

UNIVERSIDAD NACIONAL DE LA PLATA
Facultad de Ciencias Exactas
Departamento de Matemática

**Transformada Wavelet
y
Teoría de Información
en el análisis de señales complejas**



Tesis Doctoral
María Teresa Martín

Director: Dr. Angel Luis Plastino
Co-Director: Dr. Osvaldo Aníbal Rosso

Enero, 2004

*A mi Padre,
quien me trasmitió
la ilusión de lo posible.*

*A Carlos, Mariana, Andrés y Paula,
quienes hicieron posible mi ilusión.*

Índice General

1	Introducción General	1
1.1	Análisis de Señales	1
1.2	Objetivos y Organización de la Tesis	2
2	Fundamentos Teóricos	4
2.1	Análisis de Señales	4
2.1.1	Espacio-estado	4
2.1.2	Estacionariedad	5
2.1.3	Transitorios	7
2.2	Representación Tiempo-Frecuencia	7
2.2.1	Análisis de Fourier	7
2.2.2	Problema tiempo-frecuencia	8
2.2.3	Transformada de Gabor	9
2.2.4	Transformada wavelet continua	12
2.2.5	Transformadas wavelet discreta	13
2.2.6	Análisis multirresolución	14
2.3	Teoría de la Información	17
2.3.1	Entropía de Shannon	18
2.3.2	Medida de Información de Fisher	19
3	Análisis Wavelet y Medidas de Entropía	24
3.1	Introducción	24
3.2	Entropías de Series Temporales	25
3.2.1	Entropía de amplitudes	25
3.2.2	Entropía espectral	26
3.3	Entropías y Análisis Wavelet	27
3.4	Entropía de Amplitudes con Localización Frecuencial	28
3.4.1	Entropía multirresolución total (TMS)	28
3.5	Entropía Espectral con Localización Temporal	29
3.5.1	Energía y decimación ortogonal	29
3.5.2	Entropía wavelet	31
3.5.3	Entropía wavelet total (TWS)	31
3.6	Aplicaciones	32
3.6.1	Aplicaciones I: entropía para mapas discretos	33
3.6.2	Aplicaciones II: entropía para modelos continuos	36
3.7	Conclusiones	38

4 Medidas de Complejidad (I):	51
4.1 Introducción	51
4.2 Medidas de Complejidad	52
4.3 Complejidad Estadística: dos Medidas Simples	55
4.3.1 Medida de complejidad LMC	55
4.3.2 Medida de complejidad SDL	57
4.3.3 Inhomogeneidad y desequilibrio	58
4.4 Complejidad y Transformada Wavelet	59
4.4.1 Multi-escalas y complejidad	59
4.4.2 Complejidad wavelet	60
4.4.3 Complejidad wavelet total (TWC)	60
4.5 Aplicaciones III: Medidas de Complejidad	61
4.5.1 Mapa logístico	62
4.6 Conclusiones	63
5 Medidas de Complejidad (II):	69
5.1 Introducción	69
5.2 Generalización del Desequilibrio	70
5.3 Distancia Estadística de Wootters	71
5.3.1 Relación con la medida de información de Fisher	73
5.3.2 Desequilibrio según Wootters	75
5.4 Divergencias	76
5.4.1 Entropías relativas	78
5.4.2 Divergencias de Jensen	80
5.4.3 Desequilibrio según Jensen	81
5.5 Aplicaciones IV: Medidas C_W y C_{JS}	82
5.6 Conclusiones	83
6 Medidas de Complejidad (III):	88
6.1 Introducción	88
6.2 Medidas Generalizadas	89
6.2.1 Inhomogeneidad basada en entropías de orden q	89
6.2.2 Desequilibrio basado en distancias generalizadas	90
6.2.3 Complejidad generalizada	92
6.3 Análisis de la Familia de Medidas	93
6.3.1 Caracterización analítica de los extremos	93
6.3.2 Formalismo geométrico	95
6.3.3 Caracterización geométrica de los extremos	97
6.4 Aplicaciones V	102
6.5 Conclusiones	103
7 Aplicaciones a señales EEG	111
7.1 Introducción	111
7.2 Electroencefalograma (EEG)	111
7.2.1 Análisis de señales EEG	112
7.2.2 EEG y epilepsia	113

ÍNDICE GENERAL

iv

7.3 Separación de Señales y Análisis Wavelet	115
7.3.1 Resultados	116
7.4 Conclusiones	120
8 Resumen y consideraciones finales	135

Índice de Figuras

2.1	FUNCIÓN WAVELET SPLINE CÚBICA ψ UTILIZADA EN ESTE TRABAJO.	22
2.2	AMPLITUD Y FASE DE LA TRANSFORMADA DE FOURIER PARA LA FUNCIÓN WAVELET SPLINE CÚBICA ψ (VER FIG. 2.1).	23
3.1	MAPA DE HENON: SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y SU DIFERENCIA CON LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA).	39
3.2	MAPA DE HENON: ENTROPÍA DE SHANNON TAS, DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.	40
3.3	MAPA DE HENON: ENTROPÍA MULTIESCALAS TMS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE MUESTRAN LOS TRES PRIMEROS NIVELES DE RESOLUCIÓN. SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.	41
3.4	MAPA DE HENON: <i>a)</i> SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y SU DIFERENCIA CON LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). <i>b)</i> INFORMACIÓN DE FISHER TAI DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). <i>c)</i> ENTROPÍA DE SHANNON TAS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.	42
3.5	MAPA LOGÍSTICO: <i>a)</i> DIAGRAMA DE ÓRBITAS EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . <i>b)</i> EXPONENTE DE LYAPUNOV EN FUNCIÓN DEL PARÁMETRO DE CONTROL r	43
3.6	MAPA LOGÍSTICO: <i>a)</i> DIAGRAMA DE ÓRBITAS EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . <i>b)</i> MEDIDA DE INFORMACIÓN DE FISHER TAI EN FUNCIÓN DEL PARÁMETRO DE CONTROL. SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 10$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.	44
3.7	MAPA LOGÍSTICO: ENTROPÍA WAVELET TOTAL, TWS, EN FUNCIÓN DEL PARÁMETRO DE CONTROL.	45
3.8	MODELO DE LORENZ: SEÑAL CORRESPONDIENTE A LA COORDENADA y , GENERADA CON PARÁMETRO FIJO (LÍNEA PUNTEADA), Y GENERADA CON PARÁMETRO VARIABLE (LINEA LLENA).	46

3.9	MODELO DE LORENZ: ENTROPÍA DE SHANNON TAS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.	47
3.10	MODELO DE LORENZ: ENTROPÍA MULTIESCALAS TMS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). LOS TRES PRIMEROS NIVELES DE RESOLUCIÓN. SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.	48
3.11	MODELO DE LORENZ: <i>a)</i> EXPONENTES DE LYAPUNOV EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . λ_1 (LÍNEA LLENA), λ_2 (LÍNEA SEGMENTADA) Y λ_3 (LÍNEA PUNTEADA). <i>b)</i> MEDIDA DE INFORMACION DE FISHER, TAI, EN FUNCIÓN DEL PARÁMETRO DE CONTROL. SE CONSIDERARON VENTANAS DE 512 MUESTRAS Y $N = 50$ SUBINTERVALOS EN EL RANGO DE LAS AMPLITUDES.	49
3.12	MODELO DE ROSSLER: <i>a)</i> EXPONENTES DE LYAPUNOV EN FUNCIÓN DEL PARÁMETRO DE CONTROL c . λ_1 (LÍNEA LLENA), λ_2 (LÍNEA SEGMENTADA) Y λ_3 (LÍNEA PUNTEADA). <i>b)</i> MEDIDA DE INFORMACION DE FISHER, TAI, EN FUNCIÓN DEL PARÁMETRO DE CONTROL. SE CONSIDERARON VENTANAS DE 1024 MUESTRAS Y $N = 10^4$ SUBINTERVALOS EN EL RANGO DE LAS AMPLITUDES.	50
4.1	DIAGRAMA DE GRASSBERGER: COMPLETAMENTE ORDENADO (DERECHA), CAOS (CENTRO) Y COMPLETAMENTE DESORDENADO (IZQUIERDA)	65
4.2	MAPA LOGÍSTICO (WAVELET SPLINE CUBICA): <i>a)</i> DESEQUILIBRIO EUCLÍDEO (Q_E) EN FUNCIÓN DEL PARAMETRO DE CONTROL r . <i>b)</i> COMPLEJIDAD TWC EN FUNCIÓN DEL PARAMETRO DE CONTROL r	66
4.3	MAPA LOGÍSTICO (SUCESIÓN BINARIA): <i>a)</i> DESEQUILIBRIO EUCLÍDEO (Q_E) EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . <i>b)</i> COMPLEJIDAD C_{LMC} EN FUNCIÓN DEL PARÁMETRO DE CONTROL r	67
4.4	MAPA LOGÍSTICO (SUCESIÓN BINARIA): COMPLEJIDAD C_{SDL} EN FUNCIÓN DEL PARÁMETRO DE CONTROL r	68
5.1	MAPA LOGÍSTICO (SUCESIÓN BINARIA): <i>a)</i> DESEQUILIBRIO DE WOOTTERS (Q_W) Y <i>b)</i> MEDIDA DE COMPLEJIDAD C_W EN FUNCIÓN DEL PARÁMETRO DE CONTROL r	85
5.2	MAPA LOGÍSTICO (SUCESIÓN BINARIA): <i>a)</i> DESEQUILIBRIO DE JENSEN-SHANNON (Q_{JS}) Y <i>b)</i> MEDIDA DE COMPLEJIDAD C_{JS} EN FUNCIÓN DEL PARÁMETRO DE CONTROL r	86
5.3	Grafico de cantidades de interes en funcion del parametro r restringido al rango asociado a la ventana de período 3. <i>a)</i> diagrama de orbitas, <i>b)</i> medida de complejidad $C_{(LMC)}$, <i>c)</i> medida de complejidad $C_{(JS)}$	87
6.1	ESPACIO DE PROBABILIDADES Ω PARA $N = 4$: <i>a)</i> $\Omega \equiv \Delta^2$ EN EL HIPERPLANO DE DIMENSIÓN 3. LAS LÍNEAS PUNTEADAS TRAZAN LA SUBDIVISIÓN BARICÉNTRICA Y μ_3 BARICENTRO DE Ω . <i>b)</i> SUB-SIMPLEX Δ_I^3	105
6.2	GRÁFICO DE LA MEDIDA DE COMPLEJIDAD EN FUNCIÓN DE $H^{(S)}(P)$, CUANDO P VARÍA SOBRE $\mathcal{B}^{(1)}$ O SOBRE $\mathcal{B}^{(N_1)}$, PARA $N = 6$. <i>a)</i> $C_E^{(S)}(P)$ Y <i>b)</i> $C_W^{(S)}(P)$	106

6.3	SIMULACIÓN NUMÉRICA DE LA MEDIDA DE COMPLEJIDAD EN FUNCIÓN DE $H^{(S)}(P)$, CUANDO P VARÍA EN Ω , PARA $N = 6$. a) $C_E^{(S)}(P)$ Y b) $C_W^{(S)}(P)$.	107
6.4	CARA BIDIMENSIONAL: GRÁFICO DE LA MEDIDA DE COMPLEJIDAD EN FUNCIÓN DE $H^{(S)}(P)$, CUANDO P VARÍA SOBRE L_{ik} O SOBRE $L_{ij} \cup L_{jk}$. a) $C_E^{(S)}(P)$ Y b) $C_W^{(S)}(P)$.	108
6.5	a) DIAGRAMA DE ÓRBITAS b) LYAPUNOV EXPONENT Y c) ENTROPÍA NORMALIZADA H PARA EL MAPA LOGÍSTICO EN FUNCIÓN DEL PARÁMETRO r .	109
6.6	CANTIDADES DE INTERÉS SON GRAFICADAS EN FUNCIÓN DE LA ENTROPÍA NORMALIZADA H : a) EXPONENTE DE LYAPUNOV λ . NOTAR QUE LAS VENTANAS PERIÓDICAS SE DISTINGUEN CLARAMENTE PARA $H^* \leq 0.3$. b) COMPLEJIDAD C_E Y c) COMPLEJIDAD C_{JS} . EN b) Y c) SE GRAFICARON TAMBIÉN LAS COTAS C_{max} Y C_{min} (LÍNEA CONTINUA). LA LÍNEA VERTICAL PUNTEADA SEÑALA EL VALOR H^* PARA EL FINAL DE LA ZONA DE PERIÓDICIDAD.	110
7.1	SEÑAL EEG (<i>scalp</i>) CORRESPONDIENTE A UNA CRISIS EPILÉPTICA TÓNICO-CLÓNICA, REGISTRADO EN LA UBICACIÓN CENTRAL DERECHA $C4$. LA CRISIS COMIENZA A LOS 80 s Y LA FASE CLÓNICA A LOS 125 s. LA CRISIS FINALIZA A LOS 155 s.	123
7.2	SEÑAL EEG LIBRE DE ARTIFICIOS (RUIDO). SIMILAR A LA FIG. 7.1 SIN LA CONTRIBUCIÓN DE LA ACTIVIDAD MUSCULAR REPRESENTADA POR BANDAS DE FRECUENCIA B_1 Y B_2 , VER TABLA. 7.1.	124
7.3	EVOLUCIÓN TEMPORAL DE LA ENERGÍA MULTIESCALA WAVELET CORRESPONDIENTE A LA SEÑAL EEG SIN ACTIVIDAD MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2) PARA LAS BANDAS DE FRECUENCIA B_3 (LÍNEA SÓLIDA), B_4 (LÍNEA PUNTEADA), B_5 (LÍNEA PUNTO-GUIÓN), B_6 (LÍNEA GUIÓN). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.	125
7.4	EVOLUCIÓN TEMPORAL DE LA ENTROPÍA WAVELET TOTAL (TWS). LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DEL CRISIS EPILÉPTICA.	126
7.5	a) EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_{SDL} . LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). b) IGUAL AL ANTERIOR PARA LA EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_{LMC} . LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.	127
7.6	a) DIAGRAMA “COMPLEJIDAD vs ENTROPÍA NORMALIZADA” CORRESPONDIENTE A TWC_{LMC} PARA LA SEÑAL EEG LIBRE DE ARTIFICIOS (FIG. 7.2). CON LÍNEA CONTINUA SE HA REPRESENTADO LA COMPLEJIDAD MÁXIMA C_{max} Y MÍNIMA C_{min} EN FUNCIÓN DE LA ENTROPÍA WAVELET SHANNON NORMALIZADA. b) EVOLUCIÓN EN ESTE ESPACIO, PARA LAS FASES TÓNICA Y CLÓNICA. DETALLE CORRESPONDIENTE AL INTERVALO ASOCIADO CON EL RITMO DE RECLUTAMIENTO EPILÉPTICO.	128

7.7	EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_W . LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.	129
7.8	a) DIAGRAMA “COMPLEJIDAD <i>vs</i> ENTROPÍA NORMALIZADA” CORRESPONDIENTE A TWC_W PARA LA SEÑAL EEG LIBRE DE ARTIFICIOS (FIG. 7.2). CON LÍNEA CONTINUA SE HA REPRESENTADO LA COMPLEJIDAD MÁXIMA C_{max} Y MÍNIMA C_{min} EN FUNCIÓN DE LA ENTROPÍA WAVELET SHANNON NORMALIZADA. b) EVOLUCIÓN EN ESTE ESPACIO, PARA LAS FASES TÓNICA Y CLÓNICA. DETALLE CORRESPONDIENTE AL INTERVALO ASOCIADO CON EL RITMO DE RECLUTAMIENTO EPILÉPTICO.	130
7.9	EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_{JS} . LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.	131
7.10	a) DIAGRAMA “COMPLEJIDAD <i>vs</i> ENTROPÍA NORMALIZADA” CORRESPONDIENTE A TWC_{JS} PARA LA SEÑAL EEG LIBRE DE ARTIFICIOS (FIG. 7.2). CON LÍNEA CONTINUA SE HA REPRESENTADO LA COMPLEJIDAD MÁXIMA C_{max} Y MÍNIMA C_{min} EN FUNCIÓN DE LA ENTROPÍA WAVELET SHANNON NORMALIZADA. b) EVOLUCIÓN EN ESTE ESPACIO, PARA LAS FASES TÓNICA Y CLÓNICA. DETALLE CORRESPONDIENTE AL INTERVALO ASOCIADO CON EL RITMO DE RECLUTAMIENTO EPILÉPTICO.	132
7.11	TWS ($media \pm SE$) EVALUADO PARA LAS 20 CRISIS EPILEPTICAS TÓNICO-CLÓNICAS PARA LOS PERÍODOS PRE-ICTAL E ICTAL. a) SEÑAL EEG ORIGINAL. b) SEÑAL EEG SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS). CADA PACIENTE ES REPRESENTADO POR UNA LETRA Y EL NÚMERO A SU LADO IDENTIFICA LA CRISIS CORRESPONDIENTE.	133
7.12	TWC_W ($media \pm SE$) VALUADO PARA LAS 20 CRISIS EPILEPTICAS TÓNICO-CLÓNICAS PARA LOS PERÍODOS PRE-ICTAL E ICTAL. a) SEÑAL EEG ORIGINAL. b) SEÑAL EEG SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS). CADA PACIENTE ES REPRESENTADO POR UNA LETRA Y EL NÚMERO A SU LADO IDENTIFICA LA CRISIS CORRESPONDIENTE.	134

Índice de Tablas

- 7.1 INTERVALOS DE FRECUENCIA (EN Hz) ASOCIADOS CON LOS NIVELES DE RESOLUCIÓN WAVELET j Y TIEMPO DE RESOLUCIÓN ΔT (EN s) PARA UNA FRECUENCIA DE MUESTREO $\omega_s = 102.4\ Hz$. LAS BANDAS DE FRECUENCIA EEG TRADICIONALES CORRESPONDEN A LAS SIGUIENTES INTERVALOS DE FRECUENCIAS: δ (0.5-3.5 Hz); θ (3.5-7.5 Hz); α (7.5-12.5 Hz); β (12.5-30.0 Hz); γ (30.0-70.0 Hz). 122

Capítulo 1

Introducción General

1.1 Análisis de Señales

En el análisis de señales se distinguen dos principales escuelas. La más antigua proviene de la estadística aplicada [68] y prevalece entre los estadísticos, científicos sociales (especialmente economistas) e ingenieros. La escuela más moderna, desarrollada esencialmente desde 1970, tiene su origen en la física y la teoría de sistemas dinámicos. Para la primera las series temporales son vistas como muestras de un proceso estocástico, y aplica una combinación de técnicas y suposiciones estadísticas tradicionales (regresión lineal, propiedades de las distribuciones Gaussianas) junto con el Análisis de Fourier. Por otra parte, la segunda escuela considera a las series temporales como mediciones distorsionadas y ruidosas de un subyacente sistema dinámico, que intenta reconstruir [66].

