficientes de onduletas llevan la información de al-

²⁰Formalmente las condiciones se enuncian de la siguiente manera:

$$\sum_{l=0}^{L-1} h_l = 0, \quad \sum_{l=0}^{L-1} h_l^2 = 1 \quad \text{y} \\ \sum_{l=0}^{L-1} h_l h_{l+2n} = \sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0;$$

estas propiedades hacen pensar en el filtro $\{h_l\}$ como la versión discreta de la onduleta básica.

²¹Para más detalles ver Percival y Walden (2000) cap. 4.2 y 4.3

²²Más detalles en Percival y Walden (2000) p.80-85

tas frecuencias, (es decir en el intervalo $[1/4, 1/2]$) los coeficientes de escala aportan información acerca de los componentes de baja frecuencia (en el intervalo $[0, 1/4]$). Esto ocurre debido a que el filtro $\{h_l\}$, puede ser visto como un filtro *high-pass*, mientras que el filtro $\{g_l\}$ es equivalente a un filtro *low-pass*. Esta afirmación se constata através de la gráfica del módulo al cuadrado de la transformada de Fourier de cada uno de los filtros ²³. Se muestran como ejemplo las gráficas para los filtros de onduletas y de escala de Haar y Daubechies D(4).

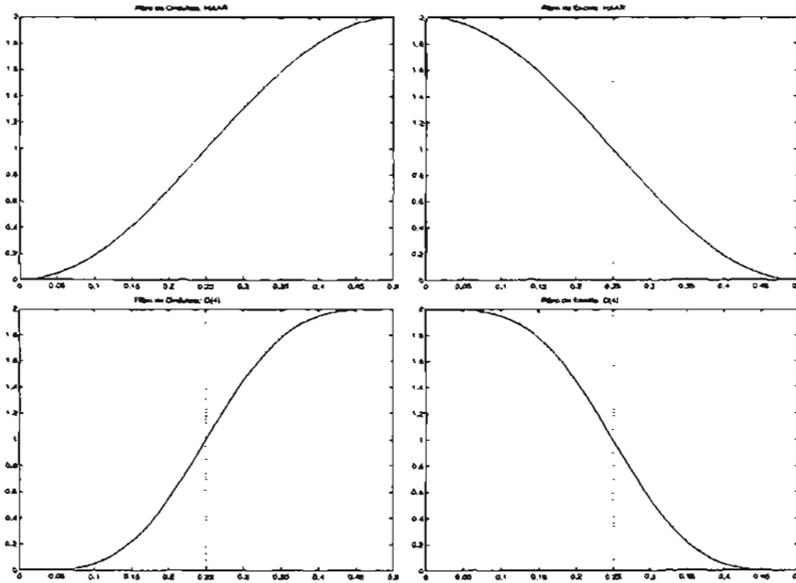


Figura 1.4: Funciones de ganancia cuadrada (módulo al cuadrado de la transformada de Fourier del filtro) para los filtros de onduletas(izquierda) y escala (derecha), para Haar (arriba) y Daubechies D(4) (abajo).

En la segunda etapa del algoritmo se toma como insumo al vector \mathbf{V}_1 resultante de la primera etapa, sólo que en esta ocasión se utiliza una versión modificada de los filtros originales (se inserta un cero entre cada elemento del filtro y se periodiza el filtro para tener una longitud conmesurable con \mathbf{V}_1 ²⁴). Como resultado se obtienen los coeficientes de onduletas \mathbf{W}_2 y de escala \mathbf{V}_2 de la segunda etapa, que permiten calcular a los detalles de segundo nivel \mathcal{D}_2 y el *smooth* de segundo nivel \mathcal{S}_2 respectivamente. El vector \mathbf{W}_2

²³Esta función puede verse como densidad del espectro de la energía de las series resultantes del filtrado

²⁴Más detalles en *ibid.* cap. 4.6

constituirá la cuarta parte del vector \mathbf{W} que viene inmediatamente después de \mathbf{W}_1 . Como se puede apreciar en la siguiente gráfica, correspondiente al módulo al cuadrado de la transformada de fourier de los filtros de la segunda etapa, el filtro de onduleta de la segunda etapa tiene ancho de banda $[1/8 \leq |f| \leq 1/4]$, mientras que el filtro de escala de la segunda etapa tiene ancho de banda $[0 \leq |f| \leq 1/8]$.

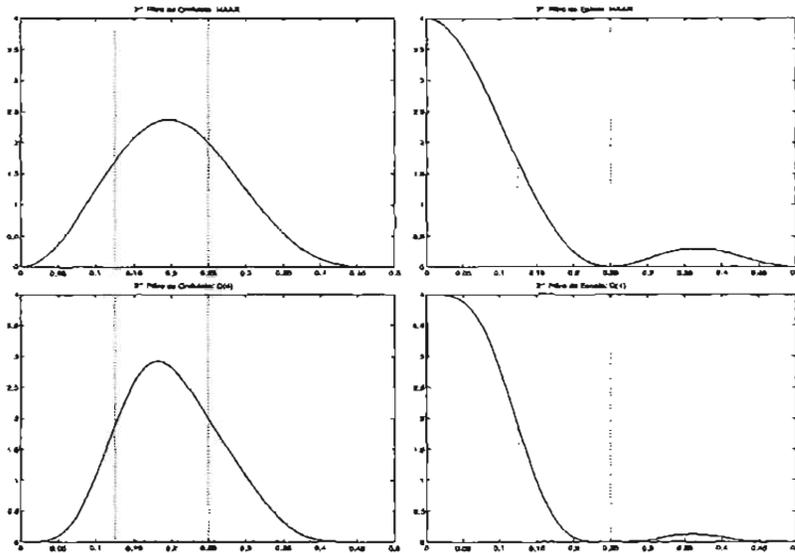


Figura 1.5: Funciones de ganancia cuadrada para los filtros de la segunda etapa: filtros de onduletas(izquierda) y escala (derecha), para Haar (arriba) y Daubechies D(4) (abajo).

En las etapas subsecuentes se toma como insumo el vector de coeficientes de escala resultante de la etapa anterior y se producen dos nuevos vectores de coeficientes de onduletas y escala (con otra versión modificada de los filtros de onduletas y escala respectivamente²⁵), donde el coeficiente de onduletas tiene la información de alta frecuencia (en el intervalo $[1/2^{j+1}, 1/2^j]$), mientras que el de escala contiene la información de baja frecuencia, (en el intervalo $[0, 1/2^{j+1}]$) es decir algo muy cercano a la tendencia de la serie. Esta descripción muy sucinta del algoritmo pirámide nos permite apreciar como el vector de *smooths* S_j resulta de refinamientos sucesivos de la señal que van retirando sus componentes de alta frecuencia, lo cual constituye la característica más relevante de la DWT para esta trabajo.

²⁵La operación de filtrado en cada etapa va agregando un cero más entre los coeficientes de los filtros originales $\{h_l\}$ y $\{g_l\}$. Más detalles en Percival y Walden (2000) p.93-99

Capítulo 2

Marco Teórico

2.1. La Ley del Precio Único y La Paridad de Poder Adquisitivo

Un punto de referencia subyacente a la validez de la teoría de la paridad de poder adquisitivo (PPA) es el concepto de la “Ley del precio único” que replica un principio importante de la teoría del comercio internacional y, basado en el sencillo arbitraje en el mercado de bienes, postula que dentro del contexto de los mercados competitivos y la ausencia de costos de transporte e impedimentos al comercio internacional, el libre comercio debe asegurar que el precio de un bien dado debe ser el mismo entre países, cuando está expresado en términos de una misma moneda. De lo anterior, la Ley del precio único, y a su vez la PPA puede ser resumida en la ecuación:

$$E = \frac{P}{P^*} \quad (2.1)$$

donde E es el tipo de cambio y P y P^* son los precios en unidades de moneda, denominados como doméstica y extranjera respectivamente¹. Cuando la ecuación (2.1) es referida a la Ley del precio único, P y P^* definen los precios de un solo bien en dos países, mientras que si se refiere a la PPA, entonces P y P^* definen los niveles de precios de los países.

Que la Ley del precio único prevalezca es consecuencia de la racionalidad de los agentes económicos, a saber, la Ley del precio único se cumple debido

¹La ecuación (2.1) no representa una relación de causalidad; en el sentido que los cambios en el tipo de cambio causan cambios en precios. De modo que la relación paridad de poder adquisitivo es sólo una condición, no es en sí misma una teoría completa sobre la determinación del nivel de precios o el tipo de cambio.

al arbitraje internacional; si el precio de un bien es mayor que el precio de ese mismo bien en otro país, expresado en una moneda común, se crea en los individuos un incentivo para comprar el bien en el país donde el precio es bajo y venderlo en el país donde el precio es mayor, por tanto, el arbitraje será el mecanismo que tiende a disminuir el precio en este último país como consecuencia del aumento en la oferta y a elevar el precio en el primer país debido al aumento de la demanda; el resultado de este mecanismo será la igualdad del precio en ambos países.

El principio postulado por la Ley del precio único, como ya se dijo, se aplica para todo bien individual y la extensión hacia la PPA puede ser hecha al considerar índices de precios construidos por la ponderación de los precios de un conjunto de bienes que conforman la cesta de referencia, ante la dificultad de conocer los niveles de precios. De manera similar a la Ley del precio único, la PPA sugiere que habrá niveles de precios idénticos entre países cuando son expresados en una misma moneda. Una reformulación de lo anterior es que si la Ley del precio único se cumple para todo bien individual, entonces debe cumplirse simultáneamente la PPA para cualquiera cesta de bienes idéntica entre países. Sin embargo, aunque existan desviaciones de la Ley del precio único para los bienes individuales, es decir, que ésta no se cumpla, estas desviaciones pueden anularse una vez construida la cesta de bienes y mantener la relación entre los niveles de precios y los tipos de cambio propuesta por la teoría de la PPA. Por otra parte, es conocido que, esta concepción de la relación PPA no es práctica dentro de la literatura empírica cuyo objetivo es su validación ; ya que i) los costos de transporte y los impedimentos al comercio existen y ii) las cestas de bienes no son normalizadas internacionalmente, es decir, los bienes que las conforman y sus ponderaciones suelen ser diferentes entre países. Por ello en la práctica, las desviaciones de la PPA que estos causarán, impedirán la comprobación de la relación PPA. Para aliviar esta falla se hizo una reclasificación del concepto de la PPA.

Lo que hasta aquí se ha descrito de la relación PPA proviene de la versión de Gustav Cassel (a quién se atribuye el concepto de PPA), la cual se denominó PPA absoluta, y según la formuló éste: “el tipo de cambio entre dos países será determinado por el cociente entre los niveles generales de precios en los dos países... en todo momento la paridad real entre dos países está representada por el cociente entre el poder adquisitivo del dinero en un país y el otro. Propongo llamar a esta paridad *“la paridad de poder adquisitivo”*. Mientras algo como el libre movimiento de mercancías y un comprensible comercio entre dos países tome lugar, el tipo de cambio presente no puede desviarse por mucho de esta paridad de poder adquisitivo

(en otras palabras,) las paridades de poder adquisitivo calculadas como una razón de precios de bienes de consumo para cualquier par de países tendería a aproximarse al tipo de cambio de equilibrio²”.

La segunda versión de la PPA retoma la relación entre los niveles de precios y tipos de cambio pero la establece en términos de cambios, y se define dentro de un intervalo de tiempo donde los movimientos en los precios relativos se compensan por los movimientos en los tipos de cambio, esto implica un ajuste necesario para mantener constante el poder adquisitivo de una moneda respecto a otras divisas; a esta definición se denominó PPA relativa, que formalmente se describe como:

$$\frac{E_{t_1}}{E_{t_2}} = \frac{P_{t_1}/P_{t_1}^*}{P_{t_2}/P_{t_2}^*} \quad \text{para todo } t_1 \neq t_2 \quad (2.2)$$

La PPA relativa puede ser descrita de mejor manera mediante una aproximación expresada en términos de variaciones porcentuales de los niveles de precios y tipos de cambio como³ :

$$\frac{E_t - E_{t-1}}{E_{t-1}} = \pi_t - \pi_t^* \quad (2.3)$$

donde π_t y π_t^* son las tasas de inflación doméstica y extranjera ⁴, respectivamente; verbi gratia, si la inflación doméstica es $x\%$, y sin que el nivel de precios extranjero varíe, entonces la relación PPA relativa sugiere que la moneda doméstica debe depreciarse ese $x\%$ respecto a la extranjera para dejar constante el poder adquisitivo, es decir, el tipo de cambio tendrá una variación de $x\%$.

Otro concepto que se debe incorporar al tratar la relación PPA, es el tipo de cambio real, el cual se define en términos de los tipos de cambio nominales y los niveles de precios; derivado de la ecuación (2.1) se escribe al tipo de cambio real, R , como:

$$R = E \frac{P^*}{P} \quad (2.4)$$

²Balassa, “The purchasing-power parity doctrine: A reappraisal,” Universidad de Yale, Diciembre 1964. P.p.: 584-585.

³P. R. Krugman y M Obstfeld (1995), “Economía Internacional Teoría y Política”, Mc Graw Hill. P. 487. Note que las ecuaciones (2.2) y (2.3) no son exactamente iguales (coinciden sólo cuando los precios del país externo son iguales en ambos periodos). Se utiliza esta aproximación sólo para explicar la PPA relativa cuando el interés se centra en los cambios porcentuales.

⁴ $\pi_t = (P_t - P_{t-1})/P_{t-1}$, análogo para el nivel de precios extranjero.

El tipo de cambio real expresa el precio en moneda doméstica del nivel de precios extranjero dividido por el nivel de precios doméstico, entonces, si los niveles de precios doméstico y extranjero están construidos en base a cierta cesta de bienes en cada uno de los países, el tipo de cambio real será el precio en moneda doméstica de la cesta de bienes extranjera respecto a la cesta de bienes doméstica. La implicación más importante de esta definición con relación a la PPA, es que si ésta se cumple en su versión absoluta, el tipo de cambio real será uno en cada momento; mientras que si la PPA relativa se cumple, el tipo de cambio real será una constante.⁵ Dado que esta afirmación tiene un sentido estricto, se puede decir entonces, que los cambios en el tipo de cambio real no deben ser permanentes o muy persistentes en el tiempo para que la PPA se cumpla.

Regresando a la cuestión sobre la validez de la PPA en la literatura empírica, una razón que fundamenta el uso de la versión de la PPA relativa en este contexto, es que ésta puede cumplirse mientras la PPA absoluta no (la PPA relativa es en cierto sentido una astucia para tener en cuenta desviaciones de la PPA absoluta, mientras los factores causantes de estas desviaciones sean más o menos estables). En un sentido lógico, la PPA relativa tiene la ventaja de que aunque los índices de precios usados para cada país no estén construidos de igual forma (o normalizados internacionalmente), si la composición permanece constante las variaciones porcentuales de los precios relativos (o diferencial de inflación), hechos con los índices de precios, hacen patentes los cambios en los niveles de precios y pueden ser comparados con las variaciones porcentuales de los tipos de cambio; esto es precisamente lo que describe la ecuación (2.3).

2.1.1. Desviaciones de la Relación PPA

Una revisión de la evidencia empírica muestra que no existe un consenso sobre la validez de la teoría de la PPA, por el contrario, existe evidencia que niega y otra que apoya que la relación PPA se cumple. La validez de la relación PPA degenera cuando no se cumplen los supuestos de la Ley del precio único, lo cual sucede en la realidad. Los costos de transporte y restricciones al comercio crean un intervalo donde el precio de los bienes en un país puede exceder el precio en el otro, expresados en la misma moneda y dado el tipo de cambio, y al mismo tiempo no se presenten oportunidades para el arbitraje; o en su lugar, donde el tipo de cambio puede fluctuar, dados los precios, antes que el arbitraje sea rentable. Así, mientras más grandes

⁵Para probar estas afirmaciones bastantan algunas operaciones algebraicas derivadas de las ecuaciones (2.1) y (2.2), respectivamente.

sean estos costos y restricciones el intervalo donde fluctuara el precio de un país o el tipo de cambio se hace más amplio. Como lo postula la Ley del precio único, el arbitraje sólo operará cuando el precio de un país supere al precio en otro (cuando está expresado en la misma moneda) más los costos de transporte, tarifas y/o otros impedimentos. Por tanto estos factores son un primer argumento por el cual se crean desviaciones de la PPA. Otras desviaciones sistemáticas de la relación PPA se dan cuando no se cumple el supuesto de los mercados competitivos e integrados⁶. Bajo el supuesto de estructuras no competitivas y segmentación de mercado, existirán perturbaciones en la Ley del precio único y en la relación PPA absoluta, y posiblemente tampoco la PPA relativa se cumpliría.

Lo que ya es sabido sobre la teoría de la PPA es que no se da como una relación de corto plazo, así como lo es que la economía en conjunto no se encuentra en equilibrio en todo momento; por tanto, en este contexto se argumenta que la existencia de factores monetarios y reales causan desviaciones de corto plazo de la relación PPA, pero se espera que en el largo plazo lo establecido por la PPA se mantenga. En adición, se ha observado respecto a los regímenes cambiarios que las desviaciones de la PPA en el corto plazo son mayores y más frecuentes en los tipos de cambio flexibles que en los tipos de cambio fijos⁷. En cuanto a los factores reales se argumenta que los precios se ajustan lentamente o son semifijos (en el corto plazo)⁸, aunque eventualmente alcanzan el de valor de equilibrio de largo plazo.

Froot y Rogoff (1994), Krugman y Obstfeld (1995) y Frankel, et. al. (1996), apoyados en la literatura empírica, señalan que las desviaciones del tipo de cambio real desaparecen lentamente y que 50 % de la desviación desaparece aproximadamente en cuatro años en los países con mayor industrialización. Además, los países con mayor crecimiento económico tienen tendencias hacia apreciaciones del tipo de cambio real; relacionado a esto, los países ricos respecto a los pobres experimentan niveles de precios más elevados y tienen, también, tendencias más altas del tipo de cambio real; estas desviaciones pueden estar asociadas con shocks, de oferta y demanda agregadas, permanentes.

En resumen, se puede entender que las desviaciones de la PPA, en

⁶Un mercado no competitivo es aquel en el cual alguno de los agentes (compradores o vendedores tienen la capacidad de alterar los precios del mercado por si solos. Por ejemplo un monopolio o un cártel.)

⁷Michael Mussa (1986) observó esto en un estudio realizado para muchos países y diferentes periodos al comparar ambos regímenes cambiarios.

⁸Esto puede ocurrir debido a la información imperfecta, contratos, inercia en hábitos de consumo, entre otros. A los precios que reaccionan de esta manera se le llama "sticky"

primera instancia, surgen cuando no se cumplen los supuestos subyacentes a la Ley del precio único; después, las desviaciones que se entienden como temporales y suceden debido a las diferentes velocidades de ajuste en los mercados, vinculadas a los precios “sticky” de los bienes; mientras que desviaciones permanentes son relacionadas a cambios en los hábitos o gustos de los consumidores, cambios en productividad, y más allá, cambios en el progreso tecnológico; por ello cuando alguno de éstos ocurre, habrá cambios permanentes en los precios relativos de equilibrio.

2.2. Evidencia Empírica: Resumen Histórico de las Pruebas de la PPA

Una vez discutidos los factores que pueden producir desviaciones sistemáticas de la PPA, examinamos la evidencia empírica que se ha aportado a través del tiempo sobre la validez de la PPA; considerando la metodología que se ha desarrollado para aumentar la potencia de las pruebas sobre el cumplimiento de la PPA y la incorporación de las causas que desvían al tipo de cambio real del equilibrio de largo plazo.

Esta sección es esencialmente una recapitulación del trabajo de Froot y Rogoff (1994)⁹; quienes describen los estudios empíricos realizados a partir de la década de los setenta. Los primeros estudios tratan de probar la relación tipos de cambio y niveles de precios con regresiones simples de mínimos cuadrados ordinarios (OLS; Ordinary Least Squares). La forma de la regresión para la prueba es

$$e_t = \alpha + \beta(p_t - p_t^*) + \epsilon_t$$

donde $e = \log(E)$, $p = \log(P)$ y $p^* = \log(P^*)$. Debido que estas pruebas sólo consideran un término de error, por lo tanto, son incapaces de distinguir entre efectos reales de corto y largo plazo, así como su dinámica de ajuste. Además, la estructura de la regresión para probar la relación PPA puede dar lugar a problemas de endogeneidad¹⁰, y hace al estimador de OLS (β) sesgado e inconsistente debido al término de error que se correlaciona con la variable endógena. Algunos estudios considerando este problema, tratan

⁹Froot y Rogoff (1994), “*Perspectives on PPP and Long-run Real Exchange Rates*” National Bureau of Economic Research.

¹⁰Es decir que los tipos de cambio y los precios son simultáneamente determinados, por tanto, es necesario otra construcción para el modelo y herramientas de estimación adicionales.

el uso de variables instrumentales ¹¹ Si bien se hace alusión a problemas de endogeneidad, también puede tratarse de un problema de variable(s) omitida(s), si éste es el caso, el parámetro estimado será sesgado e inconsistente debido a la correlación entre variable(s) omitida(s) y variables independientes incluidas.

Respecto a los resultados de estas pruebas, las mejores estimaciones por OLS fueron con datos de economías con alta inflación, pero en economías con baja inflación hubo fuertes rechazos de la PPA; estudios de este tipo fueron realizados por Frenkel (1978 y 1981). En general los pobres resultados obtenidos de estas pruebas no fueron concluyentes acerca del cumplimiento de la relación PPA, ni siquiera respecto al largo plazo.

El siguiente paso en la evolución de las técnicas para probar la teoría de la PPA fue considerar la no estacionariedad. El esquema de la prueba fue hacer al coeficiente de los precios relativos igual a uno y probar si el tipo de cambio real es estacionario. Es decir, β es restringido a tomar el valor de uno, y si el logaritmo del tipo de cambio real ($r_t = e_t + p_t^* - p_t$) es estacionario, la PPA se cumple.¹²

Para detectar si el tipo de cambio real sigue una caminata aleatoria, tres técnicas se asumen como las más usadas: la primera es la prueba Dickey-Fuller y Dickey-Fuller Aumentada¹³ (cuando se cree que los residuales están correlacionados); la segunda es la prueba "variance ratios" ¹⁴, la última está basada en integración fraccional¹⁵ que abarca bajo la hipótesis alterna-

¹¹El procedimiento de variables instrumentales o mínimos cuadrados en dos etapas, busca en primera instancia construir un modelo para las variables independientes del modelo original con variables que sean exógenas es decir que no estén correlacionadas con el término de error de la ecuación original.

¹²Froot y Rogoff (1994) señalan que los primeros autores en este contexto fueron Darby (1983), Adler y Lehman (1983), Hakkio (1984), Frankel (1986), Edison (1987), Huizinga (1987) y Meese y Rogoff (1988), todo ellos citados apropiadamente en su bibliografía.

¹³La prueba Dickey-Fuller Aumentada para el tipo de cambio real se basa en la regresión: $r_t = \alpha_0 + \alpha_1 t + \alpha_2 r_{t-1} + \phi(L)r_{t-1} + \epsilon_t$, donde $r = \log(R)$, L es el operador de rezagos, $\phi(\cdot)$ es un polinomio de orden p en L , con coeficientes ϕ_1, \dots, ϕ_p , ϵ_t es un ruido blanco. Bajo la hipótesis nula de una raíz unitaria $H_0 : \alpha_2 = 1$, mientras que en la alternativa $H_a : \alpha_1 = 0$ y $\alpha_2 < 1$

¹⁴Básicamente, si el tipo de cambio real sigue un proceso de caminata aleatoria, entonces la varianza del tipo de cambio real debe crecer linealmente sobre el tiempo.

