



UNIVERSIDAD NACIONAL DEL LITORAL  
Facultad de Ingeniería y Ciencias Hídricas  
Instituto de Desarrollo Tecnológico para la Industria Química

**Métodos guiados por los datos  
para el análisis de señales:  
contribuciones a la  
descomposición empírica en modos**

**Marcelo Alejandro Colominas**

Tesis remitida al Comité Académico del Doctorado  
como parte de los requisitos para la obtención  
del grado de  
**DOCTOR EN INGENIERIA**  
Mención Inteligencia Computacional, Señales y Sistemas  
de la  
**UNIVERSIDAD NACIONAL DEL LITORAL**

**2016**

Comisión de Posgrado, Facultad de Ingeniería y Ciencias Hídricas, Ciudad Universitaria,  
Paraje "El Pozo", S3000, Santa Fe, Argentina.





**UNIVERSIDAD NACIONAL DEL LITORAL**  
**Facultad de Ingeniería y Ciencias Hídricas**

Santa Fe, 10 de agosto de 2016.

Como miembros del Jurado Evaluador de la Tesis de Doctorado en Ingeniería titulada *“Métodos guiados por los datos para el análisis de señales: contribuciones a la descomposición empírica en modos”*, desarrollada por el Bioing. Marcelo Alejandro COLOMINAS, en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas”, certificamos que hemos evaluado la Tesis y recomendamos que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

La aprobación final de esta disertación estará condicionada a la presentación de dos copias encuadernadas de la versión final de la Tesis ante el Comité Académico del Doctorado en Ingeniería.

Dr. Marcelo Risk

Dr. Martín Hurtado

Dr. Leandro Vignolo

Santa Fe, 10 de agosto de 2016

Certifico haber leído la Tesis, preparada bajo mi dirección en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas” y recomiendo que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

  
.....  
Dra. María Eugenia Torres  
Codirectora de Tesis  
.....  
Dr. Gastón Schlotthauer  
Director de Tesis

Universidad Nacional del Litoral  
Facultad de Ingeniería y  
Ciencias Hídricas  
  
Secretaría de Posgrado

Ciudad Universitaria  
C.C. 217  
Ruta Nacional N° 168 – Km. 472,4  
(3000) Santa Fe  
Tel: (54) (0342) 4575 229  
Fax: (54) (0342) 4575 224  
E-mail: posgrado@fich.unl.edu.ar



MÉTODOS GUIADOS POR LOS DATOS  
PARA EL ANÁLISIS DE SEÑALES:  
CONTRIBUCIONES A LA DESCOMPOSICIÓN EMPÍRICA EN MODOS

**Resumen.** Las señales multicomponentes, compuestas de un número relativamente pequeño de componentes cuyas frecuencias instantáneas son únicas, sirven como modelo a un gran número de señales del mundo real, como las señales de audio, señales biológicas o series de datos económicos. El análisis tradicional de tales señales se realiza mediante las llamadas representaciones tiempo-frecuencia o tiempo-escala. Cada componente de la señal genera una «cresta» en la representación, la cual puede ser detectada y, si se desea, posteriormente antitransformada. Según este paradigma, se tendrá un componente por cada cresta detectada.

Este enfoque por supuesto no es el único posible. La descomposición empírica en modos (EMD, del inglés *Empirical Mode Decomposition*) es una técnica completamente guiada por los datos que separa una señal en oscilaciones localmente rápidas y lentas. Al final, la señal original se expresa como la suma de una cantidad pequeña de modos, que en ciertos casos pueden representarse como funciones AM-FM. Sin embargo, la EMD no se basa en una representación tiempo-frecuencia o tiempo-escala y, por lo tanto, no identifica un componente por cada cresta sino que puede agrupar varias en un mismo componente, representando modos más complejos que simples funciones circulares, y por lo tanto se constituye en un método más versátil.

Una importante limitación de la EMD es que se define como la salida de un algoritmo, sin bases teóricas sólidas. Es por ello que, como primera contribución de esta tesis, proponemos ver a EMD como la solución a un problema de optimización sin restricciones como un primer paso a dotar a esta técnica de bases teóricas. Evaluamos nuestra propuesta en señales artificiales y al comparar nuestros resultados con otros métodos de EMD basados en optimización se evidencian ventajas tanto en desempeño como en tiempo de cómputo. Nuestra propuesta permite una extensión a dos dimensiones de manera natural y sin mayores dificultades. Evaluamos esta extensión bidimensional en imágenes artificiales y reales, comparando nuestros resultados con aquellos métodos similares del estado del arte. Los resultados muestran en todos los casos ventajas para nuestra propuesta.

Como segunda contribución, exploramos otra de las líneas de trabajo tendientes a dotar a EMD de bases teóricas. Proponemos fusionar la naturaleza guiada por los datos de EMD con las sólidas bases matemáticas de las representaciones tiempo-escala. Definiendo una *escala local* determinada por los extremos locales de la señal observada, logramos un método que es tan guiado por los datos como EMD, pero con bases mucho más sólidas. Evaluamos nuestra propuesta en señales artificiales, y comparamos con métodos también basados en representaciones tiempo-escala mostrando una mayor versatilidad para nuestra propuesta.

Como consecuencia de su naturaleza local, la EMD experimenta en ciertos casos un problema que se denomina «mezcla de modos». Para reducir su efecto se han propuestos variantes de EMD asistidas por ruido, dando lugar a una nueva rama de métodos. Sin embargo, las versiones existentes presentan dificultades que aún deben ser superadas: las descomposiciones obtenidas no son completas (la suma de las componentes no reconstruyen la señal original sin errores) y dan lugar a componentes espurios o que no cumplen con ciertas condiciones buscadas.

Debido a esta situación, y como tercera contribución de este trabajo de tesis, proponemos nuestra propia versión asistida por ruido de EMD que supera los problemas persistentes en las versiones anteriores. Evaluamos su desempeño en señales artificiales, comparando nuestros resultados con las versiones asistidas por ruido del estado del arte y evidenciando ventajas para nuestra propuesta. Además, aplicamos nuestro método a señales biológicas, obteniendo información relevante relacionada a procesos fisiológicos involucrados en su producción.

# DATA-DRIVEN METHODS FOR SIGNAL ANALYSIS: CONTRIBUTIONS TO EMPIRICAL MODE DECOMPOSITION

**Abstract.** Multicomponent signals, which are made of a superposition of a relatively small number of components with unique instantaneous frequencies, are useful to model a great number of real world signals, such as audio signals, biological signals or economic data series. The traditional analysis of these signals is performed through time-frequency or time-scale representations. Each component generates a ridge on the representation, which can be detected and, if desired, inversely transformed. According to this paradigm, one has one component for each detected ridge.

This approach, of course, is not the only possible. The empirical mode decomposition (EMD) is a completely data-driven technique, which separates a signal into locally slow and fast oscillations. At the end, the original signal can be expressed as the sum of a small number of modes, which can be represented in certain cases as AM-FM functions. However, since it is not based on time-frequency or time-scale representation, and therefore it does not identify one component for each ridge, but it can aggregate several ridges in one component, then it is capable of representing modes which are more complex than simple circular functions, becoming a more versatile method.

An important limitation of EMD is that it is defined as an algorithm output, without solid theoretical bases. Because of this, and as a first contribution, we propose to see EMD as the solution of an unconstrained optimization problem as a first step to give this technique theoretical bases. We evaluate our proposal on artificial signals and when compared with other optimization-based EMD methods there are advantages both in performance and computational time. Our proposal allows for an extension to two dimensions in a natural manner and without major difficulties. We evaluate this bidimensional extension on artificial and real signals, comparing our results with those of similar state of the art methods. In all cases, the results showed advantages for our proposal.

As a second contribution, we explore another line aimed at giving EMD theoretical bases. We propose to merge the data-driven nature of EMD with the solid mathematical bases behind time-scale representations. Defining a *local scale* determined by the observed local extrema, we achieved a method that is as data-driven as EMD, but with much more solid bases. We evaluate our proposal on artificial signals, and compare with other time-scale representation-based methods showing more versatility for our method.

As a consequence of its local nature, EMD experiences in certain cases a problem called mode mixing. In order to reduce its effects, the noise-assisted variations of EMD have been proposed, giving birth to a new family of methods. However, the existing versions have difficulties still to be solved: the obtained decompositions are not complete (the sum of the components does not reconstruct the original signal without errors) and there appear spurious components or components which does not fulfill certain desired conditions. Because of this situation, and as a third contribution, we propose our own noise-assisted version of EMD. We evaluate its performance on artificial signals, comparing our results with the noise-assisted EMD versions of the state of the art and showing advantages for our proposal. Moreover,

we apply our method to biological signals, obtaining relevant information related to physiological processes involved in its production.

# Agradecimientos

Quiero agradecer en primer lugar a mi amigo y director Gastón Schlotthauer, quien me reclutó hace varios años para dedicarme a la investigación científica. Gracias por alentarme y confiar siempre en mí.

A continuación quiero agradecer al resto de los integrantes del Laboratorio de Señales y Dinámicas no Lineales, de la Facultad de Ingeniería de la Universidad Nacional de Entre Ríos, donde encontré un ámbito de trabajo cordial, de camaradería y propicio al debate e intercambio de ideas.

Finalmente, a mis padres, Marcelo y Alejandra, que me pasaron el virus de la curiosidad, la indagación y el escepticismo. A mis hermanas, Lucrecia y Renata, por estar siempre ahí. Y a mi compañera, Cecilia, por su constante paciencia, acompañamiento y amor.



*Deseamos buscar la verdad, sin importar adónde nos lleve...*

---

COSMOS  
Carl Sagan

*Pero el avance en nuestra comprensión del universo no es apenas una tarea científica o académica; está condicionada por una estructura social agotada que ha llegado a ser una amenaza al medio ambiente y a la vida de los habitantes de este planeta. Para entender ese mundo es también un requisito revolucionarlo y reconstruirlo humanamente. ¿Entender el mundo para revolucionarlo o revolucionar el mundo para entenderlo? Es un interrogante paradójico que se resuelve actuando.*

---

EL MUNDO NO COMENZÓ EN EL 4004 ANTES DE CRISTO  
Pablo Rieznik



# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Marco teórico</b>	<b>5</b>
2.1. Introducción . . . . .	7
2.2. Señales multicomponentes . . . . .	7
2.2.1. Método de la señal analítica . . . . .	8
2.2.2. Limitaciones . . . . .	10
2.3. Representaciones lineales tiempo-frecuencia y tiempo-escala . . . . .	11
2.3.1. Extracción de líneas de cresta . . . . .	16
2.3.2. De las líneas de cresta a los modos . . . . .	17
2.4. Descomposición empírica en modos . . . . .	18
2.4.1. Principio recursivo . . . . .	18
2.4.2. Proceso de <i>sifting</i> . . . . .	21
2.4.3. Espectro de Hilbert-Huang . . . . .	23
2.4.4. Definición más general de EMD . . . . .	23
2.4.5. Propiedades de la descomposición . . . . .	26
2.4.6. Funciones intrínsecas de modo . . . . .	29
2.4.7. Cuestiones prácticas de implementación . . . . .	33
2.5. Comentarios finales . . . . .	38
<b>3. Un enfoque basado en optimización para EMD</b>	<b>41</b>
3.1. Introducción . . . . .	43
3.2. Hacia un enfoque teórico . . . . .	43
3.3. EMD como un problema de optimización . . . . .	44
3.3.1. Antecedentes . . . . .	44
3.3.2. Optimización sin restricciones (UOA-EMD) . . . . .	46
3.4. Características de la descomposición . . . . .	49
3.4.1. Comparación con <i>B-splines</i> . . . . .	49
3.4.2. Interpretación como filtro variante en el tiempo . . . . .	51
3.4.3. Espacio nulo de $P_x$ . . . . .	52
3.4.4. Relación con las IMTs de Daubechies . . . . .	54
3.5. Implementación . . . . .	56
3.5.1. Selección del parámetro $\lambda$ . . . . .	57
3.5.2. Efectos de borde . . . . .	59
3.5.3. Criterios de detención . . . . .	60
3.5.4. Efectos de muestreo . . . . .	61
3.6. Experimentos y resultados . . . . .	61

3.6.1.	Separación de tonos . . . . .	62
3.6.2.	Mezclas de oscilaciones AM-FM y no lineales . . . . .	64
3.6.3.	Ruido blanco gaussiano y banco de filtros equivalente . . . . .	66
3.6.4.	Descomposición de señales de ancho de banda localmente estrecho . . . . .	69
3.6.5.	Señal real de voz . . . . .	70
3.7.	Extensión a dos dimensiones . . . . .	72
3.7.1.	Nuestro enfoque . . . . .	73
3.7.2.	Cuestiones de implementación . . . . .	75
3.7.3.	Delta de Dirac bidimensional . . . . .	76
3.7.4.	Imágenes AM-FM . . . . .	76
3.7.5.	Imagen real . . . . .	81
3.7.6.	Aplicación biomédica . . . . .	83
3.8.	Comentarios finales . . . . .	84
<b>4.</b>	<b>Un enfoque tiempo-escala para EMD</b>	<b>85</b>
4.1.	Introducción . . . . .	87
4.2.	EMD y el problema de la media local . . . . .	87
4.3.	Transformada ondita continua y <i>synchrosqueezing</i> . . . . .	89
4.4.	Propuesta de algoritmo . . . . .	91
4.4.1.	Una nueva definición para los modos . . . . .	91
4.4.2.	La escala local . . . . .	91
4.4.3.	Problemas de discontinuidad y <i>synchroqueezing</i> . . . . .	93
4.4.4.	La constante C . . . . .	93
4.4.5.	Refinamiento de la escala local . . . . .	93
4.4.6.	Tratamiento de bordes . . . . .	93
4.5.	Experimentos y resultados . . . . .	94
4.5.1.	Suma de componentes AM-FM solapados . . . . .	94
4.5.2.	Formas de onda no lineales . . . . .	94
4.6.	Comentarios finales . . . . .	97
<b>5.</b>	<b>Versiones de EMD asistidas por ruido</b>	<b>99</b>
5.1.	Introducción . . . . .	101
5.2.	El problema de la mezcla de modos . . . . .	101
5.3.	Variaciones asistidas por ruido . . . . .	102
5.3.1.	EMD por conjuntos . . . . .	103
5.3.2.	EMD por conjuntos complementaria . . . . .	105
5.4.	EMD por conjuntos completa con ruido adaptativo . . . . .	106
5.4.1.	Antecedente . . . . .	108
5.5.	Un estudio sobre la amplitud de ruido . . . . .	109
5.5.1.	Ruido de amplitud constante . . . . .	110
5.5.2.	«Salto» en la amplitud del ruido . . . . .	111
5.5.3.	Ruido de amplitud creciente . . . . .	111
5.5.4.	Discusión . . . . .	112
5.6.	Experimentos y resultados . . . . .	114
5.6.1.	Señales artificiales . . . . .	114
5.6.2.	Electroglotograma . . . . .	119
5.6.3.	Electrocardiograma con fibrilación ventricular . . . . .	121

---

5.6.4. Electroencefalograma con crisis epiléptica . . . . .	122
5.7. Comentarios finales . . . . .	129
<b>6. Conclusiones y perspectivas</b>	<b>131</b>



# Índice de figuras

2.1.	Transformada de Fourier de tiempo corto . . . . .	13
2.2.	Transformada ondita continua . . . . .	15
2.3.	Combinación lineal de tres funciones coseno vistas como una única señal AM-FM . . . . .	19
2.4.	Descomposición de una señal en oscilaciones rápidas y lentas . . . . .	20
2.5.	Diagrama de flujos del algoritmo de EMD . . . . .	24
2.6.	Señal compuesta por un coseno y un seno y estimaciones de su frecuencia instantánea . . . . .	31
3.1.	Señal oscilante, extremos locales y función interpolante . . . . .	47
3.2.	Señal artificial descrita por la ec. (3.12) . . . . .	50
3.3.	Comparación con <i>B-splines</i> . . . . .	50
3.4.	Interpretación como filtro variante en el tiempo . . . . .	52
3.5.	Elementos no nulos de $P_x$ . . . . .	58
3.6.	Efectos de muestreo . . . . .	61
3.7.	Separación de tonos . . . . .	63
3.8.	Suma de dos tonos y un átomo gaussiano . . . . .	65
3.9.	Suma de dos formas de onda triangulares y una señal sinusoidal AM . . . . .	65
3.10.	Suma de dos señales AM-FM y un átomo gaussiano . . . . .	67
3.11.	Proceso de <i>sifting</i> para el primer modo de una realización de ruido blanco gaussiano . . . . .	68
3.12.	Descomposiciones de ruido blanco gaussiano . . . . .	69
3.13.	Señales de ancho de banda local estrecho . . . . .	70
3.14.	Señales de ancho de banda local estrecho . . . . .	71
3.15.	Descomposición de una señal de voz . . . . .	72
3.16.	Función AM-FM bidimensional, extremos locales y plano interpolante . . . . .	74
3.17.	Descomposición de una Delta de Dirac bidimensional . . . . .	77
3.18.	Resultados para el Experimento 1 . . . . .	78
3.19.	Errores y tiempos de cómputo para el Experimento 1 . . . . .	78
3.20.	Resultados para el Experimento 2 . . . . .	79
3.21.	Errores y tiempos de cómputo para el Experimento 2 . . . . .	79
3.22.	Resultados para el Experimento 3 . . . . .	80
3.23.	Errores y tiempos de cómputo para el Experimento 3 . . . . .	80
3.24.	Resultados para el Experimento 4 . . . . .	82
3.25.	Errores y tiempos de cómputo para el Experimento 4 . . . . .	82
3.26.	Descomposiciones de la imagen «Barbara» . . . . .	83
3.27.	Descomposición de una imagen de resonancia magnética mediante UOA-EMD . . . . .	84

4.1. Superposición de componentes AM-FM . . . . .	95
4.2. Superposición de un <i>chirp</i> y una forma de onda triangular . . . . .	96
5.1. Mezcla de modos . . . . .	102
5.2. Señal de enmascaramiento . . . . .	103
5.3. Diagrama de flujos del algoritmo de EEMD . . . . .	104
5.4. Diagrama de flujos del algoritmo de CEEMDAN . . . . .	107
5.5. Diseño del experimento . . . . .	110
5.6. Resultados para ruido de amplitud constante . . . . .	112
5.7. Resultados para amplitud de ruido con «salto» . . . . .	113
5.8. Resultados para ruido de amplitud linealmente creciente . . . . .	114
5.9. Descomposición de la señal artificial $s$ descrita por las ecuaciones (5.16) . . . . .	115
5.10. Desempeño de los cuatro métodos asistidos por ruido aplicados sobre la señal artificial $s$ . . . . .	117
5.11. Desempeño de los cuatro métodos asistidos por ruido aplicados sobre la señal artificial $s$ . . . . .	117
5.12. Desempeño de los cuatro métodos asistidos por ruido aplicados sobre la señal artificial $s$ . . . . .	118
5.13. Señal artificial $x$ , junto a su espectrograma reasignado y espectro marginal . . . . .	118
5.14. Descomposición de la señal artificial $x$ mediante EMD, EEMD, Complementary EEMD, CEEMDAN original y CEEMDAN mejorada.	120
5.15. Señal de EGG. . . . .	121
5.16. Descomposición mediante CEEMDAN mejorada de la señal de EGG mostrada en la Fig. 5.15. $\epsilon_0 = 0,2$ . $I = 100$ . . . . .	122
5.17. Señal de EGG . . . . .	123
5.18. Descomposición mediante CEEMDAN mejorada de la señal de EGG mostrada en la 5.17. $\epsilon_0 = 0,2$ . $I = 100$ . . . . .	124
5.19. Señal de ECG. . . . .	125
5.20. Descomposición mediante CEEMDAN mejorada de la señal de ECG mostrada en la 5.19. $\epsilon_0 = 0,2$ . $I = 100$ . . . . .	126
5.21. Señal de EEG . . . . .	127
5.22. Descomposición mediante CEEMDAN mejorada de la señal de EEG mostrada en la 5.21. $\epsilon_0 = 0,2$ . $I = 100$ . . . . .	128

# Índice de tablas

3.1. Ejemplo 1. Suma de dos tonos y un átomo gaussiano (1000 muestras). Errores cuadráticos relativos y tiempos computacionales para los cuatro métodos. . . . .	66
3.2. Ejemplo 2. Suma de dos formas de onda triangulares y una señal AM (1000 muestras). Errores cuadráticos relativos y tiempos computacionales para los cuatro métodos. . . . .	67
3.3. Ejemplo 3. Suma de dos señales AM-FM y un átomo gaussiano (un millón de muestras). Errores cuadráticos relativos y tiempos computacionales para los cuatro métodos. . . . .	68



# Acrónimos

En el presente documento, utilizaremos los acrónimos en inglés, por ser la forma más difundida en la bibliografía.

AM-FM	Modulado en amplitud y modulado en frecuencia
FT	Transformada de Fourier
HT	Transformada de Hilbert
STFT	Transformada de Fourier de tiempo corto
CWT	Transformada ondita continua
EMD	Descomposición empírica en modos
IMF	Función intrínseca de modo
HHT	Transformada de Hilbert-Huang
IMT	Función tipo intrínseca de modo de Daubechies
OS-EMD	EMD basada en optimización de <i>splines</i>
Prox-EMD	EMD basada en optimización proximal
UOA-EMD	EMD basada en optimización sin restricciones
NSP	Búsqueda de espacio nulo
wGn	Ruido blanco gaussiano
RSE	Error cuadrático relativo
RRSE	Raíz del error cuadrático relativo
IEMD	EMD para imágenes de Linderhed
TV-G	Descomposición textura+geometría de Gilles-Osher
EEMD	EMD por conjuntos
CEEMDAN	EMD por conjuntos completa con ruido adaptativo
SNR	Relación señal a ruido
EGG	Electroglotograma
ECG	Electrocardiograma
EEG	Electroencefalograma



# Notación

$x(t)$	Señal en tiempo continuo
$x = x[n]$	Serie de datos discreta
$\hat{x}(f) = (\mathcal{F}x)(f)$	Transformada de Fourier de $x(t)$
$(\mathcal{H}x)(t)$	Transformada de Hilbert de $x(t)$
$ z $	Módulo del número complejo $z$
$\arg(z)$	Argumento del número complejo $z$
$z^*$	Conjugado del número complejo $z$
$\text{sgn}(\cdot)$	Función signo
$\chi_I$	Función característica del conjunto $I$
$F_x^g(t, f)$	Transformada de Fourier de tiempo corto de $x$ con la ventana $g$
$W_x^\psi(t, s)$	Transformada ondita continua de $x$ con la ondita $\psi$
$SW_x(t, s)$	Transformada ondita de $x$ a la que se aplicó <i>synchrosqueezing</i>
$L_x^\varphi(t, s)$	Análisis de $x$ con función de escalamiento $\varphi$
$\mathbb{L}_x$	Conjunto de extremos locales de la señal $x$
$\#\mathbb{A}$	Cardinalidad del conjunto $\mathbb{A}$
$\mathcal{D}x = d_1$	Primer modo de EMD de la señal $x$
$\mathcal{E}_k x = d_k$	$k$ -ésimo modo de EMD
$\mathcal{M}x = a_1$	Media local de EMD de la señal $x$ ( $a_1 = x - \mathcal{D}x$ )
$d_k^j$	$j$ -ésima iteración de <i>sifting</i> del $k$ -ésimo modo de EMD
$d_k^{(j)}$	$k$ -ésimo modo de EMD de la $j$ -ésima realización
$x(t) = O(\epsilon)$	$\exists M/ x(t)  \leq M\epsilon$ para $t$ suficientemente grande
$\mathbb{E}(W)$	Valor esperado de la variable aleatoria $W$
$\text{var}(W)$	Varianza de la variable aleatoria $W$
$\text{std}(W)$	Desvío estándar de la variable aleatoria $W$
$\ \cdot\ _p$	Norma $L^p$ o $\ell^p$
$\text{diag}(M)$	Matriz diagonal construida a partir de la diagonal principal de $M$
$\text{vec}(M)$	Vector columna que resulta de concatenar las columnas de $M$
$\ker(M)$	Núcleo de la matriz $M$
$\text{tr}(M)$	Traza de la matriz $M$



# Capítulo 1

## Introducción

*El principio siempre es duro; esto vale para todas las ciencias.*

---

PRÓLOGO A LA PRIMERA EDICIÓN ALEMANA DEL  
PRIMER TOMO DE EL CAPITAL  
Karl Marx

El procesamiento de señales ha recorrido un largo camino desde Fourier hasta nuestros días. Una señal es la variación de alguna magnitud física (medida por ejemplo con algún sensor) con respecto a otra (en general el tiempo), y su procesamiento, en términos generales, apunta a extraer alguna parte de la información que acarrea. El enfoque tradicional, siguiendo a Fourier, consiste en dejar el espacio de observación, o sea el dominio temporal, en donde los datos se presentan «crudos» y pasar a otro espacio, por ejemplo el frecuencial, para «ver» a la señal desde otra perspectiva. De esta forma, la representación frecuencial resume el contenido *global* de frecuencias presentes en la señal. Este modo de trabajo presenta rápidamente algunas limitaciones, especialmente importantes en presencia de *no estacionariedades*. Se hace necesaria entonces una representación conjunta en tiempo y en frecuencia, que pueda reflejar las variaciones temporales de las frecuencias que nuestros sentidos, en especial la audición, están tan acostumbrados a percibir.

La transformada de Fourier de tiempo corto es la representación tiempo-frecuencia por excelencia. En ella, se limita el período de observación de la señal mediante una ventana (he allí el *tiempo corto*) deslizante, logrando cuantificar un contenido frecuencial *local* capaz de reflejar los cambios frecuenciales mencionados. Por otro lado a menudo es conveniente ver a la señal en otro espacio: el de las escalas (relacionado con la inversa de la frecuencia). La transformada ondita, o *wavelet* en inglés, es la representación tiempo-escala más estudiada. En ella se compara a la señal de interés con una forma de onda particular (la ondita) que en general posee soporte compacto (o decae suficientemente rápido). Trasladando y dilatando la ondita se logra una representación que permite apreciar las variaciones temporales de las escalas presentes.

Las señales llamadas multicomponentes, compuestas de una superposición de un número relativamente pequeño de componentes cuyas frecuencias instantáneas son únicas, poseen una representación tiempo-frecuencia o tiempo-escala con una estructura particular. Cada uno de los componentes genera una «cresta» en la

---

representación, la cual se ubica alrededor de su frecuencia instantánea real. De esa forma es que las representaciones mencionadas permiten realizar un seguimiento local de las frecuencias contenidas en la señal.

Pero analizar algo es descomponerlo. Además de las frecuencias instantáneas de cada componente, suele ser de utilidad la obtención de los componentes mismos en el dominio temporal. Esta tarea, denominada comúnmente «recuperación de modos de una señal multicomponente», atrajo recientemente la atención de una parte de la comunidad científica de procesamiento de señales. La estrategia basada en una representación tiempo-frecuencia o tiempo-escala consiste en identificar las crestas presentes en ella y «antitransformar» la representación en un dominio acotado que contenga a la cresta. Así se tiene un método que identifica a un componente por cada cresta.

Una técnica con un enfoque por completo diferente es la descomposición empírica en modos (EMD, del inglés *Empirical Mode Decomposition*). Esta técnica trabaja directamente en el espacio de observación, obteniendo los modos sin ninguna transformación de por medio. De manera completamente guiada por los datos, esta técnica separa una señal en oscilaciones localmente rápidas y lentas. Al final, la señal original se expresa como la suma de una cantidad pequeña de modos, que en ciertos casos pueden representarse como funciones AM-FM. Sin embargo, al no basarse la EMD en la idea de identificar un modo por cada cresta, sino que puede agrupar varias en un mismo componente, es capaz de representar modos más complejos que simples funciones circulares, y por lo tanto se constituye en un método más versátil.

En el Capítulo 2 presentamos generalidades sobre las señales multicomponentes, frecuencia instantánea y transformada de Hilbert. Asimismo, presentamos las representaciones lineales tiempo-frecuencia (transformada de Fourier de tiempo corto) y tiempo-escala (transformada ondita continua) como las herramientas más naturales para abordar el análisis del tipo de señales que nos interesan. Pero la mayor parte del capítulo se dedica a la discusión de la descomposición empírica en modos. Presentamos un estudio de esta técnica que intenta ser exhaustivo, desde su definición y la forma en que trabaja hasta los aspectos prácticos de su implementación.

A pesar de su versatilidad, y de ser aplicada en diversos campos, la EMD posee una limitación crucial: se la define como la salida de un algoritmo, sin bases matemáticas sólidas. Tal es así, que en los últimos años se han realizado esfuerzos por dotarla de la base matemática que carece. Entre las distintas líneas exploradas, la que se basa en teoría de optimización es una de las más promisorias. La cuestión radica en definir a las soluciones que construye EMD como soluciones a problemas de optimización. Siguiendo esa lógica, y como primera contribución original de esta tesis, es que en el Capítulo 3 proponemos formular a EMD como un problema de optimización sin restricciones. Nuestra propuesta es evaluada en señales artificiales y al comparar nuestros resultados con otros métodos de EMD basados en optimización se evidencian ventajas tanto en desempeño como en tiempo de cómputo. La simpleza de nuestra propuesta permite una extensión a dos dimensiones sin mayores dificultades. La extensión bidimensional de nuestro método se evalúa en imágenes artificiales y reales, comparando nuestros resultados con aquellos métodos similares del estado del arte. Los resultados muestran en todos los casos ventajas para nuestra propuesta.

---

Otra línea de trabajo tendiente a dotar a EMD de bases teóricas es aquella iniciada por Daubechies, quien en un artículo planteó las bases de una descomposición tipo EMD en base a una representación tiempo-escala. A diferencia de EMD, en este caso se cuenta con una definición precisa de los componentes obtenidos, pudiendo estimar unívocamente sus propiedades, con una precisión *a priori* y dadas ciertas condiciones de la señal analizada. Siguiendo esta línea, proponemos en el Capítulo 4 como segunda contribución original fusionar la naturaleza guiada por los datos de EMD con las sólidas bases matemáticas de las representaciones tiempo-escala. Definiendo una *escala local* determinada por los extremos locales de la señal observada, logramos un método que es tan guiado por los datos como EMD, pero con bases mucho más sólidas. Evaluamos nuestra propuesta en señales artificiales, y comparamos con métodos también basados en representaciones tiempo-escala mostrando una mayor versatilidad para nuestra propuesta.

Aunque aplicada en varios campos, la EMD posee una dificultad que limita, pero no impide, su aplicación a señales reales. Como consecuencia de su naturaleza local, experimenta un problema que se denomina «mezcla de modos». La solución que encontraron los investigadores dio origen a una nueva rama de métodos: las variaciones de EMD asistidas por ruido. La versión más difundida es la EEMD (del inglés *Ensemble EMD*), la cual ha sido ampliamente aplicada en problemas reales. Sin embargo presenta algunas dificultades: la suma de los componentes obtenidos no reconstruye la señal original de manera perfecta; y además existe una dificultad a la hora de la obtención de los resultados finales, los cuales se hallan promediando resultados parciales. Debido a esta situación, y como tercera contribución de este trabajo de tesis, en el Capítulo 5 proponemos nuestra propia versión asistida por ruido de EMD que supera los problemas persistentes en las versiones existentes. Evaluamos su desempeño en señales artificiales, comparando nuestros resultados con las versiones asistidas por ruido del estado del arte y evidenciando ventajas para nuestra propuesta. Además, aplicamos nuestro método a señales biológicas, obteniendo información relevante relacionada a procesos fisiológicos involucrados en su producción.

En el Capítulo 6 presentamos las conclusiones de este trabajo de tesis, así como las perspectivas que deja planteadas y los trabajos futuros.



# Capítulo 2

## Marco teórico

*Sin teoría revolucionaria tampoco puede haber movimiento revolucionario.*

---

QUÉ HACER  
Vladimir I. Lenin

### Contenidos

---

<b>2.1. Introducción</b>	<b>7</b>
<b>2.2. Señales multicomponentes</b>	<b>7</b>
2.2.1. Método de la señal analítica	8
2.2.2. Limitaciones	10
<b>2.3. Representaciones lineales tiempo-frecuencia y tiempo-escala</b>	<b>11</b>
2.3.1. Extracción de líneas de cresta	16
2.3.2. De las líneas de cresta a los modos	17
<b>2.4. Descomposición empírica en modos</b>	<b>18</b>
2.4.1. Principio recursivo	18
2.4.2. Proceso de <i>sifting</i>	21
2.4.3. Espectro de Hilbert-Huang	23
2.4.4. Definición más general de EMD	23
2.4.5. Propiedades de la descomposición	26
2.4.6. Funciones intrínsecas de modo	29
2.4.7. Cuestiones prácticas de implementación	33
<b>2.5. Comentarios finales</b>	<b>38</b>

---



## 2.1. Introducción

Las variaciones en frecuencia son sensaciones de las más primitivas, ya que estamos rodeados de luz que cambia de color, de sonidos que varían su tono, y muchos otros fenómenos cuyas periodicidades cambian en el tiempo. Muchas señales del mundo físico pueden ser modeladas como una superposición de ondas cuyas amplitudes y frecuencias varían con el tiempo. Éstas incluyen señales de audio (habla y música [76]), señales biomédicas (electrocardiograma [113, 114]), aquellas procedentes de máquinas rotatorias [65] o series de datos económicos [125].

En el presente capítulo introducimos las ideas básicas para el análisis de señales multicomponentes. En la Sec. 2.2 presentamos generalidades sobre frecuencia instantánea, transformada de Hilbert y las limitaciones existentes para la determinación precisa del valor de frecuencia para un instante dado. En la Sec. 2.3 presentamos las representaciones lineales tiempo-frecuencia y tiempo-escala y el esquema general para descomponer una señal a partir de ellas. Por último, en la Sec. 2.4 presentamos la técnica a la que le prestaremos la mayor atención a lo largo de esta tesis: la descomposición empírica en modos. Analizamos sus aspectos básicos, desde su definición como la salida de un algoritmo hasta sus aspectos prácticos de implementación.

## 2.2. Señales multicomponentes

La idea de *frecuencia instantánea* [25, 37] es algo así como un oxímoron. En el sentido tradicional del análisis de Fourier, la frecuencia se define sólo para funciones seno y coseno, siendo ésta una cantidad constante para toda la duración de la señal. De acuerdo a esta lógica, nada más corto que una oscilación completa permitiría la definición de frecuencia, con lo que una función que defina la frecuencia para cada instante carecería de sentido. Ya veremos que esto no es tan así, y en muchos casos es deseable descomponer una señal *multicomponente* en sus partes constitutivas, y determinar para cada una de ellas su frecuencia instantánea, con lo que se busca resolver un problema de la siguiente forma: dada una señal multicomponente que se puede escribir como

$$x(t) = \sum_{k=1}^K x_k(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)), \quad \text{con } A_k(t), \phi'_k(t) > 0 \forall t, \quad (2.1)$$

calcular  $A_k(t)$  y  $\phi'_k(t)$  y/o describir sus propiedades. En este caso, las variaciones temporales de  $A_k(t)$  y  $\phi'_k(t)$  son muchos menores que las de  $\phi_k(t)$ , lo cual de forma más precisa quiere decir que  $|A'_k(t)|, |\phi''_k(t)| \ll \phi'_k(t)$ . Esencialmente, cada componente es una señal modulada tanto en frecuencia como en amplitud (AM-FM); como  $|A'_k(t)|, |\phi''_k(t)| \ll \phi'_k(t)$ , entonces puede interpretarse localmente (es decir, en un intervalo  $[t - \delta, t + \delta]$ , con  $\delta \approx [\phi'_k(t)]^{-1}$ ) a los componentes  $x_k(t)$  como señales armónicas con amplitud  $A_k(t)$  y frecuencia  $\phi'_k(t)$  [31].

El razonamiento para un modelo como el de la ec. (2.1) es poder codificar no-estacionariedades de manera compacta a través de variaciones temporales de las amplitudes y frecuencias de *modos* «tipo Fourier». Cada componente, o simplemente modo,  $x_k(t)$  es una señal AM-FM, para la cual su frecuencia instantánea se define como la tasa de cambio del argumento de la función coseno [37].

**Definición 2.2.1. Frecuencia instantánea ideal.** Dada  $x(t) \in \mathcal{C}^2(\mathbb{R})$  de la forma  $x(t) = A(t) \cos(2\pi\phi(t))$ , con  $|A'(t)|, |\phi''(t)| \ll \phi'(t)$ , se define su frecuencia instantánea ideal como

$$f(t) \equiv \phi'(t). \quad (2.2)$$

En la práctica, sin embargo, no suele disponerse de las expresiones de  $A(t)$  y  $\phi(t)$  a partir de las cuales describir la amplitud y frecuencia instantáneas de la señal, sino que se tiene acceso simplemente a los datos crudos  $x(t)$  (en realidad por lo general uno dispone de una versión muestreada  $x[n] = x(nT)$ , con frecuencia de muestreo  $f_s = 1/T$ , pero supondremos sin pérdida de generalidad que poseemos la versión en tiempo continuo). Entonces, dada  $x$  existen infinitos pares  $A$  y  $\phi$  tales que  $x(t) = A(t) \cos(2\pi\phi(t))$ . Una primera opción puede ser definir  $\phi(t) \equiv 0$  y  $x(t) = A(t)$ , con lo que la idea de modulación en frecuencia y en amplitud queda prácticamente de lado. En el caso más general si consideramos una función  $B(t)$ , con  $0 < B(t) < 1$ , entonces podemos escribir [37]:

$$\begin{aligned} x(t) &= A(t) \cos(2\pi\phi(t)), \\ &= \frac{A(t)}{B(t)} B(t) \cos(2\pi\phi(t)), \\ &= \tilde{A}(t) \cos(2\pi\tilde{\phi}(t)), \end{aligned} \quad (2.3)$$

con  $\tilde{A}(t) = A(t)/B(t)$  y  $\tilde{\phi}(t) = \frac{1}{2\pi} \arccos(B(t) \cos(2\pi\phi(t)))$ .

### 2.2.1. Método de la señal analítica

Podemos encontrar una solución a este problema si primero consideramos el caso más sencillo de un tono puro  $x(t) = A \cos(2\pi f_0 t)$ , llamado a veces «monocromático» [25, 37]. La idea radica en ver al tono como la parte real de una función exponencial compleja

$$A \cos(2\pi f_0 t) = \mathcal{R}e(Ae^{i2\pi f_0 t}). \quad (2.4)$$

La amplitud y la frecuencia del tono puro son el módulo y la fase (excepto por un factor de  $2\pi$ ) de la exponencial, respectivamente. Su parte imaginaria  $A \sin(2\pi f_0 t)$  se obtiene a partir de la parte real mediante un corrimiento de fase de  $\pi/2$ : decimos entonces que las partes real e imaginaria están en *cuadratura*. La operación matemática detrás de esta transformación se describe más fácilmente en el dominio frecuencial.

**Definición 2.2.2. Transformada de Fourier (FT).** Dada una señal real  $x(t)$ , se define su transformada de Fourier  $\hat{x}(f)$  como

$$(\mathcal{F}x)(f) = \hat{x}(f) = \int_{-\infty}^{+\infty} x(t) e^{-i2\pi ft} dt, \quad \text{para } f \in \mathbb{R}. \quad (2.5)$$

Podemos recuperar la señal original mediante

$$(\mathcal{F}^{-1}\hat{x})(t) = x(t) = \int_{-\infty}^{+\infty} \hat{x}(f) e^{i2\pi ft} df, \quad \text{para } t \in \mathbb{R}. \quad (2.6)$$

Si comparamos la transformada de Fourier de una función coseno

$$\frac{1}{2}[\delta(f - f_0) + \delta(f + f_0)] \quad (2.7)$$

con la correspondiente a una función seno de la misma frecuencia

$$\frac{1}{2i}[\delta(f - f_0) - \delta(f + f_0)] \quad (2.8)$$

podemos apreciar que para ir de la primera a la segunda es necesario un filtrado lineal cuya respuesta en frecuencia es  $-i \operatorname{sgn}(f)$  (entonces su respuesta al impulso es  $\text{PV}(1/\pi t)$ , donde PV indica el valor principal de Cauchy).

**Definición 2.2.3. Transformada de Hilbert.** Dada una señal real  $x(t)$ , se define su transformada de Hilbert como

$$(\mathcal{H}x)(t) \equiv \text{PV} \int_{-\infty}^{+\infty} \frac{x(u)}{\pi(t - u)} du. \quad (2.9)$$

**Definición 2.2.4. Señal analítica<sup>1</sup>.** Dada una señal real  $x(t)$ , se define su versión analítica compleja como

$$z_x(t) \equiv x(t) + i\mathcal{H}x(t). \quad (2.10)$$

De esta forma, obtenemos un par módulo-fase de manera no ambigua, y podemos definir una amplitud instantánea  $A_x(t)$  y una frecuencia instantánea  $f_x(t)$ :

$$A_x(t) \equiv |z_x(t)|, \quad (2.11)$$

$$f_x(t) \equiv \frac{1}{2\pi} \frac{d}{dt} \arg z_x(t). \quad (2.12)$$

Usando coordenadas polares, la señal analítica describe un vector girando cuyas longitud y velocidad angular dependen del tiempo. Esta señal admite una descripción sencilla en el dominio frecuencial. Por definición de señal analítica sabemos que

$$\widehat{z}_x(f) = \widehat{x}(f) + i(-i) \operatorname{sgn}(f) \widehat{x}(f) = 2\widehat{x}(f) \chi_{[0, +\infty)} \quad (2.13)$$

donde  $\chi_I$  se refiere a la función característica del conjunto  $I$ . Esto muestra que la señal analítica se obtiene de la señal real removiendo de su espectro las frecuencias negativas. Esto no modifica la información de la señal porque  $x(t)$  se supuso real y por lo tanto  $\widehat{x}(-f) = \widehat{x}^*(f)$  ( $z^*$  indica el conjugado del número complejo  $z$ ). El truncado de las frecuencias negativas «complejiza» la señal original, lo que puede interpretarse como una reubicación de la redundancia. Dividir el ancho de banda de la señal por dos permite reducir a la mitad la frecuencia de muestreo necesaria (Shannon-Nyquist [109]); sin embargo ahora es necesario calcular dos valores para cada instante de muestreo: la parte real y la parte imaginaria de la señal compleja. Así, la «dimensión» de la señal se mantiene globalmente [37]. Siempre podemos recuperar la señal original:

<sup>1</sup>Este tipo de funciones complejas son llamadas «analíticas» porque satisfacen las condiciones de diferenciabilidad de Cauchy-Riemann [25]

$$x(t) = \mathcal{R}e(z_x(t)) = \mathcal{R}e\left(\int_0^\infty \widehat{z}_x(f) e^{i2\pi ft} df\right) = 2\mathcal{R}e\left(\int_0^\infty \widehat{x}(f) e^{i2\pi ft} df\right). \quad (2.14)$$

Si consideramos sólo las frecuencias positivas, entonces la frecuencia media  $\bar{f}$  admite dos definiciones equivalentes [37]

$$\bar{f} = \frac{1}{E_x/2} \int_0^{+\infty} f |\widehat{x}(f)|^2 df = \frac{1}{E_z} \int_{-\infty}^{+\infty} f_x(t) |z_x(t)|^2 dt, \quad (2.15)$$

con  $E_y = \int |y|^2 dt < +\infty$ , para  $y : t \mapsto B$ , siendo la energía de  $y$ . Si usamos  $\widehat{z}_x(f) = 2\widehat{x}(f)\chi_{[0,+\infty)}$  y, luego,  $E_z = 2E_x$ , podemos escribir

$$\bar{f} = \frac{1}{E_z} \int_0^{+\infty} f |\widehat{z}_x(f)|^2 df. \quad (2.16)$$

Aplicando la identidad de Parseval

$$\bar{f} = \frac{1}{i2\pi E_z} \int_{-\infty}^{+\infty} \frac{d}{dt} z_x(t) z_x^*(t) dt \quad (2.17)$$

$$= \frac{1}{i2\pi E_z} \int_{-\infty}^{+\infty} \left( \frac{d}{dt} |z_x(t)| + i |z_x(t)| \frac{d}{dt} \arg z_x(t) \right) |z_x(t)| dt. \quad (2.18)$$

La igualdad con (2.15) se cumple, si suponemos que  $\lim_{t \rightarrow -\infty} |z_x(t)|^2 = \lim_{t \rightarrow +\infty} |z_x(t)|^2$ , y esto es cierto tanto para señales de duración finita como aquellas cuyas envolventes se anulan en infinito, y para señales con envolventes constantes. Entonces podemos pensar a la frecuencia media como una media pesada de la frecuencia instantánea, con el peso dado por el cuadrado de la amplitud instantánea.

## 2.2.2. Limitaciones

Hemos visto que dada  $x(t) = A(t) \cos(2\pi\phi(t))$ , podemos encontrar su versión analítica

$$z_x(t) = |z_x(t)| e^{i \arg z_x(t)} = A_x(t) e^{i2\pi\phi_x(t)} \quad (2.19)$$

cuya parte real tiene la forma

$$\mathcal{R}e(z_x(t)) = A_x(t) \cos(2\pi\phi_x(t)). \quad (2.20)$$

Es importante aquí aclarar que no podemos concluir nada en el sentido opuesto. O sea, la señal analítica asociada a la señal real  $A_x(t) \cos(2\pi\phi_x(t))$  no tiene necesariamente las mismas expresiones  $A_x(t)$  y  $\phi_x(t)$  como su módulo y fase (salvo por el factor  $2\pi$ ) respectivamente. En general existen diferencias entre la versión analítica  $z_x(t)$  y la versión compleja en cuadratura  $z_x^q(t) = A(t) e^{i2\pi\phi(t)}$ .

**Teorema 2.2.1. (Bedrosian) [14].** *Sea una señal real  $x(t) = A(t) \cos(2\pi\phi(t))$ . Si  $A(t)$  es tal que  $\widehat{A}(f) \subseteq [-B, +B]$  y  $\phi(t)$  es tal que  $\widehat{\cos(2\pi\phi)}(f) \subseteq \{(-\infty, -B'] \cup [B', +\infty)\}$ , y si  $B' > B$ , entonces*

$$\mathcal{H}(A(t) \cos(2\pi\phi(t))) = A(t) \mathcal{H}(\cos(2\pi\phi(t))). \quad (2.21)$$

El teorema de Bedrosian nos dice que si los espectros de  $A(t)$  y  $\cos(2\pi\phi(t))$  son disjuntos, y el de  $A(t)$  ocupa un rango inferior de frecuencias, entonces la amplitud instantánea  $A(t)$  puede «salir» cuando calculamos  $\mathcal{H}x(t)$ , lo cual contribuiría a que la versión analítica se aproxime a la versión en cuadratura. Sin embargo esto no es suficiente a menos que  $\mathcal{H}(\cos(2\pi\phi(t))) = \sin(2\pi\phi(t))$ . Desafortunadamente esto no es cierto para cualquier función de fase  $\phi(t)$ .

**Teorema 2.2.2. (Nuttall) [82].** *Sea una señal real  $x(t) = A(t) \cos(2\pi\phi(t))$ ,  $x^q(t) = A(t) \sin(2\pi\phi(t))$  su versión en cuadratura y  $z_x^q(t) = A(t)e^{i2\pi\phi(t)}$  su versión compleja en cuadratura. Entonces, una medida de la discrepancia entre su transformada de Hilbert y su versión en cuadratura está dada por la energía de la diferencia*

$$\int_{-\infty}^{+\infty} |\mathcal{H}x(t) - x^q(t)|^2 dt = 2 \int_{-\infty}^0 |\widehat{z_x^q}(f)|^2 df. \quad (2.22)$$

El teorema de Nuttall nos da una medida global del error que se comete cuando se supone a la versión analítica compleja de una función igual a su versión en cuadratura. Como el error es proporcional a la energía en frecuencias negativas que posee el espectro de la versión compleja en cuadratura  $z_x^q(t)$ , una manera de reducirlo (pero nunca de eliminarlo por completo) consiste en aumentar la frecuencia de la señal, de modo de desplazar su espectro hacia la derecha. De esa forma, la versión compleja en cuadratura se aproxima a la versión analítica, y se disminuye el error dado por el teorema de Nuttall.

Huang y col. [58] argumentan sobre la poca utilidad de la medida de error dada por Nuttall ya que no permite «localizar» el error (esto es, decir dónde el error es mayor o dónde es menor) debido a su naturaleza global. Un análisis más detallado del error que se comete cuando se aproxima la versión compleja en cuadratura mediante la versión analítica puede hallarse en [61].

La frecuencia instantánea caracteriza un comportamiento local de la frecuencia como función del tiempo. De manera dual, también puede indagarse sobre el comportamiento local del tiempo como una función de la frecuencia [37].

**Definición 2.2.5. Retardo de grupo.** *Dada una señal analítica  $z_x(t)$  con su correspondiente transformada de Fourier  $\widehat{z_x}(f)$ , se define su retardo de grupo como*

$$t_x(f) = -\frac{1}{2\pi} \frac{d}{df} \arg \widehat{z_x}(f). \quad (2.23)$$

Podemos ver al retardo de grupo como una medida del instante de aparición de cierta frecuencia.

## 2.3. Representaciones lineales tiempo-frecuencia y tiempo-escala

La manera más natural de abordar un análisis (o sea, una descomposición) de una señal multicomponente como el modelo de la ec. (2.1) es través de una representación tiempo-frecuencia o tiempo-escala. En cada caso, se realiza el producto interno entre la señal y un núcleo  $k_{t\nu}$  trasladado tanto en tiempo como en frecuencia/escala:

$$R_x^k(t, \nu) = \langle x, k_{t\nu} \rangle = \int x(u)k^*(t - u, \nu)du, \quad (2.24)$$

donde la variable  $\nu$  representa a la frecuencia o a la escala (ya veremos una forma de relacionar ambas). Existen distintas representaciones tiempo-frecuencia/tiempo-escala que difieren en la forma del núcleo  $k_{t\nu}$ . Probablemente la primera en ser estudiada en profundidad haya sido la transformada de Fourier de tiempo corto (STFT, del inglés *Short Time Fourier Transform*). El análisis consiste en enfatizar la señal para un tiempo  $t$  y suprimirla en los otros instantes. Esto se logra multiplicando a la señal por una función ventana  $g(u)$ , centrada en  $t$  y que decae rápidamente a ambos lados. La señal modificada  $x_t(u) = x(u)g(u - t)$  es una función de dos tiempos: el tiempo fijo, en el cual estamos interesados,  $t$ , y el tiempo auxiliar  $u$ . Como la señal modificada enfatiza a la señal original en torno a  $t$ , su transformada de Fourier reflejará la distribución de frecuencias alrededor de ese instante.

**Definición 2.3.1. Transformada de Fourier de tiempo corto (STFT).** Sea  $g(t)$  una función real normalizada en  $L^2(\mathbb{R})$  y par. La transformada de Fourier de tiempo corto de la señal  $x \in L^2(\mathbb{R})$  usando la ventana  $g$  se define como

$$F_x^g(t, f) = \int_{-\infty}^{+\infty} x(u)g(u - t)e^{-i2\pi f(u-t)}du. \quad (2.25)$$

La representación en un plano  $t - f$  del módulo al cuadrado de la STFT de  $x$ ,  $SP_x^g(t, f) = |F_x^g(t, f)|^2$ , se denomina el espectrograma de  $x$ .

La definición que acabamos de dar difiere sutilmente de la forma tradicional en que se presenta la STFT en un factor de desfase  $e^{i2\pi ft}$ . La forma (2.25) resulta muy conveniente cuando se compara a la STFT con la transformada ondita (la cual veremos más adelante), y además ajusta a la STFT a la ec. (2.24) con  $k_{tf}(u) = g(u - t)e^{i2\pi f(u-t)}$  constituyendo una familia de átomos tiempo-frecuencia. Así, podemos ver que  $F_x^g(t, f) = (x \star g_f)(t)$ , con  $g_f(u) = g(u)e^{i2\pi fu}$ , donde « $\star$ » representa al operador de convolución (recordemos que  $g$  es par). De esta manera, podemos ver que

$$F_x^g(t, f) = \mathcal{F}^{-1}(\hat{x}(u)\hat{g}^*(u - f)), \quad (2.26)$$

lo que nos dice que debemos interpretar a la STFT como un análisis realizado por un banco uniforme de filtros de ancho de banda constante [37]. En la Fig. 2.1 pueden apreciarse tanto los átomos tiempo-frecuencia como los filtros en frecuencia que éstos generan.

Aunque muy popular en la literatura reciente [62, 85], la forma (2.25) ya había sido discutida por Cohen [25] quien la presentó como una «transformada temporal de frecuencia corta» (*short-frequency time transform* en inglés) en la cual se hacía hincapié en el estudio de propiedades temporales para una frecuencia en particular. Resulta evidente que el espectrograma obtenido a partir de (2.25) es idéntico a aquél obtenido a partir de su forma más tradicional (ya que difieren en un factor  $e^{i2\pi ft}$  de módulo unitario).

La STFT es invertible en  $L^2(\mathbb{R})$ , esto es que podemos recuperar la señal original a partir de su STFT [37]:

$$x(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_x^g(u, f)g(t - u)e^{i2\pi f(t-u)}dudf. \quad (2.27)$$

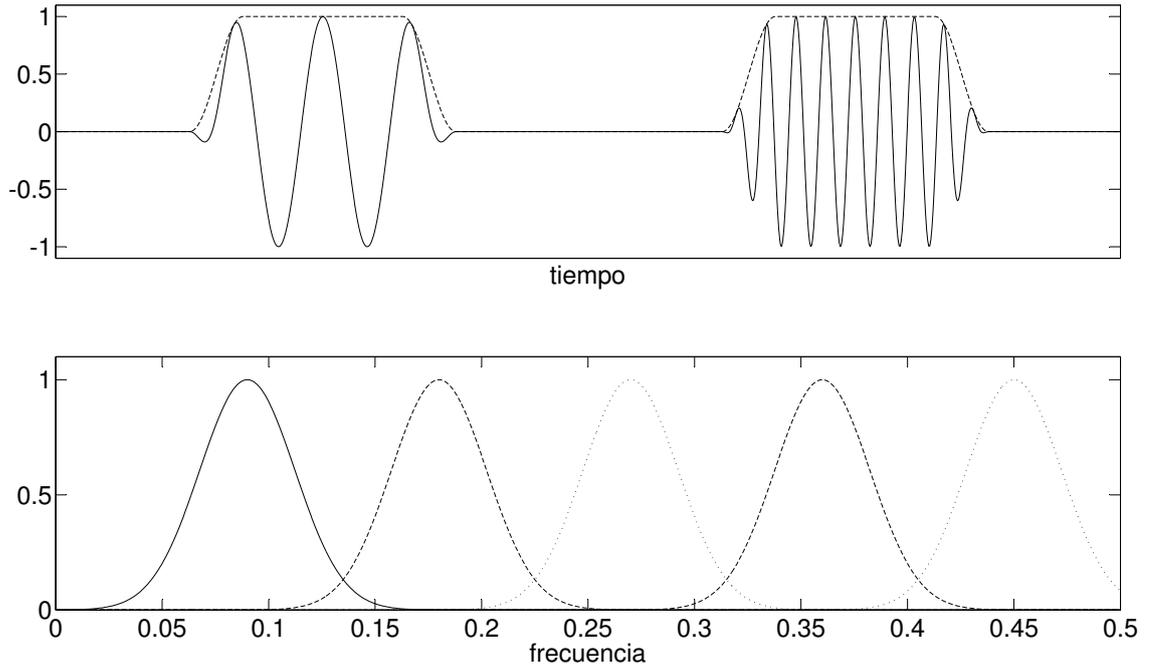


Figura 2.1: Transformada de Fourier de tiempo corto. Arriba: átomos tiempo-frecuencia. El análisis de la ec. (2.25) puede ser visto como una proyección de la señal analizada sobre átomos tiempo-frecuencia. Cada uno de estos átomos se obtiene a partir de una única ventana  $g(u)$  que se traslada en tiempo y se modula en frecuencia. La figura presenta dos ejemplos de tales átomos. La línea sólida muestra su parte real y la línea de trazos su módulo. Abajo: banco uniforme de filtros. En la interpretación de la ec. (2.26) la STFT puede verse como un análisis realizado por un banco uniforme de filtros. Cada filtro se deduce a partir de una única plantilla que se desplaza en frecuencia. La figura representa esquemáticamente algunos de esos filtros.

Esta fórmula de reconstrucción sin embargo no es única. También se puede recuperar  $x$  mediante una reconstrucción «vertical» [13]:

$$x(t) = \frac{1}{g(0)} \int_{-\infty}^{+\infty} F_x^g(t, f) df, \quad (2.28)$$

la cual es válida para una ventana  $g$  tal que  $g(0) \neq 0$  (este requisito es razonable ya que queríamos una ventana  $g$  que enfatice en torno a su centro a la señal de interés  $x$ ).

Por otro lado, la transformada ondita continua (CWT, del inglés *Continuous Wavelet Transform*) [30] usa como núcleos versiones trasladadas y escaladas de funciones oscilatorias de soporte compacto, o que decaen suficientemente rápido:  $k_{ts}(u) = \frac{1}{s} \psi\left(\frac{u-t}{s}\right)$ .

**Definición 2.3.2. Transformada ondita continua (CWT).** Sea  $x(t)$  una señal real y  $\psi$  una función que pertenece a la clase de Schwartz. La transformada ondita continua de  $x$  usando la ondita  $\psi$  se define como

$$W_x^\psi(t, s) = \frac{1}{s} \int_{-\infty}^{+\infty} x(u) \psi^*\left(\frac{u-t}{s}\right) du, \quad \text{con } s > 0. \quad (2.29)$$

Una conversión de escalas a frecuencias puede realizarse haciendo  $s = f_0/f$ , con  $f_0 = \arg \max |\hat{\psi}(f)|$ , obteniendo la siguiente representación:

$$W_x^\psi(t, f) = \frac{f}{f_0} \int_{-\infty}^{+\infty} x(u) \psi^* \left( \frac{f(u-t)}{f_0} \right) du, \quad \text{con } f > 0. \quad (2.30)$$

La representación en un plano  $t-s$  del módulo al cuadrado de la CWT de  $x$ ,  $SC_x^\psi(t, s) = |W_x^\psi(t, s)|^2$ , se denomina el escalograma de  $x$ .

Para el análisis de la evolución de tonos frecuenciales, es necesario el uso de onditas analíticas [76]. En lo que resta del documento nos referiremos siempre a onditas analíticas, a menos que digamos lo contrario. La definición de transformada ondita que damos difiere de la forma en que es presentada tradicionalmente por Daubechies [30] o Mallat [76] en un factor de  $\sqrt{s}$ , de forma que para una ondita  $\psi$  dada se tiene  $\tilde{W}_x = \sqrt{s}W_x$ . Nuestra normalización coincide con la elegida en [21, 62, 78, 85].

Así como ocurre con la STFT, la transformada ondita puede verse como una convolución entre la señal y una forma de onda particular:  $W_x^\psi(t, s) = (x \star \bar{\psi}_s)(t)$ , con  $\bar{\psi}_s = \frac{1}{s} \psi \left( \frac{-t}{s} \right)$ . Esta interpretación resulta de especial interés ya que permite ver a la CWT como un producto entre la FT de la señal analizada y versiones dilatadas de la FT de la ondita. Más precisamente

$$W_x^\psi(t, s) = \mathcal{F}^{-1} \left( \hat{x}(f) \hat{\psi}^*(sf) \right), \quad (2.31)$$

lo que sugiere ver a la CWT como un análisis mediante un banco continuo de filtros de factor  $Q$  constante [37] (donde el factor  $Q$  es el factor de *calidad* del filtro que relaciona su frecuencia central con su ancho de banda). Esta interpretación refuerza la normalización que elegimos para la transformada ondita. Al tener todos los filtros la misma ganancia, se dará que, por ejemplo, dos senos puros de la misma amplitud y distinta frecuencia estarán representados con la misma amplitud en la CWT. Puede apreciarse en la Fig. 2.2 los átomos tiempo-escala y el banco de filtros de factor  $Q$  constante producido por ellos.

La CWT también es invertible en  $L^2(\mathbb{R})$  [76, 84]:

$$x(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^{+\infty} \int_{-\infty}^{+\infty} W_x^\psi(u, s) \psi \left( \frac{t-u}{s} \right) du \frac{ds}{s^2} \right), \quad (2.32)$$

con  $C_\psi = \int_0^{+\infty} |\hat{\psi}(f)|^2 \frac{df}{f}$  denominada «condición de admisibilidad».