La separación entre las dos escuelas se debe en parte al hecho de que cuando los métodos estadísticos para el análisis de series temporales ya estaban casi formalizados, en 1920-1930, la teoría de los sistemas dinámicos recién estaba comenzando. El desarrollo fuerte de la dinámica no lineal como una potente disciplina, recién tuvo lugar a partir de 1970, generando un marcado crecimiento de la teoría estadística. En efecto, físicos dedicados a la dinámica no lineal experimental, por ejemplo, se convirtieron en firmes impulsores del desarrollo de herramientas estadísticas.

En los últimos años, se han manifestado signos evidentes de una comunióón entre estas dos escuelas. Desde 1960, estadísticos han comenzado a darse cuenta de las virtudes de los “modelos espacio-estado”, y por su parte los físicos, se muestran más sensitivos a las técnicas estadísticas. Existe actualmente un creciente trabajo interdisciplinario. En este sentido pueden mencionarse los trabajos de Grassberger et al. [51] y Abarbanel [1], donde se presentan sendas revisiones de series temporales provenientes de sistemas caóticos. Ambos trabajos ofrecen abundantes referencias orientadas a la investigación sobre esta temática.

La mayoría de las técnicas provenientes de estas escuelas clásicas, presentan serias limitaciones en su aplicación al asumir la estacionariedad de la señal. En este sentido, se podría hablar de una nueva corriente desarrollada específicamente para el análisis de señales no estacionarias que se basa en los métodos tiempo-frecuencia, con especial énfasis en la Transformada Wavelet [87]. En este caso la bibliografía es abundante, tanto

en lo que respecta a investigaciones teóricas [88, 30, 20], como en desarrollos orientados hacia las aplicaciones [2, 29, 113].

1.2 Objetivos y Organización de la Tesis

El objetivo principal de esta tesis es el desarrollo de métodos específicos para el análisis de señales no estacionarias. En particular, estos métodos tendrán la capacidad para detectar y cuantificar cambios de las dinámicas subyacentes, con el fin de caracterizar distintos aspectos del fenómeno físico asociado.

Algunos métodos clásicos, como la entropía de amplitudes y la entropía espectral, y otros recientemente desarrollados desde un enfoque tiempo-frecuencia, como la entropía multiescala [127] y la entropía wavelet [113], cuantifican aspectos ligados a dos importantes cuestiones dentro del tratamiento de señales, como son la morfología y el contenido frecuencial de una señal. Desde diversos campos de la ciencia surge actualmente el interés por medidas que cuantifiquen la llamada “complejidad” de una señal. Si bien no existe un consenso conceptual sobre este tema, se pueden inferir algunos mínimos requerimientos que deberán satisfacer tales medidas. Por ejemplo, existe coincidencia respecto a que sistemas con “orden perfecto” o “aleatoriedad total”, deberían tener medida de complejidad nula. Entre estas dos situaciones extremas, catalogadas como “triviales”, existe una amplia gama de comportamientos “no-triviales” que estas nuevas medidas deberían cuantificar. Con la idea de realizar un aporte en el contexto planteado, se discutirán en esta tesis medidas de “no-trivialidad” o “complejidad” con propiedades ventajosas para su aplicación en el análisis de señales noestacionarias, como son en general las señales provenientes del mundo natural.

La organización de la tesis es la siguiente:

En el capítulo 2, y a fin de lograr una presentación autoconsistente, se expondrán en forma breve algunas consideraciones teóricas. De este modo, se realizará una sintética presentación de conceptos ligados al análisis de señales no-estacionarias. También se desarrollarán las ideas básicas vinculadas al problema tiempo-frecuencia. Finalizando el capítulo con un breve comentario relativo a Teoría de la Información, centrado en la entropía de Shannon y la medida de Información de Fisher.

En el capítulo 3 se presentarán dos medidas entrópicas especialmente orientadas al tratamiento digital de señales no-estacionarias. Estas son la Entropía Multiresolución Total y la Entropía Wavelet Total.

Una nueva medida de información, inspirada en los comportamientos multiescalas típicos de los sistemas naturales, será presentada en el capítulo 4. Esta medida es la Complejidad Wavelet Total (TWC) y está orientada a cuantificar la “estructura” o “complejidad” de señales no-estacionarias.

En el capítulo 5 una apertura del enfoque considerado hasta el momento, permitirá incorporar las nociones de “distancia estadística” de Wootters y divergencia de Jensen-Shannon, en la formulación de la medida de complejidad estadística.

Las medidas de complejidad introducidas en los capítulos precedentes y una ampliación del marco teórico, proporcionarán una familia de medidas de complejidad es-

tadística generalizada. Sus propiedades serán estudiadas en el capítulo 6. Se determinarán cotas de variación para las medidas de complejidad en el plano *complejidad vs. entropía*.

En todos los casos se realizarán exemplificaciones con señales simuladas desde sistemas teóricos, y el capítulo 7 estará dedicado a las aplicaciones de los métodos propuestos a señales electroencefalográficas (EEG). Se discutirán las ventajas de estas nuevas técnicas, así como también las conclusiones extraídas.

En el capítulo 8 se exponen las conclusiones generales.

Capítulo 2

Fundamentos Teóricos

2.1 Análisis de Señales

El tratamiento de señales es una componente esencial de la actividad científica y tecnológica contemporánea. El concepto de señal aparece en una amplia variedad de campos, y las ideas y técnicas asociadas juegan un rol importante en áreas muy diversas como son: comunicaciones, aeronáutica y astronáutica, diseño de circuitos, acústica, sismología, ingeniería biomédica, control de procesos químicos, imágenes satelitales y reconocimiento de voces, entre otras. Aunque la naturaleza física de las señales que aparecen en disciplinas tan variadas puede ser totalmente diferente, todas tienen dos características básicas en común: son función de una o más variables independientes y típicamente contienen información acerca del comportamiento o naturaleza de algún fenómeno particular.

2.1.1 Espacio-estado

Siguiendo el enfoque provisto por la teoría de sistemas dinámicos, una señal representa un sistema cuya evolución temporal está definida en algún “espacio de estados”. El concepto de “espacio de estados” o “espacio-fase” adquiere gran importancia aún en sistemas no determinísticos. Asumiendo que el espacio-fase es el espacio vectorial de dimensión finita \mathbb{R}^m , un estado será especificado por un vector $x \in \mathbb{R}^m$. La dinámica puede describirse o bien por un mapa m -dimensional o por un sistema explícito de m ecuaciones diferenciales ordinarias de primer orden. En el primer caso, el tiempo es una variable discreta:

$$x_{k+1} = F(x_k, t_k, \epsilon_k) , \quad (2.1)$$

donde la función F depende no sólo del tiempo t_k sino también de una sucesión de variables aleatorias independientes ϵ_k , que dan lugar a la llamada “dinámica ruidosa” [66]. En el segundo caso el tiempo es una variable continua:

$$\frac{dx}{dt} = F(x(t), t, \epsilon_t) . \quad (2.2)$$

Si se pudiera disponer de la serie de estados x_k , $k \in \mathbb{Z}$, seguramente podrían investigarse muchas propiedades, lo cual convierte a x_k en una sucesión muy conveniente para ser analizada.

Lamentablemente sin embargo, la variable observada no es exactamente el estado x . La cantidad que se observa o mide, llamada s (señal), es generalmente una función ruidosa y nolineal del estado:

$$s_k = h(x_k, \eta_k) , \quad (2.3)$$

donde η_k es el ruido proveniente de la medición. En general no se conoce la función de observación h , ni la dinámica de estado F , ni aún el espacio-fase en el que realmente está x . El éxito de los métodos de análisis de series temporales, consiste precisamente en realizar aproximaciones o suposiciones apropiadas para todas estas entidades, y así entender y predecir mejor la evolución temporal del sistema que provee dicho conjunto de datos.

2.1.2 Estacionariedad

A veces las señales carecen de una clara evidencia determinística. En estos casos y desde un enfoque estadístico, las señales pueden ser vistas como realizaciones de un proceso aleatorio. Señales discretas y finitas $s = \{s_1, \dots, s_M\}$ son representadas por un vector aleatorio X , cuyas componentes $X[k]$ son variables aleatorias correspondientes a cada instante de tiempo simbolizado por k , con $1 \leq k \leq M$. De este modo, las señales tienen las mismas propiedades que las variables aleatorias, tales como media, correlación, varianza, etc., y será entonces importante saber si estas propiedades estadísticas, permanecen para el proceso aleatorio entero. Con este fin se ha desarrollado el concepto de proceso estacionario.

El concepto de estacionariedad, desde un punto de vista cualitativo, se refiere a la invariancia en el tiempo de las propiedades estadísticas del vector aleatorio X y por ende de la señal s [66]. Inversamente, procesos cuyas propiedades estadísticas cambian en el tiempo, son llamados: *no estacionarios*. Más técnicamente, se distinguen varios niveles de estacionariedad, entre los cuales los más frecuentes son:

Estacionariedad de primer-orden:

Un proceso aleatorio es *estacionario de primer-orden* si sus funciones densidad de primer orden, son independientes del ajuste temporal aplicado sobre la variable. Llamando $f_{X[k]}$ a la densidad de la variable correspondiente al tiempo k , se deberá satisfacer la siguiente ecuación:

$$f_{X[k]} = f_{X[k+\tau]} , \quad (2.4)$$

para $1 \leq k \leq M$ y valores apropiados τ . Es decir que las densidades son independientes del tiempo k y de cualquier corrimiento τ . El más importante resultado de la ecuación (2.4) y la característica que identifica a los procesos estacionarios de primer-orden, es que la media es constante, independiente de cualquier corrimiento en el tiempo.

Estacionariedad de segundo-orden:

Un proceso aleatorio es *estacionario de segundo-orden* si sus densidades de probabilidad conjunta de segundo orden f_{Y_1, Y_2} , son independientes del ajuste temporal aplicado sobre ambas variables aleatorias Y_1 y Y_2 .

En el caso del proceso aleatorio X correspondiente a la señal s , se tiene:

$$f_{X[k], X[k']} = f_{X[k+\tau], X[k'+\tau]}, \quad (2.5)$$

para $1 \leq k, k' \leq M$ y valores apropiados τ . Esta ecuación revela que las densidades conjuntas no están afectadas por el tiempo absoluto, ya que ellas en realidad sólo dependen de la diferencia de tiempo entre las dos variables. Para un proceso estacionario de segundo-orden, la función de autocorrelación manifiesta su más importante propiedad. Llamando R a la autocorrelación, y denotando por E a la esperanza o valor esperado, se cumple que:

$$R[k, k + \tau] = E\{X[k], X[k + \tau]\} = R_X[\tau] \quad \text{para } 1 \leq k \leq M. \quad (2.6)$$

Es decir que en un proceso estacionario de segundo-orden, la función correlación de dos puntos depende solamente de la distancia entre dichos puntos.

En un sentido más general, se suele hablar de dos *grados* de estacionariedad, que siguiendo la nomenclatura de Mallat [80], serán llamados: *estacionariedad en sentido estricto* (SSS), y *estacionariedad en sentido amplio* (WSS).

- Un proceso aleatorio X es llamado *estacionario en sentido estricto* cuando *todas* las funciones de distribución conjunta del proceso son independientes del ajuste temporal aplicado sobre las variables aleatorias.
- Relajando suavemente las condiciones, se dice que un proceso X es *estacionario en sentido amplio* si satisface los dos siguientes requerimientos:
 1. $E\{X[k]\} = \text{constante}, \forall 1 \leq k \leq M.$
 2. $E\{X[k], X[k + \tau]\} = R_X[\tau].$

Por otro lado, asumiendo que solamente se conocen las realizaciones s_k de un proceso, cualquier información sobre el vector aleatorio X solamente podrá inferirse a través de las propiedades estadísticas de la serie temporal asociada. De este modo, la evaluación de propiedades estadísticas relevantes servirá para testear la estacionariedad de la señal.

Las señales estacionarias han ocupado un lugar preponderante en el procesamiento clásico de señales, habiéndose dedicado mucho esfuerzo al diseño de operadores tiempo-invariante. Esto condujo a la indiscutible hegemonía de la transformada de Fourier, dejando afuera muchas aplicaciones. En efecto, llamando *no estacionarios* a los procesos que no satisfacen las condiciones 1 y 2, resultará que una clase muy amplia de señales es no estacionaria. Puede citarse al respecto la frase de S. Mallat en Ref. [80]: “El mundo de los transitorios es considerablemente mayor y más complejo que el jardín de las señales estacionarias” ¹. Esto explica el creciente interés por una base ideal tipo Fourier, que

¹The world of transients is considerably larger and more complex than the garden of stationary signals.

pudiera simplificar el procesamiento de señales no estacionarias. Ciertamente, una gran cantidad de transformadas y bases han aparecido en la literatura pertinente, entre las cuales las wavelets son un ejemplo.

2.1.3 Transitorios

Señales complejas poseen una fenomenología muy rica ligada aspectos variados del análisis. Algunos de ellos se refieren a problemas que pueden aparecer con la reproducibilidad o estacionariedad de la señal, aún cuando los parámetros del sistema permanezcan fijos. Un ejemplo claro son los llamados “comportamientos transitorios”. Desde el momento que un “comportamiento transitorio” desaparecerá después de un intervalo de tiempo, parecería no ser en sí mismo ni relevante ni interesante para el análisis de series temporales. Sin embargo estas dos impresiones son erróneas. En efecto, al observar la evolución temporal de un sistema, es posible que la trayectoria demore un “tiempo transitorio” antes de ingresar a un posible atractor. Durante este “tiempo transitorio” las propiedades de la señal pueden ser completamente diferentes a las que tendrá luego sobre el mismo atractor.

El “transitorio” puede ser extremadamente corto, aunque también se presentan casos en los cuales tiene una duración muy extensa. De este modo, un “transitorio” puede ser un verdadero problema que convierte a la señal en no estacionaria. Pero también a veces es en sí mismo de tanto interés, que justifica una observación lo suficientemente lenta como para que el registro incluya solamente al “comportamiento transitorio”. Citando nuevamente a S. Mallat en Ref. [80]: “Concentrarse en los transitorios es probablemente una estrategia para seleccionar la información importante de la sobreabundancia de datos recolectados por nuestros sentidos ” ².

2.2 Representación Tiempo-Frecuencia

Se realizará en esta sección una breve introducción a la Transformada Wavelet (WT), siguiendo en la exposición el lineamiento de su desarrollo histórico. Según el esquema propuesto por E. Serrano en Ref. [114], se presentará la WT como una natural extensión de los métodos de Fourier.

Gran número de publicaciones desarrollan este tema desde ángulos diversos, siendo particularmente importantes los trabajos realizados por Y. Meyer, I. Daubechies, S. Mallat, A. Aldroubi, M. Unser, Ch. Chui y R. Coifman [2, 132, 20, 21].

2.2.1 Análisis de Fourier

En el ámbito de las ciencias aplicadas usualmente se representa una señal física mediante una función del tiempo $s(t)$ o, alternativamente en el dominio de la frecuencia por su Transformada de Fourier (FT) $\hat{s}(\omega)$. Ambas representaciones son en cierto sentido naturales, resultantes de la habitual modalidad de enfocar el universo real. Las mismas contienen exactamente la misma información sobre la señal, respondiendo a enfoques

²Concentrating on transients is probably a strategy for selecting important information from the overwhelming amount of data recorded by our senses.

distintos y complementarios. A partir del fundamental trabajo de Fourier en el siglo XIX, se ha aprendido mucho acerca del mundo natural usando su potente herramienta.