¹⁵Un proceso integrado fraccional para probar raíz unitaria en el tipo de cambio real lleva a la ecuación:

$$\phi(L)(1-L)^d r_t = x(L)\epsilon_t,$$

donde L es el operador de rezagos; $\phi(\cdot)$ y $x(\cdot)$ son polinomios de orden p y q en L , respectivamente, con coeficientes ϕ_1, \dots, ϕ_p y x_1, \dots, x_q ; $r = \ln(R)$ y ϵ_t un ruido blanco. Luego, si el parámetro $d = 0$ entonces el tipo de cambio real es un proceso ARMA(p, q), si $d = 1$ entonces el tipo de cambio real sigue un proceso de caminata aleatoria y

tiva, además de los procesos ARMA, más procesos estacionarios haciendo mayores las posibilidades de rechazar la hipótesis nula de caminata aleatoria del tipo de cambio real.

No obstante este nuevo esquema de prueba descubre otros nuevos problemas, en particular la baja potencia de las pruebas, i.e., una baja probabilidad de rechazar la hipótesis nula de caminata aleatoria cuando de hecho es falsa. Este problema resulta de la dificultad de distinguir una convergencia lenta hacia el equilibrio de largo plazo en los tipos de cambio real. Por ejemplo, cuando estos están bajo un régimen que permite al tipo de cambio ser flexible se da un fenómeno de alta volatilidad en ellos, disfrazando un lento retorno a la media como una caminata aleatoria. En los estudios empíricos de esta clase se trata de eliminar este problema al considerar conjuntos de datos más amplios, llevando a cabo pruebas que incorporan un gran horizonte de datos históricos (que abarcan tipos de cambio flotantes y fijos) y/o el uso de más de una moneda simultáneamente.¹⁶

Finalmente al extremo de la línea de evolución sobre las técnicas en series de tiempo para probar la validez de la teoría de la PPA se hallan los estudios realizados con el uso de los modernos métodos de cointegración¹⁷, lo que implica verificar la existencia de una relación de equilibrio de largo plazo entre el conjunto de variables no estacionarias. Significa al mismo tiempo que las variables no pueden moverse independientemente unas de otras, es decir, sus tendencias estocásticas están relacionadas, dicho en otras palabras, la dinámica de las variables a través del tiempo, debe tener alguna relación con las desviaciones presentes de la relación de equilibrio de largo plazo. Los estudios empíricos, contemplados en el trabajo de Froot y Rogoff (1994), utilizan métodos de cointegración basados en la metodología de Engle y Granger (1987)¹⁸, los cuales proponen diferentes estructuras al probar cointegración entre los niveles de precios, P y P^* , y el tipo de cam-

si $0 < d < 1$ entonces el tipo de cambio real es un proceso integrado fraccionario y es estacionario.

¹⁶Sin embargo, ya que estos conjuntos de datos, utilizados en este tipo de estudios, combinan las varianzas relativamente bajas de los tipos de cambio fijos con las varianzas altamente volátiles de los tipos de cambio flotantes, la potencia de las pruebas puede ser afectada según sea el modelo especificado; así Lothian y Taylor (1994) en un estudio sobre esta cuestión, dan evidencia de que el uso de ambos tipos de cambio no sesgan la prueba de raíz unitaria cuando se estima un modelo AR(1) sobre el tipo de cambio real.

¹⁷Los métodos de cointegración se basan en la búsqueda dentro de un conjunto de series no estacionarias de alguna combinación lineal de ellas que sea estacionaria.

¹⁸Este método entra en la clasificación de métodos de estimación con una sola ecuación, a saber, sólo se estima un vector de cointegración. [Ver Maddala y Kim (1998), "Unit Roots, Cointegration, and Structural Change," Universidad de Cambridge, p. 155.]

bio, E . Algunos estudios prueban cointegración entre los niveles de precios expresados en la misma moneda, de tal modo, la ecuación (2.1) en forma logarítmica, se transforma de las siguientes maneras:

$$p_t = \alpha_0 + \alpha_1(e_t + p_t^*) + u_t, \quad (2.5)$$

$$e_t + p_t^* = \beta_0 + \beta_1 p_t + v_t, \quad (2.6)$$

donde u_t y v_t son llamados errores de equilibrio. En cada caso, el rechazo de la hipótesis nula de no cointegración implica que el error de equilibrio es estacionario (i.e. $I(0)$) y se puede concluir que la relación PPA se cumple. Aunque la teoría asintótica señala que cuando el horizonte temporal tiende a infinito, $t \rightarrow \infty$, probar caminata aleatoria en u_t y v_t deben ser equivalentes, esto no es aplicable al tamaño de los conjuntos de datos que han estado disponibles en los diversos estudios, por lo que la prueba es sensible a la elección de las variables que se utilizan como endógena e independiente. Por consecuencia, la metodología de Engle y Granger es ineficiente en este sentido y puede ser una fuente de ambigüedad acerca del resultado sobre la validez de la PPA.

En adición, otros estudios probaron cointegración entre los precios y el tipo de cambio, es decir, la ecuación (2.1) en forma logarítmica; como antes, las variables están cointegradas si el error de equilibrio, ϵ_t , es estacionario:

$$e_t = \theta_0 + \theta_1 p_t + \theta_2 p_t^* + \epsilon_t \quad (2.7)$$

Las pruebas de cointegración que usan la ecuación (2.7) en los trabajos de varios autores¹⁹ muestran como resultados comunes que los rechazos de la hipótesis nula de no cointegración se obtienen con mayor frecuencia cuando emplean datos del tipo de cambio fijo o índices de precios al productor, mientras que, con tipos de cambio flotantes o índices de precios al consumidor, los rechazos de la hipótesis nula son menos frecuentes²⁰.

Las pruebas basadas en la ecuación (2.7) son la forma general para probar cointegración entre esas variables, en particular, si se espera por alguna razón que $\theta_1, \theta_2 \neq 1$. Aunque también se presentan variaciones a la prueba al restringir los parámetros θ_1 y θ_2 , concretamente, al aplicar la restricción de simetría $\theta_1 = -\theta_2$; o bien la restricción de proporcionalidad $\theta_1 = -\theta_2 = 1$

¹⁹Tales como Corbae y Ouliaris (1988), Enders (1988), Kim (1990), Mark (1990), Fisher y Park (1991), Cheung y Lai (1993 a) y Kugler y Lenz (1993). Estos trabajos son citados por Froot y Rogoff (1994)

²⁰La ecuación (2.7) es consistente con tipos de cambio flexibles, donde el tipo de cambio nominal es quién responde para compensar las desviaciones de la PPA.

²¹. En ambos casos, la evidencia empírica para datos del tipo de cambio flexibles, resulta en menores rechazos de la hipótesis nula de no cointegración, que aquellas donde no hay restricciones sobre los parámetros.

Sin embargo, en el contexto de las pruebas de cointegración subyace la problemática del tamaño de la muestra. Muestras pequeñas llevan a resultados poco admisibles desde el punto de vista económico; por lo que es común encontrar en la literatura empírica coeficientes estimados de las pendientes que varían en un rango muy amplio y además muestran valores muy alejados de la unidad, lo cual no es consistente con la teoría postulada por la PPA.

²²

Finalmente diremos que la idea de la relación de equilibrio de largo plazo sigue postulándose como un buen marco para probar la teoría de la PPA; de modo que; aquellas metodologías que hacen una distinción más clara entre los distintos horizontes temporales parecen ser apropiadas para confrontar los problemas que plantean las pruebas ya mencionadas, buscando, al menos, aumentar la potencia de estas pruebas.

²¹Éste es el caso donde se prueba que el tipo de cambio real $r_t = e_t + p_t^* - p_t$ es estacionario.

²²Kim (1990) se señala como un ejemplo de pruebas que incluyen una muestra de datos grande. El es capaz de rechazar la hipótesis nula de no cointegración para los países considerados, además encuentra coeficientes que son cercanos a uno para casi todos los casos que examina.

Capítulo 3

Ejercicio Empírico: Metodología y Resultados

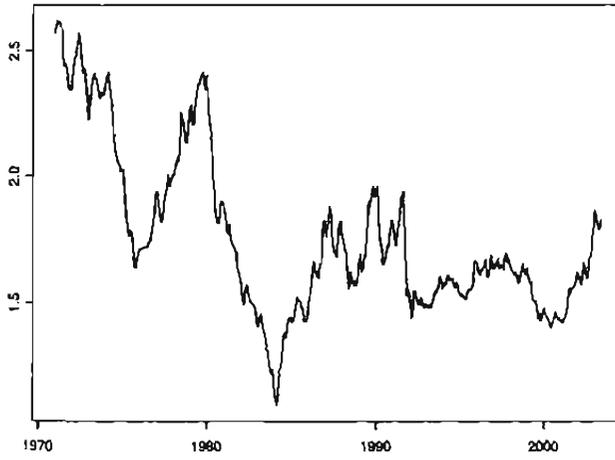
3.1. Datos

En este trabajo, cuyo objetivo es verificar la hipótesis PPA, se emplean datos mensuales acerca de los índices de precios al consumidor de Estados Unidos y el Reino Unido (en lo sucesivo abreviados como IPC.USA y IPC.UK respectivamente), así como el tipo del cambio (TC) de libras esterlinas por dólar estadounidense. Los datos abarcan un periodo de 32 años comprendidos entre enero de 1972 y junio de 2004, constituyendo una serie de 390 observaciones para cada variable. Las series de datos para los índices de precios han sido expresadas, ambas, tomando como base el mes de enero de 1987.¹ Los datos fueron obtenidos, en la base de datos del Banco de la Reserva Federal de Saint Louis a través de su página electrónica research.stlouisfed.org/fred2/, en el caso del índice de precios norteamericano y el tipo de cambio; mientras que la serie de índice de precios del Reino Unido corresponde a la base de datos del ministerio estadísticas de este país a través de su página electrónica www.statistics.gov.uk. En la página siguiente se muestran gráficos de las series, en primer lugar de TC y posteriormente IPC.USA-IPC.UK

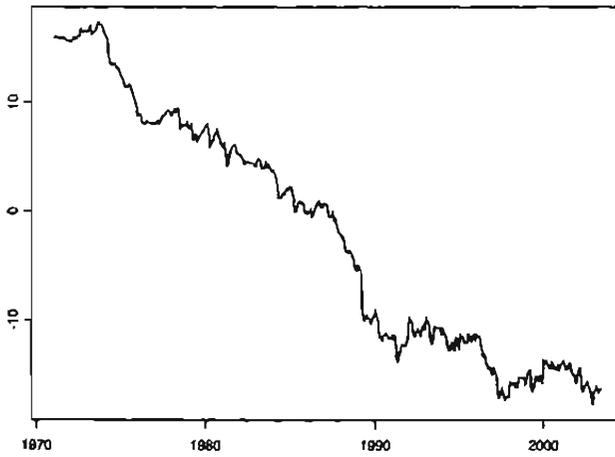
Los gráficos nos permiten apreciar fuertes comportamientos tendenciales

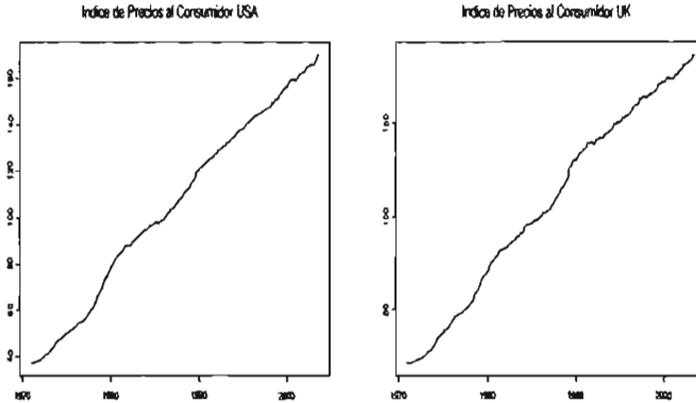
¹Es una práctica habitual en economía reexpresar las variables nominales (i.e. que están expresadas con los precios dados por el mercado), tales como el índice de precios, teniendo como punto de referencia un año en particular, en el cual la variable tomará como valor 1 o 100; esto permite hacer comparaciones más confiables entre las variables pues con ello se eliminan los efectos perturbadores de los precios de mercado.

Tipo de Cambio Dolar US vs. Libra UK



Diferencia entre IPC USA-UK





en las series; en el caso de la diferencia del tipo de cambio observamos que la serie ha caído consistentemente desde el inicio de la liberalización en los tipos de cambio (a principios de los setentas) hasta finales de la década de los ochenta, lo que nos permite afirmar que la serie IPC.USA se ha desplazado más lentamente que su contraparte británica. Por su parte la serie del TC muestra descensos muy pronunciados en las décadas de los setentes y ochentas, sin embargo a partir de la década de los noventas se observa un comportamiento muy estable de la variable.

3.2. Metodología

La metodología a seguir en este ejercicio empírico será una modificación de aquella establecida originalmente por Ramsey y Lampart (1997) [14] y que consiste en lo siguiente, primero la series son transformadas mediante la DWT, obteniendo de esta manera distintos niveles de detalles (\mathcal{D}_j), que de acuerdo a lo establecido en el capítulo 1 representan los cambios de la serie en la correspondiente escala j , esto es, el componente de frecuencia de la serie original en el intervalo $[1/2^{j+1}, 1/2^j]$. Como el objetivo de Ramsey y Lampart es mostrar la naturaleza de la relación entre las diferentes escalas de tiempo, tanto entre el dinero y el ingreso como del ingreso y el consumo, se utilizan los coeficientes de los detalles \mathcal{D}_j como *proxies* en la referida escala de tiempo para la variable correspondiente; a continuación efectúan pruebas de causalidad de Granger entre los detalles, y de esta manera se establece que la relación de causalidad entre las variables originales en cierto horizonte de tiempo esta dada por la relación entre los detalles \mathcal{D}_j de las

mismas. Ahora como el objetivo de este trabajo es verificar la existencia de una relación de largo plazo, entre el tipo de cambio y la diferencia entre los niveles de precios, modificaré la metodología anterior de la siguiente manera. Se transformarán las series mediante la DWT, sin embargo ahora el interés se centrará en el componente de largo plazo de las series, a saber, el j -ésimo nivel de *smooth* S_j , quien como se discutió en el capítulo 1 está asociado a la banda de frecuencia $[0, 1/2^{j+1}]$, es decir en él está expresado el componente de largo plazo de la serie. Posteriormente trataré de esclarecer la relación que existe entre las variables mediante, en primera instancia, regresiones (con diferentes especificaciones) entre los distintos *smooths* y posteriormente mediante otra medida de "asociación" buscando en ambos casos corroborar o bien refutar la hipótesis la PPA. A continuación explicaré con más detalle este procedimiento.

Recordemos que la PPA (paridad del poder adquisitivo) establece que existe una relación de largo plazo entre el tipo de cambio y la diferencia de los niveles de precios. Así mismo en el capítulo 2 hemos mencionado que casi todas las pruebas empíricas sobre la PPA se enfocan en el ajuste de la ecuación dada por la ley de único precio, i.e.

$$E = \frac{P}{P^*} \quad (3.1)$$

donde E es el tipo cambio entre ambas divisas (TC), P y P^* son el índice de precios local (IPC.USA) y externo (IPC.UK) respectivamente. Con el objetivo de linealizar el modelo propuesto en la ecuación se expresa a las variables en logaritmos de modo que

$$\log E = \log P - \log P^* \quad (3.2)$$

de esta manera se ajusta el siguiente modelo de regresión lineal

$$e_t = \alpha + \beta(p_t - p_t^*) + \epsilon_t \quad (3.3)$$

donde e_t es el logaritmo de TC, p_t y p_t^* es el logaritmo de los niveles de precios del país interno (IPC.USA) y externo (IPC.UK) respectivamente. Ahora bajo la hipótesis PPA se cumple que $\beta = 1$. O bien el modelo más general

$$e_t = \alpha + \beta_0 p_t + \beta_1 p_t^* + \epsilon_t \quad (3.4)$$

bajo la hipótesis $\beta_1 = -\beta_0 = -1$. Sin embargo, en el capítulo 2 establecimos que todas la pruebas econométricas de este tipo han encontrado dificultades para corroborar la hipótesis. Por tanto en este trabajo se hace uso del análisis en onduletas como alternativa en el estudio de este tema. Este tipo de

análisis es de interés en este caso pues la transformada en onduletas de una señal proporciona información en diferentes escalas de la señal (recordemos que en el primer capítulo ilustramos a través de la onduleta de Haar como los coeficientes resultantes de la DWT para cada escala son proporcionales a diferencias en promedios adyacentes a la escala correspondiente ²). En particular la hipótesis PPA requiere el estudio de los componentes de largo plazo de las variables, por tanto, y a diferencia de los estudios realizados por Ramsey y Lampart, el interés de este trabajo no se centrará en los detalles \mathcal{D}_j que resultan de la DWT sino en aquello que expresa los componentes de largo plazo de la señal, a saber, los *smooths* S_j (pues como se constata a través de la ecuación de análisis (1.18) el *smooth* es una aproximación a la señal a la que se le ha removido los cambios hasta la escala j , i.e. \mathcal{D}_i para $i = 1 : j$ y por tanto S_j es más cercano a la tendencia de la serie). De esta manera en el trabajo se propone establecer la relación de largo plazo que exista entre las variables a través de sus *smooths*. Emulando los trabajos realizados con anterioridad sobre el tema se estiman regresiones como las planteadas anteriormente, pero utilizando en lugar de las series originales se hará uso de los *smooths* resultantes del análisis (DWT) de las variables originales.³ Sin embargo ante la posibilidad de que dicha relación sea de carácter no lineal, en cuyo caso el análisis de regresión tendría numerosos defectos, entre otros heteroscedasticidad, correlación serial etc., se propone entonces el uso de otro estadístico que nos permita establecer la dependencia entre las variables, en la siguiente subsección se describirá a detalle dicho estadístico.

Finalmente se recurrirá a otra de las metodologías propuestas concerniente a las pruebas de la PPA, ⁴ en ella se prueba la estacionariedad de la serie del tipo de cambio real, que está definido (en logaritmos) como $r_t = e_t - p_t + p_t^*$. Este ejercicio, de resultar exitoso, probaría la versión relativa de la PPA. Como la DWT es una transformación lineal para calcular la DWT del tipo de cambio real basta con sumar (o restar según sea el caso) las series transformadas del tipo de cambio y los índices de precios doméstico y externo respectivamente. Una vez obtenida la DWT del tipo de cambio real se probará la hipótesis de estacionariedad para sus *smooths* mediante la prueba

²Más detalles en la p. 17

³Para ser más precisos, si denotamos como $S_j(\mathbf{X})$ al *smooth* al j -ésimo nivel de una señal \mathbf{X} ; entonces el modelo de regresión que se estimará es el siguiente:

$$S_j(e_t) = \alpha + \beta_0 S_j(p_t) + \beta_1 S_j(p_t^*) + \epsilon_t$$

con hipótesis análogas.

⁴Froot y Rogoff (1994) hacen una recapitulación de los autores que han realizado este tipo de pruebas. También se hace referencia a estos trabajos en el capítulo 2.

de Dickey-Fuller aumentada.

3.2.1. Análisis de Regresión y Dependencia

Consideremos un vector aleatorio (respuesta) Y , cuyos valores pueden depender de un vector aleatorio $\underline{X} = (X_1, X_2, \dots, X_k)$ donde k es un entero positivo. Asumamos que es cierto un modelo de la forma

$$Y = f(X_{i_1}, X_{i_2}, \dots, X_{i_m}),$$

donde $1 \leq i_1 < i_2 < \dots < i_m \leq k$, y f es cualquier función medible, no necesariamente lineal. En este caso existe una relación de dependencia entre la variable aleatoria Y y el vector aleatorio $(X_{i_1}, X_{i_2}, \dots, X_{i_m})$.

Definamos

$$\delta_{Y, (X_{i_1}, X_{i_2}, \dots, X_{i_m})} = \sup_{(y, x_{i_1}, x_{i_2}, \dots, x_{i_m}) \in \mathbb{R}^{m+1}} |F_{y, i_1, i_2, \dots, i_m}(y, x_{i_1}, x_{i_2}, \dots, x_{i_m}) - F_y(y)F_{i_1, i_2, \dots, i_m}(x_{i_1}, x_{i_2}, \dots, x_{i_m})|$$

donde $F_{y, i_1, i_2, \dots, i_m}$ denota la función de distribución conjunta de Y y todos los X_{i_i} , F_y es la función de distribución de Y , y F_{i_1, i_2, \dots, i_m} es la función de distribución conjunta de X_{i_i} , entonces $\delta_{Y, (X_{i_1}, X_{i_2}, \dots, X_{i_m})} = 0$ si y sólo si Y y el vector $(X_{i_1}, X_{i_2}, \dots, X_{i_m})$ son independientes. De otra manera $\delta_{Y, (X_{i_1}, X_{i_2}, \dots, X_{i_m})} > 0$. Más adelante veremos que dicho estadístico nos permitirá detectar la dependencia establecida entre Y y \underline{X} aún cuando la función f sea no lineal. En tanto establezcamos algunas propiedades del estadístico,⁵

- a) $\delta_{Y, (X_{i_1}, \dots, X_{i_m})} = 0$ si y sólo si la variable aleatoria Y y el vector aleatorio $(X_{i_1}, \dots, X_{i_m})$ son independientes.
- b) $0 \leq \delta_{Y, (X_{i_1}, \dots, X_{i_m})} \leq 1/4$
- c) $0 \leq \delta_{Y, (X_{i_{\sigma(1)}})} \leq \delta_{Y, (X_{i_{\sigma(1)}}, X_{i_{\sigma(2)}})} \leq \dots \leq \delta_{Y, (X_{i_{\sigma(1)}}, \dots, X_{i_{\sigma(m-1)}})} \leq \delta_{Y, (X_{i_{\sigma(1)}}, \dots, X_{i_{\sigma(m)}})}$ y $\delta_{Y, (X_{i_1}, \dots, X_{i_m})} = \delta_{Y, (X_{i_{\sigma(1)}}, \dots, X_{i_{\sigma(m)}})}$ para cualquier permutación σ de I_m .
- a) $\delta_{Y, (X_{i_1}, \dots, X_{i_m})} = \delta_{Y, (X_{i_1}, \dots, X_{i_m}, X_{i_{m+1}})}$ si $X_{i_{m+1}}$ es independiente de $(X_{i_1}, \dots, X_{i_m})$ e Y .

⁵Pruebas detalladas por Fernandez-Fernandez y González-Barrios (2001,2003) y González-Barrios y Ruiz-Velasco (2004)

Asumamos que tenemos una muestra aleatoria de dimensión $m + 1$ de tamaño j , $\mathbf{X}_l = (Y_l, X_{i_{1l}}, X_{i_{2l}}, \dots, X_{i_{ml}})$ para $l = 1, 2, \dots, j$ proveniente de una función de densidad conjunta $F_{y, i_1, i_2, \dots, i_m} = (y, x_{i_1}, x_{i_2}, \dots, x_{i_m})$ con función densidad marginal $F_y(y)$ y función de densidad marginal conjunta $F_{i_1, i_2, \dots, i_m} = (x_{i_1}, x_{i_2}, \dots, x_{i_m})$.