Esta fórmula de reconstrucción tampoco es única. Como con la STFT, la CWT también admite una reconstrucción «vertical»:

$$x(t) = \frac{2}{\widetilde{C}_\psi} \mathcal{R}e \left( \int_0^{+\infty} W_x^\psi(t, s) \frac{ds}{s} \right), \quad (2.33)$$

donde  $\widetilde{C}_\psi = \int \hat{\psi}^*(f) \frac{df}{f}$ . Vemos que en este caso la reconstrucción se logra integrando  $W_x^\psi$  respecto al logaritmo de la escala  $ds/s = d \log s$ , que es lo usual para medidas basadas en la ondita continua. La fórmula de reconstrucción (2.33) atrajo recientemente la atención de una parte de la comunidad científica, y en la actualidad muchos de los nuevos desarrollos se basan en ella [13, 63, 85, 113]. Presentada por

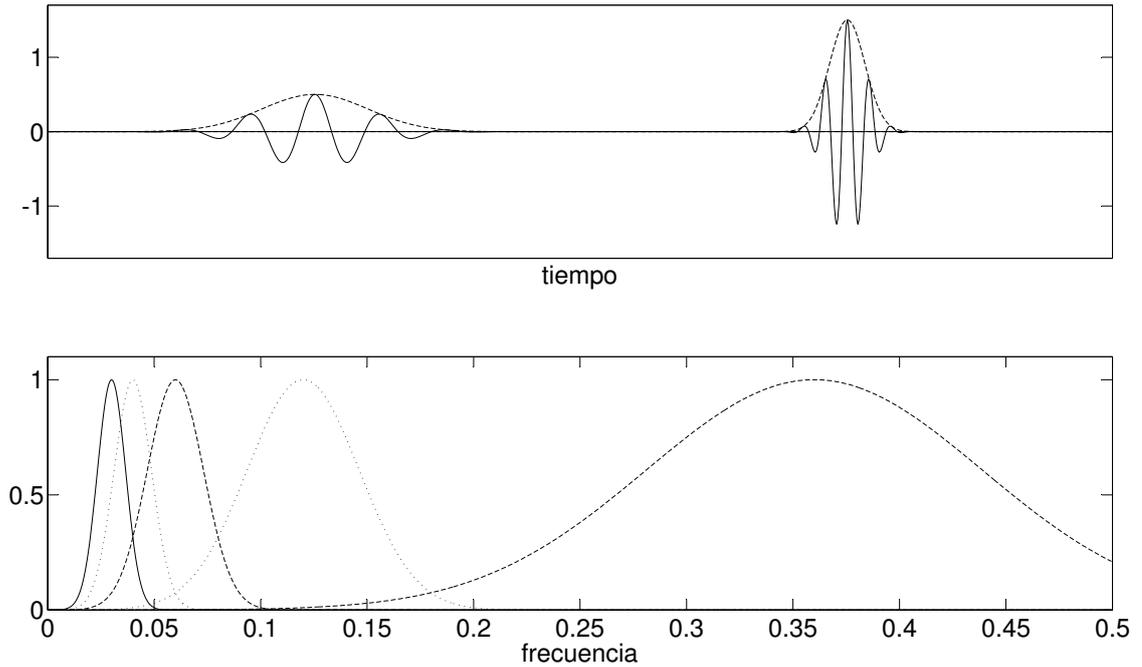


Figura 2.2: Transformada ondita continua. Arriba: átomos tiempo-escala. El análisis ondita de la ec. (2.30) puede verse como una proyección de la señal analizada sobre átomos tiempo-escala. Cada uno de estos átomos se obtiene a partir de una única ondita  $\psi(u)$  que se traslada en tiempo y se dilata o comprime. La figura presenta dos ejemplos de tales átomos. La línea sólida muestra la parte real y la línea de trazos su módulo. Abajo: banco de filtros de factor  $Q$  constante. En su interpretación de la ec. (2.31) la CWT puede ser vista como un análisis realizado por un banco de filtros  $Q$  constante. Cada filtro se deduce a partir de un único modelo que se dilata o comprime en frecuencia. La figura representa esquemáticamente algunos de esos filtros.

Daubechies en [32], ella misma la puso nuevamente en escena en [31] en el marco del problema de recuperación de modos. Sin embargo, la fórmula ya era conocida por Morlet, quien la encontró como una fórmula «empírica», pero que en realidad puede obtenerse si se usa una función delta de Dirac como núcleo de síntesis [8, 36].

Hemos visto entonces cómo el producto interno entre la señal y un núcleo tiempo-frecuencia o tiempo-escala permite un seguimiento local de la frecuencia o escala. Sin embargo, este seguimiento no puede ser «perfecto». Esto es, para un instante dado no se tiene una precisión absoluta sobre la frecuencia instantánea, y para una frecuencia dada no puede conocerse perfectamente el instante en que aparece [25, 37].

**Definición 2.3.3. Principio de incertidumbre de Heisenberg-Gabor.** Dada una señal  $x(t)$ , definimos  $\Delta_t^2 = (\int t^2 |x(t)|^2 dt - (\int t |x(t)|^2 dt)^2) / E_x$ , que mide el soporte temporal de  $x(t)$ , y  $\Delta_f^2 = (\int f^2 |\hat{x}(f)|^2 df - (\int f |\hat{x}(f)|^2 df)^2) / E_x$ , que mide el soporte frecuencial. Entonces, necesariamente debe darse que

$$\Delta_t \Delta_f \geq \frac{1}{4\pi}. \quad (2.34)$$

El principio de incertidumbre impone entonces un límite en la precisión que pueda tener la representación tiempo-frecuencia, de forma que siempre debe haber

un compromiso entre la resolución temporal (esto es saber en qué preciso momento ocurre tal frecuencia) y la resolución frecuencial (qué frecuencias aparecen para un tiempo dado).

### 2.3.1. Extracción de líneas de cresta

Recordemos la premisa inicial, expuesta justo después de la ec. (2.1): dada una señal multicomponente, deseamos describir la frecuencia instantánea para cada uno de sus componentes. Si la señal es de la forma (2.1), entonces en general tendrá una representación tiempo-frecuencia «rala», o sea, tendrá pocos coeficientes no nulos agrupados ellos en bandas alrededor de las frecuencias instantáneas teóricas de cada componente [31].

Las mencionadas bandas, en contraposición de líneas, son consecuencia del teorema de incertidumbre. Entonces, con el fin de obtener una estimación de la frecuencia instantánea, se hace necesario encontrar en cada banda una línea dominante alrededor de la cual se agrupan los coeficientes. Esta línea, denominada línea de cresta o más popularmente *ridge* en inglés, puede encontrarse de diversas maneras. Una forma muy difundida es aquella propuesta en [20] en la cual se encuentran las crestas  $\mathcal{C} = \{c_k, k = 1, \dots, K\}$  resolviendo el siguiente problema de optimización:

$$\min_{\mathcal{C}} \sum_{k=1}^K \left( - \int |R_x^k(t, c_k(t))|^2 dt + \int (\alpha c_k'(t)^2 + \beta c_k''(t)^2) dt \right), \quad (2.35)$$

donde  $R_x^k$  es una representación tiempo-frecuencia o tiempo-escala de la señal  $x$  para un núcleo  $k$ . Vemos que el problema de optimización intenta maximizar la energía a lo largo de la cresta al tiempo que favorece una cresta continua (a través del término del parámetro  $\alpha$ ) y suave (mediante el término del parámetro  $\beta$ ). Esta formulación para la obtención de crestas es muy sencilla e intuitiva, y sin embargo muy sensible a la selección de parámetros  $\alpha$  y  $\beta$ ; además es en general de difícil implementación. Debido a ello es que suelen preferirse algoritmos heurísticos como el de recocido simulado (en inglés *simulated annealing*) [20] o el algoritmo de los «escaladores locos» (en inglés *crazy climbers algorithm*) [21]. Otro enfoque simple para la resolución de (2.35) es el propuesto en [106], que consiste en determinar localmente las crestas empezando con distintas inicializaciones y promediar los resultados obtenidos.

Cualquiera sea el método seleccionado para la obtención de las líneas de cresta (todos ellos sensibles a algún conjunto de parámetros), una vez aplicado se obtienen las líneas  $c_k$  que son estimaciones para las frecuencias instantáneas ideales  $f_k$ . Estas estimaciones, sin embargo, pueden mejorarse si en lugar de extraer líneas de cresta de la representación  $R_x^k$  original, se extraen de una transformación de ésta para la cual se haya previamente «concentrado» la información que estaba dispersa como consecuencia del principio de incertidumbre. Esa es la idea principal detrás del método de reasignación o *reassignment*, desarrollado originalmente para ser aplicado sobre el espectrograma [68, 69], y luego generalizado para cualquier tipo de distribución bilineal [12]. Este método puede verse como una reubicación de la redundancia, mediante la cual se obtiene una representación mucho más «concentrada», pudiendo estimarse mejor las líneas de cresta. Así se pasa de un espectrograma o escalograma a una versión «reassignada» de ellos ( $RSP_x^g(\tilde{t}, \tilde{f})$  y

$RSC_x^\psi(\tilde{t}, \tilde{s})$  respectivamente), en la cual los coeficientes modifican su ubicación tanto en tiempo como en frecuencia o escala. La desventaja del método de *reassignment* es que no es invertible, o sea, no podemos volver del espectrograma reasignado a la señal.

### 2.3.2. De las líneas de cresta a los modos

En muchos casos, la necesidad de estimar las frecuencias instantáneas de los componentes puede ser acompañada por la necesidad de extraer tales modos  $x_k(t)$ . Este problema, conocido como recuperación de modos o componentes, se encuentra actualmente en auge, recibiendo gran atención por parte de la comunidad de procesamiento de señales [13, 31, 78].

La extracción de los modos está íntimamente relacionada con las ecuaciones de reconstrucción vertical (2.28) y (2.33). En general, dada una representación  $R_x^k$  de una señal que sigue un modelo como en la ec. (2.1), y una vez que las crestas  $\{c_k\}$  fueron obtenidas, los componentes individuales pueden ser recuperados como [13]

$$x_k(t) \approx C_R^{-1} \int_{|\nu - c_k(t)| < d_k} R_x^k(t, \nu) d\mu_\nu, \quad (2.36)$$

donde  $C_R$  es igual a  $g(0)$  en el caso de la STFT o  $\widetilde{C}_\psi$  para la CWT; y la medida  $\mu_\nu$  es igual a  $f$  para la STFT o  $\log s$  para la CWT. Vemos entonces cómo los componentes se obtienen de invertir  $R_x^k$  en una banda centrada en la cresta estimada  $c_k(t)$  y de ancho  $2d_k$ . Por supuesto que el desempeño de la recuperación de los componentes estará influido por la selección de los parámetros  $d_k$ , que necesariamente deberá estar relacionada con la separación frecuencial de los componentes, de modo que las bandas no se «solapen». Idealmente se tiene un modo por cada línea de cresta.

Con el objetivo de mejorar la concentración de la información en la CWT, en un espíritu como en *reassignment*, pero manteniendo la propiedad de invertibilidad, es que Daubechies y Maes presentaron la técnica de *synchrosqueezing* en [32]. En ella se procedía a una reasignación vertical, o sea se reubicaban los componentes sólo en escala, manteniendo su posición temporal. De esa manera se garantizaba la posibilidad de inversión de la nueva representación. Años después, esta técnica fue puesta nuevamente en escena bajo la luz del problema de recuperación de modos [31], definiendo una clase particular de funciones para las cuales se puede determinar unívocamente  $A_k(t)$  y  $f'_k(t)$ , dada una precisión definida *a priori* y bajo ciertas condiciones de  $x(t)$ , definiendo entonces condiciones de separabilidad en base a la CWT. A este trabajo siguieron otros abocados a la formalización de *synchrosqueezing* sobre la STFT, lo cual planteó otra clase de funciones que pueden ser correctamente separadas, ahora con criterios de separabilidad para la STFT. En general, en *synchrosqueezing* se pasa de la STFT o CWT a sus versiones reasignadas verticalmente:  $SF_x^g(t, \tilde{f})$  y  $SW_x^\psi(t, \tilde{s})$ . La recuperación de los componentes individuales se realiza como en la ec. (2.36), pero donde la constante de normalización ha sido incluida en la nueva representación, que es la que debe invertirse.

## 2.4. Descomposición empírica en modos

Hemos visto hasta ahora para la descomposición de  $x(t) = \sum_{k=1}^K x_k(t)$  un enfoque que puede resumirse de la siguiente manera:

$$x(t) \xrightarrow{\text{STFT o CWT}} R_x^k(t, \nu) \xrightarrow{\text{líneas de cresta}} \{x_k(t)\},$$

donde al primer paso (STFT o CWT) se puede agregar *synchrosqueezing*, como se mencionó anteriormente. Entonces la calidad de la descomposición estará influida por el tipo de representación elegido (y por los parámetros inherentes, entre ellos la ventana  $g$  o la ondita  $\psi$ ), y por el algoritmo de extracción de líneas de cresta que se utilice (que como hemos visto son sensibles a sus propios parámetros). Pero en términos generales, se tendrá que una línea de cresta implica un modo.

Esta afirmación puede ponerse en discusión considerando la siguiente señal:

$$\begin{aligned} x(t) &= \frac{1}{4} \cos((\alpha - \beta)t) + \frac{9}{4} \cos(\alpha t) + \frac{1}{4} \cos((\alpha + \beta)t) \\ &= \left( 2 + \cos^2\left(\frac{\beta}{2}t\right) \right) \cos(\alpha t), \end{aligned} \quad (2.37)$$

la cual cumple que, para  $\alpha \gg \beta$ ,  $A(t) = 2 + \cos^2(\frac{\beta}{2}t)$  oscila mucho más lentamente que  $\cos(\alpha t)$ . Si bien la señal  $x(t)$  posee idealmente una representación con tres líneas de cresta en las frecuencias  $\alpha - \beta$ ,  $\alpha$  y  $\alpha + \beta$ , como puede verse claramente en la primera línea de la ec. (2.37), éstas podrán diferenciarse entre sí sólo si la resolución elegida para la representación es menor a la separación  $\beta$  entre las crestas. Por otro lado, la segunda línea sugiere una interpretación por completo distinta: un único componente de frecuencia  $\alpha$  y una lenta modulación en amplitud. Dependiendo de las circunstancias, cualquiera de las dos interpretaciones puede ser considerada la «mejor».

La Descomposición Empírica en Modos (EMD, del inglés *Empirical Mode Decomposition*), presentada por Huang y colaboradores [58], es una técnica que prefiere la segunda interpretación cuando  $\alpha \gg \beta$  (ver Fig. 2.3), la cual tiene mucho más sentido físico si consideramos, por ejemplo, la forma en que nuestra audición percibe este tipo de señales. Esta técnica guiada por los datos separa localmente a la señal en oscilaciones rápidas y lentas, obteniendo como resultado *señales que oscilan simétricamente alrededor de cero*, como en la Fig. 2.3, y para las cuales tiene sentido definir una *frecuencia instantánea única*. De manera general, a diferencia de las representaciones vistas en la Sec. 2.3, la técnica de EMD obtiene los modos directamente de la señal analizada:

$$x(t) \xrightarrow{\text{EMD}} \{x_k(t)\}.$$

### 2.4.1. Principio recursivo

La técnica de EMD descompone señales *oscilantes* de una forma completamente local y guiada por los datos. Damos aquí nuestra propia definición de señal oscilante.

**Definición 2.4.1. Señal oscilante.** Se dice que una señal  $x(t) \in \mathbb{R}$ , con  $x \in L^\infty(\mathbb{R})$  es oscilante si y sólo si se cumplen las siguientes condiciones:

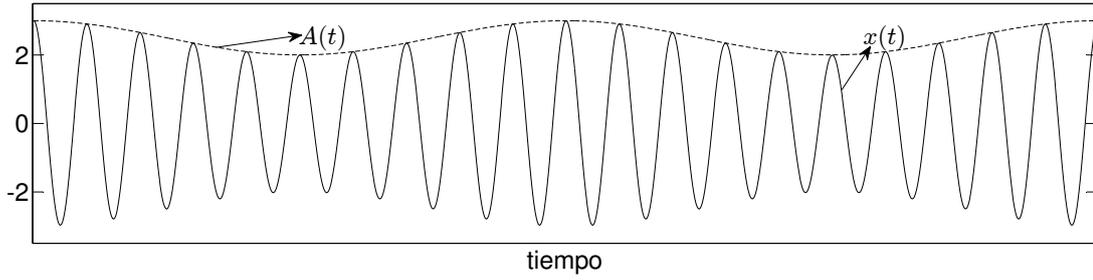


Figura 2.3: Combinación lineal de tres funciones coseno vistas como una única señal AM-FM. El modelo de la ec. (2.37) se interpreta mucho mejor como una señal de frecuencia  $\alpha$  suavemente modulada en amplitud cuando ocurre que  $\alpha \gg \beta$ . En este ejemplo consideramos  $\alpha = 40\pi$  y  $\beta = 4\pi$ .

- No existe un par  $\{t_0, \delta_0\}$  tal que  $x'(t)$  sea idénticamente nula para  $|t - t_0| < \delta_0$ .
- Existe un conjunto  $\mathbb{L}_x = \{t_\ell, \ell \in \mathbb{Z}\}$  de extremos locales para los cuales se cumple que existen  $\delta_\ell$  tales que  $x(t_\ell) > x(t)$  si  $0 < |t - t_\ell| < \delta_\ell$  o bien  $x(t_\ell) < x(t)$  si  $0 < |t - t_\ell| < \delta_\ell$ .
- $\#\mathbb{L}_x \geq 2$ .

Esta definición nos dice que para una señal oscilante no existen segmentos en los cuales la señal es constante, los extremos locales son aislados y existen al menos dos extremos locales.

El primer paso para describir cómo trabaja EMD consiste en considerar oscilaciones en la señal a un nivel local. Si se observa la evolución de una señal *oscilante*  $x(t)$  entre dos puntos extremos (por ejemplo entre dos mínimos consecutivos ubicados en  $t = a$  y  $t = b$ ) podemos definir heurísticamente una componente local de alta frecuencia o detalle local  $\mathcal{D}x_{[a,b]} = d_1(t)_{[a,b]}$  que corresponde a la oscilación entre los dos mínimos y que debe pasar necesariamente por un máximo existente entre ellos, donde  $\mathcal{D}$  es un operador que extrae de una señal oscilante la parte que oscila más rápidamente. Completando la idea, debemos identificar una componente de baja frecuencia o tendencia local  $a_1(t)_{[a,b]} = x_{[a,b]} - \mathcal{D}x_{[a,b]}$ . Si, de alguna forma, extendemos este análisis para todas las oscilaciones entre los extremos locales de la señal, podemos expresar para todo  $t$ :

$$x(t) = a_1(t) + d_1(t), \quad (2.38)$$

donde  $d_1(t)$  será la componente que oscile *localmente* más rápido, mientras que  $a_1(t)$  lo hará *localmente* más lento. En la Fig. 2.4 puede verse un ejemplo sencillo que ilustra este primer paso de la descomposición.

El proceso de descomposición continúa aplicando la misma idea a  $a_1(t)$ , esto es, extrayendo de ella una parte localmente más rápida  $\mathcal{D}a_1$  y una parte más lenta  $a_1 - \mathcal{D}a_1$ . Así se define entonces una descomposición deflacionaria, que opera recursivamente sobre sucesivos «residuos» parciales que van quedando al extraer uno a uno los modos:

$$\text{EMD}(x) = \{\mathcal{D}x, \text{EMD}(x - \mathcal{D}x)\}, \quad (2.39)$$

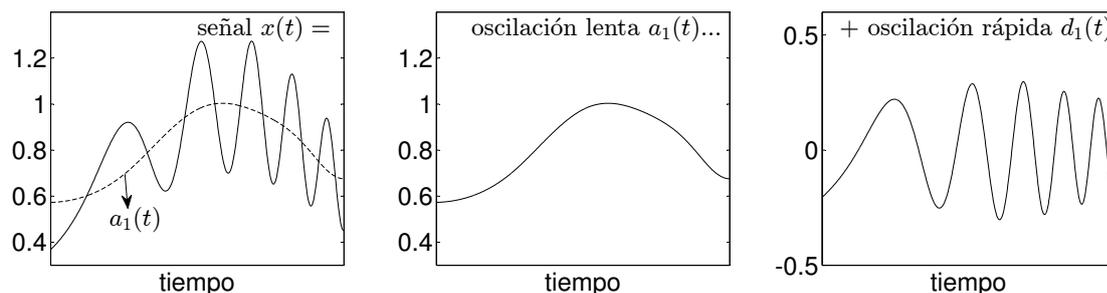


Figura 2.4: Descomposición de una señal en oscilaciones rápidas y lentas. La señal oscilante  $x(t)$  se descompone en la suma de su media local  $a_1(t)$  más el modo  $d_1(t)$ . El esquema deflacionario sigue con la descomposición de  $a_1(t)$ .

donde la descomposición se detiene cuando  $x - \mathcal{D}x$  no es más una señal oscilante (entonces ya no puede separarse en una parte localmente más rápida y una localmente más lenta).

Al finalizar, el esquema descrito produce una descomposición de la forma:

$$\begin{aligned}
 x(t) &= a_1(t) + d_1(t) \\
 &= a_2(t) + d_2(t) + d_1(t) \\
 &= \dots \\
 &= a_K(t) + \sum_{k=1}^K d_k(t),
 \end{aligned} \tag{2.40}$$

para la cual se cumple que

$$d_{k+1}(t) = \mathcal{D}a_k(t), \tag{2.41}$$

$$a_{k+1}(t) = a_k(t) - \mathcal{D}a_k(t), \tag{2.42}$$

donde usamos la convención  $a_0(t) = x(t)$ . La componente  $a_K(t)$ , que es lo que queda una vez que se extrajeron todos los componentes oscilantes, suele denominarse «residuo». En la práctica, la descomposición produce un número finito de componentes  $d_k(t)$ , aunque este resultado debe aún demostrarse teóricamente para el caso más general.

Para describir de manera más precisa a la técnica de EMD, conviene explicar aquí qué se entiende por «señal oscilante» y «la componente que oscila localmente más rápido». El primer concepto, ya definido en la Def. 2.4.1, es sencillo y corresponde a la idea intuitiva de oscilación: llamamos señal oscilante a toda aquella que es tanto creciente como decreciente. Podemos considerar entonces que una señal es oscilante si presenta al menos dos extremos locales; el caso de un solo extremo es más lejano a la idea intuitiva de oscilación.

El segundo concepto, el de componente que oscila localmente más rápido, es específico a EMD. Estos componentes, que oscilan simétricamente alrededor de cero y que son extraídos recursivamente como hemos visto, poseen una definición precisa.

**Definición 2.4.2. Función intrínseca de modo (IMF, del inglés Intrinsic mode function) [58].** Una señal real  $x \in L^\infty(\mathbb{R})$  es una IMF si y sólo si:

- (I) la cantidad de extremos locales y la cantidad de cruces por cero son iguales o difieren a lo sumo en uno; y
- (II) la media entre la envolvente superior definida por los máximos locales y la envolvente inferior definida por los mínimos locales es cero.

Hemos dado la definición de IMF tal y como la presentan Huang y col. [58]. Sin embargo puede apreciarse que la segunda condición, en exceso restrictiva, no puede satisfacerse sin la primera. Representan, entonces, no dos condiciones independientes entre sí, sino que serían una forma *débil* (la primera) y una forma *fuerte* (la segunda) de definir una IMF [102]. Rilling modifica la primera condición, enunciándola de la siguiente manera: todos los mínimos locales son estrictamente negativos, y todos los máximos locales son estrictamente positivos [95]. Esta forma, completamente equivalente a la dada por Huang, enfatiza la naturaleza *local* de la definición de IMF, en contraposición de una característica *global* usada en la definición original. Siendo la EMD un método que trabaja a nivel local (a través de los extremos locales), creemos más adecuada la definición de Rilling.

El cuadro general presentado para EMD, resumido en las ecuaciones (2.39) y (2.40), no difiere sustancialmente del marco de la transformada ondita discreta (diádica) [76]. Éste es en realidad llevado a la práctica de manera recursiva: la señal de estudio se descompone en «detalle» y «aproximación», los cuales se corresponden, respectivamente, a una parte de alta frecuencia y a una parte de baja frecuencia. Luego la misma descomposición se aplica a la «aproximación». La diferencia entre la transformada ondita discreta y EMD se halla en la manera en que se realiza la descomposición en «detalle» más «aproximación», u «oscilaciones rápidas» más «oscilaciones lentas». En la transformada ondita discreta, los dos componentes son calculados por una operación predefinida de filtrado lineal e invariante en el tiempo donde la respuesta en frecuencia se contrae en un factor igual a 2. Al contrario, y como se verá más adelante, las características de la descomposición mediante EMD son determinadas de manera local y a partir de la señal.

### 2.4.2. Proceso de *sifting*

Hemos visto cómo EMD descompone las señales de manera recursiva, sustrayendo sucesivamente medias locales. Pero si dijimos que la frecuencia instantánea era algo así como un oxímoron, no podemos decir menos de la media local. Sin embargo, al asimilar la idea de oscilaciones lentas con la media local, podríamos hablar más apropiadamente de una «media a escala local», sobre la cual están «montadas» las oscilaciones rápidas. Siguiendo ese razonamiento, puede definirse a la media local como el resultado de una operación de la forma

$$a(t) = \frac{1}{\alpha(t)} \int x(u) w \left( \frac{u-t}{\alpha(t)} \right) du, \quad (2.43)$$

con  $w(u)$  una función de ventaneo (apropiadamente normalizada) que enfatiza a la señal en torno a cero y  $\alpha(t)$  la «escala local» sobre la que se «promedia» a la señal para obtener su media local. Proponer una función de ventaneo no es difícil, pero la selección de la escala local presenta serias dificultades, y no existe una solución simple para el caso general [95].

Tal es la situación, que los creadores de EMD propusieron definir la media local de manera geométrica con la ayuda de la noción intuitiva de envolventes, y usando a los extremos locales como indicadores de la escala local [58, 66]. Más precisamente, definieron a la envolvente superior como una curva suave que interpola los máximos locales, y a la envolvente inferior como una curva suave que hace lo propio con los mínimos locales. En la práctica, la interpolación se realiza mediante *splines* cúbicos, pero esta no es la única opción disponible, como veremos más adelante. Una vez determinadas las envolventes, la media local se define simplemente como la semisuma de ellas. De esta forma, la parte que oscila localmente más rápido se encuentra sustrayendo a la señal la semisuma de sus envolventes, lo cual define un operador elemental:

**Definición 2.4.3. Operador elemental de sifting.** Dada una señal oscilante  $x(t)$ , se define el operador elemental de sifting  $\mathcal{S}$  como

$$(\mathcal{S}x)(t) = x(t) - \frac{e_{max}(t) + e_{min}(t)}{2}, \quad (2.44)$$

donde  $e_{max}(t)$  y  $e_{min}(t)$  representan, respectivamente, a las envolventes superior e inferior.

Esta acción, sin embargo, no garantiza que la señal resultante satisfaga las condiciones de IMF (Def. 2.4.2) con lo que el proceso debe *iterarse*. Tal es así que se realizan sustracciones de media local hasta que se satisfagan (al menos hasta una tolerancia dada) las condiciones de IMF. Queda entonces definido el proceso de *sifting* que permite definir a la oscilación rápida como el resultado de la aplicación sucesiva del operador elemental  $\mathcal{S}$ :

**Definición 2.4.4. Proceso de sifting.** Dada una señal oscilante  $x(t)$ , se define al proceso de sifting como la aplicación sucesiva de  $n$  veces el operador elemental de sifting (Def. 2.4.3)

$$(\mathcal{D}x)(t) = (\mathcal{S}^n x)(t), \quad (2.45)$$

donde  $n$  es la cantidad de iteraciones necesarias para satisfacer las condiciones de IMF, la cual es determinada mediante algún criterio de detención (Sec. 2.4.7.2).

Con estas definiciones, estamos en condiciones de definir el algoritmo de EMD:

---

**Algoritmo 2.1.** Descomposición empírica en modos (EMD)

---

- 1: **Inicialización:** asignar  $r = x$  (señal);  $k = 1$ ;  $j = 1$ ;
- 2: encontrar máximos locales de  $r$  e interpolar obteniendo  $e_{máx}$ ;
- 3: encontrar mínimos locales de  $r$  e interpolar obteniendo  $e_{mín}$ ;
- 4: estimar la media local como  $m = \frac{e_{máx} + e_{mín}}{2}$ ;
- 5:  $d_k^j = r - m$ ;
- 6: **if**  $d_k^j$  es IMF **then**
- 7:   asignar  $d_k = d_k^j$  como modo y guardarlo;
- 8:    $a_k = x - \sum_{n=1}^k d_n$ ;
- 9:    $r = a_k$ ;
- 10:  $k = k + 1$ ;  $j = 1$ ; e ir al paso 2;

```

11: else
12:    $r = d_k^j$ ;
13:    $j = j + 1$ ; e ir al paso 2;
14: end if

```

---

El proceso continúa hasta que  $k$  alcanza un número prefijado de modos o hasta que ya no es posible hallar las envolventes debido a que el residuo ya no cumple con la definición de señal oscilante (Def. 2.4.1). Con el fin de ofrecer una mejor explicación del algoritmo, presentamos en la Fig. 2.5 un diagrama de flujos.

### 2.4.3. Espectro de Hilbert-Huang

En la formulación original de EMD sus autores perseguían la misma premisa con la que comenzamos este capítulo: dada una señal multicomponente, calcular y/o describir las propiedades de las frecuencias y amplitudes instantáneas de cada uno de sus componentes. En este contexto, la descomposición mediante EMD se considera tan sólo el primer paso de un análisis más completo denominado transformada de Hilbert-Huang (HHT, del inglés *Hilbert-Huang Transform*). Dicho análisis consiste en:

1. dada la señal  $x(t)$ , aplicar EMD:  $x(t) \xrightarrow{EMD} \{d_k(t), k = 1, \dots, K\} \cup \{a_K(t)\}$ ,
2. para cada modo  $d_k(t)$ , calcular su frecuencia instantánea  $f_k(t)$  y su amplitud instantánea  $A_k(t)$  mediante el método de la señal analítica (ver Sec. 2.2.1).

De esta forma es posible construir una representación tiempo-frecuencia denominada el espectro de Hilbert-Huang:

$$H_x(t, f) = \sum_{k=1}^K A_k(t) \delta(f - f_k(t)). \quad (2.46)$$

Esta representación superaría al fin los límites del principio de incertidumbre, logrando una resolución perfecta tanto en tiempo como en frecuencia (cada componente aparece en la representación como una línea algebraica). Sin embargo debemos recordar las limitaciones impuestas por los teoremas de Bedrosian y Nuttall respecto a la certeza de que las expresiones  $A_k$  y  $f_k$  coincidan con las verdaderas amplitud y frecuencia instantáneas del componente respectivamente.

A partir del espectro de Hilbert-Huang, puede definirse su espectro marginal:

$$h_x(f) = \int H_x(t, f) dt. \quad (2.47)$$

Este espectro marginal ofrece una medida de la contribución total en amplitud para cada valor de frecuencia, representando la amplitud acumulada a lo largo de la señal en un sentido probabilístico [58].

### 2.4.4. Definición más general de EMD

La técnica de EMD, reducida a sus principios más elementales, se distingue por dos características: la selección de una escala basada en los extremos locales; y

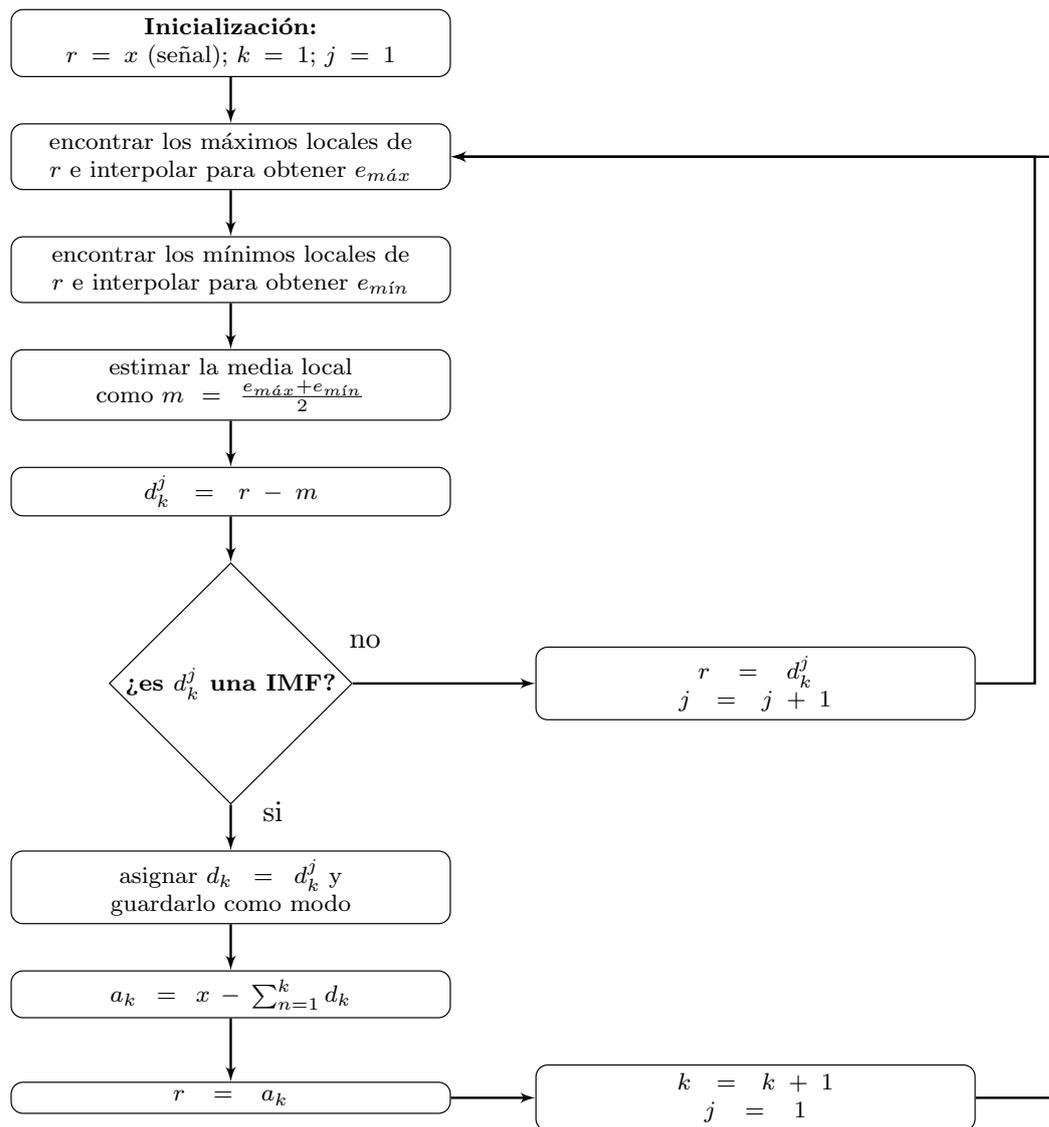


Figura 2.5: Diagrama de flujos del algoritmo de EMD. La extracción de los modos se realiza de manera deflacionaria, sustrayendo sucesivas estimaciones de medias locales. El bucle interno es el correspondiente al proceso de *sifting* (Def. 2.4.4). Se usan los contadores  $k$  para los modos y  $j$  para las iteraciones necesarias para la extracción de cada modo.

la extracción sucesiva de componentes de las escalas más finas a las más gruesas. Además de los extremos locales, se podría considerar un conjunto de puntos característicos que también puede incluir a los puntos de inflexión o los extremos de curvatura.

Una formulación más general debe retener estos elementos característicos, manteniendo el enfoque recursivo presentado anteriormente:  $\text{EMD}(x) = \{\mathcal{D}x, \text{EMD}(x - \mathcal{D}x)\}$ . Si  $\mathcal{D}x$  corresponde a la escala más pequeña de la señal (siempre a un nivel

local), entonces el enfoque recursivo lleva a la práctica el hecho de extraer las escalas de la más fina a la más gruesa.

La cuestión crucial es la definición del operador  $\mathcal{D}$  de una forma distinta a como fue presentado anteriormente. Para ello consideramos un conjunto de puntos característicos de la señal  $\{t_\ell \in \mathbb{L}\}$  correspondientes, por ejemplo, a los extremos locales de la señal. Entonces podemos definir a la versión muestreada de la señal sobre ese conjunto de puntos

$$x_{\mathbb{L}}(t) = \sum_{t_\ell \in \mathbb{L}} x(t_\ell) \delta(t - t_\ell), \quad (2.48)$$

y en base a ella definir la media local de la señal mediante una operación de «filtrado generalizado»

$$a[x](t) = (h \tilde{\star} x_{\mathbb{L}})(t), \quad (2.49)$$

donde  $h$  representa a un filtro pasa-bajos generalizado, « $\tilde{\star}$ » representa una convolución generalizada, y hemos enfatizado que la media local corresponde a la señal  $x$  con la notación  $a[x](t)$  [95]. Es importante remarcar aquí que el filtro pasa-bajos generalizado no actúa como una filtro lineal e independiente de la señal muestreada  $x_{\mathbb{L}}(t)$ , sino que realiza más bien un filtrado variante en el tiempo y dependiente de la señal sobre el cual volveremos más adelante. Esta operación de filtrado generalizado goza de un número de propiedades de invarianza que son típicamente verificadas por una operación de interpolación:

- **Multiplicación por un escalar:**

$$\forall \alpha \in \mathbb{R}, \hat{x}_{\mathbb{L}}(t) = \alpha x_{\mathbb{L}}(t), \Rightarrow (h \tilde{\star} \hat{x}_{\mathbb{L}})(t) = \alpha (h \tilde{\star} x_{\mathbb{L}})(t).$$

- **Dilatación/compresión/inversión temporal:**

$$\forall \alpha \in \mathbb{R}, \hat{x}_{\mathbb{L}}(t) = x_{\mathbb{L}}(\alpha t), \Rightarrow (h \tilde{\star} \hat{x}_{\mathbb{L}})(t) = (h \tilde{\star} x_{\mathbb{L}})(\alpha t).$$

- **Traslación temporal:**

$$\forall \alpha \in \mathbb{R}, \hat{x}_{\mathbb{L}}(t) = x_{\mathbb{L}}(\alpha + t), \Rightarrow (h \tilde{\star} \hat{x}_{\mathbb{L}})(t) = (h \tilde{\star} x_{\mathbb{L}})(\alpha + t).$$

- **Adición de una constante:**

$$\forall \alpha \in \mathbb{R}, \hat{x}_{\mathbb{L}}(t) = \alpha + x_{\mathbb{L}}(t), \Rightarrow (h \tilde{\star} \hat{x}_{\mathbb{L}})(t) = \alpha + (h \tilde{\star} x_{\mathbb{L}})(t).$$

Definiendo así la media local  $a(t)$ , podemos definir  $\mathcal{D}x$  como

$$(\mathcal{D}x)(t) = x(t) - a(t) \quad (2.50)$$

$$= x(t) - (h \tilde{\star} x_{\mathbb{L}})(t). \quad (2.51)$$

Según esta formulación, se puede definir a  $\mathcal{D}x$  de forma directa sin un proceso iterativo, contrariamente a la formulación original de EMD. Esto puede explicarse notando que el proceso de *sifting* de EMD cumple esencialmente dos funciones: hacer aparecer nuevos extremos que están ocultos debido a la superposición de escalas, y hacer más simétricas las envolventes [97]. Puede verse que la aparición de nuevos

extremos se concilia con esta formulación más general a través de una mejor selección de los puntos característicos  $\{t_\ell \in \mathbb{L}\}$ . De igual manera, el mejoramiento en la simetría de las envolventes se condice con una mejor definición para la media local.

Esta formulación más general parte de dos premisas: prescindir de las envolventes para la definición de la media local y evitar el proceso de *sifting*. El primer punto es necesario ya que no existe un consenso respecto a *qué es una envolvente* o *qué constituye una envolvente*. Además, para una señal de la forma  $A(t) \cos(2\pi\phi(t))$ , la interpolación a través de los extremos locales *nunca* devolverá  $A(t)$  a menos que ésta sea una constante [52]. Numerosas propuestas calculan la media local sin hacer un uso explícito de las envolventes. En [34] estiman la media local interpolando los puntos de inflexión. En [51] la interpolación se hace sobre unos puntos particulares estimados mediante integrales locales. En [73] se usan filtros de media móvil para la estimación de la media local. El método de *B-spline EMD* [23] estima la media local como una suma de *B-splines* cuyos nodos coinciden con el conjunto de extremos locales de la señal. Los coeficientes de la suma de *B-splines* se obtienen como medias pesadas de la señal en los extremos que caen en el soporte del *B-spline*.

El segundo punto, evitar el proceso de *sifting*, es necesario porque su convergencia no está teóricamente garantizada [110], aunque se cumple en la mayoría de los casos prácticos.

## 2.4.5. Propiedades de la descomposición

### 2.4.5.1. No linealidad

El algoritmo de EMD es globalmente no lineal. Esto puede verse fácilmente en el hecho que la descomposición mediante EMD de una suma de dos señales es en general diferente de la suma de las descomposiciones por EMD de las señales por separado:

- La cantidad de IMFs de la suma no depende de la cantidad de IMFs de las señales por separado.
- Una IMF de la suma no se describe en términos de la suma de un conjunto de IMFs de las señales por separado ni por una combinación lineal de éstas.

Existen tres causas de no linealidad. La primera, y más importante, es que el método se basa en los extremos locales. La cantidad y posición de los extremos locales de una suma de señales son en general distintos de aquellos de las señales individuales. Esta causa de no linealidad es fundamental en la EMD en la medida en que las escalas de variaciones rápidas de una señal están definidas a través de sus extremos locales.

Las otras dos causas de no linealidad son el esquema de interpolación y el criterio de detención para el proceso de *sifting*, que pueden ser no lineales según su implementación. En la práctica, la interpolación por lejos más usada es la *spline* cúbica, la cual es lineal, pero la casi totalidad de los criterios de detención propuestos son susceptibles de generar no linealidades haciendo variar la cantidad de iteraciones [95].

### 2.4.5.2. (Cuasi)ortogonalidad

Mucho se ha dicho respecto a la ortogonalidad de la EMD. En su formulación original [58] se sugiere que la descomposición en oscilaciones rápidas y lentas la haría (cuasi)ortogonal. El razonamiento consiste en asimilar la parte de oscilaciones lentas a la componente de continua (en un sentido de Fourier) de la señal descompuesta. La parte de oscilaciones rápidas entonces carecerían (localmente) de componente continua, siendo necesariamente (de forma local) ortogonales a la parte de oscilaciones lentas.

Según este razonamiento, dos aspectos son susceptibles de impedir la ortogonalidad de la descomposición:

- La ortogonalidad es una propiedad global, no local. La no estacionariedad puede comprometer la ortogonalidad global aunque aparentemente esté asegurada a una escala local.
- La media de las envolventes de una señal no coincide necesariamente con la media en un sentido integral, ni siquiera localmente.

Si bien es claro que el primer punto impide la ortogonalidad global de la descomposición, esto no trae necesariamente consecuencias negativas ya que el inconveniente es que la noción de ortogonalidad global no es adecuada para un método que analiza localmente a las señales. La noción de ortogonalidad local, en cambio, presenta problemas de definición, lo que explica que en la práctica se use la ortogonalidad global.

Una forma de medir la ortogonalidad entre dos IMFs es a través un índice de ortogonalidad definido como el producto escalar normalizado (coeficiente de correlación)

$$IO_{ij} = \frac{\langle d_i(t), d_j(t) \rangle}{\sqrt{\langle d_i(t), d_i(t) \rangle \langle d_j(t), d_j(t) \rangle}}. \quad (2.52)$$

En términos globales (esto es, para toda la descomposición), la ortogonalidad puede describirse mediante una matriz de ortogonalidad  $IO$  de elementos  $IO_{ij}$ . Sin embargo, esta representación presenta la dificultad que da la misma importancia a todas las IMFs, sin tener en cuenta su energía. Por esta razón, generalmente se prefiere medir la ortogonalidad de la descomposición sin normalizar los productos escalares, o normalizarlos a todos por la misma cantidad, por ejemplo la norma  $L^2$  de la señal analizada. De esta manera, se puede definir un único índice de ortogonalidad de la descomposición como [58, 95]:

$$IO = \frac{\sum_{i < j} |\langle d_i(t), d_j(t) \rangle|}{\langle x(t), x(t) \rangle}. \quad (2.53)$$

### 2.4.5.3. Localidad

En el marco de EMD, la escala local está definida por los extremos locales. En la formulación original se afirma que «la escala temporal característica se define como la distancia temporal entre extremos» [58]. Para ver esto mejor, es suficiente recordar que la media local se define a partir de las envolventes que interpolan los

máximos y los mínimos locales. La idea de envolvente es estrictamente local debido a que el valor, por ejemplo de la envolvente superior, en un punto dado depende esencialmente de los dos máximos locales más próximos y, en menor medida, de aquellos más alejados. De ahí que podemos considerar que las envolventes, y por lo tanto la media local, se definen localmente mediante la *escala correspondiente a la distancia entre los máximos/mínimos*, es decir *grosso modo* dos veces la escala correspondiente a la distancia entre extremos (máximos y mínimos en conjunto).

#### 2.4.5.4. Multirresolución

La técnica de EMD realiza una descomposición multiescala, o multirresolución, ya que explora sucesivamente las escalas de una señal de la más fina, representada por el primer modo, a la más gruesa, representada por el último modo o por el residuo.

Si comparamos a EMD con otros métodos de análisis multirresolución, por ejemplo la transformada ondita, veremos que la primera posee ciertas particularidades. En primer lugar, la transformada ondita discreta propone una descomposición en escalas discretas ya que la descomposición consta de un número finito de componentes [95]. En cambio, las escalas de EMD difieren significativamente de las de la ondita discreta, y se distinguen por las siguientes características:

- **Definición por los extremos locales.** La idea de escala en EMD está asociada a la distancia entre los extremos. Esta idea difiere notablemente de la idea de escala que se tiene en la transformada ondita, donde la escala está definida de manera relativa por comparación con una forma de onda dada (la ondita). El concepto de escala en EMD es tan distinto que se ha propuesto usar en lugar de «escala» el término «*empiquency*» [74], del inglés *empirical mode frequency*, que corresponde a la inversa de la escala y que está definido como

$$f_e = \frac{1}{2d}, \quad (2.54)$$

donde  $d$  es la distancia entre los dos extremos más cercanos al punto en cuestión. El enfoque de EMD permite asociar una misma escala a formas de onda distintas.

- **Adaptividad.** Las escalas de las IMFs se determinan mediante las escalas presentes en la señal analizada y no mediante una grilla predeterminada como en el caso de la transformada ondita discreta. En este último caso, las escalas son de la forma  $a2^k$ ,  $k \in \mathbb{N}$ , donde  $a$  es una escala de referencia.
- **Localidad.** En virtud de lo discutido en la Sec. 2.4.5.3, las escalas de una IMF no se definen de manera global sino local, estando el carácter local relacionado a la distancia entre los extremos locales.

La adaptividad y la localidad son asimismo una causa de dificultad para la EMD conocida como «mezcla de modos» (en inglés «*mode mixing*»). Este problema, una consecuencia de la naturaleza local y adaptativa del método, se manifiesta especialmente cuando la señal posee componentes *intermitentes*, es decir, cuya duración es menor a la duración de la señal. Volveremos a discutir este fenómeno en el Cap. 5.

### 2.4.6. Funciones intrínsecas de modo

Las funciones básicas (de algún modo *atómicas*) en las que EMD descompone a una señal son las funciones intrínsecas de modo, IMFs, que fueron definidas oportunamente (ver Def. 2.4.2). Dicha definición, aunque restrictiva (su segunda condición debe relajarse en la práctica) define una clase de funciones que es muy amplia, y sin embargo carece de la precisión suficiente. Ello se debe a que la definición persigue un único objetivo: que la IMF posea un único valor de frecuencia instantánea, esto es, que no sufra de las paradojas que menciona Cohen [25].

Sin embargo, la definición de IMF no puede separarse de la cuestión de su implementación práctica: al definirse las envolventes mediante interpolación, éstas dependerán del esquema elegido. Como se mencionó anteriormente, la opción más habitual son los *splines* cúbicos. En ese caso puede demostrarse que la única forma de satisfacer la condición de media local cero es cuando las envolventes superior e inferior son ambas iguales a un único polinomio cúbico (sólo difieren en signo) [95]. En el caso de *splines* naturales (derivada segunda igual a cero en los bordes), se tiene que sólo una envolvente lineal permitirá satisfacer las condiciones de IMF [112, 119]. Además, y como consecuencia de la definición de las envolventes, todos los modos salvo el primero son sumas de interpolaciones. Tal conclusión atenta contra la interpretación que hacemos de las IMFs. Tomemos como ejemplo una señal senoidal modulada en amplitud por una función lineal:

$$x(t) = (at + b) \sin(2\pi t), \quad (2.55)$$

con  $a, b \in \mathbb{R}$  constantes. La señal  $x(t)$  debería ser a todas luces una IMF, con envolvente superior igual a  $at + b$  y con una única frecuencia instantánea igual 1 para todo  $t$ . Sin embargo sus extremos locales no se encuentran sobre la recta  $at + b$  que modula en amplitud a la señal. Si  $x(t_0)$  es un extremo local, tenemos:

$$\begin{aligned} x'(t_0) &= a \sin(2\pi t_0) + 2\pi(at_0 + b) \cos(2\pi t_0) = 0 \\ a \sin(2\pi t_0) &= -2\pi(at_0 + b) \cos(2\pi t_0) \\ t_0 + \frac{b}{a} &= -\frac{1}{2\pi} \tan(2\pi t_0) \\ t_0 + \frac{1}{2\pi} \tan(2\pi t_0) &= -\frac{b}{a}. \end{aligned} \quad (2.56)$$

La solución puede hallarse numéricamente. Para que los extremos estén sobre la recta  $at + b$  debe darse  $\cos(2\pi t_0) = 1$ , y eso es imposible debido a que los  $\{t_0\}$  no están equiespaciados. Entonces podemos confirmar que los extremos no se ubican sobre la recta  $at + b$ , ni sobre ninguna otra recta. Siguen más bien una curva que tiende a una recta sólo para  $t \rightarrow \infty$ . Por lo tanto, una señal tan sencilla como  $x(t)$  dada por la ec. (2.55) no es estrictamente una IMF.

Una forma de aliviar este problema para las IMFs es redefinir las envolventes. En [52] se propone que la envolvente no pase por los extremos locales sino que sea tangente a la señal, en puntos distintos a los extremos locales, para lo cual los autores dan una condición que debe cumplir la fase de una señal AM-FM de la forma  $A(t) \cos(2\pi\phi(t))$ . Esto lleva a una nueva definición de la media local, y en definitiva a una nueva definición del proceso de *sifting*. Para más detalles referimos al lector a [52].

Otra dirección han tomado Daubechies y col. [31]. En su trabajo, proponen emular a EMD, pero desde una perspectiva del análisis con CWT. Para ello definieron de forma mucho más precisa una nueva clase de funciones: las funciones tipo intrínsecas de modo (funciones IMT, del inglés *Intrinsic Mode Type functions*), las cuales son el resultado de la descomposición propuesta.

**Definición 2.4.5. Función tipo intrínseca de modo (IMT, del inglés Intrinsic Mode Type)<sup>2</sup> [31].** Una señal continua  $x(t) \in \mathbb{R}$ ,  $x \in L^\infty(\mathbb{R})$ , es una IMT con precisión  $\epsilon > 0$  si  $x(t) = A(t) \cos(2\pi\phi(t))$ , con  $A$  y  $\phi$  que tienen las siguientes propiedades:

$$\begin{aligned} A &\in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \phi \in C^2(\mathbb{R}) \\ \inf_{t \in \mathbb{R}} A(t) &> 0, \inf_{t \in \mathbb{R}} \phi'(t) > 0, \sup_{t \in \mathbb{R}} \phi'(t) < \infty \\ |A'(t)|, |\phi''(t)| &\leq \epsilon \phi'(t), \quad \forall t \in \mathbb{R} \\ \sup_{t \in \mathbb{R}} |\phi''(t)| &< \infty \end{aligned}$$

En este caso se introduce el parámetro de «calidad»  $\epsilon$ . La definición por lo tanto ya no está dada en términos binarios, sino que distintas funciones serán IMT con distintas precisiones  $\epsilon$ . Incluso en [78] se las denomina  $\epsilon$ -IMT, incorporando el factor de precisión a la clase. De esta forma, si reescribimos a la señal de la ec. (2.55) como

$$x(t) = (at + b) \cos(2\pi t - \pi/2), \quad (2.57)$$

podemos ver que  $x(t)$  es una IMT con  $\epsilon = a$ . La precisión de la IMT mejora con la disminución de  $\epsilon$ , de forma que cuanto menor sean las modulaciones en amplitud y en frecuencia, comparados con la frecuencia instantánea, la IMT es más «precisa». Esto está de acuerdo con la idea que planteáramos al comienzo del capítulo: las señales *monocomponentes* que forman la señal compuesta están moduladas *suavemente* en amplitud y frecuencia.

En el caso particular de la función de la ec. (2.57), disminuir  $\epsilon$  implica disminuir la pendiente de la modulación en amplitud, lo que a su vez está de acuerdo con los teoremas de Bedrosian y Nuttall (ver Teoremas 2.2.1 y 2.2.2). El caso ideal,  $\epsilon = 0$ , se logra en ausencia de modulación en amplitud, con lo que se resuelve la dificultad planteada por Bedrosian. Si en cambio se mantiene la modulación en amplitud, el factor  $\epsilon$  puede disminuirse al aumentar la frecuencia de la señal. El teorema de Nuttall nos muestra claramente cómo ocurre esto: al considerar la versión compleja en cuadratura  $z_x^q(t) = (at + b)e^{i(2\pi t - \frac{\pi}{2})}$ , un aumento en la frecuencia «corre» a la derecha su espectro, acercando más a la señal compleja en cuadratura a la versión analítica. Si bien las clases IMF e IMT no coinciden, tienen similitudes. Una 0-IMT es claramente una IMF, pero en general no toda IMF puede ser expresada como una IMT para un  $\epsilon$  dado.

Esto último puede verse para una señal de la forma

$$x(t) = \cos(2\pi t) + \frac{1}{2} \sin(4\pi t), \quad (2.58)$$

<sup>2</sup>La definición dada en [31] es para señales complejas. En nuestro caso estamos interesados en señales reales, por lo que modificamos levemente la definición que, sin embargo, mantiene toda su esencia.

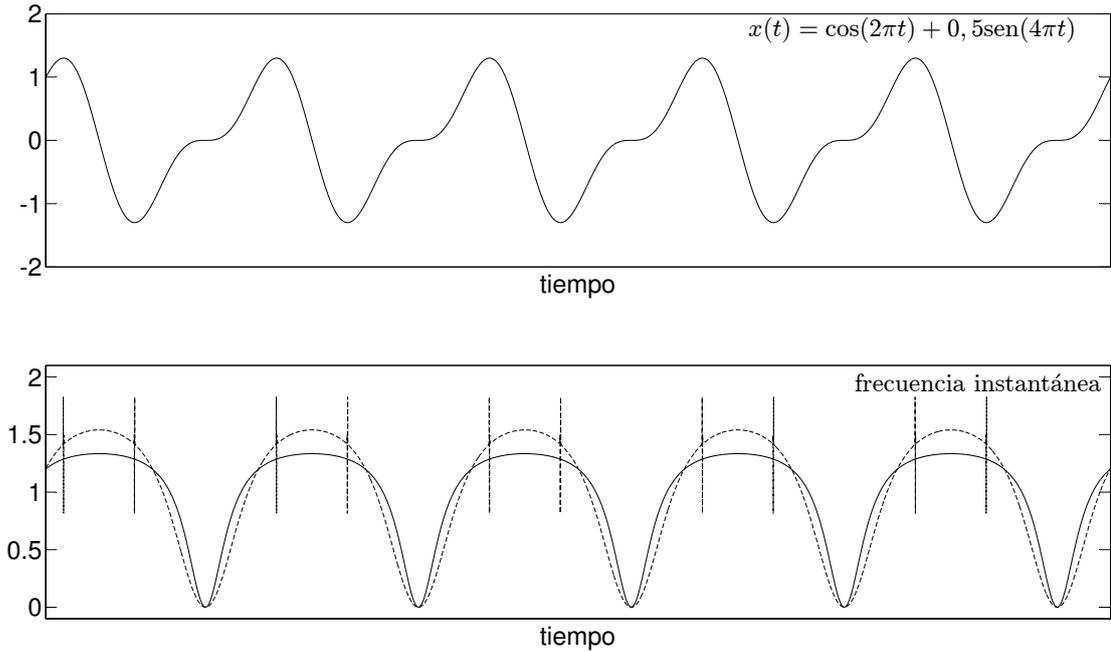


Figura 2.6: Señal compuesta por un coseno y un seno y estimaciones de su frecuencia instantánea. La señal de la ec. (2.58) es claramente una IMF y sin embargo las estimaciones de sus frecuencias instantáneas distan de ser satisfactorias. En línea sólida puede verse la estimación mediante transformada de Hilbert; en línea de trazos mostramos la estimación usando la función  $\arccos(\cdot)$ .

la cual puede apreciarse en la Fig. 2.6. Si bien  $x(t)$  es evidentemente una IMF, su frecuencia instantánea (obtenida con el método de la señal analítica) está lejos de ser satisfactoria. Además, si escribimos  $x(t) = A \cos(2\pi\phi(t))$ , con  $A \in \mathbb{R}$ , de forma de verla como una IMT, tendríamos una fase  $\phi(t) = \frac{1}{2\pi} \arccos\left(\frac{x(t)}{A}\right)$ , lo cual puede llevar a valores negativos de la frecuencia instantánea sin cumplir siquiera con la definición de IMT (Def. 2.4.5), si no se tiene el cuidado de corregir las discontinuidades que produce la función  $\arccos$ . Casos similares a este, donde se presentan IMFs «patológicas», fueron estudiados en [102]. Una propuesta para superar esta dificultad fue la de Hau-Tieng Wu [113], donde, en el marco de una descomposición basada en la CWT, se propone una señal multicomponente más general, de la forma

$$x(t) = \sum_{k=1}^K x_k(t) = \sum_{k=1}^K A_k(t) s(2\pi\phi_k(t)), \quad \text{con } A_k(t), \phi'_k(t) > 0 \forall t. \quad (2.59)$$

Este caso reemplaza la función coseno por una forma de oscilación más general  $s(t)$  periódica de período  $2\pi$ , con representación en serie de Fourier  $s(t) = \sum_k \hat{s}(k)e^{ikt}$ . Este nivel extra de generalidad permite reflejar mejor las funciones que entrega EMD a su salida, las cuales ya no están asociadas a funciones circulares y permiten representar de una mejor manera la naturaleza no lineal de los sistemas que generan las señales analizadas [38]. Sin embargo también trae aparejado un nivel extra de dificultad: a la estimación de las frecuencia y amplitud instantáneas se suma la estimación de la *forma de onda*  $s(t)$ .

**Definición 2.4.6. Clase de señales de forma de onda**<sup>3</sup> [113]. Dados  $\delta \geq 0$ ,  $D \in \mathbb{N}$  y  $\theta \geq 0$ , se define la clase de señales de forma de onda  $\mathbb{S}^{\delta,D,\theta}$  como aquella que consiste de señales periódicas  $s(t)$  de media cero, es decir con  $\hat{s}(0) = 0$ , y norma  $L^2$  unitaria que satisfacen:

- $\forall k \in \mathbb{Z}$ , con  $|k| \neq 1$ , se cumple  $|\hat{s}(k)| \leq \delta |\hat{s}(1)|$ .
- $\sum_{|n|>D} |n\hat{s}(n)| \leq \theta$ .

La primera condición nos dice que la amplitud de la frecuencia «base» o «fundamental»  $\hat{s}(1)$  no puede ser cero y que es su modulación la que será definida como la frecuencia instantánea  $\phi'(t)$ . La segunda condición nos dice que esencialmente la forma de onda no oscila demasiado rápido [113].

Al modular en amplitud y en frecuencia las formas de onda  $s(t)$  se obtienen las señales componentes del modelo de la ec. (2.59). La similitud con las funciones intrínsecas de modo de EMD es tal que Hau-Tieng Wu llama a sus señales de la misma forma.