Suponiendo que la señal es aperiódica y de energía finita, es decir $s \in L^2(\mathbb{R})$, las representaciones en tiempo y en frecuencia se relacionan mediante las siguientes fórmulas de Fourier:

$$\widehat{s}(\omega) = \int_{-\infty}^{\infty} s(t)e^{-i\omega t} dt , \quad (2.7)$$

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{s}(\omega)e^{i\omega t} d\omega , \quad (2.8)$$

donde t representa el tiempo y ω la frecuencia angular. Esto indica que, en teoría, la información en uno de los dominios puede recuperarse a partir de la información desplegada en el otro.

También expresa (2.8) que un proceso temporal puede considerarse como la superposición integral de una colección de ondas monocromáticas que oscilan con amplitud constante. Estas ondas, de frecuencia definida, están representadas por las funciones $s_\omega(t) = \widehat{s}(\omega)e^{i\omega t}$. Al interferirse entre sí conforman los distintos fenómenos y estructuras, localizadas en el tiempo. Recíprocamente se observa que, para cada frecuencia, la amplitud de la onda $s_\omega(t)$ es la resultante de la información de la señal integrada sobre el dominio tiempo. En suma, ambas representaciones están relacionadas en forma global y en principio la totalidad de la información en un dominio es requerida para recuperar la información en el otro.

2.2.2 Problema tiempo-frecuencia

Al procesar señales es común encontrar fenómenos oscilantes, casi monocromáticos, localizados en el tiempo. También ondas perdurables, que con amplitudes casi estacionarias, exhiben patrones de frecuencias variables en el tiempo. Surge entonces, en forma natural, la noción de los fenómenos localizados en tiempo y frecuencia, es decir de objetos que para su descripción requieren información conjunta de ambos dominios. Claramente, la Transformada de Fourier (FT) no es la herramienta adecuada para expresar explícitamente este tipo de información conjunta, dado que las funciones elementales de representación son las ondas estacionarias y monocromáticas $e^{i\omega t}$.

En contraposición, para representar tales fenómenos se requiere de patrones elementales capaces de localizar conjuntamente la información de ambos dominios. Esto plantea el problema de las *representaciones tiempo-frecuencia*. El problema es arduo, sobre todo, cuando se está frente a señales donde conviven múltiples fenómenos localizados en el tiempo y en la frecuencia, superponiéndose bajo complejas estructuras. Entonces, ¿cómo detectar esos fenómenos y caracterizar las estructuras? Claramente, por las razones expuestas no existe una respuesta única. Entonces el desafío consiste en definir una apropiada representación acorde con las características de la señal y a los objetivos de su procesamiento.

Las señales biológicas, por ejemplo, plantean este tipo de requerimientos. Típicamente estas señales son no-estacionarias y a menudo se presentan como sucesivos eventos transitorios, que contienen en general, la información más relevante. Frecuentemente tales eventos están parcialmente superpuestos en el dominio temporal, o afectados de ruido,

en el sentido que se interfieren con eventos secundarios y no significativos, pero que afectan su estructura. Uno de los problemas comunes en el procesamiento de señales biológicas es caracterizar parámetros significativos de la señal, cuando éstos involucran variaciones conjuntas en el dominio tiempo-frecuencia.

Para medir el contenido de información tiempo-frecuencia de una señal, en 1946 el físico D. Gabor [40], introdujo una “función ventana” con localización temporal, $g(t - b)$ donde el parámetro b es usado para trasladar la ventana en el tiempo, cubriendo el dominio temporal entero y extraer así información local de la Transformada de Fourier (FT) de la señal. En paralelo a la contribución de Gabor, en 1948 Ville [133] propuso analizar las propiedades tiempo-frecuencia de señales mediante una densidad de energía bidimensional. Desde la física teórica la distribución de Ville ya había sido introducida en 1932 por Wigner [136] en el contexto de la mecánica cuántica.

2.2.3 Transformada de Gabor

Una clase de representaciones tiempo-frecuencia ampliamente difundida en el ámbito del procesamiento de señales se basa en el empleo de “ventanas temporales”, es decir funciones suaves y bien localizadas en un intervalo. La ventana $g(t)$ enmarca una porción de la señal que se supone estacionaria, y permite aplicar localmente la Transformada de Fourier (FT). De este modo, se releva la información en frecuencia localizada temporalmente en el dominio efectivo de la ventana. Desplazando la ventana se cubre el dominio temporal de la señal obteniéndose una completa información tiempo-frecuencia de la misma.

Si $s \in L^2(\mathbb{R})$, se define la Transformada de Fourier por “ventanas” (WFT), como:

$$\mathcal{G}s(b, \omega) = \int_{-\infty}^{\infty} s(t)e^{-i\omega t}g(t - b) dt , \quad (2.9)$$

donde $b \in \mathbb{R}$ y $g(t - b)$ localiza la integral de Fourier en el entorno de $t = b$.

La Transformada de Gabor es una WFT, que utiliza como función ventana una Gaussiana. Por varias razones tales como eficiencia computacional o conveniencia en la implementación, otras funciones “ventana”, diferentes a la Gaussiana, pueden ser utilizadas en su lugar. Para que una función no trivial $g \in L^2(\mathbb{R})$ sea calificada como “función ventana”, debe satisfacer el siguiente requerimiento [20]:

$$t \cdot g(t) \in L^2(\mathbb{R}) . \quad (2.10)$$

Si $\|g\| = 1$, se define el centro de la función ventana g , como:

$$t^* = \int_{-\infty}^{\infty} t |g(t)|^2 dt . \quad (2.11)$$

La dispersión alrededor de t^* estará medida por la varianza Δ_g :

$$\Delta_g = \left[\int_{-\infty}^{\infty} (t - t^*)^2 |g(t)|^2 dt \right] . \quad (2.12)$$

Por otra parte, la transformada $\mathcal{G}s(b, \omega)$ puede ser interpretada también, en un sentido diferente al presentado en (2.9). Si se llama:

$$g_{b,\omega}(t) = e^{i\omega t} g(t - b) , \quad (2.13)$$

se tiene que:

$$\mathcal{G}s(b, \omega) = \langle s, g_{b,\omega} \rangle = \int_{-\infty}^{\infty} s(t) \overline{g_{b,\omega}(t)} dt . \quad (2.14)$$

Es decir que, en lugar de considerar $\mathcal{G}s(b, \omega)$ como una localización de la FT de s , se la puede interpretar como una transformada que correlaciona la señal con una familia de funciones $g_{b,\omega}$ dadas por la ec. (2.13), bien localizadas en tiempo y en frecuencia, llamadas “átomos tiempo-frecuencia”. Un “átomo” es construido con una función ventana $g(t)$, trasladada por b y modulada en la frecuencia ω , como indica la ec. (2.13). De este modo, observando (2.13) se ve que $\mathcal{G}s(b, \omega)$ da información local de s en la ventana temporal:

$$[t^* + b - \Delta_g, t^* + b + \Delta_g] , \quad (2.15)$$

centrada en $t^* + b$ y “ancho” Δ_g .

Supongamos ahora que su FT \widehat{g} , es también una función ventana, es decir cumple la condición (2.10). Se podrá entonces determinar su centro ω^* y su dispersión $\Delta_{\widehat{g}}$. La FT de $g_{b,\omega}$, que denotamos por $\widehat{g}_{b,\omega}$, cumple que:

$$\frac{1}{2\pi} \widehat{g}_{b,\omega}(\eta) = \left(\frac{e^{ib\eta}}{2\pi} \right) e^{-ib\eta} \widehat{g}(\eta - \omega) . \quad (2.16)$$

Por lo tanto, $\widehat{g}_{b,\omega}$ es también una función ventana con centro en $\omega^* + \omega$ y dispersión $\Delta_{\widehat{g}}$. Aplicando la Identidad de Parseval:

$$\mathcal{G}s(b, \omega) = \langle s, g_{b,\omega} \rangle = \langle \widehat{s}, \widehat{g}_{b,\omega} \rangle . \quad (2.17)$$

Esto muestra que $\mathcal{G}s(b, \omega)$ también da información espectral de s , localizada en la ventana de frecuencias:

$$[\omega^* + \omega - \Delta_{\widehat{g}}, \omega^* + \omega + \Delta_{\widehat{g}}] . \quad (2.18)$$

En síntesis, eligiendo una función ventana g , tal que \widehat{g} también satisfaga la condición (2.10), $g_{b,\omega}$ corresponde a un rectángulo de “ancho” $2\Delta_g$ (el ancho de la ventana temporal) y área constante $4\Delta_g \Delta_{\widehat{g}}$, centrado en (b, ω) . El tamaño de este rectángulo es independiente del par de índices (b, ω) , lo que significa que la WFT tiene la misma resolución a lo largo de todo el plano tiempo-frecuencia.

Definición 2.1:

Si $g \in L^2(\mathbb{R})$ es tal que ambas g y su transformada de Fourier \widehat{g} satisfacen (2.10), entonces la transformada de Fourier por ventanas (WFT) es llamada *Transformada de Fourier “tiempo-corto”* (STFT).

- **Completitud y Estabilidad**

Cuando los índices (b, ω) varían a través de \mathbb{R}^2 , los rectángulos de localización de los átomos $g_{b,\omega}(t)$ cubren el plano tiempo-frecuencia entero. Es de esperar entonces que s pueda ser recuperada desde su transformada $\mathcal{G}s(b, \omega)$. Efectivamente se cumple la siguiente fórmula de reconstrucción y se prueba además que se conserva la energía.

Si $s \in L^2(\mathbb{R})$ entonces:

$$s(t) = \frac{1}{2\pi} \int \int \mathcal{G}s(b, \omega) g_{b,\omega}(t) d\omega db , \quad (2.19)$$

y

$$\int |s(t)|^2 dt = \frac{1}{2\pi} \int \int |\mathcal{G}s(b, \omega)|^2 d\omega db . \quad (2.20)$$

La ecuación (2.19) sintetiza la señal como la superposición integral de las funciones elementales $g_{b,\omega}$. Si bien esta forma recuerda la descomposición de la señal en una base ortonormal, esto no es así, porque las funciones $\{g_{b,\omega}\}_{\mathbb{R}^2}$ son altamente redundantes en \mathbb{R}^2 . $\mathcal{G}s(b, \omega)$ representa una señal unidimensional $s(t)$ por una función bidimensional $\mathcal{G}s(b, \omega)$. La conservación de la energía prueba que $\mathcal{G}s \in L^2(\mathbb{R}^2)$.

- **Elección de la ventana**

Para la implementación numérica de la transformada, es necesario discretizar los parámetros b y ω en una red apropiada, que equilibre la eficiencia del mapeo de la información con la complejidad computacional. La resolución en tiempo y en frecuencia de la WFT depende de la dispersión de la ventana en tiempo y en frecuencia, y la redundancia depende la superposición de los “átomos tiempo-frecuencia”. Esto es función del área de los rectángulos de localización: $4\Delta_g \Delta_{\hat{g}}$. Justamente, el Teorema de Incertezza [80, 2] dice que es imposible reducir indefinidamente esta dimensión, existiendo una constante positiva universal C tal que: $4 \Delta_g \Delta_{\hat{g}} > C$. Dicho Teorema también prueba que $\Delta_g \Delta_{\hat{g}}$ es mínimo e igual a $\frac{1}{2}$ si y solo si g es una Gaussiana. La ventana “óptima” para localización temporal se logra utilizando como función ventana una Gaussiana, que está máximamente concentrada en el dominio tiempo y frecuencia lo que asegura buenas propiedades de localización.

Sin embargo como se puntualiza en [20, 80] la transformada de Gabor presenta algunas dificultades. Las dimensiones de los rectángulos de localización son constantes. En las altas frecuencias el número de oscilaciones en el dominio temporal de las ventanas es elevado y la información frecuencial del mapeo es nítida. En contraposición, en las bajas frecuencias, las oscilaciones son relativamente largas y no pueden caracterizarse apropiadamente. En otras palabras, para lograr una buena resolución temporal, es necesario usar una ventana temporalmente pequeña. En este caso sin embargo, se tendrá una pobre resolución frecuencial, ya que el número de muestras dentro de la ventana es pequeño. Por otra parte, buena resolución frecuencial requiere una banda amplia pero entonces se tendrá una larga respuesta temporal. En general, la Transformada de Gabor resulta eficiente cuando se trata de localizar y caracterizar eventos con patrones de frecuencia bien definida, no superpuestos y relativamente largos, respecto de la ventana de análisis. En contraposición es totalmente inapropiada para detectar detalles de corta duración, oscilaciones largas asociadas a las bajas frecuencias, o caracterizar patrones autosimilares presentes en fenómenos a distintas escalas.

2.2.4 Transformada wavelet continua

Se ha visto que al analizar una señal con una STFT, la ventana tiempo-frecuencia es rígida, en el sentido que su ancho no cambia al observar cualquier banda de frecuencia,

$$[\omega^* + \omega - \Delta_\omega, \omega^* + \omega + \Delta_\omega], \quad (2.21)$$

con centro en $\omega^* + \omega$.

Una alternativa a la STFT es utilizar ventanas moduladas, pero de dimensión variable, ajustada a la frecuencia de oscilación. Esto sugiere, contar con una única ventana modulada y generar una completa familia de funciones elementales mediante sus dilataciones o contracciones y traslaciones en el tiempo,

$$\psi(t) \rightarrow \psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (2.22)$$

donde $a \neq 0$ y b son los parámetros de *escala* y *traslación* respectivamente.

Si la función $\psi \in L^2(\mathbb{R})$ satisface la siguiente condición de “admisibilidad”:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\widehat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad (2.23)$$

entonces ψ se llama “wavelet básica” o “wavelet madre”.

Dada una señal $s(t)$ de energía finita, la *Transformada Wavelet Continua* (CWT) de s , relativa a la wavelet básica ψ , está definida por:

Definición 2.2:

La *Transformada Wavelet Continua* de s , relativiva a la wavelet básica ψ , ($W_\psi s$), está dada por:

$$\begin{aligned} (W_\psi s)(a, b) &= \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} s(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \\ &= \frac{1}{\sqrt{|a|}} \langle s(t), \psi_{a,b}(t) \rangle \end{aligned} \quad (2.24)$$

donde $a, b \in \mathbb{R}$ con $a \neq 0$.

Si además ψ y $\widehat{\psi}$ están bien localizadas, satisfaciendo la condición (2.10), entonces la wavelet básica ψ provee una ventana tiempo-frecuencia de área $4 \Delta_\psi \Delta_{\widehat{\psi}}$. Además se deduce que $\widehat{\psi}$ es una función continua, de modo tal que la finitud de C_ψ en (2.23) implica $\widehat{\psi} = 0$, o equivalentemente,

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (2.25)$$

De este modo se ve que ψ es oscilatoria y de rápido decrecimiento, o sea que su gráfico es una “pequeña onda”, razón por la cual es llamada: “wavelet”.

Si el centro y la dispersión de la “función ventana” ψ están dados por t^* y Δ_ψ , respectivamente, la función $\psi_{a,b}$ es una ventana con centro en $b + at^*$ y dispersión igual a: $a\Delta_\psi$. De aquí que la CWT da información local de una señal s con una ventana temporal, dada por:

$$[b + at^* - a\Delta_\psi, b + at^* + a\Delta_\psi] . \quad (2.26)$$

Esta ventana se afina para pequeños valores de a y se amplía para valores grandes. Supongamos que su transformada $\hat{\psi}$ se localiza en una banda bilátera $0 < \omega_1 < |\omega| < \omega_2$, de ancho Δ_ψ . Entonces, la CWT da información local de una señal s con una ventana frecuencial $0 < \omega_1/a < |\omega| < \omega_2/a$. Se observa que la ventana rectangular tiempo-frecuencia automáticamente se afina para detectar fenómenos de alta frecuencia (es decir, pequeño $a > 0$), y se amplía para investigar comportamientos de baja frecuencia (es decir, grandes $a > 0$). Es decir que la precisión en tiempo es inversamente proporcional a la precisión en frecuencia, manteniéndose constante la relación $\Delta_\psi\Delta_{\hat{\psi}}$. Esta es la diferencia fundamental con la STFT. Más aún, para cada valor de a , la familia de wavelets parametrizada por b , se comporta como una ventana deslizante, de ancho de banda constante, pero el número de oscilaciones de estas ondas elementales es siempre el mismo, en el marco efectivo de la ventana.

La CWT preserva la energía de la señal, y posee una fórmula integral de inversión. Si la wavelet básica es real, la reconstrucción se realiza como:

$$s(t) = \frac{1}{C_\psi} \int_0^\infty \frac{1}{\sqrt{a}} \int_{-\infty}^\infty (W_\psi s)(a, b) \psi\left(\frac{t-b}{a}\right) db da , \quad (2.27)$$

donde C_ψ es la constante positiva definida en (2.2.4).

Otra propiedad relevante de la transformada wavelet continua es su invariancia respecto de traslaciones o cambios de escala de la señal. Estructuras similares serán detectadas de la misma forma, independientemente de su localización temporal o de escala. En suma, por sus propiedades la Transformada Wavelet Continua, constituye una promisoria y ventajosa alternativa para el procesamiento de señales, en particular las señales biológicas. Sin embargo, para estas aplicaciones debe resolverse la cuestión de la implementación numérica de esta transformada integral. Más aún, debe formularse un apropiado esquema para la interpretación de la información contenida en los valores $(W_\psi s)(a, b)$. Estas cuestiones son análogas a las planteadas en el caso de la STFT y conducen a analizar el caso de la llamada Transformada Wavelet Discreta.