Denotamos por $F_{y, i_1, i_2, \dots, i_m}^j = (y, x_{i_1}, x_{i_2}, \dots, x_{i_m})$ la función de distribución empírica de \mathbf{X}_l , $l = 1, 2, \dots, j$, sea $F_y^j(y)$ la función de distribución empírica de Y , y $F_{i_1, i_2, \dots, i_m}^j = (x_{i_1}, x_{i_2}, \dots, x_{i_m})$ la función de distribución conjunta empírica de $(X_{i_{1l}}, X_{i_{2l}}, \dots, X_{i_{ml}})$ para $l = 1, 2, \dots, j$. Se usará como medida de dependencia multidimensional muestral a:

$$\delta_{Y, (X_{i_1}, \dots, X_{i_m})}^j := \sup_{(y, x_{i_1}, \dots, x_{i_m}) \in \mathbb{R}^{m+1}} \left| F_{y, i_1, \dots, i_m}^j (y, x_{i_1}, \dots, x_{i_m}) - F_y^j(y) F_{i_1, \dots, i_m}^j (x_{i_1}, \dots, x_{i_m}) \right|$$

Esta versión muestral imita a la versión poblacional de la medida de población definida anteriormente, lo cual tiene más sentido cuando $j \rightarrow \infty$, de tal suerte que el estadístico muestral satisface una ley de grandes números bajo la hipótesis de independencia, es decir si $(Y_l, X_{i_{1l}}, X_{i_{2l}}, \dots, X_{i_{ml}})$ es una muestra independiente de tamaño j , proveniente de función una distribución conjunta de dimensión $m + 1$, $F_{y, i_1, i_2, \dots, i_m} (y, x_{i_1}, x_{i_2}, \dots, x_{i_m})$ cuyas funciones de distribución marginales correspondientes están dadas por $F_y(y)$ y función de distribución marginal conjunta $F_{i_1, i_2, \dots, i_m} (x_{i_1}, x_{i_2}, \dots, x_{i_m})$, entonces $\delta_{Y, (X_{i_1}, \dots, X_{i_m})}^j$ se aproxima a $\delta_{Y, (X_{i_1}, X_{i_2}, \dots, X_{i_m})}$ casi seguramente cuando $j \rightarrow \infty$.

Además si Y es independiente del vector $(X_{i_1}, X_{i_2}, \dots, X_{i_m})$ tenemos que:

$$\delta_{Y, (X_{i_1}, \dots, X_{i_m})}^j \rightarrow 0$$

casi seguramente cuando $j \rightarrow \infty$.

El estadístico tiene entre otras propiedades, que su valor no cambia si se transforman las coordenadas de X (en los datos) mediante transformaciones monotonas crecientes o si transformamos Y a través de una función monótona creciente. También es posible establecer un orden creciente de las estadísticas añadiendo coordenadas.

Más importantes para los propósitos de este trabajo son los siguientes resultados sobre la distribución de $\delta_{Y, (X_{i_1}, \dots, X_{i_m})}^j$, en ellos se asume como hipótesis que X es una variable aleatoria univariada y no un vector como hasta ahora habíamos supuesto.

- a) Si X e Y son variables aleatorias continuas entonces $\delta_{Y,(X)}^j \neq 0$ casi seguramente.
- b) Si se asume que $F_{y,x}(y, x) = xy$, es decir X e Y son variables aleatorias independiente y uniformes en el intervalo $(0, 1)$, entonces siempre es posible hallar la distribución de $\delta_{Y,(X)}^j$.
- c) Si se asume que $F_{y,x}(y, x) = F_y(y)F_x(x)$, es decir que X e Y son variables aleatorias independientes cuyas funciones de distribución son $F_x(x)$ y $F_y(y)$ correspondientemente, entonces

$$\delta_{Y,(X)}^j \stackrel{\text{dist}}{=} \delta_{F_Y(Y),(F_X(X))}^j$$

Los resultados anteriormente establecidos permiten utilizar al estadístico $\delta_{Y,(X_{i_1}, \dots, X_{i_m})}^j$ en una prueba de independencia incluso si es necesario aproximar su distribución mediante simulaciones.

Dada una muestra aleatoria, de tamaño j , de vectores con dimensión $m+1$, $\mathbf{X}_l = (Y_l, X_{i_1}, X_{i_2}, \dots, X_{i_m})$ para $l = 1 \dots, j$, donde cada \mathbf{X}_l proviene de una función de distribución común $F_{y,i_1, \dots, i_m}(\cdot)$; González-Barrios y Ruiz-Velasco proponen los siguientes pasos para realizar la prueba:

1. Hallar la distribución exacta o una aproximación mediante simulaciones de $\delta_{Y,(X_{i_1}, \dots, X_{i_m})}^j$ en el caso de $m+1$ variables aleatorias independientes uniformes y para un tamaño de muestra j .
2. Evaluar al estadístico $\delta_{Y,(X_{i_1}, \dots, X_{i_m})}^j$ para la muestra dada.
3. Rechazar la hipótesis si el estadístico es mayor que el cuantil $1 - \alpha$ de la distribución exacta o aproximada. En el caso contrario no se rechaza la hipótesis de independencia.

La implementación práctica del ejercicio antes propuesto se realizará mediante el software (escrito en lenguaje FORTRAN) diseñado por Contreras-Cristán, González-Barrios y Ruiz-Velasco, quienes han permitido su uso para este trabajo. En la siguiente sección se presentan los resultados.

3.3. Estimaciones

En esta sección se presentan los resultados obtenidos mediante la metodología anteriormente propuesta. En primera instancia, se muestra un gráfico de las series entre las cuales se intenta “descubrir” una relación de dependencia.

Se muestra en el eje de las ordenadas el valor del TC, mientras que en el eje de las abscisas se encuentra la diferencia IPC.USA-IPC.UK (cada punto en la gráfica corresponde a un año de las observaciones).

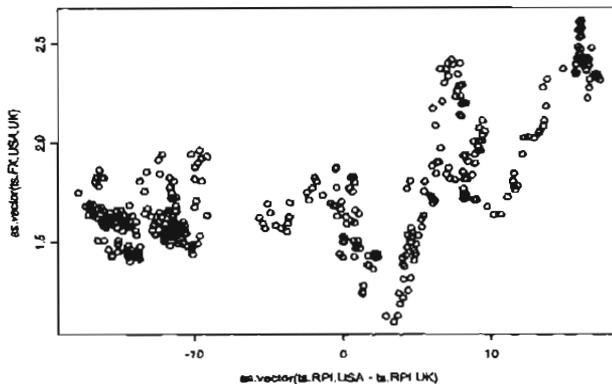


Figura 3.1:

También se muestra la estimación de la regresión de la ecuación (3.4) antes de calcular la DWT.

Dependent Variable: DATOS

Method: Least Squares

Sample: 1 390

Included observations: 390

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.544275	0.174158	3.125176	0.0019
USADATOS	0.676858	0.115520	5.859229	0.0000
-UKDATOS	0.683193	0.079519	8.591532	0.0000

R-squared	0.549362	Mean dependent var	0.551200
Adjusted R-squared	0.547033	S.D. dependent var	0.174721
S.E. of regression	0.117592	Akaike info criterion	-1.435524
Sum squared resid	5.361415	Schwarz criterion	-1.405015
Log likelihood	282.9272	F-statistic	235.8914
Durbin-Watson stat	0.046411	Prob(F-statistic)	0.000000

Como ya se había señalado anteriormente esta regresión pese a estimar coeficientes significativos, estos no son significativamente iguales a 1, de acuerdo a la prueba de Wald que se muestra después de este párrafo. Además el estadístico Durbin-Watson evidencia un problema de correlación serial de los errores.

Wald Test:

Equation: EQDATOS

Test	Statistic Value	df	Probability
F-statistic	182.7442	(2,387)	0.0000
Chi-square	385.4885	2	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-1 + C(2)	-0.323142	0.115520
-1 + C(3)	-0.316807	0.079519

Restrictions are linear in coefficients.

Ahora se procede propiamente con la metodología propuesta, primero se obtiene la DWT de cada una de las series, con el número de niveles que $S+$ usa por default ($j = 6$) utilizando como onduleta básica a las *Symmlets* o LA(8),⁶ lo que da lugar al siguiente análisis multiresolución.

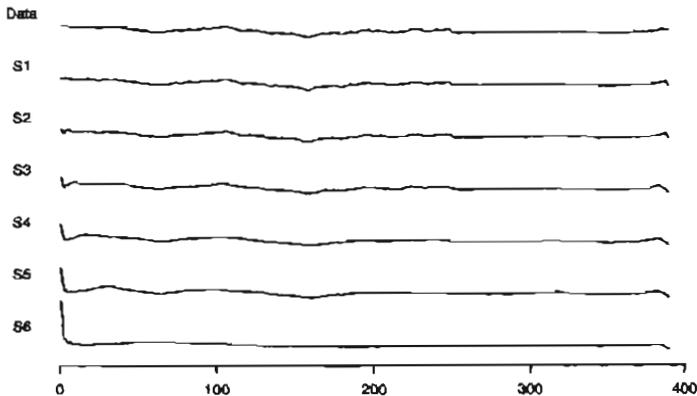


Figura 3.2: *Smooths* para la serie del TC desde el nivel 1 hasta 6.

⁶Se utiliza esta onduleta por las mismas razones que argumentan Ramsey y Lampart, a saber, la función es “esencialmente” distinta de cero en un conjunto compacto y relativamente “angosto”, también es casi simétrica y diferenciable en la mayor parte de su dominio. Además Lampart y Ramsey establecen que el ejercicio es bastante robusto al cambio en la onduleta básica, por lo tanto el presente trabajo se concreta a hacer el análisis con una sola función.

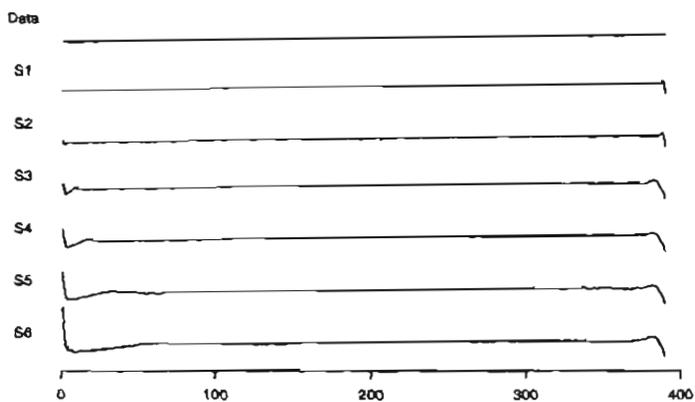


Figura 3.3: *Smooths* para la serie del IPC.USA desde el nivel 1 hasta 6.

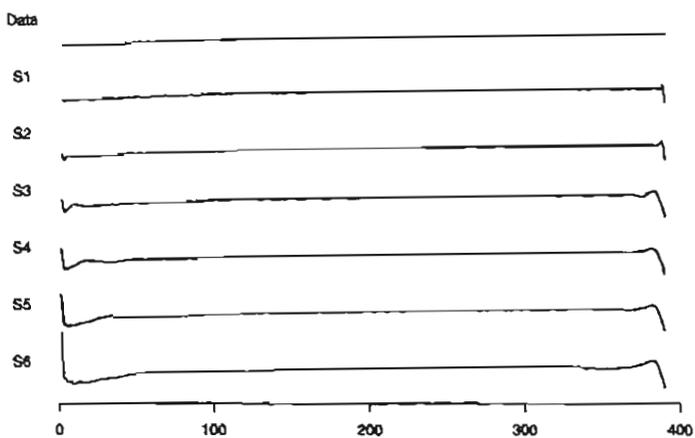


Figura 3.4: *Smooths* para la serie del IPC.UK desde el nivel 1 hasta el 6.

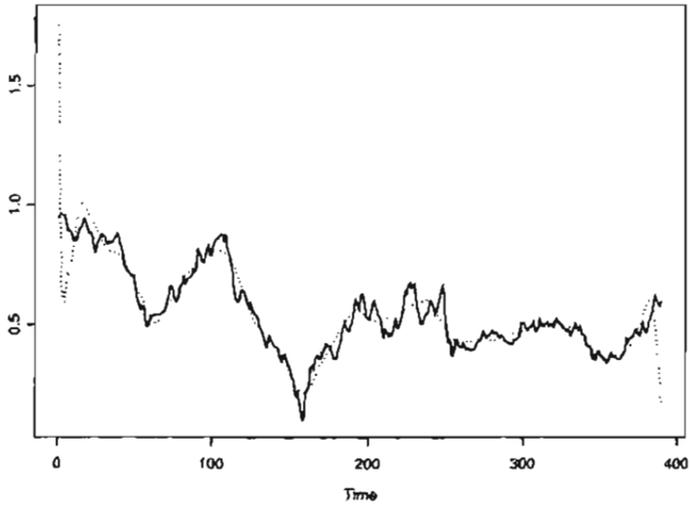


Figura 3.5: Serie del TC (línea sólida) comparada con su *smooths* de cuarto nivel S_4

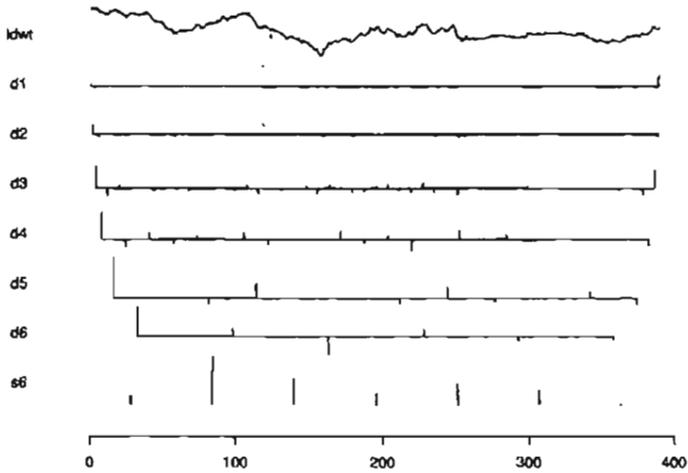


Figura 3.6: DWT del logaritmo del TC

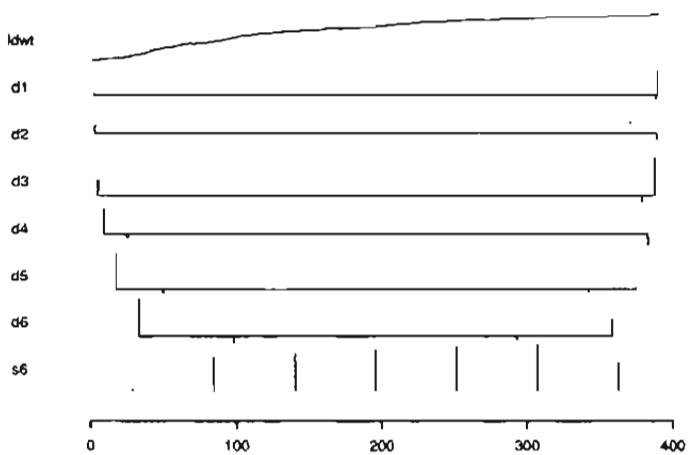


Figura 3.7: DWT del logaritmo de la serie del IPC.UK

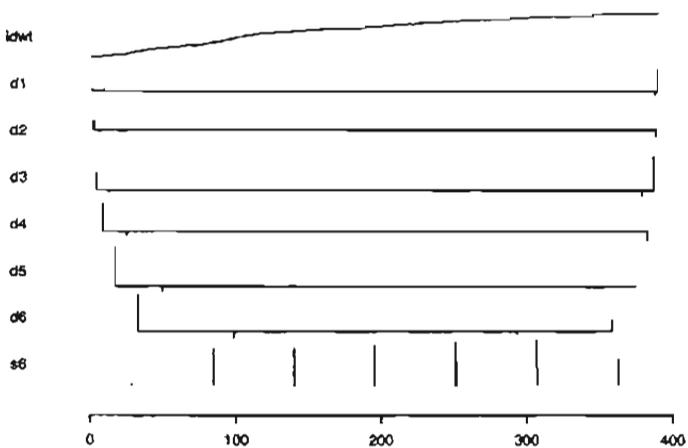
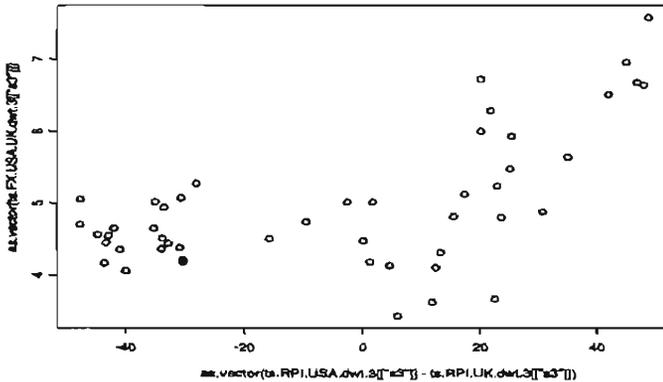


Figura 3.8: DWT del logaritmo de la serie del IPC.USA

Entre las gráficas anteriores se incluyó una figura comparativa de la serie del TC con su *smooth* de cuarto nivel, en la que es posible apreciar el modo en el que la DWT ofrece una “versión” más suave de la señal a través de sus *smooths*. También se incluyen gráficos de los coeficientes de la DWT para cada una de las variables. En la figura correspondiente a la serie del TC se evidencian perturbaciones en las escalas menores (d_1, d_2, d_3), mientras que en las series IPC.USA y IPC.UK solo muestran coeficientes grandes para s_6 .

Ahora para corroborar que exista alguna posible relación de largo plazo realizo graficas de la diferencia de los *smooths* de los índices de precios versus los del TC (siempre en el mismo nivel). Se muestra en seguida, como ejemplo, la gráfica entre los *smooths* de tercer nivel S_3 .



como es posible apreciar en el gráfico anterior existen buenas posibilidades de encontrar alguna especie de “asociación” entre las variables. Enseguida se estima una regresión lineal entre los “*smooth*” s_3 de las series, de lo que resulta lo siguiente,

Dependent Variable: S3

Method: Least Squares

Sample: 1 390

Included observations: 390

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.105160	0.053656	1.959904	0.0507
USAS3	0.987973	0.037434	26.39211	0.0000
-UKS3	0.900866	0.027659	32.57019	0.0000

R-squared	0.583604	Mean dependent var	0.548773
Adjusted R-squared	0.581452	S.D. dependent var	0.177018
S.E. of regression	0.114522	Akaike info criterion	-1.488433
Sum squared resid	5.075637	Schwarz criterion	-1.457924
Log likelihood	293.2445	F-statistic	271.2014
Durbin-Watson stat	0.013434	Prob(F-statistic)	0.000000

en la regresión anterior los coeficientes estimados para ambos "smooths" resultan significativos y los signos parecen correctos de acuerdo a la PPA, i.e. el nivel de precios doméstico (IPC.USA) tiene asociado un coeficiente positivo y el índice de precios externos, (IPC.UK) tiene un coeficiente de signo negativo y son bastante cercanos a 1. No obstante las pruebas de Wald rechazan la hipótesis de que el coeficiente asociado al nivel de precios externo (Uk) es igual a -1, mientras que no se rechazó la prueba en el caso del nivel de precios para USA. La prueba F rechaza la hipótesis nula de que todos los coeficientes sean iguales a cero.

Wald Test: Equation: EQS3

Test Statistic	Value	df	Probability
F-statistic	0.103218	(1,387)	0.7482
Chi-square	0.103218	1	0.7480

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-1 + C(2)	-0.012027	0.037434

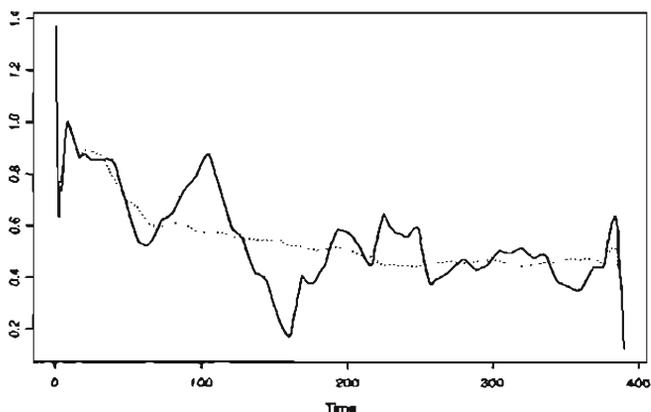
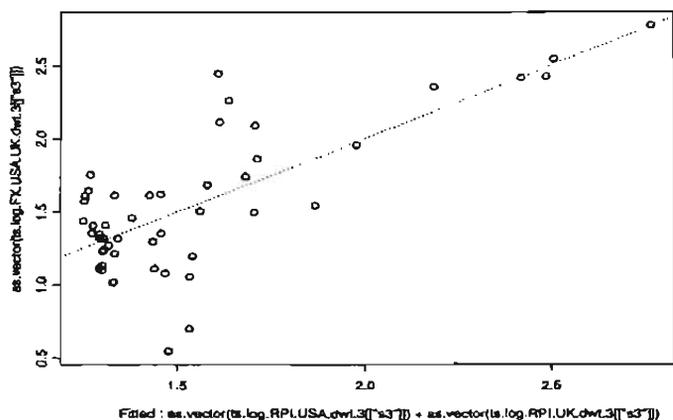
Test Statistic	Value	df	Probability
F-statistic	12.84589	(1,387)	0.0004
Chi-square	12.84589	1	0.0003

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-1 + C(3)	-0.099134	0.027659

La regresión ajustada da lugar la siguiente gráfica. Se muestra en el eje de las ordenadas el valor de logaritmo del *smooth* de TC, mientras que en el de las abscisas se muestra el valor ajustado por el modelo; las línea punteada representa a la recta identidad, de esta manera mientras más cercanos a la línea se encuentren los puntos tendremos una estimación más certera, en este caso particular las desviaciones graves del modelo estimado se dan entre 1.4 y 1.6.

También se muestra la gráfica através del tiempo de la serie del *smooth* del tercer nivel (línea sólida) así como el valor ajustado mediante la regresión (línea punteada).



Enseguida se estima la regresión bajo la especificación dada por el modelo restringido, es decir, suponiendo que $\beta_1 = -\beta_0$ en la ecuación 3.4, de lo que resulta lo siguiente.