**Definición 2.4.7. Clase de funciones intrínsecas de modo de Hau-Tieng Wu** [113]. Dados  $\epsilon, \delta, D > 0$ , con  $\epsilon \ll 1$ , se define el espacio  $\mathbb{A}_\epsilon^{\delta,D}$  de funciones intrínsecas de modo a funciones  $x(t) \in \mathbb{R}$ , con  $x \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$  que tienen la forma  $x(t) = A(t)s(2\pi\phi(t))$ , donde  $s \in \mathbb{S}^{\delta,D,\theta}$ , con  $A$  y  $\phi$  con las siguiente propiedades:

$$\begin{aligned} A &\in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \phi \in C^2(\mathbb{R}) \\ \inf_{t \in \mathbb{R}} A(t) &> 0, \inf_{t \in \mathbb{R}} \phi'(t) > 0, \sup_{t \in \mathbb{R}} \phi'(t) < \infty \\ |A'(t)|, |\phi''(t)| &\leq \epsilon \phi'(t), \quad \forall t \in \mathbb{R} \\ \sup_{t \in \mathbb{R}} |\phi''(t)| &< \infty \end{aligned}$$

Las condiciones que cumplen la amplitud  $A$  y la fase instantánea  $\phi$  son idénticas a las de las funciones IMT. Sin embargo, este modelo más compacto permite que  $\phi'(t)$  sea la frecuencia fundamental instantánea, y que el «espectro instantáneo» esté compuesto de frecuencias  $\{k\phi'(t), k \in \mathbb{Z}\}$ . Este enfoque merece ser estudiado con mayor profundidad para establecer relaciones con el espectro evolutivo de Priestley [91].

A la luz de esta nueva definición, la señal  $x(t)$  de la ec. (2.58) pertenece a la clase  $\mathbb{S}^{\frac{1}{2},2,0}$ , con  $\phi'(t) = 1$ . En su representación en un plano tiempo frecuencia aparecerá además una segunda línea en  $2\phi'(t)$ .

Podemos ver cómo una señal  $s(t)$  que pertenece a la clase  $\mathbb{S}^{\delta,D,\theta}$  será una IMF en el sentido de Huang, salvo por una constante, si posee sólo un máximo local y un mínimo local por período. De esa forma,  $d(t) = s(t) + C_s$  será una IMF, con  $C_s = (\max(s) + \min(s))/2$ . La modulación del argumento de  $s$ , esto es escribirla como  $s(2\pi\phi(t))$ , no alterará su condición de IMF si se cumplen las condiciones de la Def. 2.4.7. La modulación en amplitud en cambio, impedirá la satisfacer las condiciones de IMF de Huang (Def. 2.4.2), que como vimos es una definición muy restrictiva.

<sup>3</sup>Al igual que con las funciones IMT, la definición dada en [113] es para señales complejas. Damos aquí una definición equivalente para señales reales.

A pesar de la versatilidad del modelo de la ec. (2.59) su utilidad es aún limitada. En la propuesta original [113] se supone conocida la forma de onda  $s(t)$ , y las demostraciones se realizan bajo la hipótesis de que ésta está suficientemente cerca de un tono puro (el cual pertenece a  $\mathbb{S}^{0,1,0}$ ). Esta situación da aún más valor a la técnica de EMD, la cual no realiza tantas suposiciones sobre los datos, y entrega modos que oscilan simétricamente alrededor de cero, pero no son necesariamente señales circulares, pudiendo así reflejar las no-linealidades de los sistemas involucrados en la producción de las señales.

Para finalizar esta subsección, damos una definición general propia de lo que entendemos debiera ser una IMF. La definición de Hau-Tieng Wu, si bien muy general, aún no abarca todas las posibilidades para una IMF, ya que la forma de onda se mantiene para toda la duración de la señal.

**Definición 2.4.8. Definición general de IMF.** Una señal  $x(t) \in \mathbb{R}$ , con  $x \in L^\infty(\mathbb{R})$ , es una IMF si puede escribirse como

$$x(t) = A(t)\psi(t), \quad (2.60)$$

donde  $\psi(t)$  es una señal oscilante (Def. 2.4.1) con todos sus máximos locales iguales a 1 y todos sus mínimos locales iguales a  $-1$ . La modulación es amplitud  $A(t)$  es tal que no induce nuevos extremos locales, esto es  $\#\mathbb{L}_\psi = \#\mathbb{L}_{A\psi}$ .

Esta forma permite ver a una IMF como un par amplitud-oscilación. La modulación en amplitud no genera extremos locales, y la oscilación puede «modularse» en frecuencia, de modo de oscilar más «rápido» o más «lento» (disminuir o aumentar la distancia entre extremos) al tiempo que puede modificar su forma de onda.

## 2.4.7. Cuestiones prácticas de implementación

Para llevar a la práctica la técnica de EMD es necesario considerar varias cuestiones referentes su implementación. Ello ha dado pie a numerosas investigaciones planteando distintas propuestas para cada aspecto de la implementación práctica del algoritmo.

Los efectos de bordes «dañan» a todos los métodos de procesamiento de señales en la medida que la cantidad de datos es finita, y EMD no es ajena a ello. Es necesario entonces hacer «suposiciones» sobre cómo es la señal antes y después de su registro. Otras cuestiones son más propias de EMD: el criterio de detención para saber hasta cuándo realizar el proceso de *sifting*, el esquema de interpolación usado para encontrar las envolventes y los problemas relacionados al muestreo.

### 2.4.7.1. Condiciones de borde

El algoritmo de EMD se basa en los extremos locales de la señal, a los que utiliza para estimar las envolventes que son definidas como la interpolación entre los máximos locales (para la envolvente superior) y entre los mínimos locales (para la envolvente inferior). Mediante esta definición, una envolvente (por ejemplo la superior) está perfectamente definida para todos los puntos entre dos extremos

(máximos en este caso), pero presenta problemas a la hora de definir la envolvente en los bordes de la señal.

Existen tres enfoques generales para este problema:

- **Multiplicación por una función de ventaneo.** Este enfoque consiste en multiplicar a la señal analizada por una función de ventaneo que tome el valor 1 en la parte central y que decaiga suavemente hacia cero en los bordes. Mediante esta operación se puede resolver el problema de la definición de las envolventes en los bordes imponiendo un valor de cero en éstos. Esta propuesta puede verse en [33, 94]. Esta idea es sin embargo limitada, ya que la simple adición de una constante o una tendencia no estacionaria a la señal haría que la multiplicación por la ventana induzca grandes variaciones en los bordes.
- **Prolongación de la señal.** Este enfoque consiste en prolongar la señal más allá de los bordes de manera de hacer aparecer nuevos extremos locales que permitan una correcta definición de las envolventes en los bordes de la señal. Este problema constituye un área en sí mismo y se encuadra en las investigaciones sobre predicción de series temporales. Se han propuesto métodos basados en redes neuronales [35] y en máquinas de soporte vectorial [24, 111].
- **Prolongación de los extremos.** Más sencillo que prolongar la señal es prolongar el conjunto de extremos. Desde un punto de vista teórico, el problema posee una dificultad similar a la prolongación de la señal. Sin embargo el costo computacional de la prolongación de extremos es menor al de la prolongación de la señal. Un método de predicción de extremos consiste en realizar un espejado en los bordes de la señal [99]. Otros dos métodos se proponen en [29]: uno similar a un espejado, y el otro a una predicción lineal de los extremos.

#### 2.4.7.2. Criterios de detención

Existen distintos enfoques para determinar cuándo hay que detener el proceso de *sifting*. De forma general, persiguen dos propósitos: detener el proceso de *sifting* cuando se cumple con la definición de IMF (Def. 2.4.2); y que el proceso de *sifting* no realice demasiadas iteraciones ya que se corre el riesgo de desnaturalizar la información contenida en las IMFs [97].

El enfoque original [58] del criterio de detención se focalizaba principalmente en el primer objetivo, deteniendo el proceso de *sifting* cuando:

1. todos los máximos locales son estrictamente positivos y todos los mínimos locales son estrictamente negativos; y
2. la diferencia entre la IMF actual y su versión en la iteración anterior, medida por

$$SD = \int \left( \frac{d_1^n(t) - d_1^{n-1}(t)}{d_1^{n-1}(t)} \right)^2 dt, \quad (2.61)$$

con  $d_1^n(t)$  la  $n$ -ésima iteración de *sifting* para el primer modo, sea inferior a un umbral, por ejemplo del orden de  $0,2 - 0,3$  para una señal de 1024 muestras.

El problema de este enfoque es que el criterio propuesto para la diferencia entre la IMF actual y su iteración anterior no cumple su función deseada. De hecho, el integrando diverge generalmente para los puntos donde el denominador se anula, lo que hace que el valor de  $SD$  esté controlado más por el comportamiento de  $d_1^{n-1}(t)$  en la vecindad de sus cruces por cero que por el valor del numerador. Enmendando esta dificultad, los mismos autores propusieron

$$SD = \frac{\int (d_1^n(t) - d_1^{n-1}(t))^2 dt}{\int (d_1^{n-1}(t))^2 dt}, \quad (2.62)$$

como criterio en [55, 56, 117]. El mismo criterio fue discutido en [118], sugiriendo detener el proceso de *sifting* cuando este valor cae por debajo de  $10^{-4}$ . Podemos ver que en el numerador, se integra el cuadrado de la media local de  $d_1^{n-1}$ , el cual es igual a  $d_1^{n-1} = a_1^n + d_1^n$ . Esta definición cumple mejor el objetivo buscado pero tiene la dificultad de ser *global*. De hecho, para un mismo valor de  $SD$ , la media puede alejarse mucho de cero muy localmente, o alejarse moderadamente de cero para toda la duración de la señal. Estas dos situaciones son bien distintas: se puede tolerar un error que es fuerte pero localizado, pero no un error moderado distribuido por toda la señal.

Si indagamos más en la interpretación podemos constatar que la definición compara, de hecho, la amplitud de la media de  $d_1^{n-1}$  (el numerador) con su propia amplitud (el denominador), para saber si detener o no el proceso de *sifting* de  $d_1^{n-1}$ . Más aún, no se verifica que la media de la IMF final sea pequeña sino la media de la iteración anterior. Por lo tanto, no se está verificando la definición de IMF sino un criterio de convergencia tipo Cauchy [95, 118].

Una propuesta significativamente distinta es considerar un criterio *local*. Esta idea, desarrollada en [99], propone que la media local (semisuma de envolventes) sea evaluada localmente contra una cantidad que también sea local. El criterio se resume de la siguiente forma:

1. todos los máximos locales son estrictamente positivos y todos los mínimos locales son estrictamente negativos; y
2. las envolventes de la IMF actual verifican

$$|e_{m\acute{a}x}(t) + e_{m\acute{i}n}(t)| \leq \epsilon |e_{m\acute{a}x}(t) - e_{m\acute{i}n}(t)|$$

para un subconjunto del dominio de medida mayor o igual a  $(1 - \alpha)T$ , con  $T$  la duración total, y

$$|e_{m\acute{a}x}(t) + e_{m\acute{i}n}(t)| \leq \epsilon_2 |e_{m\acute{a}x}(t) - e_{m\acute{i}n}(t)|$$

para todo punto  $t$  del dominio.

Al contrario del enfoque global, esta definición del criterio de detención está concebida para controlar lo más localmente posible la calidad de la IMF en construcción. Idealmente, uno desearía imponer para todo  $t$  la misma restricción  $|e_{m\acute{a}x}(t) + e_{m\acute{i}n}(t)| \leq \epsilon |e_{m\acute{a}x}(t) - e_{m\acute{i}n}(t)|$ , pero en la práctica se verifica que

$|e_{m\acute{a}x}(t) + e_{m\grave{m}n}(t)|/|e_{m\acute{a}x}(t) - e_{m\grave{m}n}(t)|$  tiende a tener picos bien localizados y a estar relativamente cerca de cero allí donde no hay picos. Imponiendo una condición  $|e_{m\acute{a}x}(t) + e_{m\grave{m}n}(t)| \leq \epsilon |e_{m\acute{a}x}(t) - e_{m\grave{m}n}(t)|$  con  $\epsilon$  cercano a 0, 1 para todo  $t$ , se obtiene una cantidad de iteraciones de *sifting* demasiado grande [95]. Para relajar esta restricción se pide que se cumpla sólo para el 95 % de la señal, mientras que la restricción que debe cumplirse para toda la duración es con  $\epsilon_2 = 0,5$ . Este enfoque local, entonces, controla que la media local sea suficientemente pequeña para la mayor parte de la señal, permitiendo amplitudes mayores para la porción restante. Con ello se asegura que la segunda condición de IMF (Def. 2.4.2) se satisfaga suficientemente bien sin usar demasiadas iteraciones de *sifting* que pudieran desnaturalizar la información de modulación en amplitud de la IMF.

Con el objetivo de no iterar el proceso de *sifting* más de lo necesario, los autores de EMD propusieron un método para determinar la cantidad de iteraciones que no verifica explícitamente la definición de IMF [57, 59]. Este criterio consiste en detener el proceso de *sifting* cuando por una cantidad  $S$  de iteraciones sucesivas se cumpla que:

1. todos los máximos locales son estrictamente positivos y todos los mínimos locales son estrictamente negativos; y
2. la cantidad de extremos no cambia.

Dada la dificultad que vimos para evaluar qué tan cerca de cero está la media local de una IMF, esta estrategia propone simplemente no hacerlo. En vez de ello se supone que  $S$  iteraciones son suficientes para que la media local sea suficientemente pequeña. De hecho, se observa generalmente que luego de una iteración de *sifting* la media local siempre es más pequeña que antes de la iteración, excepto cuando aparecen nuevos extremos locales [95], lo que justifica el punto de vista según el cual una iteración de *sifting* durante la cual no aparecen extremos llevan a la media más cerca de cero para toda la duración de la señal.

En la práctica se observa que un valor de  $S$  entre 3 y 5 permite en la mayoría de los casos obtener una IMF con una cantidad razonable de iteraciones [57]. Más aún, y siempre con el objetivo de evitar iterar demasiado, los autores recomiendan siempre limitar a un máximo la cantidad de iteraciones permitidas bajo toda eventualidad.

Finalmente, el enfoque más sencillo es determinar *a priori* una cantidad fija de iteraciones, al margen de la definición de IMF. Esta es la estrategia utilizada en una variante denominada EMD por conjuntos (EEMD, del inglés *Ensemble EMD*) sobre la cual volveremos más adelante (Cap. 5).

### 2.4.7.3. Esquemas de interpolación

La obtención de las envolventes, a partir de los máximos y mínimos locales, se hace usualmente mediante una interpolación. La elección de la técnica de interpolación es de vital importancia, ya que como vimos todos los modos, salvo el primero, son suma de interpolaciones. Además, la cantidad de iteraciones de *sifting* necesarias depende del esquema de interpolación elegido [95].

En el contexto de EMD, las propiedades deseadas de la interpolación son esencialmente dos:

- **Localidad.** Uno de los objetivos de EMD es ser un método que permita el procesamiento de señales no estacionarias, de modo de adaptarse a las evoluciones de la señal. Para lograr esto, es necesario que el valor de una envolvente en un instante dado dependa esencialmente de la evolución de la señal en las proximidades de ese instante, y menos de instantes más alejados donde la señal pueda ser significativamente diferente. Por lo tanto, las interpolaciones particularmente no locales como la interpolación de Lagrange, son por completo inadecuadas. A la inversa, es deseable que el valor dependa sólo de un número finito de puntos de interpolación, o nodos, a ambos lados del instante considerado. Como mínimo, es necesario que la influencia de un nodo sobre el valor de la interpolación en un instante dado decrezca en función de su distancia al instante considerado y que tienda a cero cuando ésta tienda a infinito. En la práctica, la mayor parte de los esquemas de interpolación usados cumplen con esto último, pero la manera en que la influencia de un nodo decrece con la distancia puede variar según el método.
  
- **Regularidad.** La importancia de esta propiedad puede verse en el hecho que las aproximaciones (medias locales) sucesivas de la señal son sumas de interpolaciones. Si calculamos la primera IMF con distintos esquemas de interpolación, estamos restando a la señal diferentes primeras aproximaciones que poseen distinta cantidad de extremos. Entre esos extremos, algunos corresponden a las escalas características de la señal, pero otros son introducidos artificialmente por la interpolación. Por consiguiente, es deseable que la interpolación no oscile demasiado: a lo sumo un extremo o una inflexión entre dos nodos parece razonable. Más aún, si la interpolación tiene extremos, es necesario que éstos no sean demasiado «picudos».

Si buscamos un esquema de interpolación que satisfaga ambas propiedades, caeremos en la cuenta que de hecho ambas son contradictorias: para obtener una interpolación regular es necesario que su valor en un intervalo entre dos extremos dependa al menos de la evolución de la señal en sus intervalos vecinos. Esta dependencia conduce al hecho que el valor de la interpolación en un punto dependa de todos los puntos de interpolación, pero afortunadamente esta dependencia decrece con la distancia y tiende a cero cuando ésta tiende a infinito. Si tomamos como ejemplo la familia de *splines* polinomiales, ellos no dependen de intervalos vecinos cuando el grado es igual a 0 o 1 (interpolación constante y lineal a trozos respectivamente). Para órdenes superiores, la dependencia con los vecinos decrece exponencialmente con la distancia, crece con el grado y tiende a  $1/k$ , siendo  $k$  la distancia expresada en cantidad de nodos, cuando el grado tiende a infinito [108]. Al mismo tiempo, la regularidad de los *splines* polinomiales aumenta con el grado: un *spline* polinomial de grado  $n$  es por definición un polinomio de grado a lo sumo  $n$  a trozos con derivadas de hasta orden  $n - 1$  continuas.

Por estas razones, en la práctica la interpolación más usada es la *spline* cúbica. No es ciertamente la óptima desde un punto de vista de la localidad, pero es en cambio particularmente buena desde el punto de vista de la regularidad.

#### 2.4.7.4. Problemas de muestreo

Si bien EMD es siempre presentada en un marco de tiempo continuo, es decir que las señales son consideradas como funciones de una variable de tiempo continua, en la práctica trabaja sobre señales discretas (como todos los métodos de procesamiento de señales). En tanto señales discretas, es evidente que el desempeño de EMD se verá afectado por la frecuencia con la cual fueron muestreadas.

La influencia de la frecuencia de muestreo de las señales fue estudiada en [96, 98, 104]. La cuestión radica fundamentalmente en las ubicaciones de los extremos locales, las cuales se ven influidas por la frecuencia de muestreo de la señal. Particularmente en [98] deducen una cota superior para la norma  $L_1$  del error del operador elemental de *sifting*:

$$\|\delta\mathcal{S}x\|_1 \leq \frac{\alpha}{f_s} + \frac{\beta}{f_s^2} + \frac{\gamma}{f_s^3}, \quad (2.63)$$

con  $\alpha$ ,  $\beta$  y  $\gamma$  parámetros extraídos de la misma señal, y  $f_s$  la frecuencia de muestreo.

## 2.5. Comentarios finales

En este capítulo introdujimos los conceptos de señal multicomponente y frecuencia instantánea, los cuales se usan para construir modelos de fenómenos físicos en los que puede identificarse un número relativamente pequeño de componentes cuya frecuencia es única para un instante de tiempo dado. Sin embargo, nuestra capacidad para estimar tales frecuencias se encuentra limitada por los teoremas de Bedrosian y Nutall.

Luego presentamos el que quizá sea el enfoque más natural a la hora de analizar una señal multicomponente: las representaciones lineales tiempo-frecuencia y tiempo-escala. En este caso, el análisis se realiza mediante el producto interno entre la señal y un núcleo tiempo-frecuencia o tiempo-escala. Si la señal es multicomponente, entonces aparecerán «crestas» en su representación las cuales pueden usarse para estimar las frecuencias instantáneas de cada componente. Por último, para descomponer la señal, i.e. obtener los componentes en el dominio temporal, se antitransforma la representación en dominios disjuntos que rodean a cada cresta. Según este paradigma, se tiene que una cresta implica un componente. El principio de incertidumbre de Heisenberg-Gabor limita nuestra capacidad para describir instantáneamente las frecuencias presentes en la señal.

Finalmente presentamos un técnica que trabaja directamente en el dominio de observación y a partir de él extrae los componentes: la descomposición empírica en modos. Esta técnica, completamente guiada por los datos, descompone una señal en oscilaciones rápidas y lentas. Al final, se tiene que la señal puede representarse como la suma de una cantidad relativamente pequeña de componentes que, en algunos casos, pueden representarse como funciones AM-FM. Un procesamiento posterior mediante transformada de Hilbert permite obtener una representación tiempo-frecuencia.

Abordamos en este capítulo tanto los aspectos teóricos de la EMD, como sus aspectos prácticos. En los próximos capítulos, que constituyen las contribuciones

---

originales de este trabajo de tesis, profundizaremos en aspectos relacionados con la EMD, tomando como base lo aquí desarrollado.



# Capítulo 3

## Un enfoque basado en optimización para EMD

*En la teoría del conocimiento, como en todos los otros dominios de la ciencia, hay que razonar dialécticamente, o sea, no suponer jamás a nuestro conocimiento acabado e invariable, sino analizar el proceso gracias al cual el conocimiento nace de la ignorancia o gracias al cual el conocimiento incompleto e inexacto llega a ser más completo y más exacto.*

---

MATERIALISMO Y EMPIRIO-CRITICISMO  
Vladimir I. Lenin

### Contenidos

---

<b>3.1. Introducción</b>	<b>43</b>
<b>3.2. Hacia un enfoque teórico</b>	<b>43</b>
<b>3.3. EMD como un problema de optimización</b>	<b>44</b>
3.3.1. Antecedentes	44
3.3.2. Optimización sin restricciones (UOA-EMD)	46
<b>3.4. Características de la descomposición</b>	<b>49</b>
3.4.1. Comparación con <i>B-splines</i>	49
3.4.2. Interpretación como filtro variante en el tiempo	51
3.4.3. Espacio nulo de $P_x$	52
3.4.4. Relación con las IMTs de Daubechies	54
<b>3.5. Implementación</b>	<b>56</b>
3.5.1. Selección del parámetro $\lambda$	57
3.5.2. Efectos de borde	59
3.5.3. Criterios de detención	60
3.5.4. Efectos de muestreo	61
<b>3.6. Experimentos y resultados</b>	<b>61</b>

---

3.6.1. Separación de tonos . . . . .	62
3.6.2. Mezclas de oscilaciones AM-FM y no lineales . . . . .	64
3.6.3. Ruido blanco gaussiano y banco de filtros equivalente . . . . .	66
3.6.4. Descomposición de señales de ancho de banda localmente estrecho . . . . .	69
3.6.5. Señal real de voz . . . . .	70
<b>3.7. Extensión a dos dimensiones . . . . .</b>	<b>72</b>
3.7.1. Nuestro enfoque . . . . .	73
3.7.2. Cuestiones de implementación . . . . .	75
3.7.3. Delta de Dirac bidimensional . . . . .	76
3.7.4. Imágenes AM-FM . . . . .	76
3.7.5. Imagen real . . . . .	81
3.7.6. Aplicación biomédica . . . . .	83
<b>3.8. Comentarios finales . . . . .</b>	<b>84</b>

---

## 3.1. Introducción

Como hemos visto en el capítulo anterior, la Descomposición Empírica en Modos (EMD, del inglés *Empirical Mode Decomposition*) es un método adaptativo para el análisis de señales no estacionarias que pueden provenir de sistemas no lineales. De manera local y completamente guiada por los datos, separa una señal en oscilaciones rápidas y lentas. Al final de la descomposición, la señal original se expresa como una suma de funciones AM-FM llamadas funciones intrínsecas de modo (IMF, del inglés *Intrinsic Mode Function*) más una tendencia final, o *residuo*, que es o bien monótono o tiene a lo sumo dos extremos locales.

La EMD ha sido aplicada exitosamente en los campos más diversos. Pueden encontrarse ejemplos de su aplicación en la oceanografía [57], geofísica [60], neurofisiología [112], identificación de sistemas no lineales [67] o el análisis vibracional [89], por citar sólo algunos casos. Sin embargo, y tal como describimos en el capítulo anterior, aún posee una dificultad importante: carece de sólidas bases matemáticas, y se define como la salida de un algoritmo.

En este capítulo indagaremos en los esfuerzos más recientes para dotar a EMD de fundamentos matemáticos. Prestaremos especial atención a una línea particular dentro de esta tendencia: las propuestas basadas en teoría de optimización. Tal es así, que propondremos como primera contribución original de este trabajo de tesis un nuevo enfoque para EMD, estudiando sus propiedades, relacionándola con propuestas anteriores y deduciendo características particulares propias.

Extensas simulaciones soportan la validez de nuestra propuesta. Revisitamos experimentos pioneros en el análisis del desempeño de EMD: separación de tonos puros y descomposición de ruido blanco. Además descomponemos señales AM-FM solapadas en tiempo y en frecuencia, y comparamos nuestros resultados con aquellos del estado del arte. Para finalizar, proponemos una extensión a datos 2D (imágenes) que por su sencillez invita a la extensión a  $n$  dimensiones.

En la Sec. 3.2 repasamos brevemente los esfuerzos realizados para dotar a EMD de bases teóricas. En la Sec. 3.3 primero describimos los métodos del estado del arte de EMD inspirada en optimización para luego presentar nuestra propia propuesta: un enfoque para EMD basado en optimización sin restricciones. Las secciones siguientes analizan las propiedades de nuestra propuesta y discuten su implementación práctica. En la Sec. 3.6 se presentan simulaciones con señales artificiales y reales, comparando nuestros resultados con los del estado del arte. En la Sec. 3.7 presentamos una extensión a dos dimensiones de nuestro método.

## 3.2. Hacia un enfoque teórico

En los últimos años se han realizado varios esfuerzos por dotar a EMD de bases matemáticas más sólidas. Deléchelle y col. [34] estimaron las envolventes de la señal resolviendo una ecuación diferencial parabólica. Xu y col. [123] modificaron la definición de envolvente obteniendo una expresión analítica más simple, para la cual las variaciones en los extremos durante el proceso iterativo pueden ser investigadas en detalle. Hawley y col. [49] reemplazaron la interpolación *spline* cúbica por interpolaciones trigonométricas para la estimación de envolventes. Debido a ello, dedujeron interesantes propiedades y garantías de convergencia, aunque

sus resultados difieren significativamente de aquellos que otorga la EMD clásica. Daubechies y col. [31, 115] compararon a EMD con la teoría de onditas, reasignando la CWT mediante la técnica de *synchrosqueezing*.

Un enfoque distinto, basado en teoría de optimización, atrajo recientemente la atención de la comunidad científica de EMD. Este es el enfoque que desarrollaremos en lo que resta del capítulo.

**Notación.** Sin pérdida de generalidad, consideraremos ahora señales discretas de la forma  $x = x(nT)$ ,  $n = 1, 2, \dots, N$ ,  $x \in \mathbb{R}^N$ , con  $1/T$  siendo la frecuencia de muestreo. Recordamos aquí la ec. (2.40), que describe la forma de descomposición que entrega EMD:

$$x = a_K + \sum_{k=1}^K d_k, \quad (3.1)$$

donde además se tiene que  $a_k = a_{k+1} + d_{k+1}$ , usando la convención  $a_0 = x$ .

Usaremos letras mayúsculas  $A$  para denotar a las matrices pertenecientes a  $\mathbb{R}^{N \times N}$ , de forma tal que el producto  $Ax$ , con  $x \in \mathbb{R}^N$  vector columna, también sea un vector en  $\mathbb{R}^N$ .

Para los problemas de optimización Argmin denotará un conjunto de minimizadores, mientras que arg min hará lo propio para un único minimizador.

## 3.3. EMD como un problema de optimización

### 3.3.1. Antecedentes

B. Huang y Kunoth [54] propusieron una implementación de EMD en la cual reemplazan la interpolación explícita a través de extremos locales para la estimación de las envolventes por la solución a un problema de optimización. Sin embargo, esta propuesta mantiene el enfoque basado en envolventes y el proceso de *sifting*. Hemos visto en el capítulo anterior que el concepto de envolventes como interpolación a través de extremos locales es cuanto menos cuestionable. Es necesario entonces evitar las envolventes obteniendo una media local de manera *directa* y prescindir, si se puede, del proceso de *sifting*. A continuación presentaremos dos métodos del estado del arte de EMD basados en optimización, con los cuales contrastaremos nuestras propias propuestas.

#### 3.3.1.1. Optimización de *splines* (OS-EMD)

La primera propuesta que intentó emular a EMD dentro de un marco de optimización se debe a Meignen y Perrier [79]. Dada una señal  $x$ , su objetivo era encontrar una media local, de tal forma que la diferencia entre la señal y esta última, fuera una IMF. Para ello diseñaron un problema de optimización convexo *con restricciones* tanto de igualdad como de desigualdad. Esta propuesta fue mejorada en [83], donde los autores extraen  $a_k$ , la tendencia de orden  $k$ , resolviendo el siguiente problema de optimización:

$$a_k = \arg \min_{a \in \mathbb{R}^N} \|a''\|_2^2, \text{ sujeto a } a \in \Pi \cap \mathbb{B}_{a_{k-1}}, \quad (3.2)$$

donde la derivada segunda se aproxima mediante una matriz de diferencias,  $\Pi$  representa el espacio de *splines* cúbicos y  $\mathbb{B}_{a_{k-1}}$  representa las restricciones de simetría impuestas a  $a_{k-1} - a_k$  (que sería el modo  $d_k$ ) en las posiciones de los extremos locales de  $a_{k-1}$ . Lo que se busca entonces es una media local suave que al restarla a la señal produzca una IMF. Siendo estrictos, la variable a optimizar en (3.2) no es la media local sino los coeficientes asociados a la interpolación de Hermite de  $a_k$ . Este enfoque, sin embargo, posee una limitación importante: es necesaria una primera aproximación  $\hat{a}$  de  $a_k$  para que el problema sea convexo. Los autores usan para ello el enfoque basado en integrales locales, el cual presenta dificultades en presencia de señales fuertemente moduladas en frecuencia [51].

Para este método, el usuario debe definir *a priori* dos parámetros: la tolerancia  $\alpha$ , que es usada como cota superior al modelar la simetría del modo; y el orden  $\sigma$  de la familia de *splines*. Los parámetros sugeridos son  $\alpha = 0,05$  y  $\sigma = 8$  [83]. Los autores proporcionan una implementación en MATLAB<sup>®</sup> disponible en: <http://www-ljk.imag.fr/membres/Thomas.Oberlin/EMDOS.tar.gz>.

### 3.3.1.2. Optimización proximal multicomponente (Prox-EMD)

Una propuesta más reciente se debe a Pustelnik y col. [92, 93]. Este enfoque proximal se caracteriza por una estimación *simultánea* de media local y modo. Dada  $a_{k-1}$ , el problema de optimización a resolver es:

$$(a_k, d_k) \in \underset{a, d \in \mathbb{R}}{\text{Argmin}} \|a_{k-1} - a - d\|_2^2 \quad (3.3)$$

$$\text{sujeto a } \begin{cases} \|Da\|_p^p \leq \eta_k, \\ \|P_k d\|_q^q \leq \varepsilon_k, \\ (\forall j < k), |\langle d, d_j \rangle| \leq \zeta_{k,j}. \end{cases} \quad (3.4)$$

Puede verse cómo la media local  $a$  y el modo  $d$  se buscan simultáneamente, al tiempo que se imponen restricciones de desigualdad sobre ambos. La restricción  $\|Da\|_p^p$ , donde se debe elegir la norma  $p$ , modela la suavidad de la media local debido a que  $D$  es un operador de derivadas, mientras que  $\|P_k d\|_q^q$ , fijando una norma  $q$ , modela la simetría del modo controlando los valores que toma en sus extremos locales (veremos en breve cómo es esta restricción). Las restricciones  $|\langle d, d_j \rangle| \leq \zeta_{k,j}$  tienen como objetivo imponer a los modos una condición de (cuasi)ortogonalidad.

Esta propuesta presenta una serie de dificultades. En primer lugar no se garantiza una reconstrucción perfecta, ya que se *minimiza* la diferencia entre señal y media local más modo pero no se asegura que esa diferencia sea *cero*. Por otro lado, ya hemos discutido lo cuestionable que es suponer la ortogonalidad de los modos producidos por EMD. La dificultad de problemas de este tipo, con una cantidad importante de *restricciones*, es que el usuario debe fijar parámetros sin la garantía de que la solución sea factible.

Esto se hace patente en la propia implementación en MATLAB<sup>®</sup> de los autores, disponible en: [http://perso.ens-lyon.fr/nelly.pustelnik/Software/Prox-EMD\\_v1.0.zip](http://perso.ens-lyon.fr/nelly.pustelnik/Software/Prox-EMD_v1.0.zip). En vez de resolver el problema (3.3), la *toolbox* resuelve el siguiente problema:  $(a_k, d_k) = \arg \min_{a, d} \|a_{k-1} - a - d\|_p^p + c_1 \|Da\|_q^q + c_2 \|P_k d\|_r^r$ . Reemplazan el problema con restricciones por su versión *regularizada*, dejando de lado las restricciones de ortogonalidad. Los parámetros por defecto son:  $p = 2, q = 2,$

$r = 1$  y una matriz de diferencias  $D$  que estima la derivada segunda. (Debe notarse que las normas  $p$ ,  $q$  y  $r$  de la *toolbox* difieren de las normas  $p$  y  $q$  del problema (3.3).) En este caso, la cantidad de modos y los valores de los parámetros son determinados mediante una descomposición por EMD clásica.

### 3.3.2. Optimización sin restricciones (UOA-EMD)

Hemos visto dos propuestas del estado del arte que intentan emular EMD bajo un marco de teoría de optimización. La propuesta de Oberlin y col. [83] necesita de una primera aproximación que se encuentra interpolando *explícitamente* un conjunto de puntos hallados mediante el método de integrales locales [51]. El método de Pustelnik y col. [92, 93] no garantiza una reconstrucción perfecta, y la selección de los parámetros puede ser muy complicada. Tal es así, que su *toolbox* los estima a partir de la descomposición EMD clásica.

Propondremos aquí un nuevo enfoque para EMD, basado en un problema de optimización *sin restricciones*. De manera general, el problema de extracción de media local se puede resumir de la siguiente forma: dada una señal oscilante  $x$ , encontrar una media local  $a$  tal que  $d = x - a$  sea IMF (al menos en forma aproximada, como discutimos en la Sec. 2.4.6). Recuperamos así un esquema deflacionario para la descomposición, junto con una reconstrucción perfecta  $x = a + d$ .

Un aspecto común a los métodos de Oberlin y col. y Pustelnik y col. es que las condiciones de IMF (simetría y media local nula, Def. 2.4.2) no son consideradas en la función objetivo del problema de optimización sino que aparecen a través de restricciones de desigualdad, con las correspondientes cotas que deben ser fijadas por el usuario. Recordemos las condiciones que debe cumplir una función para ser considerada IMF:

- (I) máximos locales estrictamente positivos y mínimos locales estrictamente negativos; y
- (II) semisuma de envolventes nula.

Es claro que la condición (II) no puede ser satisfecha sin cumplir la condición (I). Un método que busque la satisfacción de (II) logrará, como subproducto, satisfacer (I).

Usaremos aquí, para esta última cuestión, un enfoque derivado de [83]. Consideremos que, por ejemplo, la señal  $x$  tiene un mínimo local en  $t_\ell$ . El máximo local más próximo a su izquierda se encuentra en  $t_{\ell-1}$  y el más próximo a su derecha en  $t_{\ell+1}$ . Entonces, una medida de qué tan «simétrica» es la señal en torno a  $t_\ell$  es la expresión

$$x(t_\ell) + i_\ell(t) \Big|_{t=t_\ell} = x(t_\ell) + \left( \frac{x(t_{\ell+1})(t - t_{\ell-1})}{t_{\ell+1} - t_{\ell-1}} + \frac{x(t_{\ell-1})(t_{\ell+1} - t)}{t_{\ell+1} - t_{\ell-1}} \right) \Big|_{t=t_\ell}, \quad (3.5)$$

la cual será tanto más cercana a cero cuanto más «simétrica» sea la señal. Puede verse que  $i_\ell(t)$  es una función que interpola linealmente entre  $t_{\ell-1}$  y  $t_{\ell+1}$ . La situación se describe en la Fig. 3.1.

Si extendemos el mismo razonamiento para toda la señal, entonces podremos definir una media local. La simetría debe evaluarse sobre el modo  $d = x - a$ , pero

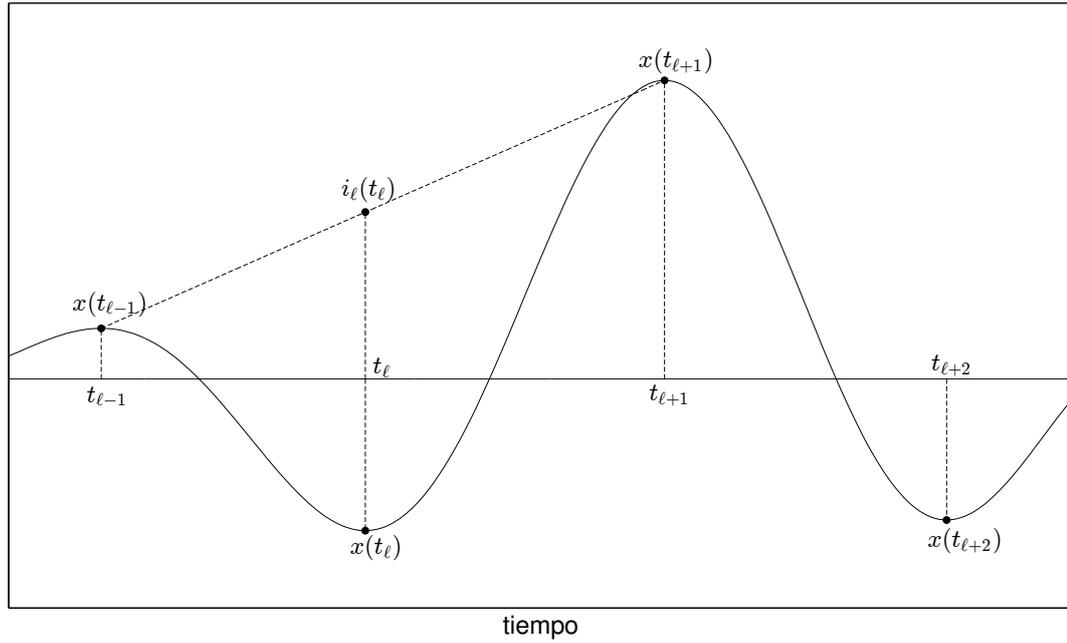


Figura 3.1: Señal oscilante, extremos locales y función interpolante. A partir de los extremos locales se construyen las funciones  $i_\ell(t)$ , de tal forma que  $x(t_\ell) + i_\ell(t_\ell)$  sea una medida de la simetría de  $x(t)$  en  $t = t_\ell$ .

no disponemos de él para conocer las posiciones de sus extremos locales. Entonces las estimaremos a partir de los extremos de la señal. Sea  $x$  es una señal oscilante, y  $\mathbb{L}_x = \{t_\ell, \ell = 1, \dots, L\}$ , con  $L = \#\mathbb{L}_x$  el conjunto de sus extremos locales. Podemos reproducir la ec. 3.5 definiendo un vector  $p_{t_\ell}$  tal que

$$p_{t_\ell}^T x = x(t_\ell) + \left( \frac{x(t_{\ell+1})(t_\ell - t_{\ell-1})}{t_{\ell+1} - t_{\ell-1}} + \frac{x(t_{\ell-1})(t_{\ell+1} - t_\ell)}{t_{\ell+1} - t_{\ell-1}} \right), \quad (3.6)$$

con lo que resulta

$$p_{t_\ell} = \left[ \dots \frac{t_{\ell+1} - t_\ell}{t_{\ell+1} - t_{\ell-1}} \dots 1 \dots \frac{t_\ell - t_{\ell-1}}{t_{\ell+1} - t_{\ell-1}} \dots \right]^T, \quad (3.7)$$

donde mostramos sólo los elementos distintos de cero, los cuales se hallan, respectivamente, en las posiciones  $t_{\ell-1}$ -ésima,  $t_\ell$ -ésima y  $t_{\ell+1}$ -ésima. A partir de estas definiciones, podemos construir una matriz *cuadrada*  $P_x \in \mathbb{R}^{N \times N}$  (el subíndice nos indica que la matriz se construye a partir de los extremos locales de  $x$ ) que aloje a  $p_{t_\ell}^T$  en su  $t_\ell$ -ésima fila. La matriz se ve de la siguiente forma (mostramos sólo los elementos no nulos más cercanos a la posición  $t_\ell, t_\ell$ ):



5: calcular la media local  $a_k$  para  $\lambda_k > 0$

$$a_k = (P_{a_{k-1}}^T P_{a_{k-1}} + \lambda_k D^T D)^{-1} (P_{a_{k-1}}^T P_{a_{k-1}}) a_{k-1}; \quad (3.11)$$

6: calcular el modo  $d_k = a_{k-1} - a_k$ ;

7: **end for**

### 3.4. Características de la descomposición

Una vez descrito el funcionamiento del método, es necesario indagar en sus características. En tanto enfoque teórico para EMD, el método debe entonces brindar la posibilidad de un estudio más detallado de la descomposición. Dada una señal oscilante  $x$ , deseamos analizar las características de la descomposición, esto es, describir cómo trabaja el método y cómo «construye» tanto la media local como el modo de la señal.

Para ejemplificar las características del método propuesto que se describirán a continuación, usaremos una señal artificial compuesta de dos componentes modulados en frecuencia y amplitud:

$$x = x_1 + x_2 = e^{-\pi\left(\frac{n-250}{500}\right)^2} \cos\left(2\pi\frac{6}{25}n - 60 \operatorname{sen}\left(\frac{2\pi}{500}n\right)\right) + e^{-\pi\left(\frac{n-250}{500}\right)^2} \cos\left(2\pi\frac{3}{25}n - 40 \operatorname{sen}\left(\frac{2\pi}{500}n\right)\right), \quad n = 1, 2, \dots, 500. \quad (3.12)$$

Ambos componentes se solapan en frecuencia, como puede verse en la Fig. 3.2. Es necesario entonces una acción de filtrado variante en el tiempo para separarlos.

#### 3.4.1. Comparación con *B-splines*

Reescribamos la ec. 3.10 de la siguiente forma:

$$\begin{aligned} a &= (P_x^T P_x + \lambda D^T D)^{-1} P_x^T P_x x \\ &= \underbrace{\left(2(P_x^T P_x + \lambda D^T D)^{-1} P_x^T\right)}_{\Psi} \underbrace{\frac{P_x x}{2}}_{\beta} \\ &= \sum_{\ell} \psi_{\ell} \beta(t_{\ell}), \end{aligned} \quad (3.13)$$

Dada la naturaleza rala de  $P_x$  (la cual discutiremos más adelante), la matriz  $\Psi$  posee columnas no nulas sólo para las posiciones  $\{t_{\ell}\}$  de extremos locales de  $x$ . Dichas columnas han sido denotadas por  $\psi_{\ell}$ . Además, y debido a la misma causa, el vector  $\frac{P_x x}{2}$  es igual a

$$\beta = \sum_{\ell} \left( \frac{x(t_{\ell-1})(t_{\ell+1} - t_{\ell})}{2(t_{\ell+1} - t_{\ell-1})} + \frac{x(t_{\ell})}{2} + \frac{x(t_{\ell+1})(t_{\ell} - t_{\ell-1})}{2(t_{\ell+1} - t_{\ell-1})} \right) \delta(n - t_{\ell}), \quad (3.14)$$

es decir, posee elementos no nulos sólo en las posiciones  $t_{\ell}$ . De esta manera, la ec. 3.13 enmarca nuestra propuesta en un marco más general para EMD discutido en

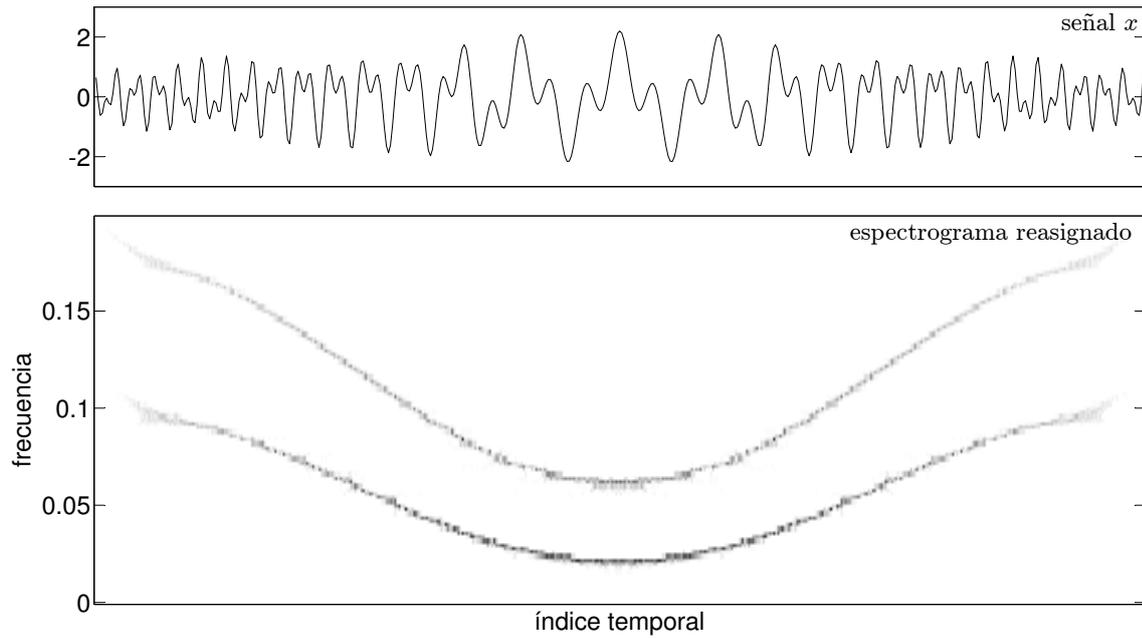


Figura 3.2: Señal artificial descrita por la ec. (3.12). A partir de su espectrograma reasignado puede apreciarse claramente el solapamiento frecuencial de ambas componentes. Un filtrado lineal convencional no sería capaz de separarlos, con lo que se hace necesaria una acción de filtrado variante en el tiempo.

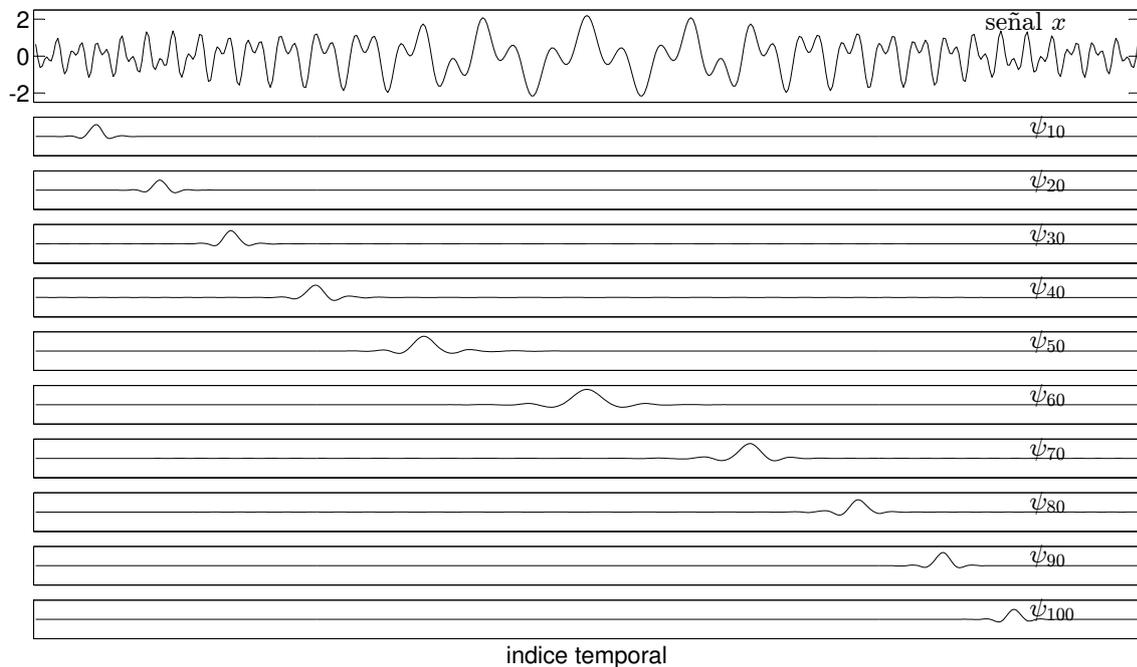


Figura 3.3: Comparación con *B-splines*. Se muestran 10 funciones  $\psi_\ell$  generadas. Éstas decaen rápidamente cuando  $t$  se aleja de  $t_\ell$ .

[23, 64, 122]. En [23], se propuso una versión de EMD basada en *B-splines*. En su

propuesta, la media local de una señal oscilante  $x$  se define como

$$a = \sum_{\ell} \left( \frac{x(t_{\ell+1})}{4} + \frac{x(t_{\ell+2})}{2} + \frac{x(t_{\ell+3})}{4} \right) \tilde{\psi}_{\ell}, \quad (3.15)$$

donde  $\tilde{\psi}_{\ell}$  indica la  $\ell$ -ésima *B-spline* cúbica definida en la partición dada por el conjunto  $\mathbb{L}_x$  de los extremos locales de la señal.

Las similitudes entre las ecuaciones (3.13), con (3.14), y (3.15) son evidentes. Sin embargo, nuestra propuesta es totalmente guiada por los datos. Tanto las formas de onda  $\psi_{\ell}$  como los coeficientes  $\beta$  son encontrados a partir de los datos, mientras que el esquema es mucho más rígido para la propuesta de *B-splines*: las formas de onda están determinadas *a priori* y los coeficientes se encuentran como combinación lineal de la señal en sus extremos con coeficientes fijos, sin considerar la distancia entre ellos. Además, nuestras funciones  $\psi_{\ell}$  cumplen que

$$\sum_{\ell} \psi_{\ell} = \chi_{[1\dots N]}, \quad (3.16)$$

o sea, están normalizadas. Esta propiedad se comparte con las *B-splines*.

En la Fig. 3.3 se muestran algunas de las  $\psi_{\ell}$  generadas con nuestro método. Las formas de onda, si bien no poseen soporte compacto, decaen suficientemente rápido. Su «duración efectiva» es proporcional a la distancia entre los extremos cercanos a  $t_{\ell}$ .

El esquema general para EMD, discutido en [64] y en donde se hace un análisis en  $N$  dimensiones, define la media local  $A$  para una señal  $X : \mathbb{R}^N \mapsto \mathbb{R}$  como

$$A = \sum_{\ell} \mu(T_{\ell}) \varphi_{\ell}, \quad (3.17)$$

donde los coeficientes  $\mu(T_{\ell})$  se encuentran como combinación lineal de  $X$  en sus extremos locales  $\{T_{\ell}, \ell \in \mathbb{Z}\}$ , y las funciones  $\varphi_{\ell}$  constituyen una base del espacio  $N$ -dimensional. Queda claro ahora que nuestra propuesta implementa de manera sencilla, y completamente guiada por los datos, este enfoque para señales unidimensionales.

### 3.4.2. Interpretación como filtro variante en el tiempo

Una forma más sencilla de analizar la solución de nuestro problema de optimización consiste en escribir

$$\begin{aligned} a &= (P_x^T P_x + \lambda D^T D)^{-1} P_x^T P_x x \\ &= \underbrace{\left( (P_x^T P_x + \lambda D^T D)^{-1} P_x^T P_x \right)}_{\Phi} x \\ &= \sum_{\ell} \phi_{\ell}(n - t_{\ell}) x(t_{\ell}). \end{aligned} \quad (3.18)$$

con  $n = 1 \dots N$ . Como en el caso anterior, la matriz  $\Phi$  posee columnas no nulas en las posiciones  $t_{\ell}$ , las cuales fueron denotadas como  $\phi_{\ell}$ , y en este caso fueron «trasladadas» para centrarlas en  $t_{\ell}$ . El objetivo de tal acción es poder ver a la

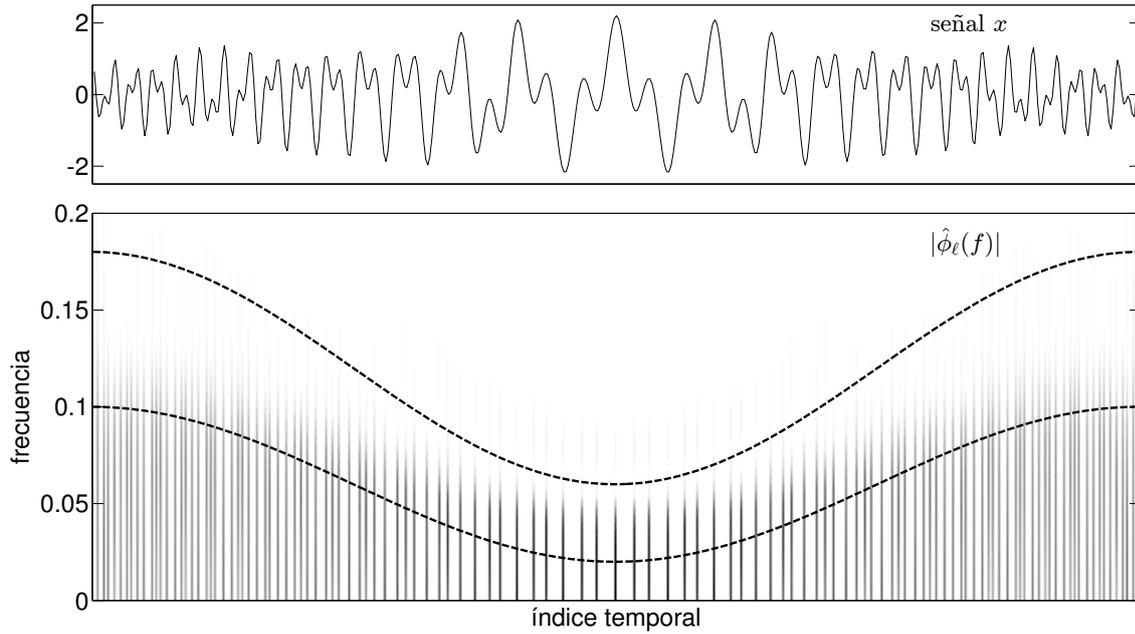


Figura 3.4: Interpretación como filtro variante en el tiempo. Vemos los módulos de las respuestas en frecuencia  $|\hat{\phi}_\ell(f)|$ , ubicados en las posiciones de los  $\ell$ -ésimos extremos, «desde arriba» en tonos de gris. Superponemos en líneas de trazo las frecuencias instantáneas de los componentes verdaderos.

ec. (3.18) como un filtrado variante en el tiempo, donde la respuesta en frecuencia  $\hat{\phi}_\ell(f)$  varía de extremo a extremo de la señal (aquí, claro está,  $\hat{\phi}_\ell(f)$  representa a la transformada *discreta* de Fourier de  $\phi_\ell(n)$ ). La ec. (3.18) resulta entonces una implementación de la ec. (2.49) que oportunamente discutiéramos en la Sec. 2.4.4:  $a(t) = (h \hat{\star} x_{\mathbb{L}})(t)$ , donde « $\hat{\star}$ » representa a una convolución generalizada y  $h$  es un filtro pasa-bajos generalizado. Podemos ver en la Fig. 3.4 cómo varía la respuesta en frecuencia  $\hat{\phi}_\ell(f)$  de cada una de las columnas no nulas de la matriz  $\Phi$ . La figura se presenta de modo que vemos las amplitudes de  $|\hat{\phi}_\ell(f)|$  en tonos de gris. Puede apreciarse cómo los «átomos»  $\phi_\ell$ , localizados alrededor de  $t_\ell$ , poseen un espectro tipo pasa-bajos con un ancho de banda cercano a la máxima frecuencia local. De esta forma podemos interpretar que nuestro método genera la media local de la señal mediante un filtrado tipo pasa-bajos con una «frecuencia de corte local» cercana a la máxima frecuencia instantánea.

Si queremos dar un paso más, debemos comparar a la ec. (3.18) con la ec. (2.43) de la Sec. 2.4.2:

$$a(t) = \int x(u) \frac{1}{\alpha(t)} w \left( \frac{u-t}{\alpha(t)} \right) du, \quad (3.19)$$

donde  $\frac{1}{\alpha(t)} w \left( \frac{u-t}{\alpha(t)} \right)$  serían los átomos. Nuestro método implementa de forma *discreta* y *local* una versión de la media local *ideal*, la cual debería ser *continua* e *instantánea*.

### 3.4.3. Espacio nulo de $P_x$

Con el objetivo de descomponer señales en componentes más «básicos» es que en [53, 87, 88] se propuso el método de Búsqueda de Espacio Nulo (NSP, del inglés

*Null Space Pursuit*). En dicho método, los autores llegan a una ecuación similar a nuestra ec. (3.10), donde nuestra matriz  $P_x$  ocupa el lugar de un operador particular. Continúan iterando tal ecuación hasta un criterio de detención dado, de forma que la solución pertenezca al *espacio nulo* del operador.

Resulta conveniente entonces estudiar el espacio nulo de nuestra matriz  $P_x$ . Tenemos que  $x \in \ker(P_x) \Leftrightarrow P_x x \equiv \bar{0}$ , con  $\bar{0}$  el vector nulo.

**Proposición 3.4.1.** *Sea  $x$  una señal oscilante y  $P_x$  construida a partir de sus extremos locales según las ecs. (3.7) y (3.8). Entonces,  $P_x x \equiv \bar{0}$  si los máximos locales se encuentran sobre  $At + B$  y los mínimos locales sobre  $-At - B$ , con  $A, B \in \mathbb{R}^+$ .*

Vemos que una señal oscilante  $x$  será tal que  $P_x x \equiv \bar{0}$  si sus máximos locales se encuentran sobre una recta, y los mínimos locales sobre otra recta simétrica a la anterior respecto al eje de las abscisas.

*Demostración.* Tenemos  $\mathbb{L}_x = \{t_\ell\}$  el conjunto de extremos locales de  $x$  y junto a  $x_\ell = x(t_\ell)$  forman los pares  $\{x_\ell, t_\ell\}$ . Entonces, para que sea  $P_x x \equiv \bar{0}$ , debe darse

$$x_\ell + \left( \frac{x_{\ell+1} - x_{\ell-1}}{t_{\ell+1} - t_{\ell-1}} \right) (t_\ell - t_{\ell-1}) + x_{\ell-1} = 0, \quad \forall \ell. \quad (3.20)$$

Aquí estamos suponiendo que  $\text{sgn}(x_\ell) = -\text{sgn}(x_{\ell-1}) = -\text{sgn}(x_{\ell+1})$ . Como los máximos y mínimos locales están intercalados, supondremos, sin pérdida de generalidad, que  $x$  tiene un máximo local en  $t_{2\ell}$  y un mínimo local en  $t_{2\ell+1}$ . Entonces

$$x_{2\ell} + \left( \frac{x_{2\ell+1} - x_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) (t_{2\ell} - t_{2\ell-1}) + x_{2\ell-1} = 0 \quad (3.21)$$

$$x_{2\ell+1} + \left( \frac{x_{2\ell+2} - x_{2\ell}}{t_{2\ell+2} - t_{2\ell}} \right) (t_{2\ell+1} - t_{2\ell}) + x_{2\ell} = 0 \quad (3.22)$$

De (3.21) en (3.22) tenemos

$$x_{2\ell+1} + \left( \frac{x_{2\ell+2} - x_{2\ell}}{t_{2\ell+2} - t_{2\ell}} \right) (t_{2\ell+1} - t_{2\ell}) - \left( \frac{x_{2\ell+1} - x_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) (t_{2\ell} - t_{2\ell-1}) - x_{2\ell-1} = 0 \quad (3.23)$$

$$(x_{2\ell+1} - x_{2\ell-1}) \left( 1 - \frac{t_{2\ell} - t_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) + (x_{2\ell+2} - x_{2\ell}) \left( \frac{t_{2\ell+1} - t_{2\ell}}{t_{2\ell+2} - t_{2\ell}} \right) = 0 \quad (3.24)$$

$$\left( \frac{x_{2\ell+1} - x_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) = - \left( \frac{x_{2\ell+2} - x_{2\ell}}{t_{2\ell+2} - t_{2\ell}} \right) \quad (3.25)$$

La ec. (3.25) se cumple para todo  $\ell$  y nos dice que la pendiente de la recta donde se encuentran los máximos locales es igual pero de signo opuesto a la recta donde se encuentran los mínimos locales. A partir de la ec. (3.21) podemos escribir:

$$x_{2\ell} + \left( \frac{x_{2\ell+1} - x_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) t_{2\ell} = -x_{2\ell-1} + \left( \frac{x_{2\ell+1} - x_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) t_{2\ell-1}, \quad (3.26)$$

y dada la ec. (3.25) tenemos

$$x_{2\ell} - \left( \frac{x_{2\ell+2} - x_{2\ell}}{t_{2\ell+2} - t_{2\ell}} \right) t_{2\ell} = -x_{2\ell-1} + \left( \frac{x_{2\ell+1} - x_{2\ell-1}}{t_{2\ell+1} - t_{2\ell-1}} \right) t_{2\ell-1}, \quad (3.27)$$

con lo que las ordenadas al origen son opuestas. De esta forma, se demuestra la proposición.  $\square$

La proposición 3.4.1 nos dice que si iteramos la ec. (3.10) de la forma

$$x \leftarrow x - (P_x^T P_x + \lambda D^T D)^{-1} P_x^T P_x x, \quad (3.28)$$

en el sentido de NSP, terminaremos con una señal cuyas «envolventes» son rectas. Esto concuerda con la formulación original para EMD, dado que para *splines* naturales, sólo señales con envolventes constantes cumplen estrictamente las condiciones de IMF [97, 112, 119]. Nuestra propuesta resulta más versátil, en la medida que la envolvente puede ser una recta con pendiente no nula.