2.2.5 Transformadas wavelet discreta

La transformada wavelet continua mapea una función $s(t)$ de una variable independiente t , en una función $(W_\psi s)(a, b)$ de dos variables independientes a y b . Desde un punto de vista computacional esta transformación resulta redundante para muchos propósitos prácticos. Si bien la transformada wavelet continua tiene algunas ventajas, hay muchas aplicaciones donde es deseable un análisis wavelet computacionalmente más eficiente y simple.

La versión discreta de la Transformada Wavelet (WT), esencialmente consiste en definir una apropiada red discreta de parámetros $\{(a_j ; b_{jk})\}$, de escalas y traslaciones,

respectivamente. Por ejemplo la red *diádica*:

$$a_j = 2^{-j} ; \quad b_{jk} = 2^{-j} k \quad j, k \in \mathbb{Z}. \quad (2.28)$$

Con esta elección de los parámetros, las wavelets serán:

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) \quad j, k \in \mathbb{Z}. \quad (2.29)$$

Asumiendo que la wavelet madre es real y la señal es de energía finita se define la Transformada Wavelet Discreta asociada a las $\psi_{j,k}$, como:

Definición 2.3:

La *Transformada Wavelet Discreta* (DWT) relativa a las wavelets diádicas $\psi_{j,k}$, ($DWT_\psi s$), está dada por:

$$\begin{aligned} (DWT_\psi s)(j, k) &= \int_{-\infty}^{+\infty} s(t) \psi_{j,k}(t) dt \\ &= \langle s, \psi_{j,k} \rangle \quad j, k \in \mathbb{Z}. \end{aligned} \quad (2.30)$$

Por otra parte, se tiene la fórmula de síntesis:

$$s(t) = \sum_j \sum_k d_j(k) \psi_{j,k}(t) \simeq \sum_j \sum_k \langle s, \psi_{j,k} \rangle \psi_{j,k}(t), \quad (2.31)$$

para apropiados coeficientes wavelets $d_j(k)$.

Si $\psi(t)$ genera una base ortonormal de wavelets, entonces $d_j(k) = \langle s, \psi_{j,k} \rangle$ y la fórmula (2.31) es exacta. Es decir que los valores de la transformada, o sea los coeficientes wavelets resumen la información de la señal, en forma análoga al caso continuo. Entonces, el espectro de coeficientes $d_j(k)$ parametrizado por (j, k) , reemplaza al mapeo continuo en el dominio.

A semejanza del caso continuo, se ve ahora que las wavelets $\psi_{jk}(t)$ están localizadas en el intervalo centrado en $(t^* + k)2^j$, de longitud $2^j\Delta_\psi$ y en la banda bilátera, de ancho $2^j\Delta_\psi$. Estas bandas representan una partición en niveles o en octavas del dominio de las frecuencias. Bajo este punto de vista, analizar una señal por medio de la DWT consiste en descomponer la misma en un banco de filtros pasabanda y en cada octava, caracterizar el comportamiento en el tiempo.

2.2.6 Análisis multirresolución

Es posible desarrollar una forma de definir wavelets a partir del concepto de análisis de multirresolución introducido por Mallat [79] y Meyer [88]. Se llama análisis de multirresolución de $L^2(\mathbb{R})$ a una sucesión decreciente de subespacios cerrados V_j , $j \in \mathbb{Z}$,

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots, \quad (2.32)$$

que satisface las siguientes propiedades:

1. $\bigcap_{-\infty}^{+\infty} V_j = \{0\}$.
2. $\bigcup_{-\infty}^{+\infty} V_j$ es denso en $L^2(\mathbb{R})$.
3. $s(t) \in V_0 \Leftrightarrow s(2^j t) \in V_j; \quad \forall j \in \mathbb{Z}$ y $s \in L^2(\mathbb{R})$.
4. existe $\phi(t) \in V_0$ tal que $\{\phi(t - k)\}, k \in \mathbb{Z}$ es una base Riesz de V_0 . En el sentido que el subespacio lineal generado por $\{\phi(t - k)\}, k \in \mathbb{Z}$ es denso en V_0 y además existen constantes positivas $0 < A \leq B$, tales que

$$A \|c\|_{l^2}^2 \leq \left\| \sum_k c_k \phi(t - k) \right\|^2 \leq B \|c\|_{l^2}^2 , \quad (2.33)$$

para todo $c = \{c_k\} \in l^2$.

Se demuestra que si $\{\phi(t - k), k \in \mathbb{Z}\}$ es una base Riesz de V_0 , entonces cada $f \in V_0$ puede desarrollarse en serie

$$s(t) = \sum_{k=-\infty}^{\infty} \langle s, \phi(t - k) \rangle \phi^*(t) , \quad (2.34)$$

donde $\{\phi^*\}$ es la base dual de $\{\phi\}$. La función ϕ que da origen a la base de Riesz indicada en el punto 4) es llamada *función de escala* del análisis de multiresolución.

En este esquema una señal $s(t) \in L^2(\mathbb{R})$ puede aproximarse tanto como se quiera por elementos de los subespacios V_j , según la propiedad enunciada en el punto 2). Pero por otra parte, cuando j decrece, las proyecciones sobre V_j podrán tener energía arbitrariamente pequeña, como dice 1). Lo que no está descripto, es la más importante propiedad intrínseca de estos espacios: más y más “variaciones” de la proyección de f sobre V_j son removidas a medida que $j \rightarrow -\infty$. En efecto estas “variaciones” son extraídas, nivel por nivel en orden decreciente de la “variación” y almacenadas en subespacios complementarios.

Una pequeña modificación de lo descripto hasta ahora permitirá interpretar este proceso que conducirá a una descomposición del espacio $L^2(\mathbb{R})$: si llamamos W_j al complemento ortogonal de V_j relativo a V_{j-1} , es decir:

$$W_j \oplus V_j = V_{j+1} , \quad (2.35)$$

resulta

$$L^2(\mathbb{R}) = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots . \quad (2.36)$$

Si existe una función $\psi(t) \in L^2(\mathbb{R})$ tal que

$$\{\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k)\} , \quad (2.37)$$

sea una base de Riesz de los subespacios W_j , se obtiene entonces un desarrollo en serie de wavelets:

$$s(t) = \sum_j \sum_k d_j(k) \psi_{j,k}(t), \quad (2.38)$$

para un conjunto de coeficientes $d_j(k)$.

Interpretamos esta descomposición: variar el índice j significa pasar de la aproximación a la señal que se puede obtener en V_j , a la que puede obtenerse en V_{j-1} ($V_{j-1} \subset V_j$). Como $W_{j-1} \oplus V_{j-1} = V_j$, las frecuencias más altas están en la parte W_{j-1} de V_j , mientras que las más bajas están en V_{j-1} . Este proceso puede continuarse lo que da una interpretación de la descomposición mencionada. En algunos casos $\{\psi_{j,k}\}$ resulta ser además una familia ortonormal, en cuyo caso:

$$\sum_{j,k \in \mathbb{Z}} |d_j(k)|^2 = \|f\|_2^2 \quad f \in L^2(\mathbb{R}). \quad (2.39)$$

- Supongamos ahora que se debe analizar una señal discreta. Veamos cómo se interpreta en este caso el análisis multirresolución:

Sea $s = \{s_k; k = 1, \dots, M\}$ la señal muestreada, que corresponde a una señal en tiempo continuo en el nivel de resolución más fino, es decir, al determinado por el intervalo de muestreo, y que por convención tomaremos como $j = 0$. En el análisis wavelet de la señal discreta $s \equiv s_0$ se le puede asociar también una sucesión de señales s_j de resolución cada vez menor a medida que j aumenta. Si llamamos s_j a la proyección de s sobre V_j , la podemos escribir como combinación lineal de funciones de escala:

$$s_j = \sum_k \langle s, \phi_{j,k} \rangle \phi_{j,k} = \sum_k c_j(k) \phi_{j,k}, \quad (2.40)$$

en donde,

$$c_j(k) = \langle s, \phi_{j,k} \rangle, \quad (2.41)$$

son los coeficientes de escala para la serie temporal s . Los s_j , para los distintos valores de j , son las versiones suavizadas de la señal a diferentes niveles de resolución. La diferencia entre dos aproximaciones sucesivas s_{j-1} y s_j es lo que se denomina *residual* o *residuo* de la señal. En esta forma, el residuo será la señal que “vive” en el espacio W_j y por lo tanto se puede escribir como combinación lineal de wavelets:

$$\begin{aligned} s_{j-1} - s_j &= \sum_k \langle s, \psi_{j,k} \rangle \psi_{j,k} \\ r_j &= \sum_k d_j(k) \psi_{j,k}, \end{aligned} \quad (2.42)$$

donde

$$d_j(k) = \langle s, \psi_{j,k} \rangle, \quad (2.43)$$

son los coeficientes wavelets y la señal original puede luego escribirse como

$$s = \sum_j \sum_k d_j(k) \psi_{j,k} = \sum_j r_j . \quad (2.44)$$

Por otra parte, el mapeo del factor de escala $a = 2^j$ en la frecuencia $\omega : a \rightarrow \omega = c \cdot 2^{-j}$, permite pensar a W_j como el espacio que gobierna a la octava j -ésima en el dominio frecuencial, es decir, pasar de 2^j a 2^{j+1} implica bajar una octava de frecuencias, y desde el punto de vista de los subespacios del análisis en resolución múltiple, implica pasar de la aproximación s_j a la s_{j+1} , dejando en el camino al residual r_{j+1} . El residual del nivel j está asociado con la banda de frecuencia $\Delta\omega_j \equiv B_j$, que corresponde al rango de frecuencias:

$$2^{-(j+1)} \omega_s \leq B_j \leq 2^{-j} \omega_s , \quad (2.45)$$

en que ω_s es la frecuencia de muestreo de la señal. De esta manera el análisis wavelet en resolución múltiple permite separar las diferentes bandas de frecuencias involucradas en una señal s .

En este trabajo se utilizarán las funciones spline cúbicas para definir el análisis multiresolución de $L^2(\mathbb{R})$. Este análisis multiresolución que da origen a una familia wavelet $\{\psi_{j,k}\}$ ortogonal, será el usado en las aplicaciones. Las funciones spline cúbicas son simétricas y combinan una adecuada proporción de suavidad con ventajas numéricas que las hacen las herramientas más recomendables, entre varias alternativas, para representar señales naturales. En la Fig. 2.1 se presenta una función wavelet spline cónica típica ψ y la Fig. 2.2 muestra la variación en frecuencia de la amplitud y fase de su transformada de Fourier $\hat{\psi}$ [132, 20].

2.3 Teoría de la Información

En investigaciones experimentales de procesos físicos, que serán el objeto del presente trabajo, típicamente sólo se tiene acceso al sistema a través de un dispositivo de medición que produce registros temporales de un cierto observable, es decir a través de una sucesión de datos. Desde este punto de vista un sistema genera “mensajes” y puede ser visto entonces, como una *fuente de información*. Esta observación abre la posibilidad de estudiar sistemas dinámicos desde un punto de vista muy interesante: la Teoría de Información (IT).

En su fundamental trabajo, realizado en 1948, C. E. Shannon [115] fundó la teoría de comunicación, motivado en parte por sus desarrollos en criptografía durante la Segunda Guerra Mundial. Esto lo condujo al estudio de compresión y trasmisión de señales. Su concepción básica consistió de: una *fuente de información* productora de mensajes, un *canal de comunicación* trasmisor de mensajes codificados. Posteriormente un receptor decodifica y recupera los mensajes originales. Este análisis tiene dos componentes esenciales:

- tasa de producción de información, llamada *entropía*
- tasa de trasmisión de información, llamada *capacidad del canal*

Varios años antes, en 1930, Hartley [54] ya había medido la producción de información a través del logaritmo del número de posibles mensajes.

La definición de entropía debida a Shannon, adaptó la medida de Hartley a una estructura probabilística: algunos mensajes serán mas o menos probables que otros. Shannon interpretó el logaritmo negativo de la probabilidad de un mensaje como una medida de sorpresa: cuanto más improbable sea un mensaje más informativo será cuando se lo reciba. Esta sorpresa, promediada sobre una fuente de mensajes, será la entropía. La forma funcional de la entropía de Shannon ya había sido desarrollada por Boltzman 1800 como una medida de desorden en sistemas termodinámicos. Teniendo en cuenta que IT se origina en la teoría de comunicaciones, la semejanza entre la entropía termodinámica de Boltzman y la entropía de Shannon, es absolutamente notable. Boltzman claramente no anticipó el uso que Shannon dio a la entropía.

Rápidamente surgieron los fundamentos matemáticos de la teoría de comunicación de Shannon, así como también un gran número de aplicaciones e importantes extensiones. Por ejemplo Jaynes [64] introdujo la Teoría de Información en la mecánica estadística, reformulando conceptos propios en términos de una máxima entropía de Shannon. Efectivamente, Jaynes mostró que si se elige la constante de Boltzmann como unidad de información y se identifica la entropía de Shannon con la entropía termodinámica, entonces toda la Mecánica Estadística puede ser reformulada, sin ninguna referencia a la noción de “ensamble”. Esto se logra maximizando la entropía de Shannon sujeta a las restricciones provenientes de la información *a priori* del sistema en estudio (*Principio de Máxima Entropía*) [67, 64]. Esto ayudó a explicar el rol de la probabilidad en mecánica estadística.

Las cantidades básicas utilizadas en IT son: la *entropía* S de una distribución, la *ganancia de información* de una distribución respecto de otra y la *información mutua* entre dos distribuciones.

2.3.1 Entropía de Shannon

Sea X una variable aleatoria tomando valores x_1, \dots, x_N . Para $i = 1, \dots, N$, se indicará con p_i , a la probabilidad de que X adquiera el valor particular x_i ; así mismo, se llamará P a la distribución de probabilidad $P = \{p_1, \dots, p_N\}$.

- **Definición Axiomática**

La noción de probabilidad para describir una situación implica, claramente, una “incertezza”, la que es descripta a través de una distribución de probabilidad. Algunas distribuciones indican obviamente, mayor incertezza que otras; esto es, algunas generan más dudas respecto al resultado del experimento que otras. Se busca una medida S que “cuantifique” esta noción de incertezza o duda.

Considerar las propiedades que naturalmente debería satisfacer la medida S , conduce a tres requerimientos básicos, a saber:

1. $S(P)$ alcanza su máximo para la distribución uniforme.
2. $S(P)$ es función continua de las probabilidades p_i .

3. $S(P)$ es independiente de la forma en que se agrupen las probabilidades de eventos individuales.

Puede demostrarse que los tres requerimientos anteriores limitan la función S , a menos de una constante multiplicativa, a la siguiente forma:

$$S(P) = - \sum_{i=1}^N p_i \ln p_i , \quad (2.46)$$

donde se supone que, $u \ln(u) = 0$ si $u = 0$. La elección de la base del logaritmo se usa para fijar las unidades de información, y si la base es 2 entonces las unidades se llamarán *bits*.

La definición axiomática de S dada, justifica que S sea considerada la *medida* cuantitativa de la cantidad de incertezza asociada a una distribución de probabilidades P . Pero la entropía de Shannon admite ser interpretada, también, desde otros puntos de vista.

• Promedio de Información

En el enfoque de teoría de comunicación, la expresión (2.46) cuantifica la cantidad de información contenida en un mensaje. En este contexto, cada mensaje i transmitido de un emisor a un receptor, es visto como un evento con una probabilidad p_i . La cantidad de información ganada por el receptor al recibir el mensaje i está dada por $I_i = \ln(1/p_i)$. La cantidad I_i es a veces referida como la “sorpresa” asociada con la salida i . Cuanto más raro sea el evento (menos probable el mensaje, más pequeño p_i), mayor será la ganancia de información cuando éste ocurre. En este sentido, S puede interpretarse como el valor de expectación de la “sorpresa”.

La expresión (2.46) será entonces la información media ganada cuando un evento ocurre. Los eventos menos probables (los mensajes menos esperados) son los que llevan más información, pero su contribución a la información media (2.46) es pequeña, precisamente porque hay poca chance de que ellos ocurran (poca chance de recibirlos): así, cuando $p_i \rightarrow 0$, entonces $\ln(1/p_i) \rightarrow \infty$, pero $(p_i \ln p_i) \rightarrow 0$.

2.3.2 Medida de Información de Fisher

Teniendo en cuenta el interés que ha despertado en los últimos años la medida de Información (I) creada por Fisher en 1925 [35, 37, 38], se han estudiado en el presente trabajo, diversas aplicaciones de la misma al análisis de señales complejas. El trabajo de Frieden, Soffer, Silver [38], así como también el de otros autores [36, 118], mostró la utilidad de la medida de información de Fisher I en las aplicaciones físicas. La importancia de I radica en su vinculación con variados aspectos de la descripción teórica de la naturaleza, y más específicamente con respecto a problemas de evolución. Así, I está involucrada en la descripción de ecuaciones de evolución de distinto tipo, que van desde la ecuación de Dirac, hasta la ecuación de Fokker-Planck. Siguiendo el enfoque de Hilgevoord [57], a continuación se brindará una breve descripción de la medida de información de Fisher I .