Dependent Variable: S3

Method: Least Squares

Sample: 1 390

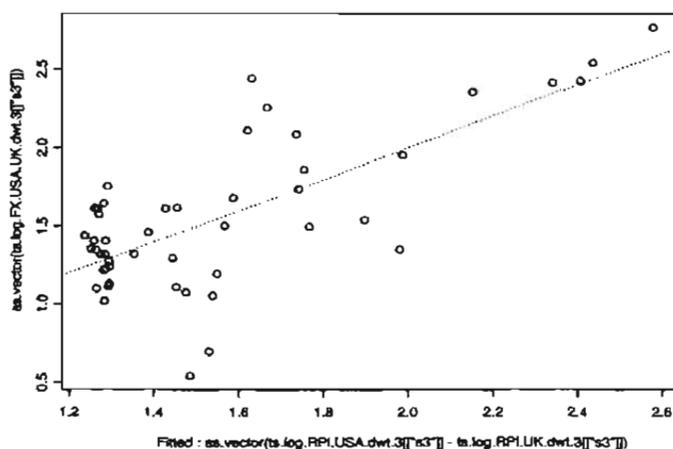
Included observations: 390

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.511328	0.006118	83.57134	0.0000
USAS3-UKS3	0.711482	0.023233	30.62437	0.0000

R-squared	0.558082	Mean dependent var	0.548773
Adjusted R-squared	0.556943	S.D. dependent var	0.177018
S.E. of regression	0.117828	Akaike info criterion	-1.434075
Sum squared resid	5.386731	Schwarz criterion	-1.413735
Log likelihood	281.6445	F-statistic	489.9909
Durbin-Watson stat	0.045343	Prob(F-statistic)	0.000000

En esta regresión observamos que el coeficiente estimado β es igual a 0,711482, que aunque no es cercano a 1 es positivo y significativo a cualquier nivel de significancia. Así mismo la regresión en su conjunto es significativa de acuerdo a la prueba F. La regresión ajustada da lugar al siguiente gráfico, en donde se muestra en el eje de las ordenadas el valor de la serie del *smooth* del TC, mientras que en el eje de las abscisas se muestra el valor estimado por el modelo de regresión, de nueva cuenta la línea punteada es la recta identidad y cualquier punto sobre ella indica que no existe error en la estimación.



A continuación muestro los coeficientes que se obtienen en las regresiones con los *smooths* desde el nivel S1 hasta S6. En estas estimaciones es posible apreciar que mientras en los primeros niveles los coeficientes estimados parecen estar lejos de los postulados por la hipótesis PPA, en cuanto se estima la regresión en los *smooths* de los niveles mayores los coeficientes se aproximan a 1. Es necesario mencionar que todas las regresiones estiman coeficientes significativos y las pruebas F rechazan que las regresiones no sea significativas. No obstante las pruebas de Wald solo avalan la hipótesis PPA, es decir que los coeficientes estimados sean significativamente iguales a 1 y

-1 respectivamente, en el tercer y cuarto nivel (marcados con asterisco).

	β_0	β_1
S1	0.8447	-0.8001
S2	0.8777	-0.8232
S3	0.9880*	-0.9009
S4	1.0225*	-0.9261
S5	1.0452	-0.9438
S6	1.0730	-0.9666

Otro problema recurrente encontrado en las regresiones es la presencia de estadísticos de Durbin-Watson que se encuentran en la región que no rechaza la hipótesis de autocorrelación serial, lo cual hace inconsistente el estimador de la varianza de los estimadores y en consecuencia algunas pruebas. El hallazgo de correlación serial muestra que, en este caso en particular, la DWT no removió la correlación serial.⁷

Buscando alternativas a los problemas antes mencionados en la regresiones estimé otras especificaciones, en primera instancia estimé los modelos restringidos en los que $\beta_0 = -\beta_1$. En la tabla siguiente se muestran los coeficientes estimados. Una vez más se observa que los coeficientes se acercan a 1 a medida que se avanza en el nivel de la transformación pero de acuerdo a la prueba de Wald sólo en el caso de S6 no se rechaza la hipótesis de que el coeficiente sea igual a 1.

	β
S1	0.6991
S2	0.7007
S3	0.7115
S4	0.7322
S5	0.7848
S6	0.9335*

Sin embargo estas regresiones también padecen los problemas de correlación anteriormente descritos e incluso se presentan problemas de heteroscedasticidad. Se especificaron algunas otras especificaciones de tipo polinomial que lamentablemente no resolvieron el problema.⁸ Pese a todo las

⁷Percival y Walden (2000) afirman que esto ocurre para una clase de procesos estocásticos conocidos como procesos de memoria larga.

⁸Los trabajos citados en el capítulo 2 salvan este escollo realizan modelos de corrección de errores luego de probar la existencia de cointegración entre las variables en cuestión

covariancia estimada (muestral) entre los términos de error y la variables explicativas, i.e. IPC.USA e IPC.UK es del orden de 10^{-12} (que puede considerarse como cero) para los distintos modelos estimados, lo cual permite conceder cierta validez a los estimadores pues bajo la hipótesis de no correlación entre las variables explicativas y los errores, los estimadores serán consistentes, no así las correspondientes pruebas de hipótesis que no son del todo válidas. Ahora como el objetivo original del trabajo es mostrar la existencia de una relación de largo plazo que avale la hipótesis PPA bastará con mostrar que no existe independencia entre los smooths de las variables, a pesar de que no se conozca la relación específica que guardan entre sí las variables. De esta manera se procede a probar la hipótesis de independencia entre el tipo de cambio y la diferencia de los índices de precios (doméstico menos externo) mediante el estadístico δ introducido en la sección 3.2.1.

Identificaremos a las variables del presente ejercicio con las del marco teórico en la sección antes mencionada de la siguiente manera: Y es la variable dependiente y corresponde a el tipo de cambio, mientras que X es la variable independiente y corresponde al *spread* o diferencia del índice de precios de Estados Unidos menos el del Reino Unido. Entonces se postula la existencia de una relación de dependencia del tipo $Y = f(X)$ donde la función f no es conocida.

Como primer paso para probar la hipótesis se estimó la función de densidad del estadístico, para el caso particular de 390 observaciones. Se realizó este procedimiento mediante simulaciones del estadístico que permitieron generar el siguiente histograma que aproxima la verdadera función de densidad de δ .⁹

A partir del gráfico se puede establecer una región de rechazo¹⁰ para la hipótesis nula de **independencia**¹¹ cuando $\delta^j \geq 0,042$. En seguida se calcula el valor del estadístico para cada uno de los niveles de *smooths* en los que se busca establecer la relación de dependencia. Como es posible apreciar en la table subsecuente, para todos los niveles desde S1 hasta S6 el estadístico estimado es mayor que 0,1 por tanto se rechaza la hipótesis de independencia; inclusive se encuentra una suerte de “convergencia” hacia el valor que sirve como cota superior del estadístico, i.e. 0.25.

⁹Para realizar la simulaciones y el calculo del estadístico mismo se utilizó el programa escrito originalmente por Contreras-Cristan, González-Barrios y Ruiz-Velasco

¹⁰El cuantil de 0,95 de distribución estimada se halló en 0,041362

¹¹ $H_0 : \delta^j = 0$

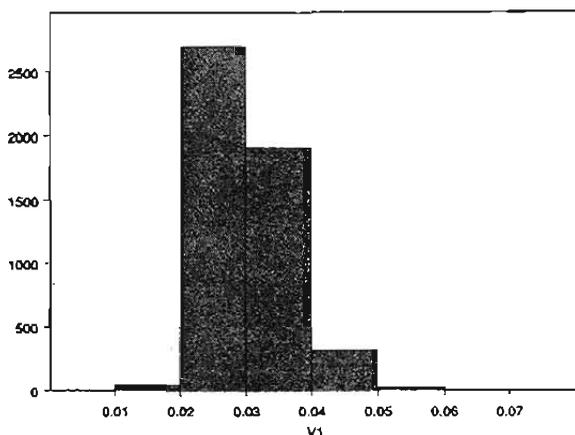


Figura 3.9: Histograma para el estadístico δ

	Estadístico δ
S1	0.173438527
S2	0.170085470
S3	0.174950690
S4	0.181485865
S5	0.192925707
S6	0.209072978

Ahora, buscando probar la validez de la hipótesis de la PPA en su versión relativa se calcula la DWT de la serie del tipo de cambio real como se describió en la sección anterior, esto es, calculando directamente $r = e - p_{USA} + p_{UK}$ utilizando las series transformadas del tipo de cambio y los índices de precios respectivos. Como el interés está centrado en los movimientos de largo plazo se obtienen los *smooths* hasta el sexto nivel (S1 a S6) de la serie del tipo de cambio real. Luego, con el objetivo de probar si estas series son estacionarias se aplica la prueba Dickey-Fuller aumentada cuya hipótesis nula postula que la serie sigue una caminata aleatoria (o bien que posee una raíz unitaria). En caso de que no sea posible el rechazo de la hipótesis nula se concluirá entonces que la serie no es estacionaria. Se presenta a continuación, a modo de ilustración, la pruebas realizadas mediante el paquete *E-views* para la series del primer y sexto *smooths*.

Null Hypothesis: REAL1 has a unit root
 Exogenous: Constant
 Lag Length: 4 (Automatic based on SIC, MAXLAG=16)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.680288	0.0784
Test critical values:		
1% level	-3.447125	
5% level	-2.868829	
10% level	-2.570719	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(REAL1)
 Method: Least Squares
 Sample(adjusted): 6 390
 Included observations: 385 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
REAL1(-1)	-0.026688	0.009957	-2.680288	0.0077
D(REAL1(-1))	1.046983	0.070043	14.93349	0.0000
D(REAL1(-2))	-0.938511	0.089537	-10.48185	0.0000
D(REAL1(-3))	0.616246	0.093233	6.609750	0.0000
D(REAL1(-4))	-0.224984	0.076398	-2.983956	0.0030
C	0.012840	0.005111	2.512186	0.0124
R-squared	0.384083	Mean dependent var		-0.000182
Adjusted R-squared	0.375958	S.D. dependent var		0.030935
S.E. of regression	0.024438	Akaike info criterion		-4.569908
Sum squared resid	0.226342	Schwarz criterion		-4.508299
Log likelihood	885.7072	F-statistic		47.26859
Durbin-Watson stat	1.589513	Prob(F-statistic)		0.000000

Null Hypothesis: REAL6 has a unit root
 Exogenous: Constant
 Lag Length: 10 (Automatic based on SIC, MAXLAG=16)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.455433	0.8965
Test critical values:		
1% level	-3.447395	
5% level	-2.868948	
10% level	-2.570783	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(REAL6)
 Method: Least Squares
 Sample(adjusted): 12 390
 Included observations: 379 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
REAL6(-1)	-0.002041	0.004483	-0.455433	0.6491
D(REAL6(-1))	0.570709	0.057039	10.00564	0.0000
D(REAL6(-2))	0.287950	0.064157	4.488215	0.0000
D(REAL6(-3))	0.294839	0.068593	4.298399	0.0000
D(REAL6(-4))	0.287098	0.071283	4.027581	0.0001
D(REAL6(-5))	0.096492	0.071934	1.341397	0.1806
D(REAL6(-6))	-0.000638	0.070632	-0.009033	0.9928
D(REAL6(-7))	-0.247870	0.065473	-3.785849	0.0002
D(REAL6(-8))	-0.203295	0.063087	-3.222427	0.0014
D(REAL6(-9))	-0.022585	0.039237	-0.575598	0.5652
D(REAL6(-10))	0.022497	0.021644	1.039417	0.2993
C	0.000480	0.002260	0.212280	0.8320

R-squared	0.712460	Mean dependent var	-0.000343
Adjusted R-squared	0.703841	S.D. dependent var	0.012813
S.E. of regression	0.006973	Akaike info criterion	-7.062454
Sum squared resid	0.017843	Schwarz criterion	-6.937783
Log likelihood	1350.335	F-statistic	82.66747
Durbin-Watson stat	1.937144	Prob(F-statistic)	0.000000

Este par de ejemplos nos muestran que en el caso de la serie del *smooth* de primer del tipo de cambio real, sólo es posible rechazar la hipótesis nula con un 10% de significancia. Los subsecuentes *smooths* (S2 hasta S6) no arrojan la misma conclusión, en todos ellos no se rechaza la hipótesis nula, lo cual permite concluir que dichas series siguen un esquema de caminata aleatoria y por tanto son no estacionarias. Estos resultados no permiten avalar la hipótesis de la PPA en su versión relativa. Se realizaron más pruebas con las diferencias de las series y se encontró que sólo es posible rechazar la hipótesis con las segundas diferencias de las series. A continuación se presenta a modo de ejemplo la prueba para el *smooth* S6.

Null Hypothesis: D(REAL6,2) has a unit root
 Exogenous: Constant
 Lag Length: 6 (Automatic based on SIC, MAXLAG=16)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.455741	0.0097
Test critical values:		
1% level	-3.447304	
5% level	-2.868908	
10% level	-2.570761	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(REAL6,3)

Method: Least Squares

Sample(adjusted): 10 390

Included observations: 381 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(REAL6(-1),2)	-0.628372	0.181834	-3.455741	0.0006
D(REAL6(-1),3)	-0.747240	0.162984	-4.584739	0.0000
D(REAL6(-2),3)	-0.845674	0.132055	-6.403944	0.0000
D(REAL6(-3),3)	-0.714079	0.100678	-7.092688	0.0000
D(REAL6(-4),3)	-0.393808	0.066397	-5.931106	0.0000
D(REAL6(-5),3)	-0.090475	0.025212	-3.588633	0.0004
D(REAL6(-6),3)	0.019553	0.019771	0.988958	0.3233
C	-0.000466	0.000373	-1.250349	0.2120

R-squared	0.655938	Mean dependent var	-0.000236
Adjusted R-squared	0.649481	S.D. dependent var	0.012213
S.E. of regression	0.007231	Akaike info criterion	-7.000218
Sum squared resid	0.019501	Schwarz criterion	-6.917429
Log likelihood	1341.542	F-statistic	101.5868
Durbin-Watson stat	1.976019	Prob(F-statistic)	0.000000

A continuación realicé algunas pruebas Dickey-Fuller más buscando encontrar una relación de cointegración de acuerdo a la ecuación (2.7). En la tabla siguiente se muestran los resultados de estas pruebas. Como se puede

	TC	IPC.USA	IPC.UK
Datos	NR	R(1)	R(1)
S1	NR	R(5)	R(1)
S2	NR	NR	NR
S3	NR	R(5)	R(5)
S4	NR	NR	NR
S5	NR	NR	R(5)
S6	NR	NR	NR

Cuadro 3.1: Pruebas de Dickey-Fuller aumentadas para los diferentes *smooths*. NR: no se rechaza la hipótesis nula de raíz unitaria (i.e. $I(1)$); R se rechaza la hipótesis nula (al nivel de significancia señalado entre paréntesis)

apreciar en la tabla anterior los *smooths* de segundo, cuarto y sexto niv-

el para todas las variables son integrados de orden uno, $I(1)$; este es un primer indicio de cointegración entre dichas variables. Ahora se busca hallar

	TC	IPC.USA - IPC.UK
Datos	NR	R(1)
S1	NR	R(5)
S2	NR	NR
S3	NR	NR
S4	NR	NR
S5	NR	R(1)
S6	NR	NR

Cuadro 3.2: Pruebas de Dickey-Fuller aumentadas para los *smooths* de primer a sexto nivel.

una relación de cointegración para las variables en el modelo restringido, la tabla anterior nos da un primer indicio de existencia de cointegración entre las variables de segundo, tercero, cuarto y sexto nivel. Enseguida se realizan las pruebas Dickey-Fuller para los residuales asociados a las regresiones de las variables entre las que se sospecha que existe cointegración. En este caso la prueba requiere variaciones pues la distribución del estadístico t depende del número de variables. Para los residuales de la regresiones de los *smooths* de sexto nivel en el primer modelo se obtuvo un estadístico t de $-3,871157$, mientras que el cuantil de 95 por ciento de la distribución es de $-3,77$ por tanto no se rechaza la hipótesis de no estacionariedad para los residuales de la regresión de sexto nivel y en consecuencia existe un vector de cointegración entre los *smooths* de sexto nivel. En este caso los coeficientes estimados para la relación de largo plazo es $e = 0,0141 + 1,073p_{USA} - 0,9666p_{UK}$, lo cual es muy cercano a lo postulado por la hipótesis postulado por la PPA. Las pruebas para el resto de los residuales rechazan la hipótesis nula de estacionariedad y por tanto no existe cointegración en los demás niveles.

Pese a que los resultados mostrados hasta aquí no nos permiten hacer alguna afirmación concluyente respecto a la validez de la hipótesis de la PPA, rescatemos como resultados relevantes la no independencia hallada entre los *smooths* del tipo de cambio y la diferencia de los niveles de precios; esto nos permite afirmar que existe entre estas variables una relación, aunque la naturaleza de esta no es necesariamente lineal. En este contexto adquiere relevancia el hecho de que los coeficientes estimados por los modelos lineales muestran una especie de "convergencia" hacia los valores postulados por la PPA (i.e. $\beta_0 = 1 = -\beta_1$ o bien $\beta = 1$). Es necesario observar que la no

estacionariedad de los *smooths* es culpable en buena medida de los problemas con las regresiones estimadas. A partir de esta observación se pueden plantear alternativas de solución sobre este mismo problema; por ejemplo es posible realizar el análisis de la señal mediante “wavelet packets”, que es una versión más general de análisis de señal que generaliza a la DWT realizando proyecciones mediante filtros asociados a conjuntos de onduletas básicas que además de localizar el tiempo y la escala también localizan la frecuencia, de manera que siempre es posible obtener series estacionarias a través del filtrado.

Conclusiones

Como hemos podido apreciar a través de este trabajo el reto econométrico postulado por la teoría de la Paridad del Poder Adquisitivo, no fue resuelto (suponiendo que en verdad la PPA sea en verdad cierta) de manera contundente por la metodología aquí propuesta, sin embargo todo el ejercicio arrojó resultados que pueden resultar sumamente interesantes con respecto a la validez de la hipótesis, e incluso en lo que concierne a la estimación de relaciones de largo plazo entre variables.

En primera instancia las regresiones estimadas con los *smooths* dieron como resultado coeficientes muy cercanos a los que la PPA requiere, (particularmente en las estimaciones realizadas con los *smooths* a partir del segundo nivel) lo cual no es una característica común en este tipo de trabajos. Y aún cuando se puede cuestionar la validez de estas estimaciones, en el análisis hecho a través del estadístico δ se establece que en verdad existe “un tipo” de dependencia (relación) de largo plazo entre las variables, aunque ésta no sea de carácter necesariamente lineal (y que quizá sea la razón por la cual las estimaciones lineales sean tan defectuosas). Por tanto no es aventurado afirmar que las regresiones presentadas en el capítulo 3, pese a presentar problemas de correlación serial no se encuentran dentro de la categoría de regresiones espurias.¹² Incluso las pruebas de raíz unitaria (Dickey-Fuller aumentada) parecen indicar que existe un vector de cointegración entre los *smooths* de sexto nivel de las variables, ya que las versiones suavizadas de todas las variables son $I(1)$ (i.e. poseen un raíz unitaria) y los residuales rechazan la hipótesis de raíz unitaria en favor de la estacionariedad. De acuerdo a toda esta evidencia no es descabellado afirmar que se cumple una relación de largo plazo (entre el tipo de cambio y los índices de precios del país doméstico y externo) como la PPA, aunque esta no esté dada por una relación lineal.

¹²Formalmente para postula que una regresión es espuria si no existe alguna estimación de los parametros que haga que los errores sean estacionarios.

Por lo que respecta a la estimación de relaciones de largo plazo mediante onduletas es necesario decir que se requiere aún mucha investigación en dicho campo (tanto empírica como teórica) para poder apreciar todas sus posibles ventajas así como para establecer sus alcances y limitantes.

Bibliografía

- [1] **Almasri, Abdullah y Shukur, Ghazi.** (2003) *An illustration of the causality relation between government spending and revenues using wavelet analysis on Finnish data*, Journal of Applied Statistics, Vol 30, No. 5, 571-584
- [2] **Balassa, Bela.** (1964) *The purchasing-power parity doctrine: A reappraisal*, Universidad de Yale, Diciembre 1964. p.: 584-585
- [3] **Bruce, Andrew y Gao, Hong-Ye** (1996) *Applied Wavelet Analysis with S-PLUS*, Springer-Verlag
- [4] **Daubechies, Ingrid.** (1992) *Ten Lecture on Wavelets*, Volumen 61 de la *CBMS-NFS Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia.
- [5] **Froot, Kenneth y Rogoff, Kenneth.** (1994) *Perspectives on PPP and Long-run Real Exchange Rates* National Bureau of Economic Research.
- [6] **Goffe, William L..** (1994) *Wavelets in Macroeconomics: An Introduction en Computational Techniques for Econometrics and Economic Analysis*, 137-149. Kluwer Academic Publishers. Netherlands
- [7] **Hamilton, James D.** (1994) *Time Series Analysis* Princeton UNiversity Press.
- [8] **Kaiser, Gerald.** (1994) *A Friendly Guide to Wavelets*, Birkhäuser Boston.
- [9] **Krugman, Paul R. y Obstfeld, Maurice.** (1995), *Economía Internacional Teoría y Política*, Mc Graw Hill.
- [10] **Ogden, R. Todd.** (1997) *Essential Wavelets for Statistical Applications and Data Analysis*, Birkhäuser, Boston

- [11] **Michael, Mussa.** (1986) *Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates*. Carnegie-Rochester Conference Series on Public Policy 25 (Autumn): 117-219.
- [12] **Percival, Donald y Walden, Andrew.** (2000) *Wavelet Methods for Time Series Analysis*, Cambridge University Press.
- [13] **Ramsey, B. James.** (2002) *The contribution of wavelets to the analysis of economic and financial data*
- [14] **Ramsey, B James y Lampart, Camille.** (1997) *The Decomposition of Economic Relationships by Time Scale using Wavelets*, Economic Research Reports Number 97-08 January-1997 C.V. Starr Center for Applied Economics New York University.
- [15] **Ramsey, B. James.** (1999) *Regression over Timescale Decompositions: A sampling Analysis of Distributional Properties*, Economic System Research, Vol. 11. No. 2, 1999.
- [16] **Shumway, Robert H. y Stoffer, David S.** (2000) *Time Series Analysis and Its Applications*, Springer-Verlag, New York.
- [17] **Vidakovic, Brani.** (1999) *Statistical Modeling by Wavelets*, John Wiley and Sons.

Apéndices

1. Aspectos técnicos de la DWT

El Filtro de Onduleta

Ahora vamos a estudiar en detalle la construcción de los renglones de la matriz \mathcal{W} que define a la DWT, para el caso general y no sólo para la onduleta de Haar. Dicha construcción estará formulada en términos de un algoritmo que permite factorizar a \mathcal{W} en términos de matrices muy simples. Este algoritmo es conocido como el *algoritmo pirámide* y fue introducido en el contexto de onduletas por Mallat (1989). Este permite que $\mathbf{W} = \mathcal{W}\mathbf{X}$ sea calculado utilizando solamente $O(N)$ multiplicaciones, mientras que calcular directamente el producto de la matriz \mathcal{W} de $N \times N$ y el vector \mathbf{X} requiere N^2 multiplicaciones (la notación $O(a_N)$ significa que existe una constante C tal que el verdadero número de multiplicaciones es menor o igual a Ca_N para todo N). Este algoritmo es, inclusive, más rápido que el algoritmo de la transformada rápida de Fourier que requiere $O(N \log_2 N)$ multiplicaciones.