Si bien establecimos aquí un paralelismo entre nuestra propuesta y NSP [87, 88], vale decir que los operadores propuestos por los autores no contemplan las características particulares de las IMFs. De esa forma, su método no emula a EMD sino que plantea una descomposición en otros términos (y no en oscilaciones localmente más rápidas y más lentas). Más adelante presentaremos (y discutiremos) señales que no pueden ser correctamente separadas por NSP y sí por nuestro enfoque.

### 3.4.4. Relación con las IMTs de Daubechies

Hemos visto en la Sec. 2.4.6 la relación entre las IMFs en las cuales EMD descompone una señal dada con la clase IMT (Def. 2.4.5) definida por Daubechies y col. [31]. Esta clase de señales puede usarse para definir un subconjunto de la clase IMF. Dado un  $\epsilon$  suficientemente pequeño, una  $\epsilon$ -IMT puede ser considerada una IMF, pero sólo una IMT sin modulación en amplitud (es decir, una señal puramente FM) satisface rigurosamente la definición. Sin embargo, la relación inversa no se cumple: no toda IMF puede ser vista como una  $\epsilon$ -IMT (sin importar el valor de  $\epsilon$ ). Enunciaremos, y demostraremos, a continuación una importante relación entre las IMTs y nuestro enfoque basado en optimización sin restricciones.

**Proposición 3.4.2.** *Si  $x(t)$  es una  $\epsilon$ -IMT con  $\epsilon$  suficientemente pequeño, entonces*

1. *Todos los máximos (mínimos) locales son positivos (negativos).*
2. *La penalización de simetría sobre  $x$  está acotada por  $\epsilon$ :  $\frac{\|P_x x\|_1}{L} = O(\epsilon)$ , con  $L = \#\mathbb{L}_x$  siendo la cantidad de extremos locales de  $x$ .*

La primera consecuencia satisface la condición (I) de IMF. La segunda consecuencia acota el «error» que cometen las envolventes al no ser estrictamente simétricas, satisfaciendo así la condición (II) de IMF (ver Def. 2.4.2).

*Demostración.* Comencemos con la primera consecuencia. Supongamos que  $x(t) = A(t) \cos(2\pi\phi(t))$  admite un máximo local en  $t = t_0$ . Entonces  $x'(t_0) = 0$  y  $x''(t_0) < 0$ .

Necesitamos que  $x(t_0) > 0 \iff x''(t_0) < 0$  y  $x(t_0) < 0 \iff x''(t_0) > 0$ . Para la derivada primera de  $x$  tenemos

$$\begin{aligned} x'(t_0) &= A'(t_0) \cos(2\pi\phi(t_0)) - 2\pi A(t_0)\phi'(t_0) \operatorname{sen}(2\pi\phi(t_0)) = 0 \Rightarrow \\ &\Rightarrow \cos(2\pi\phi(t_0)) = \frac{2\pi A(t_0)\phi'(t_0)}{A'(t_0)} \operatorname{sen}(2\pi\phi(t_0)). \end{aligned} \quad (3.29)$$

La derivada segunda es

$$\begin{aligned} x''(t_0) &= A''(t_0) \cos(2\pi\phi(t_0)) - 4\pi A'(t_0)\phi'(t_0) \operatorname{sen}(2\pi\phi(t_0)) \\ &\quad - 2\pi A(t_0)\phi''(t_0) \operatorname{sen}(2\pi\phi(t_0)) - 4\pi^2 A(t_0)[\phi'(t_0)]^2 \cos(2\pi\phi(t_0)). \end{aligned} \quad (3.30)$$

De (3.29) en (3.30) obtenemos

$$\begin{aligned} x''(t_0) &= A''(t_0) \cos(2\pi\phi(t_0)) - \frac{2[A'(t_0)]^2}{A(t_0)} \cos(2\pi\phi(t_0)) - \frac{A'(t_0)\phi''(t_0)}{\phi'(t_0)} \cos(2\pi\phi(t_0)) \\ &\quad - 4\pi^2 A(t_0)[\phi'(t_0)]^2 \cos(2\pi\phi(t_0)) \\ &= \underbrace{\left( A''(t_0) - \frac{2[A'(t_0)]^2}{A(t_0)} - \frac{A'(t_0)\phi''(t_0)}{\phi'(t_0)} - 4\pi^2 A(t_0)[\phi'(t_0)]^2 \right)}_{B(t_0)} \cos \phi(t_0) \\ &= B(t_0) \cos \phi(t_0). \end{aligned} \quad (3.31)$$

Ahora, si  $x(t_0) > 0$  entonces  $\cos \phi(t_0) > 0$  (debido a que  $A(t) > 0 \forall t$  por hipótesis (Def. 2.4.5)). Por lo tanto, para que se cumpla  $x''(t_0) < 0$ , debe darse  $B(t_0) < 0$ . Un razonamiento similar puede hacerse para un mínimo local. Para ambos casos, debe ser  $B(t_0) < 0$ . Ello implica

$$A''(t_0) - \frac{A'(t_0)\phi''(t_0)}{\phi'(t_0)} < \frac{2[A'(t_0)]^2}{A(t_0)} + 4\pi^2 A(t_0)[\phi'(t_0)]^2. \quad (3.32)$$

El miembro derecho es siempre positivo (ver Def. 2.4.5). Para el miembro izquierdo se tiene que  $|A'(t_0)\phi''(t_0)/\phi'(t_0)| \leq \epsilon^2 \phi'(t_0)$  y  $|A''(t_0)| \rightarrow 0$  a medida que  $\epsilon \rightarrow 0$  (como  $|A'(t)| \leq \epsilon \phi'(t) \forall t$ , entonces conforme  $\epsilon \rightarrow 0$  se tiene que  $|A'(t)| \rightarrow 0$  para todo  $t$ ). Por lo tanto, la primera consecuencia queda demostrada.

Para la segunda consecuencia, comenzaremos probando que para  $\epsilon$  suficientemente pequeño se tiene  $A(t_0) \approx |x(t_0)|$  o  $|A(t_0) - |x(t_0)|| \leq C\epsilon$  (ya que  $A(t) > 0$  y  $x(t) \leq A(t)$ ). Usando la ec. (3.29), podemos escribir

$$\begin{aligned} |A(t_0) - |x(t_0)|| &= \left| \frac{A'(t_0) \cos(2\pi\phi(t_0))}{2\pi\phi'(t_0) \operatorname{sen}(2\pi\phi(t_0))} \right| |1 - |\cos(2\pi\phi(t_0))|| \\ &= \left| \frac{A'(t_0)}{2\pi\phi'(t_0)} \right| |\cos(2\pi\phi(t_0))| \left| \frac{1 - |\cos(2\pi\phi(t_0))|}{\operatorname{sen}(2\pi\phi(t_0))} \right| \\ &\leq \frac{\epsilon}{2\pi} \end{aligned} \quad (3.33)$$

debido a que  $0 \leq \left| \frac{1 - |\cos(2\pi\phi(t_0))|}{\operatorname{sen}(2\pi\phi(t_0))} \right| \leq 1$ .

Ahora consideremos tres extremos sucesivos con ubicaciones  $t_{\ell-1}$  (máximo local),  $t_\ell$  (mínimo local) y  $t_{\ell+1}$  (máximo local), como en la Fig. 3.1. Gracias a las ecuaciones (3.5), (3.6) y (3.7) podemos escribir

$$\begin{aligned}
(|P_x x|)(t_\ell) &= \left| x(t_\ell) + \frac{x(t_{\ell+1})(t_\ell - t_{\ell-1})}{t_{\ell+1} - t_{\ell-1}} + \frac{x(t_{\ell-1})(t_{\ell+1} - t_\ell)}{t_{\ell+1} - t_{\ell-1}} \right| \\
&= \left| \frac{\Delta_\ell^+ [x(t_{\ell-1}) - |x(t_\ell)|] + \Delta_\ell^- [x(t_{\ell+1}) - |x(t_\ell)|]}{\Delta_\ell^+ + \Delta_\ell^-} \right| \\
&\leq \frac{\max\{\Delta_\ell^-, \Delta_\ell^+\}}{\Delta_\ell^+ + \Delta_\ell^-} (|x(t_{\ell-1}) - |x(t_\ell)|| + |x(t_{\ell+1}) - |x(t_\ell)||) \\
&\leq (||x(t_\ell)| - x(t_{\ell-1})| + |x(t_{\ell+1}) - |x(t_\ell)||) \\
&\leq (|A(t_\ell) - A(t_{\ell-1})| + |A(t_{\ell+1}) - A(t_\ell)|) + O(\epsilon) \\
&\leq \int_{t_{\ell-1}}^{t_{\ell+1}} |A'(t)| dt + O(\epsilon) \leq \epsilon \int_{t_{\ell-1}}^{t_{\ell+1}} |\phi'(t)| dt + O(\epsilon) \\
&\leq \epsilon(\phi(t_{\ell+1}) - \phi(t_{\ell-1})) + O(\epsilon) = M\epsilon + O(\epsilon), \tag{3.34}
\end{aligned}$$

donde usamos la ec. (3.33) para reemplazar a  $x$  por  $A$ ,  $\Delta_\ell^- = t_\ell - t_{\ell-1}$ ,  $\Delta_\ell^+ = t_{\ell+1} - t_\ell$ , y el hecho que  $\phi$  es una función de variación acotada [18]. Entonces, para la norma  $L_1$  podemos afirmar

$$\begin{aligned}
\|P_x x\|_1 &= \sum_\ell \left| x(t_\ell) + \frac{x(t_{\ell+1})(t_\ell - t_{\ell-1})}{t_{\ell+1} - t_{\ell-1}} + \frac{x(t_{\ell-1})(t_\ell - t_{\ell+1})}{t_{\ell+1} - t_{\ell-1}} \right| \\
&\leq L(M\epsilon + O(\epsilon)) = LO(\epsilon), \tag{3.35}
\end{aligned}$$

con  $L$  la cardinalidad del conjunto  $\mathbb{L}_x$  de extremos locales. De esta manera, la demostración queda completa.  $\square$

La Proposición 3.4.2 nos da una relación entre las IMTs de Daubechies y nuestro enfoque. En primer lugar, vemos que a medida que mejora la «calidad» de la IMT, o sea a medida que  $\epsilon$  disminuye, se garantiza la condición (I) de IMF, cumpliendo con la definición de *weak*-IMF [102]. La segunda consecuencia es mucho más importante para nosotros: a medida que mejora la calidad de la IMT, el «error» en la simetría (medido como  $\|P_x x\|_1$ ) disminuye. Funciones que son «mejores» desde el punto de vista de Daubechies también son mejores para nuestro enfoque.

## 3.5. Implementación

En lo que respecta a la implementación práctica de cualquier método de procesamiento de señales, siempre es deseable hacerlo de la forma más eficiente posible. El problema definido solamente a partir de las ecs. (3.5), (3.6) y (3.7) está mal condicionado: plantea condiciones *sólo* en las posiciones donde  $x$  tiene un extremo local pero pide una solución  $a$  definida para *todos* los puntos. Más específicamente, la matriz  $P_x \in \mathbb{R}^{N \times N}$  tiene rango igual a  $L$ , con  $L = \#\mathbb{L}_x$  la cantidad de extremos locales, y sabemos que  $L \ll N$ .

Tal es la situación que es necesario *regularizar* el problema, agregando un término que dependa de la norma (o una pseudo-norma) de la solución  $a$ . Para favorecer

soluciones suaves, tal y como lo hace la EMD original, elegimos una regularización de Tikhonov de segundo orden que usa como matriz de suavizado

$$D = \begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}, \quad (3.36)$$

la cual aproxima a una derivada segunda mediante una diferencia central de segundo orden [41].

Este proceso de regularización es crucial para que el problema tenga solución. Como la matriz  $P_x$  no es invertible, es esencial agregar el término de regularización al problema. Gracias a él, se obtiene un problema con una solución única ya que la matriz simétrica  $P_x^T P_x + \lambda D^T D$  (ver ec. (3.10)) es definida positiva si  $P_x$  se construye a partir de una señal oscilante:

$$\begin{aligned} y^T (P_x^T P_x + \lambda D^T D) y &= y^T P_x^T P_x y + \lambda y^T D^T D y \\ &= (P_x y)^T P_x y + \lambda (D y)^T D y > 0, \forall y \neq \bar{0}, \end{aligned} \quad (3.37)$$

debido a que  $\ker(P_x) \cap \ker(D) = \{\bar{0}\}$ :  $D y = \bar{0}$  sólo si  $y$  es una función afín; y el espacio nulo de  $P_x$  (Sec. 3.4.3) es el de un tipo particular de señales oscilantes. El caso de descomponer una función afín no nos interesa ya que no tendríamos extremos locales, la matriz  $P$  construida a partir de ella sería nula y no tendría siquiera sentido plantear nuestro problema. Si, en cambio,  $x$  es una señal oscilante (Def. 2.4.1), tendremos al menos dos extremos locales, nunca se daría  $D x = \bar{0}$ , y nuestro problema *siempre* tendría solución.

Para resolver de manera eficiente la inversión matricial presente en la ec. (3.10), debemos explorar la naturaleza de la matriz  $P_x^T P_x + \lambda D^T D$ . Para una señal de longitud  $N$ , la matriz  $D^T D$  tiene  $5N - 6$  elementos no nulos y la matriz  $P_x^T P_x$  tiene  $5L$  elementos no nulos, pero  $L$  de ellos se encuentran en la diagonal principal. Entonces, los elementos distintos de cero de la matriz  $P_x^T P_x + \lambda D^T D$  son

$$5N - 6 + 4L \leq 9N - 6 < 9N, \quad (3.38)$$

y así su proporción está acotada por  $9/N$ , lo cual deja a la vista que la matriz es *rala*. Otra característica importante es que la matriz es *diagonal en banda*, con ancho de banda de matriz igual  $bw = \max_{\ell} (t_{\ell} - t_{\ell+2})$  [44]. Todo esto se ilustra en la Fig. 3.5, donde mostramos los elementos no nulos de la matriz  $P_x$ , con  $x$  según la ec. (3.12). La rareza de las matrices  $P_x$  y  $D$  puede ser explotada en la implementación ahorrando una cantidad significativa de tiempo de cómputo. Existen rutinas de MATLAB<sup>®</sup> para el manejo eficiente de matrices ralas, y gracias a ellas logramos tiempos de cómputo similares a los de EMD clásica, y dos órdenes de magnitud menores a los tiempos de las otras propuestas de EMD basadas en optimización.

### 3.5.1. Selección del parámetro $\lambda$

A partir de la ec. (3.37) puede verse que sólo es necesario que  $\lambda$  sea positiva para que el problema tenga solución. Sin embargo, un valor demasiado pequeño llevaría

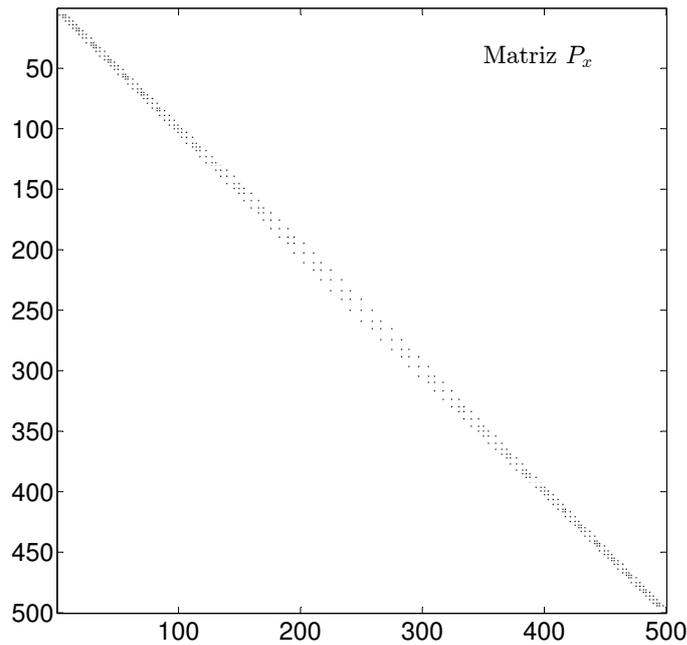


Figura 3.5: Elementos no nulos de  $P_x$ . Puede apreciarse la naturaleza rala de la matriz, y cómo los elementos no nulos se concentran alrededor de la diagonal principal. Tomamos como ejemplo la señal  $x$  de la ec. (3.12).

a un número de condición muy grande, y los resultados podrían no ser precisos. Por otro lado, a medida que  $\lambda$  tiende a infinito, la solución  $a$  se vuelve una función afín.

Una forma habitual para la selección del parámetro de regularización en un problema como el que nos compete es elegir su valor de acuerdo con los grados de libertad deseados. Los grados de libertad ( $GL$ ) de un problema de regularización como el de la ec. (3.18) están dados por la traza de la matriz que multiplica a los datos [44]:

$$GL_\lambda = \text{tr}(\Phi) = \text{tr}((P^T P + \lambda D^T D)^{-1} P^T P). \quad (3.39)$$

Una descomposición en valores singulares generalizada [11] de las matrices  $P$  y  $D$ , de la forma  $P = U\Lambda Y^T$  y  $D = VM^T$ , con  $U$  y  $V$  ortogonales,  $\Lambda$  y  $M$  diagonales, y  $Y$  no singular, nos permite escribir

$$GL_\lambda = \text{tr}(\Phi) = \text{tr}(Y^{-T}(\Lambda^T \Lambda + \lambda M^T M)^{-1} \Lambda^T \Lambda Y^T) = \sum_i \frac{b_i^2}{b_i^2 + \lambda c_i^2}, \quad (3.40)$$

donde  $b_i$  son los elementos de la matriz diagonal  $\Lambda$  y  $c_i$  los de la matriz diagonal  $M$ . La ec. (3.40) nos dice que los grados de libertad de la regularización están dados por una función *monótona* decreciente. La monotonía de la función nos permite optimizar su valor de forma sencilla, a través de métodos como los de la sección dorada y sus variaciones, o aquellos basados en la serie de Fibonacci [15]. Una opción lógica puede ser establecer los grados de libertad de nuestra regularización iguales a aquellos que resultan de una interpolación *spline* a través de los extremos locales. Si vemos a la interpolación como una regularización, puede demostrarse que los grados de libertad coinciden con la cantidad de nodos utilizados. En el caso de EMD, que usa dos interpolaciones (una para cada envolvente), los grados de libertad

resultan iguales a  $L/2$ , con  $L = \#\mathbb{L}_x$  la cantidad de extremos locales de la señal analizada.

Sin embargo, no debemos dejar de lado que EMD utiliza un proceso de *sifting*, en el cual existen sucesivas sustracciones de medias locales y por lo tanto la media local resultante es una *spline* con más nodos que cantidad de extremos de la señal original. Además, a medida que aumenta la cantidad de iteraciones de *sifting*, el ancho de banda marginal del modo es más estrecho. Ello es congruente con nuestro enfoque: a medida que disminuye  $\lambda$  aumenta la cantidad de grados de libertad y las «frecuencias de corte» de las columnas no nulas de  $\Phi$  también aumentan, con lo que el ancho de banda marginal del modo se hace más estrecho. De esto concluimos que el valor de  $\lambda$  debiera ser menor que aquél que otorga grados de libertad iguales a  $L/2$ .

Pueden usarse otros métodos para la selección del parámetro de regularización  $\lambda$ . El método de la curva L o el principio de discrepancia de Morozov [48] pueden ser empleados, pero tales opciones llevarían un alto costo computacional. Debido a ello y de acuerdo a nuestros estudios que muestran que nuestra propuesta se comporta de forma robusta a la opción de  $\lambda$ , elegimos en todos los casos  $\lambda = 1$ .

### 3.5.2. Efectos de borde

Como todos los métodos de análisis de señales, y debido a la longitud finita de éstas, nuestra propuesta presenta efectos de borde. Sin embargo, la finitud de los datos no es el único inconveniente: la función interpolante  $i_\ell(t)$  de la ec. (3.5) no está definida entre el borde izquierdo de la señal y el segundo extremo local, para penalizar al primer extremo local; ni entre el penúltimo extremo local y su borde derecho para penalizar al último extremo local (ver Fig. 3.1). Esta dificultad es compartida en su totalidad con el algoritmo tradicional de EMD (Sec. 2.4.7.1), debido a que las envolventes no están bien definidas para los mismos intervalos. Mencionaremos dos estrategias que pueden usarse: prolongación de los extremos, a través de un espejado en los bordes; y no penalizar el primero y el último extremo local.

Además de las cuestiones mencionadas en el párrafo anterior, que refieren únicamente a los extremos locales y por lo tanto a la matriz  $P$ , también debe observarse lo que ocurre con la matriz  $D$  (ver ec. (3.36)), la cual estima una derivada segunda. La estimación de una derivada mediante diferencias finitas presenta siempre «problemas» en algún borde (o en ambos) debido a la falta de datos. La matriz  $D$ , tal como se la presenta en la ec. (3.36), ofrece una estimación distinta en las muestras primera y última respecto de las demás (las filas primera y última de  $D$  poseen sólo dos elementos, mientras que el resto posee tres). Esta «deficiencia» en realidad supone que la solución es nula en ambos bordes (condiciones de borde de Dirichlet). Pero esta no es la única opción posible, ya que podrían desearse condiciones de borde distintas. Las siguientes son las posibles elecciones para  $D$ :

$$\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} -2 & 1 & \dots & \dots & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 1 & \dots & \dots & 1 & -2 \end{bmatrix}.$$

(Dirichlet)                      (Neumann)                      (Periódica)

### 3.5.3. Criterios de detención

Cualquier método que utilice los extremos locales de la señal para estimar los del modo extraído puede necesitar alguna forma de *sifting*. Esto es tan cierto para nuestro método como para el resto de las propuestas para EMD basadas en optimización [83, 92, 93]. Si algunos de los extremos del modo están «ocultos» (i.e. no aparecen en la señal compuesta), éstos pueden aparecer luego de una iteración de *sifting* [97, 98]. Entonces, el proto-modo tendría más extremos locales que la señal original, y la matriz  $P_x$  (construida a partir de los extremos de la señal original) no sería capaz de modelar la simetría. Sería necesaria la construcción de una nueva matriz  $P$  y por lo tanto una nueva sustracción de media local.

El enfoque más sencillo para establecer la cantidad de iteraciones de *sifting* es considerar un número fijo, lo cual no garantiza que se cumplan las condiciones de IMF (Def. 2.4.2). Propondremos aquí dos criterios, basándonos en aquellos anteriormente propuestos para EMD y discutidos en la Sec. 2.4.7.2, pero haciendo una pequeña salvedad. Como la versión original de EMD está basada en interpolaciones *spline*, luego la media local está enteramente determinada por las posiciones de los extremos locales y el valor de la función en tales posiciones. Dadas las condiciones de borde entonces, el *spline* es único para un conjunto de nodos (extremos locales en este caso) dados, y por lo tanto la única forma de tener una media local idénticamente nula es que todos los nodos de la *spline* tengan valor cero en su ordenada. De esto se desprende que para evaluar si la media local es nula es suficiente observar el comportamiento de ella en los nodos.

Como criterio global para evaluar si  $d_k$  es una IMF, puede implementarse la siguiente expresión:

$$GC = \frac{\|P_{d_k^i} d_k^i\|_2^2}{\|A_{d_k^i} d_k^i\|_2^2}, \quad (3.41)$$

donde  $d_k^i$  es la  $i$ -ésima iteración, y  $A_x$  es la matriz de «amplitud» que se define de forma tal que  $A_x x$  evalúe en cada extremo local la diferencia

$$x(t_\ell) - i_\ell(t_\ell), \quad (3.42)$$

donde  $i_\ell(t)$  es la función interpolante de la ec. (3.5). La construcción de la matriz  $A_x$  es tal que resulta  $A_x = 2\text{diag}(P_x) - P_x$ , donde el operador  $\text{diag}(\cdot)$  es aquel que construye una matriz diagonal con la diagonal principal de la matriz argumento (y donde hemos omitido los subíndices a los efectos de simplificar la notación). En definitiva, el criterio de la ec. (3.41) evalúa la simetría global de  $d_k$  normalizándola por una «amplitud global».

Un criterio de tipo global/local puede implementarse considerando la función

$$LC(\ell) = \frac{[P_{d_k^i} d_k^i](t_\ell)}{[A_{d_k^i} d_k^i](t_\ell)}, \quad (3.43)$$

la cual está definida sólo allí donde  $d_k^i$  posee un extremo local. Luego se pueden aplicar criterios locales similares a aquellos propuestos en [99]:  $LC(\ell) \leq \epsilon$  para una fracción  $1 - \alpha$  de los extremos locales; y  $LC(\ell) \leq \epsilon_2$  para la fracción  $\alpha$  restante.

### 3.5.4. Efectos de muestreo

Tal como ocurre con la EMD tradicional, cuyos problemas de muestreo fueron oportunamente discutidos en la Sec. 2.4.7.4, nuestra propuesta sufrirá de inconvenientes similares. El hecho radica nuevamente en la dificultad para ubicar correctamente los extremos locales, ya que a medida que disminuye la frecuencia de muestreo, los extremos locales se desplazan respecto a sus ubicaciones teóricas en la señal de tiempo continuo. En nuestro caso, se aplican en su totalidad los hallazgos reportados en [98], donde se ofrece una cota para el error del operador elemental de *sifting* (ver ec. (2.63)). Podemos corroborarlo empíricamente en la Fig. 3.6, donde puede verse la norma  $\|P_x(x - a)\|_1$  (que sirve como una medida de «error») en función del logaritmo de la frecuencia de muestreo  $f_s$ . Nuevamente, hemos usado como ejemplo la señal de la ec. (3.12).

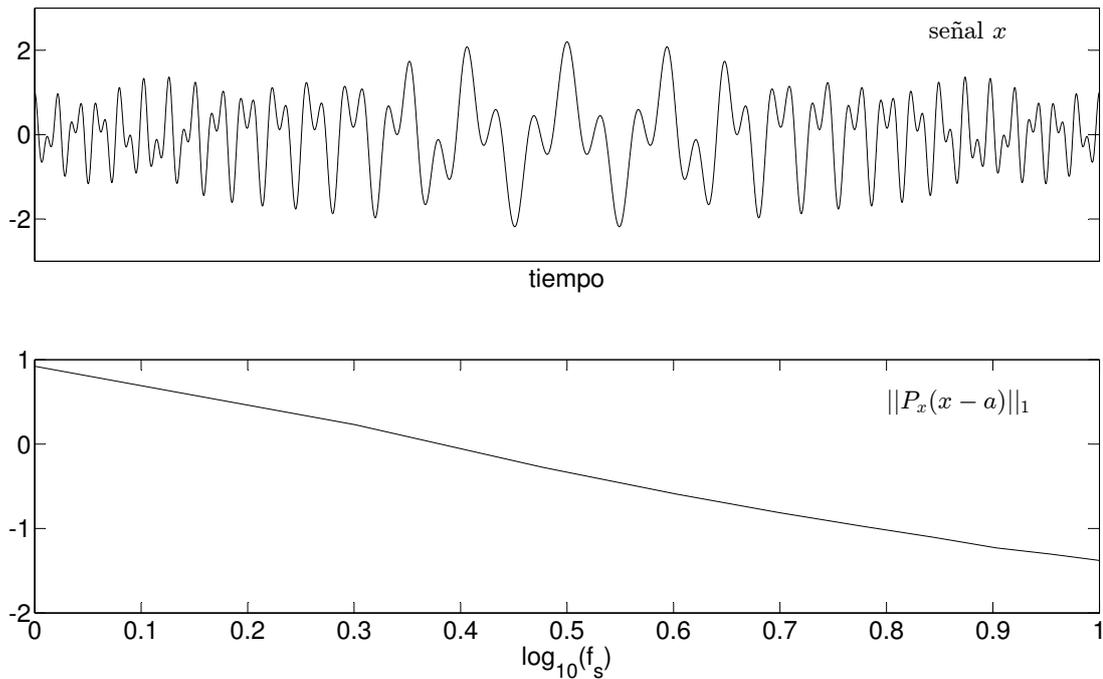


Figura 3.6: Efectos de muestreo. La norma  $\|P_x(x - a)\|_1$  disminuye a medida que aumenta la frecuencia de muestreo de la señal. La señal  $x$  está dada por la ec. (3.12).

## 3.6. Experimentos y resultados

Con el objetivo de comprender mejor las capacidades de este nuevo enfoque para EMD que presentamos, es que proponemos en esta sección revisar ejemplos clásicos de EMD con señales artificiales y comparar nuestros resultados con aquellos obtenidos mediante EMD tradicional, y las dos propuestas basadas en optimización descriptas en la Sec. 3.3.1. Además, ofrecemos un experimento con datos reales (señal

de voz) y comparamos los resultados obtenidos con nuestro método con aquellos obtenidos mediante EMD tradicional y versiones de EMD asistidas por ruido.

### 3.6.1. Separación de tonos

Como primer experimento, revisitamos el trabajo pionero de Rilling y Flandrin [97]. Definimos la suma de dos tonos puros como

$$x(t; \alpha, f) = s_1 + s_2 = \underbrace{\text{sen}(2\pi t)}_{s_1} + \underbrace{\alpha \text{sen}(2\pi ft)}_{s_2} \quad (3.44)$$

con  $f \in (0, 1)$  y  $\alpha > 0$ . Aunque muy sencillo, este modelo probó ser muy efectivo a la hora de evaluar el desempeño de EMD. Nuestra meta aquí es deducir (y luego medir) el desempeño de nuestro método cuando se lo aplica a  $x$ . La naturaleza «fina a tosca» de nuestra propuesta nos sugiere que el primer modo  $d_1$  debe ser similar a  $s_1$  (el tono más rápido). Nos interesa responder la siguiente pregunta: ¿bajo qué condiciones se da que  $d_1 \approx s_1$ ?

Podemos ver fácilmente que  $d_1 \approx s_1$  es equivalente a  $a_1 \approx s_2$ . Si aplicamos el Algoritmo 3.1, y por consiguiente la ec. (3.11) a nuestra señal  $x$ , obtenemos

$$a_1 = (P_x^T P_x + \lambda D^T D) x = (P_x^T P_x + \lambda D^T D) s_1 + (P_x^T P_x + \lambda D^T D) s_2, \quad (3.45)$$

donde  $P_x$  se construye a partir de los extremos locales de  $x$ . Resulta evidente que una condición necesaria para la separación de los dos tonos (i.e.  $a_1 \approx s_2$ ) es

$$(P_x^T P_x + \lambda D^T D)^{-1} P_x^T P_x s_1 \approx \bar{0} \quad (3.46)$$

y

$$\begin{aligned} (P_x^T P_x + \lambda D^T D)^{-1} P_x^T P_x s_2 &= (P_x^T P_x + \lambda D^T D)^{-1} (P_x^T P_x + \lambda D^T D - \lambda D^T D) s_2 \\ &= (I - (P_x^T P_x + \lambda D^T D)^{-1} \lambda D^T D) s_2 \\ &= s_2 - (P_x^T P_x + \lambda D^T D)^{-1} \lambda D^T D s_2 \approx s_2. \end{aligned} \quad (3.47)$$

Rilling y Flandrin [97] demostraron que si el producto  $\alpha f$  en la ec. (3.44) es suficientemente pequeño, entonces los extremos locales de  $x$  están cercanos a los de  $s_1$  y, dado que  $s_1$  es un tono puro con media local cero,  $P_x s_1 \approx \bar{0}$ , satisfaciendo (3.46). Además  $\alpha f$  menor que 1 asegura que la derivada cuarta de  $s_2$  (i.e.,  $D^T D s_2$ ) es pequeña, y si  $\lambda$  es también pequeño, entonces se satisface (3.47).

Ahora tenemos una respuesta a la pregunta que formulamos más arriba. Para  $\alpha f$  pequeño ( $\alpha f < 1$  es suficiente), y  $\lambda$  también pequeño, la separación será satisfactoria y tendremos  $d_1 \approx s_1$ . Para medir el desempeño de nuestro método seguimos a [97] y proponemos calcular

$$c(\alpha, f) = \frac{\|d_1(t; \alpha, f) - s_1\|_2}{\|s_2\|_2}, \quad (3.48)$$

donde  $d_1(t; \alpha, f)$  indica el primer modo de  $x(t; \alpha, f)$ . Un valor de  $c(\alpha, f)$  cercano a cero indica una separación correcta, mientras que un valor cercano a uno ocurre cuando los tonos son mal separados (o no separados en absoluto).

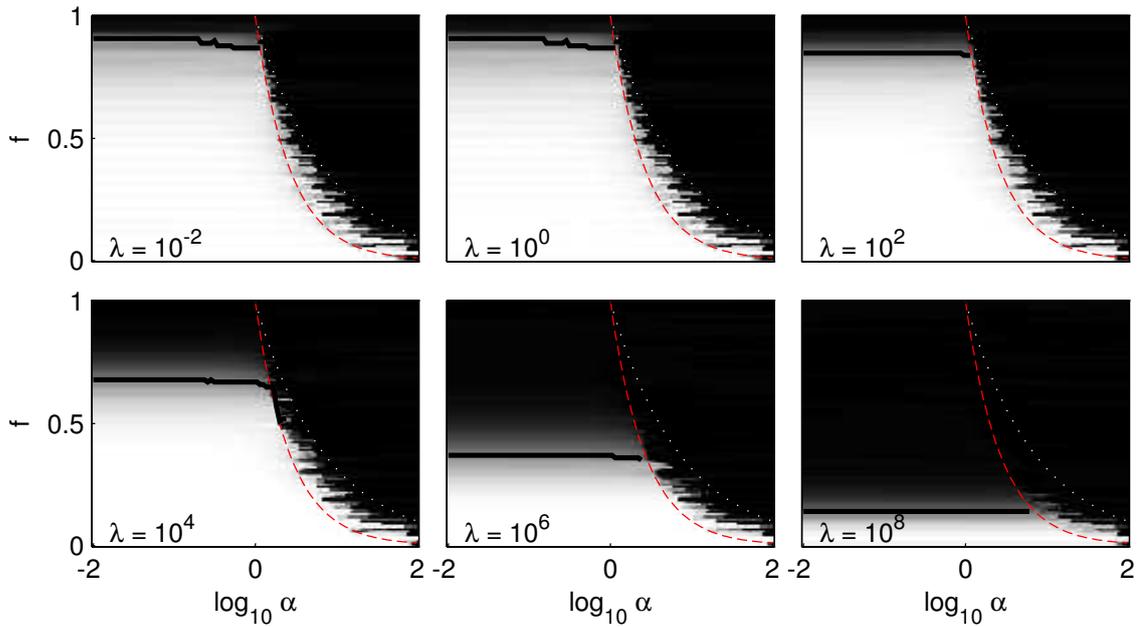


Figura 3.7: Separación de tonos. Graficamos en escala de grises los valores que toma la ec. (3.48). Las curvas críticas se superpusieron en línea de trazos roja ( $\alpha f = 1$ ) y línea punteada blanca ( $\alpha f^2 = 1$ ). La línea negra gruesa para  $\alpha f < 1$  representa la «frecuencia de corte» para la cual  $c(\alpha, f) = 0,5$ . En las imágenes, el color blanco significa un valor de 0 y el color negro un valor de 1.

Aunque el modelo (3.44) fue formulado en tiempo continuo, debemos realizar una implementación discreta por razones obvias. Usamos entonces señales de 2000 muestras, con una frecuencia de muestreo de 200 Hz. El parámetro  $f$  fue dividido en 100 particiones, y la ec. (3.48) fue evaluada sobre las muestras 501 a 1500 con el fin de evitar cualquier tipo de efecto de borde. La Fig. 3.7 presenta los resultados, con  $f \in (0, 1)$  y  $10^{-2} \leq \alpha \leq 10^2$ , para seis valores de  $\lambda$ . La semejanza con los resultados presentados en [97] es notable. En la fila de arriba pueden identificarse dos zonas bien diferenciadas. Para  $\alpha f < 1$  los dos tonos se separan perfectamente, excepto por encima de una frecuencia «de corte»  $f_c \approx 0,9$  donde es imposible separar los dos componentes, sin importar el valor de  $\alpha$ . Este resultado es consistente con nuestras expectativas de que los dos tonos podrían separarse correctamente si  $\alpha f$  es suficientemente pequeño. Por otro lado, para  $\alpha f^2 > 1$ , los dos tonos son separados de manera incorrecta. Rilling y Flandrin mostraron que, para este dominio, el tono más lento es «dominante» en el sentido que la tasa de extremos de la señal es igual a la tasa de extremos del tono de baja frecuencia. Entonces, resulta lógico que nuestro método (al igual que el EMD tradicional) sea incapaz de aislar en su primer modo al tono más rápido si los extremos de éste están ausentes de la señal compuesta. Existe una zona de transición (una especie de «tierra de nadie») entre los dominios mencionados donde el desempeño de la separación parece estar principalmente condicionado por la frecuencia  $f$ . Resulta conveniente mencionar aquí que para alcanzar una frecuencia de corte  $f_c \approx 0,9$  con EMD tradicional, son necesarias más de 100 iteraciones de *sifting* [97]. En nuestro caso, usamos solamente una iteración.

En la fila de abajo de la Fig. 3.7, los resultados son todavía similares a los

anteriores, pero también distintos en algunos aspectos. Las dos áreas (junto con la zona de transición) pueden reconocerse aún, pero la frecuencia de corte es más pequeña a medida que  $\lambda$  crece. Estos resultados sugieren que podemos seleccionar  $\lambda$  a partir de un amplio rango de valores, lo cual evidencia cierta robustez del método. En lo que sigue, y sólo por cuestiones de simpleza, usaremos  $\lambda = 1$ , un valor para el cual obtuvimos muy buenos resultados en situaciones muy diferentes, como se verá en los siguientes experimentos.

### 3.6.2. Mezclas de oscilaciones AM-FM y no lineales

Presentaremos en esta subsección tres ejemplos para analizar el desempeño del método propuesto en comparación con otros del estado del arte.

**Ejemplo 1.** El primer experimento de esta subsección lo realizamos sobre una señal muy similar a una estudiada en [92]. Discretizada para 1000 muestras, consiste de la suma de un tono rápido y un tono lento de amplitudes diferentes y un átomo gaussiano con una frecuencia intermedia. Presentamos en la Fig. 3.8 los resultados de las descomposiciones obtenidas con EMD tradicional, Prox-EMD, OS-EMD y nuestro método UOA-EMD. En todos los casos, la cantidad de modos  $K = 3$  se supuso conocida. El parámetro de regularización usado para UOA-EMD fue  $\lambda = 1$ , y se realizó sólo una iteración de *sifting* para la extracción de cada componente. Para comparar los resultados con los componentes originales de la señal, usamos el error cuadrático relativo (RSE, del inglés *Relative Square Error*):

$$RSE_b(a) = \frac{\|a - b\|_2^2}{\|b\|_2^2}, \quad (3.49)$$

donde  $b$  es una señal dada y  $a$  su estimación. Tanto Prox-EMD como OS-EMD fueron usados con sus parámetros por defecto. Para el EMD tradicional, usamos la *toolbox* disponible en <http://perso.ens-lyon.fr/patrick.flandrin/emd.html>, con los criterios de detención por defecto. En la Tabla 3.1 se presentan los RSEs y tiempos de cómputo para los cuatro métodos. Para los tres componentes, los errores de EMD y UOA-EMD son de magnitudes comparables, mientras que los errores de los otros métodos son al menos un orden de magnitud mayores. Los tiempos de cómputo también merecen ser observados detenidamente. El tiempo para nuestro método UOA-EMD es seis veces menor que el de EMD, y dos órdenes de magnitud más bajo que aquellos de los otros dos métodos. (Aclaremos que ejecutamos todas las simulaciones en una PC con procesador Intel(R) Core(TM) i7-4770K CPU @ 3.50GHz).

**Ejemplo 2.** Como segundo experimento de esta subsección, descomponemos mediante los cuatro métodos una señal muy similar a una estudiada en [38]. La señal se compone de la suma de tres componentes: una forma de onda triangular de «frecuencia alta» con su amplitud creciendo linealmente y de forma muy lenta, un tono de «frecuencia media» cuya amplitud decrece rápidamente de forma lineal, y una forma de onda triangular de «frecuencia baja». Al igual que en el caso anterior, la señal discreta posee 1000 muestras. Los resultados se muestran en la Fig. 3.9. Nuevamente, usamos  $\lambda = 1$  y sólo una iteración de *sifting* para UOA-EMD. Los errores y tiempos de cómputo se presentan en la Tabla 3.2. Los valores recuerdan a aquellos del ejemplo previo. Los errores de EMD y UOA-EMD son de magnitudes

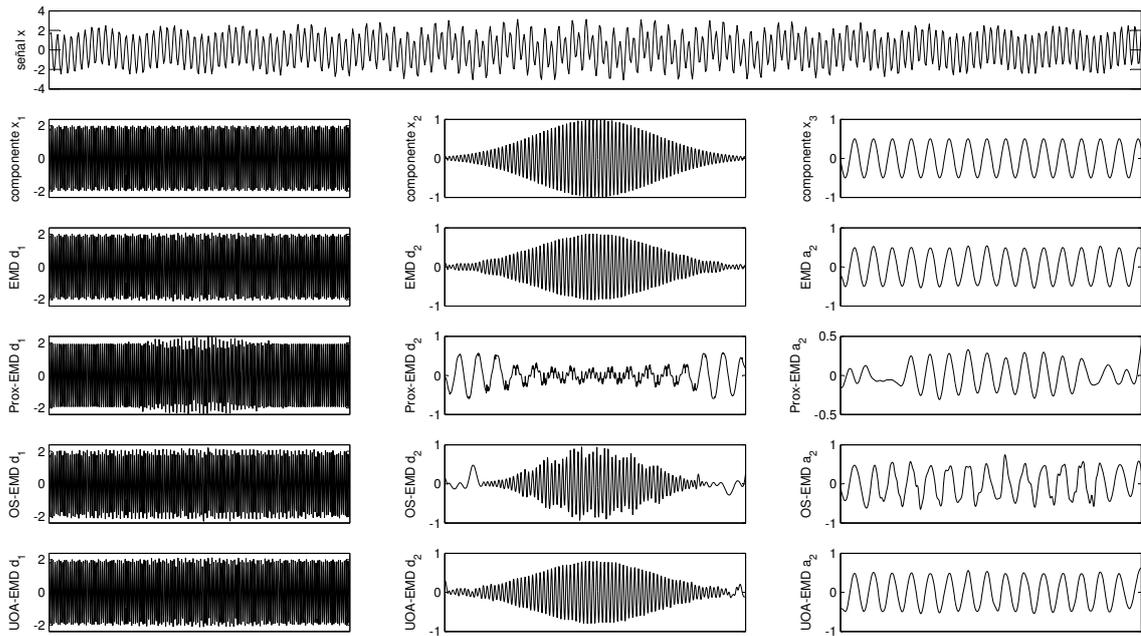


Figura 3.8: Ejemplo 1. Suma de dos tonos y un átomo gaussiano. La señal original  $x$  se muestra en la fila superior, y sus componentes  $x_1$ ,  $x_2$  y  $x_3$  en la segunda fila. En las filas siguientes se muestran los modos obtenidos mediante EMD, Prox-EMD, OS-EMD y UOA-EMD respectivamente.

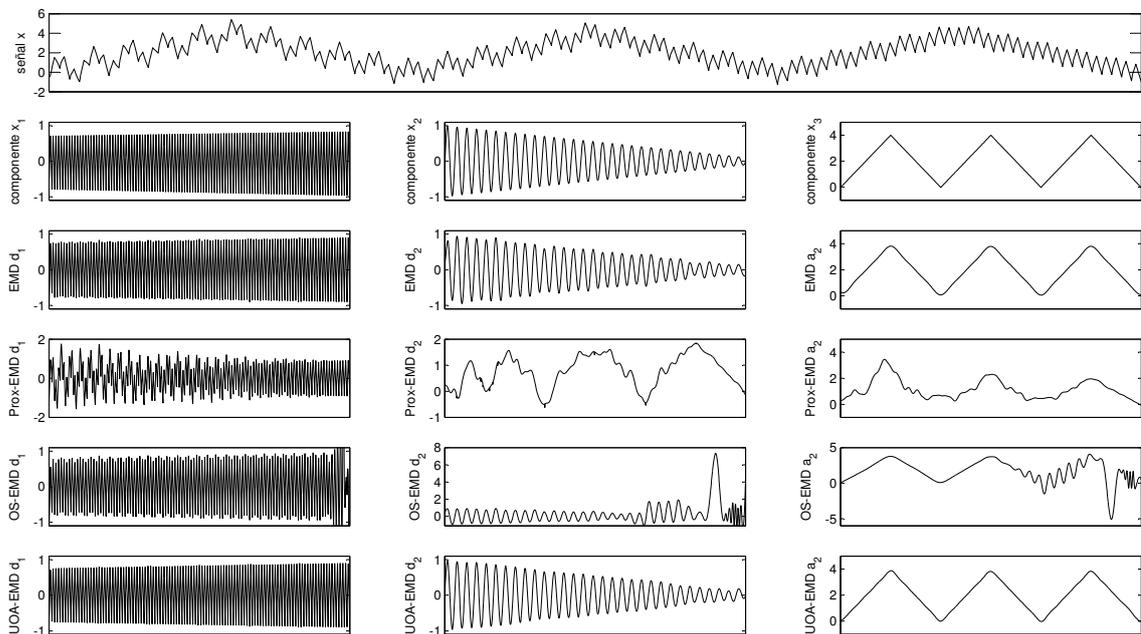


Figura 3.9: Ejemplo 2. Suma de dos formas de onda triangulares y una señal sinusoidal FM. La señal original  $x$  se muestra en la fila superior, y sus componentes  $x_1$ ,  $x_2$  y  $x_3$  en la segunda fila. En las filas siguientes se muestran los modos obtenidos mediante EMD, Prox-EMD, OS-EMD y UOA-EMD respectivamente.

comparables y al menos un orden de magnitud menores que aquellos de Prox-EMD y OS-EMD (con esta diferencia creciendo a dos órdenes de magnitud para dos modos

de Prox-EMD). Los tiempos de cómputo también son similares a los del ejemplo previo. UOA-EMD necesitó un sexto del tiempo de EMD, y dos órdenes de magnitud menos que el necesitado por Prox-EMD y OS-EMD.

**Ejemplo 3.** Como tercer y último experimento de esta subsección, analizamos una señal muy similar a la usada en [99]. Es una suma de dos señales moduladas senoidalmente en frecuencia y un átomo gaussiano, superpuestas todas ellas tanto en tiempo como en frecuencia. Con el objetivo de evaluar nuestro método en señales de longitud extensa es que decidimos muestrear la señal con un millón de muestras. Los resultados de las descomposiciones pueden apreciarse en la Fig. 3.10. Como en los ejemplos anteriores, usamos  $\lambda = 1$  y sólo una iteración de *sifting*. La Tabla 3.3 presenta los RSEs y tiempos de cómputo. Nuestro método UOA-EMD logra errores un orden de magnitud menores a los de EMD para dos de los tres componentes, y al menos dos órdenes de magnitudes menores a los de Prox-EMD. El método OS-EMD no logró terminar la descomposición, imprimiendo un mensaje de error luego de casi 12 minutos de ejecución. Los tiempos de cómputo merecen algunos comentarios. EMD necesitó poco más de un segundo para realizar la descomposición, mientras que UOA-EMD tomó casi tres segundos. Contrastando esta situación, Prox-EMD necesitó más de 47 minutos.

En esta subsección ilustramos las capacidades de nuestro método con señales artificiales clásicas, descomponiendo exitosamente señales con componentes AM-FM y oscilaciones no lineales superpuestas en tiempo y en frecuencia. En todos los casos los resultados son similares a los obtenidos con EMD. Los errores cuadráticos relativos son del mismo orden de magnitud. El costo computacional de nuestro método es similar al de EMD. Por el contrario, los otros dos métodos basados en optimización lograron errores al menos un orden de magnitud por encima y tiempos de al menos dos órdenes de magnitud mayores. Estas diferencias se incrementaron significativamente cuando analizamos una señal de un millón de muestras.

### 3.6.3. Ruido blanco gaussiano y banco de filtros equivalente

Exploraremos ahora un esquema iterativo con nuestro método aplicado a una señal de ruido blanco gaussiano (wGn, del inglés *white Gaussian noise*) con  $\mathcal{N} \sim (0, 1)$  y de 512 muestras de largo. Realizamos iteraciones de *sifting* hasta que se cumple la condición de *weak-IMF* (condición (I), Def. 2.4.2) [102]. En la Fig. 3.11 ilustramos el proceso de *sifting* para el primer modo. Usamos  $\lambda = 1$ . Los puntos más notables donde el candidato a primer modo viola la condición (I) están marcados con círculos; se puede observar cómo nuestro método corrige esas violaciones a la

	Error cuadrático relativo			Tiempo de cómputo (segundos)
	$x_1(t)$	$x_2(t)$	$x_3(t)$	
EMD	$2,30 \times 10^{-3}$	$2,42 \times 10^{-2}$	$3,70 \times 10^{-3}$	0,0717
Prox-EMD	$5,29 \times 10^{-2}$	1,1615	$5,88 \times 10^{-1}$	3,1490
OS-EMD	$1,19 \times 10^{-2}$	$1,59 \times 10^{-1}$	$9,39 \times 10^{-2}$	7,0712
UOA-EMD	$4,00 \times 10^{-3}$	$4,27 \times 10^{-2}$	$8,30 \times 10^{-3}$	<b>0,0110</b>

Tabla 3.1: Ejemplo 1. Suma de dos tonos y un átomo gaussiano (1000 muestras). Errores cuadráticos relativos y tiempos computacionales para los cuatro métodos.

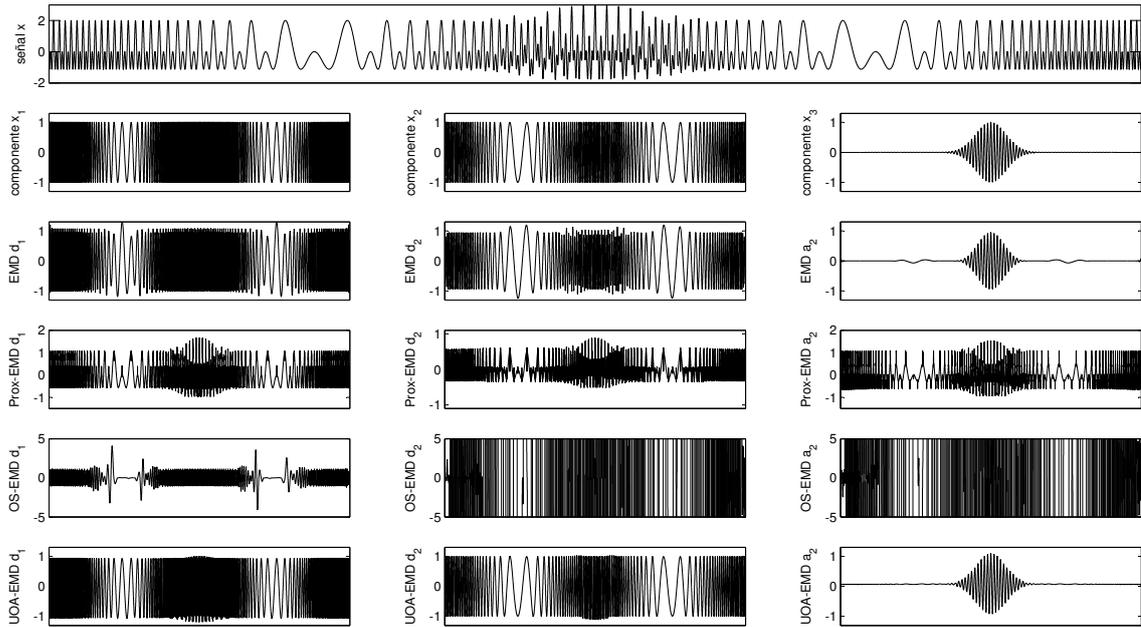


Figura 3.10: Ejemplo 3. Suma de dos señales AM-FM y un átomo gaussiano. La señal original  $x$  se muestra en la fila superior, y sus componentes  $x_1$ ,  $x_2$  y  $x_3$  en la segunda fila. En las filas siguientes se muestran los modos obtenidos mediante EMD, Prox-EMD, OS-EMD y UOA-EMD respectivamente.

	Error cuadrático relativo			Tiempo de cómputo (segundos)
	$x_1(t)$	$x_2(t)$	$x_3(t)$	
EMD	$1,50 \times 10^{-2}$	$8,10 \times 10^{-3}$	$8,00 \times 10^{-4}$	0,0679
Prox-EMD	$7,14 \times 10^{-1}$	6,2173	$2,05 \times 10^{-1}$	3,0362
OS-EMD	$4,06 \times 10^{-1}$	7,366	$2,53 \times 10^{-1}$	5,4287
UOA-EMD	<b><math>1,27 \times 10^{-2}</math></b>	<b><math>3,30 \times 10^{-3}</math></b>	<b><math>7,00 \times 10^{-4}</math></b>	<b>0,0104</b>

Tabla 3.2: Ejemplo 2. Suma de dos formas de onda triangulares y una señal AM (1000 muestras). Errores cuadráticos relativos y tiempos computacionales para los cuatro métodos.

quinta iteración. Luego de cada iteración, mostramos la diferencia entre la cantidad de extremos y cruces por cero (ver Def. 2.4.2). Usamos la notación  $d_k^j$  del Algoritmo 2.1 para indicar la  $j$ -ésima iteración de *sifting* del  $k$ -ésimo modo.

En el pasado se llevaron a cabo extensas simulaciones sobre ruido blanco gaussiano con el objetivo de evaluar el comportamiento de EMD. Se reveló así una estructura de banco de filtros cuasi-diádico cuando se aplica EMD a ruido blanco gaussiano [38, 40, 116]. Las mismas características aparecen cuando se analiza ruido blanco gaussiano con nuestro método. Siguiendo a Flandrin y col. [38], calculamos funciones de transferencia equivalentes para nuestro método cuando se lo aplica a ruido blanco gaussiano. Descompusimos cinco mil realizaciones ( $\lambda = 1$  y un número fijo de diez iteraciones de *sifting*) y estimamos el comportamiento promedio como en [38]:

	Error cuadrático relativo			Tiempo de cómputo (segundos)
	$x_1(t)$	$x_2(t)$	$x_3(t)$	
EMD	$2,04 \times 10^{-2}$	$2,40 \times 10^{-2}$	<b><math>5,67 \times 10^{-2}</math></b>	<b>1,1526</b>
Prox-EMD	$5,20 \times 10^{-1}$	$6,78 \times 10^{-1}$	2,6531	2854
OS-EMD	1,1321*	$5,35 \times 10^{8*}$	$6,05 \times 10^{9*}$	711*
UOA-EMD	<b><math>8,20 \times 10^{-3}</math></b>	<b><math>2,60 \times 10^{-3}</math></b>	$1,20 \times 10^{-1}$	2,8793

Tabla 3.3: Ejemplo 3. Suma de dos señales AM-FM y un átomo gaussiano (un millón de muestras). Errores cuadráticos relativos y tiempos computacionales para los cuatro métodos. (\* El algoritmo no logró terminar la descomposición e imprimió un mensaje de error.)

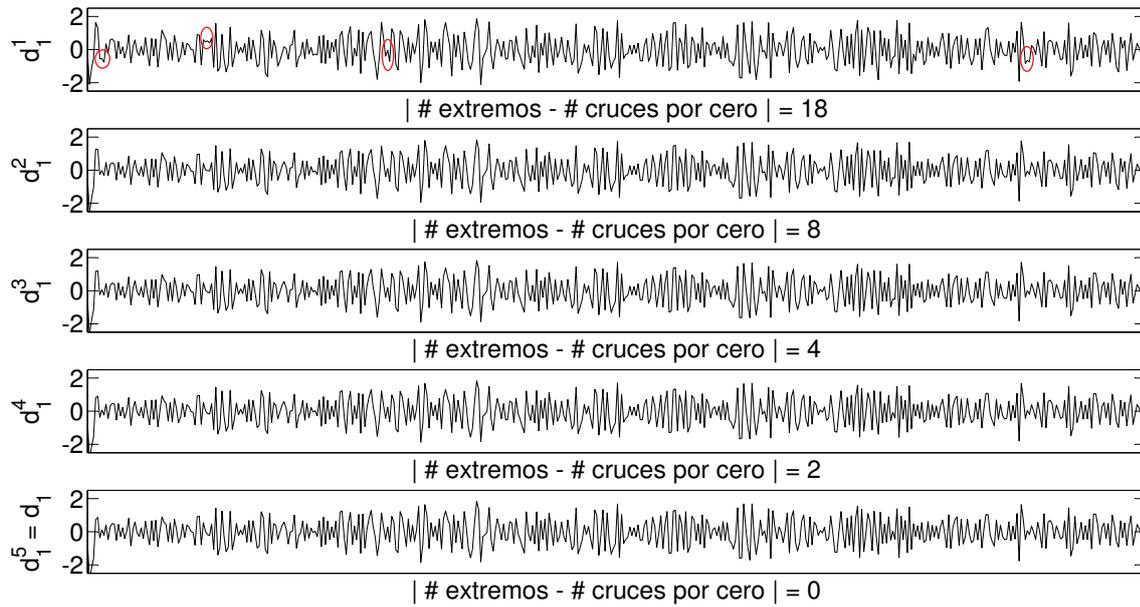


Figura 3.11: Proceso de *sifting* para el primer modo de una realización de ruido blanco gaussiano. Marcamos en rojo aquellos puntos más notables donde el candidato a primer modo viola la condición (I) de la definición de IMF (Def. 2.4.2). Puede apreciarse que nuestro método corrige tales violaciones luego de 5 iteraciones. Usamos la notación  $d_k^j$  del Algoritmo 2.1.

$$\hat{S}_k(f) = \sum_{m=-N+1}^{N-1} r_k[m] w[m] e^{-i2\pi f m}, \quad |f| \leq \frac{1}{2}, \quad (3.50)$$

con  $w[m]$  siendo una ventana de Hamming,  $N$  la cantidad de muestras, y

$$r_k[m] = \frac{1}{J} \sum_{j=1}^J \left( \frac{1}{N} \sum_{n=1}^{N-|m|} d_k^{(j)}[n] d_k^{(j)}[n+|m|] \right), \quad |m| \leq N-1 \quad (3.51)$$

el promedio (sobre las  $J$  realizaciones) de las estimaciones empíricas de la función de autocorrelación donde  $d_k^{(j)}$  es el  $k$ -ésimo modo de la  $j$ -ésima realización de ruido blanco gaussiano. Los resultados se presentan en la Fig. 3.12. Puede apreciarse que éstos son muy similares a los informados en [40, 116]. El fenómeno de autosimilaridad parece emerger. En el panel derecho graficamos  $\log_2 f_k^*$  vs.  $k$ , donde

$f_k^* = \arg \max_f |\hat{S}_k(f)|$ , junto a un ajuste lineal superpuesto. La pendiente  $\rho = 1,53$  del ajuste verifica la relación  $f_k^* \propto \rho^{-k}$  y podemos afirmar (como en [39, 40]) que

$$\hat{S}_{k'}(f) = \hat{S}_k(\rho^{k'-k} f) \quad (3.52)$$

para  $k' > k \geq 2$ . En [95] se mostró que el coeficiente de autosimilaridad  $\rho$  decrece a medida que la cantidad de iteraciones de *sifting* crece. El valor relativamente bajo de  $\rho$  que obtuvimos en nuestro ejemplo (lejos del caso cuasi-diádico donde  $\rho \approx 2$  para diez iteraciones de *sifting* [95, 116]) indica que una iteración de nuestro método es equivalente a varias iteraciones con EMD. Esta característica ya había sido observada en la Sec. 3.6.1.