Dada una clase de distribuciones sobre un conjunto arbitrario X , se asignará únicamente a cada distribución un parámetro $\theta \in \Theta$, donde Θ es un espacio de parámetros. Así, cada

valor del parámetro θ especifica una distribución de probabilidad $P_\theta(x)$, con $x \in X$. Si se ha observado un conjunto de salidas $x = (x_1, \dots, x_N)$, luego cabe la pregunta: ¿qué puede inferirse acerca del parámetro desconocido θ ?

La respuesta es proporcionada por el método del soporte para inferencia estadística que se basa en los siguientes principios básicos:

- a) Toda la información provista por el dato x acerca del valor de θ está contenida en la función

$$L_x(\theta) \equiv P_\theta(x) . \quad (2.47)$$

- b) La relación $L_x(\theta_0)/L_x(\theta_1)$ puede interpretarse como un grado de soporte relativo, en el sentido que el dato provee soporte más fuerte para θ_0 que para θ_1 si, y tanto como, esta relación excede la unidad.

La función $L_x(\theta)$ fue introducida por Fisher en 1925 [35], quien realizó una distinción entre el estudio de $P_\theta(x)$ como una distribución de x para un dado θ y como una función de θ para un dado x . Este último punto de vista es el que corresponde a inferencia estadística.

Como cualquier función estrictamente monótona de $L_x(\theta)$ serviría al fin propuesto, puede considerarse en particular:

$$Sop_x(\theta) = \ln L_x(\theta) . \quad (2.48)$$

Ahora bien, si se supone que el experimento es descripto por la distribución $P_{\theta_0}(x)$, ¿qué forma de la función soporte se debería esperar? La función soporte esperado puede escribirse como:

$$\langle Sop(\theta) \rangle_{\theta_0} = \int P_{\theta_0}(x) Sop_x(\theta) dx = \int P_{\theta_0}(x) \ln P_\theta(x) dx . \quad (2.49)$$

Una medida *inversa* de la amplitud de la función, es la curvatura de la misma alrededor de su valor máximo:

$$I(\theta_0) = -\frac{d^2}{d\theta^2} \langle Sop(\theta) \rangle_{\theta_0} \Big|_{\theta=\theta_0} . \quad (2.50)$$

Por lo tanto $1/\sqrt{I(\theta_0)}$ da una medida *directa* de la incertezza esperada en nuestra inferencia. Fisher la interpretó como “la cantidad de información anticipada desde nuestra observación”, o en otras palabras, la “precisión intrínseca de la distribución”. Actualmente esta cantidad es conocida como la “información de Fisher”.

Bajo simples condiciones de regularidad, la medida de información de Fisher $I(\theta_0)$, puede escribirse como:

$$I(\theta_0) = \int \frac{1}{P_\theta} \left(\frac{dP_\theta}{d\theta} \right)^2 dx \Big|_{\theta=\theta_0} = 4 \int \left(\frac{d\sqrt{P_\theta}}{d\theta} \right)^2 dx \Big|_{\theta=\theta_0} . \quad (2.51)$$

El caso especial de familias de traslación merece una mención especial. Efectivamente, estas son familias monoparamétricas con distribución de la forma: $P(x-\theta)$ las cuales son

conocidas a menos del parámetro de corrimiento θ . Todos los miembros de esta familia poseen una forma idéntica y entonces llamando $P(x)$ a $P_\theta(x)$, la medida de información de Fisher, que ahora se denotará I , adopta la siguiente forma:

$$I = \int (P'(x))^2 / P(x) dx . \quad (2.52)$$

La información de Fisher I está ligada, en algún sentido, a la medida del error “esperado” en un proceso de medición. Si la estimación es corregida “sobre el promedio”, entonces el error medio cuadrático ε^2 de las mediciones satisface la relación:

$$\varepsilon^2 \cdot I \geq 1 . \quad (2.53)$$

La ecuación (2.53), es la conocida desigualdad de Cramer-Rao y expresa una reciprocidad entre el error ε y la información de Fisher I . La cantidad I por lo tanto puede considerarse una calidad métrica del procedimiento de estimación: como la calidad aumenta (ε disminuye) a medida que I aumenta, I es llamada “medida de información”.

La medida I también puede pensarse como una *entropía*, ya que efectivamente es una medida del grado de desorden de un sistema. Un desorden grande implica falta de predictibilidad, es decir una función densidad de probabilidad $P(x)$, uniforme o “no sesgada”. Por lo tanto, su gradiente será pequeño, así como también la información de Fisher I . Inversamente, si la curva densidad $P(x)$ muestra un sesgo hacia valores particulares de x , implica poco desorden. Pero la fuerte inclinación de la curva alrededor de estos valores de x , implica un valor mayor de I .

Si comparamos I con la entropía de Shannon S , podemos hacer la siguiente observación: S , en el caso discreto, es una suma de valores de la forma $(p_i \ln p_i)$, donde $p_i = P(x_i)$, sobre los estados $i = 1, \dots, N$. Estos valores pueden ser sumados en cualquier orden, sin alterar el resultado final. Gráficamente significa que si la curva $P(x)$ sufre un reordenamiento de los puntos (x_i, p_i) , aunque la forma de la curva seguramente cambiará, el valor de S permanece constante. S es entonces una medida “global” del comportamiento de $P(x)$.

Inversamente, las características de localización de la información de Fisher se ven fácilmente a partir de su forma discreta:

$$I = \sum_n \frac{[p_{i+1} - p_i]^2}{p_i} . \quad (2.54)$$

De este modo, al ser I sensitiva a los reordenamientos, se dice que I tiene la propiedad de “localidad” [38]. Por lo tanto, mientras S es una medida *global*, I es una medida *local* del comportamiento de la distribución de probabilidad $P(x)$ [38].

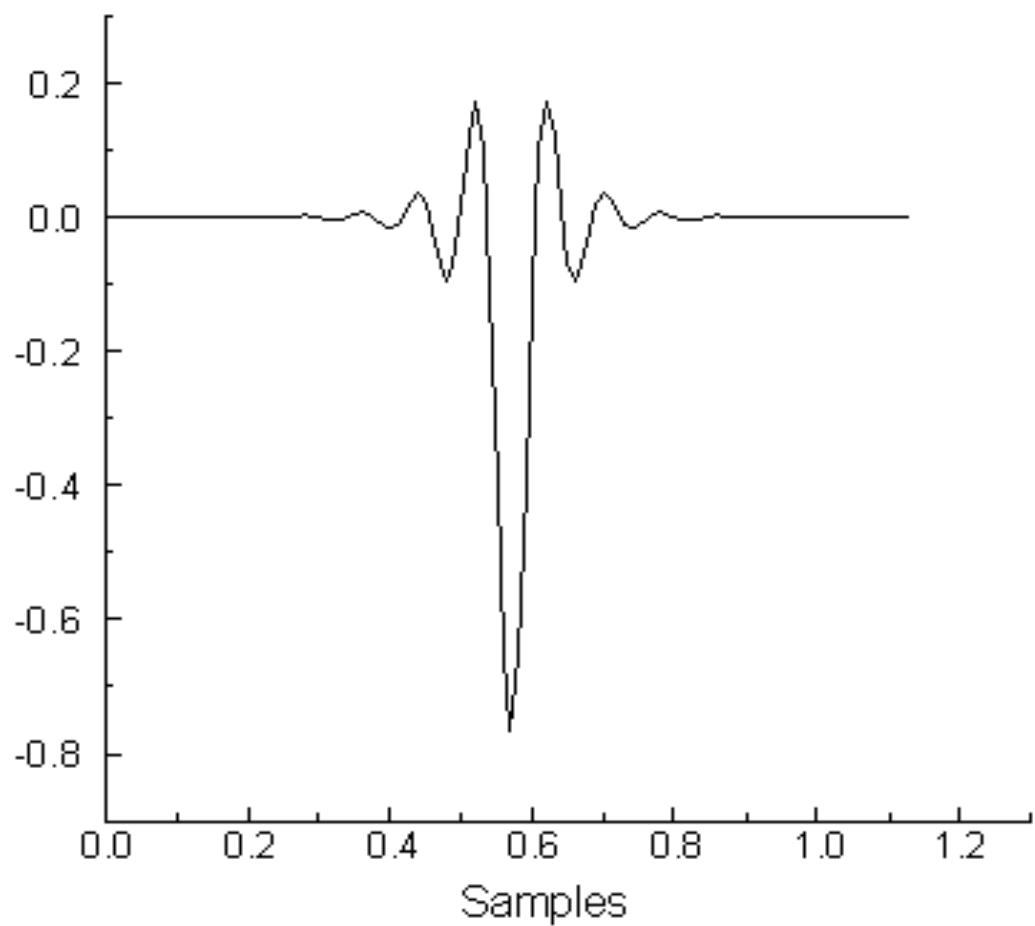


Figura 2.1: FUNCIÓN WAVELET SPLINE CÚBICA ψ UTILIZADA EN ESTE TRABAJO.

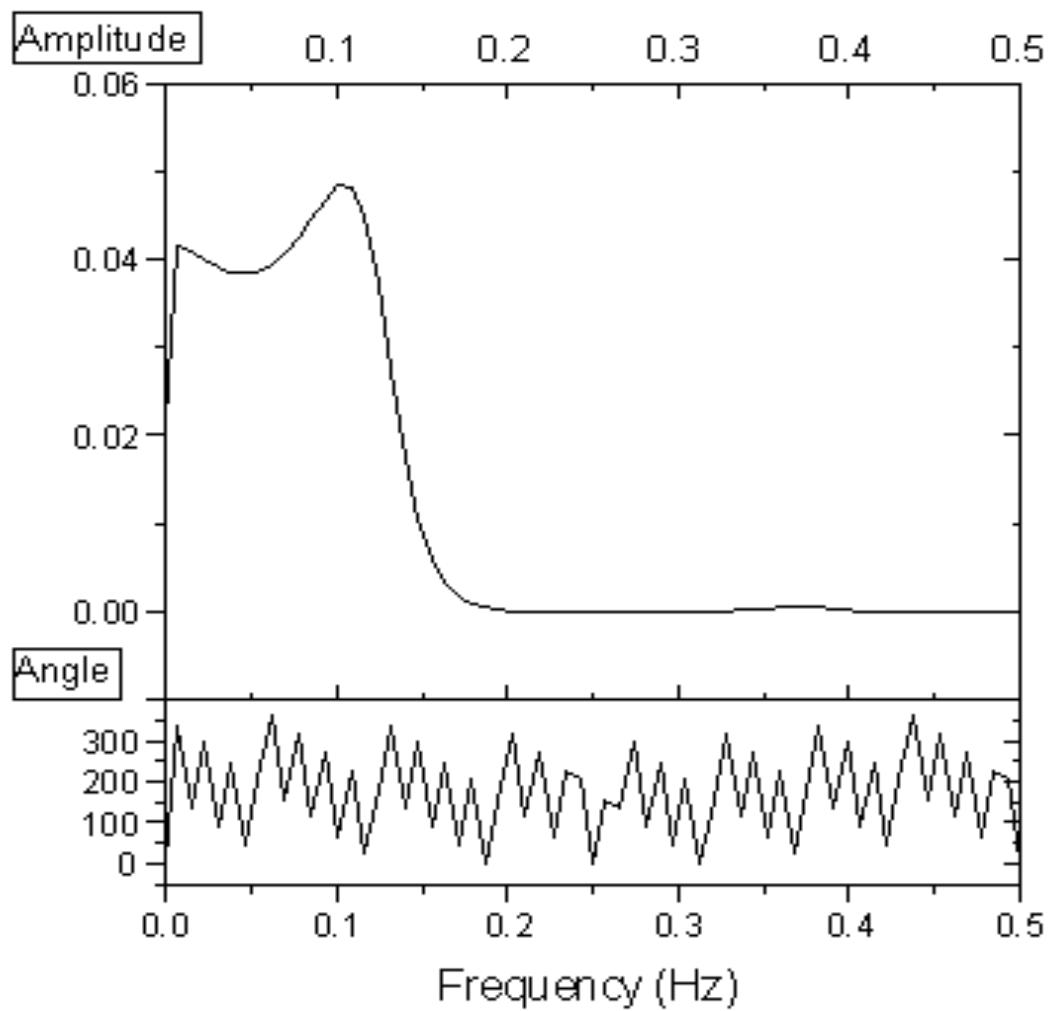


Figura 2.2: AMPLITUD Y FASE DE LA TRANSFORMADA DE FOURIER PARA LA FUNCIÓN WAVELET SPLINE CÚBICA ψ (VER FIG. 2.1).

Capítulo 3

Análisis Wavelet y Medidas de Entropía

3.1 Introducción

En este capítulo se introducirán dos recientes métodos para el tratamiento de señales complejas desarrollados en el marco del análisis multirresolución wavelet. Ellos son: la Entropía Multirresolución Total (TMS) [127] y la Entropía Wavelet Total (TWS) [113]. Estas técnicas mostraron ser especialmente convenientes para el tratamiento digital de señales no estacionarias, y en particular, para el análisis de señales electroencefalográficas (EEG).

Tanto TMS como TWS, proveen soluciones a algunos problemas básicos en la detección de eventos transitorios. Ambos métodos permiten la localización de eventos en el plano tiempo/frecuencia, y además, al usar algoritmos rápidos, posibilitan una ventajosa implementación en tiempo real. Si bien ambas técnicas se basan en la transformada wavelet, TMS lo hace a través de la señal de “detalles” correspondiente a cada nivel. Mientras que TWS introduce el concepto de entropía ligado a la energía relativa de los niveles de resolución.

Medidas clásicas para el cálculo de entropías a partir de series temporales, como son la Entropía de Amplitudes (AS) y la Entropía Espectral (SS), pueden ser vistas como inspiradoras de los métodos antes mencionados. La presentación de estas medidas en la sección 3.2 y la generalización de las mismas en el contexto de la transformada wavelet, realizada en las secciones 3.4.1 y 3.5, conducen directamente a la definición de TMS [127] y TWS [113].

Las últimas secciones del capítulo estarán dedicadas a la discusión de los resultados obtenidos cuando se aplicaron estas técnicas a series temporales provistas por sistemas dinámicos no lineales. Las aplicaciones se dividirán en dos clases. La primera se referirá a series provenientes de mapas discretos y será presentada en la sección 3.6.1, mientras que en la segunda se tratarán señales originadas a partir de modelos continuos como puede verse en la sección 3.6.2.

3.2 Entropías de Series Temporales

Desde distintos campos de la ciencia, ha surgido en los últimos años la necesidad de contar con indicadores que reflejen el comportamiento de una señal en relación a determinados aspectos de la misma. Para satisfacer este requerimiento, se han utilizado frecuentemente técnicas clásicas que cuantifican información específica contenida en la señal. Entre ellas se encuentran: la *entropía de amplitudes* (AS) y la *entropía espectral* (SS).

Para la aplicación de ambas medidas, ningún dato adicional más que la serie misma, es necesario. Así por ejemplo, la entropía de amplitudes es calculada directamente en el dominio temporal. Mientras que la entropía espectral es evaluada en el dominio frecuencial, luego de realizar la transformada de Fourier de la señal.

El concepto subyacente a estas medidas es muy simple. La entropía de amplitudes AS cuantifica el grado de orden/desorden presente en la distribución de amplitudes de la señal. La entropía espectral SS, por su parte, estima el grado de orden/desorden de la distribución de frecuencias de su espectro de potencias. Si la potencia está concentrada en una única frecuencia, la entropía espectral alcanzará su mínimo, mientras que adquirirá valores elevados si la potencia se dispersa en varias frecuencias. En ambos casos, la información provista por estas medidas tiene un carácter “global”, son cuantificadores que carecen de “localización”: ya sea frecuencial, como en el primer caso, o temporal, como en el segundo.

3.2.1 Entropía de amplitudes

Dada una serie temporal discreta:

$$s = \{s_k; k = 1, \dots, M\}, \quad (3.1)$$

donde $k, M \in \mathbb{N}$ se definirá una distribución de probabilidad sobre sus amplitudes. Sea $\mathcal{A}_N = \{I_1, \dots, I_N\}$ una partición uniforme del rango de amplitudes de s . Si s_{min} y s_{max} denotan la mínima y máxima amplitud de s respectivamente, entonces I_j , $j = 1, \dots, N$, serán subintervalos disjuntos de la recta real de longitud $\epsilon = (s_{max} - s_{min})/N$, cuya unión es igual $[s_{min}, s_{max}]$.

Se llamará p_j^A a la frecuencia de ocurrencia de la señal s en el j -ésimo intervalo de la partición \mathcal{A}_N , para $j = 1, \dots, N$. Es decir que:

$$p_j^A = \frac{\#\{s_k, \text{ tal que } s_k \in I_j\}}{M} \quad j = 1, \dots, N, \quad (3.2)$$

donde $\#$ simboliza el cardinal del conjunto.

Definición 3.1:

Se llama *distribución de probabilidad de amplitudes* de la señal finita s , a la probabilidad discreta $P_A = (p_1^A, \dots, p_N^A)$. Donde p_j^A está definido por la ec. (3.2), para $j = 1, \dots, N$.