Para describir el algoritmo pirámide se usarán tanto operaciones de filtrado lineal como manipulación de matrices. Comenzamos, con el enfoque de filtros, el cual está basado en el filtro de onduletas $\{h_l : l = \dots, L-1\}$, donde L es el ancho del filtro¹³ y tiene que ser un número entero par. Para que $\{h_l\}$ tenga ancho L , es necesario que $h_0 \neq 0$ y $h_{L-1} \neq 0$, definimos $h_l = 0$ para $l < 0$ y $l \geq L$ de tal manera que $\{h_l\}$ es una sucesión infinita con a lo más L valores distintos de cero. Un filtro de onduletas debe satisfacer las siguientes tres propiedades:

$$\sum_{l=0}^{L-1} h_l = 0, \quad (3.5)$$

¹³Se entiende aquí por ancho del filtro el número de elementos consecutivos en la sucesión que define al filtro tales que el primero y el último son distintos de cero y cualquier elemento en la sucesión anterior y posterior a ellos respectivamente es igual a cero.

$$\sum_{l=0}^{L-1} h_l^2 = 1 \quad (3.6)$$

y

$$\sum_{l=0}^{L-1} h_l h_{l+2n} = \sum_{l=-\infty}^{\infty} h_l h_{l+2n} = 0, \quad (3.7)$$

para todo entero n distinto de cero. En otras palabras un filtro de onduletas tiene que sumar cero; debe tener energía unitaria; y debe ser ortogonal con respecto a sus traslaciones pares. La primera condición se realiza en el espíritu de la noción básica de onduleta. Nos referiremos a las ecuaciones (3.6) y (3.7) como las *propiedad de ortogonormalidad* de los filtros de onduletas. Como ejemplo de estos filtros tenemos al filtro de onduletas de Haar (de la sección anterior), $\left\{ h_0 = \frac{1}{\sqrt{2}}, h_1 = \frac{-1}{\sqrt{2}} \right\}$ para el cual $L = 2$ y el filtro de D(4) para el cual $L = 4$.

$$h_0 = \frac{1 - \sqrt{3}}{4\sqrt{2}}, h_1 = \frac{-3 + \sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, \text{ y } h_3 = \frac{-1 - \sqrt{3}}{4\sqrt{2}} \quad (3.8)$$

Sea $H(\cdot)$ la función de transferencia para $\{h_l\}$, i.e.

$$H(f) \equiv \sum_{l=-\infty}^{\infty} h_l e^{-2i\pi fl} = \sum_{l=0}^{L-1} h_l e^{-2i\pi fl} \quad (3.9)$$

y sea $\mathcal{H}(\cdot)$ que denota la función de ganancia cuadrada asociada:

$$\mathcal{H}(f) \equiv |H(f)|^2$$

La siguiente igualdad¹⁴, expresada en términos de la función de ganancia cuadrada $\mathcal{H}(\cdot)$, es equivalente a las ecuaciones (3.6) y (3.7)

$$\mathcal{H}(f) + \mathcal{H}(f + \frac{1}{2}) = 2 \quad \text{para todo } f \in [-1/2, 1/2] \quad (3.10)$$

Para obtener los coeficientes de onduletas asociados con la escala unitaria, filtraremos circularmente la serie de tiempo $\{X_t : t = 0, \dots, N-1\}$ con $\{h_l\}$ y retendremos solo los valores con índice impar del resultado, donde $N \equiv 2^J$ para algún entero positivo J . Denotemos el resultado de filtrar circularmente a $\{X_t : t = 0, \dots, N-1\}$ con $\{h_l\}$ como:

$$2^{1/2} \widetilde{W}_{1,t} \equiv \sum_{l=0}^{L-1} h_l X_{t-l \bmod N} \quad t = 0, \dots, N-1 \quad (3.11)$$

¹⁴Para su demostración remitase a Percival y Walden pp. 69-70, 512.

Definimos el coeficiente de onduletas para la escala unitaria como:

$$W_{1,t} \equiv 2^{1/2} \widetilde{W}_{1,2t+1} = \sum_{l=0}^{L-1} h_l X_{2t+1-l \bmod N} \quad t = 0, \dots, \frac{N}{2} - 1 \quad (3.12)$$

El primero de los dos subíndices en $W_{1,t}$ (y $\widetilde{W}_{1,t}$) se refiere a la escala $\tau_j = 2^{j-1}$ asociada con esos $N/2$ coeficientes, en este caso $j = 1$ es el índice para la escala unitaria. Noten que los coeficientes de onduletas $\{W_{1,t}\}$ están dados por los $N/2$ valores con índices impares en el resultado del filtro reescalado $\{2^{1/2} \widetilde{W}_{1,t}\}$. Al procedimiento de tomar valores “saltados” del resultado del filtro se llama submuestreo por dos o bien *downsampling* por dos. La raíz cuadrada de dos que aparece en la ecuación (3.12) se necesita para preservar la energía después del submuestreo.

Podemos conectar la definición de $\{W_{1,t}\}$ con la formulación a través de matrices (de la transformada en onduletas) $\mathbf{W} = \mathcal{W}\mathbf{X}$, donde como antes \mathbf{W} es el vector columna de los coeficientes de la DWT, \mathcal{W} es la matriz de $N \times N$ que define a la DWT y \mathbf{X} es el vector columna de dimensión N que contiene a la serie de tiempo $\{X_t\}$. Los primeros $N/2$ elementos de \mathbf{W} , i.e. el subvector \mathbf{W}_1 está definido como $W_{1,t}$ para $t = 0, \dots, \frac{N}{2} - 1$. Como $\mathbf{W}_1 = \mathcal{W}_1 \mathbf{X}$, donde \mathcal{W}_1 es la matriz de $\frac{N}{2} \times N$ que consiste de las $N/2$ primeras filas de \mathcal{W} , la definición de \mathbf{W}_1 implica una definición para las filas de \mathcal{W}_1 . Para ver como son precisamente esas filas, reescribamos la ecuación (3.12) como:

$$W_{1,t} = \sum_{l=0}^{N-1} h_l^\circ X_{2t+1-l \bmod N} \quad t = 0, \dots, \frac{N}{2} - 1 \quad (3.13)$$

donde $\{h_l^\circ : l = 0, \dots, N-1\}$ es $\{h_l\}$ periodizado a longitud N . Para $0 \leq t \leq \frac{N}{2} - 1$ la t -ésima fila $\mathcal{W}_{t\bullet}^T$ de \mathcal{W} o \mathcal{W}_1 tenemos que:

$$W_{1,t} = \mathcal{W}_{t\bullet}^T \mathbf{X} = \sum_{l=0}^{N-1} h_l^\circ X_{2t+1-l \bmod N} = \sum_{l=0}^{N-1} h_{2t+1-l \bmod N}^\circ X_l \quad (3.14)$$

Para $t = 0$ tenemos

$$W_{1,0} = \mathcal{W}_{0\bullet}^T \mathbf{X} = \sum_{l=0}^{N-1} h_{1-l \bmod N}^\circ X_l \quad (3.15)$$

entonces

$$\mathcal{W}_{0\bullet}^T = \{h_{1}^\circ, h_0^\circ, h_{N-1}^\circ, h_{N-2}^\circ, \dots, h_2^\circ\} \quad (3.16)$$

es evidente de la ecuación (3.14) que la $\frac{N}{2} - 1$ filas restantes de \mathcal{W}_1 pueden expresarse como una versión trasladada circularmente de $\mathcal{W}_{0\bullet}^T$, esto es:

$$\mathcal{W}_{t\bullet}^T = [\mathcal{T}^{2t} \mathcal{W}_{0\bullet}]^T, \quad t = 1, \dots, \frac{N}{2} - 1$$

donde \mathcal{T} es la matriz de permutaciones circulares de $N \times N$. Por ejemplo para $t = 1$ tenemos que:

$$\mathcal{W}_{1\bullet}^T = [h_3^\circ, h_2^\circ, h_1^\circ, h_0^\circ, h_{N-1}^\circ, h_{N-2}^\circ, \dots, h_4^\circ] \quad (3.17)$$

Ahora estamos en posición de probar que las filas de \mathcal{W}_1 constituyen un conjunto de $N/2$ vectores ortonormales. Cuando $L \leq N$, el filtro periodizado toma la forma:

$$h_l^\circ = \begin{cases} h_l, & 0 \leq l \leq L - 1; \\ 0, & L \leq l \leq N - 1. \end{cases}$$

entonces la primera columna de \mathcal{W}_1 es

$$\mathcal{W}_{0\bullet}^T = [h_1, h_0, \underbrace{0, \dots, 0}_{N-L \text{ ceros}}, h_{L-1}, \dots, h_2] \quad (3.18)$$

En este caso, como \mathcal{T}^{2t} es una transformación ortonormal, tenemos que:

$$\langle \mathcal{W}_{t\bullet}, \mathcal{W}_{t\bullet} \rangle = \|\mathcal{W}_{t\bullet}\|^2 = \|\mathcal{T}^{2t} \mathcal{W}_{0\bullet}\|^2 = \|\mathcal{W}_{0\bullet}\|^2 = \sum_{l=0}^{L-1} h_l^2 = 1$$

por que el filtro tiene energía unitaria. Para $t' \neq t$ también tenemos que

$$\langle \mathcal{W}_{t'\bullet}, \mathcal{W}_{t\bullet} \rangle = \mathcal{W}_{t'\bullet}^T \mathcal{W}_{t\bullet} = \mathcal{W}_{0\bullet}^T \mathcal{T}^{-2t'} \mathcal{T}^{2t} \mathcal{W}_{0\bullet} = \sum_{l=0}^{L-1} h_l h_{l+2(t-t')} = 0$$

por que el filtro de onduletas es ortogonal con sus traslaciones pares. Entonces, cuando $L \leq N$ la prueba de ortonormalidad de las filas de \mathcal{W}_1 se sigue directamente de dos de las tres propiedades básicas de un filtro de ondeleta

La prueba anterior no puede ser fácilmente adaptada para el caso en el que $L > N$, pero el siguiente enfoque funciona aún cuando $L \leq N$ y $L > N$. Como las últimas $\frac{N}{2} - 1$ filas de \mathcal{W}_1 están formadas por traslaciones circulares en múltiplos de 2 de $\mathcal{W}_{0\bullet}^T$ y como $\mathcal{W}_{0\bullet}^T$, está dado por la ecuación (3.16), podemos establecer la ortonormalidad si mostramos que: ¹⁵

$$h^\circ \star h_l^\circ \equiv \sum_{n=0}^{N-1} h_n^\circ h_{n+l \bmod N}^\circ = \begin{cases} 1, & \text{si } l = 0; \\ 0, & \text{si } l = 2, 4, \dots, N - 2 \end{cases} \quad (3.19)$$

¹⁵Los detalles de la demostración se proporcionan en Percival y Walden (2000) p. 72, 512.

teniendo en cuenta que

$$\{h_l^\circ : l = 0, \dots, N-1\} \longleftrightarrow \{H(\frac{k}{N}) : k = 0, \dots, N-1\}$$

¹⁶ entonces $\{h_l^\circ \star h_l^\circ\} \longleftrightarrow \{|H(\frac{k}{N})|^2 = \mathcal{H}(\frac{k}{N})\}$.

Debido a que hemos definido los coeficientes de onduletas de escala unitaria en términos del resultado de un filtro, podemos tener otra interpretación de estos coeficientes usando algunas herramientas de la teoría de filtros. Sea

$$\mathcal{X}_k \equiv \sum_{t=0}^{N-1} X_t e^{-i2\pi tk/N}, \quad k = 0, \dots, N-1$$

la transformada de Fourier discreta DFT de $\{X_t\}$. Una aplicación del teorema de Parseval establece que¹⁷:

$$\mathcal{E}_X = \sum_{t=0}^{N-1} X_t^2 = \frac{1}{N} \sum_{k=0}^{N-1} |\mathcal{X}_k|^2$$

de tal manera que $|\mathcal{X}_k|^2/N$ define un espectro de energía sobre la frecuencia k/N . Ahora consideramos el resultado de filtrar $\{X_t\}$ con $\{h_l\}$ para producir

$$2^{1/2} \widetilde{W}_{1,t} = \sum_{l=0}^{L-1} h_l X_{t-l \bmod N} = \sum_{l=0}^{N-1} h_l^\circ X_{t-l \bmod N} \quad t = 0, \dots, N-1$$

La sucesión $\{2^{1/2} \widetilde{W}_{1,t}\}$ se obtiene entonces convolucionando circularmente $\{X_t\}$ con la sucesión $\{h_l^\circ\}$ de longitud N . La DFT de esta última sucesión es $\{H(\frac{k}{N})\}$, por tanto,

$$\{2^{1/2} \widetilde{W}_{1,t}\} \longleftrightarrow \{H(\frac{k}{N}) \mathcal{X}_k\}$$

del teorema de Parseval tenemos que:

$$2 \sum_{t=0}^{N-1} \widetilde{W}_{1,t}^2 = \frac{1}{N} \sum_{k=0}^{N-1} |H(\frac{k}{N}) \mathcal{X}_k|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |H(\frac{k}{N})|^2 |\mathcal{X}_k|^2 = \frac{1}{N} \sum_{k=0}^{N-1} \mathcal{H}(\frac{k}{N}) |\mathcal{X}_k|^2$$

¹⁶Una prueba de este hecho así como el corolario que se enuncia enseguida se encuentra en Percival y Walden (2000) p.29-30

¹⁷Algunas demostraciones de este resultado se encuentran en Wei(1989) p.213 y Percival y Walden (2000) p. 23, 502.

La ecuación anterior dice que el espectro de energía en la frecuencia k/N del resultado del filtro está dado por $\mathcal{H}(\frac{k}{N})|\mathcal{X}_k|^2/N$ que es justamente $\mathcal{H}(\frac{k}{N})$ veces el espectro de energía en la frecuencia k/N de lo que se filtró, i.e. $\{X_t\}$.

El efecto de filtrar $\{X_t\}$ para obtener $\{2^{1/2}\widetilde{W}_{1,t}\}$ puede ser evaluado estudiando $\mathcal{H}(\frac{k}{N})$ versus la frecuencia k/N (debido a que la función de ganancia cuadrada satisface que $\mathcal{H}(-f) = \mathcal{H}(f)$ y es periódica con periodo igual a la unidad, sólo es necesario graficarla para $0 \leq f \leq \frac{1}{2}$), así los filtros de onduletas pueden ser vistos como aproximaciones a un filtro *high pass* con paso de banda definido por $\frac{1}{4} \leq |f| \leq \frac{1}{2}$. En seguida se muestran las gráficas de $\mathcal{H}(\cdot)$ para el filtro de Haar y para D(4).

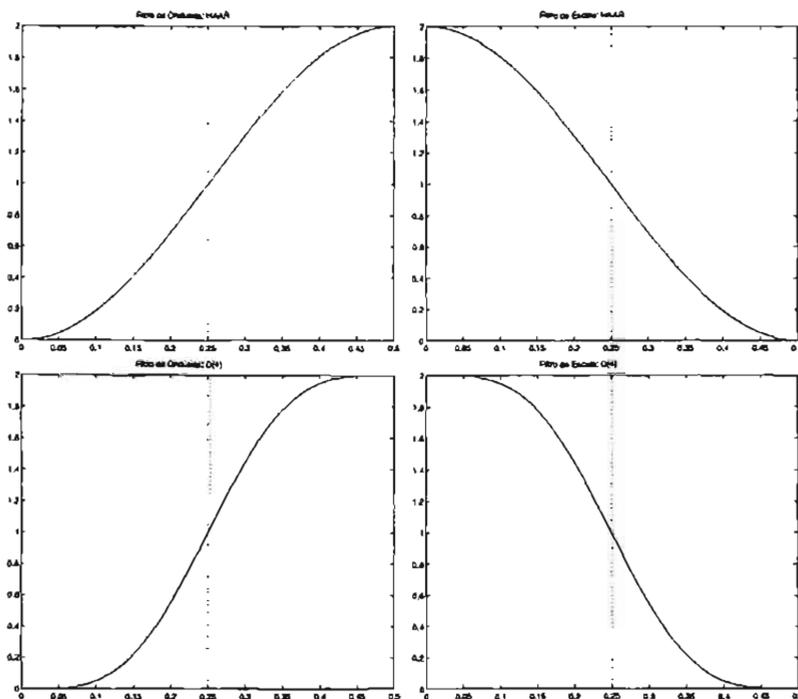


Figura 3.10: Funciones de ganancia cuadrada (modulo al cuadrado de la transformada de Fourier del filtro) para los filtros de onduletas(izquierda) y escala (derecha), para Haar (arriba) y Daubechies D(4) (abajo).

El Filtro de Escala

En la sección anterior utilizamos el filtro de onduletas $\{h_t\}$ para construir los primeros $N/2$ renglones de la matriz W de la DWT estas filas constituyen

la matriz \mathcal{W}_1 en la descomposición de \mathcal{W} dada por la ecuación (1.17). Para formar los últimos $N/2$ filas de \mathcal{W} via el algoritmo pirámide, definiremos un segundo filtro que será utilizado para construir una matriz \mathcal{V}_1 de $\frac{N}{2} \times N$. Mostraremos que esta matriz genera el mismo subespacio que la últimas $N/2$ filas de la matriz \mathcal{W} . Excepto para el caso $N = 2$ que implica $J = 1$ en la ecuación (1.17), las filas de \mathcal{V}_1 no son iguales a la últimas $N/2$ filas de \mathcal{W} , pero se pueden obtener esas filas mediante la subsecuente manipulación de \mathcal{V}_1 . El segundo filtro requerido es el filtro de “cuadratura espejo” (*quadrature mirror filter* QMF) $\{g_l\}$ que corresponde a $\{h_l\}$:

$$g_l \equiv (-1)^{l+1} h_{L-1-l} \quad (3.20)$$

para futura referencia, establecemos la relación inversa.

$$h_l = (-1)^l g_{L-1-l} \quad (3.21)$$

El filtro $\{g_l\}$ es conocido como el *filtro de escala*. Tomemos como ejemplo al filtro con la ondeleta de Haar dada por $h_0 = 1/\sqrt{2}$ y $h_1 = -1/\sqrt{2}$, la porción no cero del correspondiente filtro de escala es:

$$g_0 = -h_1 = 1/\sqrt{2} \quad \text{y} \quad g_1 = h_0 = 1/\sqrt{2} \quad (3.22)$$

La función de transferencia $G(\cdot)$ para $\{g_l\}$ está dada por:

$$G(f) \equiv \sum_{l=0}^{L-1} g_l e^{-i2\pi fl} = \sum_{l=0}^{L-1} (-1)^{l+1} h_{L-1-l} e^{-i2\pi fl} \quad (3.23)$$

es posible probar que¹⁸

$$\mathcal{G}(f) \equiv |G(f)|^2 = \left| e^{-i2\pi f(L-1)} H\left(\frac{1}{2} - f\right) \right|^2 = \left| H\left(\frac{1}{2} - f\right) \right|^2 = \mathcal{H}\left(\frac{1}{2} - f\right)$$

entonces

$$\mathcal{G}(f) = \mathcal{H}\left(\frac{1}{2} - f\right)$$

donde $\mathcal{G}(f) \equiv |G(f)|^2$ es la función de ganancia cuadrada. Como $\mathcal{G}(f) = \mathcal{H}\left(\frac{1}{2} - f\right)$ y haciendo uso de la ecuación (3.10) tenemos que:

$$\mathcal{G}(f) + \mathcal{G}\left(f + \frac{1}{2}\right) = 2 \quad \text{o} \quad \mathcal{G}(f) + \mathcal{H}(f) = 2 \quad \text{para toda } f \in [-1/2, 1/2] \quad (3.24)$$

(recordemos el hecho de que las funciones $\mathcal{G}(\cdot)$ y $\mathcal{H}(\cdot)$ son periódicas de periodo uno). Observemos que la segunda relación puede corroborarse fácilmente a través de las gráficas en la figura (1.4). Esto implica que si el filtro

¹⁸Percival y Walden (2000) p.76, 513

de onduleta es similar a un filtro *high-pass*, entonces el filtro de escala debe ser similar a un filtro *low-pass*.¹⁹ Para el caso de la onduleta de Haar y D(4) sus filtros de escala pueden ser vistos como aproximaciones a filtros *low-pass* con paso de banda definido por $0 \leq |f| \leq 1/4$. Los filtros *high-pass* $\{h_l\}$ y *low-pass* $\{g_l\}$ a veces son llamados filtros *half-band* por que seccionan la banda frecuencia $[0, 1/2]$ por mitades.

Enumeremos algunas propiedades del filtro:²⁰

$$\sum_{l=0}^{L-1} g_l = \sqrt{2} \quad \circ \quad \sum_{l=0}^{L-1} g_l = -\sqrt{2}$$

$$\sum_{l=0}^{L-1} g_l^2 = 1 \quad \text{y} \quad \sum_{l=0}^{L-1} g_l g_{l+2n} = \sum_{l=-\infty}^{\infty} g_l g_{l+2n} = 0$$

para n entero distinto de cero. De estas igualdades podemos inferir que los filtros de escala son similares a los de onduletas en el sentido de que ambos satisfacen la condición de ortonormalidad. Realizando un procedimiento similar al que se usó para construir la matriz \mathcal{W}_1 en base a $\{h_l\}$ se construye \mathcal{V}_1 en base a $\{g_l\}$ y como consecuencia de la ortonormalidad de este último filtro se sigue que las filas de \mathcal{V}_1 son ortonormales por pares. A continuación describimos la construcción de la matriz \mathcal{V}_1 . Denotemos de la siguiente manera a lo que resulta de filtrar $\{X_t\}$ con $\{g_l\}$:

$$2^{1/2} \tilde{V}_{1,t} \equiv \sum_{l=0}^{L-1} g_l X_{t-l \bmod N} \quad t = 0, \dots, N-1 \quad (3.25)$$

Definamos los coeficientes de escala de primer nivel como:

$$V_{1,t} \equiv 2^{1/2} \tilde{V}_{1,2t+1} = \sum_{l=0}^{L-1} g_l X_{2t+1-l \bmod N} = \sum_{l=0}^{N-1} g_l^\circ X_{2t+1-l \bmod N} \quad (3.26)$$

para $t = 0, \dots, \frac{N}{2} - 1$, donde $\{g_l^\circ\}$ es $\{g_l\}$ periodizado para tener longitud N . Sea \mathbf{V}_1 el vector de longitud $N/2$ cuyo t -ésimo elemento es $V_{1,t}$. Sea \mathcal{V}_1 la matriz de $\frac{N}{2} \times N$ cuya primera fila está dada por:

$$[g_1^\circ, g_0^\circ, g_{N-1}^\circ, g_{N-2}^\circ, \dots, g_2^\circ] \equiv \mathcal{V}_{0\bullet}^T$$

¹⁹Ibid. p. 76

²⁰Se ofrecen demostraciones de estas propiedades en Percival y Walden (2000) p.513

y cuyas $\frac{N}{2} - 1$ filas están dadas por $[\mathcal{T}^{2t}\mathcal{V}_{0\bullet}]^T, t = 1, \dots, \frac{N}{2} - 1$. Entonces $\mathcal{V}_1 = \mathcal{V}_1\mathbf{X}$ y las filas de \mathcal{V}_1 constituyen un conjunto de $N/2$ vectores ortonormales.²¹

Ahora probaremos que las filas de \mathcal{V}_1 y de \mathcal{W}_1 constituyen un conjunto de N vectores ortonormales. Como la t -ésima fila de \mathcal{V}_1 y de \mathcal{W}_1 son respectivamente $[\mathcal{T}^{2t}\mathcal{V}_{0\bullet}]^T$ y $[\mathcal{T}^{2t}\mathcal{W}_{0\bullet}]^T$ es necesario mostrar que:

$$\langle \mathcal{T}^{2t}\mathcal{V}_{0\bullet}, \mathcal{T}^{2t'}\mathcal{W}_{0\bullet} \rangle = 0 \quad \text{para } 0 \leq t \leq t' \leq \frac{N}{2} - 1$$

Definiendo $n = t' - t$, tenemos que para $n = 0, \dots, \frac{N}{2} - 1$,

$$\langle \mathcal{T}^{2t}\mathcal{V}_{0\bullet}, \mathcal{T}^{2t'}\mathcal{W}_{0\bullet} \rangle = \mathcal{V}_{0\bullet}^T \mathcal{T}^{-2t} \mathcal{T}^{2t'} \mathcal{W}_{0\bullet} = \mathcal{V}_{0\bullet}^T \mathcal{T}^{2n} \mathcal{W}_{0\bullet} = \sum_{l=0}^{N-1} g_l^\circ h_{l+2n \bmod N}^\circ$$

cuando $L \leq N$ se cumple que $g_l^\circ = g_l$ y $h_l^\circ = h_l$ y lo anterior se reduce a

$$\langle \mathcal{T}^{2t}\mathcal{V}_{0\bullet}, \mathcal{T}^{2t'}\mathcal{W}_{0\bullet} \rangle = \sum_{l=0}^{L-1} g_l h_{l+2n}$$

para el caso especial en que $L \leq N$ la ortonormalidad se sigue de la siguiente igualdad²²

$$\sum_{l=0}^{L-1} g_l h_{l+2n} = \sum_{l=-\infty}^{\infty} g_l h_{l+2n} = 0$$

en otras palabras la ecuación anterior dice que los filtros de escala y onduleta y cualquier traslación par de ellos son ortogonales entre sí.