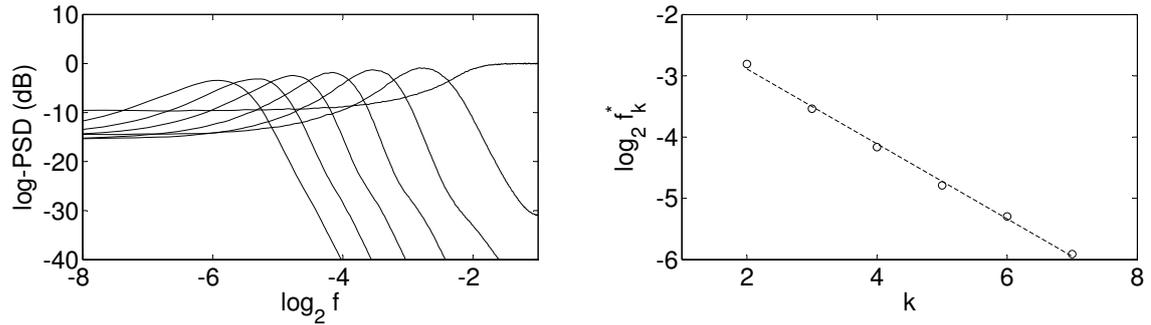


Figura 3.12: Descomposiciones de ruido blanco gaussiano. Izquierda: Funciones de transferencia equivalentes  $\hat{S}_k(f)$ ,  $1 \leq k \leq 7$ . Descompusimos 5000 realizaciones de wGn mediante nuestro método y estimamos el comportamiento promedio de acuerdo a las ecuaciones (3.50) y (3.51), siguiendo a [39, 40]. Derecha: Posiciones de los máximos de las funciones de transferencia equivalentes. El logaritmo en base 2 de  $f_k^*$  se grafica contra el índice de modo  $k$ . Se superpuso un ajuste lineal y el factor de escalamiento estimado fue  $\rho = 1,53$ .

### 3.6.4. Descomposición de señales de ancho de banda localmente estrecho

En la descomposición basada en operadores presentada en [87], y oportunamente discutida en la Sec. 3.4.3, los componentes resultantes se denominan «señales de ancho de banda localmente estrecho» (en inglés *local narrow band signals*). En su trabajo, los autores indican la incapacidad de su método para separar la señal  $x(t) = \sin(2t) - \cos(4t)/3$ . Presentamos en la Fig. 3.13 la descomposición de tal señal mediante el método aquí propuesto UOA-EMD y EMD tradicional. Usamos  $\lambda = 1$  y dos iteraciones de *sifting*. Nuestro método logró los siguientes RSEs:  $1,24 \times 10^{-2}$  y  $1,40 \times 10^{-3}$  para el primer y segundo modo respectivamente. Los RSEs para EMD fueron mayores:  $5,37 \times 10^{-2}$  y  $5,90 \times 10^{-3}$ .

Los autores en [87] también afirman que la señal  $x(t) = \sin(2t) - (1 + \cos(4t))/4 + \sin(8t) - \cos(16t)/3$  no puede ser descompuesta ni por su propio enfoque ni por EMD tradicional. Debido a esta razón, propusieron un enfoque híbrido para realizar la tarea. Al contrario, nuestro método es capaz de descomponer  $x(t)$  por sí mismo, sin necesidad de un procesamiento posterior. Los resultados de las descomposiciones obtenidas por EMD y por nuestro método UOA-EMD pueden apreciarse en la Fig.

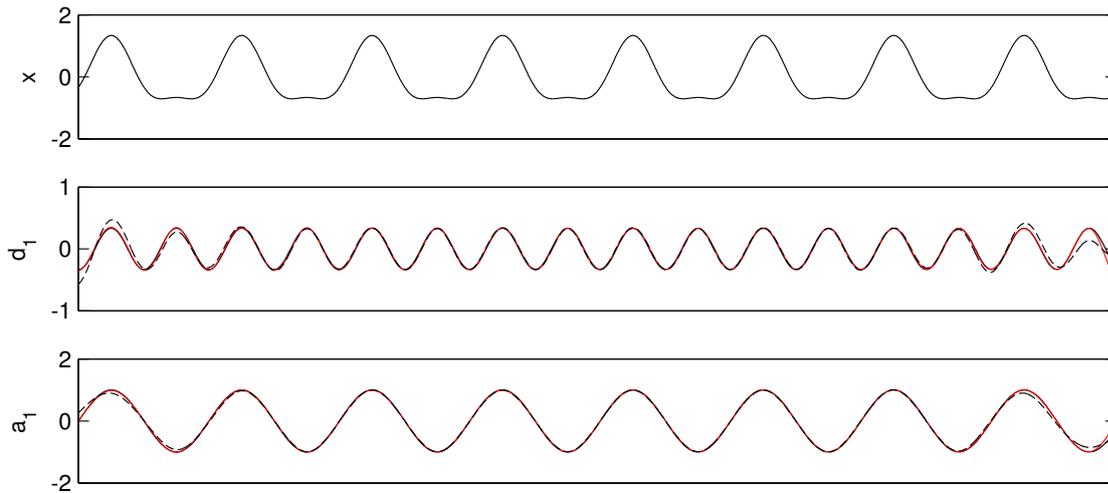


Figura 3.13: Señales de ancho de banda local estrecho. La señal  $x(t) = \sin(2t) - \cos(4t)/3$  se muestra en la fila superior. En la segunda y tercera fila se presentan los componentes originales (línea negra sólida), los resultados con UOA-EMD (línea roja sólida) y los resultados con EMD tradicional (línea negra de trazos). La señal  $x(t)$  no puede descomponerse correctamente mediante la propuesta de [87].

3.14. Mientras nuestra propuesta devuelve los cuatro componentes (excepto por un *offset* irrelevante en  $d_3$  y  $a_3$ ), EMD no puede separar los dos componentes más lentos en dos modos diferentes. Se necesitaron tres iteraciones de *sifting* y, como en los ejemplos previos, usamos  $\lambda = 1$  (al contrario de [87], donde el valor del parámetro de regularización debe ser cuidadosamente elegido para cada uno de los experimentos).

### 3.6.5. Señal real de voz

Cuando EMD se aplica a señales reales, experimenta frecuentemente problemas de mezcla de modos, especialmente en los casos ruidosos (Cap. 5). Aunque EEMD (del inglés *Ensemble EMD*) [118] alivia este problema, también crea algunas dificultades nuevas. El nuevo algoritmo de CEEMDAN (del inglés *Complete EEMD with Adaptive Noise*) [27, 107] soluciona los problemas persistentes en EEMD. Discutiremos en profundidad estas versiones de EMD asistidas por ruido en el Cap. 5; aquí sólo las mencionamos ya que compararemos con ellas los resultados de UOA-EMD.

Para ilustrar las ventajas de nuestro método cuando se lo aplica a señales reales, descompusimos una señal de voz (vocal /a/ sostenida). Usamos  $\lambda = 1$ , y un número máximo de iteraciones de *sifting* igual a 30. Los resultados obtenidos pueden apreciarse en la columna de la izquierda de la Fig. 3.15. También descompusimos la misma señal mediante EMD, EEMD y CEEMDAN. Usamos el mismo criterio de detención que para EMD, sin limitar las iteraciones de *sifting*. En ambos métodos asistidos por ruido (EEMD y CEEMDAN) se usaron los mismos parámetros: 0, 2 para la amplitud de ruido agregado [28] y 20 realizaciones de ruido.

Cuando se analizan señales cuasiperiódicas como la aquí estudiada, uno espera extraer componentes cuasiperiódicos también. Ese es justamente el caso para nuestro

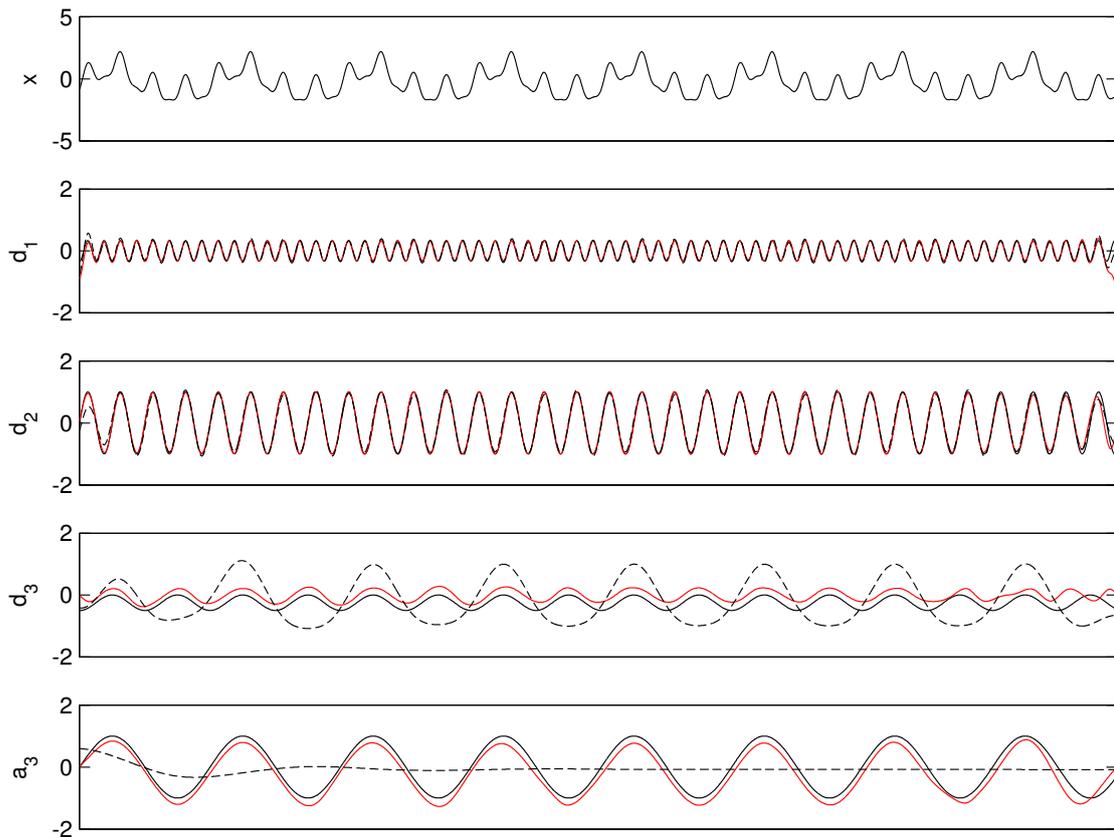


Figura 3.14: Señales de ancho de banda local estrecho. La señal  $x(t) = \sin(2t) - (1 + \cos(4t))/4 + \sin(8t) - \cos(16t)/3$  se muestra en la fila superior. En la segunda y tercera fila se presentan los componentes originales (línea negra sólida), los resultados con UOA-EMD (línea roja sólida) y los resultados con EMD tradicional (línea negra de trazos).

método UOA-EMD. En la columna de la izquierda de la Fig. 3.15 los dos primeros modos parecen contener información relacionada principalmente al ruido, mientras que los modos tercero a sexto contienen la mayor parte de la información relevante de la señal. Más aún, el sexto modo parece capturar casi perfectamente a  $F_0$ , la frecuencia fundamental de la señal de voz (inversa de la distancia entre impulsos glóticos sucesivos). Al contrario de lo recién expuesto, EMD (en la segunda columna de la Fig. 3.15) presenta mezcla de modos entre el segundo y tercer modo, y entre el tercer y cuarto modo. Además, EMD no es capaz de capturar la frecuencia fundamental. En la tercera columna, los modos cuarto a séptimo de EEMD son regulares y cuasiperiódicos, aunque se observan mínimos positivos y máximos negativos en el cuarto modo, en clara violación de la definición de IMF (Def. 2.4.2). La tendencia de la señal se reparte entre los modos noveno a undécimo, mostrando claramente una dificultad de EEMD (volveremos a esto en el Cap. 5). Los modos de CEEMDAN (última columna de la Fig. 3.15) también son regulares, pero el modo cuarto viola la definición de IMF. Puede apreciarse entonces, que en el caso de esta aplicación, los resultados se muestran a favor del método que aquí hemos propuesto.

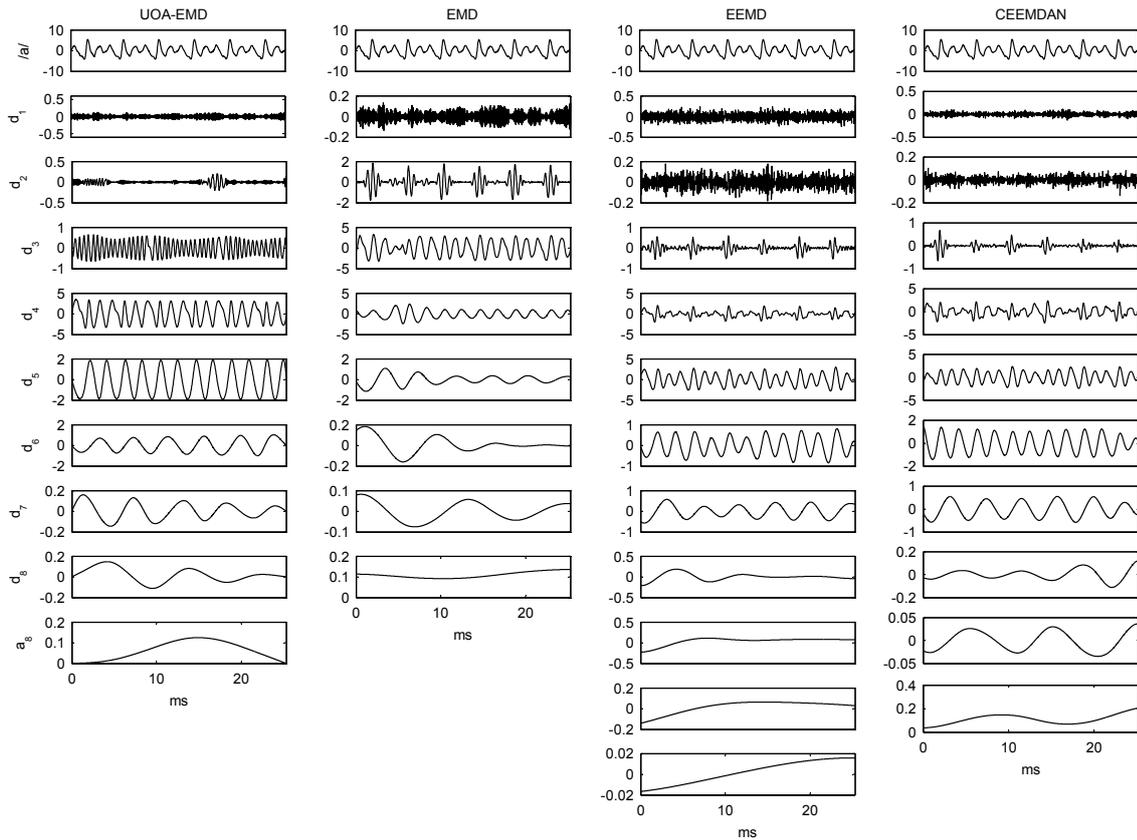


Figura 3.15: Descomposición de una señal de voz. Nuestro método UOA-EMD logra la descomposición más regular de todas, capturando la frecuencia fundamental  $F_0$  de manera casi perfecta. EMD no es capaz de hacerlo. El séptimo modo de EEMD captura  $F_0$  de manera aceptable, pero pierde el último ciclo. El séptimo modo de CEEMDAN captura a  $F_0$ .

### 3.7. Extensión a dos dimensiones

La separación de imágenes no estacionarias en componentes frecuenciales localmente disjuntos, que pueden no estarlo globalmente, es a menudo necesaria para aplicaciones como clasificación de texturas [19, 47, 75], reconocimiento de rostros [7, 72, 105], y marcas de agua digitales [103]. Se considera entonces un modelo similar al de la ec. (2.1) pero en dos dimensiones:

$$X(x, y) = \sum_{k=1}^K \alpha_k(x, y) \cos \phi_k(x, y) + A_K(x, y), \quad (3.53)$$

donde cada *detalle local* o modo es  $D_k(x, y) = \alpha_k(x, y) \cos \phi_k(x, y)$ . Al igual que en el caso unidimensional, uno de los objetivos principales es la extracción de cantidades locales como amplitudes, fases y frecuencias. Sin embargo, la descomposición de la imagen original en sus modos es un fin en sí mismo.

Dado que EMD provee una separación de la forma de la ec. (3.53), y debido a su adaptividad y flexibilidad, es que diversos investigadores se vieron motivados para adaptarla para datos bidimensionales. Nunes y col. [81] desarrollaron un *bidimensional EMD* para el cual usaron operadores morfológicos para detectar

máximos regionales y funciones de base radial para la interpolación de superficies. Linderhed [74] presentó una *image EMD* donde las envolventes se encuentran a partir de *thin plate splines*. Wu y col. [120] propusieron un enfoque multidimensional en el cual descomponen filas y columnas de manera independiente y combinan los resultados basándose en una estrategia de mínima escala.

Todos estos métodos, sin embargo, se asientan en procesos iterativos, como el *sifting*. Los enfoques basados en optimización, como los discutidos en la Sec. 3.3, y desarrollados para señales 1D evitan este proceso iterativo [83, 93]. Nuestro enfoque para señales 1D, basado en un problema de optimización *sin restricciones*, y desarrollado en el presente capítulo, es significativamente más sencillo en su formulación y más rápido que los mencionados. Es por ello que propondremos aquí una extensión de nuestro método para datos 2D. Logramos un método independiente de la orientación (isótropo) que implementa una descomposición EMD para imágenes. Mostraremos que, en comparación con otros enfoques de EMD en 2D, nuestra propuesta es computacionalmente eficiente, isótropa (no depende de la orientación) y logra un mejor desempeño al descomponer imágenes AM-FM.

### 3.7.1. Nuestro enfoque

Consideremos ahora una función real de dos variables reales  $X = X(x, y)$ , la cual en el caso discreto se ve como una matriz  $X \in \mathbb{R}^{N_1 \times N_2}$ . El conjunto contable de extremos locales es

$$\mathbb{L}_X = \left\{ z_\ell = (x_\ell, y_\ell) / X(z_\ell) > X(z) \vee X(z_\ell) < X(z), \right. \\ \left. 0 < |z - z_\ell| < \delta_\ell \right\}. \quad (3.54)$$

Esta definición de extremos locales puede variar dependiendo de si se requiere o no una desigualdad estricta. Nuestro objetivo es encontrar una *aproximación*  $A \in \mathbb{R}^{N_1 \times N_2}$  tal que la penalización (en un sentido de simetría, como en  $\|P_x(x-a)\|_2^2$  para señales 1D) del modo  $D = X - A$  sea minimizada. Para hacerlo, debemos aplicar la ec. (3.6) en un sentido 2D. Si  $z_\ell$  es un máximo (resp. mínimo) local, definamos  $z_i, i = 1, 2, 3$ , como los tres mínimos (resp. máximos) locales no colineales más cercanos [101]. Entonces, la penalización sobre  $z_\ell$  es

$$\langle P_\ell, D \rangle = D(z_\ell) + c_1 D(z_1) + c_2 D(z_2) + c_3 D(z_3), \quad (3.55)$$

donde  $\langle \cdot, \cdot \rangle$  indica el producto interno entre matrices,  $P_\ell \in \mathbb{R}^{N_1 \times N_2}$  es una matriz con coeficientes no nulos sólo en  $P_\ell(z_i) = c_i, i = 1, 2, 3$ , y los coeficientes  $c_i$  son los necesarios para comparar la función evaluada en el punto  $D(z_\ell)$  con el plano que intersecta a  $D(z_i)$ :

$$c_2 = \frac{(x_\ell - x_1)(y_3 - y_1) - (x_3 - x_1)(y_\ell - y_1)}{(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)} \quad (3.56a)$$

$$c_3 = \frac{(x_2 - x_1)(y_\ell - y_1) - (x_\ell - x_1)(y_2 - y_1)}{(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)} \quad (3.56b)$$

$$c_1 = 1 - c_2 - c_3. \quad (3.56c)$$

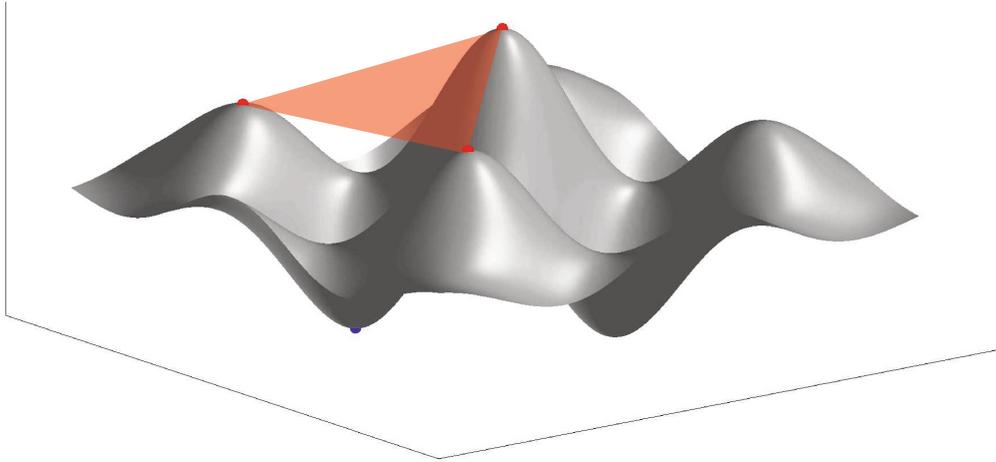


Figura 3.16: Función AM-FM bidimensional, extremos locales y plano interpolante. Tal como planteáramos para funciones 1D, en este caso comparamos el mínimo local (en azul) con el plano que definen los tres máximos locales más cercanos (en rojo). De esta forma obtenemos una medida de la simetría de la función alrededor del mínimo local.

La situación descrita se ilustra en la Fig. 3.16.

Para favorecer soluciones suaves, debemos usar una matriz de suavizado (*roughening matrix* en inglés). Elegimos una aproximación discreta del operador Laplaciano  $L_{i,j} \in \mathbb{R}^{N_1 \times N_2}$  la cual consiste en una matriz del mismo tamaño que  $X$  con todos sus coeficientes nulos excepto por el *kernel*

$$K = \begin{bmatrix} 0,5 & 1 & 0,5 \\ 1 & -6 & 1 \\ 0,5 & 1 & 0,5 \end{bmatrix} \quad (3.57)$$

centrado en el punto  $i, j$ . Cuando  $i, j$  está en un borde de la matriz, simplemente cortamos el *kernel*, y modificamos el valor del elemento inicialmente en el centro para que la suma de todos sus elementos dé cero.

Gracias a la ec. (3.55) y al operador Laplaciano, podemos formular el siguiente problema de optimización:

$$\min_A \sum_{\ell/z_\ell \in \mathbb{L}_X} \langle P_\ell, (X - A) \rangle^2 + \lambda \sum_{i,j} \langle L_{i,j}, A \rangle^2, \quad (3.58)$$

en el cual el primer término modela la penalización de simetría en cada extremo local y el segundo término favorece soluciones suaves. Sin embargo, el problema de la ec. (3.58) puede ser de difícil resolución, con lo que se hace necesario modificarlo. El producto interno entre matrices puede calcularse como  $\langle A, B \rangle = \text{tr}(A^T B) = (\text{vec}(A))^T (\text{vec}(B))$ , donde  $\text{vec}(\cdot)$  es la clásica operación de vectorización, en la cual las columnas de una matriz son concatenadas verticalmente para formar un vector; y  $\text{tr}(\cdot)$  indica la traza de la matriz. Definamos los vectores  $\hat{a} \in \mathbb{R}^{N_1 N_2} = \text{vec}(A)$  y  $\hat{x} \in \mathbb{R}^{N_1 N_2} = \text{vec}(X)$ , la matriz  $\hat{P}_X \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$  que contiene a  $(\text{vec}(P_\ell))^T$  como

únicas filas no nulas, y  $\hat{L} \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$  que contiene a  $(\text{vec}(L_{i,j}))^T$  como filas. El siguiente problema de optimización

$$\min_{\hat{a}} \|\hat{P}_X(\hat{x} - \hat{a})\|_2^2 + \lambda \|\hat{L}\hat{a}\|_2^2, \quad (3.59)$$

es equivalente al de la ec. (3.58) y posee una única solución en

$$\hat{a}^* = (\hat{P}_X^T \hat{P}_X + \lambda \hat{L}^T \hat{L})^{-1} \hat{P}_X^T \hat{P}_X \hat{x}, \quad (3.60)$$

a partir de la cual se puede construir la media local  $A$  deseada.

El problema de la ec. (3.59) surge como una extensión natural de aquél de la ec. (3.9) para señales 1D, y marca el camino para extensiones a datos  $n$ D, simplemente definiendo el tensor de penalización de simetría para cada extremo local, y el tensor de suavizado en  $n$ D. Constituye también este enfoque un verdadero enfoque 2D *genuino*, en contraposición con los enfoques *pseudo* 2D que tratan separadamente las filas, columnas y diagonales. El resultado es una herramienta de procesamiento de imágenes independiente de la orientación.

### 3.7.2. Cuestiones de implementación

Los tamaños de las matrices involucradas en el problema de la ec. (3.59) pueden parecer abrumadores a primera vista, y por lo tanto la inversión requerida en la ec. (3.60) ser muy costosa en términos computacionales. Como en el caso unidimensional, la rareza debe ser tenida en cuenta para reducir el costo. Las matrices  $P_\ell$  tienen a lo sumo 4 elementos no nulos, y por lo tanto  $\hat{P}_X$  tiene una proporción de  $4\#\mathbb{L}_X/N_1^2 N_2^2$  elementos no nulos, con  $\#\mathbb{L}_X$  siendo la cardinalidad del conjunto  $\mathbb{L}_X$ , y  $\#\mathbb{L}_X \ll N_1 N_2$ . Lo mismo se aplica a  $L_{i,j}$ , la cual tiene a lo sumo 9 elementos distintos de cero, y en consecuencia  $\hat{L}$  tiene una proporción de  $9/N_1 N_2$  elementos no nulos. Existen rutinas eficientes para invertir matrices ralas lo cual hace a nuestra propuesta un método de bajo costo computacional cuando se la compara con otras implementaciones bidimensional de EMD. A partir de las ecuaciones (3.55)-(3.57) y (3.60), construimos un esquema deflacionario para descomponer una imagen  $X$ :

---

#### Algoritmo 3.2. UOA-EMD para datos 2D

---

- 1: dada  $X \in \mathbb{R}^{N_1 \times N_2}$  (imagen), asignar  $A_0 = X$  y  $\hat{a}_0 = \text{vec}(X) \in \mathbb{R}^{N_1 N_2}$ ;
- 2: definir la matriz de suavizado  $\hat{L} \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$ ;
- 3: **for**  $k = 1, 2, \dots, K$  **do**
- 4:   determinar el conjunto  $\mathbb{L}_k$  de extremos locales de  $A_k$ ;
- 5:   definir la matriz rala  $\hat{P}_{A_0} \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}$ ;
- 6:   **for**  $\ell = 1, 2, \dots, \#\mathbb{L}_k$  **do**
- 7:     definir la matriz rala  $P_\ell \in \mathbb{R}^{N_1 \times N_2}$  y sus coeficientes  $P(z_\ell) = 1$  y  $P(z_i) = c_i$  según la ec. (3.56);
- 8:      $\hat{P}_{A_0}(N_1(y_\ell - 1) + x_\ell, \cdot) \leftarrow \text{vec}(P_\ell)^T$ ;
- 9:   **end for**
- 10:   calcular la media local  $\hat{a}_k$  para  $\lambda_k > 0$

$$\hat{a}_k = (\hat{P}_{A_{k-1}}^T \hat{P}_{A_{k-1}} + \lambda_k \hat{L}^T \hat{L})^{-1} \hat{P}_{A_{k-1}}^T \hat{P}_{A_{k-1}} \hat{a}_{k-1}; \quad (3.61)$$

11: construir  $A_k$  a partir de  $\hat{a}_k$  y calcular el modo  $D_k = A_{k-1} - A_k$ ;  
 12: **end for**

*Notación:* usamos la notación  $M(n, \cdot)$  para indicar la  $n$ -ésima fila de la matriz  $M$ .

### 3.7.3. Delta de Dirac bidimensional

Los enfoques 2D genuinos deben ser independientes de cualquier orientación (isotrópicos). Para evaluar esta propiedad en nuestra propuesta UOA-EMD para dos dimensiones, presentamos en primer lugar la descomposición de cien copias ruidosas de una función delta de Dirac discreta bidimensional y promediamos los resultados, en un estilo similar a EEMD [39, 118]. Dada  $\Delta \in \mathbb{R}^{64 \times 64}$ , con

$$\Delta(i, j) = \begin{cases} 1, & \text{para } i = j = 32 \\ 0, & \text{en otro caso} \end{cases}, \quad (3.62)$$

descomponemos cien realizaciones de  $\Delta^{(i)} = \Delta + \beta W^{(i)}$ , con  $W^{(i)} \sim \mathcal{N}(0, 1)$  y  $\beta = 0,001$ , en dos modos y un residuo.

Los resultados se muestran en la Fig. 3.17 y confirman la independencia respecto a la orientación de esta propuesta UOA-EMD para datos 2D.

### 3.7.4. Imágenes AM-FM

Presentamos aquí cuatro experimentos diferentes en los cuales descomponemos imágenes AM-FM, midiendo tanto el desempeño en la recuperación de componentes como el tiempo de cómputo en función del único parámetro  $\lambda$ . Además, comparamos nuestros resultados con aquellos obtenidos por métodos del estado del arte: Image EMD (IEMD) [74], Prox-EMD [101] en su implementación *pseudo-2D*, y la descomposición TV-G texture-geometry [45].

IEMD usa un proceso de *sifting* para extraer cada modo, iterando hasta que  $|\bar{\epsilon}(x, y)| < \epsilon$ ,  $\forall(x, y)$ , donde  $\bar{\epsilon}(x, y)$  indica la semisuma de envolventes. La *toolbox* disponible en <http://aquador.vovve.net/IEMD/> requiere  $0 < \epsilon \leq 1$ . A medida que  $\epsilon$  se hace más pequeño, se realizan más iteraciones de *sifting*.

Prox-EMD logra una descomposición resolviendo un problema de optimización convexo no suave que posee términos relacionados con la *textura* y con la *geometría*. La *toolbox* está disponible en <http://perso.ens-lyon.fr/nelly.pustelnik/logicielE.html>.

La descomposición TV-G texture-geometry provee una separación en la cual el componente identificado con la *textura* está relacionado con las escalas más finas y por lo tanto puede ser considerado como *detalle* y el componente identificado con la *geometría* contiene escalas más gruesas y se relaciona con la *tendencia*. La *toolbox* está disponible en <http://www-rohan.sdsu.edu/jegilles>.

Tanto Prox-EMD como la descomposición TV-G texture-geometry contienen dos parámetros que deben ser establecidos:  $\nu$  para la *textura* y  $\rho$  para la *geometría*. Basándonos en [101], usamos  $\nu = \{1, 5, 10, 50\}$  y  $1 \leq \rho \leq 20$ .

Para los cuatro métodos (UOA-EMD, IEMD, Prox-EMD, y descomposición TV-G texture-geometry), descomponemos la imagen dada  $X = A + D$  en dos componentes:  $X \mapsto [\hat{A}, \hat{D}]$ . Para UOA-EMD e IEMD, tenemos que  $X = \hat{A} + \hat{D}$ ; para Prox-EMD y TV-G tenemos  $X \approx \hat{A} + \hat{D}$ . Por lo tanto, para medir la capacidad

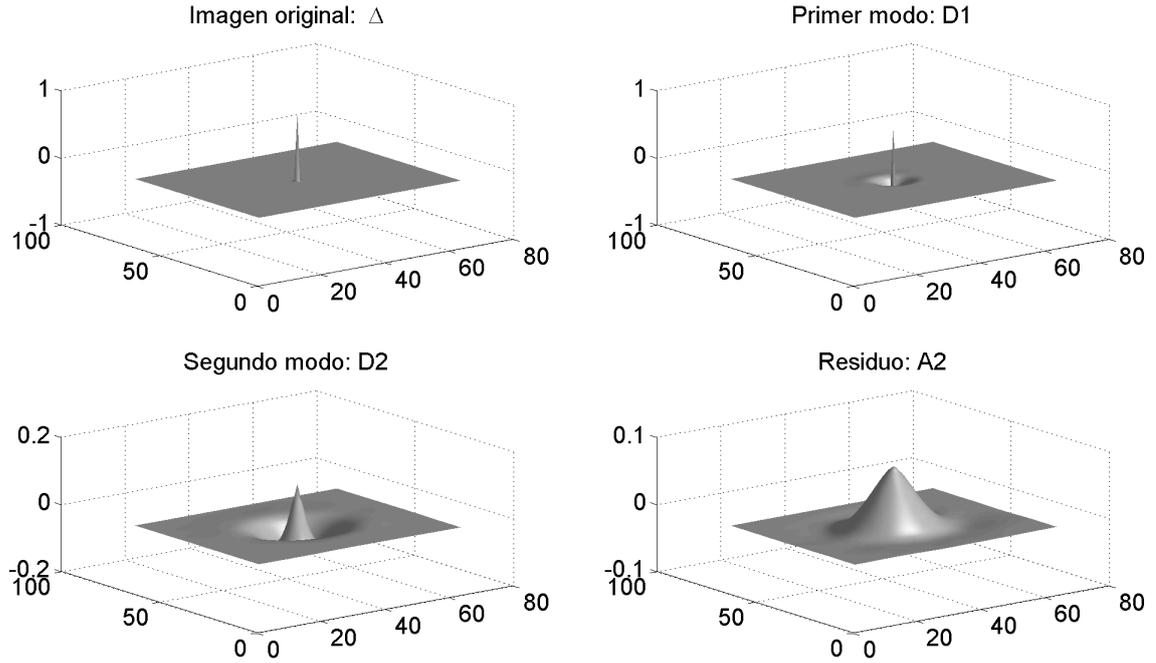


Figura 3.17: Descomposición de una Delta de Dirac bidimensional. Se muestran los resultados promedio de 100 descomposiciones de Deltas de Dirac ruidosas. Puede notarse la independencia respecto de la orientación del método aquí propuesto.

de recuperación de componentes de UOA-EMD e IEMD podemos usar tanto el error para  $A$  como el error para  $D$ , dado que  $\|A - \hat{A}\|_F = \|D - \hat{D}\|_F$ , donde  $\|\cdot\|_F$  indica la norma Frobenius. Para Prox-EMD y TV-G usamos  $(\|D - \hat{D}\|_F + \|A - \hat{A}\|_F)/2$  como medida del error cometido. En todos los casos, mostramos también los resultados de la descomposición para los parámetros que minimizan el error, con las imágenes escaladas para usar todo el rango de la escala de grises.

#### 3.7.4.1. Experimento 1

Este experimento consiste en descomponer  $X = X_1 + X_2$ , donde

$$\begin{aligned} X_1 &= \cos(6\pi(x+y)^2) + \cos(18\pi(x-y)) \\ X_2 &= \cos(1,5\pi(x+y)^2), \end{aligned} \quad (3.63)$$

con  $-1 \leq x \leq 1$  y  $-1 \leq y \leq 1$ . La imagen tiene un tamaño de  $128 \times 128$  píxeles. Los resultados de los cuatro métodos se muestran en la Fig. 3.18. UOA-EMD e IEMD logran resultados similares, evidenciando un buen desempeño a la hora de recuperar componentes. Al contrario, Prox-EMD y TV-G muestran un desempeño más bien pobre. La Fig. 3.19 presenta los errores al estimar los componentes y los tiempos de cómputo. UOA-EMD da el menor error para dos tercios de los rangos de parámetros considerados además de otorgar el menor error en términos absolutos, si elegimos el mejor resultado de cada método, y un tiempo de cómputo constante. IEMD evidencia tres regímenes distintos: cuando el valor de umbral  $\varepsilon$  va de 0 a 1, se usan tres cantidades distintas de iteraciones de *sifting*. Este primer experimento muestra que Prox-EMD y TV-G son incapaces de recuperar los componentes AM-

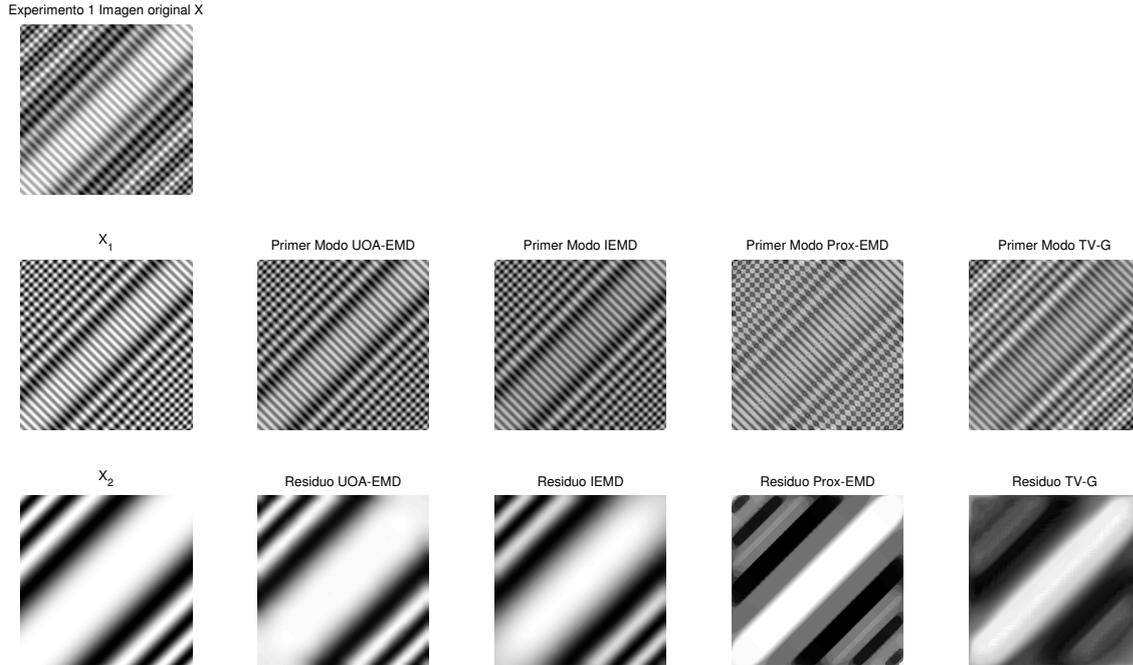


Figura 3.18: Resultados para el Experimento 1. Se muestran los mejores resultados para cada uno de los cuatro métodos analizados cuando descomponen la imagen cuyos componentes se describen en la ec. (3.63).

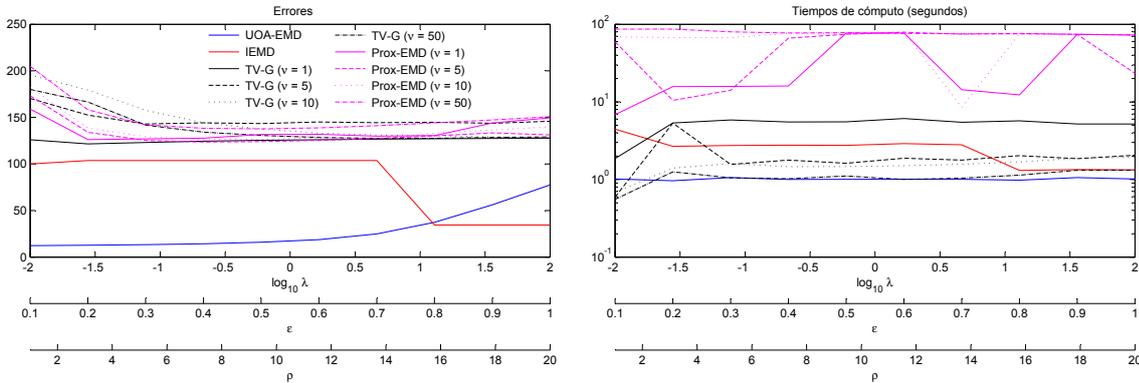


Figura 3.19: Errores y tiempos de cómputo para el Experimento 1 (ver el texto para los detalles).

FM de la imagen procesada. Más aún, la selección de los valores de sus parámetros puede ser un problema bastante complicado.

### 3.7.4.2. Experimento 2

Para este experimento consideramos la descomposición de  $X = X_1 + X_2$ , con

$$\begin{aligned}
 X_1 &= \begin{cases} \cos(6\pi(x+y)^2) + \cos(18\pi(x-y)), & \text{si } 0 \leq y \leq 1 \\ \cos(2\pi(x+y)^2), & \text{si } -1 \leq y < 0 \end{cases} \\
 X_2 &= \begin{cases} \cos(2\pi(x+y)^2), & \text{si } 0 \leq y \leq 1 \\ \cos(0,5\pi(x+y)^2), & \text{si } -1 \leq y < 0 \end{cases}
 \end{aligned} \tag{3.64}$$

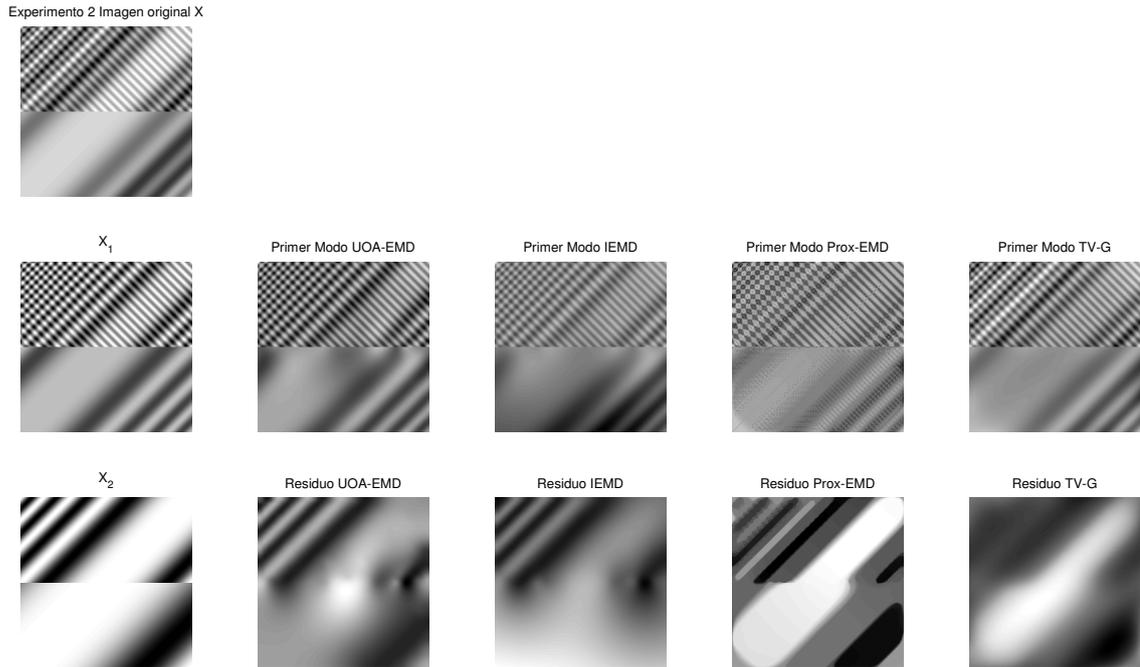


Figura 3.20: Resultados para el Experimento 2. Se muestran los mejores resultados para cada uno de los cuatro métodos analizados cuando descomponen la imagen cuyos componentes se describen en la ec. (3.64).

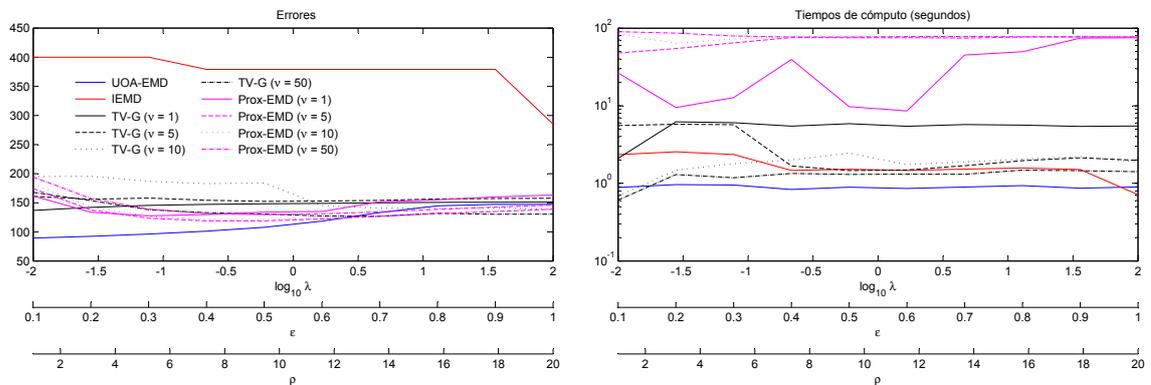


Figura 3.21: Errores y tiempos de cómputo para el Experimento 2 (ver el texto para los detalles).

La mitad inferior de  $X_1$  es igual a la mitad superior de  $X_2$ . Dado que  $X_2$  es *tendencia* para  $X_1$ , las mismas frecuencias espaciales son *tendencia* en una parte de la imagen y *detalle* en la otra. Un simple filtrado lineal no sería capaz de separar los componentes. Los resultados obtenidos con los cuatro métodos se muestran en la Fig. 3.20. UOA-EMD presenta los mejores resultados para la mitad inferior del residuo. Los errores para Prox-EMD son cercanos a los de UOA-EMD (Fig. 3.21). Sin embargo, cuando analizamos el residuo, Prox-EMD parece favorecer soluciones constantes a trozos, las cuales no son adecuadas para representar las características AM-FM. Los mejores desempeños en tiempos de cómputo son los obtenidos con UOA-EMD (ver Fig. 3.21).

Experimento 3 Imagen original X

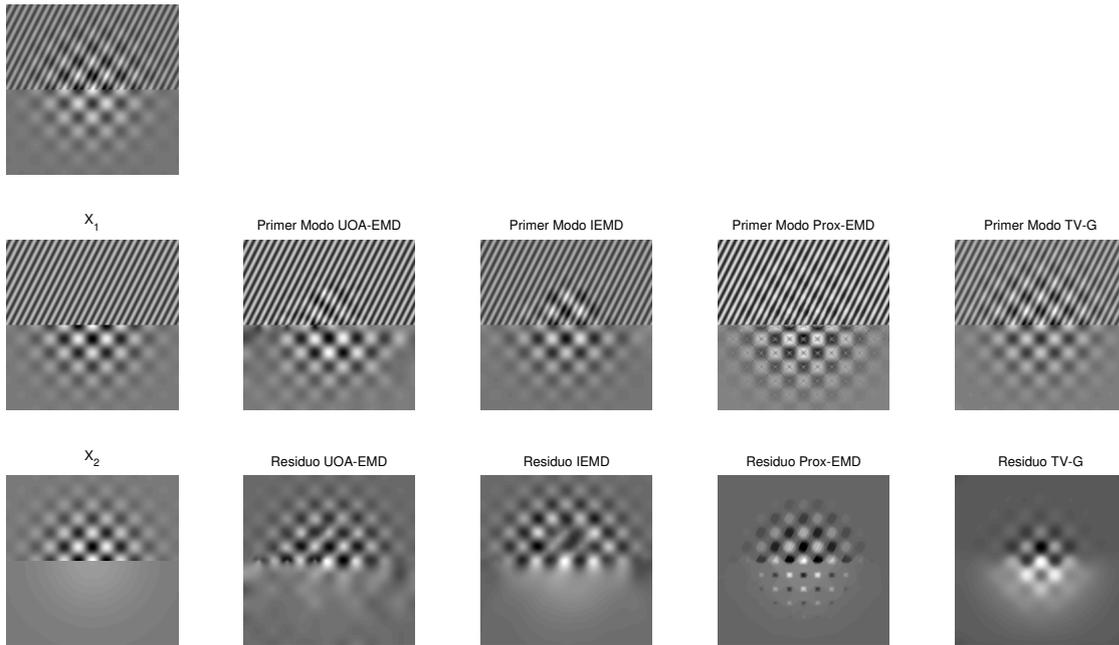


Figura 3.22: Resultados para el Experimento 3. Se muestran los mejores resultados para cada uno de los cuatro método analizados cuando descomponen la imagen cuyos componentes se describen en la ec. (3.65).

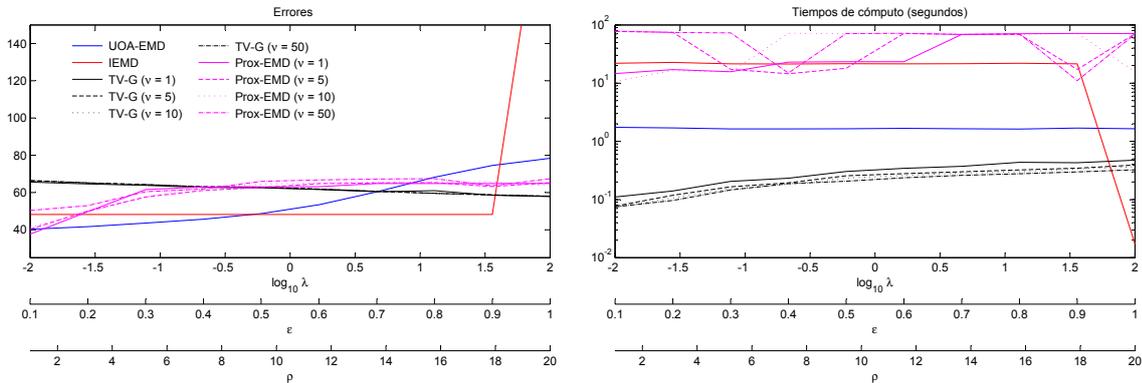


Figura 3.23: Errores y tiempos de cómputo para el Experimento 3 (ver el texto para los detalles).

### 3.7.4.3. Experimento 3

En este experimento introducimos modulaciones en amplitud:

$$\begin{aligned}
 X_1 &= \begin{cases} \cos(24\pi(x + 0,1y)), & \text{si } 0 \leq y \leq 1 \\ e^{-3(x^2+y^2)} (\cos(6\pi(x+y)) + \cos(6\pi(x-y))), & \text{si } -1 \leq y < 0 \end{cases} \\
 X_2 &= \begin{cases} e^{-3(x^2+y^2)} (\cos(6\pi(x+y)) + \cos(6\pi(x-y))), & \text{si } 0 \leq y \leq 1 \\ e^{-3(x^2+y^2)}, & \text{si } -1 \leq y < 0 \end{cases}
 \end{aligned} \tag{3.65}$$

Como en el ejemplo anterior, la mitad inferior de  $X_1$  es igual a la mitad superior de  $X_2$ . Las descomposiciones otorgadas por los cuatro métodos se presentan en la Fig. 3.22. Los errores y tiempos de cómputo pueden apreciarse en la Fig. 3.23. Para este tercer experimento, los resultados de UOA-EMD e IEMD son similares. Los errores son del mismo orden de magnitud, excepto para  $\varepsilon = 1$  para IEMD donde el error es significativamente mayor. Aunque los errores para Prox-EMD son bajos para valores pequeños del parámetro de geometría  $\rho$ , los resultados merecen una atención especial. El residuo de Prox-EMD presenta cuadrados aislados de un valor casi constante, los cuales no están presentes en el componente original. El primer modo de Prox-EMD presenta en su mitad inferior un patrón de líneas verticales, horizontales y diagonales que no se encuentran en  $X_1$ . Este efecto es una clara consecuencia de un enfoque *pseudo-2D*, en el cual filas, columnas y diagonales son tratadas por separado. La mitad superior del primer modo de TV-G es muy similar a la de la imagen original  $X$ . Más aún, los tiempos de cómputo permanecen casi constantes para UOA-EMD, siendo trece veces menores a los de IEMD.

#### 3.7.4.4. Experimento 4

Como cuarto y último experimento de esta subsección, enmascaramos las imágenes del Experimento 3 con

$$M = \begin{cases} 1, & \text{si } x^2 + y^2 < \frac{16}{25} \\ 0, & \text{en otro caso} \end{cases}, \quad (3.66)$$

con el objetivo de evaluar el comportamiento de los métodos en una imagen con una zona constante.

Los resultados pueden apreciarse en la Fig. 3.24. Aparecen las mismas características que en el ejemplo anterior. Los errores (ver Fig. 3.25) para IEMD aumentan, mientras que el desempeño de UOA-EMD se mantiene prácticamente inalterado.

#### 3.7.5. Imagen real

Con frecuencia los investigadores en procesamiento de imágenes utilizan imágenes reales estándares para evaluar los desempeños de los distintos métodos. Presentamos aquí la descomposición de una versión de  $512 \times 512$  píxeles de la imagen “Barbara” *via* los cuatro métodos. En cada caso, la imagen se descompuso en dos modos más un residuo. Usamos  $\lambda_1 = \lambda_2 = 1$  para UOA-EMD,  $\varepsilon = 0,5$  y  $K = 2$  para IEMD, y  $\nu_1 = \nu_2 = 1$ ,  $\rho_1 = \rho_2 = 10$  y  $K = 2$  tanto para Prox-EMD como para TV-G. La descomposición necesitó alrededor de 5 minutos para UOA-EMD, cinco horas y media para IEMD, 17 minutos para Prox-EMD y 50 segundos para TV-G. Los resultados de las descomposiciones con los cuatro métodos se muestran en la Fig. 3.26.

El primer modo de UOA-EMD captura las frecuencias más altas. Las franjas de la ropa de la mujer y del mantel se conservan perfectamente, junto con los bordes del resto de la imagen. El segundo modo de UOA-EMD captura frecuencias más bajas pero conserva características importantes como los pliegues del pantalón de la mujer, de la tela alrededor de su cabeza y del mantel. Se llama la atención sobre

Experimento 4 Imagen original X

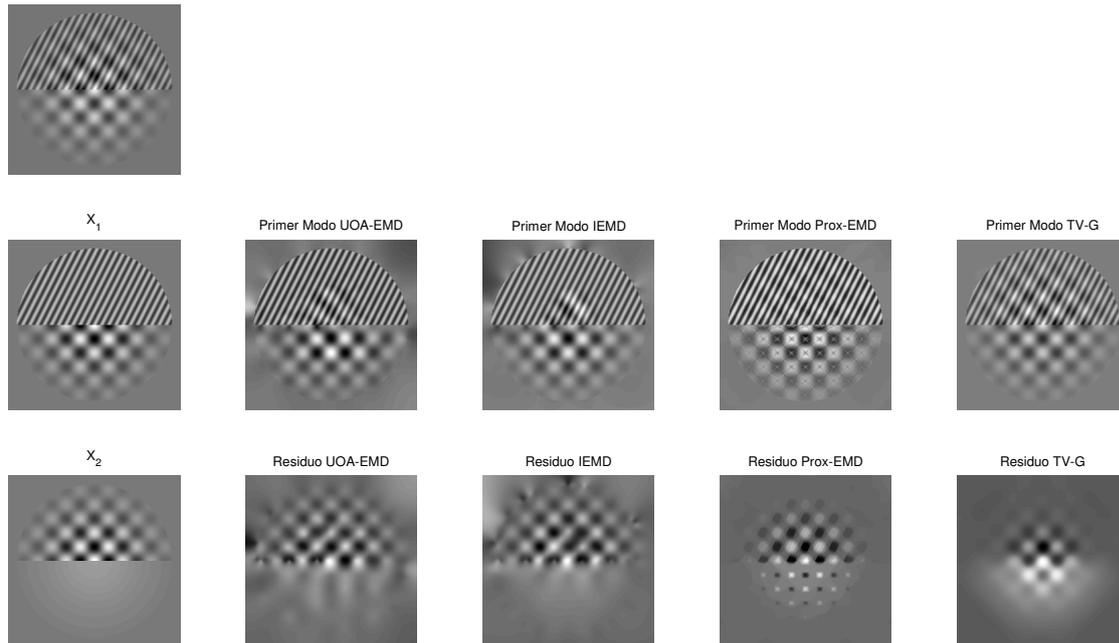


Figura 3.24: Resultados para el Experimento 4. Se muestran los mejores resultados para cada uno de los cuatro métodos analizados cuando descomponen la imagen cuyos componentes se describen en la ec. (3.66).

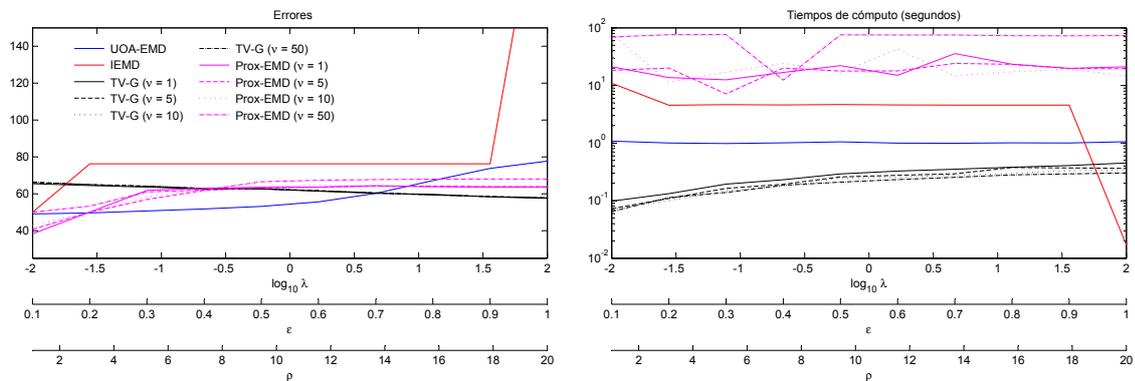


Figura 3.25: Errores y tiempos de cómputo para el Experimento 4 (ver el texto para los detalles).

una porción de la imagen original. La misma zona es aumentada en los dos primeros modos. En la zona aumentada pueden apreciarse las franjas de la tela, sin rastros de las áreas más oscuras. Las franjas de la tela están ausentes en esta zona del segundo modo, y sólo pueden verse los pliegues. Todo el contraste de la imagen está presente en el residuo.