Utilizando ahora el concepto de entropía estadística desarrollado en la sección 2.3.1, se aplicará la medida de Shannon (2.46) para definir la entropía de amplitudes de s , AS:

Definición 3.2:

Se define la *entropía de amplitudes* de s , AS, como:

$$\text{AS}(s) \triangleq S(P_A), \quad (3.3)$$

donde S es la entropía de Shannon (2.46).

Claramente AS, no es una función del tiempo, sino que asigna un valor numérico a la señal completa, al dominio temporal *entero*. Será entonces, una herramienta inadecuada para la detección de eventos temporalmente transitorios. Es decir que, para el análisis de señales no estacionarias, será conveniente su generalización hacia una medida de información temporalmente localizada.

Con este objetivo se definirá en s una ventana temporal deslizante W_i , sin superposición y dependiente de un parámetro fijo $L \in \mathbb{N}$ que representa su ancho:

$$W_i^L = \{s_k; k = 1 + (i-1)L, \dots, iL\} \quad i = 1, 2, \dots, I_L, \quad (3.4)$$

siendo $I_L = [M/L]$ el número de ventanas a considerar, donde $[.]$ indica la parte entera. El centro de la ventana, estará dado por $s_{(i-1)L+[L/2]}$. Para simplificar la notación omitiremos el superíndice L .

La evolución temporal de la medida AS se logrará al considerar la señal s vista a través de ventanas deslizantes W_i . La consecuente medida de información, denominada TAS, admitirá la siguiente definición:

Definición 3.3:

Dada una serie temporal discreta s y un parámetro ancho-ventana L , se define la *evolución temporal de la entropía de amplitudes* de s , TAS $_s^L$, como:

$$\text{TAS}_s^L(i) \triangleq \text{AS}(W_i) = S(P_A^{(i)}) \quad i = 1, 2, \dots, I_L, \quad (3.5)$$

donde $P_A^{(i)}$ es la distribución de probabilidad de amplitudes de W_i y I_L es el número de ventanas a considerar.

3.2.2 Entropía espectral

La entropía espectral se define a partir del espectro de potencias de Fourier [100, 24]. Ella cuantifica el grado de orden/desorden en frecuencias y es de este modo, una medida del grado de concentración o expansión del espectro de potencias de la señal. Una actividad ordenada, representada por una señal sinusoidal monofrecuente o que varíe levemente (rango de variación pequeño) con el tiempo, se manifestará con un pico estrecho en el dominio de frecuencias. Esta concentración del espectro de frecuencias en un solo pico corresponderá a un valor de la entropía bajo. Por otra parte, si se analiza un comportamiento dinámico de actividad totalmente desordenada (generado por un

proceso aleatorio o un sistema dinámico caótico), la respuesta será una banda ancha de frecuencias, resultando así con un alto valor de la entropía.

El cálculo de esta medida espectral se basa en la estimación de la potencia relativa $\mathcal{P}(w)$, i.e. el porcentaje de la potencia total. Si llamamos \widehat{s} a la transformada de Fourier de s , la potencia total estará dada por

$$\mathcal{P}_{total} = \int \widehat{s}(w) \cdot \overline{\widehat{s}(w)} dw \quad (3.6)$$

y la potencia relativa será entonces

$$\mathcal{P}(w) = \widehat{s}(w) \cdot \overline{\widehat{s}(w)} / \mathcal{P}_{total}. \quad (3.7)$$

En 1979 Powell and Percival [100] definieron la entropía espectral, SS, como:

Definición 3.4:

Se define la *entropía espectral* de s , SS, como:

$$SS = - \int \mathcal{P}(w) \ln \mathcal{P}(w) dw. \quad (3.8)$$

donde $\mathcal{P}(w)$ es la potencia relativa de s definida en (3.7).

En el caso de una señal discreta s , la entropía espectral se escribe como:

$$SS = - \sum_i \mathcal{P}(w_i) \ln \mathcal{P}(w_i). \quad (3.9)$$

La aplicabilidad de este método tiene importantes restricciones. La transformada de Fourier (FT) requiere la periodicidad de la señal y además, no contempla la evolución temporal de los patrones de frecuencia; en consecuencia, la entropía espectral no está definida como una función del tiempo. Estas desventajas pueden ser parcialmente resueltas por el uso de una Transformada de Fourier “tiempo-corto” (STFT) o Transformada de Fourier con ventana (WFT) (ver Sec. 2.2.3). Powell and Percival [100] definieron la evolución temporal de la entropía espectral a partir de la WFT utilizando una ventana Hanning. Con esta aproximación, se aplica la FT a ventanas temporales de unos pocos segundos de duración, de modo tal que los datos son re-definidos con una función apropiada. Luego, la evolución de las frecuencias puede ser seguida y el requisito de estacionariedad queda parcialmente satisfecho. Una limitación crítica aparece, cuando se definen las ventanas de datos, debido al Principio de Incerteza [80, 2]. Si la ventana es muy angosta, la resolución en frecuencia será muy pobre; si la ventana es muy ancha, la localización temporal será menos precisa. Esta limitación resulta importante cuando la señal tiene componentes transitorias localizadas en el tiempo.

3.3 Entropías y Análisis Wavelet

Una serie temporal puede pensarse como la representación de un sistema en una frecuencia determinada, la correspondiente frecuencia de muestreo. Por lo tanto, si

una medida de información es calculada directamente sobre la representación temporal de la señal, como sucede con la entropía de amplitudes TAS, la interpretación de sus resultados se referirá exclusivamente a la mencionada frecuencia. Si por el contrario, se aplica la entropía espectral SS, surgirán serias dificultades sobre todo al estudiar su evolución temporal en series no estacionarias. Estas observaciones sugieren dos naturales interrogantes:

- a) ¿Será posible definir una medida que cuantifique temporalmente el orden/desorden de las amplitudes de la señal, como lo hace TAS, pero que a su vez tenga localización en *frecuencias*?
- b) ¿Será posible definir una medida que cuantifique el orden/desorden en frecuencias de la señal, como lo hace SS, pero que a su vez tenga localización en el *tiempo*?

Es importante resaltar las ventajas que en este sentido proporciona, como herramienta alternativa, la transformada wavelet. La aplicación de la transformada wavelet necesita como dato de entrada, solamente la serie temporal misma y por otra parte no requiere estacionariedad del registro. Además, el análisis multirresolución al permitir estudiar la evolución de patrones de frecuencias con óptima resolución tiempo/frecuencia, crea el marco propicio para la discusión de los anteriores interrogantes. En efecto:

- La aplicación de la evolución temporal de la entropía de amplitudes TAS, a la serie temporal asociada con cada nivel de resolución, proveerá una nueva herramienta que adiciona localización frecuencial a la ya lograda, en TAS, localización temporal.
- La potencia relativa de cada frecuencia, concepto básico en SS, se asociará en el contexto wavelet, con la energía relativa de cada nivel de resolución. Se mostrará que esto provee una nueva herramienta que adiciona localización temporal a la ya lograda, en SS, localización frecuencial.

3.4 Entropía de Amplitudes con Localización Frecuencial

De acuerdo con lo propuesto, se presentará una generalización de la medida de información definida en la sección 3.2.1, TAS. Esta generalización basada en la transformada wavelet discreta, estará ligada a los métodos presentados en 1995 por Torres y coautores bajo el nombre de Entropía Multirresolución [127, 128, 129, 130, 17].

3.4.1 Entropía multirresolución total (TMS)

Dada una señal $s(t) \in L^2(\mathbb{R})$, se indicará con s su evolución discreta y se considerarán los coeficientes wavelet correspondientes a un análisis multirresolución de la misma. Se supone que la cantidad total de datos a analizar, correspondientes al nivel de resolución 0, es 2^M . El número de coeficientes en el nivel j será entonces $K_j = 2^{(M+j)}$, para $j = -1, \dots, -M$, debido a la decimación en la descomposición wavelet.

Sea:

$$D_j = \{d_j(k); k = 1, \dots, K_j\} \quad (3.10)$$

el conjunto de coeficientes wavelet correspondientes al nivel j del análisis multirresolución. Para cada nivel j , se considerarán ventanas deslizantes dependientes de un parámetro L y definidas por:

$$W_{(i,j)}^L = \{d_j(k); k = 1 + (i-1)2^{L+j}, \dots, i2^{L+j}\} \quad (3.11)$$

donde $i = 1, 2, \dots, I_L$, siendo $I_L = 2^{M-L}$ el número de ventanas a considerar. Queda claro en la ec. (3.11), que el ancho-ventana en el nivel j es $L_j = 2^{L+j}$. Por lo tanto, la ventana deslizante $W_{(i,j)}^L$ “mira en cada nivel j ” los coeficientes wavelet correspondientes al i -ésimo segmento de la señal original.

La Entropía Multirresolución Total (TMS) de la señal discreta s se obtiene considerando la evolución temporal de la entropía de amplitudes (TAS), definida según Def. 3.3, en cada uno de los niveles del análisis multirresolución [127, 128, 129, 130, 17]. Este método puede expresarse según la siguiente definición:

Definición 3.5:

Dada una señal discreta s y un parámetro L , que determinará ventanas de ancho 2^L , se define la *entropía multirresolución total* de s , TMS_s^L , como:

$$\text{TMS}_s^L(i, j) \triangleq \text{TAS}_{D_j}^{L_j}(i) = \text{AS}(W_{(i,j)}) = S(P_A^{(i,j)}), \quad \text{para } i = 1, \dots, I_L, \quad (3.12)$$

donde $j = -1, -2, \dots$ y D_j es el conjunto de coeficientes wavelet correspondiente al nivel j del análisis multirresolución. $I_L = 2^{M-L}$ es el número de ventanas a considerar, $L_j = 2^{L+j}$ y $P_A^{(i,j)}$ es la distribución de probabilidad de amplitudes de $W_{(i,j)}$.

La medida $\text{TMS}_s^L(i, j)$ representa la evolución temporal de la entropía de amplitudes de los coeficientes wavelet en una escala fija $a = 2^{-j}$ y en ventanas temporales deslizantes de ancho 2^L . Si para cada nivel j se grafica la TMS vs. el tiempo, considerando los puntos $(i2^L + 2^{L-1}, \text{TMS}_s^L(i, j))$, para $i = 1, \dots, I_L$, se obtiene un enfoque que tiene las propiedades de localización de las wavelets satisfaciendo así el planteo presentado en el interrogante a) de la sección 3.3.

3.5 Entropía Espectral con Localización Temporal

De acuerdo con lo propuesto en la sección 3.3 se presentará aquí una generalización de la entropía espectral SS, definida en la sección 3.2.2, que está basada en la transformada wavelet discreta y que su versión original fue introducida por E. Serrano en 1996 [113].

3.5.1 Energía y decimación ortogonal

Cuando la familia wavelet $\{\psi_{j,k}(t) = \psi(2^j t - k)\}$ es una base ortonormal en $L^2(\mathbb{R})$, el concepto de energía está ligado a la noción usual proveniente de la teoría de Fourier. Si $s(t) \in L_2(\mathbb{R})$ sus coeficientes wavelet estarán dados por $d_j(k) = \langle s(t), \psi_{j,k}(t) \rangle$, y la energía de la señal será:

$$\|s(t)\|_2^2 = \sum_{j,k} |d_j(k)|^2 = \sum_{j,k} |\langle s(t), \psi_{j,k}(t) \rangle|^2, \quad (3.13)$$

cuando la descomposición wavelet está dada por

$$s(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_j(k) \psi(2^j t - k). \quad (3.14)$$

Se supone que la cantidad total de datos a analizar, correspondientes al nivel de resolución 0 de una señal discreta s es 2^M . El número de coeficientes en el nivel j será $K_j = 2^{(M+j)}$, para $j = -1, \dots, -M$, debido a la decimación en la descomposición wavelet. Sea:

$$D_j = \{d_j(k); k = 1, \dots, K_j\} \quad (3.15)$$

la sucesión de coeficientes wavelet correspondientes al nivel j del análisis multirresolución. Entonces la energía de cada nivel j estará dada por:

$$E_j = \sum_{k=1}^{K_j} |d_j(k)|^2 = \|D_j\|_2^2. \quad (3.16)$$

Si ahora se consideran los coeficientes wavelet correspondientes a los M niveles de resolución, se tendrá el siguiente conjunto D^M :

$$D^M = \bigcup_{j=-1}^{-M} D_j = \{d_j(k); k = 1, \dots, K_j; j = -1, \dots, -M\}. \quad (3.17)$$

En consecuencia la energía total podrá ser obtenida por:

$$E_{tot} = \sum_{j=-1}^{-M} \sum_{k=1}^{K_j} |d_j(k)|^2 = \sum_{j=-1}^{-M} E_j = \|D\|_2^2. \quad (3.18)$$

Luego, normalizando los valores de la energía se obtiene:

$$p_j^{\text{ME}} = E_j / E_{tot}, \quad j = -1, \dots, -M, \quad (3.19)$$

que representan la *Energía Multiescala* (ME). Obviamente se cumple que:

$$\sum_{j=-1}^{-M} p_j^{\text{ME}} = 1. \quad (3.20)$$

O sea que, conceptos vinculados a la energía, permiten asociar una distribución de probabilidades al conjunto de coeficientes wavelet de una señal discreta s . Ella responde a la siguiente definición:

Definición 3.6:

Sea D^M el conjunto de coeficientes wavelet correspondientes a los M niveles de resolución de una señal discreta s de longitud 2^M . Se llama distribución de *energía multiescala asociada a D^M* a la probabilidad discreta $P_{\text{ME}} = \{p_{-1}^{\text{ME}}, \dots, p_{-M}^{\text{ME}}\}$, definida por:

$$p_j^{\text{ME}} = \frac{\|D_j\|_2^2}{\|D^M\|_2^2}, \quad j = -1, \dots, -M. \quad (3.21)$$

donde D_j representa al conjunto de coeficientes del nivel j .

3.5.2 Entropía wavelet

El concepto de entropía estadística definido en la sección 2.3.1, da una vez más, un criterio útil para analizar y caracterizar la distribución de probabilidad definida en la sección anterior. Aplicando entonces la medida de Shannon (2.46) se define la entropía wavelet, WS, de la siguiente forma:

Definición 3.7: Sea D^M el conjunto de coeficientes wavelet correspondientes a los M niveles de resolución de una señal discreta s de longitud 2^M . Se define la *entropía wavelet* asociada a D^M , WS, como:

$$\text{WS}(D^M) \triangleq S(P_{\text{ME}}), \quad (3.22)$$

donde S es la entropía de Shannon y P_{ME} es la distribución de energía multiescala asociada a D^M .

3.5.3 Entropía wavelet total (TWS)

Supongamos como siempre, que la cantidad total de datos a analizar es 2^M . Se considerarán sobre el conjunto de coeficientes D^M (ec.(3.17)), ventanas deslizantes dependientes de un parámetro L , que a diferencia de las consideradas anteriormente en la sección 3.4.1, son ahora *bidimensionales*. Si el ancho de las ventanas es 2^L , cada una de ellas admitirá L niveles de resolución. Pueden pensarse entonces, como ventanas bidimensionales de altura L y ancho 2^L , definidas por:

$$\begin{aligned} W_{(i)}^L &= \bigcup_{j=-1}^{-L} W_{(i,j)} \\ &= \{d_j(k) ; k = 1 + (i-1) 2^{L+j}, \dots, i 2^{L+j} ; \\ &\quad j = -1, \dots, -L\}, \end{aligned} \quad (3.23)$$

para $i = 1, \dots, I_L$ donde $I_L = 2^{M-L}$ es el número de ventanas a considerar. Las $W_{(i,j)} \equiv W_{(i,j)}^L$ están definidas según la ec. (3.11). Se puede observar que, la ventana bidimensional deslizante $W_{(i)}^L$, de ancho 2^L , “mira simultáneamente los coeficientes wavelet de los L ” niveles de resolución, correspondientes al i -ésimo segmento de la señal original. La Entropía Wavelet Total (TWS) de $s(t)$ se obtiene considerando la evolución temporal de la entropía wavelet (WS), definida según Def. 3.7, en el conjunto de niveles del análisis multirresolución [113]. Este método puede expresarse mediante la siguiente definición:

Definición 3.8:

Dada una señal s y un parámetro L , que determinará ventanas de ancho 2^L , se define la *entropía wavelet total*, TWS_s^L , como:

$$\text{TWS}_s^L(i) \triangleq \text{WS}(W_{(i)}^L) = S(P_{\text{ME}}^{(i)}), \quad (3.24)$$

donde $P_{\text{ME}}^{(i)}$ es la distribución de energía multiescala asociada a $W_{(i)}^L$.

$P_{\text{ME}}^{(i)}$ contiene información de la señal correspondiente a una ventana de ancho 2^L , centrada en $(2^{L-1}(2i - 1))$, por lo cual $\text{TWS}_s^L(i)$, se podría decir que es una medida entrópica de frecuencias o escalas, localizada temporalmente.

Como TWS se calcula aplicando la entropía wavelet WS a las ventanas bidimensionales de coeficientes wavelet, $W_{(i)}^L$, podemos decir que TWS cuantifica el orden/desorden correspondiente a las energías relativas de cada *banda de frecuencia* en un dominio temporal *limitado al ancho* 2^L . En este sentido se diferencia de la entropíapectral, ya que SS cuantifica el orden/desorden de las energías relativas correspondientes a cada *frecuencia individual* en el dominio temporal *completo*.