La prueba anterior de ortonormalidad no se puede extender fácilmente para el caso en que $L > N$. Una manera que es válida en ambos casos (i.e $L \leq N$ y $L > N$) es la siguiente. Es necesario probar que

$$\sum_{l=0}^{N-1} g_l^\circ h_{l+2n \bmod N}^\circ = g^\circ \star h_{2n}^\circ = 0 \quad \text{para } n = 0, \dots, \frac{N}{2} - 1 \quad (3.27)$$

donde $\{g^\circ \star h_l^\circ\}$ es la correlación cruzada circular de $\{g_l^\circ\}$ y $\{h_l^\circ\}$, (recordando que $g_l^\circ = [g_l^\circ]^*$ debido a que son reales) y haciendo uso del hecho que $\{g_l^\circ\} \longleftrightarrow \{G(\frac{k}{N})\}$ y $\{h_l^\circ\} \longleftrightarrow \{H(\frac{k}{N})\}$ implican que $\{g^\circ \star h_l^\circ\} \longleftrightarrow \{G^*(\frac{k}{N})H(\frac{k}{N})\}$.²³

²¹Ibid. p. 77

²²Se ofrece una prueba de esta igualdad en Percival y Walden(2000) p.514

²³La demostración detallada de la ecuación 3.27 se encuentra en Percival y Walden (2000) p. 78 y 515.

En notación de matrices, el hecho de que cada fila de \mathcal{V}_1 es ortogonal a cada fila de \mathcal{W}_1 se puede expresar como:

$$\mathcal{W}_1 \mathcal{V}_1^T = \mathcal{V}_1 \mathcal{W}_1^T = 0_{N/2}$$

donde $0_{N/2}$ es la matriz de $\frac{N}{2} \times \frac{N}{2}$ cuyos elementos son todos cero. Como sabemos que $\mathcal{W}_1 \mathcal{W}_1^T = I_{N/2}$ y $\mathcal{V}_1 \mathcal{V}_1 = I_{N/2}$ donde $I_{N/2}$ es la matriz identidad de $\frac{N}{2} \times \frac{N}{2}$, la matriz

$$\mathcal{P}_1 \equiv \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{V}_1 \end{bmatrix} \quad (3.28)$$

es ortonormal por que

$$\mathcal{P}_1 \mathcal{P}_1^T = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{V}_1 \end{bmatrix} \begin{bmatrix} \mathcal{W}_1^T & \mathcal{V}_1^T \end{bmatrix} = \begin{bmatrix} \mathcal{W}_1 \mathcal{W}_1^T & \mathcal{W}_1 \mathcal{V}_1^T \\ \mathcal{V}_1 \mathcal{W}_1^T & \mathcal{V}_1 \mathcal{V}_1^T \end{bmatrix} = I_N \quad (3.29)$$

Como la matriz de la DWT \mathcal{W} es también ortonormal y como las primeras $N/2$ filas de \mathcal{W} y \mathcal{P}_1 son idénticas se sigue que las últimas $N/2$ filas de \mathcal{P}_1 , a saber \mathcal{V}_1 , tienen que generar el mismo subespacio que las últimas $N/2$ filas de \mathcal{W} . Excepto en el caso en el que $N = 2$, las filas de \mathcal{V}_1 y las últimas $N/2$ filas de \mathcal{W} no son las mismas, pero podemos manipular \mathcal{V}_1 para obtener estas filas.

El Algoritmo Pirámide

La primera etapa

La primera etapa del algoritmo pirámide para calcular la DWT consiste simplemente en transformar la serie de tiempo \mathbf{X} de longitud $N = 2^J$ en los $N/2$ coeficientes de ondulatas de primer nivel \mathbf{W}_1 y los $N/2$ coeficientes de escala de primer nivel \mathbf{V}_1 . Habrá $J - 1$ etapas subsiguientes del algoritmo pirámide. En la j -ésima etapa para $j = 2, \dots, J$ se transforma al vector \mathbf{V}_{j-1} de longitud $N/2^{j-1}$ en los vectores \mathbf{W}_j y \mathbf{V}_j cada uno de longitud $N/2^j$. En la j -ésima etapa se trata a \mathbf{V}_{j-1} exactamente de la misma manera que a \mathbf{X} en la primera: los elementos de \mathbf{V}_{j-1} son filtrados separadamente con $\{h_l\}$ y $\{g_l\}$ y los resultados de los filtros son submuestreados para formar, respectivamente \mathbf{W}_j y \mathbf{V}_j . Los elementos de \mathbf{V}_j son llamados los coeficientes de escala para el nivel j , mientras que \mathbf{W}_j contiene los coeficientes de ondulatas para el nivel j . Al final de la J -ésima etapa se pueden formar el vector de coeficientes de la DWT \mathbf{W} concatenando a los $J + 1$ vectores $\mathbf{W}_1, \dots, \mathbf{W}_J$ y \mathbf{V}_J .

Consideremos la síntesis de \mathbf{X} . Utilizando la matriz \mathcal{P}_1 podemos expresar la primera etapa del algoritmo como:

$$\mathcal{P}_1 \mathbf{X} = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{V}_1 \end{bmatrix} \mathbf{X} = \begin{bmatrix} \mathcal{W}_1 \mathbf{X} \\ \mathcal{V}_1 \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{V}_1 \end{bmatrix}$$

y como \mathcal{P}_1 es una matriz ortonormal, podemos recuperar (sintetizar) \mathbf{X} utilizando la ecuación:

$$\mathbf{X} = \mathcal{P}_1^T \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{V}_1 \end{bmatrix} = \begin{bmatrix} \mathcal{W}_1^T & \mathcal{V}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{V}_1 \end{bmatrix} = \mathcal{W}_1^T \mathbf{W}_1 + \mathcal{V}_1^T \mathbf{V}_1 \quad (3.30)$$

recordemos que el primer nivel de detalles está definido por $\mathcal{D}_1 \equiv \mathcal{W}_1^T \mathbf{W}_1$ y que el primer nivel del *smooth* \mathcal{S}_1 es tal que $\mathcal{S}_1 + \mathcal{D}_1 = \mathbf{X}$. Ahora comparando con la ecuación anterior tenemos que

$$\mathcal{S}_1 = \mathcal{V}_1^T \mathbf{V}_1 = \mathcal{V}_1^T \mathcal{V}_1 \mathbf{X},$$

de tal manera que la matriz $\mathcal{V}_1^T \mathcal{V}_1$ de $N \times N$ puede ser vista como un operador que extrae el *smooth* de onduleta de primer nivel de \mathbf{X} . Del mismo modo podemos escribir

$$\mathcal{D}_1 = \mathcal{W}_1^T \mathcal{W}_1 \mathbf{X},$$

lo cual muestra como formar \mathcal{D}_1 via una operación de matrices sobre \mathbf{W} .

Si definimos la “versión *upsampled*” \mathbf{W}_1 cuyos elementos son

$$W_{1,t}^\uparrow \equiv \begin{cases} 0, & t = 0, 2, \dots, N-2; \\ W_{1, \frac{t-1}{2}} & t = 1, 3, \dots, N-1; \end{cases}$$

es posible escribir²⁴

$$\mathcal{D}_{1,t} = \sum_{l=0}^{L-1} h_l W_{1,t+l \bmod N}^\uparrow = \sum_{l=0}^{N-1} h_l^\circ W_{1,t+l \bmod N}^\uparrow, \quad t = 0, 1, \dots, N-1$$

De aquí es posible ver que los elementos de \mathcal{D}_1 se obtienen *cross-correlating* el filtro $\{h_l^\circ\}$ con la versión *upsampled* de \mathbf{W}_1 . Como $\{h_l^\circ\} \longleftrightarrow \{H(\frac{k}{N})\}$ una interpretación alternativa es que los elementos de \mathcal{D}_1 son el resultado de filtrar la versión *upsampleada* de \mathbf{W}_1 con un filtro circular cuya representación en el dominio de frecuencia es $\{H^*(\frac{k}{N})\}$.²⁵

²⁴Percival y Walden (2000) p.81-82

²⁵Idem

De manera similar, los elementos de S_1 puede ser generado filtrando una versión *upsamplada* de V_1 con el filtro circular $\{G^*(\frac{k}{N})\}$. Si definimos $V_{1,t}^\dagger$ en una manera similar a $W_{1,t}^\dagger$ podemos escribir²⁶

$$\begin{aligned} X_t &= \sum_{l=0}^{L-1} h_l W_{1,t+l \bmod N}^\dagger + \sum_{l=0}^{L-1} g_l V_{1,t+l \bmod N}^\dagger \\ &= \sum_{l=0}^{N-1} h_l^\circ W_{1,t+l \bmod N}^\dagger + \sum_{l=0}^{N-1} g_l^\circ V_{1,t+l \bmod N}^\dagger, \quad t = 0, 1, \dots, N-1 \end{aligned} \quad (3.31)$$

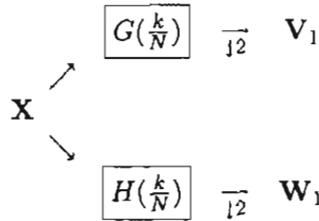
esta ecuación nos muestra como X puede sintetizarse mediante las versiones *upsampled* de las series W_1 y V_1 respectivamente. Ahora recordemos que para definir los primeros $N/2$ coeficientes de la transformada en onduletas, lo hacíamos mediante el filtro de onduletta $\{h_l\}$ y posteriormente se realizaba un *downsampling* por dos de la serie resultante. Percival y Walden (2000) muestran que mientras $\widetilde{W}_{1,t}$ es una serie *half-band* que adolesce de elementos de baja frecuencia (por que es lo que resulta de aplicar un filtro *high-pass*), la serie submuestreada es *full-band*, es decir, puede tener coeficientes de Fourier significativos sobre todas las frecuencias. Así mismo muestran también que mientras que la serie $\widetilde{V}_{1,t}$ puede ser llamada una serie de *half-band* por que es el resultado de un filtro *low-pass* (i.e. $\{g_l\}$), con paso nominal $[-1/4, 1/4]$ y que entonces adolesce de elementos de alta frecuencia, la serie submuestreada $V_{1,t}$ es una serie *full-band* por que puede tener coeficientes de Fourier significativos sobre todas las frecuencias en $[-1/2, 1/2]$.²⁷

En síntesis la primera etapa del algoritmo pirámide toma una serie *full-band* $\{X_t\}$ de longitud N y las descompone ortogonalmente en dos nuevas series *half-band*, a saber, los coeficientes de escala y de onduletas de primer nivel $\{V_{1,t}\}$ y $\{W_{1,t}\}$, cada uno con longitud $N/2$. Los coeficientes de escala capturan aproximadamente el contenido de baja frecuencia de $\{X_t\}$, mientras que los coeficientes de onduletas capturan el contenido de alta frecuencia (aunque con orden reverso en las frecuencias). Así mismo, la serie $\{V_{1,t}\}$ es proporcional a promedios ponderados en la escala de *dos* de $\{X_t\}$, mientras que $\{W_{1,t}\}$ está asociado con cambios en promedios ponderados en la escala de *uno*. Entonces las escalas las escalas asociadas con los resultados de los filtros de onduletas y de escala difieren por un factor de dos. Para recordar esta distinción en etapas posteriores del algoritmo pirámide, utilizamos $\lambda_j \equiv 2^j$ para denotar la escala asociada al resultado del filtro de

²⁶Percival y Walden (2000) p.83

²⁷Ibid. p. 83-85

escala, mientras que utilizaremos $\tau_j \equiv 2^{j-1}$ para denotar la escala asociada con el filtro de onduletas (en la primera etapa $\lambda_1 = 2$ mientras que $\tau_1 = 2$). El gráfico siguiente ilustra el proceso de síntesis hasta ahora descrito.



Recapitulemos enseguida la propiedades de los coeficientes resultantes de esta primera etapa. Los coeficientes de onduletas de primer nivel $\{\mathbf{W}_{1,t} : t = 0, \dots, \frac{N}{2} - 1\}$ están asociados con:

1. \mathcal{W}_1 , una matriz de $\frac{N}{2} \times N$ que satisface que $\mathcal{W}_1 \mathcal{W}_1^T = I_{N/2}$ y consiste de las $N/2$ primeras filas de la matriz de la DWT \mathcal{W} (cada fila de \mathcal{W}_1 contiene los elementos de $\{h_l^\circ\} \leftrightarrow \{H(\frac{k}{N})\}$, i.e. el filtro de onduletas $\{h_l\}$ periodizado a la longitud N ; la primera fila es $[h_1^\circ, h_0^\circ, h_{N-1}^\circ, h_{N-2}^\circ, \dots, h_2^\circ]$ y las $\frac{N}{2} - 1$ filas restantes se forman trasladando circularmente la primera fila hacia la derecha, en unidades de $2, 4, 6, \dots, N - 2$ respectivamente;)
2. Los elementos del vector $\mathbf{W}_1 = \mathcal{W}_1 \mathbf{X}$ de longitud $N/2$;
3. El primer nivel de detalle $\mathcal{D}_1 = \mathcal{W}_1^T \mathbf{W}_1 = \mathcal{W}_1^T \mathcal{W}_1 \mathbf{X}$, que puede ser formado por *upsampling* por dos de \mathbf{W}_1 y luego filtrando con $\{H^*(\frac{k}{N})\}$;
4. Cambios en los promedio de la escala $\tau_1 = 1$;
5. Submuestreo por dos de un proceso *half-band* asociado con las altas frecuencias $[1/4, 1/2]$ en la serie de tiempo $\{X_t\}$; y
6. La porción de la DFT de $\{X_t\}$ con frecuencias en el intervalo $[1/4, 1/2]$, con un orden invertido respecto a las frecuencias.

En constraste, los coeficientes de escala de primer nivel $\{\mathbf{V}_{1,t} : t = 0, \dots, \frac{N}{2} - 1\}$ están asociados con:

1. \mathcal{V}_1 , una matriz de $\frac{N}{2} \times N$ que satisface que $\mathcal{V}_1 \mathcal{V}_1^T = I_{N/2}$ y $\mathcal{W}_1 \mathcal{V}_1^T = \mathcal{V}_1 \mathcal{W}_1^T = 0_{N/2}$ y cuyas filas general el mismo subespacio de \mathbb{R}^N que es generado por las $N/2$ primeras filas de la matriz de la DWT \mathcal{W} (\mathcal{V}_1 tiene la misma estructura que \mathcal{W}_1 : sólo es necesario reemplazar a h_l°

con g_l° , donde $\{g_l^\circ\}$ es el filtro de escala $\{g_l\}$ periodizado a la longitud N , i.e. $\{g_l^\circ\} \leftrightarrow \{G^*(\frac{k}{N})\}$;

2. Los elementos del vector $\mathbf{V}_1 = \mathcal{V}_1 \mathbf{X}$ de longitud $N/2$;
3. El primer nivel de *smooth* $\mathcal{S}_1 = \mathcal{V}_1^T \mathbf{V}_1 = \mathcal{V}_1^T \mathcal{V}_1 \mathbf{X}$, que puede ser formado por *upsampling* por dos de \mathbf{V}_1 y luego filtrando con $\{G^*(\frac{k}{N})\}$;
4. Cambios en los promedio de la escala $\lambda_1 = 2$;
5. Submuestreo por dos de un proceso *half-band* asociado con las bajas frecuencias $[0, 1/4]$ en la serie de tiempo $\{X_t\}$; y
6. La porción de la DFT de $\{X_t\}$ con frecuencias en el intervalo $[0, 1/4]$.

La serie $\{W_{1,t}\}$ constituye la primera mitad del vector de coeficientes de onduletas \mathbf{W} ; los coeficientes restantes son obtenidos a partir de $\{V_{1,t}\}$ en las etapas sucesivas del algoritmo.

La segunda etapa

Como se describió anteriormente, la segunda etapa del algoritmo pirámide consiste en tratar a $\{V_{1,t}\}$ de la misma manera que $\{X_t\}$ fue tratado en la primera etapa. Intuitivamente este es un modo razonable de proceder por que $\{X_t\}$ y $\{V_{1,t}\}$ son similares en el sentido que $\{X_t\}$ puede ser visto como un promedio en la escala unitaria, mientras que los coeficientes de de escala de primer nivel (i.e. $\{V_{1,t}\}$) pueden ser vistos como promedio en la escala de dos. Entonces filtrando circularmente y por separado a $\{V_{1,t}\}$ con $\{h_l\}$ y $\{g_l\}$ y submuestreándola para producir dos nuevas series tenemos

$$W_{2,t} \equiv \sum_{l=0}^{L-1} h_l V_{1,2t+1-l \bmod \frac{N}{2}} \quad \text{y} \quad V_{2,t} \equiv \sum_{l=0}^{L-1} g_l V_{1,2t+1-l \bmod \frac{N}{2}} \quad (3.32)$$

para $t = 0, \dots, \frac{N}{4} - 1$. Notemos que la ecuación anterior es muy similar a las ecuaciones (3.12) y (3.26), excepto que en este caso se ha reemplazado $\bmod N$ por $\bmod \frac{N}{2}$. El coeficiente de onduleta de la DWT para la escala de dos y para el nivel $j = 2$ están dados por las $W_{2,t}$'s, i.e.

$$\begin{aligned} \mathbf{W}_2 &\equiv \left[W_{\frac{N}{2}}, W_{\frac{N}{2}+1}, \dots, W_{\frac{3N}{4}-1} \right]^T \\ &= \left[W_{2,0}, W_{2,1}, \dots, W_{2, \frac{N}{4}-1} \right]^T \end{aligned}$$

Definamos un vector que contenga los coeficientes de escala para el nivel $j = 2$ como:

$$\mathbf{V}_2 = \left[V_{2,0}, V_{2,1}, \dots, V_{2, \frac{N}{4}-1} \right]^T$$

Sean \mathcal{B}_2 y \mathcal{A}_2 dos matrices de $\frac{N}{4} \times \frac{N}{4}$ cuyas filas consisten en versiones trasladadas circularmente de $\{h_l\}$ y $\{g_l\}$ periodizadas para tener longitud $N/2$ de tal manera que podemos expresar a la transformación en función de \mathbf{V}_1 y \mathbf{W}_2 como:

$$\begin{bmatrix} \mathbf{W}_2 \\ \mathbf{V}_2 \end{bmatrix} = \mathcal{P}_2 \mathbf{V}_1 = \begin{bmatrix} \mathcal{B}_2 \\ \mathcal{A}_2 \end{bmatrix} \mathbf{V}_1 \quad \text{con} \quad \mathcal{P}_2 \equiv \begin{bmatrix} \mathcal{B}_2 \\ \mathcal{A}_2 \end{bmatrix}$$

Al igual que con \mathcal{W}_1 y \mathcal{V}_1 , las filas de \mathcal{B}_2 y de \mathcal{A}_2 son ortonormales:

$$\mathcal{B}_2 \mathcal{A}_2^T = \mathcal{A}_2 \mathcal{B}_2^T = 0_{\frac{N}{4}} \quad \text{y} \quad \mathcal{B}_2 \mathcal{B}_2^T = \mathcal{A}_2 \mathcal{A}_2^T = I_{\frac{N}{4}}$$

Entonces la matriz \mathcal{P}_2 de $\frac{N}{2} \times \frac{N}{2}$ es ortonormal y por tanto podemos recuperar a \mathbf{V}_1 utilizando la siguiente ecuación

$$\mathbf{V}_1 = \mathcal{P}_2^T \begin{bmatrix} \mathbf{W}_2 \\ \mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathcal{B}_2^T & \mathcal{A}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{W}_2 \\ \mathbf{V}_2 \end{bmatrix} = \mathcal{B}_2^T \mathbf{W}_2 + \mathcal{A}_2^T \mathbf{V}_2$$

sustituyen la ecuación anterior en (3.30) tenemos

$$\mathbf{X} = \mathcal{W}_1^T \mathbf{W}_1 + \mathcal{V}_1^T \mathcal{B}_2^T \mathbf{W}_2 + \mathcal{V}_1^T \mathcal{A}_2^T \mathbf{V}_2 = \mathcal{B}_1^T \mathbf{W}_1 + \mathcal{A}_1^T \mathcal{B}_2^T \mathbf{W}_2 + \mathcal{A}_1^T \mathcal{A}_2^T \mathbf{V}_2$$

donde $\mathcal{A}_1 \equiv \mathcal{V}_1$ y $\mathcal{B}_2 \equiv \mathcal{W}_1$, esta notación no será de mucha utilidad en la descripción del algoritmo en su etapa más general. Comparando la ecuación (1.17) tenemos que si $\mathcal{W}_2 \equiv \mathcal{B}_2 \mathcal{A}_1$ entonces

$$\mathcal{D}_2 = \mathcal{W}_2^T \mathbf{W}_2 = \mathcal{A}_1^T \mathcal{B}_2^T \mathbf{W}_2$$

como $\mathcal{D}_1 = \mathcal{B}_1^T \mathbf{W}_1$ tenemos que

$$\mathbf{X} = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{A}_1^T \mathcal{A}_2^T \mathbf{V}_2 \quad \text{y} \quad \mathbf{X} = \mathcal{D}_1 + \mathcal{D}_2 + \mathcal{S}_2$$

se sigue que:

$$\mathcal{S}_2 = \mathcal{A}_1^T \mathcal{A}_2^T \mathbf{V}_2 \equiv \mathcal{V}_2^T \mathbf{V}_2$$

donde $\mathcal{V}_2 \equiv \mathcal{A}_2 \mathcal{A}_1$.