Por el contrario, los otros métodos aquí analizados parecen ser incapaces de capturar dichas características. El primer modo de IEMD incluye no sólo los bordes y las franjas de la tela sino también oscilaciones más gruesas como los pliegues de las telas. El segundo modo está demasiado borroso como para identificar estructuras finas. También contiene la mayor parte del contraste de la imagen, dejando al residuo casi sin información relevante. Prox-EMD devuelve un primer modo donde los bordes

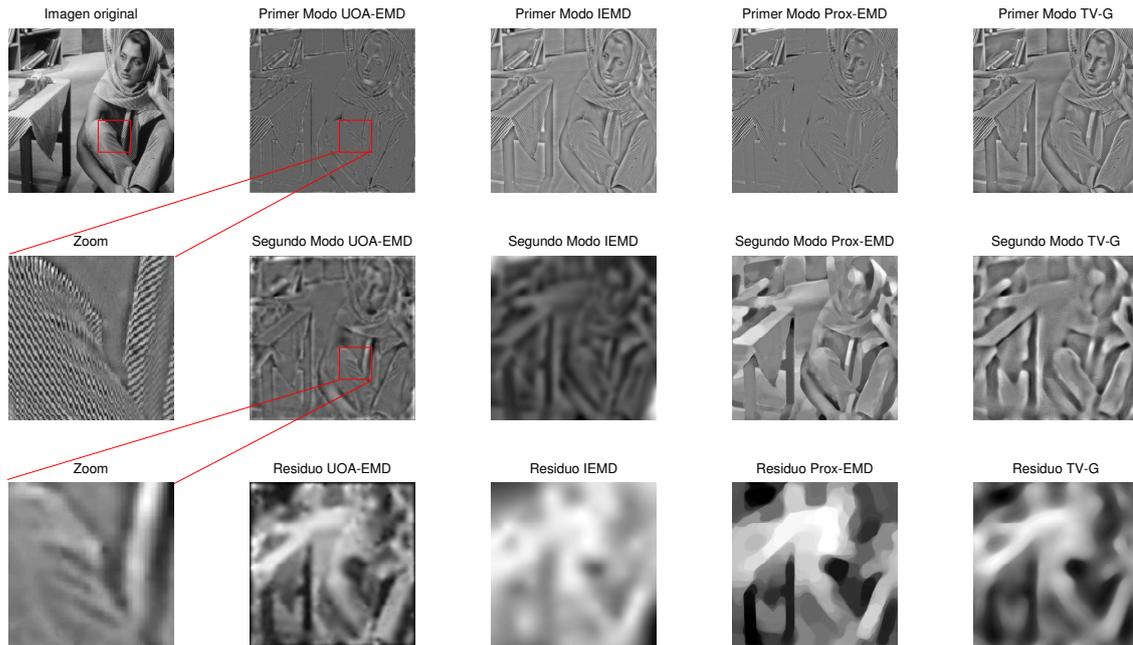


Figura 3.26: Descomposiciones de la imagen «Barbara». Se muestran los resultados obtenidos mediante los cuatro métodos.

son difíciles de detectar, como los de las patas de la mesa. El segundo modo y el residuo son difíciles de analizar. Mientras el segundo modo se asemeja a una versión en acuarela de la imagen sin detalles finos, el residuo está demasiado difuminado y no se identifican fácilmente estructuras (debemos recordar aquí que Prox-EMD no provee una reconstrucción exacta de la imagen a partir de los modos). Finalmente, el primer modo de TV-G es el más similar, excepto por el brillo, a la imagen entre los primeros modos extraídos mediante los métodos estudiados. Las patas de la mesa son significativamente más oscuras que el piso, detectando más que simples bordes; lo mismo puede decirse sobre el torso de la mujer. El segundo modo y el residuo dan poca información sobre la imagen original (como con Prox-EMD, la reconstrucción de TV-G tampoco es exacta).

### 3.7.6. Aplicación biomédica

Presentamos ahora la descomposición de una imagen médica de  $256 \times 256$  píxeles. Fue extraída del *website* de la Harvard Medical School for Radiology en <http://www.med.harvard.edu/AANLIB/cases/case1/mr1-tl4/029.html> [120]. Esta imagen, una resonancia magnética T2-pesada, se muestra en la Fig. 3.27. Fue registrada en una mujer que padece de glioma. La descomposición de la imagen en tres componentes obtenidos mediante UOA-EMD se muestra en la Fig. 3.27, donde se usó  $\lambda_k = 1, 1 \leq k \leq 3$ . Observamos que los tres componentes representan distintos rangos de escala: el primer componente corresponde a las frecuencias más altas (escalas más finas); a medida que el índice de la descomposición crece, se descubren escalas más gruesas. La tendencia puede observarse en el residuo. El primer modo revela claramente el contorno de la lesión. Este resultado ilustra la potencialidad de UOA-EMD en el campo médico.



Figura 3.27: Descomposición de una imagen de resonancia magnética mediante UOA-EMD. Los resultados se presentan desde las escalas finas a las gruesas.

### 3.8. Comentarios finales

Hemos presentado en este capítulo la primera contribución original de este trabajo de tesis: definir a EMD como un problema de optimización sin restricciones. Este enfoque se enmarca en una corriente novedosa que intenta dotar a EMD de bases teóricas tomando la teoría de optimización como punto de partida. Nuestra propuesta permite analizar *a priori* el comportamiento de la descomposición. La solución que construye puede compararse con aquella obtenida mediante el método de *B-spline* EMD, pero con las funciones «básicas» y los coeficientes completamente determinados por los datos. Una interpretación de filtro variante en el tiempo, revela que nuestro método construye una media local de manera discreta y local, cuando el caso ideal es continuo e instantáneo.

Realizamos extensas simulaciones con señales artificiales, comparando nuestros resultados con aquellos obtenidos con métodos del estado del arte de EMD basada en optimización. En todos los casos, se evidencian ventajas notorias para nuestro método, tanto en desempeño como en tiempo de cómputo. Su aplicación en una señal real de voz ilustra la potencialidad de nuestra propuesta para el análisis de este tipo de señales.

El enfoque presentado resulta tan sencillo en comparación con los del estado del arte, que su extensión a dos dimensiones no presenta mayores obstáculos. Tal es así que extendimos nuestro método al campo de las imágenes. En este caso, también lo comparamos con métodos del estado del arte, mostrando ventajas de nuestra propuesta.

La técnica guiada por los datos de EMD ha conseguido atraer la atención de una parte de la comunidad científica de procesamiento de señales. Desde sus primeros años, existió una corriente abocada a explicar de manera teórica el comportamiento de esta técnica que fue definida originalmente como la salida de un algoritmo. En los últimos años, los enfoques basados en optimización se han mostrado promisorios para lograr una descripción de la técnica. Nuestra contribución es parte de esa tendencia.

Los resultados desarrollados en este capítulo dieron origen a dos publicaciones en *journals* internacionales. Una publicación en *Digital Signal Processing* de Elsevier [5] y otra en *IEEE Transactions on Image Processing* [1].

# Capítulo 4

## Un enfoque tiempo-escala para EMD

*No está fuera de lo posible la repetición sobre una escala más extensa del mismo fenómeno.*

---

MI VIDA  
León Trotsky

### Contenidos

---

<b>4.1. Introducción</b>	<b>87</b>
<b>4.2. EMD y el problema de la media local</b>	<b>87</b>
<b>4.3. Transformada ondita continua y <i>synchrosqueezing</i></b>	<b>89</b>
<b>4.4. Propuesta de algoritmo</b>	<b>91</b>
4.4.1. Una nueva definición para los modos	91
4.4.2. La escala local	91
4.4.3. Problemas de discontinuidad y <i>synchrosqueezing</i>	93
4.4.4. La constante C	93
4.4.5. Refinamiento de la escala local	93
4.4.6. Tratamiento de bordes	93
<b>4.5. Experimentos y resultados</b>	<b>94</b>
4.5.1. Suma de componentes AM-FM solapados	94
4.5.2. Formas de onda no lineales	94
<b>4.6. Comentarios finales</b>	<b>97</b>

---



## 4.1. Introducción

Ya habíamos mencionado en el Cap. 2 que las representaciones tiempo-frecuencia y tiempo-escala son probablemente la manera más natural de abordar el análisis de una señal multicomponente. La forma habitual de proceder suele ser buscar en las representaciones las distintas «crestas» que los modos generan en ellas, y luego antitransformar «franjitas» que rodean a tales crestas de forma de obtener los modos en el dominio temporal. Siguiendo este razonamiento, se tiene que una cresta implica un modo.

La descomposición empírica en modos (EMD) considera oscilaciones a nivel local, descomponiendo una señal en un número pequeño de funciones intrínsecas de modo (IMFs), las cuales son básicamente señales que oscilan simétricamente alrededor de cero, que pueden estar moduladas tanto en amplitud como en frecuencia. En algunos casos, la descomposición EMD concuerda mejor con la física y/o percepción cuando dos tonos cercanos en frecuencia son considerados como una sola onda, como en el efecto de *beating* [97]. Entonces, la EMD no sigue la idea de que debe haber un modo por cada cresta. En esta técnica, los componentes pueden ser algo más complejo que funciones circulares, ganando por lo tanto en versatilidad.

Sin embargo, y como mencionamos en el Cap. 3, la EMD se define como la salida de un algoritmo, sin una definición analítica. Debido a ello, y para dotarla con una base matemática más sólida, es que aquí propondremos fusionar la naturaleza guiada por los datos de EMD con las sólidas bases matemáticas de las representaciones tiempo-escala. Esta relación que proponemos no es arbitraria sino que responde a algunas similitudes entre EMD y el análisis *wavelet* ya estudiadas en la literatura [38]. Basta recordar el esquema deflacionario de la ondita diádica y otros aspectos ya discutidos en la Sec. 2.4.1. El comportamiento de EMD cuando se la aplica a ruido blanco no hace más que reforzar esta idea [39, 40].

En este capítulo presentamos en la Sec. 4.2 un breve repaso de los fundamentos de EMD y la definición de una media local. En la Sec. 4.3 recordamos los conceptos de transformada ondita y presentamos las ideas básicas detrás de *synchrosqueezing*. En la Sec. 4.4 formalizamos nuestra propuesta para la descomposición de señales multicomponentes, proponiendo una nueva definición para la extracción de los modos, la escala local y demás cuestiones inherentes al método. En la Sec. 4.5 presentamos simulaciones con señales artificiales.

## 4.2. EMD y el problema de la media local

La EMD es un método adaptativo (guiado por los datos) diseñado para analizar señales no estacionarias que pueden provenir de sistemas no lineales. Separa a las señales, de manera local y completamente guiada por los datos, en oscilaciones rápidas y lentas. Trabajando en un esquema deflacionario, en un primer paso la señal  $x(t)$  se descompone en

$$x(t) = d_1(t) + a_1(t), \quad (4.1)$$

con  $d_1(t)$  siendo una IMF, la parte de la señal que oscila *localmente* más rápido, y  $a_1(t)$  siendo la primera *media local*. El método continúa descomponiendo la media local en  $a_1(t) = d_2(t) + a_2(t)$ , y así sucesivamente. Al final de la descomposición, la

señal original puede ser expresada como una suma de IMFs ( $d_k(t)$ ,  $k = 1 \dots K$ ) más una tendencia final ( $a_K(t)$ ):

$$x(t) = \sum_{k=1}^K d_k(t) + a_K(t). \quad (4.2)$$

El algoritmo de EMD puede verse en el Algoritmo 2.1. Lo resumimos aquí a modo de recordatorio [58, 99]:

1. identificar los extremos locales de  $x(t)$ ,
2. interpolar entre los mínimos (resp. máximos) para obtener la envolvente inferior (resp. superior)  $e_{min}(t)$  (resp.  $e_{max}(t)$ ),
3. calcular la media local  $a(t) = (e_{min}(t) + e_{max}(t))/2$ ,
4. extraer el detalle o modo  $d(t) = x(t) - a(t)$ ,
5. iterar en el residuo o media local  $a(t)$ .

Todo el problema se reduce a la separación de señal en modo más media local, como en la ec. (4.1) (o como en el paso 4), y luego repetir el mismo procedimiento sobre la media local (paso 5). Dadas las dificultades para obtener una definición satisfactoria de la media local, los creadores de la EMD la definieron como la media de las envolventes superior e inferior obtenidas mediante interpolaciones *spline* de los extremos locales correspondientes. Esta definición geométrica, aunque útil en la práctica, carece de la precisión suficiente para realizar un estudio analítico de la EMD. Hemos visto en la ec. (2.43) una propuesta analítica para la media local dada por Rilling [95], que recordamos aquí, donde la media local  $a(t)$  de una señal  $x(t)$  se obtendría de la siguiente manera:

$$a(t) = \frac{1}{\alpha(t)} \int x(u)w\left(\frac{u-t}{\alpha(t)}\right) du, \quad (4.3)$$

donde  $w(u)$  es una función de ponderación de masa unitaria concentrada alrededor de cero y  $\alpha(t)$  es la *escala local*. El autor consideró que la elección de  $w(u)$  no es necesariamente difícil pero la elección de la escala local  $\alpha(t)$  es un problema delicado. Debemos notar, sin embargo, la similitud entre la ec. (4.3) y un análisis de función de escalamiento (*scaling function* en inglés) evaluada en la escala local.

**Definición 4.2.1. Función de escalamiento (scaling function) [76].** Cuando la transformada ondita  $W_x^\psi(t, s)$  se conoce sólo para escalas  $s < s_0$ , para recuperar la señal  $x(t)$  se hace necesaria una información complementaria que corresponde a  $W_x^\psi(t, s)$  para  $s > s_0$ . Dicha información se obtiene usando una función de escalamiento  $\varphi$ , la cual es una agregación de onditas a escalas mayores que 1. El módulo de su transformada de Fourier se define como

$$|\hat{\varphi}(f)|^2 = \int_1^{+\infty} |\hat{\psi}(sf)|^2 \frac{ds}{s}, \quad (4.4)$$

siendo  $\hat{\psi}(f)$  la transformada de Fourier de la ondita. La función de escalamiento puede ser interpretada como la respuesta al impulso de un filtro pasa-bajos. Entonces, la aproximación de baja frecuencia de  $x$  a la escala  $s$  es

$$L_x^\varphi(t, s) = \frac{1}{s} \int x(u)\varphi^*\left(\frac{u-t}{s}\right) du. \quad (4.5)$$

Vemos que en la ec. (4.5) la función de escalamiento  $\varphi(u)$  ocupa el lugar de la función de ponderación de la ec. (4.3). Entonces, podemos ver a la media local como

$$a(t) = L_x^\varphi(t, \alpha(t)). \quad (4.6)$$

### 4.3. Transformada ondita continua y *synchrosqueezing*

Recordemos la transformada ondita continua (Def. 2.3.2):

$$W_x^\psi(t, s) = \frac{1}{s} \int x(u) \psi^* \left( \frac{u-t}{s} \right) du, \quad s > 0, \quad (4.7)$$

donde  $\psi(u)$  es una ondita *analítica* y  $z^*$  indica el complejo conjugado de  $z$ . La fórmula de reconstrucción vertical es

$$x(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^\infty W_x^\psi(t, s) \frac{ds}{s} \right), \quad (4.8)$$

con  $\widetilde{C}_\psi = \int \widehat{\psi}^*(f) \frac{df}{f}$ , donde  $\widehat{\psi}(f) = \int \psi(u) e^{-2\pi i f u} du$  es la transformada de Fourier de  $\psi(u)$ .

La técnica de *synchrosqueezing* [31, 32], que es un caso especial de *reassignment* [12], tiene como objetivo «afilar» la representación, mejorando su legibilidad, mientras se mantiene invertible.

**Definición 4.3.1. Synchrosqueezing** [31]. Dada  $W_x^\psi(t, s)$ , la transformada ondita continua de  $x(t)$ , se define para  $A_W = \{(t, s) / W_x^\psi(t, s) \neq 0\}$

$$\tilde{f}(t, s) = \frac{\partial_t W_x^\psi(t, s)}{2\pi i W_x^\psi(t, s)}. \quad (4.9)$$

La versión de  $W_x^\psi$  con *synchrosqueezing*, o *reassignada verticalmente*, es

$$SW_x^\psi(t, f) = \int_{A_W} W_x^\psi(t, v) \delta(f - \tilde{f}(t, s)) \frac{dv}{v}. \quad (4.10)$$

La señal  $x(t)$  puede recuperarse mediante

$$x(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^\infty SW_x^\psi(t, f) df \right). \quad (4.11)$$

Entonces vemos que *synchrosqueezing* no sólo provee una mejora en la representación en cuanto a su legibilidad, sino que es usado al mismo tiempo para hacer una conversión de escala a frecuencia e interpretar la representación en una forma más clásica (ya que en ingeniería se suele preferir hablar de frecuencias en vez de escalas). Sin embargo, y como hemos visto más arriba, la idea de escalas es preferible en EMD a la idea de frecuencia. Tal es así que aquí, en lugar de la más popular transformación de escala a frecuencia, usaremos una transformación escala a escala, reemplazando la ec. (4.9) por

$$\tilde{s}(t, s) = \frac{2\pi i W_x^\psi(t, s)}{\partial_t W_x^\psi(t, s)}, \quad (4.12)$$

definida sólo para los puntos en  $B_W = \{(t, s)/\partial_t W_x^\psi(t, s) \neq 0\}$ . Entonces, la CWT reasignada verticalmente de la ec. (4.10) se transforma en

$$SW_x(t, s) = \int_{B_W} W_x^\psi(t, v) \delta(s - \tilde{s}(t, s)) \frac{dv}{v}, \quad (4.13)$$

mientras que la fórmula de reconstrucción resulta

$$x(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^\infty SW_x(t, s) ds \right). \quad (4.14)$$

Para las implementaciones prácticas, seguimos a [78]. Se define una partición del eje de escalas  $\{s_k\}_{k=0}^\infty$  y luego los intervalos  $S_k = [\frac{s_k+s_{k-1}}{2}, \frac{s_k+s_{k+1}}{2}]$ . La CWT con *synchrosqueezing* se aproxima mediante  $\widetilde{SW}_x^\psi(t, s_k) = \int_{s:|\tilde{s}(t,s)| \in S_k} W_x^\psi(t, v) \frac{dv}{v}$  y la reconstrucción es  $x(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \sum_k \widetilde{SW}_x^\psi(t, s_k) \right)$ .

El marco de trabajo de *synchrosqueezing* define una clase especial de componentes (IMTs, ver Def. 2.4.5) y cómo se combinan entre sí para formar las señales [31].

**Definición 4.3.2. Superposición de IMTs bien separadas.** Una función  $x(t) \in \mathbb{R}$  está compuesta de funciones tipo intrínsecas de modo bien separadas, de precisión hasta  $\epsilon$ , y con separación  $D$ , si existe un número natural finito  $K$  tal que  $x(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t))$ , donde todos los componentes son IMT y sus funciones de fase  $\phi_k$  satisfacen para todo  $t$ :  $\phi'_k(t) > \phi'_{k-1}$  y  $|\phi'_k(t) - \phi'_{k-1}(t)| \geq D[\phi'_k(t) + \phi'_{k-1}(t)]$ .

Además definiremos aquí la *escala instantánea ideal* simplemente como el recíproco de la frecuencia instantánea (ver Def. 2.2.1).

**Definición 4.3.3. Escala instantánea ideal.** Sea  $d(t) = A(t) \cos(2\pi\phi(t))$ . Se define su escala instantánea ideal como  $1/\phi'(t)$ .

Un importante teorema de [31] asegura que la CWT de una IMT posee la mayoría de sus coeficientes no nulos concentrados en una «franja» alrededor de su escala instantánea.

**Teorema 4.3.1.** Si  $d(t) = A(t) \cos(2\pi\phi(t))$  es una IMT con precisión  $\epsilon$ , y  $\psi(t)$  es una ondita analítica con soporte compacto en la frecuencia  $\text{supp } \hat{\psi} \subseteq [f_\psi - \Delta, f_\psi + \Delta]$ , entonces  $|W_d^\psi(t, s)| > \epsilon^{1/3}$  sólo cuando  $(t, s) \in Z = \left\{ (t, s) / |s - \frac{f_\psi}{\phi'(t)}| < \frac{\Delta}{\phi'(t)} \right\}$ .

(La franja está en realidad alrededor de  $f_\psi/\phi'(t)$ , donde  $f_\psi$  es la ubicación del máximo del módulo del espectro de Fourier de la ondita.)

Finalmente, la extracción del modo se reduce a encontrar la línea de cresta resolviendo <sup>1</sup>

$$\max_{c(t)} \int E_x(t, s) dt - \lambda \int [c'(t)]^2 dt, \quad (4.15)$$

<sup>1</sup>En la práctica, la implementación real se hace mediante una cresta discretizada  $c(n)$ . Usamos la *Synchrosqueezing Toolbox* [17] que implementa un enfoque heurístico (voraz) que maximiza la función objetivo para cada índice temporal, suponiendo que ya ha sido maximizada para todos los índices previos.

donde  $E_x(t, s) = \log(|SW_x(t, s)|^2)$ , y  $\lambda > 0$  es un parámetro de regularización. Entonces, el modo es estimado como

$$d(t) = \int_{|s-c(t)| < \frac{\Delta}{f_\psi} c(t)} SW_x(t, s) ds. \quad (4.16)$$

A diferencia de otros enfoques basados en *synchrosqueezing*, para los cuales la transformación de escala a frecuencia conlleva que la franja en el plano tiempo-escala que ocupa un modo sea de ancho constante, como por ejemplo en la ec. (2.36) donde seguíamos a [13], definimos aquí un «ancho de franja» variable debido a nuestra transformación escala a escala (ecuaciones (4.12) y (4.13)). Para la extracción de más de un modo, se aplica usualmente una estrategia deflacionaria: una vez extraído el modo, se hace cero su franja en la representación tiempo-escala, y se procede a la siguiente extracción de cresta [16]. Con el criterio de la ec. (4.15), los modos no aparecen necesariamente ordenados por escala o frecuencia, sino por una combinación de mayor energía total y menor ancho de banda marginal. Claramente, estos no son los criterios jerárquicos por los cuales un método tipo EMD ordena los modos encontrados.

## 4.4. Propuesta de algoritmo

### 4.4.1. Una nueva definición para los modos

Propondremos aquí definir un modo, i.e. la parte que oscila localmente más rápido, de una señal dada tomando la ec. (4.6) como inspiración pero enfocándonos más en el modo que en la media local. El razonamiento es el siguiente: si la media local es el resultado de un análisis con función de escalamiento evaluado a una escala local específica, entonces el modo resultará igual a la integración de un análisis *wavelet* desde la escala más fina posible hasta la escala local mencionada. Combinando las ecuaciones (4.1) y (4.6) con la fórmula de reconstrucción vertical (4.8), obtenemos

$$d(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^{C \alpha(t)} W_x^\psi(t, s) \frac{ds}{s} \right), \quad (4.17)$$

donde  $\alpha(t)$  es la escala local, y  $C$  una constante real que explicaremos en detalle más adelante. Esta idea implementa de manera explícita la extracción del modo como un filtrado pasa-altos dependiente del tiempo, lo cual ha sido ampliamente discutido en la literatura de EMD [38, 99].

### 4.4.2. La escala local

Para la definición de la escala local, usamos la siguiente estimación:

**Estimación 4.4.1.** *Sea  $d(t)$  una IMT (Def. 2.4.5), y  $t_0$  y  $t_1$  dos extremos locales consecutivos de  $d(t)$ . Entonces, si  $\epsilon$  es suficientemente pequeño, se cumple que*

$$\frac{1}{2\phi'(t_0)} + \frac{1}{2\phi'(t_1)} \approx 2(t_1 - t_0). \quad (4.18)$$

*Demostración.* La estimación se verifica fácilmente para el caso de un tono puro. Para el caso más general, empezamos tomando una aproximación de Taylor de primer orden de  $\phi(t)$  alrededor tanto de  $t_0$  como de  $t_1$  (dado que  $\epsilon$  es suficientemente pequeño podemos ignorar los términos cuadráticos en adelante):

$$\begin{aligned}\phi(t) &\approx \phi(t_0) + \phi'(t_0)(t - t_0) \\ \phi(t) &\approx \phi(t_1) + \phi'(t_1)(t - t_1).\end{aligned}\tag{4.19}$$

Evaluable cada aproximación en el otro extremo, y sumando miembro a miembro, obtenemos

$$(\phi(t_1) - \phi(t_0)) \left( \frac{1}{\phi'(t_0)} + \frac{1}{\phi'(t_1)} \right) \approx 2(t_1 - t_0).\tag{4.20}$$

Luego, aplicamos la condición  $d'(t^*) = 0$ , que se cumple para  $t^* = t_0$  y  $t^* = t_1$  por ser extremos locales:

$$\begin{aligned}d'(t^*) &= A'(t^*) \cos(2\pi\phi(t^*)) - A(t^*)2\pi\phi'(t^*) \sin(2\pi\phi(t^*)) = 0 \Rightarrow \\ \Rightarrow \tan(2\pi\phi(t^*)) &= \frac{A'(t^*)}{A(t^*)2\pi\phi'(t^*)} \\ \phi(t^*) &= \frac{1}{2\pi} \arctan \left( \frac{A'(t^*)}{A(t^*)2\pi\phi'(t^*)} \right).\end{aligned}\tag{4.21}$$

Esto nos lleva a

$$\begin{aligned}\phi(t_1) - \phi(t_0) &= \frac{1}{2\pi} \left[ \pi + \arctan \left( \frac{A'(t_1)}{2\pi A(t_1)\phi'(t_1)} \right) \right. \\ &\quad \left. - \arctan \left( \frac{A'(t_0)}{2\pi A(t_0)\phi'(t_0)} \right) \right].\end{aligned}\tag{4.22}$$

Dado que los dos argumentos de las funciones arcotangente son  $O(\epsilon)$  debido a la cota para  $|A'(t)|$  dada en la definición de IMT (Def. 2.4.5), se demuestra que  $\phi(t_1) - \phi(t_0) \rightarrow 1/2$  cuando  $\epsilon \rightarrow 0$ .  $\square$

Esta estimación nos dice que podemos aproximar la media de la escala local en dos extremos locales consecutivos mediante el doble de la distancia entre ellos.

Además, y al igual que antes, usamos los extremos locales de  $x(t)$  como estimaciones de los extremos locales del modo buscado  $d(t)$  [83, 93]. Por lo tanto, podemos definir

$$\alpha(t) = 2(t_{\ell+1} - t_\ell), \quad t_\ell < t < t_{\ell+1},\tag{4.23}$$

donde  $\{t_\ell\}$  son los extremos locales de  $x(t)$ . Esta idea ya había sido esbozada en la contribución original de EMD [58], donde se afirma que “...la escala de tiempo característica está definida por la distancia temporal entre los extremos...”. Discutimos oportunamente esta idea en la Sec. 2.4.5.4, donde mencionamos el término *empiquency*, contracción del inglés *empirical mode frequency*, acuñado en [74].

### 4.4.3. Problemas de discontinuidad y *synchroqueezing*

Nuestra estimación constante a trozos de la escala local (ec. (4.23)) puede causar algunos problemas de discontinuidad para el modo  $d(t)$ . Para que  $d(t)$  sea continua en  $t_0$  (extremo local de  $x(t)$ ) debe darse  $d(t_0) = \lim_{t \rightarrow t_0} \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^{C\alpha(t)} W_x^\psi(t, s) \frac{ds}{s} \right)$ , lo cual puede garantizarse si  $W_x^\psi(t_0, s) = 0$ , para  $\min(\alpha(t_0^-), \alpha(t_0^+)) < s < \max(\alpha(t_0^-), \alpha(t_0^+))$ . Difícilmente sea este el caso general. Debido a ello, debemos usar *synchrosqueezing*. La CWT reasignada verticalmente, o sea con *synchrosqueezing*, al ser una versión «concentrada» contiene menos coeficientes no nulos, mientras que el dominio de cada modo (su franja) se hace «más pequeño». De esta manera, la obtención de nuestros modos resulta de la siguiente manera:

$$d(t) = \frac{2}{C_\psi} \mathcal{R}e \left( \int_0^{C\alpha(t)} SW_x(t, s) ds \right). \quad (4.24)$$

### 4.4.4. La constante C

Para la constante  $C$ , debemos mirar en el Teo. 4.3.1. Para una IMT, la «franja» en su CWT es  $[\frac{f_\psi - \Delta}{\phi'(t)}, \frac{f_\psi + \Delta}{\phi'(t)}]$ . Dado que sólo estamos interesados en la frontera de escalas toscas (y por lo tanto frecuencias bajas), y que  $\alpha(t)$  es una aproximación de la escala instantánea  $1/\phi'(t)$ , entonces  $C = f_\psi + \Delta$  es la opción correcta.

### 4.4.5. Refinamiento de la escala local

No debemos perder de vista el hecho de que los extremos locales de  $x(t)$  son estimaciones de los extremos locales del modo buscado  $d(t)$ . Esto es así en todos los métodos basados en extremos locales, como nuestra propuesta para EMD basada en optimización (Cap. 3) y las presentadas en [83, 92, 93]. Luego de aplicar la ec. (4.24), los extremos de este proto-modo están más cerca de los del modo real que los extremos de  $x(t)$ . Entonces podemos usarlos para realizar un refinamiento de  $\alpha(t)$ , y volver a estimar el modo  $d(t)$  a partir de la CWT reasignada verticalmente de  $x(t)$ . Este procedimiento es similar al proceso de *sifting* de EMD (Def. 2.4.4), pero mientras en EMD consiste en la sustracción sucesiva de medias locales, aquí lo usamos sólo como una manera de mejorar la selección de puntos característicos para definir la escala local. Esto está en congruencia con algunas interpretaciones del *sifting* de EMD como una forma de mejorar la selección de puntos característicos [95].

Este refinamiento puede realizarse hasta que se cumpla cierto criterio, tal como que  $\|\frac{2}{C_\psi} \mathcal{R}e(\int_0^{C\alpha_d(t)} SW_d(t, s) ds)\|^2 / \|d(t)\|^2$  sea suficientemente cercano a la unidad y la cantidad de extremos de  $d(t)$  permanezca igual a la de  $x(t)$ . Hemos enfatizado aquí que la escala local fue estimada a partir del proto-modo.

### 4.4.6. Tratamiento de bordes

Nuestra propuesta presenta efectos de bordes por las mismas causas que los de EMD, oportunamente discutidos en la Sec. 2.4.7.1. La escala local definida en la ec. (4.23) es incompleta, debido a que no está definida entre el borde izquierdo y el

primer extremo local, ni entre el último extremo local y el borde derecho de la señal. Como en el algoritmo clásico de EMD, una estrategia de espejado de los bordes lleva a buenos resultados.

## 4.5. Experimentos y resultados

Para nuestros experimentos usamos la *bump wavelet*, caracterizada por su transformada de Fourier

$$\hat{\psi}(f) = e^{1 - \frac{1}{1 - \left(\frac{f - f_\psi}{\Delta}\right)^2}} \chi_{[f_\psi - \Delta, f_\psi + \Delta]}, \quad (4.25)$$

la cual posee soporte compacto y admite un único pico en  $f = f_\psi$ . Para la frecuencia central, usamos  $f_\psi = 0,5$ , que es la máxima frecuencia normalizada para señales de tiempo discreto. De esa forma, la escala más fina posible será  $s = 1$ , que es la escala más pequeña para señales de tiempo discreto. La resolución en escala está condicionada por el valor del parámetro  $\Delta$ . En nuestro caso, logramos buenos resultados con  $\Delta = 0,25$ . Aplicamos un esquema deflacionario: una vez que la señal se descompone como  $x(t) = d_1(t) + a_1(t)$ , la descomposición continúa sobre  $a_1(t)$ . Comparamos nuestros resultados con los obtenidos mediante detección de crestas. Para esta última tarea usamos el algoritmo de [17] con los parámetros por defecto. Para la estimación de los modos usamos la ec. (4.16).

### 4.5.1. Suma de componentes AM-FM solapados

Como primer ejemplo presentamos un típico caso de componentes AM-FM que se solapan en tiempo y en frecuencia. La señal de tiempo discreto  $x(n) = x_1(n) + x_2(n) + x_3(n)$  se compone de  $x_1(n) = \cos\left(2\pi\left(\frac{15}{256}n + \frac{9000}{512\pi}\sin\left(\frac{2\pi}{1000}n\right)\right)\right)$ ,  $x_2(n) = \cos\left(2\pi\left(\frac{15}{768}n + \frac{9000}{1536\pi}\sin\left(\frac{2\pi}{1000}n\right)\right)\right)$ , y  $x_3(n) = \exp\left(-\pi\left(\frac{n-1000}{250}\right)^2\right)\sin\left(2\pi\frac{3}{256}n\right)$ , para  $n = 1, 2, \dots, 2000$ , con separación  $D \approx 1/3$  según la Def. 4.3.2.

Los resultados se muestran en la Fig. 4.1. El refinamiento de la escala local es evidente luego de tan sólo tres iteraciones. Los errores cuadráticos medios son:  $6,9 \times 10^{-4}$ ,  $2,3 \times 10^{-3}$  y  $1,3 \times 10^{-3}$  para los modos primero y segundo y el residuo respectivamente. Superponemos los modos con los componentes verdaderos, los cuales se muestran en línea de trazos roja. Los errores para el método basado en detección de crestas son:  $2 \times 10^{-2}$ ,  $2,5 \times 10^{-2}$  y  $4,9 \times 10^{-3}$ . Para los dos primeros modos, nuestra propuesta obtiene un error que es al menos un orden de magnitud menor. El método basado en crestas presenta efectos de borde notables, además de «saltos» en la detección de las crestas.

### 4.5.2. Formas de onda no lineales

Como segundo ejemplo ofrecemos el estudio de una forma de onda triangular cuyos «períodos» sucesivos son cada vez más cortos en un estilo de *chirp* lineal, y un *chirp* propiamente dicho de la forma  $c(n) = \cos(2\pi 2 \times 10^{-5}n^2)$  para  $n = 1, 2, \dots, 2000$ . Los resultados pueden apreciarse en la Fig. 4.2. La forma de onda triangular genera más de una cresta en la representación tiempo-escala. El algoritmo

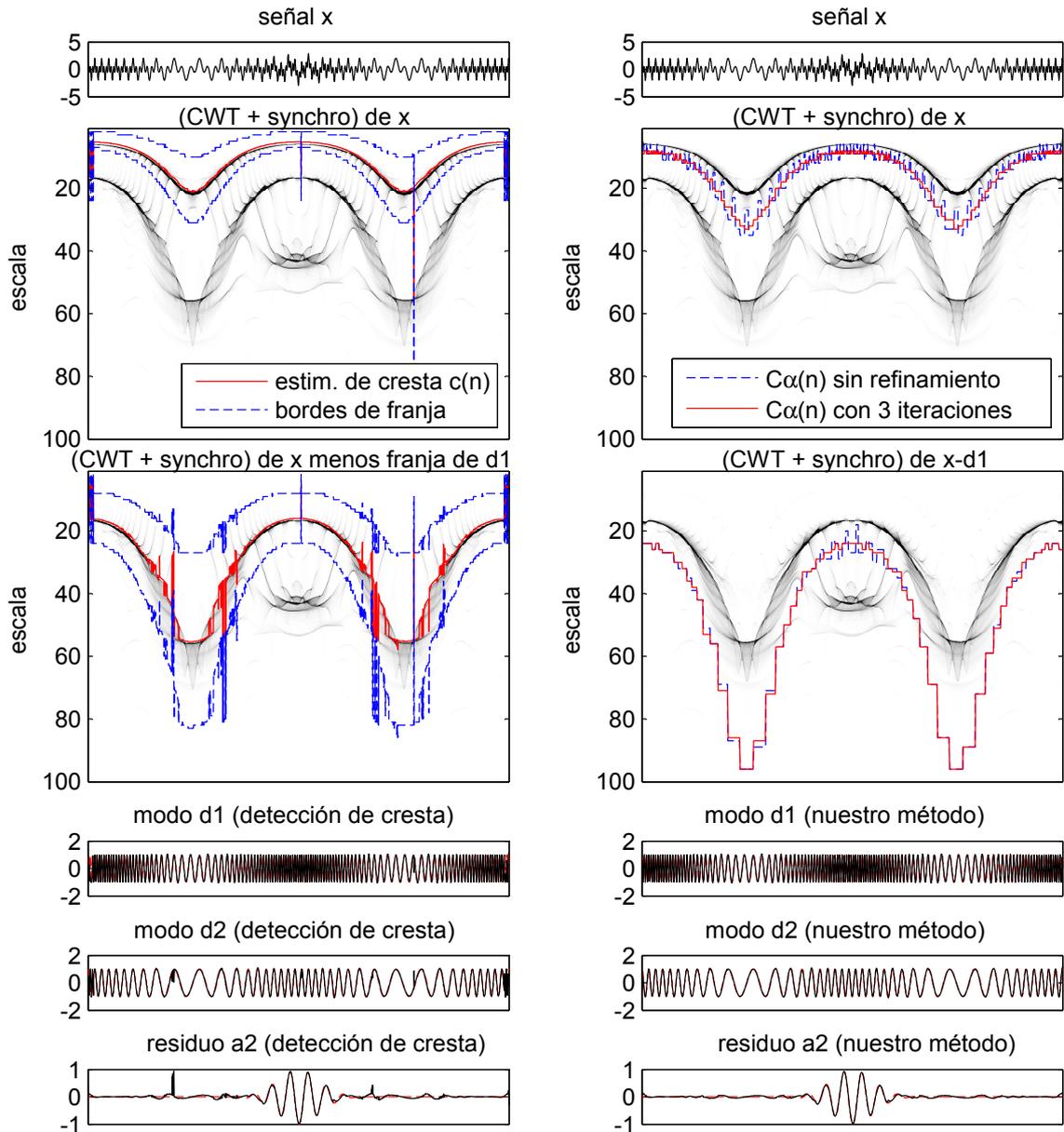


Figura 4.1: Superposición de componentes AM-FM. Izquierda: método basado en crestas. Derecha: nuestra propuesta. De arriba abajo: señal  $x(n)$ , CWT+synchrosqueezing de  $x(n)$  y el necesario para el segundo paso (depende de cada método), y los dos primeros modos y el residuo en negro superpuestos con los componentes verdaderos en línea de trazos roja.

basado en crestas para la extracción de modos necesariamente repartirá esta forma de onda en varios modos (ya que se encontrarán al menos tres crestas) y además puede ignorar la información de las escalas más finas (frecuencias altas), como podemos ver en el panel izquierdo de la Fig. 4.2. Estas limitaciones de los algoritmos que se basan en el paradigma de una cresta - un modo están siendo superadas actualmente mediante propuestas nuevas, tales como *nonlinear mode decomposition* [63] e *instantaneous wave shape functions* [113]. Nuestra propuesta, basada en los extremos locales observados en el dominio temporal, es capaz de recuperar

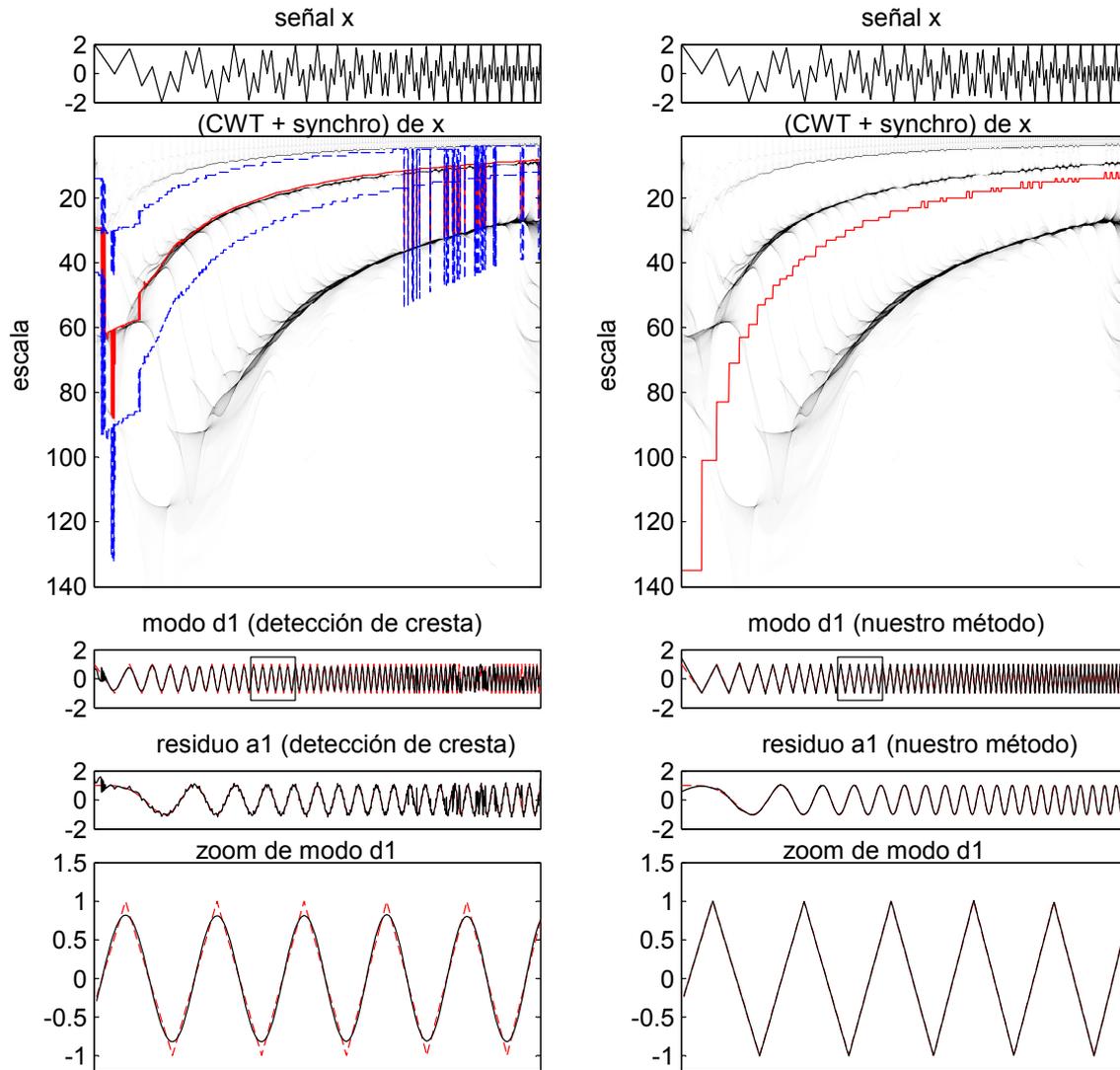


Figura 4.2: Superposición de un *chirp* y una forma de onda triangular. Izquierda: método basado en crestas. Derecha: nuestra propuesta. De arriba abajo: señal compuesta, CWT+synchrosqueezing de la señal, y primer modo y residuo en negro superpuestos con los componentes verdaderos en línea de trazos roja. Un *zoom* de la porción recuadrada de  $d_1(n)$  se muestra en la última fila.

por completo la forma de onda triangular ya que considera la información desde las escalas más finas hasta aquellas escalas definidas mediante la distancia entre extremos sucesivos. Mostramos en la última fila un *zoom* de la porción recuadrada del modo  $d_1(n)$  para una mejor apreciación. Puede notarse que el método basado en crestas ofrece una estimación suave de la forma de onda triangular (básicamente una senoidal modulada en frecuencia), incapaz de reflejar la no linealidad del fenómeno. Vemos nuevamente la aparición de saltos al momento de detectar la cresta. La recuperación provista por nuestra propuesta en cambio es satisfactoria.

Debido a que los extremos locales ocurren para esta señal allí donde ella no es derivable, las posiciones de los extremos de la señal compuesta coinciden con las posiciones del modo verdadero. Entonces, no fue necesario el refinamiento de la escala local. El error cuadrático medio cometido es:  $2,9 \times 10^{-3}$ . Como en el ejemplo

anterior, superpusimos los modos con los componentes verdaderos. El error para el algoritmo basado en detección de crestas es:  $4,9 \times 10^{-2}$ . Aquí obtenemos un error de un orden de magnitud menor.

## 4.6. Comentarios finales

En este capítulo continuamos persiguiendo el objetivo trazado al inicio del capítulo anterior: intentar dotar a EMD de bases teóricas sólidas. A diferencia del capítulo anterior, donde exploramos las técnicas de optimización, aquí seguimos la otra tendencia para analizar teóricamente a EMD, iniciada por Daubechies [31]: la de las representaciones tiempo-escala.

Desarrollamos un método de descomposición de señales fusionando las ideas de *synchrosqueezing* con las de EMD. La localidad y naturaleza guiada por los datos de esta última técnica, las cuales se basan en los extremos locales de la señal como puntos característicos para definir una escala local, se encontraron con las sólidas bases teóricas detrás de las onditas y *synchrosqueezing*. El resultado es un método tan guiado por los datos como EMD, evitando las *splines* y cualquier tipo de interpolación en el dominio temporal.

Evaluamos el método propuesto en dos señales artificiales con componentes solapados en tiempo y en frecuencia. Más aún, una de las señales analizadas contenía una forma de onda triangular que genera varias crestas en el plano tiempo-escala. Nuestro método fue capaz de aislarla por completo en un único modo.

Nuestra idea de realizar una descomposición EMD en un marco tiempo-escala debería ayudar a comprender mejor la forma en la que trabaja EMD. Las ideas planteadas en este capítulo dieron origen a un artículo presentado en el congreso europeo de procesamiento de señales EUSIPCO 2016 [2].



# Capítulo 5

## Versiones de EMD asistidas por ruido

*Creo en el alba oír un atareado rumor de multitudes que se alejan...*

---

LÍMITES  
Jorge Luis Borges

### Contenidos

---

<b>5.1. Introducción</b>	<b>101</b>
<b>5.2. El problema de la mezcla de modos</b>	<b>101</b>
<b>5.3. Variaciones asistidas por ruido</b>	<b>102</b>
5.3.1. EMD por conjuntos	103
5.3.2. EMD por conjuntos complementaria	105
<b>5.4. EMD por conjuntos completa con ruido adaptativo</b>	<b>106</b>
5.4.1. Antecedente	108
<b>5.5. Un estudio sobre la amplitud de ruido</b>	<b>109</b>
5.5.1. Ruido de amplitud constante	110
5.5.2. «Salto» en la amplitud del ruido	111
5.5.3. Ruido de amplitud creciente	111
5.5.4. Discusión	112
<b>5.6. Experimentos y resultados</b>	<b>114</b>
5.6.1. Señales artificiales	114
5.6.2. Electroglotograma	119
5.6.3. Electrocardiograma con fibrilación ventricular	121
5.6.4. Electroencefalograma con crisis epiléptica	122
<b>5.7. Comentarios finales</b>	<b>129</b>

---



## 5.1. Introducción

En este capítulo discutiremos las versiones de EMD asistidas por ruido. Todas ellas, a pesar de sus diferencias, tienen en común que la señal analizada no se descompone directamente. En este tipo de métodos, lo que EMD procesa es una suma de la señal más un ruido particular. De esta manera, se «fuerza» a la técnica a concentrarse algunas escalas en particular, que resultan de la combinación de la señal analizada y el tipo especial de ruido agregado.

En la Sec. 5.2 se describe brevemente el problema de la mezcla de modos, y en la Sec. 5.3 se presentan las variaciones asistidas por ruido de EMD, ideadas para paliar dicho problema. En la Sec. 5.4 presentamos nuestra propia versión de EMD asistida por ruido. En la Sec. 5.5 estudiamos la influencia de un parámetro crucial: la amplitud de ruido agregado. Finalmente, en la Sec. 5.6 pueden encontrarse experimentos con señales reales y artificiales.

## 5.2. El problema de la mezcla de modos

Posiblemente una de las principales dificultades de EMD en su aplicación a señales reales sea el fenómeno de *mezcla de modos*. Dicho fenómeno está descrito de una manera algo vaga: presencia de escalas similares en diferentes modos, o escalas muy disímiles presentes en un modo. Tal resultado es una consecuencia natural de la forma en la que EMD trabaja: dada una señal, el primer modo contiene la frecuencia *local* más alta. Esto no es un problema de implementación algorítmica, dado que la capacidad de separar señales en oscilaciones localmente rápidas y lentas de una manera guiada por los datos es el propósito principal del método. Resumiendo, la mezcla de modos es una consecuencia de la naturaleza local de EMD, y esta naturaleza local se logra usando los extremos locales como puntos característicos.

Un ejemplo muy simple de mezcla de modos puede apreciarse en la Fig. 5.1. Mientras una interpretación puede ser que el método falló en aislar el tono más lento cuya duración coincide con la de la señal compuesta, otra puede ser que el método tuvo éxito al capturar la frecuencia local más alta en su primer modo. Este tipo de señales *intermitentes*, tal como la señal de alta frecuencia que es capturada en el primer modo de la Fig. 5.1, son una de las principales causas, junto con el ruido, de la mezcla de modos.

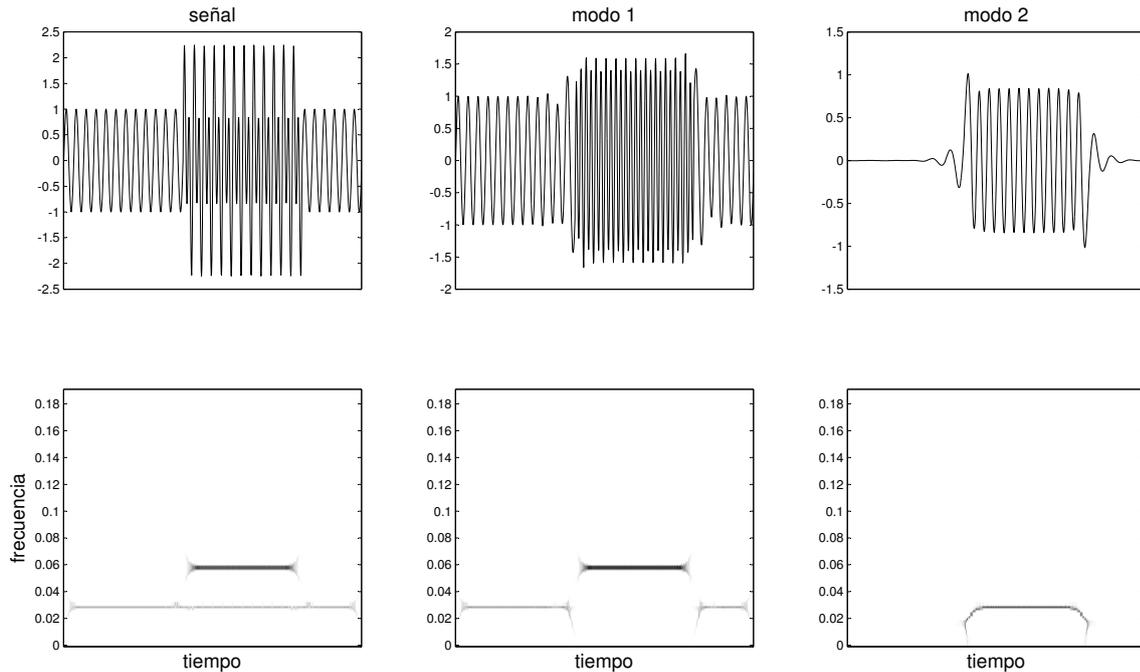


Figura 5.1: Mezcla de modos. La señal que se compone de un tono puro sostenido y de uno intermitente se analiza mediante EMD, resultando de tal forma que el primer modo captura la frecuencia *localmente* más alta. Las representaciones tiempo-frecuencia son espectrogramas reasignados.

### 5.3. Variaciones asistidas por ruido

Una de las primeras ideas para aliviar este fenómeno fue el uso de una *señal de enmascaramiento*, propuesto por Deering y Kaiser hace una década [33]. La señal de enmascaramiento es una señal de alta frecuencia que debe ser sumada a la señal original con el fin de crear nuevos extremos. La creación de estos extremos fuerza a la *media local* a «pegarse» a la señal original en esos lugares donde la frecuencia y/o la amplitud son relativamente bajas (cuando se las compara con las de la señal de enmascaramiento), pero manteniéndose cerca de la media local real en el resto de la señal, donde no hubo creación de nuevos extremos. El proceso es ilustrado en la Fig. 5.2. A la señal original se le suma una senoidal pura, con lo que se generan nuevos extremos (a un tasa igual a la frecuencia de la señal de enmascaramiento) en las zonas de baja frecuencia de la señal analizada. En las zonas de alta frecuencia, en cambio, no ocurre creación de nuevos extremos. Se hace evidente la necesidad de fijar entonces los dos parámetros de la señal de enmascaramiento: su frecuencia y su amplitud.

**Notación.** Usaremos aquí una notación conveniente para la descripción de los distintos métodos. El operador  $\mathcal{E}_k(\cdot)$  es tal que dada una señal  $x$  entrega su  $k$ -ésimo modo obtenido mediante EMD. Por supuesto que los modos dependen de distintos factores del algoritmo, como por ejemplo los criterios de detención. Además, dado que la cantidad de modos es finito, no podemos garantizar *a priori* que dada una señal  $x$  exista su  $k$ -ésimo modo para cualquier valor de  $k$ . No obstante, veremos la utilidad de este operador.

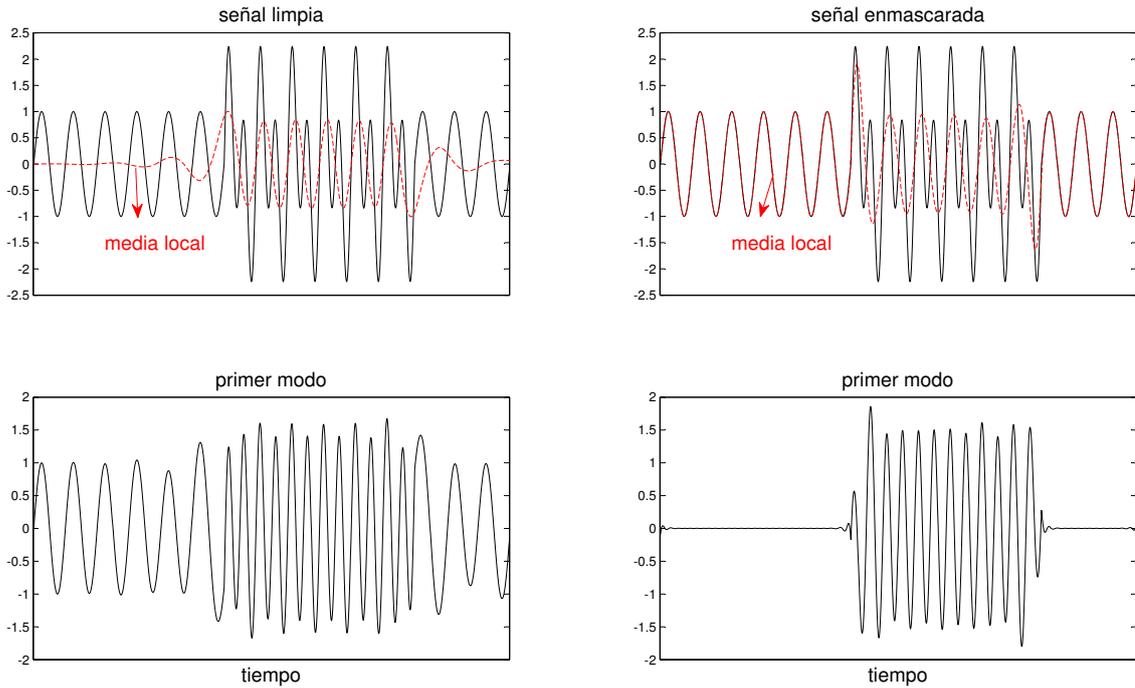


Figura 5.2: Señal de enmascaramiento. A la izquierda se muestra la descomposición con EMD de la señal limpia. A la derecha, a tal señal se le ha sumado una señal de enmascaramiento que hizo que la media local se pegue a la señal original allí donde se crearon nuevos extremos, pero que permanezca cerca de la media local real en el resto de la señal. Cada primer modo se obtiene de restar la media local a la señal original.

Además, definimos el operador  $\mathcal{M}(\cdot)$ , el cual dada una señal  $x$  devuelve la media local obtenida mediante EMD. Es tal que resulta  $\mathcal{M}(x) = x - \mathcal{E}_1(x)$ .

### 5.3.1. EMD por conjuntos

La mencionada estrategia de señal de enmascaramiento dio nacimiento a la EMD por conjuntos (EEMD, del inglés *Ensemble EMD*), presentada por Wu y Huang [118]. En lugar de preocuparse por cuál debía ser la frecuencia exacta de la señal de enmascaramiento, propusieron sumar una señal que posea todas las frecuencias posibles: ruido blanco. Para cancelar el efecto perjudicial del ruido, dejando sólo su efecto beneficioso en este caso, definieron los resultados de EEMD como el promedio del conjunto (*ensemble average* según los autores) de varias descomposiciones EMD de señal más ruido. De esa forma se estima mejor la media local que varía ligeramente entre una realización de señal más ruido y otra. Sea  $w^{(i)}$  una realización de ruido blanco con  $\mathbb{E}(w^{(i)}) = 0$  y  $\text{var}(w^{(i)}) = 1$ , donde  $\mathbb{E}(\cdot)$  y  $\text{var}(\cdot)$  indican, respectivamente, los operadores esperanza y varianza. El algoritmo resulta:

---

**Algoritmo 5.1.** Descomposición empírica en modos por conjuntos (EEMD)

---

- 1: generar  $I$  realizaciones de ruido blanco  $w^{(i)}$ ;
- 2: descomponer independientemente mediante EMD las copias ruidosas  $x + \epsilon w^{(i)}$ ;

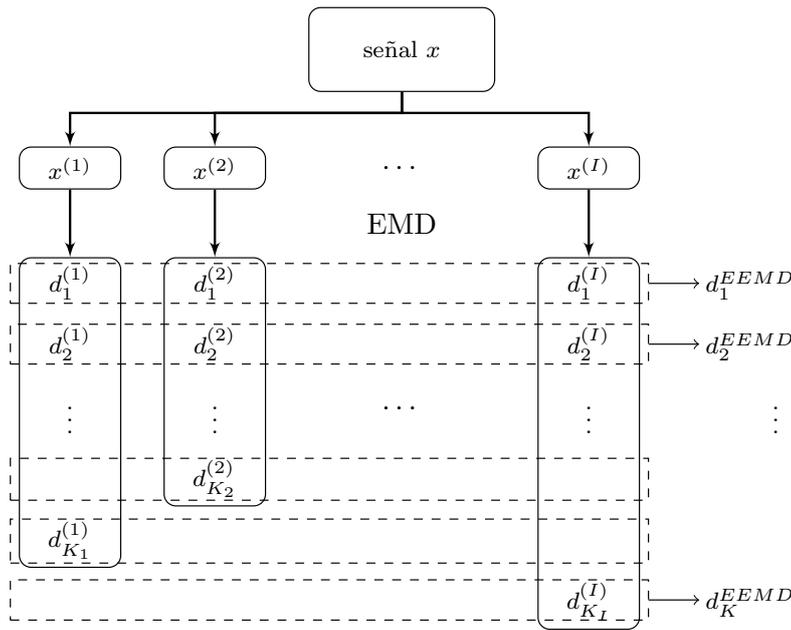


Figura 5.3: Diagrama de flujos del algoritmo de EEMD. La extracción de los modos se realiza de manera deflacionaria, sustrayendo sucesivas estimaciones de medias locales. El bucle interno es el correspondiente al proceso de *sifting* (Def. 2.4.4). Se usan los contadores  $k$  para los modos y  $j$  para las iteraciones necesarias para la extracción de cada modo.

3: definir los modos verdaderos como el promedio:

$$d_k^{EEMD} = \frac{1}{I} \sum_{i=1}^I \mathcal{E}_k(x + \epsilon w^{(i)}). \quad (5.1)$$

Si bien EEMD soluciona el problema de la frecuencia de enmascaramiento, al usar una señal que posee todas las frecuencias, genera ahora un nuevo parámetro: la cantidad de realizaciones de ruido. Si tomamos el caso particular del primer modo de EEMD, y usamos el operador  $\mathcal{M}(\cdot)$ , podemos ver que

$$\begin{aligned} d_1^{EEMD} &= \frac{1}{I} \sum_{i=1}^I \mathcal{E}_1(x + \epsilon w^{(i)}) \\ &= \frac{1}{I} \sum_{i=1}^I (x + \epsilon w^{(i)} - \mathcal{M}(x + \epsilon w^{(i)})) \\ &= x + \frac{1}{I} \sum_{i=1}^I \epsilon w^{(i)} - \frac{1}{I} \sum_{i=1}^I \mathcal{M}(x + \epsilon w^{(i)}), \end{aligned} \quad (5.2)$$

de lo cual pueden sacarse dos conclusiones relacionadas con el segundo término del miembro derecho: se debe usar la mayor cantidad de realizaciones posibles (parámetro  $I$ ), y el menor  $\epsilon$ , que resulta igual al desvío estándar del ruido agregado. Esta conclusión, que podía intuirse de antemano, nos permite independizarnos del parámetro  $I$  si usamos una cantidad suficientemente grande de realizaciones. Queda entonces sólo el parámetro  $\epsilon$ , sobre el cual volveremos más adelante.

Este método de EEMD es ampliamente utilizado en los más diversos campos [22, 70, 71, 100, 121, 126]. Sin embargo, presenta una serie de dificultades. Como primer punto puede verse claramente que  $\sum_k d_k^{EEMD} \neq x$ , con lo que existe un error de reconstrucción cuando se quiere recuperar la señal original como suma de los modos. Además, y probablemente más importante, distintas realizaciones de señal más ruido pueden producir distinta cantidad de modos, con lo que se genera una dificultad al momento de la promediación final. Si estamos particularmente interesados en un fenómeno cuyo contenido frecuencial está acotado a una banda específica de frecuencias, entonces puede ocurrir que en una realización de señal más ruido este fenómeno sea capturado por un modo dado, y que en otra realización el fenómeno aparezca en un modo de distinto índice. Estaríamos entonces «repartiendo» el fenómeno de interés entre distintos modos, promediando cosas distintas y en presencia finalmente de un caso de mezcla de modos. Estas dificultades se hacen evidentes en la Fig. 5.3, que ilustra el algoritmo de EEMD.