Si se grafica TWS *vs.* el tiempo, considerando los puntos $(i2^L - 2^{L-1}, \text{TWS}_s^L(i))$, para $i = 1, 2, \dots, I_L$, las propiedades de localización de las wavelets proporcionan un útil método para detectar cambios dinámicos en la señal.

Las informaciones brindadas por TMS (sección 3.4.1) y la actual TWS, lejos de ser redundantes son entre sí complementarias. $\text{TMS}_s^L(i, j)$ proporciona una información bidimensional usando ventanas unidimensionales ($W_{(i,j)}$), que se i -deslizan sobre cada nivel j . $\text{TWS}_s^L(i)$, por su parte, brinda información unidimensional pero relativa a ventanas bidimensionales ($W_{(i)}$), que se i -deslizan abarcando todos los niveles de coeficientes wavelet a la vez. Puede observarse que para cada intervalo temporal, mientras TMS analiza cada nivel de resolución por separado, TWS analiza a todos simultáneamente, lo cual permitiría captar la inter-relación que existe entre ellos.

3.6 Aplicaciones

Para investigar las perspectivas y limitaciones de los métodos que se han introducido, será ventajoso contar con la ayuda de datos artificiales donde los resultados correctos son conocidos. Por eso los métodos entrópicos presentados en este capítulo, TMS y TWS, serán aplicados a series temporales provenientes de conocidos sistemas dinámicos no lineales.

Se mostrará que las técnicas propuestas pueden determinar tanto estados de equilibrio (variación del número de puntos fijos), como también cambios topológicos. Se analizará además la caracterización de transitorios y bifurcaciones. Pequeños cambios realizados en los valores de los parámetros del modelo conducen a no estacionariedades que serán investigadas para evaluar la bondad de las medidas, siendo la detección y localización decisivos aspectos a tener en cuenta.

La variación de parámetros lleva a “transiciones de fase dinámicas”, lo cual sugiere que las transiciones de fase físicas aparecen como un producto final de estas transiciones dinámicas microscópicas. Este punto de vista es particularmente atractivo, sobre todo en el caso de la medida de información de Fisher I (sección 2.3.2) que constituye un principio unificado para la física. En este sentido, sería interesante investigar el comportamiento de I como detector de los mencionados cambios dinámicos. Según muestran numerosas publicaciones, la cantidad I está ligada a variados aspectos de la descripción teórica de la naturaleza, siendo de este modo su aplicación a sistemas dinámicos no lineales absolutamente promisoria.

Si bien las medidas entrópicas fueron presentadas en base a la entropía estadística de Shannon, es evidente que esta puede ser reemplazada por cualquier otra medida de

información adecuada. Esto conduce a generalizar la medida TAS, reemplazando en (3.5) la entropía de Shannon S , por la información de Fisher I (ver Ref. [81, 82]). Se denotará TAI a la evolución temporal de esta medida cuando esta basada en la información de Fisher I .

El comportamiento de TAS y TMS, definidas en las secciones 3.2.1 y 3.4.1 respectivamente, así como también el de la medida de información de Fisher TAI introducida en el párrafo anterior, será analizado ante:

- a) Modificación del valor de uno de los parámetros del modelo durante un *breve* período de tiempo,
- b) Modificación del valor de uno de los parámetros del modelo con *continuidad* a lo largo del dominio temporal completo.

Las series temporales serán generadas a partir de modelos clásicos, agrupándolos en: I) mapas discretos y II) modelos continuos.

3.6.1 Aplicaciones I: entropía para mapas discretos

Mapa de Henon

Se considera un mapa bi-paramétrico presentado por M. Hénon en 1976 [56]. En el presente ejemplo se estudiará la evolución temporal del sistema con referencia al comportamiento de la variable x_n , cuyos valores son registrados a intervalos discretos [17]. Las ecuaciones que definen el sistema son:

$$\begin{aligned} x_n &= 1 + y_{n-1} - a_n x_{n-1}^2 \\ y_n &= b x_{n-1}. \end{aligned} \quad (3.25)$$

En el estudio de este modelo, usualmente ambos parámetros a_n y b , son considerados constantes. En este caso, se supondrá que el parámetro a_n evoluciona de modo tal que se produce un cambio dinámico lineal del mismo, dado por:

$$a_n = \begin{cases} a_1 & \text{si } n < n_1 \\ a_1 + (n - n_1)(a_2 - a_1)/(n_2 - n_1) & \text{si } n_1 \leq n \leq n_2 \\ a_2 & \text{si } n > n_2 \end{cases} \quad (3.26)$$

En la Fig. 3.1 se muestra en línea de puntos, la evolución temporal de la variable x_n cuando los dos parámetros en (3.25) son constantes. En la misma figura pero en línea llena, se muestra la diferencia entre la anterior señal (a_n constante) y aquella obtenida imponiendo la variación lineal (3.26) del parámetro a_n , con $a_1 = 1.062371838$, $a_2 = 1.080744879$, $n_1 = 1000$ y $n_2 = 1030$, mientras que b se mantendrá constante ($b = 0.3$). Debe resaltarse que la ruptura lineal del parámetro no es perceptible en el gráfico de la señal, por lo cual se optó por mostrar en la Fig. 3.1 junto con la señal original, la diferencia de ambas señales.

En la Fig. 3.2, se muestran los resultados obtenidos cuando se aplica la medida entrópica TAS, sobre ambas señales (x_n parámetro a_n fijo y x_n parámetro a_n variable).

La evolución temporal de TAS, basada en la entropía de Shannon, fue evaluada con ventanas de ancho $L = 64$ y $N = 100$ particiones del rango de amplitudes. La Fig. 3.2 permite apreciar un incremento en el valor de TAS coincidente con la localización de la ruptura, el cual comienza precisamente cuando la ventana empieza a “ver” la ruptura en las proximidades de $n \sim 1000$. Una mínima variación del parámetro, imperceptible al comparar los gráficos de ambas señales (parámetro fijo y parámetro variable), es categóricamente detectado por la medida de amplitudes TAS.

En la Fig. 3.3, se muestran los resultados obtenidos cuando se aplica la entropía multiescalas, TMS, sobre ambas señales (x_n parámetro a_n fijo y x_n parámetro a_n variable) en tres niveles de resolución wavelet. Según el esquema propuesto por TMS, los ejemplos fueron realizados usando un análisis multirresolución basado en la *spline wavelet ortogonal*, ventanas temporales de 128 muestras ($L = 128$) y $N = 100$ particiones del rango de amplitudes. En la Fig. 3.3 se puede apreciar que TMS es un indicador de los subyacentes cambios dinámicos producidos en los distintos niveles de resolución wavelet, observando en este caso que la variación del parámetro se manifiesta casi exclusivamente en el primer nivel de resolución wavelet. Podría inferirse de dicha observación que la ruptura del parámetro a_n se traduce en cambios morfológicos de la señal asociados principalmente con el primer nivel de resolución wavelet. Es decir que TMS proporciona información adicional respecto de TAS: además de *localizar* la ruptura la *caracteriza* de acuerdo a los niveles de resolución wavelet.

En la Fig. 3.4.b se muestran los resultados obtenidos cuando se aplicó a las mismas señales de los ejemplos anteriores, la medida entrópica TAI basada en la información de Fisher I [81]. Notar el salto en las proximidades de $n_1 = 1000$. Los resultados obtenidos con la asociada medida de Shannon, ya vistos en la Fig. 3.2, son mostrados ahora en la Fig. 3.4.c en una escala diferente. En la Fig. 3.4.c se consideró la misma escala vertical que en la Fig. 3.4.a. Comparando las dos Figs. 3.4.b y c, y prestando atención a la relación entre la escala del “salto” y la de la señal, puede inferirse que la medida de información de Fisher es un mejor indicador del cambio en la dinámica subyacente. Como ya fue dicho, en todos los ejemplos se han seleccionado pequeños cambios lineales en el parámetro, que no puede detectarse por una mera inspección visual de la señal.

Mapa logístico

El mapa logístico [66] constituye un paradigmático ejemplo generalmente empleado para ilustrar nuevos conceptos en el tratamiento de sistemas dinámicos. Es un simple y bien estudiado sistema que presenta una rica variedad de comportamientos. Dependiendo del valor del parámetro r el mapa puede mostrar un punto estable, duplicación de período o caos.

Se trabajará con el mapa $F : x_n \rightarrow x_{n+1}$, descripto por el ecológicamente motivado sistema disipativo que responde a la ecuación en diferencia de primer-orden

$$x_{n+1} = r x_n (1 - x_n), \quad 0 \leq x_n \leq 1, \quad 0 < r \leq 4. \quad (3.27)$$

Se revisarán brevemente algunos resultados bien conocidos de este mapa [93, 9], que brindará claridad a la hora de analizar las propiedades de las nuevas medidas.

A medida que r crece, una muy rica y bien conocida estructura comienza a aparecer. Para poder apreciar mejor el comportamiento asintótico en r , la Fig. 3.5.a se muestra el

diagrama de órbitas. En este diagrama se han graficado los primeros 5000 valores $x_n(r)$, para cada valor r comprendido entre 3.4 y 4.0, con paso $\Delta r = 0.001$. Más explícitamente, para cada r se generá mediante (3.27) una sucesión numérica x_n (que depende obviamente de r) cuyos valores son graficados (ver Fig. 3.5.a) en la correspondiente línea vertical de abscisa r .

Para valores del parámetro de control $1 < r < 3$ existe un único estado. Aumentar el parámetro de control r por encima de $r = 3$ fuerza el sistema a una bifurcación de duplicación de período. Esto es, ciclos de período 2, 4, 8, 16, 32, ... ocurren y si r_n denota el valor de r donde primero aparece un período 2^n , el valor r_n converge hacia un valor límite $r_\infty \approx 3.569946$ [94]. Cuando r crece aun más, surge una estructura muy rica (caos), que además está muy bien estudiada. Para $r > r_\infty$ el diagrama de órbitas revela una mezcla extraña de orden y caos.

Una gran ventana comienza cerca de $r = 3.8284$ conteniendo un ciclo estable de período 3 y luego un nuevo proceso de duplicación de período comienza a aparecer, seguido por ciclos de período 6, 12, 24, Se observa caos en la etapa final del diagrama de órbitas que está representado por una irregular distribución de los puntos que verticalmente expanden el rango entero entre 0 y 1. Observando el diagrama de órbitas, se puede apreciar en la región caótica $r_\infty \leq r \leq 4$, la existencia de una estructura subyacente formando “bandas”, las cuales revelan que los puntos no están uniformemente distribuidos a lo largo de cada línea vertical. Ciertamente, los puntos tienden a reunirse alrededor de ciertas líneas (“líneas críticas” [9]) que, delimitan las “bandas” que caracterizan la región de dinámica caótica. Se puede notar también que para $r \approx 3.67$, se intersecan varias líneas críticas, y fácilmente se pueden reconocer regiones con “densidad de puntos” diferente.

En la Fig. 3.5.b se presenta el exponente de Lyapunov en función del parámetro r . Se ve que el exponente de Lyapunov λ permanece negativo para $r \leq r_\infty$. En el comienzo del caos cerca de $r \cong 3.57$, λ comienza a ser positivo, reflejando una divergencia exponencial de las trayectorias. Para $r > 3.57$, λ crece globalmente, excepto en las regiones correspondientes a las ventanas de comportamiento periódico.

Las Figs. 3.6 y 3.7 muestran la capacidad discriminatoria las medidas consideradas [82], TAI y TWS, al ser aplicadas a la señal generada por la variable x_n del mapa logístico cuando el parámetro r varía con continuidad. Más explícitamente, para cada valor de r comprendido entre 3.4 y 4.0 con paso $\Delta r = 0.001$, se generan $M = 3000$ valores x_n de acuerdo a la ec. (3.27) y estos valores se acumulan formando la señal a analizar. La evolución temporal de la señal significa entonces, la evolución del parámetro de control r , y por lo tanto la aparición de una rica variedad de comportamientos. Esto hace que sea una señal útil para caracterizar la performance de nuestras medidas TAI y TWS. Valores particulares del parámetro r correspondientes a regímenes periódicos son claramente identificados en las figuras. Respecto a la medida de información de Fisher se ve en la Fig. 3.6.b que los picos corresponden a regímenes regulares, en los que prevalece el orden por lo cual la información TAI es grande. Valles son asociados al caos (información escasa). Para $r = 3.5$ la dinámica es regular de período 4, como se ve en Fig. 3.6.a. La sucesión de duplicaciones de período que continúa es nítidamente detectada por TAI. Comenzando en $r = r_\infty \simeq 3.569$ una zona estocástica aparece, visible en Fig. 3.6.a, y traducida en una abrupta caída de la medida entrópica TAI. Adicionales detalles son detectados por TAI, como por ejemplo las tres más notables

ventanas regulares, correspondientes a órbitas de período 6, 5, y 3. El correspondiente rango de valores de r es satisfactoriamente predicho, así como el tamaño de las ventanas es adecuadamente descripto. La entropía wavelet total TWS es mostrada en la Fig. 3.7, donde se ve que los valores permanecen bajos hasta alcanzar la zona correspondiente a la sucesión de duplicaciones de período donde la entropía aumenta, continuando esta tendencia hasta la ventana de período 3. Esta ventana es claramente localizada, al igual que otras ventanas regulares menores. El comportamiento de la medida TWS se manifiesta globalmente más adecuado que el evidenciado por TAI, al distinguir diferentes grados de entropía en las zonas caóticas. TAI por el contrario, le otorga prácticamente el mismo grado de información casi nulo, a toda la zona caótica por igual.

3.6.2 Aplicaciones II: entropía para modelos continuos

Modelo de Lorenz

El modelo de Lorenz [78] fue presentado en 1963 por el meteorólogo Edward Lorenz del Instituto Tecnológico de Massachusetts. Lorenz desarrolló las ecuaciones como un modelo simple para el llamado problema de convección de Rayleigh-Bénard que gobierna la formación de remolinos de convección entre dos superficies paralelas a diferentes temperaturas. Su interés fue el de predecir la dinámica atmosférica con el objetivo final de desarrollar predicciones del tiempo de largo alcance. Las soluciones numéricas obtenidas para estas ecuaciones de apariencia muy simple, ha tenido importantes efectos en la comprensión de particulares comportamientos de las ecuaciones no lineales en general. El modelo de Lorenz fue deducido de las ecuaciones de Navier-Stokes luego de varias simplificaciones, siendo un sistema de tres ecuaciones diferenciales acopladas ordinarias de primer orden no-lineales:

$$\begin{aligned}\frac{dx}{dt} &= -\sigma(x - y) \\ \frac{dy}{dt} &= r x - y - x z \\ \frac{dz}{dt} &= -b z + x y\end{aligned}\tag{3.28}$$

donde σ es el número de Prandtl, b está relacionado con la forma de los remolinos de convección, y r está relacionado con el número de Rayleigh (o la diferencia de temperatura entre las superficies) [112]. Las variables dependientes $x(t)$, $y(t)$ y $z(t)$ contienen esencialmente dependencia temporal y distribución de temperatura expresadas como expansiones de Fourier. Una detallada discusión de la deducción de este modelo se encuentra en *Deterministic Chaos* de H.Schuster [112]. Mientras que un minucioso análisis de la sucesión de bifurcaciones puede encontrarse en Ref. [62].

En el ejemplo, con $\sigma = 16.0$ y $b = 4.0$, el parámetro de control r será variado de acuerdo a la versión continua de la ecuación (3.26), donde $r_1 = 45.92$ y $r_2 = 55.92$, $t_1 = 980$ y $t_2 = 1020$. Los exponentes de Lyapunov para el valor inicial r_1 son $\lambda_1 = 2.16$, $\lambda_2 = 0.0$ y $\lambda_3 = -32.4$ [137], es decir que nos encontramos con una dinámica caótica para el valor inicial r_1 del parámetro r . Para la integración del sistema de ecuaciones (3.28) se empleó el método de Runge-Kutta con un paso $\Delta t = 0.001$ y valores iniciales

$x(0) = 10$, $y(0) = 1$ y $z(0) = 0$. Se pueden comparar en la Fig. 3.8 la evolución temporal de dos señales $y(t)$: una (línea de puntos) con $r = r_1 = \text{cte.}$ y la otra (sólida) con r variando como fue descripto más arriba, ambas correspondientes a la coordenada y . Las Figs. 3.9 y 3.10 muestran los resultados de las medidas de información de Shannon TAS y TMS, al ser aplicadas para el análisis de las mencionadas señales. Para la evaluación de TMS se consideró el análisis multirresolución basado en la *spline wavelet ortogonal*, y se usaron ventanas temporales de 128 muestras ($L = 128$) con $N = 100$ particiones del rango de amplitudes. Los resultados multiescala mostrados corresponden a los tres primeros niveles de resolución. Notar en la Fig. 3.9 que la entropía de Shannon TAS es casi insensitiva a los cambios dinámicos. En cambio en el primer y segundo nivel de la entropía multiescala TMS, Fig. 3.10 se nota un típico *valle* en la región variación del parámetro que marca el cambio en la dinámica.