Debido a que $\{W_{2,t}\}$ y $\{V_{2,t}\}$ son obtenidos filtrando $\{V_{1,t}\}$ y submuestreando y como $\{V_{1,t}\}$ es obtenido filtrando $\{X_t\}$ y submuestreando, podemos

usar la noción de un filtro cascada²⁸ para describir como $\{W_{2,t}\}$ y $\{V_{2,t}\}$ puede ser obtenido filtrando directamente $\{X_t\}$, si podemos, de alguna manera, tener en cuenta la operación de submuestreo. Para hacerlo definamos $\{h_l^\uparrow\}$ como el filtro amplitud $2L-1$ con coeficientes $h_0, 0, h_1, 0, \dots, h_{L-2}, 0, h_{L-1}$; i.e. formamos $\{h_l^\uparrow\}$ insertando un cero entre cada uno de los L elementos de $\{h_l\}$. Sea

$$2\widetilde{W}_{2,t} \equiv \sum_{l=0}^{2L-1} h_l^\uparrow 2^{1/2} \widetilde{V}_{1,t-l \bmod N} \quad t = 0, \dots, N-1 \quad (3.33)$$

la ecuación (3.26) no dice que $\{2^{1/2}\widetilde{W}_{1,t}\}$ se obtiene filtrando $\{X_t\}$ con $\{g_l\}$, entonces $\{2\widetilde{W}_{2,t}\}$ es el resultado de aplicar un filtro en cascada a $\{X_t\}$, quien es sujeto a filtración primero por $\{g_l\}$ y luego por $\{h_l^\uparrow\}$. Afirmamos que

$$W_{2,t} = 2\widetilde{W}_{2,4t+3}, \quad t = 0, \dots, \frac{N}{4} - 1 \quad (3.34)$$

i.e. que $\{W_{2,t}\}$ puede ser formado tomado cada cuarto valor del $\{2\widetilde{w}_{2,t}\}$ comenzando con $2\widetilde{W}_{2,3}$ (en la terminología de ingeniería $\{W_{2,t}\}$ es el resultado del *downsampling* de $\{2\widetilde{W}_{2,t}\}$ por cuatro) ²⁹.

Ahora vemos como se construyen los renglones de la matriz de la DWT correspondientes a esta etapa, i.e. W_2 . Sea $\{h_{2,l}\} \equiv \{g * h_l^\uparrow\}$ la convolución de $\{g_l\}$ con $\{h_l^\uparrow\}$. Como $\{g_l\}$ tiene ancho L mientras que $\{h_l^\uparrow\}$ tiene ancho $2L-1$, entonces el ancho de $\{h_{2,l}\}$ es ³⁰

$$L_2 \equiv L + (2L-1) - 1 = 3L-2$$

Ahora supongamos que $\{h_l : l = 0, \dots, L-1\}$ es un filtro con función de transferencia $H(\cdot)$. Definamos un nuevo filtro insertando m ceros entre cada uno de los elementos de $\{h_l\}$

$$\underbrace{h_0, 0, \dots, 0}_m \text{ CEROS}, h_1, \dots, h_{L-2}, \underbrace{0, \dots, 0}_m \text{ CEROS}, h_{L-1}$$

²⁸Un filtro cascada es un "arreglo" de un conjunto de M filtros de tal manera que el resultado del primer filtro es el insumo para el segundo filtro y así sucesivamente.

²⁹Se ofrece una prueba en Percival y Walden (2000) p.90

³⁰Como ya se había anticipado el filtro $\{h_{2,l}\}$ puede ser visto como un filtro cascada y para este tipo de filtros se cumple que su ancho es igual a la suma de los anchos de los filtros "simples" que lo constituyen menos la cantidad de estos filtros mas uno. (Véase una prueba de este hecho en Percival y Walden (2000) p.28, 503-504.)

entonces este filtro tiene función de transferencia $H((m+1)f)$. Ahora aplicando este resultado para el caso en que $m = 1$, podemos concluir que la función de transferencia de $\{h_{2,l}\}$ está dada por

$$H_2(f) \equiv H(2f)G(f)$$

En una manera similar podemos argumentar que el filtro equivalente digamos $\{g_{2,l}\}$, que relaciona $\{V_{2,t}\}$ con $\{X_t\}$ es el filtro de amplitud $L - 2$ obtenido convolucionando $\{g_l\}$ con

$$\{g_l^{\uparrow}\} \equiv \{g_0, 0, g_1, 0, \dots, g_{L-2}, 0, g_{L-1}\}$$

y cuya función de transferencia está dada por ³¹

$$G_2(f) \equiv G(2f)G(f)$$

$$\mathcal{H}_2(f) \equiv |H_2(f)|^2 \quad \text{y} \quad \mathcal{G}_2(f) \equiv |G_2(f)|^2$$

las igualdades anteriores nos permiten calcular a las funciones de ganancia cuadrada para los filtros de la segunda etapa. Se presentan las gráficas para los casos de Haar y Daubechies D(4).

Las gráficas anteriores ilustran el hecho de que los filtros onduletas de la segunda etapa son equivalentes a un filtro cuyo paso de banda es $1/8 \leq |f| \leq 1/4$, mientras que el filtro de escala es una "aproximación" a un filtro *low-pass* cuyo paso de banda es $0 \leq |f| \leq 1/8$.³²

Por conveniencia llamaremos $\{h_{2,l}\}$ y $\{g_{2,l}\}$ a los filtros de onduletas y de escala de segundo nivel por que ellos producen los coeficientes de onduletas y de escala de segundo nivel. En la siguiente sección generalizamos esta noción para los j -ésimos filtros de onduletas y de escala.³³ Es también conveniente definir los filtros de onduletas y de escala de primer nivel como $\{h_{1,l}\}$ y $\{g_{1,l}\}$ idénticos a $\{h_l\}$ y $\{g_l\}$ respectivamente.

Es necesario recalcar que aunque el filtro $\{h_{2,l}\}$ suma cero y tiene energía unitaria, no satisface la propiedad definitoria de un filtro de onduleta, a saber la ortogonalidad con respecto a sus traslaciones pares, en efecto, el filtro de onduleta de segundo nivel no es necesariamente un filtro de onduleta. Una situación similar ocurre con el filtro de escala del segundo nivel.

Como los elementos de \mathbf{W}_2 son obtenidos filtrando circularmente $\{X_t\}$ con $\{h_{2,l} : l = 0, \dots, L_2 - 1\}$ también pueden ser obtenidos filtrando circularmente con la versión periodizada de $\{h_{2,l}\}$, al cual denotaremos como

³¹Percival y Walden (2000) p.91

³²Idem

³³Ibid. p. 91-92, 515-516.

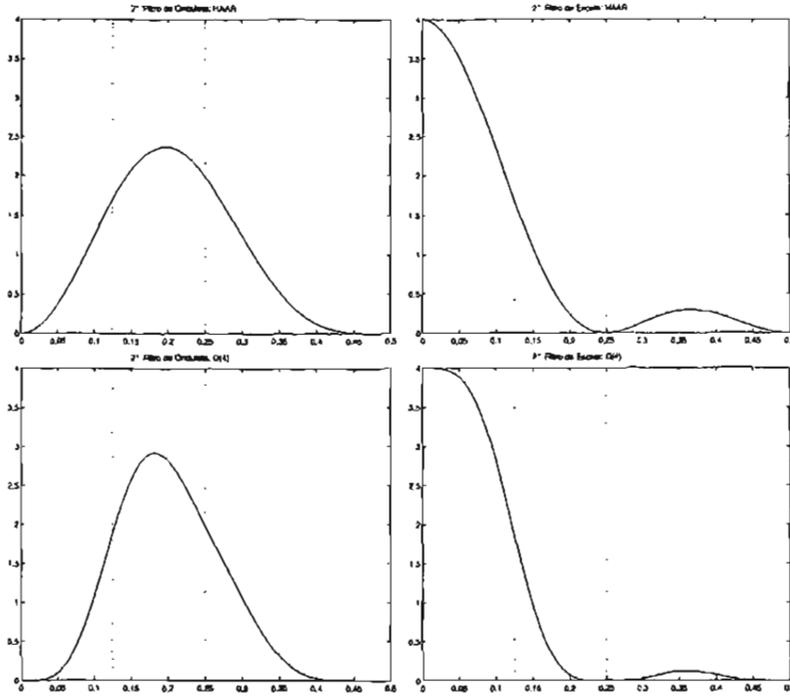


Figura 3.11: Funciones de ganancia cuadrada para los filtros de la segunda etapa: filtros de ondulatas(izquierda) y escala (derecha), para Haar (arriba) y Daubechies D(4) (abajo).

$\{h_{2,l}^{\circ} : l = 0, \dots, N-1\}$, i.e. tenemos que

$$W_{2,t} = \sum_{l=0}^{N-1} h_{2,l}^{\circ} X_{4(t+1)-1+l \bmod N} \quad t = 0, \dots, \frac{N}{4} - 1$$

Como también tenemos que $\mathbf{W}_2 = \mathcal{W}_2 \mathbf{X}$, las filas de \mathcal{W}_2 tiene que contener versiones trasladadas circularmente de $\{h_{2,l}^{\circ}\}$. Si notamos que el elemento cero de \mathbf{W}_2 es

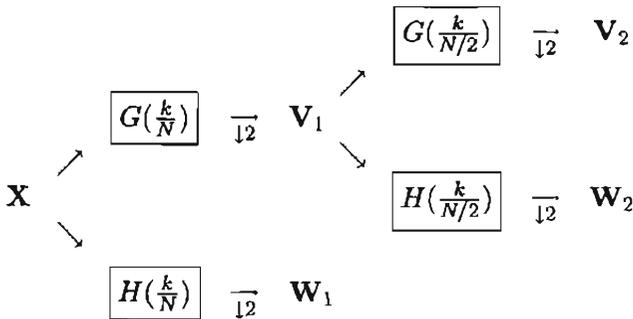
$$W_{2,0} = \sum_{l=0}^{N-1} h_{2,l}^{\circ} X_{3-l \bmod N}$$

es evidente que la primera fila de \mathcal{W}_2 (o equivalentemente, la fila $N/2$ de \mathcal{W}) está dada por:

$$[h_{2,3}^{\circ}, h_{2,2}^{\circ}, h_{2,1}^{\circ}, h_{2,0}^{\circ}, h_{2,N-1}^{\circ}, h_{2,N-2}^{\circ}, \dots, h_{2,5}^{\circ}, h_{2,4}^{\circ}] = \mathcal{W}_{\frac{N}{2}}^T. \quad (3.35)$$

las $\frac{N}{4} - 1$ filas restantes están dadas por $\mathcal{T}^{4k} \mathcal{W}_{\frac{N}{2}}^T$, $k = 1, \dots, \frac{N}{4} - 1$, es decir trasladando circularmente las primeras filas en múltiplos de cuatro.³⁴ También sabemos que la transformada de Fourier (DFT) de $\{h_{2,t}^o\}$ está dada por el submuestreo de la función de transferencia de $\{h_{2,t}\}$ en las frecuencias $f_k = k/N$, $k = 0, \dots, N - 1$. Como la función de transferencia está definida por $H(2f)G(f)$ tenemos que $\{h_{2,t}^o\} \leftrightarrow \{H(\frac{k}{N/2})G(\frac{k}{N})\}$. De manera similar el filtro periodizado $\{g_{2,t}^o\} \leftrightarrow \{G(\frac{k}{N/2})G(\frac{k}{N})\}$ provee de las filas de \mathcal{V}_2 .

En síntesis, en la segunda etapa del algoritmo pirámide se rotan los $N/2$ vectores base en \mathcal{V}_1 en dos conjuntos de $N/4$ vectores, a saber, $\mathcal{W}_2 = \mathcal{B}_2 \mathcal{V}_1 = \mathcal{B}_2 \mathcal{A}_1$ y $\mathcal{V}_2 = \mathcal{A}_2 \mathcal{V}_1 = \mathcal{A}_2 \mathcal{A}_1$, donde hemos definido $\mathcal{A}_1 = \mathcal{V}_1$. Cada fila de \mathcal{B}_2 contiene los elementos del filtro de ondulatas $\{h_t\}$ periodizado a longitud $N/2$ (i.e. la transformada de Fourier inversa de $\{H(\frac{k}{N/2}), k = 0, \dots, \frac{N}{2} - 1\}$), con filas adyacentes difiriendo por una traslación circular de 2 unidades; así mismo, cada fila de \mathcal{A}_2 contiene al filtro de escala $\{g_t\}$ periodizado a longitud $N/2$. Esta etapa transforma $\{V_{1,t} : t = 0, \dots, \frac{N}{2} - 1\}$ en dos nuevas series, a saber, $\{W_{2,t} : t = 0, \dots, \frac{N}{4} - 1\}$ y $\{V_{2,t} : t = 0, \dots, \frac{N}{4} - 1\}$. Adicionalmente, justo como $\{X_t\}$ fue particionado ortogonalmente en un componente *high-pass* $\{W_{1,t}\}$ y un componente *low-pass* $\{V_{1,t}\}$ por los filtros de ondulatas y escala; de mismo modo $\{V_{1,t}\}$ es particionado en $\{W_{2,t}\}$ y $\{V_{2,t}\}$; sin embargo la partición en *high-pass* y *low-pass* de $\{V_{1,t}\}$ es equivalente a la partición de la parte *low-pass* de $\{X_t\}$ en dos, una que corresponde a $1/8 \leq |f| \leq 1/4$ (asociado con los coeficientes de ondulatas en la escala de dos) y la otra que corresponde a $0 \leq |f| \leq 1/8$. A continuación se muestra un diagrama que ilustra el proceso de síntesis de la señal hasta la segunda etapa.



Recapitulemos las propiedades de los coeficientes resultantes de la segunda etapa de algoritmo. Los coeficientes de ondulatas del segundo nivel $\{W_{2,t}\}$ están asociados con:

³⁴Ibid. p.92

1. $\mathcal{W}_2 = \mathcal{B}_2 \mathcal{A}_1$ una matriz de $\frac{N}{4} \times N$ que satisface que $\mathcal{W}_2 \mathcal{W}_2^T = I_{\frac{N}{4}}$ y que comprende las filas $\frac{N}{2}$ hasta $\frac{3N}{4} - 1$ de la matriz de la DWT \mathcal{W} , cada fila de \mathcal{W}_2 conteniendo $\{h_{2,l}^\circ\} \longleftrightarrow \{H(\frac{k}{N/2})G(\frac{k}{N})\}$, i.e. el filtro $\{h_{2,l}\}$ periodizado a longitud N , la primera fila se muestra en la ecuación (3.35) y las restantes $\frac{N}{4} - 1$ filas se forman mediante traslaciones circulares de la primera en las unidades de 4, 8, 12, ..., $N - 4$ respectivamente.
2. Los elementos del vector $\mathbf{W}_2 = \mathcal{W}_2 \mathbf{X}$ de longitud $N/4$,
3. El segundo nivel detalle $\mathcal{D}_2 = \mathcal{W}_2^T \mathbf{W}_2 = \mathcal{W}_2^T \mathcal{W}_2 \mathbf{X}$;
4. Cambio en promedios en la escala de $\tau_2 = 2$;
5. Submuestreo por dos de un proceso *half-band* asociado con altas frecuencias $[1/4, 1/2]$, en los coeficientes de escala de primer nivel $\{V_{1,t}\}$, y
6. La porción de la DFT de $\{X_t\}$ con frecuencias en el intervalo $[1/8, 1/4]$, en orden reverso con respecto a la frecuencia.

Por comparación los coeficientes de escala del segundo nivel $\{V_{2,t}\}$ están asociados con:

1. $\mathcal{V}_2 = \mathcal{A}_2 \mathcal{A}_1$ una matriz de $\frac{N}{4} \times N$ que satisface tanto $\mathcal{V}_2 \mathcal{V}_2^T = I_{N/4}$ como $\mathcal{W}_2 \mathcal{V}_2^T = \mathcal{V}_2 \mathcal{W}_2^T = 0_{N/4}$, cuyas filas generan el mismo subespacio de \mathbb{R}^N que la últimas $N/4$ filas de la matriz de la DWT \mathcal{W} , (\mathcal{V}_2 tiene una construcción similar a \mathcal{W}_2 : sólo es necesario reemplazar cada $h_{2,l}^\circ$ por $g_{2,l}^\circ$, donde $\{g_{2,l}^\circ\}$ es $g_{2,l}$ periodizado a longitud N , i.e. $\{g_{2,l}^\circ\} \longleftrightarrow \{G(\frac{k}{N/2})G(\frac{k}{N})\}$ (la sucesión $\{g_{j,l}^\circ\}$ aparece en orden invertido en las filas de la matriz \mathcal{V}_j)
2. Los elementos del vector $\mathbf{V}_2 = \mathcal{V}_2 \mathbf{X}$ de longitud $N/4$,
3. El segundo nivel de *smooth* $\mathcal{S}_2 = \mathcal{V}_2^T \mathbf{V}_2 = \mathcal{V}_2^T \mathcal{V}_2 \mathbf{X}$;
4. Promedios en la escala de $\lambda_2 = 4$;
5. Submuestreo por dos de un proceso *half-band* asociado con bajas frecuencias $[0, 1/4]$, en los coeficientes de escala de primer nivel $\{V_{1,t}\}$, y
6. La porción de la DFT de $\{X_t\}$ con frecuencias en el intervalo $[0, 1/8]$.

La etapa general

Dadas las discusiones anteriores, podemos establecer fácilmente la j -ésima etapa del algoritmo pirámide, donde $j = 1, \dots, J$ (recordemos que el tamaño de la muestra asumido es N y es igual a 2^J). Con $\{V_{0,t}\}$ definido como $\{X_t\}$, el j -ésimo input es $\{V_{j-1,t} : t = 0, \dots, N_{j-1}-1\}$, donde $N_j \equiv N/2^j$. Este input es el coeficiente de escala asociado con promedio en la escala de $\lambda_{j-1} = 2^{j-1}$. Los j -ésimos outputs son los coeficientes del j -ésimo nivel de onduletas y de escalas:

$$W_{j,t} = \sum_{l=0}^{L-1} h_l V_{j-1,2t+1-l \bmod N_{j-1}}, \quad V_{j,t} = \sum_{l=0}^{L-1} g_l V_{j-1,2t+1-l \bmod N_{j-1}} \quad (3.36)$$

$t = 0, \dots, N_j - 1$. Los coeficientes de onduletas para la escala $\tau_j = 2^{j-1}$ están dados por los $W_{j,t}$'s,³⁵ i.e.

$$\mathbf{W}_j \equiv [W_{N-N_{j-1}}, W_{N-N_{j-1}+1}, \dots, W_{N-N_{j-1}}]^T = [W_{j,0}, W_{j,1}, \dots, W_{j,N_j-1}]^T$$

Si definimos \mathbf{V}_j^T similarmente para contener los $V_{j,t}$'s entonces la transformación de \mathbf{V}_{j-1} en \mathbf{W}_j y \mathbf{V}_j puede ser expresada como:

$$\begin{bmatrix} \mathbf{W}_j \\ \mathbf{V}_j \end{bmatrix} = \mathcal{P}_j \mathbf{V}_{j-1} = \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix} \mathbf{V}_{j-1} \quad \text{con} \quad \mathcal{P}_j \equiv \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix}, \quad (3.37)$$

donde \mathcal{B}_j y \mathcal{A}_j son matrices de $N_j \times N_{j-1}$ cuyas columnas contienen, respectivamente, versiones trasladadas circularmente de los filtros de onduletas y de escala $\{h_l\}$ y $\{g_l\}$ periodizados con longitud N_{j-1} . La matriz \mathcal{P}_j de $N_{j-1} \times N_{j-1}$ es ortonormal, entonces

$$\mathcal{P}_j \mathcal{P}_j^T = \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix} \begin{bmatrix} \mathcal{B}_j^T & \mathcal{A}_j^T \end{bmatrix} = \begin{bmatrix} \mathcal{B}_j \mathcal{B}_j^T & \mathcal{B}_j \mathcal{A}_j^T \\ \mathcal{A}_j \mathcal{B}_j^T & \mathcal{A}_j \mathcal{A}_j^T \end{bmatrix} = \begin{bmatrix} I_{N_j} & 0_{N_j} \\ 0_{N_j} & I_{N_j} \end{bmatrix} = I_{N_{j-1}}$$

y

$$\mathcal{P}_j^T \mathcal{P}_j = \begin{bmatrix} \mathcal{B}_j^T & \mathcal{A}_j^T \end{bmatrix} \begin{bmatrix} \mathcal{B}_j \\ \mathcal{A}_j \end{bmatrix} = \mathcal{B}_j^T \mathcal{B}_j + \mathcal{A}_j^T \mathcal{A}_j = I_{N_{j-1}}$$

Recordando que $\mathbf{V}_1 = \mathcal{A}_1 \mathbf{X}$ aplicando recursivamente $\mathbf{V}_j = \mathcal{A}_j \mathbf{V}_{j-1}$ tenemos que

$$\mathbf{W}_j = \mathcal{B}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 \mathbf{X} = \mathcal{W}_j \mathbf{X} \quad \text{y} \quad \mathbf{V}_j = \mathcal{A}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 \mathbf{X} = \mathcal{V}_j \mathbf{X},$$

³⁵Percival y Walden (2000) p.94

donde

$$\mathbf{W}_j \equiv \mathcal{B}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1 \quad \text{y} \quad \mathbf{V}_j = \mathcal{A}_j \mathcal{A}_{j-1} \cdots \mathcal{A}_1$$

como la matriz \mathcal{P}_j de $N_{j-1} \times N_{j-1}$ es ortonormal, podemos recuperar \mathbf{V}_{j-1} utilizando la ecuación (3.37)

$$\mathbf{V}_{j-1} = \mathcal{P}_j^T \begin{bmatrix} \mathbf{W}_j \\ \mathbf{V}_j \end{bmatrix} = \begin{bmatrix} \mathcal{B}_j^T & \mathcal{A}_j^T \end{bmatrix} \begin{bmatrix} \mathbf{W}_j \\ \mathbf{V}_j \end{bmatrix} = \mathcal{B}_j^T \mathbf{W}_j + \mathcal{A}_j^T \mathbf{V}_j \quad (3.38)$$

En términos de operaciones con filtros el t -ésimo elemento de \mathbf{V}_{j-1} puede ser reconstruido vía un análogo a la ecuación (3.31)

$$\mathbf{V}_{j-1,t} = \sum_{l=0}^{L-1} h_l W_{j,t+l \bmod N_{j-1}}^\uparrow + \sum_{l=0}^{L-1} g_l V_{j,t+l \bmod N_{j-1}}^\uparrow, \quad (3.39)$$

para $t = 0, \dots, N_{j-1} - 1$ donde

$$W_{j,t+l}^\uparrow \equiv \begin{cases} 0, & t = 0, 2, \dots, N_{j-1} - 2 \\ W_{j, \frac{t-1}{2}} & t = 1, 3, \dots, N_{j-1} - 1 \end{cases}$$

(se define en manera similar $V_{j,t+l}^\uparrow$).³⁶

Aplicando recursivamente las ecuaciones (3.30) y (3.38) tenemos³⁷

$$\mathbf{X} = \mathcal{B}_1^T \mathbf{W}_1 + \mathcal{A}_1^T \mathcal{B}_2^T \mathbf{W}_2 + \mathcal{A}_1^T \mathcal{A}_2^T \mathcal{B}_3^T \mathbf{W}_3 + \cdots + \mathcal{A}_1^T \cdots \mathcal{A}_{j-1}^T \mathcal{B}_j^T \mathbf{W}_j + \mathcal{A}_1^T \cdots \mathcal{A}_{j-1}^T \mathcal{A}_j^T \mathbf{V}_j$$

Los detalles \mathcal{D}_j y los *smooths* \mathcal{S}_j están definidos por los

$$\mathcal{D}_j = \mathcal{A}_1^T \cdots \mathcal{A}_{j-1}^T \mathcal{B}_j^T \mathbf{W}_j = \mathcal{W}_j^T \mathbf{W}_j \quad (3.40)$$

$$\mathcal{S}_j = \mathcal{A}_1^T \cdots \mathcal{A}_{j-1}^T \mathcal{A}_j^T \mathbf{V}_j = \mathcal{V}_j^T \mathbf{V}_j \quad (3.41)$$

La ortonormalidad de \mathcal{P}_j nos dice que

$$\|\mathbf{V}_{j-1}\|^2 = \|\mathbf{W}_j\|^2 + \|\mathbf{V}_j\|^2$$

como $\mathbf{V}_0 = \mathbf{X}$ la aplicación recursiva de lo anterior nos dice que:

$$\|\mathbf{X}\|^2 = \sum_{k=1}^j \|\mathbf{W}_k\|^2 + \|\mathbf{V}_j\|^2$$

³⁶Percival y Walden (2000) p. 95

³⁷Idem.

usando las ecuaciones anteriores tenemos que

$$\|\mathbf{X}\|^2 = \sum_{k=1}^j \|\mathcal{D}_k\|^2 + \|\mathcal{S}_j\|^2$$

El filtro equivalente $\{h_{j,l}\}$ que relaciona $\{W_{j,t}\}$ con $\{X_t\}$ está formado por la convolución de los siguientes j filtros:

$$\begin{aligned} \text{filtro 1 : } & g_0, g_1, \dots, g_{L-2}, g_{L-1}; \\ \text{filtro 2 : } & g_0, 0, g_1, 0, \dots, g_{L-2}, 0, g_{L-1}; \\ \text{filtro 3 : } & g_0, 0, 0, 0, g_1, \dots, g_{L-2}, 0, 0, 0, g_{L-1}; \\ & \vdots \end{aligned} \quad (3.42)$$

$$\begin{aligned} \text{filtro } j-1 : & g_0, \underbrace{0, \dots, 0}_{2^{j-2}-1 \text{ ceros}}, g_1, \underbrace{0, \dots, 0}_{2^{j-2}-1 \text{ ceros}}, \dots, g_{L-2}, \underbrace{0, \dots, 0}_{2^{j-2}-1 \text{ ceros}}, g_{L-1}; \\ \text{filtro } j : & h_0, \underbrace{0, \dots, 0}_{2^{j-1}-1 \text{ ceros}}, h_1, \underbrace{0, \dots, 0}_{2^{j-1}-1 \text{ ceros}}, \dots, h_{L-2}, \underbrace{0, \dots, 0}_{2^{j-1}-1 \text{ ceros}}, h_{L-1}; \end{aligned} \quad (3.43)$$

El filtro $\{h_{j,l}\}$ tiene ancho ³⁸

$$L_j \equiv (2^{j-1})(L-1) + 1 \quad (3.44)$$

y función de transferencia dada por³⁹:

$$H_j(f) \equiv H(2^{j-1}f) \prod_{l=0}^{j-2} G(2^l f) \quad (3.45)$$

el filtro $\{h_{j,l}\}$ tiene paso de banda nominal dado por: $1/2^{j+1} \leq |f| \leq 1/2^j$ con:

$$2^{j/2} \widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} h_{j,l} X_{t-l \bmod N}, \quad t = 0, \dots, N-1, \quad (3.46)$$

tenemos

$$W_{j,t} = 2^{j/2} \widetilde{W}_{j,2^j(t+1)-1} \equiv \sum_{l=0}^{L_j-1} h_{j,l} X_{2^j(t+1)-l-l \bmod N}, \quad (3.47)$$

³⁸Se ofrece una prueba en Percival y Walden (2000) p.96, 517.