### 5.3.2. EMD por conjuntos complementaria

La EEMD marcó el camino para las llamadas variaciones asistidas por ruido de EMD. La EMD por conjuntos complementaria (*Complementary EEMD* en inglés), presentada por Yeh y col. [124], usa «pares» complementarios de realizaciones de ruido (una con signo positivo y otra con signo negativo) las cuales se suman a la señal antes de descomponer cada suma de señal más ruido mediante EMD. Esta adición y substracción de la misma realización de ruido alivia el ruido residual cuando se reconstruye la señal como suma de modos, aunque duplica el costo computacional para la misma cantidad de realizaciones de ruido, ya que deben realizarse dos descomposiciones EMD por cada una de ellas. Resumimos el algoritmo:

---

**Algoritmo 5.2.** Descomposición empírica en modos por conjuntos complementaria (Complementary EEMD)

---

- 1: generar  $I$  realizaciones de ruido blanco  $w^{(i)}$ ;
- 2: descomponer independientemente mediante EMD las copias ruidosas  $x + \epsilon w^{(i)}$  y  $x - \epsilon w^{(i)}$ ;
- 3: definir los modos verdaderos como el promedio de las semisumas correspondientes:

$$d_k^{CEEMD} = \frac{1}{I} \sum_{i=1}^I \frac{\mathcal{E}_k(x + \epsilon w^{(i)}) + \mathcal{E}_k(x - \epsilon w^{(i)})}{2}. \quad (5.3)$$


---

Si bien no se garantiza que la suma de los modos devuelva la señal original, en la práctica se verifica que el error es significativamente menor a aquel cometido por EEMD. Sin embargo, la EEMD complementaria posee una ventaja adicional, la cual aparentemente no fue notada por sus propios autores. Si aplicamos el operador  $\mathcal{M}(\cdot)$

en la ec. (5.3), obtenemos para el primer modo:

$$\begin{aligned}
d_1^{CEEMD} &= \frac{1}{I} \sum_{i=1}^I \frac{\mathcal{E}_1(x + \epsilon w^{(i)}) + \mathcal{E}_1(x - \epsilon w^{(i)})}{2} \\
&= \frac{1}{I} \sum_{i=1}^I \frac{x + \epsilon w^{(i)} - \mathcal{M}(x + \epsilon w^{(i)}) + x - \epsilon w^{(i)} - \mathcal{M}(x - \epsilon w^{(i)})}{2} \\
&= \frac{1}{I} \sum_{i=1}^I \left( x - \frac{\mathcal{M}(x + \epsilon w^{(i)}) + \mathcal{M}(x - \epsilon w^{(i)})}{2} \right) \\
&= x - \frac{1}{I} \sum_{i=1}^I \frac{\mathcal{M}(x + \epsilon w^{(i)}) + \mathcal{M}(x - \epsilon w^{(i)})}{2}.
\end{aligned} \tag{5.4}$$

Resulta así una reducción del ruido residual presente en el primer modo, ya que no aparece en la ec. (5.4) ningún término proporcional a  $\sum_i w^{(i)}$ , como sí ocurre en la ec. (5.2). Este resultado puede extrapolarse al resto de los modos. Sin embargo, y a pesar que el error de reconstrucción se reduce significativamente, no existen garantías de que siquiera  $x + \epsilon w^{(i)}$  y  $x - \epsilon w^{(i)}$  produzcan la misma cantidad de modos, con lo que sigue existiendo la dificultad a la hora de realizar la promediación final.

## 5.4. EMD por conjuntos completa con ruido adaptativo

Si queremos solucionar el problema de la promediación final en el que distintas realizaciones de señal más ruido producen distinta cantidad de modos, el primer paso que debemos dar es plantear, como en EMD tradicional, un esquema deflacionario. Esto es, descomponer  $x = a_1 + d_1$ , y continuar la descomposición sobre  $a_1$ , de forma que  $a_1 = a_2 + d_2$ , y así sucesivamente. En este enfoque, no se descompone cada realización de señal más ruido *independientemente* y *por completo* sino que se propone enfocarse en el primer paso de la descomposición: dada la señal establecer su modo y media local.

Para realizar este primer paso, debemos recordar la premisa del paradigma de señal de enmascaramiento: dicha señal sirve para crear extremos locales de forma que la media local se pegue a la señal allí donde se crearon y permanezca inalterada donde no ocurrió tal fenómeno. Planteamos aquí entonces un primer modo de la forma

$$d_1 = x - \frac{1}{I} \mathcal{M}(x + \epsilon v_1^{(i)}), \tag{5.5}$$

donde  $v_1^{(i)}$  es un ruido especial de *media cero*. Esta propuesta difiere sutilmente de aquella planteada por Deering y Kaiser [33] en el hecho que eliminamos el término de promediación de la señal de enmascaramiento (el ruido especial  $v_1^{(i)}$  en este caso) de la definición del primer modo, logrando una reducción del contenido de ruido residual en éste. Usamos entonces la señal de enmascaramiento sólo para lo que nos es útil: crear nuevos extremos, y la desechamos cuando ya no nos sirve.

A continuación, la descomposición sigue con  $a_1 = \frac{1}{I} \mathcal{M}(x + \epsilon v_1^{(i)})$ , descomponiendo copias de  $a_1 + \epsilon v_2^{(i)}$  (con  $v_2^{(i)}$  ruido especial de media cero) para obtener el segundo modo  $d_2$ . Entonces, estamos en condiciones de formalizar un método guiado por

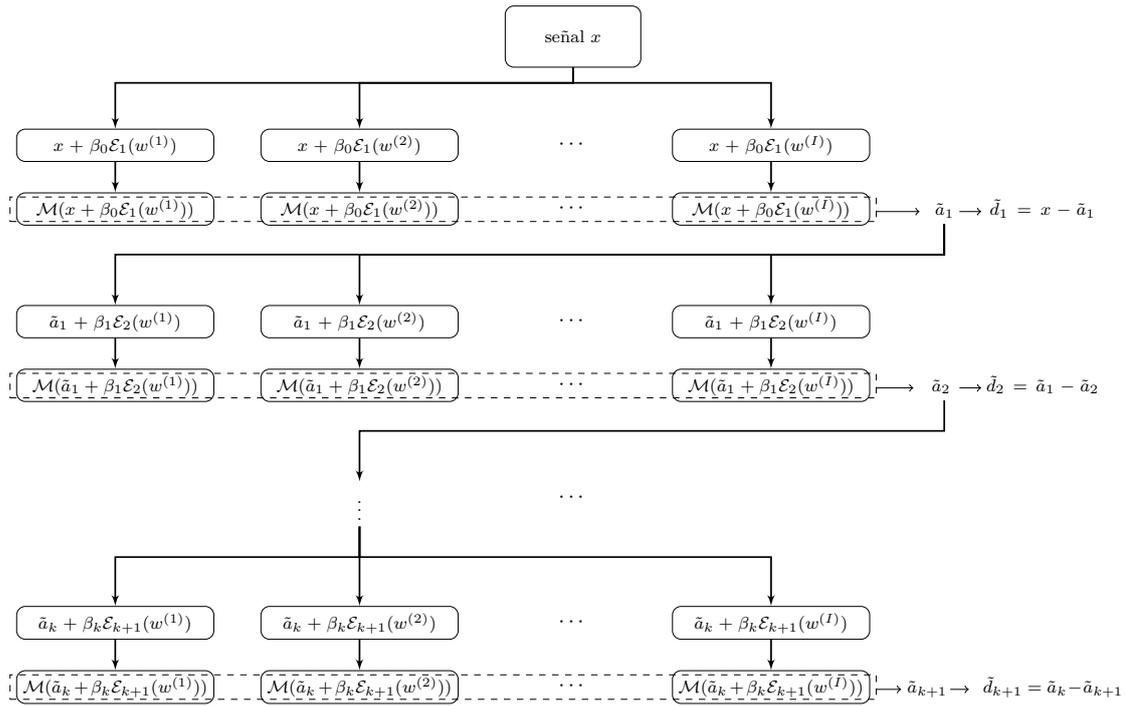


Figura 5.4: Diagrama de flujos del algoritmo de CEEMDAN. La extracción de los modos se realiza de manera deflacionaria, sustrayendo sucesivas estimaciones de medias locales. Se logra así un método con nulo error de reconstrucción.

los datos y asistido por ruido que supere las dificultades que mencionamos para los métodos existentes. Llamaremos a nuestro método EMD por conjuntos completa con ruido adaptativo (CEEMDAN, del inglés *Complete Ensemble EMD with Adaptive Noise*). Resumimos su algoritmo de la siguiente manera:

---

**Algoritmo 5.3.** Descomposición empírica en modos por conjuntos completa con ruido adaptativo (CEEMDAN)

---

- 1: generar una colección de modos de ruido blanco descomponiendo mediante EMD  $I$  realizaciones  $w^{(i)}$ , con  $\mathbb{E}(w^{(i)}) = 0$  y  $\text{var}(w^{(i)}) = 1$ ;
- 2: obtener mediante EMD las medias locales de  $x + \beta_0 \mathcal{E}_1(w^{(i)})$  para obtener el primer residuo o media local

$$\tilde{a}_1 = \frac{1}{I} \sum_{i=1}^I \mathcal{M}(x + \beta_0 \mathcal{E}_1(w^{(i)})); \quad (5.6)$$

- 3: en la primera etapa ( $k = 1$ ) calcular el primer modo como:  $\tilde{d}_1 = x - \tilde{a}_1$ ;
- 4: estimar el segundo residuo como el promedio de medias locales obtenidas mediante EMD de realizaciones  $\tilde{a}_1 + \beta_1 \mathcal{E}_2(w^{(i)})$  y definir el segundo modo:

$$\tilde{d}_2 = \tilde{a}_1 - \tilde{a}_2 = \tilde{a}_1 - \frac{1}{I} \sum_{i=1}^I \mathcal{M}(\tilde{a}_1 + \beta_1 \mathcal{E}_2(w^{(i)})); \quad (5.7)$$

5: para  $k = 3, \dots, K$  calcular el  $k$ -ésimo residuo

$$\tilde{a}_k = \frac{1}{I} \sum_{i=1}^I \mathcal{M}(\tilde{a}_{k-1} + \beta_{k-1} \mathcal{E}_k(w^{(i)})); \quad (5.8)$$

6: calcular el  $k$ -ésimo modo

$$\tilde{d}_k = \tilde{a}_{k-1} - \tilde{a}_k. \quad (5.9)$$

La cantidad de modos  $K$  se determina por completo mediante los datos y los criterios de detención. El esquema deflacionario permite evaluar un criterio de detención global (i.e. hasta cuándo seguir descomponiendo) luego de cada extracción de modo. Esto lo diferencia notablemente de EEMD y *Complementary EEMD*, en la medida que éstos descomponían totalmente cada realización de señal más ruido, sin poder imponer un criterio de detención global. Las constantes  $\beta_k$  permiten definir la amplitud de ruido agregado en cada etapa. Los valores de ruido agregado se dan por lo general en forma relativa al desvío estándar de la señal a la cual se agrega ruido [118]. Entonces las constantes resultan  $\beta_k = \epsilon_k \text{std}(a_k)$ , siendo de esa manera  $\epsilon_k$  los parámetros a fijar. Describimos de manera gráfica el algoritmo propuesto en la Fig. 5.4.

Puede verse que los ruidos especiales que mencionamos anteriormente no son otra cosa que modos EMD de ruido blanco. Dichos modos son de media cero, y su contenido frecuencial es de tipo «pasa-banda» [39, 40]. Más aún, al ser el producto de una descomposición por EMD, poseen *media local cero*, lo cual es deseable para una señal de enmascaramiento. Estamos entonces forzando al método a que se enfoque en bandas específicas de frecuencia, aprovechando el comportamiento tipo banco de filtros diádico de EMD cuando se lo aplica a ruido blanco<sup>1</sup>. Pero, lejos de ser un filtrado fijo, constante, definido *a priori*, lo que implementa el método es un banco de filtros espontáneo, *a posteriori*, en función de la señal analizada y la amplitud de ruido agregado.

Si comparamos la esencia de EEMD con la de CEEMDAN podremos ver que esta última implementa mejor la idea original de señal de enmascaramiento. EEMD no estima *medias locales* sino *modos*. Esto es así porque descompone de forma independiente cada realización de señal más ruido, entonces luego de la primera etapa de la descomposición de cada realización existe una media local y un modo. Es imposible proceder de otra manera en EEMD, entonces el *modo verdadero* es el *promedio de los modos* de copias ruidosas de la señal, conteniendo él mismo algo de ruido residual. Al contrario, CEEMDAN usa cada modo final para el cálculo del modo siguiente. Cada modo se estima secuencialmente, en un esquema deflacionario. Entonces podemos proceder de otra manera y estimar las medias locales de cada realización de señal más ruido definiendo el *modo verdadero* como la diferencia entre el *residuo actual* y el *promedio de sus medias locales*.

### 5.4.1. Antecedente

En la sección anterior hemos presentado una versión mejorada del algoritmo original de CEEMDAN, el cual fuera propuesto por nosotros en [107]. Aquél

<sup>1</sup>El comportamiento de EMD en ruido blanco es equivalente al de un banco de filtros autosimilar. El factor de autosimilaridad es 2 (o sea, banco diádico) sólo para diez iteraciones de *sifting* [77, 95]

algoritmo, que nació con el objetivo de superar algunas dificultades remanentes en la EEMD, presentaba sin embargo algunos puntos a mejorar.

En primer lugar, los modos de CEEMDAN original se estimaban de manera distinta. El primer modo era idéntico a aquél de EEMD (ec. (5.2)), y el resto de ellos se obtenía como  $d_k = \mathcal{E}_1(a_{k-1} + \epsilon_{k-1} \mathcal{E}_k(w^{(i)}))$ . Entonces existía un ruido residual en los modos (esto puede verse en la ec. (5.2) como ya lo mencionamos), lo cual fue solucionado en esta versión mejorada de CEEMDAN que presentamos. Esto podrá confirmarse más adelante.

Como segundo punto, en la versión original de CEEMDAN existían modos espurios en las etapas tempranas de la descomposición. Esto se debía principalmente a una superposición de escalas entre el ruido blanco agregado para la obtención del primer modo (como en EEMD, recordemos) y los modos EMD de ruido blanco, que poseen espectros tipo pasa-banda. En la versión mejorada de CEEMDAN, no hacemos uso de ruido blanco sino sólo de sus modos EMD, evitando entonces tal solapamiento frecuencial y los modos espurios.

En lo sucesivo, nos referiremos a CEEMDAN original [26, 107] y CEEMDAN mejorada.

## 5.5. Un estudio sobre la amplitud de ruido

Mencionamos anteriormente la necesidad de fijar dos parámetros para las versiones de EMD asistidas por ruido: la cantidad de realizaciones  $I$  y la amplitud de ruido agregado  $\epsilon$ . También mencionamos una certeza respecto a  $I$ : debe ser lo más grande posible, idealmente infinito. Sin embargo, se obtienen buenos resultados con un número relativamente pequeño (unas pocas decenas). De esta manera, la atención recae sobre el parámetro restante: la amplitud de ruido agregado  $\epsilon$ .

Con el fin de entender mejor la influencia de la amplitud de ruido agregado tanto en EEMD como en CEEMDAN mejorada, corrompimos un tono puro de la forma

$$s(n) = \sin(0,2n), \quad n = 1, 2, \dots, 256, \quad (5.10)$$

con tres tipos diferentes de ruido (los cuales describiremos en breve). Cada realización de ruido contaminante se construyó a partir de una misma realización para los tres casos, variando su «envolvente», y las mismas realizaciones se usaron para los dos métodos. La Fig. 5.5 resume este enfoque.

Para el ruido contaminante usamos distintas relaciones señal a ruido (SNR, del inglés *Signal-to-Noise Ratio*), las que fueron denominadas como «SNR de entrada», y también diferentes valores para los parámetros de los algoritmos. Dado un conjunto de parámetros (tanto de entrada como del algoritmo), descompusimos  $J = 100$  copias ruidosas del tono puro mediante los dos métodos, y calculamos una distancia  $e$  para cada una de ellas:

$$e(\eta, \epsilon, I)_j = \left\| s(n) - d_{k_j^*}(\eta, \epsilon, I)(n) \right\|_2, \quad (5.11)$$

donde  $j = 1 \dots J$  indica la realización de ruido contaminante,  $\eta$  es la SNR de entrada,  $\epsilon$  la amplitud de ruido agregado,  $I$  la cantidad de realizaciones,  $d_{k_j^*}$  el  $k^*$ -ésimo modo obtenido a partir del  $j$ -ésimo tono ruidoso (mediante cualquiera de los dos métodos),

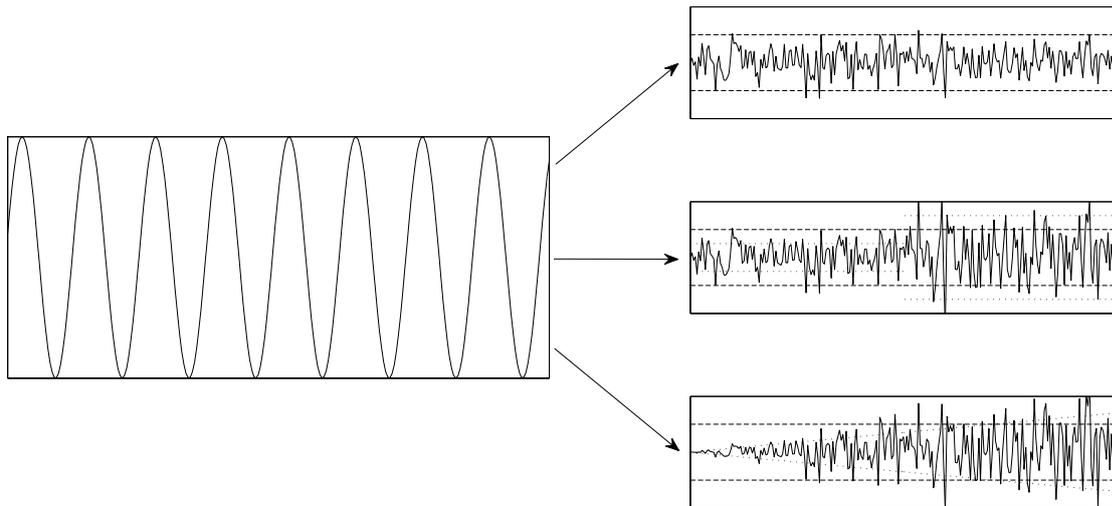


Figura 5.5: Diseño del experimento. La señal de prueba es un tono puro (izquierda) contaminada con 3 tipos diferentes de ruido (derecha). Para la misma energía total, la amplitud de las realizaciones de ruido es constante (arriba), con una duplicación abrupta de su valor a la mitad de la duración de la señal (en el medio), o linealmente creciente (abajo). Se indica esquemáticamente en línea de trazos la amplitud instantánea de la realización de ruido.

y

$$k_j^* = \arg \min_k \|s(n) - d_{k_j}(\eta, \varepsilon, I)(n)\|_2 \quad (5.12)$$

indica el modo más cercano (en norma 2) al tono original libre de ruido.

Los resultados finales se obtuvieron promediando sobre las  $J$  distancias obtenidas de las descomposiciones de tonos ruidosos:  $\bar{e}(\eta, \varepsilon, I) = \langle e(\eta, \varepsilon, I)_j \rangle$ . Un valor pequeño de  $\bar{e}$  indica una buena recuperación del tono y por ello una baja presencia de mezcla de modos. Por el contrario, si  $\bar{e}$  posee un valor grande, esto significa que la mezcla de modos fue fuerte y el algoritmo no fue capaz de aislar el tono con suficiente energía en un solo modo.

Nos valimos del error de reconstrucción nulo de CEEMDAN, y usamos pocas realizaciones de ruido de algoritmo (hasta 20). Para realizar una comparación justa entre los dos métodos, utilizamos exactamente las mismas realizaciones de ruido de algoritmo para ambos. Las iteraciones de *sifting* no fueron limitadas y se usó el criterio de detención por defecto [99]. Se usaron varios valores para el parámetro de ruido de algoritmo  $\varepsilon$ . En el intervalo  $[0,025; 0,5]$  se avanzó con un paso pequeño de 0,025, mientras que para el intervalo  $[0,6; 2,0]$  el avance fue con un paso grande de 0,1.

### 5.5.1. Ruido de amplitud constante

Como primer experimento, el ruido fue agregado con la misma amplitud para toda la duración de la señal. Entonces, descompusimos 100 copias ruidosas del tono de la forma

$$s_{\sigma_0}(n) = s(n) + \sigma_0 \text{std}(s) w^j(n), \quad (5.13)$$

utilizando diferentes valores para los parámetros de ambos algoritmos, donde  $w^j(n)$  representa a una realización de ruido blanco gaussiano de media cero y varianza unitaria. La constante  $\sigma_0$  se definió de manera de obtener las siguientes SNRs: 9 dB, 6 dB, 3 dB y 0 dB.

En la Fig. 5.6 presentamos los resultados del promedio de las distancia obtenidas a partir de las descomposiciones con ambos métodos. En estas gráficas, la SNR de entrada se mantuvo fija, y se superpusieron los resultados para cuatro cantidades de realizaciones distintas. En general, los resultados de ambos algoritmos son similares. Sin embargo CEEMDAN presenta un menor error, y una mayor uniformidad entre los resultados obtenidos con distintas cantidades de realizaciones. Los resultados que EEMD ofrece con 20 realizaciones, CEEMDAN parece obtenerlos con tan sólo 5. En las dos primeras filas, los métodos presentan una situación de dos regímenes, con un mínimo local cerca de 0,2 para el ruido de algoritmo. Esta situación desaparece para las filas tercera y cuarta, para las cuales el mínimo se trasladó considerablemente a la derecha.

### 5.5.2. «Salto» en la amplitud del ruido

Como segundo caso, el ruido contaminante fue agregado con una amplitud en la primera mitad de la señal, y con el doble de amplitud en la segunda mitad. Se descompusieron tonos ruidosos de la forma

$$s_{\sigma_s}(n) = s(n) + \text{std}(s) w_s^j(n) \quad (5.14)$$

con

$$w_s^j(n) = \begin{cases} \sigma_s w^j(n), & \text{si } n = 1, 2, \dots, 128, \\ 2 \sigma_s w^j(n), & \text{si } n = 129, 130, \dots, 256. \end{cases}$$

En este caso,  $w^j(n)$  representa a la misma  $j$ -ésima realización de ruido usada en el caso de ruido de amplitud constante. Para lograr las mismas SNRs que en el ejemplo anterior, las constantes se relacionan de la siguiente forma:  $\sigma_s = \sqrt{2/5} \sigma_0$ .

Los resultados manteniendo la SNR de entrada constante pueden apreciarse en la Fig. 5.7. Como en el caso anterior, las dos primeras filas presentan una situación de dos regímenes, mientras que esta situación desaparece en las filas tercera y cuarta. Nuevamente, CEEMDAN presenta un menor error y una mayor uniformidad en el desempeño para distintas cantidades de realizaciones.

### 5.5.3. Ruido de amplitud creciente

Como último ejemplo de esta sección, el ruido fue agregado con una amplitud linealmente creciente. Los tonos que se descompusieron fueron

$$s_\alpha(n) = s(n) + \alpha n \text{std}(s) w^j(n). \quad (5.15)$$

Para mantener las mismas SNRs que en los dos ejemplos anteriores, la pendiente usada fue  $\alpha = \sqrt{\frac{N}{\sum_{n=1}^N n^2}} \sigma_0$ , con  $N = 256$  la longitud de la señal. Igualmente,  $w^j(n)$  es la misma  $j$ -ésima realización usada en los ejemplos previos.

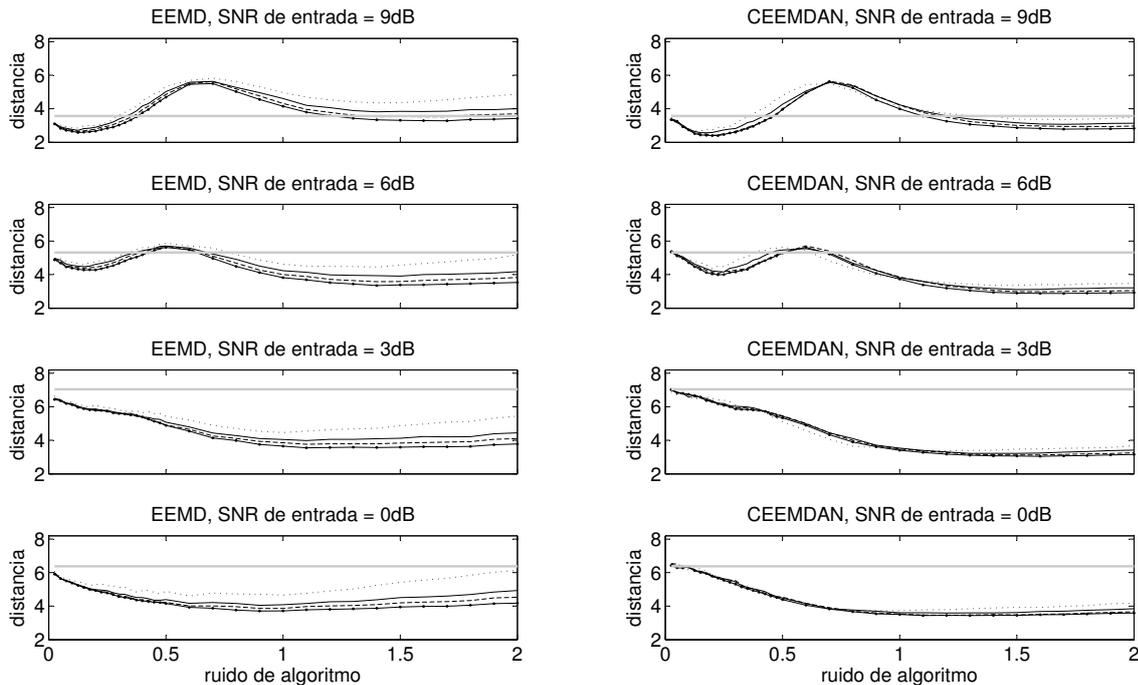


Figura 5.6: Resultados para ruido de amplitud constante. Cantidad de realizaciones: 5 (línea de puntos), 10 (línea sólida), 15 (línea de trazos) y 20 (línea sólido y puntos). El desempeño de EMD se muestra en línea gris gruesa a modo comparativo.

Los resultados se presentan en la Fig. 5.8. La SNR de entrada se mantuvo constante, variando la cantidad de realizaciones. En este caso, la situación de dos regímenes parece ser algo más fuerte en las dos primeras filas. En la tercera fila se nota la presencia de una especie de inflexión, hasta llegar a la cuarta fila donde el mínimo se ha corrido hacia la derecha. Como en los casos anteriores, el desempeño de CEEMDAN parece muy poco condicionado por la cantidad de realizaciones usadas.

#### 5.5.4. Discusión

El razonamiento detrás de las versiones de EMD asistidas por ruido es que un nivel adecuado de ruido agregado (ruido de algoritmo) debería mejorar la identificación de señales bien estructuradas contaminadas en ruido de observación (ruido de entrada). Intuitivamente, una cantidad muy pequeña de ruido agregado no debería tener prácticamente ningún efecto y el resultado sería muy similar al de EMD (esto se observa en las figuras cuando la amplitud del ruido agregado tiende a cero), mientras que una amplitud muy grande nos llevaría a formas de onda tan erráticas que la identificación seguiría siendo difícil. En el medio, es esperable encontrar algún óptimo, por ejemplo un mínimo local cuando la medida de desempeño es la distancia entre la forma de onda real (conocida) y la estimada (el modo más cercano). Es por ello notable ver que tal mínimo es en efecto observado en muchos de los ejemplos considerados, especialmente para SNRs de entrada moderadas. Más aún, esto se da más a menudo para un valor de desvío estándar de ruido de algoritmo que es cercano a 0,2, el valor sugerido por Wu y Huang en [118].

Sin embargo, existen algunas diferencias entre EEMD y CEEMDAN: los errores

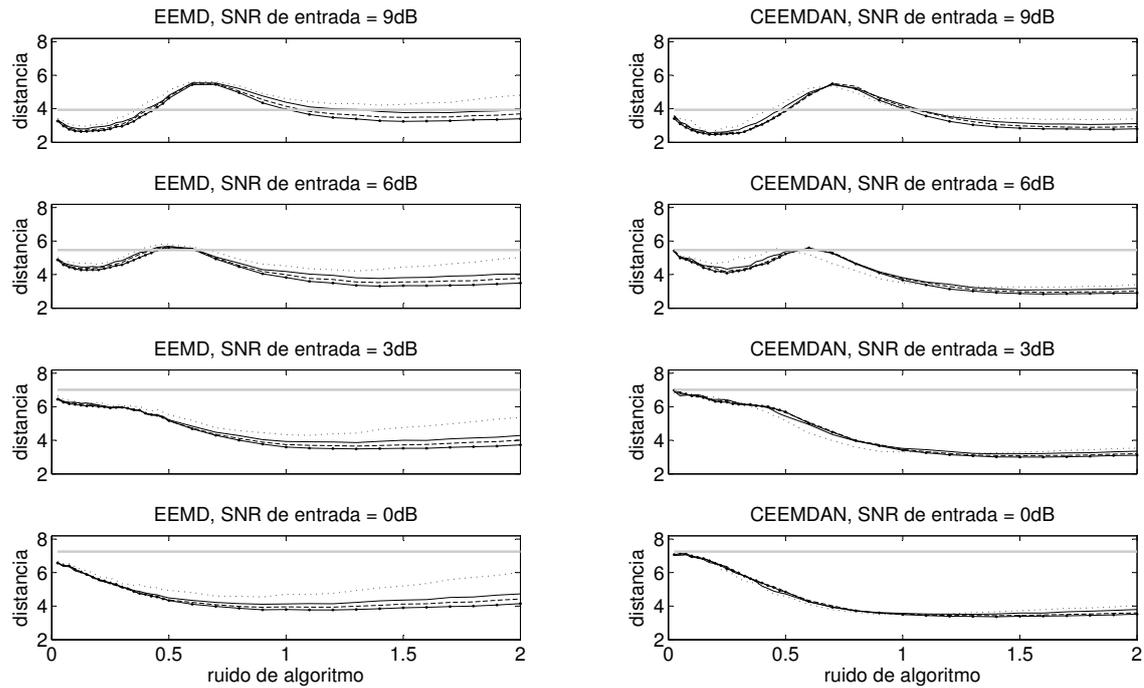


Figura 5.7: Resultados para amplitud de ruido con «salto». Cantidad de realizaciones: 5 (línea de puntos), 10 (línea sólida), 15 (línea de trazos) y 20 (línea sólido y puntos). El desempeño de EMD se muestra en línea gris gruesa a modo comparativo.

que comete CEEMDAN son menores, y además este método presenta una menor variabilidad de los resultados cuando se usan distintas cantidades de realizaciones. Es por lo tanto notable que CEEMDAN obtenga con 5 realizaciones los resultados que EEMD recién consigue con 20.

Este estudio que hemos presentado abordó la influencia de los dos parámetros clave (amplitud de ruido agregado y cantidad de realizaciones) de dos versiones de EMD asistidas por ruido en situaciones típicas, aunque no universales. Como se esperaba, ambos métodos evidenciaron la presencia de alguna forma de resonancia estocástica [43], mostrando un valor de ruido de algoritmo para el cual el desempeño es el mejor.

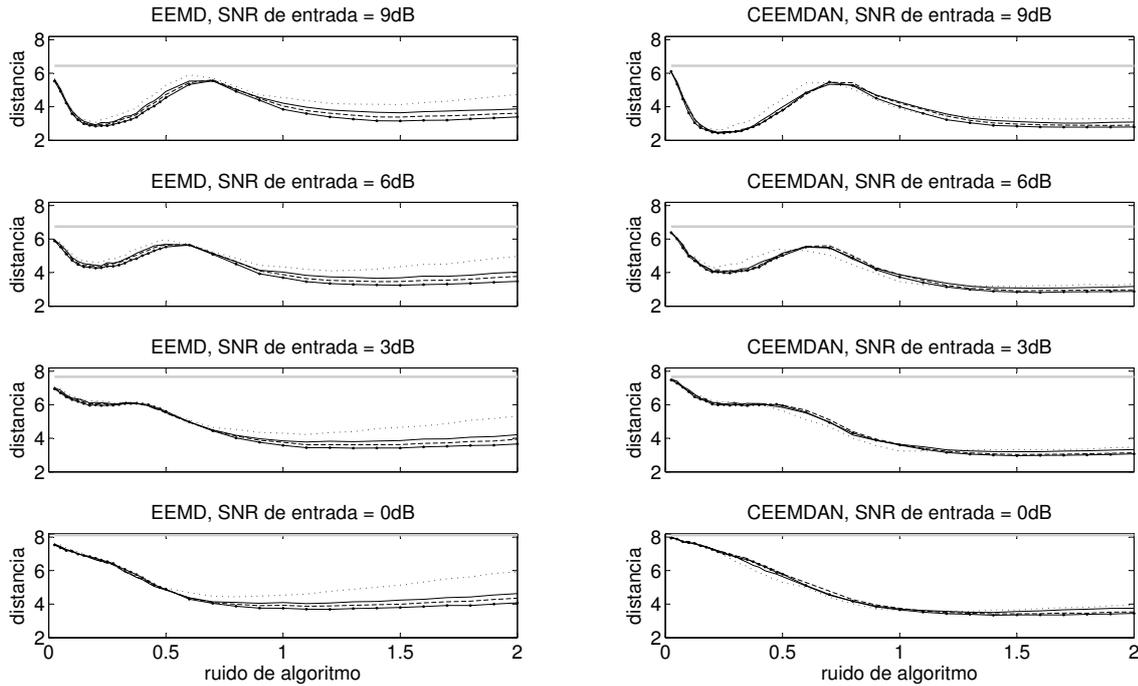


Figura 5.8: Resultados para ruido de amplitud linealmente creciente. Cantidad de realizaciones: 5 (línea de puntos), 10 (línea sólida), 15 (línea de trazos) y 20 (línea sólido y puntos). El desempeño de EMD se muestra en línea gris gruesa a modo comparativo.

## 5.6. Experimentos y resultados

En esta sección ilustramos las capacidades de la versión mejorada de CEEMDAN que hemos propuesto. Descomponemos dos señales artificiales y comparamos los resultados de nuestro nuevo método con los obtenidos mediante EMD, EEMD y la versión original de CEEMDAN. Adicionalmente, descomponemos tres señales biomédicas mediante la versión mejorada de CEEMDAN (electroglotograma, electrocardiograma y electroencefalograma) con el fin de mostrar algunas de sus aplicaciones potenciales.

### 5.6.1. Señales artificiales

Como primer ejemplo, retomamos aquí el ejemplo clásico de mezcla de modos. Un tono puro sostenido sumado a otro tono de mayor frecuencia pero de menor duración, nos conducirá inevitablemente a un fenómeno de mezcla de modos si lo analizamos mediante EMD, debido a la naturaleza local del método. La señal analizada es  $s = s_1 + s_2$ , con

$$s_1 = \begin{cases} 0 & \text{si } 1 \leq n \leq 500 \\ \text{sen}(2\pi 0,255 (n - 501)) & \text{si } 501 \leq n \leq 750 \\ 0 & \text{si } 751 \leq n \leq 1000 \end{cases} \quad (5.16)$$

$$s_2 = \text{sen}(2\pi 0,065 (n - 1)).$$

Presentamos una descomposición típica para los cinco métodos aquí analizados

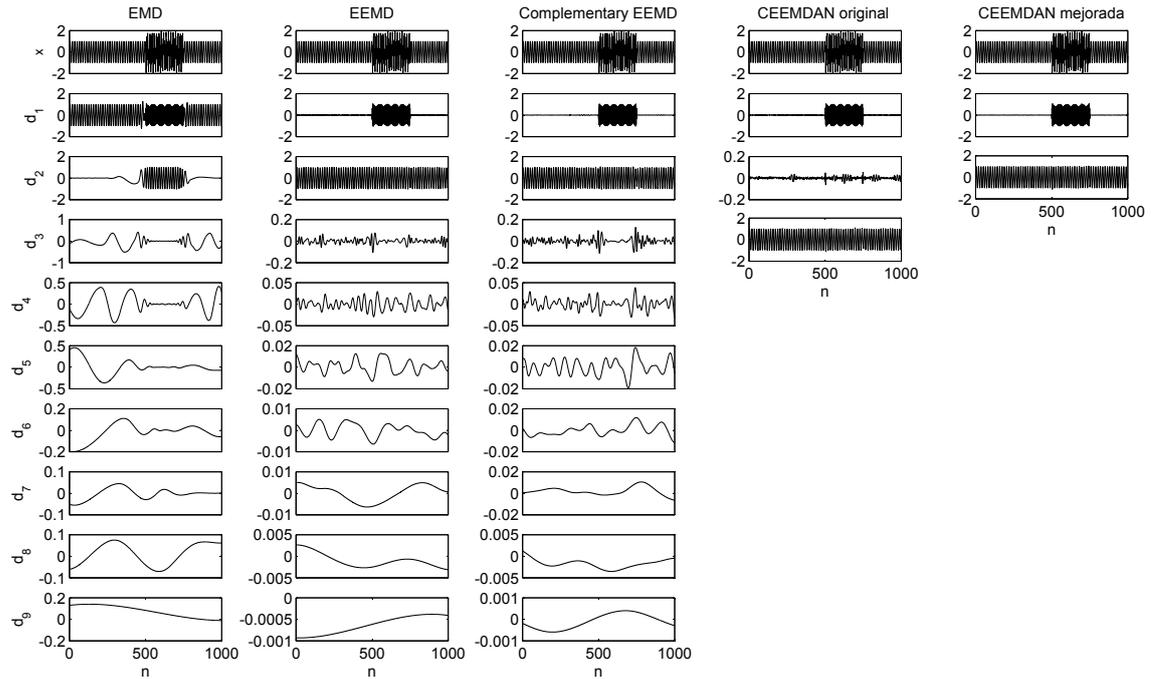


Figura 5.9: Descomposición de la señal artificial  $s$  (descrita por las ecuaciones (5.16)) mediante EMD, EEMD, Complementary EEMD, CEEMDAN original y CEEMDAN mejorada.

en la Fig. 5.9. La mezcla de modos resulta evidente para descomposición con EMD, para la cual los modos tercero en adelante poseen muy poca energía y presentan oscilaciones indeseadas. Los cuatro métodos asistidos por ruido recuperan correctamente el componente de mayor frecuencia, evitando la mezcla de modos. Sin embargo, para EEMD cada realización de señal más ruido se descompuso completamente de manera independiente y debido a ello se obtuvo un total de nueve modos, aunque desde el tercero en adelante ellos poseen poca energía, sin representar información relevante de la señal original. El mismo problema experimenta el método de Complementary EEMD. El esquema deflacionario de las versiones «completas» (CEEMDAN original y mejorada) nos permite evaluar un criterio de detención global (i.e. hasta cuando seguir extrayendo modos) en cada etapa de extracción y por eso detener antes la descomposición, una vez que las condiciones de IMF (Def. 2.4.2) son satisfechas. Un segundo modo «espurio» aparece en CEEMDAN original. La versión mejorada de CEEMDAN es capaz de aislar ambos componentes de manera satisfactoria.

Debido al ingrediente estocástico de los métodos asistidos por ruido, distintas descomposiciones de la misma señal son de hecho sutilmente diferentes. Además, la energía del ruido residual decrece a medida que aumenta la cantidad de realizaciones  $I$ . Con el objetivo de cuantificar los desempeños de los métodos, llevamos a cabo 100 descomposiciones para las cuatro variaciones asistidas por ruido de EMD (EEMD, Complementary EEMD, CEEMDAN original y mejorada), y obtuvimos resultados estadísticamente significativos. En cada caso usamos  $I = 50, 100, 200, 400, 800$  ( $I/2$  pares para Complementary EEMD) realizaciones y un valor recomendado de amplitud de ruido de  $\varepsilon_0 = 0,2$  [28, 118].

Debe notarse en la ec. (5.16) que  $s_1 = 0$  para  $1 \leq n \leq 500$  y  $751 \leq n \leq 1000$ .

La naturaleza «fino a grueso» de todos los métodos basados en EMD nos sugiere que el componente más rápido ( $s_1$  en este caso) debería ser capturado por el primer modo ( $d_1$ ). Medimos el ruido residual presente en el primer modo como la energía media de los primeros modos (de los cuatro métodos asistidos por ruido) en esos lugares donde se supone que valen cero. Para reducir la posible influencia de los efectos de bordes, consideramos los intervalos entre las muestras 11 y 490, y entre 761 y 990. Nos referimos a las energías medias de los intervalos como  $e_l(d_1)$  y  $e_r(d_1)$  respectivamente.

Para evaluar las capacidades de los métodos para recuperar componentes conocidos de una señal compuesta, usamos la raíz cuadrada del error cuadrático relativo (RRSE, del inglés *Root Relative Square Error*) de una señal recuperada  $a$  con respecto a una señal de referencia  $b$

$$RRSE_b(a) = \frac{\|a - b\|_2}{\|b\|_2}. \quad (5.17)$$

Los resultados para  $e_l(d_1)$  y  $e_r(d_1)$  se presentan en gráficos de cajas en la Fig. 5.10. El desempeño de la versión mejorada de CEEMDAN se superpuso a los desempeños de los otros tres métodos asistidos por ruido. Como era de esperarse, CEEMDAN mejorada contiene menos ruido residual en el primer modo que EEMD y CEEMDAN original. Puede apreciarse que aunque el contenido de ruido del primer modo de Complementary EEMD es menor al de EEMD, es mayor al que presenta CEEMDAN mejorada. Vale la pena notar que CEEMDAN mejorada logra para  $I = 50$  resultados que son mejores a los que los otros métodos logran para  $I = 800$ .

Los resultados de los desempeños en la recuperación de componentes conocidos se presentan en la Fig. 5.11. CEEMDAN mejorada presenta resultados mejores que los otros métodos. Aquí, la diferencia con Complementary EEMD es más evidente, especialmente en la recuperación de  $s_2$ . Finalmente, presentamos los errores de reconstrucción en la Fig. 5.12. Como se esperaba, el error para CEEMDAN mejorada es despreciable cuando se lo compara con el de EEMD. Lo que resulta algo más sorprendente es el error creciente con la cantidad de realizaciones  $I$  que presenta Complementary EEMD. Esto puede ser explicado por el hecho que diferentes copias ruidosas de la señal producen distintas cantidades de modos. Entonces, a medida que la cantidad de realizaciones aumenta, también lo hace la probabilidad de tener distintas cantidades de modos, y con ello aumenta el error de reconstrucción.

Como segundo ejemplo usamos una señal artificial de mayor complejidad. Se compone de dos átomos gaussianos y una señal FM  $x = x_1 + x_2 + x_3$ , con

$$\begin{aligned} x_1 &= 3 e^{-\pi \left(\frac{n-500}{100}\right)^2} \cos\left(2\pi \frac{5}{16} (n - 1000)\right), \\ x_2 &= \cos\left(2\pi \frac{f_{max} + f_{min}}{2} (n - 1000) + \right. \\ &\quad \left. \frac{f_{max} - f_{min}}{2} 1000 \left(\sin \frac{2\pi n}{1000} + \phi - \sin \phi\right)\right), \\ x_3 &= e^{-\left(\frac{n-1000}{200}\right)^2} \pi \cos\left(2\pi \frac{7}{256} (n - 1000)\right), \end{aligned} \quad (5.18)$$

para  $1 \leq n \leq 1000$ ,  $f_{max} = 3/32$ ,  $f_{min} = 9/128$  y  $\phi = -\arccos\left(\frac{3f_{min} + f_{max}}{f_{max} - f_{min}}\right)$ .

Descompusimos la señal *via* EMD, EEMD, Complementary EEMD y las versiones original y mejorada de CEEMDAN, usando en todos los casos asistidos por ruido un valor recomendado de amplitud de ruido de  $\varepsilon_0 = 0,2$  [28, 118] y una

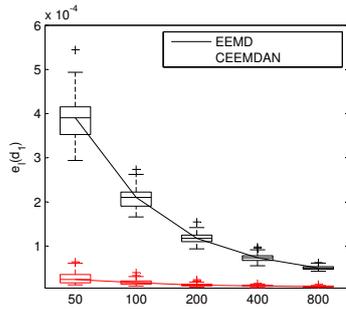


Figura 5.10: Desempeño de los cuatro métodos asistidos por ruido aplicados sobre la señal artificial  $s$  de la ec. (5.16) presentados en gráficos de cajas. Arriba: energía de la primera mitad de  $d_1$ . Abajo: energía del último cuarto de  $d_1$ .

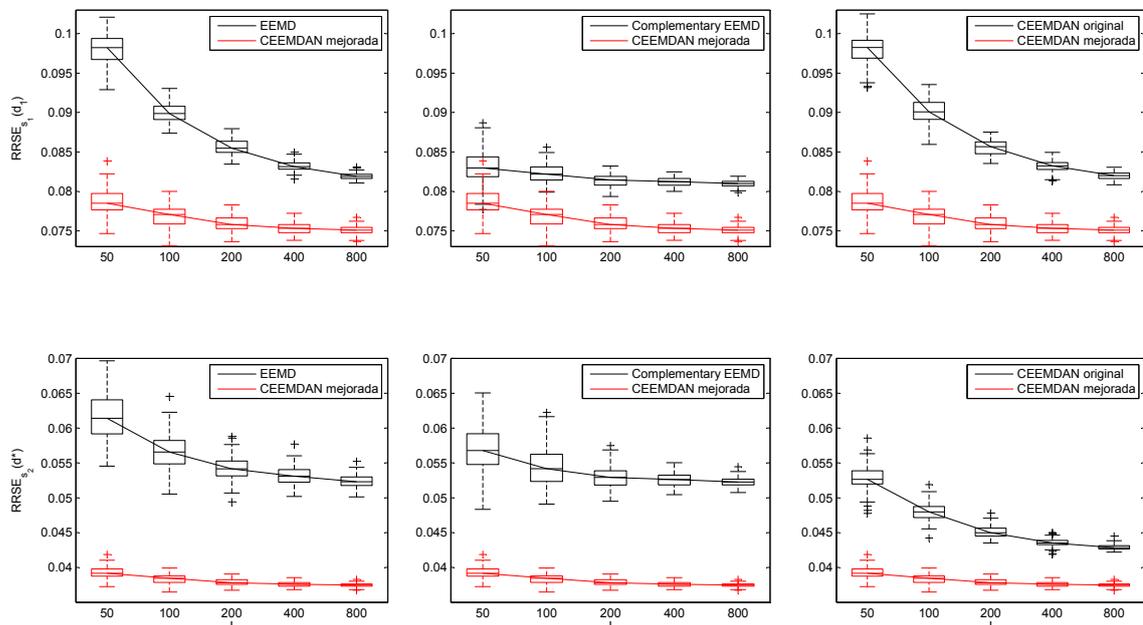


Figura 5.11: Desempeño de los cuatro métodos asistidos por ruido aplicados sobre la señal artificial  $s$  de la ec. (5.16) presentados en gráficos de cajas. Arriba: error al recuperar  $s_1$ . Abajo: error al recuperar  $s_2$ .

cantidad de realizaciones  $I = 50$  ( $I/2 = 25$  pares para Complementary EEMD).

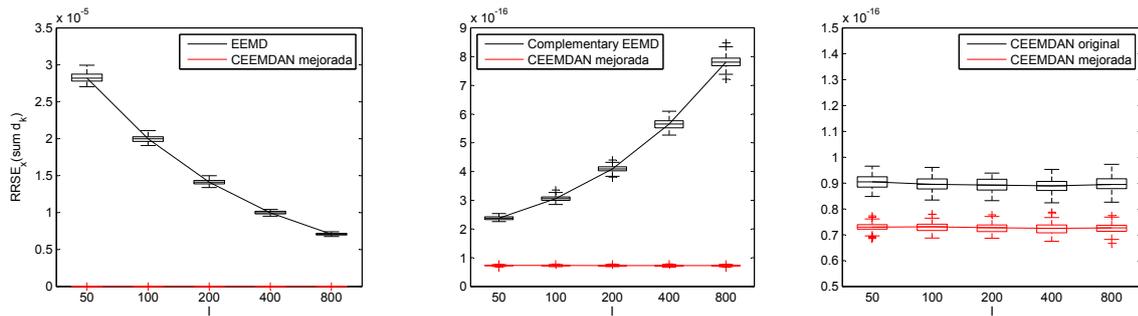


Figura 5.12: Desempeño de los cuatro métodos asistidos por ruido aplicados sobre la señal artificial  $s$  de la ec. (5.16) presentados en gráficos de cajas: error de reconstrucción.

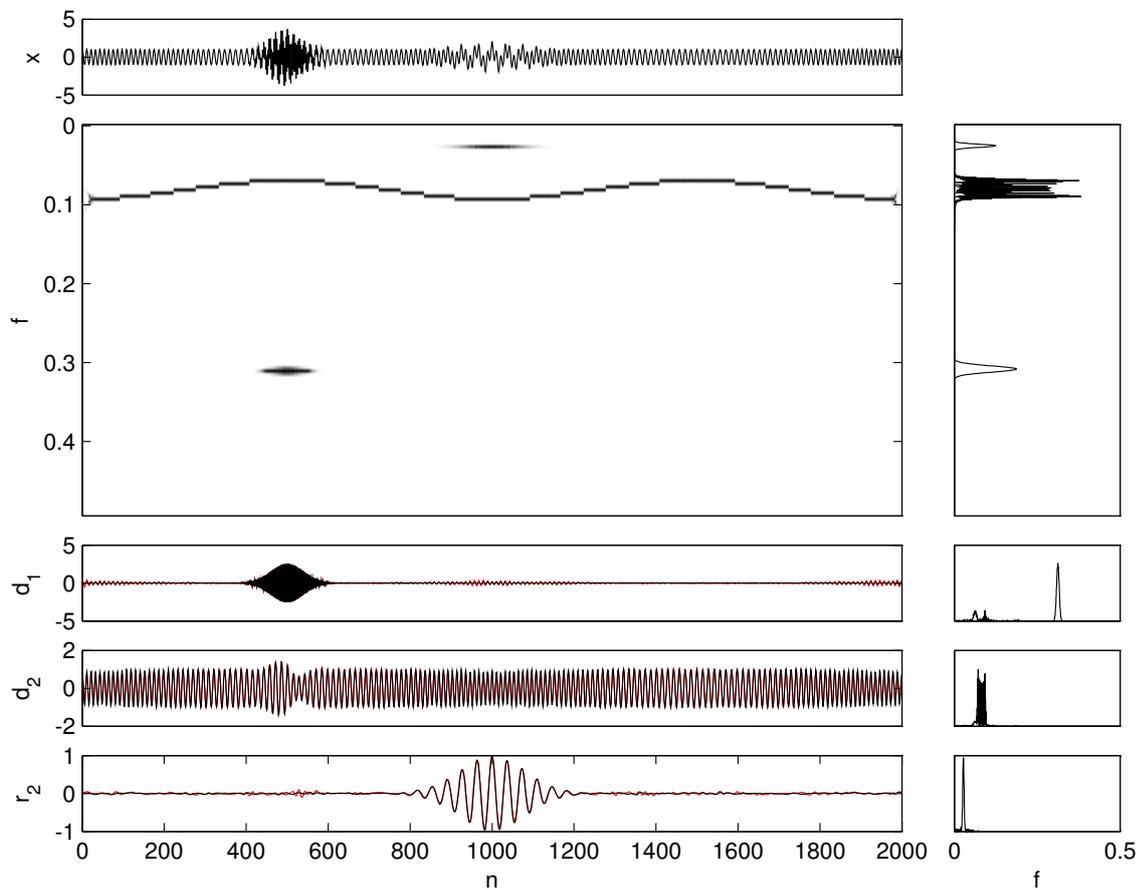


Figura 5.13: Señal artificial  $x$  de la ec. (5.18), junto a su espectrograma reasignado y espectro marginal. Los resultados de EEMD (en rojo) y CEEMDAN mejorada (en negro) se muestran junto a los espectros marginales de CEEMDAN.

Los resultados de CEEMDAN mejorada (en negro) y EEMD (en rojo) se presentan en la Fig. 5.13, en donde fueron solapados para facilitar su comparación. Además se presentan los espectros frecuenciales de los modos de CEEMDAN mejoradas. En el caso de EEMD, el segundo residuo fue creado *artificialmente*, sumando los

modos tercero a último, sólo a efectos comparativos. Esto fue así porque es imposible imponer un criterio de detención global, debido a que cada realización de señal más ruido se descompone de manera independiente una de la otra.

Puede apreciarse el comportamiento de banco de filtros guiado por los datos de EEMD y CEEMDAN mejorada. La señal  $x_1$  se aísla correctamente, así como  $x_2$  y  $x_3$ , las cuales son muy cercanas en frecuencia. Aunque podrían obtenerse resultados similares con un banco de filtros lineales, en este caso no fue necesario especificar la cantidad de filtros, ni sus frecuencias centrales o anchos de banda. El hecho de generar «espontáneamente» un banco de filtros que separa componentes con espectros distintos sin conocer su distribución *a priori* es una propiedad muy valiosa de estos métodos.

En la Fig. 5.14 se presentan las descomposiciones de la señal  $x$  mediante EMD y los cuatro métodos asistidos por ruido. Como en el ejemplo anterior, EMD presenta fuertemente el fenómeno de mezcla de modos. Para EEMD y Complementary EEMD, los modos cuarto en adelante poseen poca energía y no representan información relevante de la señal original. Sólo aparecen por el hecho ya comentado de las distintas cantidades de modos que son producidas por las distintas copias ruidosas de la señal. La versión original de CEEMDAN produjo un segundo modo espurio. La versión mejorada en cambio, recuperó correctamente los tres componentes de la señal. Como en el caso anterior, se puede apreciar que los modos primero y tercero de CEEMDAN mejorada contienen menos ruido que aquellos obtenidos con los otros tres métodos asistidos por ruido. Para ambas versiones de CEEMDAN detuvimos la descomposición una vez que las condiciones de IMF se hubieron cumplido.

Los ejemplos con señales artificiales mostraron que hemos obtenido un método que provee modos con menos contenido de ruido y que evita la presencia de modos espurios. Aplicaremos a continuación la versión mejorada de CEEMDAN a señales biomédicas, ilustrando su potencial como herramienta de análisis para este tipo de señales.

### 5.6.2. Electroglotograma

El electroglotograma (EGG) es una técnica muy difundidas en nuestros días que permite indagar en el área de contacto de las cuerdas vocales en fonación de una manera simple y no invasiva. Se caracteriza por un aumento en la amplitud de la señal en el instante de cierre glótico [50]. Su estudio permite un seguimiento de la frecuencia fundamental en fonemas sonoros. Por lo tanto, sería muy deseable poseer una señal libre de ruido y tendencia para estimar mejor los instantes de apertura y cierre glótico.

Un segmento de una señal de EGG de la base de datos Keele [90] puede ser apreciado en la Fig. 5.15. Se ha superpuesto el quinto modo de CEEMDAN mejorada. Los parámetros usados fueron  $\varepsilon_0 = 0,2$  y  $I = 100$ . Además, los espectros de Fourier de ambas señales son mostrados, donde puede notarse que la energía presente en la parte de bajas frecuencias de la señal (tendencia) es tan alta como la energía presente en la frecuencia fundamental. La versión mejorada de CEEMDAN captura en un único modo la mayor parte de la información de la señal, preservando características importantes de la forma de onda como una leve asimetría, donde el aumento de

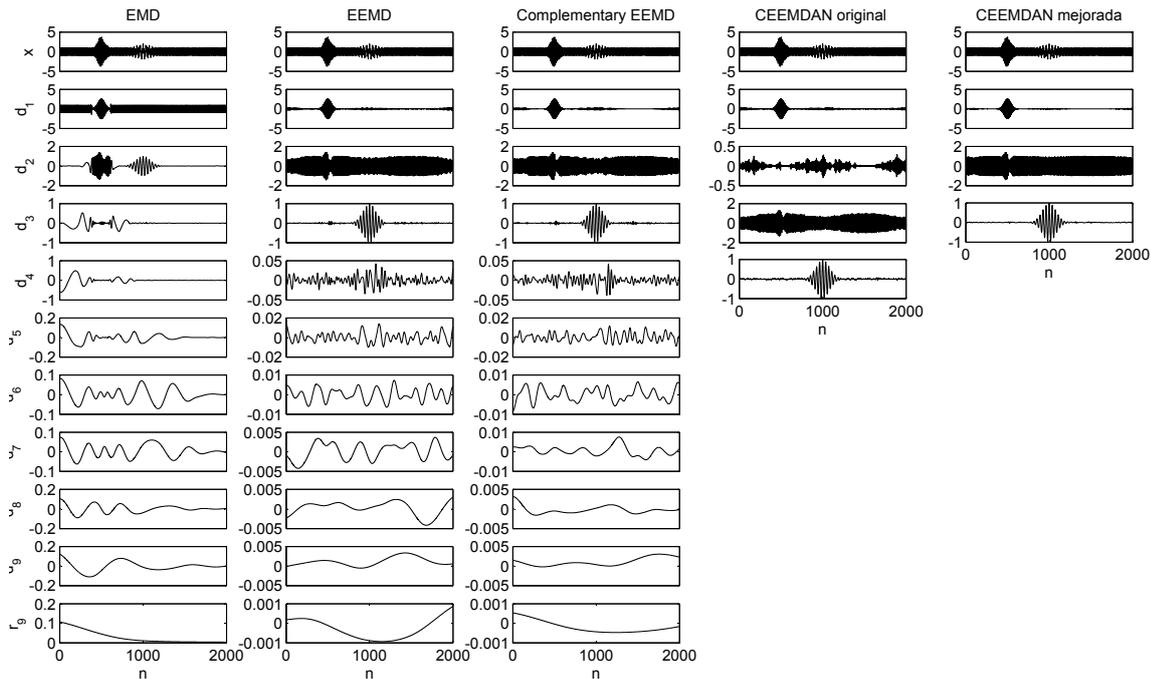


Figura 5.14: Descomposición de la señal artificial  $x$  mediante EMD, EEMD, Complementary EEMD, CEEMDAN original y CEEMDAN mejorada.

amplitud debido al cierre glótico es más rápido que el decremento de amplitud debido a la apertura [50]. Todos los modos de la descomposición se muestran en la Fig. 5.16. Los dos primeros modos contienen principalmente ruido, y del sexto modo en adelante capturan el contenido de bajas frecuencias presente en la tendencia de la señal. En el tercer modo aparece información valiosa de alta frecuencia asociada al cierre de las cuerdas vocales. La misma información, aunque en menor relevancia, puede apreciarse en el cuarto modo.

Una señal de EGG diferente de la misma base de datos que presenta una clara tendencia se muestra en la Fig. 5.17. Como en el caso anterior, hemos superpuesto el quinto modo de CEEMDAN mejorada. Usamos los mismos valores para los parámetros del algoritmo. En este caso, la energía de la frecuencia fundamental es mucho menor que la energía del ruido de baja frecuencia, y así y todo CEEMDAN mejorada captura en un único modo la información más importante de la señal, preservando características de la forma de onda. Más aún, un descenso de la frecuencia de la señal se aprecia en el segmento final del quinto modo. Todos los modos pueden apreciarse en la Fig. 5.18. Mientras el primer modo contiene principalmente ruido, los modos sexto en adelante contienen la tendencia de la señal. El tercer modo contiene información de alta frecuencia relacionada con el cierre glótico. La tendencia a la que hacíamos mención se captura satisfactoriamente en el último modo.

Con estos dos ejemplos ilustramos las capacidades de CEEMDAN mejorada para realizar una tarea de filtrado guiado por los datos de señales de EGG. La información relacionada con la frecuencia fundamental aparece en un único modo. Los modos obtenidos evidencian información de alta frecuencia relacionada con el cierre de cuerdas vocales.

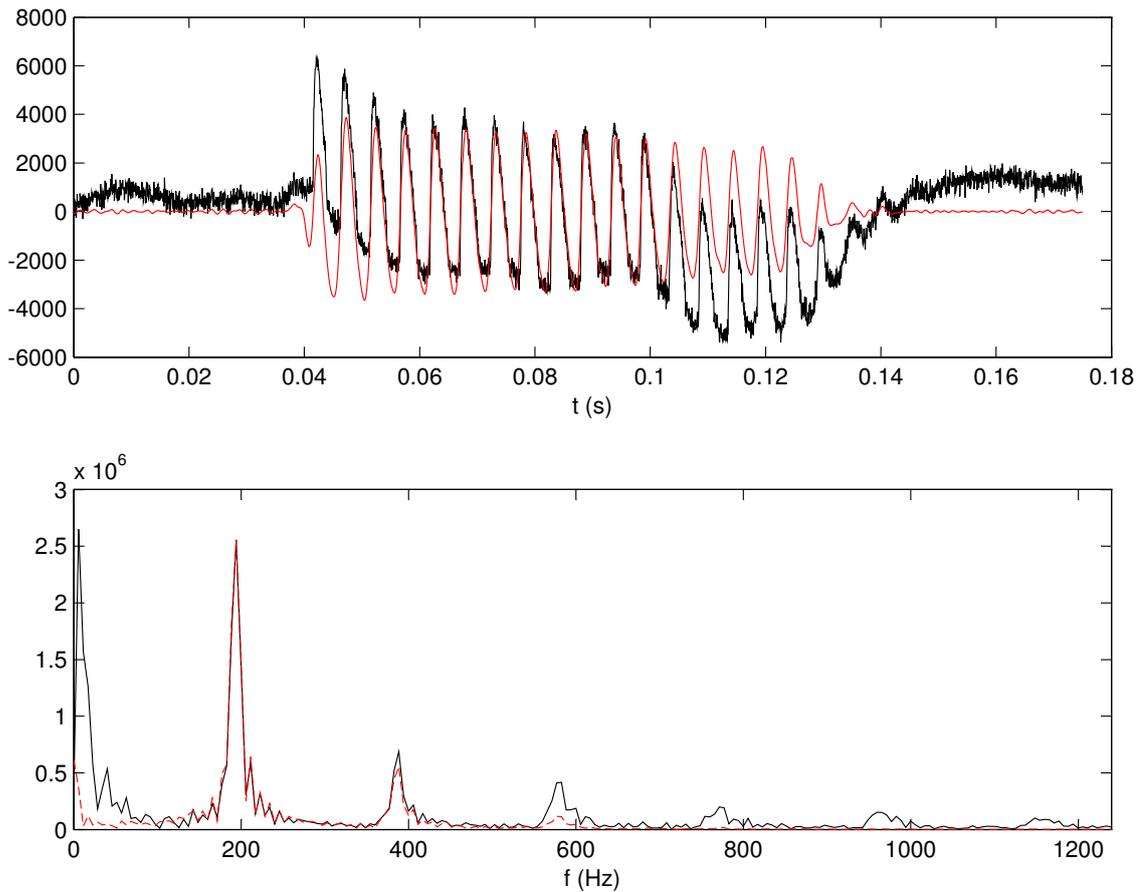


Figura 5.15: Señal de EGG. Arriba: señal de EGG (en negro) y quinto modo de CEEMDAN mejorada (en rojo). Abajo: espectro de amplitud de la señal de EGG (en negro) y del quinto modo de CEEMDAN (en rojo).

### 5.6.3. Electrocardiograma con fibrilación ventricular

En la mayoría de los casos, el mecanismo que dispara un paro cardíaco súbito es una taquicardia ventricular que progresa rápidamente hacia una fibrilación ventricular [10]. Un tercio de esos pacientes pueden sobrevivir con la aplicación a tiempo de un desfibrilador. Entonces, la detección automática de tales eventos analizando las señales de electrocardiograma (ECG) es de una importancia vital [9].

Un segmento de un registro de ECG de la Creighton University (CU) *ventricular tachyarrhythmia database* [46, 80] se muestra en la Fig. 5.19. Puede apreciarse el evento de fibrilación que comienza a la mitad de la duración de la señal. El segundo modo de CEEMDAN mejorada ha sido superpuesto. Debe notarse cómo este modo identifica los latidos cardíacos (complejos QRS) para la primera mitad de la señal, mientras que permanece casi igual a cero para la segunda mitad donde no ocurren latidos. Este modo es un detector de latidos cardíacos y puede usarse para identificar los momentos de fibrilación ventricular, caracterizados justamente por la falta de dichos latidos. Los espectros de la señal de ECG y del segundo modo de CEEMDAN mejorada se muestran en la Fig. 5.19. Aunque el modo contiene principalmente frecuencias altas, su energía se reparte por todo el espectro, y un simple filtrado pasa-altos o pasa-banda de la señal original no podría aislarlo.

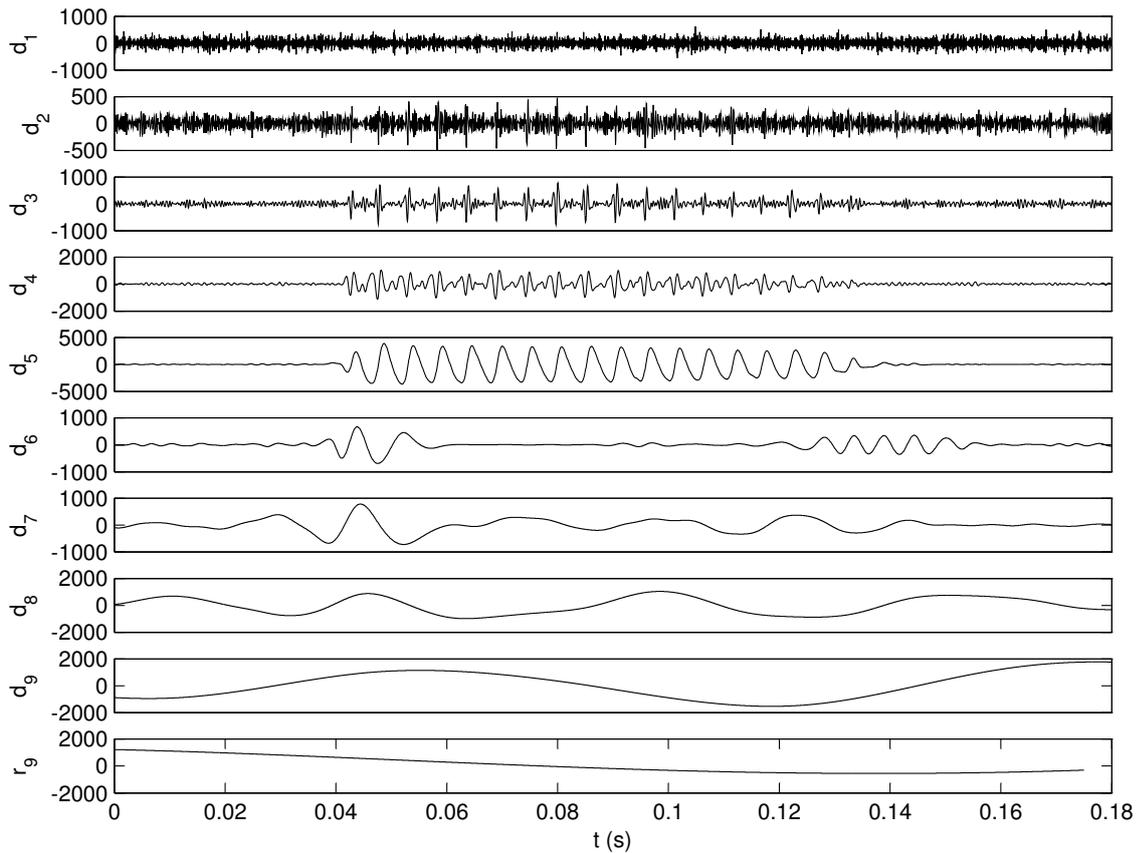


Figura 5.16: Descomposición mediante CEEMDAN mejorada de la señal de ECG mostrada en la Fig. 5.15.  $\epsilon_0 = 0,2$ .  $I = 100$ .

Todos los modos de la descomposición se presentan en la Fig. 5.20. El cuarto modo contiene la mayor parte de la información del evento de fibrilación. La primera mitad del séptimo modo captura casi perfectamente la frecuencia fundamental de latidos, mientras que su segunda mitad evidencia la pérdida de ritmo sinusal.

#### 5.6.4. Electroencefalograma con crisis epiléptica

La electroencefalografía (EEG) es un método de evaluación que contiene información muy útil relacionada con los diferentes estados fisiológicos del cerebro [86]. Las crisis súbitas, como los ataques epilépticos, se reflejan en los registros de EEG, y es por ello que pueden ser usados para una detección automática.

Estudiamos aquí un registro fisiológico real usando electroencefalografía estéreo con ocho electrodos multiderivacionales (2 mm de largo y separación de 1,5 mm). Se amplificó y filtró en la banda de 1 a 40 Hz. Como filtro *antialiasing* se usó un filtro pasa-bajos Butterworth de cuarto orden. La señal fue digitalizada a 256 Hz con un conversor analógico-digital de 10 bits. Un médico profesional realizó el análisis de regiones preictales e ictales mediante inspección visual del registro de EEG. De acuerdo a tal inspección, el paciente presentó una área epileptogénica en el hipocampo con propagación inmediata al *girus cingular* y al área motora complementaria, en el hemisferio izquierdo.

El registro de EEG intracraneal se muestra en la Fig. 5.21, donde se indican los

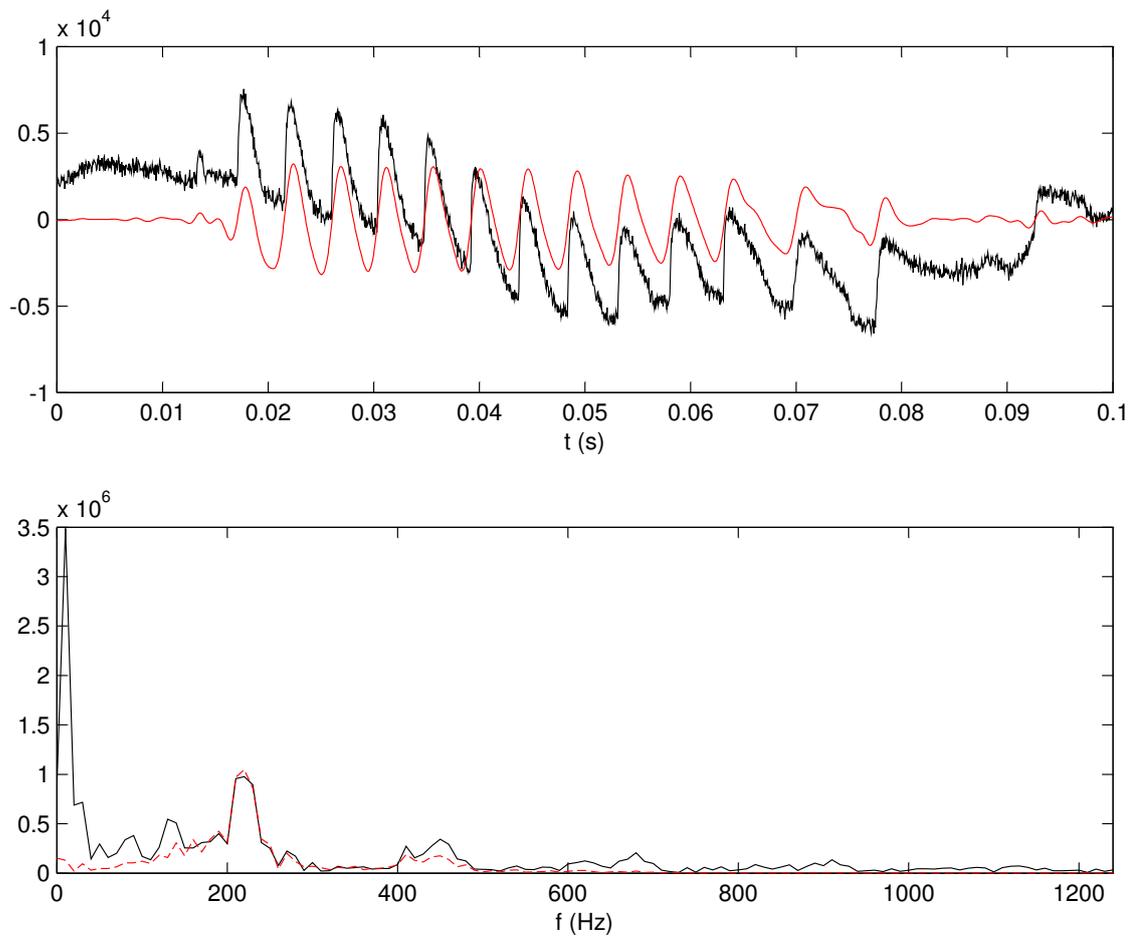


Figura 5.17: Señal de EGG. Arriba: señal de EGG (en negro) y quinto modo de CEEMDAN mejorada (en rojo). Abajo: espectro de amplitud de la señal de EGG (en negro) y del quinto modo de CEEMDAN (en rojo).

momentos de comienzo y fin de tres eventos de crisis epilépticas marcados con líneas verticales. El primer modo de CEEMDAN mejorada ha sido superpuesto en rojo. Este modo parece identificar las crisis, y su análisis puede llevar a una detección automática. A partir de los espectros, se puede apreciar que el modo no puede ser obtenido a partir de un simple filtrado pasa-altos. En la Fig. 5.22 se muestran todos los modos de la descomposición.

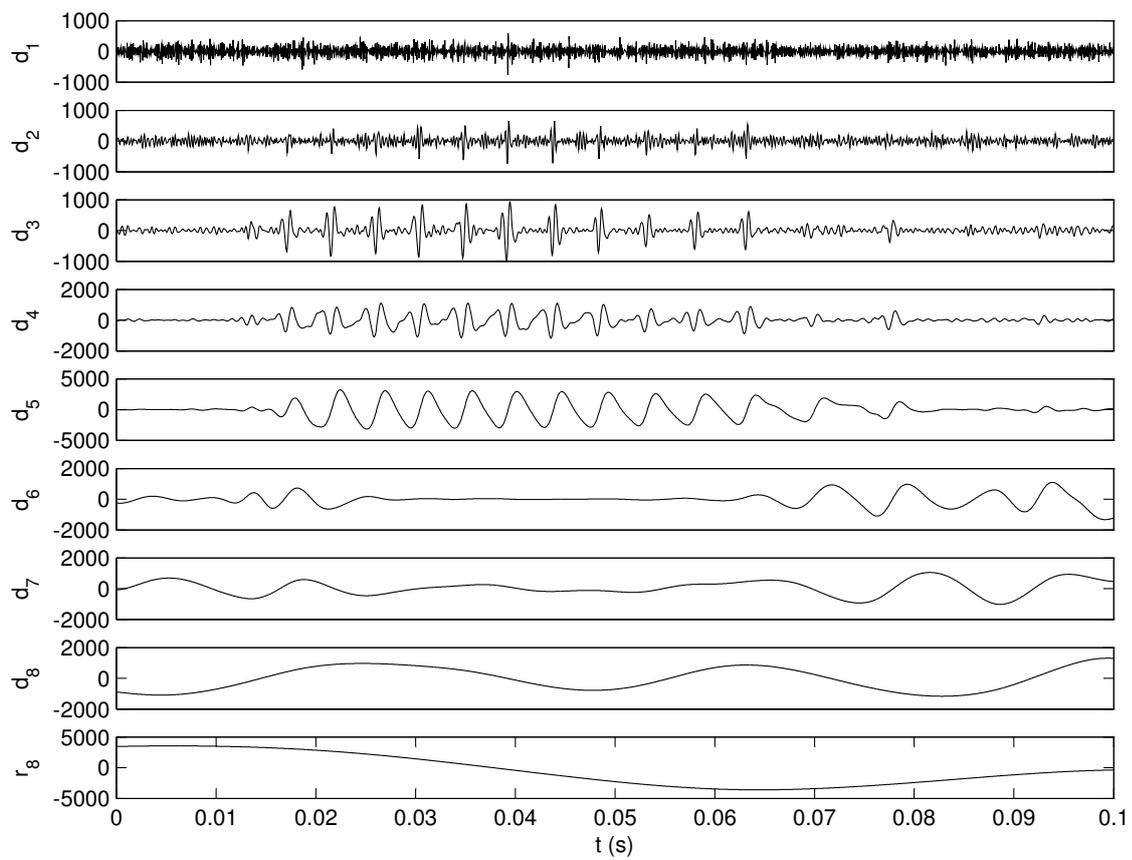


Figura 5.18: Descomposición mediante CEEMDAN mejorada de la señal de EGG mostrada en la 5.17.  $\epsilon_0 = 0,2$ .  $I = 100$ .

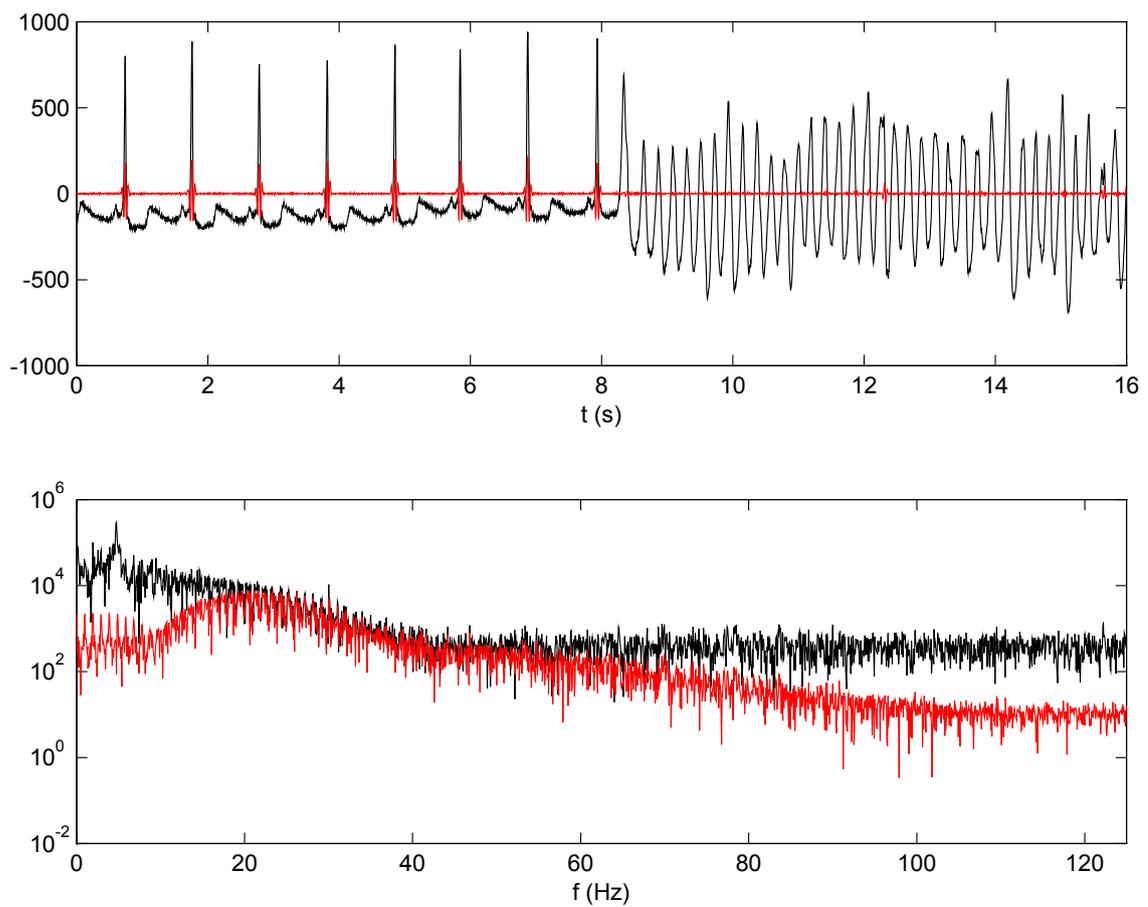


Figura 5.19: Señal de ECG. Arriba: señal de ECG (en negro) y segundo modo de CEEMDAN mejorada (en rojo). Abajo: espectro de amplitud de la señal de ECG (en negro) y del segundo modo de CEEMDAN (en rojo).

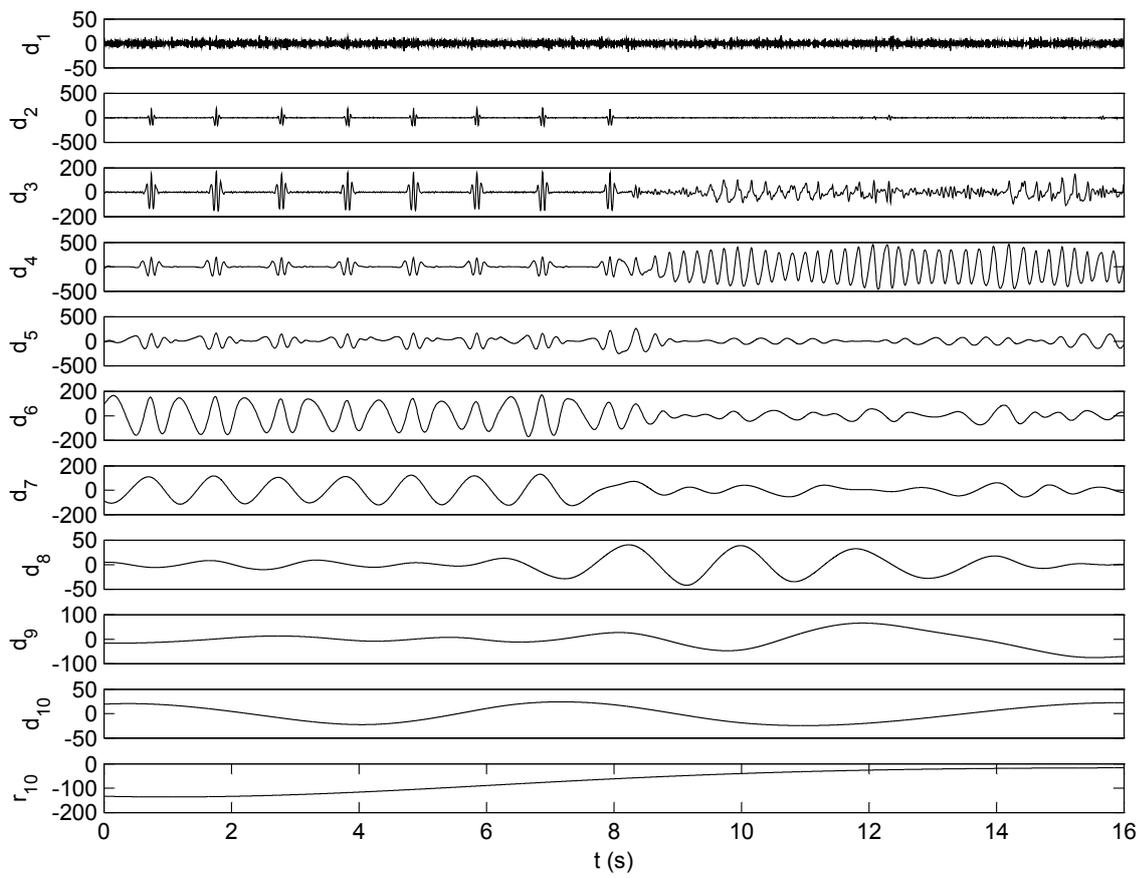


Figura 5.20: Descomposición mediante CEEMDAN mejorada de la señal de ECG mostrada en la 5.19.  $\epsilon_0 = 0,2$ .  $I = 100$ .

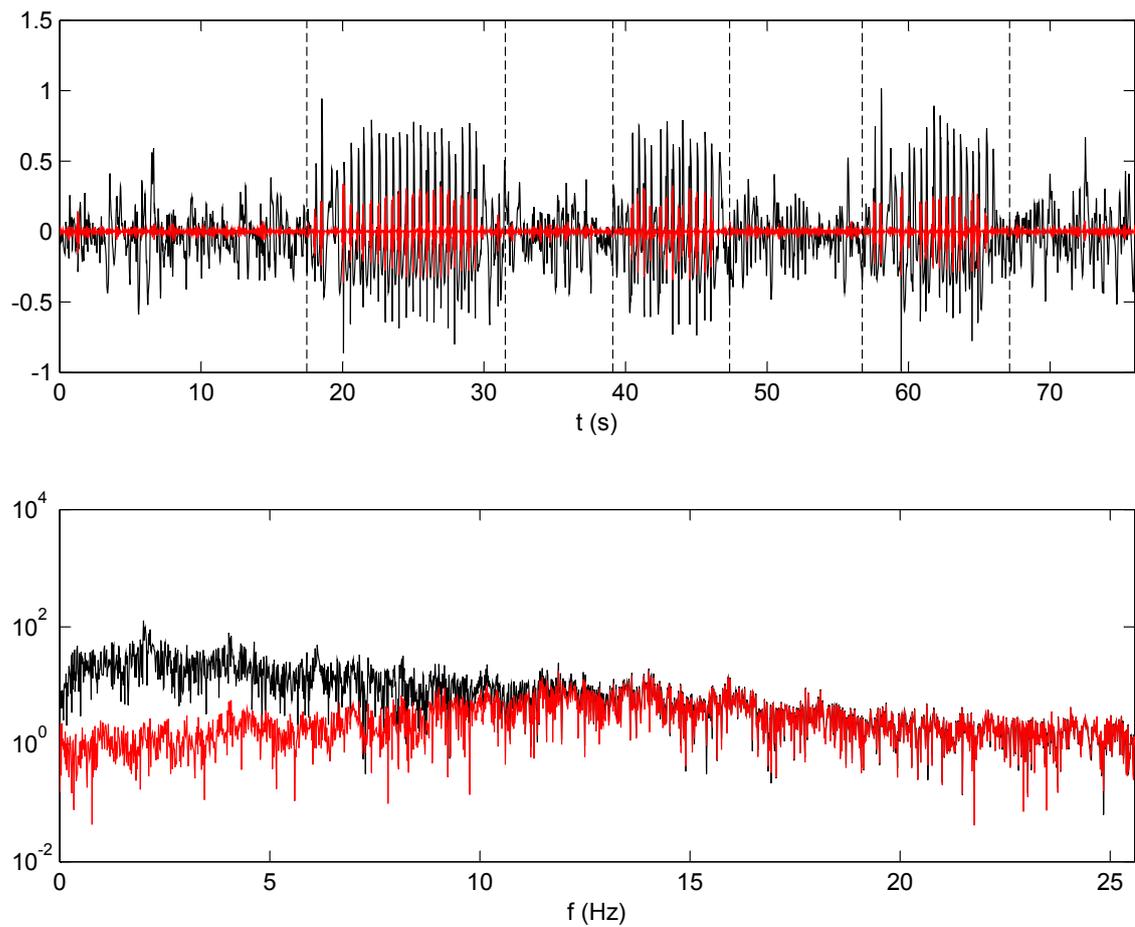


Figura 5.21: Señal de EEG. Arriba: señal de EEG (en negro) y primer modo de CEEMDAN mejorada (en rojo). Abajo: espectro de amplitud de la señal de EEG (en negro) y del primer modo de CEEMDAN (en rojo).

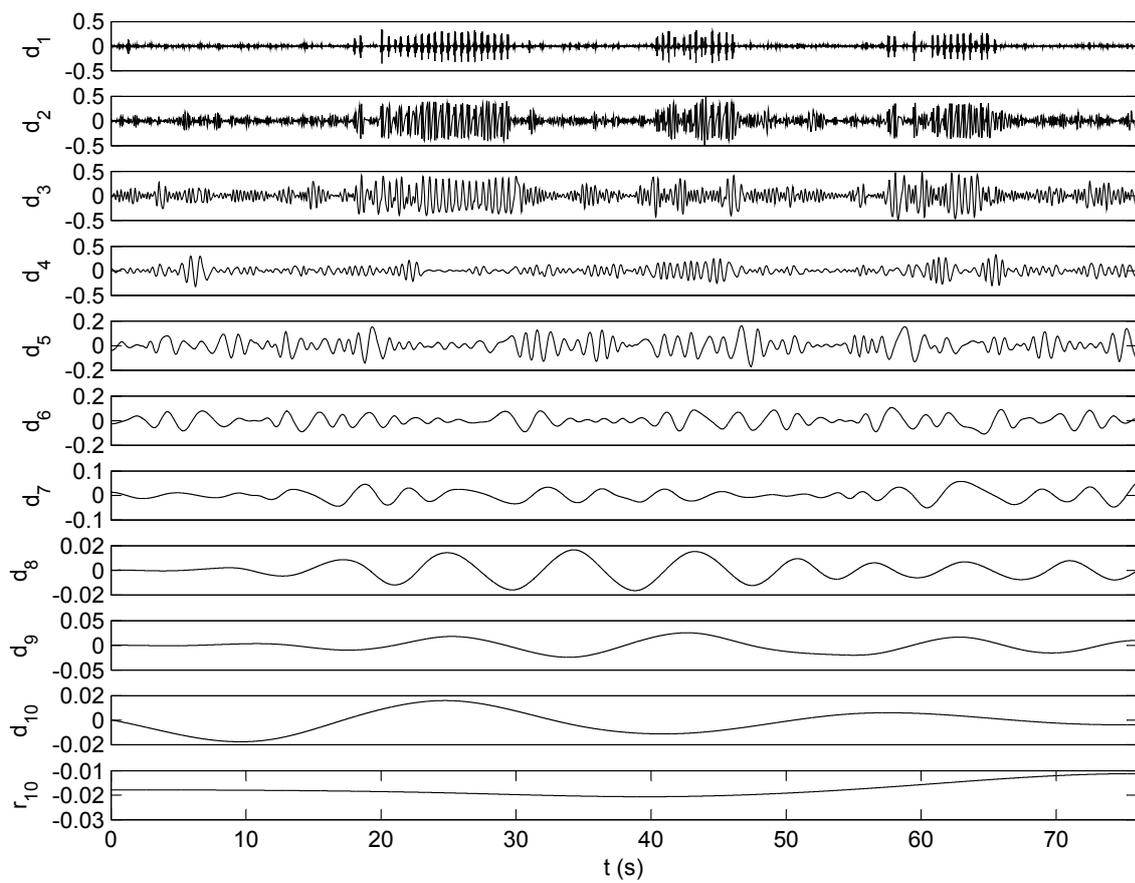


Figura 5.22: Descomposición mediante CEEMDAN mejorada de la señal de EEG mostrada en la 5.21.  $\epsilon_0 = 0,2$ .  $I = 100$ .

## 5.7. Comentarios finales

La mezcla de modos dificultó la aplicación directa de EMD a un conjunto importante de señales reales. Sin embargo, las variaciones asistidas por ruido, entre ellas la EEMD, aliviaron este problema y se difundieron rápidamente. En este capítulo analizamos las ideas fundamentales de las versiones de EMD asistidas por ruido, describiendo sus ventajas y dificultades. Además, propusimos un nuevo método que mantiene tales ventajas y supera tales dificultades: la CEEMDAN.

Debemos mencionar, no obstante, que en general el agregado de ruido define una referencia *global* para el método de EMD. Entonces, las variaciones asistidas por ruido son menos locales y menos guiadas por los datos (porque además están guiadas por el tipo de ruido y su amplitud) que la versión original libre de ruido.

En este capítulo, propusimos un nuevo método, evaluando su desempeño tanto en señales artificiales como en reales. Comparamos nuestros resultados con dos métodos asistidos por ruido del estado del arte. Evidenciamos ventajas a favor de nuestra propuesta. Además realizamos un estudio sobre la amplitud de ruido óptima que debe agregarse en los algoritmos, obteniendo un valor cercano a aquél recomendado por la práctica común. La aplicación de nuestro método a señales biológicas permitió obtener información relevante relacionada a procesos fisiológicos involucrados en su producción.

Los desarrollos presentados en este capítulo dieron lugar a dos artículos de *journals* internacionales: uno en una publicación relativamente reciente y especializada en métodos guiados por los datos *Advances in Adaptive Data Analysis* [28]; y otro en *Biomedical Signal Processing and Control* de Elsevier [27]. Además, una publicación nuestra en el Congreso Latinoamericano de Ingeniería Biomédica [3] logró una mención especial en la competición de investigador joven.



# Capítulo 6

## Conclusiones y perspectivas

*Si el autor de estas líneas ha conseguido dar algún material para poner en claro estas cuestiones, puede considerar que su trabajo no ha sido vano.*

---

EL DESARROLLO DEL CAPITALISMO EN RUSIA  
Vladimir I. Lenin

En esta tesis estudiamos el análisis de señales multicomponentes prestando especial atención a la técnica de descomposición empírica en modos, proponiendo contribuciones tanto desde el punto de vista teórico (reformulando el algoritmo original) como algorítmico (mejorando las versiones asistidas por ruido).

Comenzamos repasando los conceptos de señal multicomponente y frecuencia instantánea, presentes en numerosos modelos de fenómenos físicos en los que pueden identificarse un número relativamente pequeño de componentes cuya frecuencia es única para un instante de tiempo dado. Luego presentamos la transformada de Hilbert, que permite la estimación de la frecuencia instantánea de una señal AM-FM. Señalamos, sin embargo, las limitaciones que para ello imponen los teoremas de Bedrosian y Nutall.

Continuamos nuestra exposición con los enfoques más naturales para analizar señales multicomponentes: las representaciones lineales tiempo-frecuencia y tiempo-escala. El enfoque tradicional suele ser buscar las crestas que cada componente genera en las representaciones, estimando así las frecuencias instantáneas alrededor de las cuales éstas se disponen. La extracción de los componentes en el dominio temporal se realiza antitransformando la representación elegida en dominios acotados que contengan a las crestas.

A continuación describimos de una manera exhaustiva a la técnica de descomposición empírica en modos, motivación fundamental de este trabajo de tesis, tanto en sus aspectos teóricos como prácticos. Comenzamos describiendo cómo esta técnica representa componentes más complejos, basándose sólo en la simetría de sus envolventes, para luego describir el esquema recursivo que aplica. En la descripción de las funciones intrínsecas de modo establecimos las relaciones existentes con los trabajos recientes de Daubechies y Hau-Tieng Wu. Además, propusimos nuestra propia definición de las IMFs, conforme a nuestros criterios. Describimos también cómo un procesamiento posterior mediante transformada de Hilbert permite obtener una representación tiempo-frecuencia.

En la descripción de EMD quedaba claro su naturaleza empírica, en la medida que está definido como la salida de un algoritmo. Entonces, luego de ello presentamos la primera contribución original de esta tesis: la formulación de EMD como un problema de optimización sin restricciones. Este enfoque forma parte de una corriente orientada a dotar a EMD de las bases teóricas que carece tomando la teoría de optimización como punto de partida. Nuestra propuesta permite analizar *a priori* el comportamiento de la descomposición. La solución que construye puede compararse con la que se obtiene mediante el método de *B-spline* EMD, pero con las funciones «básicas» y los coeficientes completamente determinados por los datos. Una interpretación de filtro variante en el tiempo, revela que nuestro método construye una media local de manera discreta y local, cuando el caso ideal es continuo e instantáneo. Extensas simulaciones con señales artificiales, comparando nuestros resultados con aquellos obtenidos con métodos del estado del arte de EMD basada en optimización, sustentan que nuestra propuesta posee ventajas tanto en desempeño como en tiempo de cómputo. Su aplicación a una señal real de voz ilustra la potencialidad de nuestra propuesta para el análisis de este tipo de señales. A continuación extendimos nuestro método para su aplicación a imágenes. En este caso, también lo comparamos con métodos del estado del arte, mostrando ventajas en nuestra propuesta.

La cuestión más importante que queda planteada respecto a este tema es la posibilidad de profundizar el estudio de la influencia del proceso de *sifting*, que puede generar nuevos extremos y modificar los operadores presentes en nuestro problema de optimización (más específicamente la matriz  $P_x$ ). Como mencionamos, este proceso puede verse como una mejora en los puntos característicos que definen la escala local.

A continuación avanzamos sobre el mismo objetivo de dotar a EMD de bases teóricas más sólidas. Entonces exploramos la que quizá sea la otra gran tendencia dentro de esta tarea: el uso de representaciones tiempo-escala. Planteamos un método de descomposición de señales que fusiona las ideas de *synchrosqueezing* con las de EMD, constituyendo la segunda contribución original de esta tesis. La localidad y naturaleza guiada por los datos de esta última técnica, que se basan en los extremos locales de la señal como puntos característicos para definir una escala local, se encontraron con las sólidas bases teóricas detrás de las onditas y *synchrosqueezing*. El resultado es un método tan guiado por los datos como EMD, evitando las *splines* y cualquier tipo de interpolación en el dominio temporal. Creemos que esta idea de realizar una descomposición EMD en un marco tiempo-escala debería ayudar a comprender mejor la forma en la que trabaja EMD.

Esta línea, que consideramos la más promisoría, debe continuarse de inmediato. Los primeros problemas que se deben analizar son aquellos que se usaron para el estudio pionero de EMD: separación de tonos puros, y descomposición de ruido blanco. Además es necesario evaluar este método en conjunción con los desarrollos más recientes sobre las representaciones tiempo-escala invertibles: *reassignment* invertible y *synchrosqueezing* de segundo orden [85]; y el *synchrosqueezing* recursivo de Levenberg-Marquardt [42].

Finalmente, discutimos las versiones asistidas por ruido de EMD. Estas variaciones de EMD permitieron su plena aplicación a señales reales, lo que se hallaba limitado por la mezcla de modos, y pueden encontrarse ejemplos de ello en

---

los campos más diversos. Sin embargo, estas variaciones (y principalmente la más extendida de todas: la EEMD) poseían dificultades que necesitaban ser superadas. Debido a ello, y como tercera contribución original de esta tesis, propusimos un nuevo método: la CEEMDAN.

Evaluamos su desempeño tanto en señales artificiales como en reales. Comparamos nuestros resultados con dos métodos asistidos por ruido del estado del arte, evidenciándose ventajas a favor de nuestra propuesta. Además realizamos un estudio sobre la amplitud de ruido óptima que debe agregarse en los algoritmos, obteniendo un valor cercano a aquél recomendado por la práctica común. Al aplicar nuestro método a señales biológicas pudimos obtener información relevante relacionada a procesos fisiológicos involucrados en su producción.

Es necesario remarcar aquí que el agregado de ruido define una referencia *global* para el método de EMD. Entonces, las variaciones asistidas por ruido son menos locales y menos guiadas por los datos (porque además están guiadas por el tipo de ruido y su amplitud) que la versión original libre de ruido. Es necesario continuar el estudio de estas variaciones. Creemos importante poder relacionar de manera más estrecha los desarrollos en los capítulos 3 y 4 con el agregado de ruido a las señales. Esto puede hacerse, en el caso basado en optimización, estudiando cómo se modifica la matriz  $P_x$  cuando se agrega ruido a la señal. Será necesario estimar un valor esperado para la matriz, de forma tal de obtener una media local esperada que emule a la media de las versiones asistidas por ruido. En el caso basado en representaciones tiempo-escala será necesario estudiar cómo el agregado de ruido afecta a la escala local definida mediante los extremos locales. Deberá estimarse entonces un valor esperado para la escala local, que resultará de la conjunción entre la señal analizada, y la amplitud y tipo de ruido agregado.

Como conclusión general, consideramos que las contribuciones desarrolladas en esta tesis tienen la posibilidad de permitir una mejor comprensión de la técnica de EMD, dentro del campo del análisis de señales multicomponentes. Esta mejor comprensión, junto al desarrollo de nuevos y más robustos algoritmos, como los asistidos por ruido, deberían ayudar a adoptar esta técnica en un número mayor de aplicaciones.



# Lista de publicaciones

- [1] M. A. Colominas, A. Humeau-Heurtier, and G. Schlotthauer. Orientation-independent empirical mode decomposition for images based on unconstrained optimization. *IEEE Transactions on Image Processing*, 25:2288–2297, 2016.
- [2] M. A. Colominas and G. Schlotthauer. Empirical mode decomposition in a time-scale framework. *enviado*, 2016.
- [3] M. A. Colominas, G. Schlotthauer, and M. Torres. Complete ensemble EMD and Hilbert transform for heart beat detection. In *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014*, pages 496–499. Springer, 2015.
- [4] M. A. Colominas, G. Schlotthauer, and M. E. Torres. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 14:19–29, 2014.
- [5] M. A. Colominas, G. Schlotthauer, and M. E. Torres. An unconstrained optimization approach to empirical mode decomposition. *Digital Signal Processing*, 40:164–175, 2015.
- [6] M. A. Colominas, G. Schlotthauer, M. E. Torres, and P. Flandrin. Noise-assisted EMD methods in action. *Advances in Adaptive Data Analysis*, 4(04):1250025, 2012.



# Referencias

- [7] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä. Recognition of blurred faces using local phase quantization. In *Proc. 19th Int. Conf. Pattern Recognit. (ICPR 2008)*, pages 1–4. IEEE, 2008.
- [8] S. M. Alessio. *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*. Springer, 2016.
- [9] A. Amann, R. Tratnig, and K. Unterkofler. Detecting ventricular fibrillation by time-delay methods. *IEEE Trans. on Biomed. Eng.*, 54(1):174–177, 2007.
- [10] A. Amann, R. Tratnig, K. Unterkofler, et al. Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators. *BioMedical Engineering OnLine*, 4(60), 2005.
- [11] R. C. Aster, B. Borchers, and C. H. Thurber. *Parameter estimation and inverse problems*. Academic Press, 2011.
- [12] F. Auger and P. Flandrin. Improving the readability of time-frequency and time-scale representations by the reassignment method. *Signal Processing, IEEE Transactions on*, 43(5):1068–1089, 1995.
- [13] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu. Time-frequency reassignment and synchrosqueezing: An overview. *IEEE Signal Processing Magazine*, 30(6):32–41, 2013.
- [14] E. Bedrosian. A product theorem for Hilbert transforms. *Proceedings of the IEEE*, 51:868–869, 1963.
- [15] R. P. Brent. *Algorithms for minimization without derivatives*. Courier Corporation, 2013.
- [16] E. Brevdo, N. S. Fuckar, G. Thakur, and H.-T. Wu. The synchrosqueezing algorithm: a robust analysis tool for signals with time-varying spectrum. *Arxiv preprint*, 2011.
- [17] E. Brevdo and H. Wu. The synchrosqueezing toolbox. 2011.
- [18] A. M. Bruckner, J. B. Bruckner, and B. S. Thomson. *Real analysis*. Prentice-Hall, 1997.
- [19] J. Bruna and S. Mallat. Classification with scattering operators. *arXiv preprint arXiv:1011.3023*, 2010.

- [20] R. Carmona, W. L. Hwang, B. Torr sani, et al. Characterization of signals by the ridges of their wavelet transforms. *Signal Processing, IEEE Transactions on*, 45(10):2586–2590, 1997.
- [21] R. Carmona, W. L. Hwang, B. Torr sani, et al. Multiridge detection and time-frequency reconstruction. *Signal processing, IEEE Transactions on*, 47(2):480–492, 1999.
- [22] K.-M. Chang. Arrhythmia ECG noise reduction by ensemble empirical mode decomposition. *Sensors*, 10(6):6063–6080, 2010.
- [23] Q. Chen, N. Huang, S. Riemenschneider, and Y. Xu. A B-spline approach for empirical mode decompositions. *Advances in Computational Mathematics*, 24(1-4):171–195, 2006.
- [24] J. Cheng, D. Yu, and Y. Yang. Application of support vector regression machines to the processing of end effects of Hilbert–Huang transform. *Mechanical Systems and Signal Processing*, 21(3):1197–1211, 2007.
- [25] L. Cohen. *Time-frequency analysis*, volume 1. Prentice hall, 1995.
- [26] M. A. Colominas. Desarrollo de algoritmos guiados por los datos para el an lisis de se ales. Proyecto Final, Fac. de Ingenier a, Universidad Nacional de Entre R os, 2011.
- [27] M. A. Colominas, G. Schlotthauer, and M. E. Torres. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control*, 14:19–29, 2014.
- [28] M. A. Colominas, G. Schlotthauer, M. E. Torres, and P. Flandrin. Noise-assisted EMD methods in action. *Advances in Adaptive Data Analysis*, 4(04):1250025, 2012.
- [29] M. D tig and T. Schlurmann. Performance and limitations of the Hilbert–Huang transformation (HHT) with an application to irregular water waves. *Ocean Engineering*, 31(14):1783–1834, 2004.
- [30] I. Daubechies. *Ten lectures on wavelets*, volume 61. SIAM, 1992.
- [31] I. Daubechies, J. Lu, and H.-T. Wu. Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Applied and computational harmonic analysis*, 30(2):243–261, 2011.
- [32] I. Daubechies and S. Maes. A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models. *Wavelets in medicine and biology*, pages 527–546, 1996.
- [33] R. Deering and J. F. Kaiser. The use of a masking signal to improve empirical mode decomposition. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on*, volume 4, pages iv–485. IEEE, 2005.

- [34] E. Deléclhelle, J. Lemoine, and O. Niang. Empirical mode decomposition: an analytical approach for sifting process. *Signal Processing Letters, IEEE*, 12(11):764–767, 2005.
- [35] Y. Deng, W. Wang, C. Qian, Z. Wang, and D. Dai. Boundary-processing-technique in EMD method and Hilbert transform. *Chinese Science Bulletin*, 46(11):954–960, 2001.
- [36] M. Farge. Wavelet transforms and their applications to turbulence. *Annual review of fluid mechanics*, 24(1):395–458, 1992.
- [37] P. Flandrin. *Time-frequency/time-scale analysis*, volume 10. Academic press, 1999.
- [38] P. Flandrin and P. Goncalves. Empirical mode decompositions as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(04):477–496, 2004.
- [39] P. Flandrin, P. Gonçaves, and G. Rilling. EMD equivalent filter banks, from interpretation to applications. *Hilbert-Huang transform and its applications*, pages 57–74, 2005.
- [40] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *Signal Processing Letters, IEEE*, 11(2):112–114, 2004.
- [41] B. Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of computation*, 51(184):699–706, 1988.
- [42] D. Fourer, F. Auger, and P. Flandrin. Recursive versions of the Levenberg-Marquardt reassigned spectrogram and of the synchrosqueezed STFT. In *IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-16, Shanghai (PRC)*., page to appear. IEEE, 2016.
- [43] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Reviews of modern physics*, 70(1):223, 1998.
- [44] J. E. Gentle. *Matrix algebra: theory, computations, and applications in statistics*. Springer Science & Business Media, 2010.
- [45] J. Gilles and S. Osher. Bregman implementation of Meyer’s G-norm for cartoon+ textures decomposition. *UCLA Cam Report*, pages 11–73, 2011.
- [46] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.
- [47] Z. Guo, L. Zhang, and D. Zhang. Rotation invariant texture classification using LBP variance (LBPV) with global matching. *Pattern recognit.*, 43(3):706–719, 2010.

- [48] P. C. Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, volume 4. Siam, 1998.
- [49] S. D. Hawley, L. E. Atlas, and H. J. Chizeck. Some properties of an empirical mode type signal decomposition algorithm. *IEEE Signal Processing Letters*, 1(17):24–27, 2010.
- [50] N. Henrich, C. d’Alessandro, B. Doval, and M. Castellengo. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J. of the Acoust. Soc. of Am.*, 115(3):1321–1332, 2004.
- [51] H. Hong, X. Wang, and Z. Tao. Local integral mean-based sifting for empirical mode decomposition. *Signal Processing Letters, IEEE*, 16(10):841–844, 2009.
- [52] X. Hu, S. Peng, and W.-L. Hwang. EMD revisited: A new understanding of the envelope and resolving the mode-mixing problem in AM-FM signals. *Signal Processing, IEEE Transactions on*, 60(3):1075–1086, 2012.
- [53] X. Hu, S. Peng, and W.-L. Hwang. Multicomponent AM-FM signal separation and demodulation with null space pursuit. *Signal, Image and Video Processing*, 7(6):1093–1102, 2013.
- [54] B. Huang and A. Kunoth. An optimization based empirical mode decomposition scheme. *Journal of Computational and Applied Mathematics*, 240:174–183, 2013.
- [55] N. E. Huang et al. Introduction to the Hilbert-Huang transform and its related mathematical problems. *Hilbert-Huang transform and its applications, interdisciplinary mathematical sciences*, 5:1–24, 2005.
- [56] N. E. Huang and S. S. Shen. *Hilbert-Huang transform and its applications*, volume 5. World Scientific, 2005.
- [57] N. E. Huang, Z. Shen, and S. R. Long. A new view of nonlinear water waves: the Hilbert spectrum. *Annual review of fluid mechanics*, 31(1):417–457, 1999.
- [58] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 454, pages 903–995. The Royal Society, 1998.
- [59] N. E. Huang, M.-L. C. Wu, S. R. Long, S. S. Shen, W. Qu, P. Gloersen, and K. L. Fan. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 459, pages 2317–2345. The Royal Society, 2003.
- [60] N. E. Huang and Z. Wu. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics*, 46(2), 2008.

- [61] D. Iatsenko. *Nonlinear Mode Decomposition: Theory and Applications*. Springer, 2015.
- [62] D. Iatsenko, P. V. McClintock, and A. Stefanovska. Linear and synchrosqueezed time-frequency representations revisited. *Digital Signal Processing*, 42(C):1–26, 2015.
- [63] D. Iatsenko, P. V. McClintock, and A. Stefanovska. Nonlinear mode decomposition: A noise-robust, adaptive decomposition method. *Physical Review E*, 92(3):032916, 2015.
- [64] G. Jager, R. Koch, A. Kunoth, and R. Pabel. Fast empirical mode decompositions of multivariate data based on adaptive spline-wavelets and a generalization of the Hilbert–Huang–transformation (HHT) to arbitrary space dimensions. *Advances in Adaptive Data Analysis*, 2(03):337–358, 2010.
- [65] C. Junsheng, Y. Dejie, and Y. Yu. The application of energy operator demodulation approach based on EMD in machinery fault diagnosis. *Mechanical systems and signal processing*, 21(2):668–677, 2007.
- [66] B. Kedem. Spectral analysis and discrimination by zero-crossings. *Proceedings of the IEEE*, 74(11):1477–1493, 1986.
- [67] G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval. Past, present and future of nonlinear system identification in structural dynamics. *Mechanical systems and signal processing*, 20(3):505–592, 2006.
- [68] K. Kodera, C. De Villedary, and R. Gendrin. A new method for the numerical analysis of non-stationary signals. *Physics of the Earth and Planetary Interiors*, 12(2):142–150, 1976.
- [69] K. Kodera, R. Gendrin, and C. Villedary. Analysis of time-varying signals with small BT values. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 26(1):64–76, 1978.
- [70] Y. Lei, Z. He, and Y. Zi. Application of the EEMD method to rotor fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 23(4):1327–1338, 2009.
- [71] Y. Lei, J. Lin, Z. He, and M. J. Zuo. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 35(1):108–126, 2013.
- [72] Z. Lei, T. Ahonen, M. Pietikäinen, and S. Z. Li. Local frequency descriptor for low-resolution face recognition. In *Proc. IEEE Int. Conf. Autom. Face & Gesture Recognit. and Workshops (FG 2011)*, pages 161–166. IEEE, 2011.
- [73] L. Lin, Y. Wang, and H. Zhou. Iterative filtering as an alternative algorithm for empirical mode decomposition. *Advances in Adaptive Data Analysis*, 1(04):543–560, 2009.

- [74] A. Linderhed. Image empirical mode decomposition: A new tool for image processing. *Adv. Adapt. Data Anal.*, 1(02):265–294, 2009.
- [75] R. Maani, S. Kalra, and Y.-H. Yang. Rotation invariant local frequency descriptors for texture classification. *IEEE Trans. Image Process.*, 22(6):2409–2419, 2013.
- [76] S. Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, 2009.
- [77] R. N. Meeson Jr. HHT sifting and filtering. *Hilbert-Huang transform and its applications*, pages 75–105, 2005.
- [78] S. Meignen, T. Oberlin, and S. McLaughlin. A new algorithm for multicomponent signals analysis based on synchrosqueezing: With an application to signal sampling and denoising. *Signal processing, iee transactions on*, 60(11):5787–5798, 2012.
- [79] S. Meignen and V. Perrier. A new formulation for empirical mode decomposition based on constrained optimization. *Signal Processing Letters, IEEE*, 14(12):932–935, 2007.
- [80] F. Nolle, F. Badura, J. Catlett, R. Bowser, and M. H. Sketch. CREI-GARD, a new concept in computerized arrhythmia monitoring systems. *Computers in Cardiology*.
- [81] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel. Image analysis by bidimensional empirical mode decomposition. *Image and Vis. Comput.*, 21(12):1019–1026, 2003.
- [82] A. Nuttall. On the quadrature approximation to the Hilbert transform of modulated signals. *Proceedings of the IEEE*, 54:1458–1459, 1966.
- [83] T. Oberlin, S. Meignen, and V. Perrier. An alternative formulation for the empirical mode decomposition. *IEEE Transactions on Signal Processing*, 60(5):2236–2246, may 2012.
- [84] T. Oberlin, S. Meignen, and V. Perrier. On the mode synthesis in the synchrosqueezing method. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 1865–1869. IEEE, 2012.
- [85] T. Oberlin, S. Meignen, and V. Perrier. Second-order synchrosqueezing transform or invertible reassignment? Towards ideal time-frequency representations. *Signal Processing, IEEE Transactions on*, 63(5):1335–1344, 2015.
- [86] H. Ocak. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Syst. with Appl.*, 36(2):2027–2036, 2009.
- [87] S. Peng and W.-L. Hwang. Adaptive signal decomposition based on local narrow band signals. *Signal Processing, IEEE Transactions on*, 56(7):2669–2676, 2008.

- [88] S. Peng and W.-L. Hwang. Null space pursuit: An operator-based approach to adaptive signal separation. *Signal Processing, IEEE Transactions on*, 58(5):2475–2483, 2010.
- [89] Z. Peng, W. T. Peter, and F. Chu. An improved Hilbert–Huang transform and its application in vibration signal analysis. *Journal of sound and vibration*, 286(1):187–205, 2005.
- [90] F. Plante, G. F. Meyer, and W. A. Ainsworth. A pitch extraction reference database. In *4th Eur. Conf. on Speech Communic. and Technol.*, pages 837–840, Madrid, Spain, 1995.
- [91] M. B. Priestley. *Spectral analysis and time series*. Academic Press, 1981.
- [92] N. Pustelnik, P. Borgnat, and P. Flandrin. A multicomponent proximal algorithm for empirical mode decomposition. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 1880–1884. IEEE, 2012.
- [93] N. Pustelnik, P. Borgnat, and P. Flandrin. Empirical mode decomposition revisited by multicomponent non-smooth convex optimization. *Signal Processing*, 102:313–331, 2014.
- [94] K. Qi, Z. He, and Y. Zi. Cosine window-based boundary processing method for EMD and its application in rubbing fault diagnosis. *Mechanical Systems and Signal Processing*, 21(7):2750–2760, 2007.
- [95] G. Rilling. *Décompositions modales empiriques-contributions à la théorie, l’algorithmie et l’analyse de performances*. PhD thesis, École Normale Supérieure de Lyon, France, 2007.
- [96] G. Rilling and P. Flandrin. On the influence of sampling on the empirical mode decomposition. In *ICASSP (3)*, pages 444–447, 2006.
- [97] G. Rilling and P. Flandrin. One or two frequencies? The Empirical Mode Decomposition answers. *IEEE Transactions on Signal Processing*, 56:85–95, 2008.
- [98] G. Rilling and P. Flandrin. Sampling effects on the empirical mode decomposition. *Advances in Adaptive Data Analysis*, 1(01):43–59, 2009.
- [99] G. Rilling, P. Flandrin, P. Goncalves, et al. On empirical mode decomposition and its algorithms. In *IEEE-EURASIP workshop on nonlinear signal and image processing*, volume 3, pages 8–11. NSIP-03, Grado (I), 2003.
- [100] G. Schlotthauer, M. E. Torres, and H. L. Rufiner. A new algorithm for instantaneous F0 speech extraction based on ensemble empirical mode decomposition. In *Signal Processing Conference, 2009 17th European*, pages 2347–2351. IEEE, 2009.

- [101] J. Schmitt, N. Pustelnik, P. Borgnat, P. Flandrin, and L. Condat. 2D Prony-Huang transform: A new tool for 2D spectral analysis. *IEEE Trans. Image Process.*, 23(12):5233–5248, Dec 2014.
- [102] R. C. Sharpley and V. Vatchev. Analysis of the intrinsic mode functions. *Constructive Approximation*, 24(1):17–47, 2006.
- [103] S. Stanković, I. Orović, and N. Zarić. An application of multidimensional time-frequency analysis as a base for the unified watermarking approach. *IEEE Trans. Image Process.*, 19(3):736–745, 2010.
- [104] N. Stevenson, M. Mesbah, and B. Boashash. A sampling limit for the empirical mode decomposition. In *ISSPA 2005: The 8th International Symposium on Signal Processing and its Applications, Vols 1 and 2, Proceedings*, volume 2, pages 647–650. IEEE, 2005.
- [105] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.*, 19(6):1635–1650, 2010.
- [106] G. Thakur, E. Brevdo, N. S. Fučkar, and H.-T. Wu. The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications. *Signal Processing*, 93(5):1079–1094, 2013.
- [107] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin. A complete ensemble empirical mode decomposition with adaptive noise. In *Acoustics, speech and signal processing (ICASSP), 2011 IEEE international conference on*, pages 4144–4147. IEEE, 2011.
- [108] M. Unser. Splines: A perfect fit for signal and image processing. *Signal Processing Magazine, IEEE*, 16(6):22–38, 1999.
- [109] M. Unser. Sampling-50 years after Shannon. *Proceedings of the IEEE*, 88(4):569–587, 2000.
- [110] J.-L. Wang and Z.-J. Li. What about the asymptotic behavior of the intrinsic mode functions as the sifting times tend to infinity? *Advances in Adaptive Data Analysis*, 4(01n02):1250008, 2012.
- [111] W. Wang, X. Li, and R. Zhang. Boundary processing of HHT using support vector regression machines. In *Computational Science-ICCS 2007*, pages 174–177. Springer, 2007.
- [112] X.-J. Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, 90(3):1195–1268, 2010.
- [113] H.-T. Wu. Instantaneous frequency and wave shape functions (I). *Applied and Computational Harmonic Analysis*, 35(2):181–199, 2013.
- [114] H.-T. Wu, Y.-H. Chan, Y.-T. Lin, and Y.-H. Yeh. Using synchrosqueezing transform to discover breathing dynamics from ECG signals. *Applied and Computational Harmonic Analysis*, 36(2):354–359, 2014.

- [115] H.-T. Wu, P. Flandrin, and I. Daubechies. One or two frequencies? The synchrosqueezing answers. *Advances in Adaptive Data Analysis*, 3(01n02):29–39, 2011.
- [116] Z. Wu and N. E. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. R. Soc. Lond. A*, 460:1597–1611, 2004.
- [117] Z. Wu and N. E. Huang. Statistical significance test of intrinsic mode functions. *Hilbert–Huang Transform and Its Applications*, pages 107–127, 2005.
- [118] Z. Wu and N. E. Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01):1–41, 2009.
- [119] Z. Wu and N. E. Huang. On the filtering properties of the empirical mode decomposition. *Advances in Adaptive Data Analysis*, 2(04):397–414, 2010.
- [120] Z. Wu, N. E. Huang, and X. Chen. The multi-dimensional ensemble empirical mode decomposition method. *Adv. Adapt. Data Anal.*, 1(03):339–372, 2009.
- [121] Z. Wu, E. K. Schneider, B. P. Kirtman, E. S. Sarachik, N. E. Huang, and C. J. Tucker. The modulated annual cycle: an alternative reference frame for climate anomalies. *Climate Dynamics*, 31(7-8):823–841, 2008.
- [122] Y. Xu, B. Liu, J. Liu, and S. Riemenschneider. Two-dimensional empirical mode decomposition by finite elements. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 462, pages 3081–3096. The Royal Society, 2006.
- [123] Z. Xu, B. Huang, and K. Li. An alternative envelope approach for empirical mode decomposition. *Digital Signal Processing*, 20(1):77–84, 2010.
- [124] J.-R. Yeh, J.-S. Shieh, and N. E. Huang. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, 2(02):135–156, 2010.
- [125] X. Zhang, K. K. Lai, and S.-Y. Wang. A new approach for crude oil price analysis based on empirical mode decomposition. *Energy economics*, 30(3):905–918, 2008.
- [126] X. Zhang, L. Yu, S. Wang, and K. K. Lai. Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method. *Energy Economics*, 31(5):768–778, 2009.