³⁹Idem.

para $t = 0, \dots, N_j - 1$. Como también tenemos que $\mathbf{W}_j = \mathcal{W}_j \mathbf{X}$, Percival Y Walden (2000) muestran que las filas de \mathcal{W}_j contienen a $\{h_{j,t}^o\}$ el cual por definición es $\{h_{j,t}\}$ periodizado para tener longitud N .⁴⁰

En forma similar el filtro equivalente $\{g_{j,t}\}$ que relaciona $\{V_{j,t}\}$ con $\{X_t\}$ está formado por la convolución de los siguientes j filtros:

$$\begin{aligned}
 \text{filtro 1 : } & g_0, g_1, \dots, g_{L-2}, g_{L-1}; \\
 \text{filtro 2 : } & g_0, 0, g_1, 0, \dots, g_{L-2}, 0, g_{L-1}; \\
 \text{filtro 3 : } & g_0, 0, 0, 0, g_1, \dots, g_{L-2}, 0, 0, 0, g_{L-1}; \\
 & \vdots \\
 \text{filtro } j-1 : & g_0, \underbrace{0, \dots, 0}_{2^{j-2}-1 \text{ ceros}}, g_1, \underbrace{0, \dots, 0}_{2^{j-2}-1 \text{ ceros}}, \dots, g_{L-2}, \underbrace{0, \dots, 0}_{2^{j-2}-1 \text{ ceros}}, g_{L-1}; \\
 \text{filtro } j : & g_0, \underbrace{0, \dots, 0}_{2^{j-1}-1 \text{ ceros}}, g_1, \underbrace{0, \dots, 0}_{2^{j-1}-1 \text{ ceros}}, \dots, g_{L-2}, \underbrace{0, \dots, 0}_{2^{j-1}-1 \text{ ceros}}, g_{L-1};
 \end{aligned} \tag{3.48}$$

$$\tag{3.49}$$

Este filtro equivalente tiene ancho L_j y su función de transferencia está dada por⁴¹:

$$G_j(f) \equiv \prod_{l=0}^{j-1} G(2^l f) \tag{3.50}$$

El filtro $\{g_{j,t}\}$ es una aproximación a un filtro *low-pass* con paso de banda dado por $0 \leq |f| \leq 1/2^{j+1}$. Con

$$2^{j/2} \tilde{V}_{j,t} \equiv \sum_{l=0}^{L_j-1} g_{j,t} X_{t-l \bmod N}, \quad t = 0, \dots, N-1, \tag{3.51}$$

tenemos

$$V_{j,t} = 2^{j/2} \tilde{V}_{j, 2^{j(t+1)-1}} \equiv \sum_{l=0}^{L_j-1} g_{j,t} X_{2^{j(t+1)-l-1} \bmod N}, \tag{3.52}$$

para $t = 0, \dots, N_j - 1$. Como también tenemos que $\mathbf{V}_j = \mathcal{V}_j \mathbf{X}$ por analogía con el filtro de onduleta las filas de \mathcal{V}_j contienen a $\{g_{j,t}^o\}$ el cual por definición es $\{g_{j,t}\}$ periodizado para tener longitud N .

Recordando nuestro supuesto que $N = 2^J$, encontramos que la J -ésima repetición del algoritmo pirámide arroja vector $\mathbf{W}_J = [W_{J,0}]$ y $\mathbf{V}_J = [V_{J,0}]$,

⁴⁰Ibid. p. 96

⁴¹Ibid. p. 97

ambos unidimensionales. El algoritmo termina en este punto, con $W_{J,0}$ y $V_{J,0}$ constituyendo los dos coeficientes finales de la DWT, W_{N-1} y V_{N-1} en el vector $\mathbf{W} = \mathbf{W}\mathbf{X}$. El coeficiente W_{N-2} está asociado con un cambio en $\{X_t\}$ en la escala de $\tau_J = 2^{J-1} = N/2$, mientras que el último coeficiente es $W_{N-1} = \bar{X}\sqrt{N}$ donde \bar{X} es la media muestral de $\{X_t\}$.⁴²

En la etapa final, también tenemos una expresión para todas las filas de la matriz de la DWT, a saber,

$$\mathbf{W} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_j \\ \vdots \\ W_J \\ V_J \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 A_1 \\ \vdots \\ B_j A_{j-1} \cdots A_1 \\ \vdots \\ B_J A_{J-1} \cdots A_1 \\ A_J A_{J-1} \cdots A_1 \end{bmatrix}$$

En síntesis, para $j = 0, \dots, J$ con $N_j \equiv N/2^j$ la j -ésima etapa del algoritmo piramidal rota las N_{j-1} vectores fila en la matriz V_{j-1} de $N_{j-1} \times N$ en dos conjuntos de N_j vectores fila, a saber $W_j = B_j V_{j-1}$ y $V_j = A_j V_{j-1}$, donde B_j y A_j son matrices de $N_j \times N_{j-1}$ que contienen a los filtros de escala y onduletas $\{h_l\}$ y $\{g_l\}$ respectivamente, periodizados a la longitud N_{j-1} (donde definimos $V_0 = I_N$). Juntos los vectores fila B_j y A_j forman un conjunto de N_{j-1} vectores ortonormales. Con $V_{0,t} \equiv X_t$, la j -ésima etapa transforma $\{V_{j-1} : t = 0, \dots, N_{j-1} - 1\}$ en dos nuevas series, a saber, los coeficientes de onduletas del j -ésimo nivel $\{W_{j,t} : t = 0, \dots, N_{j-1} - 1\}$ y los coeficientes de escala del j -ésimo nivel $\{V_{j,t} : t = 0, \dots, N_{j-1} - 1\}$.⁴³

Recapitulemos las propiedades de los coeficientes resultantes de la j -ésima etapa de algoritmo. Los coeficientes de onduletas del j -ésimo nivel $\{W_{j,t}\}$ están asociados con:

1. $W_j = B_j A_{j-1}, \dots, A_1$ una matriz de $N_j \times N$ que satisface que $W_j W_j^T = I_{N_j}$ y que comprende las filas $\sum_{k=1}^{j-1} N_k$ hasta $\sum_{k=1}^j N_k - 1$ de la matriz de la DWT \mathbf{W} (la primera de esas sumas es interpretada como 0 cuando $j = 1$), con las filas de W_j conteniendo versiones trasladadas circularmente de $\{h_{j,l}\}$ periodizado a longitud N , i.e.

$$\{h_{j,l}^o\} \longleftrightarrow \left\{ H\left(\frac{2^{j-1}k}{N}\right) G\left(\frac{2^{j-2}k}{N}\right) \cdots G\left(\frac{k}{N}\right) \right\}$$

⁴²Ibid. p.97-98

⁴³Ibid. p. 99

(la sucesión $\{h_{j,t}^{\circ}\}$ aparece en orden invertido en las filas de la matriz \mathcal{W}_j , y las filas adyacentes difieren en traslaciones circulares de 2^j unidades)

2. Los elementos del vector $\mathbf{W}_j = \mathcal{W}_j \mathbf{X}$ de longitud N_j ,
3. El j -ésimo nivel detalle $\mathcal{D}_j = \mathcal{W}_j^T \mathbf{W}_j = \mathcal{W}_j^T \mathcal{W}_j \mathbf{X}$;
4. Promedios en la escala de $\tau_j = 2^{j-1}$;
5. Submuestreo por dos de un proceso *half-band* asociado con altas frecuencias $[1/4, 1/2]$, en los coeficientes de escala $\{V_{j-1,t}\}$, y
6. La porción de la DFT de $\{X_t\}$ con frecuencias en el intervalo $[1/2^{j+1}, 1/2^j]$, en orden reverso con respecto a la frecuencia.

Por comparación los coeficientes de escala del j -ésimo nivel $\{V_{j,t}\}$ están asociados con:

1. $\mathcal{V}_j = \mathcal{A}_j \mathcal{A}_{j-1}, \dots, \mathcal{A}_1$ una matriz de $N_j \times N$ que satisface tanto $\mathcal{V}_j \mathcal{V}_j^T = I_{N_j}$ como $\mathcal{W}_j \mathcal{V}_j^T = \mathcal{V}_j \mathcal{W}_j^T = 0_{N_j}$, cuyas filas generan el mismo subespacio de \mathbb{R}^N que la últimas N_j filas de la matriz de la DWT \mathcal{W} , con las filas de \mathcal{V}_j conteniendo versiones trasladadas circularmente de $\{g_{j,t}\}$ periodizado a longitud N , i.e.

$$\{g_{j,t}^{\circ}\} \longleftrightarrow \left\{ G\left(\frac{2^{j-1}k}{N}\right) G\left(\frac{2^{j-2}k}{N}\right) \cdots G\left(\frac{k}{N}\right) \right\}$$

(la sucesión $\{g_{j,t}^{\circ}\}$ aparece en orden invertido en las filas de la matriz \mathcal{V}_j)

2. Los elementos del vector $\mathbf{V}_j = \mathcal{V}_j \mathbf{X}$ de longitud N_j ,
3. El j -ésimo nivel de *smooth* $\mathcal{S}_j = \mathcal{V}_j^T \mathbf{V}_j = \mathcal{V}_j^T \mathcal{V}_j \mathbf{X}$;
4. Promedios en la escala de $\lambda_j = 2^j$;
5. Submuestreo por dos de un proceso *half-band* asociado con bajas frecuencias $[0, 1/4]$, en los coeficientes de escala $\{V_{j-1,t}\}$, y
6. La porción de la DFT de $\{X_t\}$ con frecuencias en el intervalo $[0, 1/2^{j+1}]$.

2. Regresiones

*** Linear Model ***

```
Call: lm(formula = as.vector(ts.log.FX.USA.UK) ~
as.vector(ts.log.RPI.USA) + as.vector(ts.log.RPI.UK))
```

Residuals:

Min	1Q	Median	3Q	Max
-0.4511	-0.06274	0.005066	0.05879	0.2968

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5443	0.2015	2.7005	0.0072
as.vector(ts.log.RPI.USA)	0.6769	0.1442	4.6931	0.0000
as.vector(ts.log.RPI.UK)	-0.6832	0.1028	-6.6483	0.0000

Residual standard error: 0.1176 on 387 degrees of freedom

Multiple R-Squared: 0.5494

F-statistic: 235.9 on 2 and 387 degrees of freedom, the p-value is 0

Modelos Completos

*** Linear Model ***

```
Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S1"]]) ~
as.vector(ts.log.RPI.USA.mra[["S1"]]) +
as.vector(ts.log.RPI.UK.mra[["S1"]]))
```

Residuals:

Min	1Q	Median	3Q	Max
-0.4257	-0.06392	0.003371	0.05772	0.2979

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.3051	0.1489	2.0481	0.0412
as.vector(ts.log.RPI.USA.mra[["S1"]])	0.8447	0.1086	7.7804	0.0000
as.vector(ts.log.RPI.UK.mra[["S1"]])	-0.8001	0.0788	-10.1573	0.0000

Residual standard error: 0.1177 on 387 degrees of freedom

Multiple R-Squared: 0.5511

F-statistic: 237.5 on 2 and 387 degrees of freedom, the p-value is 0

*** Linear Model ***

```
Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S2"]]) ~
as.vector(ts.log.RPI.USA.mra[["S2"]]) +
as.vector(ts.log.RPI.UK.mra[["S2"]]))
```

Residuals:

Min	1Q	Median	3Q	Max
-0.4136	-0.0625	0.004658	0.05882	0.3046

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.2590	0.1372	1.8880	0.0598
as.vector(ts.log.RPI.USA.mra[["S2"]])	0.8777	0.1004	8.7416	0.0000
as.vector(ts.log.RPI.UK.mra[["S2"]])	-0.8232	0.0732	-11.2406	0.0000

Residual standard error: 0.117 on 387 degrees of freedom

Multiple R-Squared: 0.5557

F-statistic: 242 on 2 and 387 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S3"]]) ~
as.vector(ts.log.RPI.USA.mra[["S3"]]) +
as.vector(ts.log.RPI.UK.mra[["S3"]]))

Residuals:

Min	1Q	Median	3Q	Max
-0.3512	-0.06023	0.00135	0.05879	0.3069

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.1052	0.0836	1.2577	0.2093
as.vector(ts.log.RPI.USA.mra[["S3"]])	0.9880	0.0648	15.2468	0.0000
as.vector(ts.log.RPI.UK.mra[["S3"]])	-0.9009	0.0499	-18.0606	0.0000

Residual standard error: 0.1145 on 387 degrees of freedom

Multiple R-Squared: 0.5836

F-statistic: 271.2 on 2 and 387 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S4"]]) ~
as.vector(ts.log.RPI.USA.mra[["S4"]]) +
as.vector(ts.log.RPI.UK.mra[["S4"]]))

Residuals:

Min	1Q	Median	3Q	Max
-0.3225	-0.04415	-0.001566	0.05264	0.2377

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.0613	0.0635	0.9652	0.3350
as.vector(ts.log.RPI.USA.mra[["S4"]])	1.0225	0.0507	20.1668	0.0000
as.vector(ts.log.RPI.UK.mra[["S4"]])	-0.9261	0.0405	-22.8584	0.0000

Residual standard error: 0.1103 on 387 degrees of freedom
 Multiple R-Squared: 0.6269
 F-statistic: 325.2 on 2 and 387 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S5"]]) ~
 as.vector(ts.log.RPI.USA.mra[["S5"]]) +
 as.vector(ts.log.RPI.UK.mra[["S5"]]))

Residuals:
 Min 1Q Median 3Q Max
 -0.316 -0.04683 0.006067 0.0645 0.2302

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.0377	0.0435	0.8676	0.3862
as.vector(ts.log.RPI.USA.mra[["S5"]])	1.0452	0.0367	28.4622	0.0000
as.vector(ts.log.RPI.UK.mra[["S5"]])	-0.9438	0.0314	-30.0137	0.0000

Residual standard error: 0.1049 on 387 degrees of freedom
 Multiple R-Squared: 0.7023
 F-statistic: 456.5 on 2 and 387 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S6"]]) ~
 as.vector(ts.log.RPI.USA.mra[["S6"]]) +
 as.vector(ts.log.RPI.UK.mra[["S6"]]))

Residuals:
 Min 1Q Median 3Q Max
 -0.1093 -0.01839 0.0007107 0.0294 0.09488

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.0141	0.0125	1.1313	0.2586
as.vector(ts.log.RPI.USA.mra[["S6"]])	1.0730	0.0116	92.7418	0.0000
as.vector(ts.log.RPI.UK.mra[["S6"]])	-0.9666	0.0110	-87.5196	0.0000

Residual standard error: 0.04467 on 387 degrees of freedom
 Multiple R-Squared: 0.9576
 F-statistic: 4366 on 2 and 387 degrees of freedom, the p-value is 0

Modelos Restringidos

*** Linear Model ***

Call: `lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S1"]]) -
as.vector(ts.log.RPI.USA.mra[["S1"]] - ts.log.RPI.UK.mra[["S1"]]))`
Residuals:

Min	1Q	Median	3Q	Max
-0.4277	-0.06278	0.007252	0.05791	0.2919

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5139	0.0062	82.8872	0.0000
<code>as.vector(ts.log.RPI.USA.mra[["S1"]]</code> <code>- ts.log.RPI.UK.mra[["S1"]])</code>	0.6991	0.0322	21.7238	0.0000

Residual standard error: 0.1178 on 388 degrees of freedom

Multiple R-Squared: 0.5488

F-statistic: 471.9 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: `lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S2"]]) -
as.vector(ts.log.RPI.USA.mra[["S2"]] - ts.log.RPI.UK.mra[["S2"]]))`
Residuals:

Min	1Q	Median	3Q	Max
-0.4157	-0.06309	0.006866	0.06009	0.2969

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5138	0.0062	83.1448	0.0000
<code>as.vector(ts.log.RPI.USA.mra[["S2"]]</code> <code>- ts.log.RPI.UK.mra[["S2"]])</code>	0.7007	0.0321	21.8536	0.0000

Residual standard error: 0.1174 on 388 degrees of freedom

Multiple R-Squared: 0.5517

F-statistic: 477.6 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: `lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S3"]]) -
as.vector(ts.log.RPI.USA.mra[["S3"]] - ts.log.RPI.UK.mra[["S3"]]))`
Residuals:

Min	1Q	Median	3Q	Max
-0.3736	-0.06409	0.002767	0.0603	0.2958

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5113	0.0062	82.4609	0.0000
<code>as.vector(ts.log.RPI.USA.mra[["S3"]]</code> <code>- ts.log.RPI.UK.mra[["S3"]])</code>	0.7115	0.0321	22.1357	0.0000

- ts.log.RPI.UK.mra[["S3"]])

Residual standard error: 0.1178 on 388 degrees of freedom
Multiple R-Squared: 0.5581
F-statistic: 490 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S4"]]) -
as.vector(ts.log.RPI.USA.mra[["S4"]] - ts.log.RPI.UK.mra[["S4"]]))

Residuals:

Min	1Q	Median	3Q	Max
-0.3785	-0.06408	0.004752	0.06119	0.4894

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5098	0.0062	82.7757	0.0000
as.vector(ts.log.RPI.USA.mra[["S4"]] - ts.log.RPI.UK.mra[["S4"]])	0.7322	0.0317	23.0776	0.0000

Residual standard error: 0.1171 on 388 degrees of freedom
Multiple R-Squared: 0.5785
F-statistic: 532.6 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S5"]]) -
as.vector(ts.log.RPI.USA.mra[["S5"]] - ts.log.RPI.UK.mra[["S5"]]))

Residuals:

Min	1Q	Median	3Q	Max
-0.3822	-0.06227	0.01917	0.07687	0.7489

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5064	0.0063	80.6195	0.0000
as.vector(ts.log.RPI.USA.mra[["S5"]] - ts.log.RPI.UK.mra[["S5"]])	0.7848	0.0318	24.7048	0.0000

Residual standard error: 0.1198 on 388 degrees of freedom
Multiple R-Squared: 0.6113
F-statistic: 610.3 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.mra[["S6"]]) -
as.vector(ts.log.RPI.USA.mra[["S6"]] - ts.log.RPI.UK.mra[["S6"]]))

Residuals:

Min	1Q	Median	3Q	Max
-0.3771	-0.04394	0.01595	0.04966	1.046

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.4974	0.0052	95.8417	0.0000
as.vector(ts.log.RPI.USA.mra[["S6"]] - ts.log.RPI.UK.mra[["S6"]])	0.9335	0.0246	37.9773	0.0000

Residual standard error: 0.09971 on 388 degrees of freedom

Multiple R-Squared: 0.788

F-statistic: 1442 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.waveshrink) ~
as.vector(ts.log.RPI.USA.waveshrink - ts.log.RPI.UK.waveshrink))

Residuals:

Min	1Q	Median	3Q	Max
-0.392	-0.05366	0.003275	0.05679	0.2764

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5142	0.0060	86.2569	0.0000
as.vector(ts.log.RPI.USA.waveshrink - ts.log.RPI.UK.waveshrink)	0.6991	0.0309	22.5903	0.0000

Residual standard error: 0.1133 on 388 degrees of freedom

Multiple R-Squared: 0.5681

F-statistic: 510.3 on 1 and 388 degrees of freedom, the p-value is 0

*** Linear Model ***

Call: lm(formula = as.vector(ts.log.FX.USA.UK.waveshrink) ~
as.vector(ts.log.RPI.USA.waveshrink) + as.vector(ts.log.RPI.UK.waveshrink))

Residuals:

Min	1Q	Median	3Q	Max
-0.3923	-0.05369	0.002907	0.06674	0.2755

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.5416	0.1951	2.7759	0.0058
as.vector(ts.log.RPI.USA.waveshrink)	0.6800	0.1397	4.8687	0.0000
as.vector(ts.log.RPI.UK.waveshrink)	-0.6858	0.0995	-6.8918	0.0000

Residual standard error: 0.1134 on 387 degrees of freedom

Multiple R-Squared: 0.5681

F-statistic: 254.5 on 2 and 387 degrees of freedom, the p-value is 0