El modelo de Lorenz, con $\sigma = 10.0$ y $b = 8/3$, será tratado ahora con la medida de información de Fisher (TAI), pero con modificación del valor de r con *continuidad* a lo largo del rango $75 \leq r \leq 325$. Los resultados son mostrados en la Fig. 3.11.b y comparados con los exponentes de Lyapunov obtenidos por Froyland y Alfsen (Fig. 3.11.a) [39]. Como se puede apreciar, los cambios en la dinámica son bien detectados por TAI. Los resultados obtenidos usando la medida de Fisher TAI, en función de r , al ser comparados con aquellos obtenidos por medio de los exponentes de Lyapunov por Froyland y Alfsen [39], resultan detectar los cambios dinámicos con claridad.

Modelo de Rossler

A continuación del descubrimiento del caos de Lorenz, Rossler desarrolló un número de modelos matemáticos más simples que brindan información sobre las propiedades dinámicas mínimas que producirán caos de varias formas [104]. Estos modelos son reales construcciones matemáticas, y como consecuencia de ello, no están sujetos a intentos de asociaciones inmediatas con sistemas físicos. El primer modelo de Rossler consiste de tres ecuaciones diferenciales acopladas de primer orden:

$$\begin{aligned} \frac{dx}{dt} &= -y - x \\ \frac{dy}{dt} &= x + a y \\ \frac{dz}{dt} &= b + z (x - c) \end{aligned} \quad (3.29)$$

Puede observarse que, si $x < c$, $z(t)$ está en la vecindad del valor $b/(c - x)$, mientras que si $x > c$, $z(t)$ aumenta exponencialmente. Por lo tanto la única nolinealidad de esta dinámica involucra este “planchado” y reinyección de la dinámica $z(t)$ relativa al movimiento lineal (x, y) . Rossler ha ilustrado alguno de los aspectos de este atractor en Ref. [104].

En la Fig. 3.12.a se muestran los exponentes de Lyapunov asociados, mientras que la Fig. 3.12.b muestra la evolución temporal de la información de Fisher (TAI), para $3 < c < 6$. Vemos que el máximo exponente de Lyapunov es positivo para $4.2 \leq c \leq 5.2$. La correspondiente caoticidad es claramente detectada por la medida de Fisher (un típico valle), que también captura adicionales detalles de esta dinámica.

3.7 Conclusiones

En este capítulo se presentaron medidas entrópicas formuladas en el marco de la teoría wavelet, dando origen a los métodos de Entropía Multirresolución Total (TMS) [127] y Entropía Wavelet Total (TWS) [113]. Las propiedades de estos métodos, hacen que sean especialmente aptos para el análisis de señales no estacionarias generadas por sistemas dinámicos nolineales o provenientes de sistemas físicos experimentales.

Ambos métodos se basan en la descomposición wavelet discreta y el concepto de entropía de Shannon. El análisis multirresolución brinda la posibilidad de introducir una diferente y quizás más elaborada medida de información que incorpora de manera natural todas las ventajas del análisis wavelet. Estos métodos son apropiados para la detección de eventos transitorios en señales complejas, ya que presentan la ventaja de poder procesar la información directamente sobre la señal experimental “on line”, es decir a medida que la misma está siendo registrada. Esto les confiere una especial capacidad para funcionar adecuadamente como detector automático de eventos transitorios. Aplicaciones de las nuevas medidas a series electroencefalográficas (EEG) serán presentadas en el capítulo 7, como una ilustración experimental de las ventajosas características compartidas por ambas medidas.

Es importante resaltar una vez más, que las informaciones brindadas por TMS y TWS, lejos de ser redundantes, son entre sí complementarias. TMS proporciona una información bidimensional usando ventanas unidimensionales, que se deslizan sobre cada nivel wavelet. TWS, por su parte, brinda información unidimensional pero relativa a ventanas bidimensionales que se deslizan temporalmente sobre todos los niveles de coeficientes wavelet a la vez. Dicho de otra forma, para cada intervalo temporal (ventana), mientras TMS analiza *cada* nivel wavelet por separado, TWS analiza a *todos* los niveles simultáneamente, lo cual permite captar la interrelación que existe entre ellos.

Aunque las características de la Transformada Wavelet son especiales para su aplicación al análisis de señales experimentales, con el fin de estudiar y comparar los nuevos métodos introducidos, se realizaron aplicaciones al análisis de señales provenientes de sistemas dinámicos teóricos con resultados ya conocidos. Las nuevas técnicas detectaron en la mayoría de los casos variaciones breves introducidas en el parámetro de control del sistema, y también mostraron un razonable comportamiento ante la variación continua del parámetro.

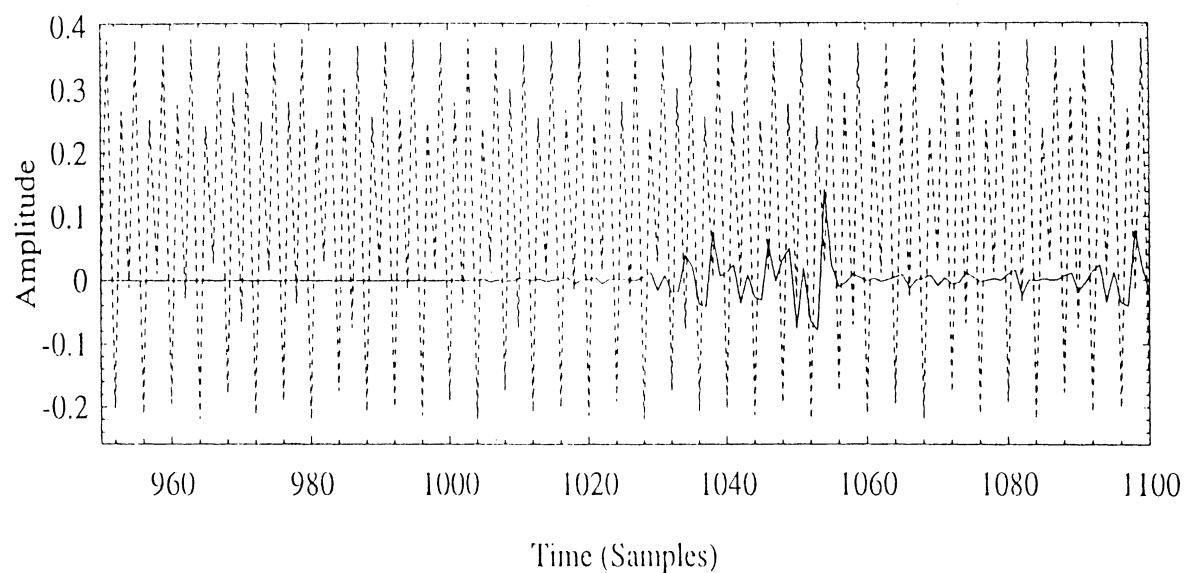


Figura 3.1: MAPA DE HENON: SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y SU DIFERENCIA CON LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA).

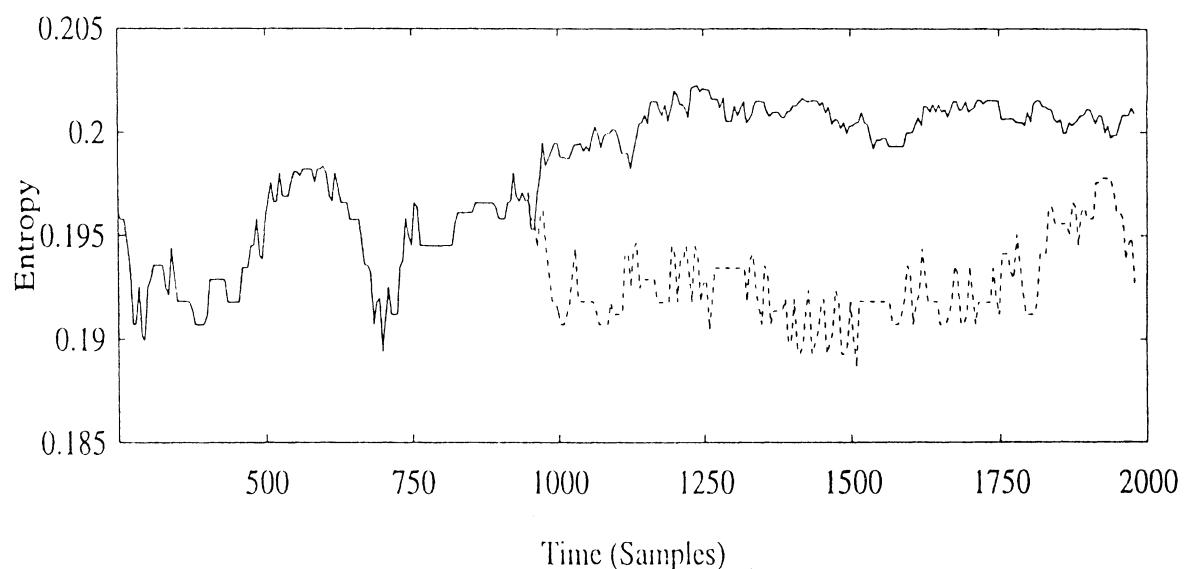


Figura 3.2: MAPA DE HENON: ENTROPIA DE SHANNON TAS, DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.

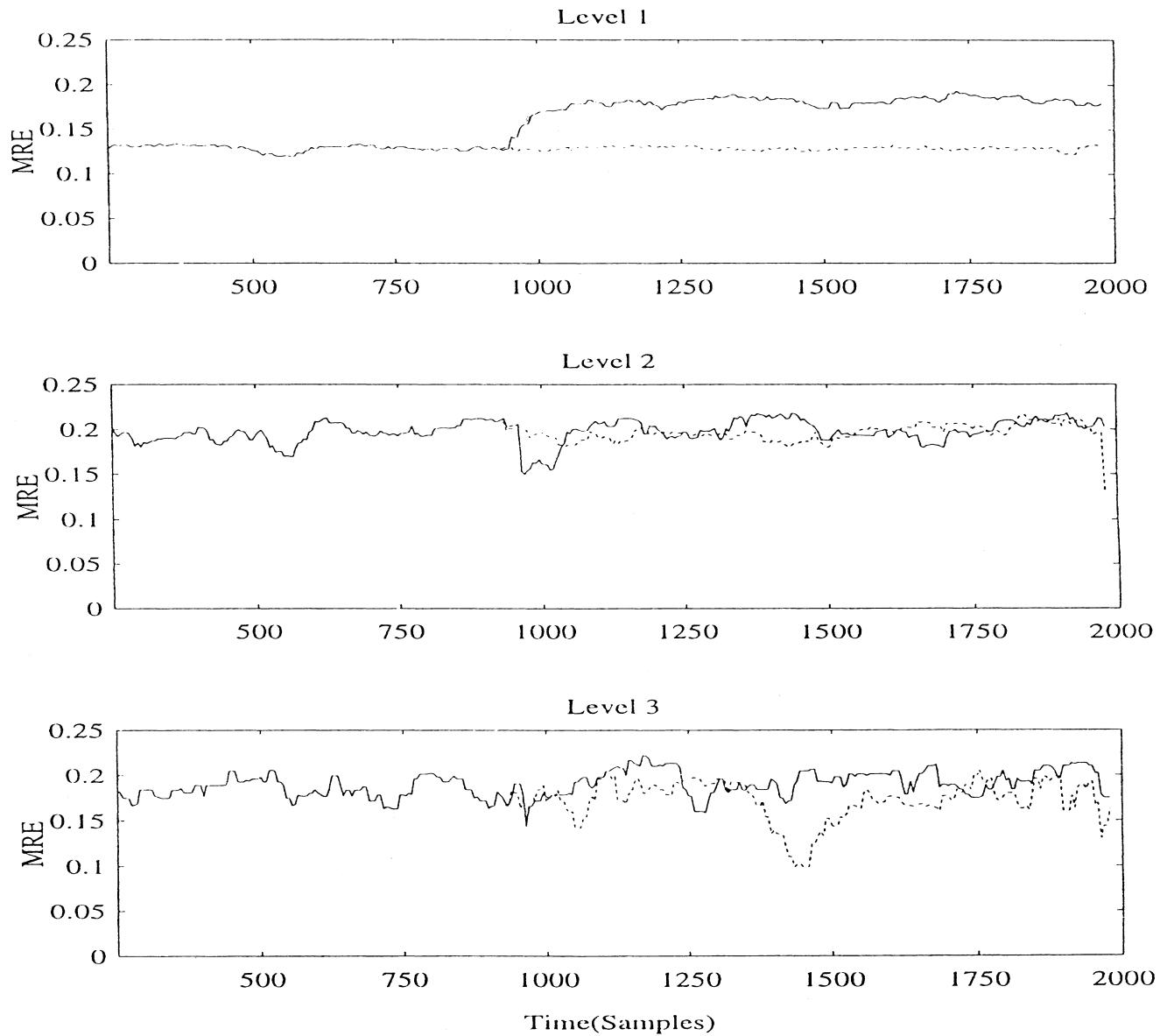


Figura 3.3: MAPA DE HENON: ENTROPÍA MULTIESCALAS TMS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE MUESTRAN LOS TRES PRIMEROS NIVELES DE RESOLUCIÓN. SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.

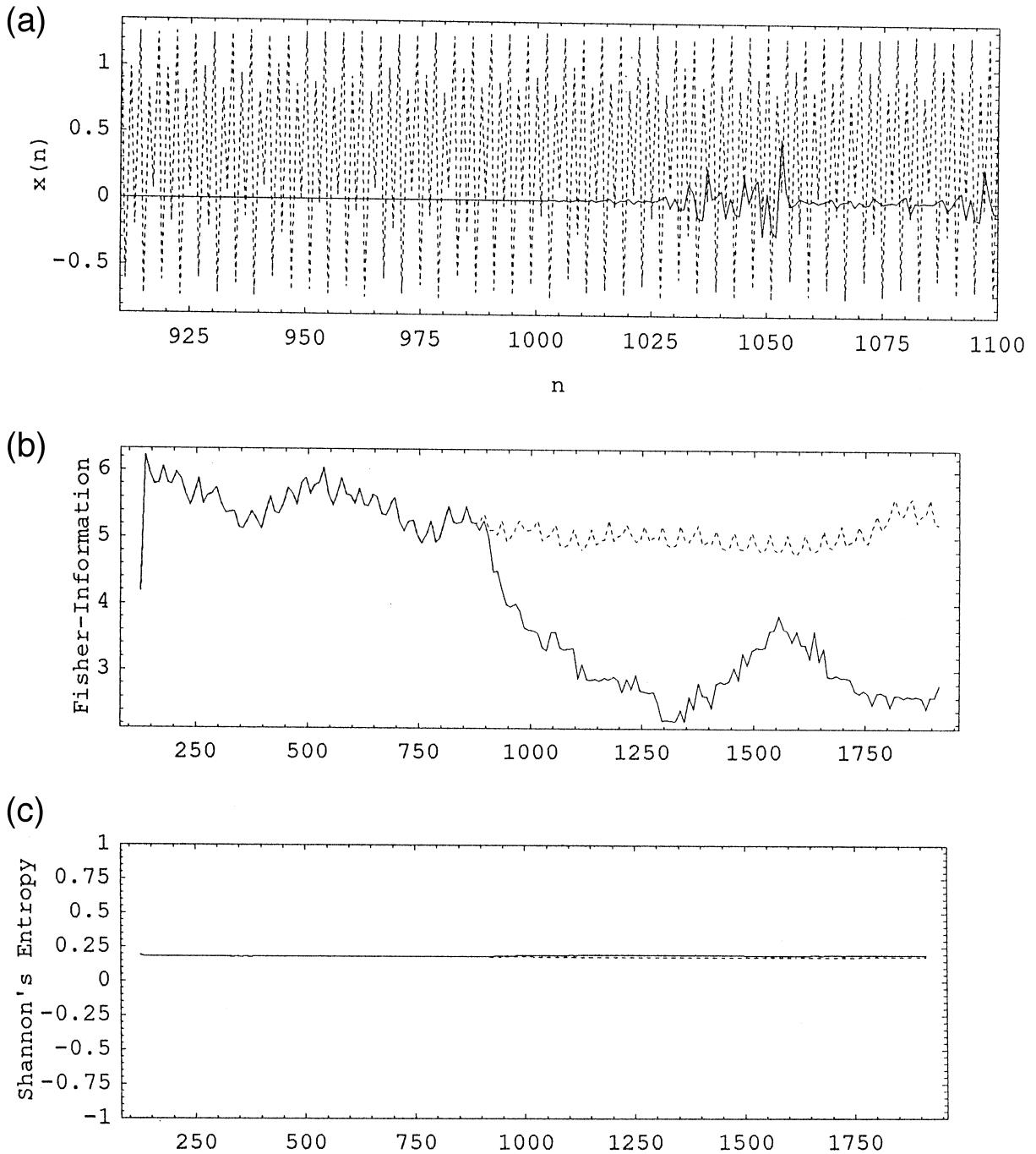


Figura 3.4: MAPA DE HENON: *a)* SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y SU DIFERENCIA CON LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). *b)* INFORMACIÓN DE FISHER TAI DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). *c)* ENTROPÍA DE SHANNON TAS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.

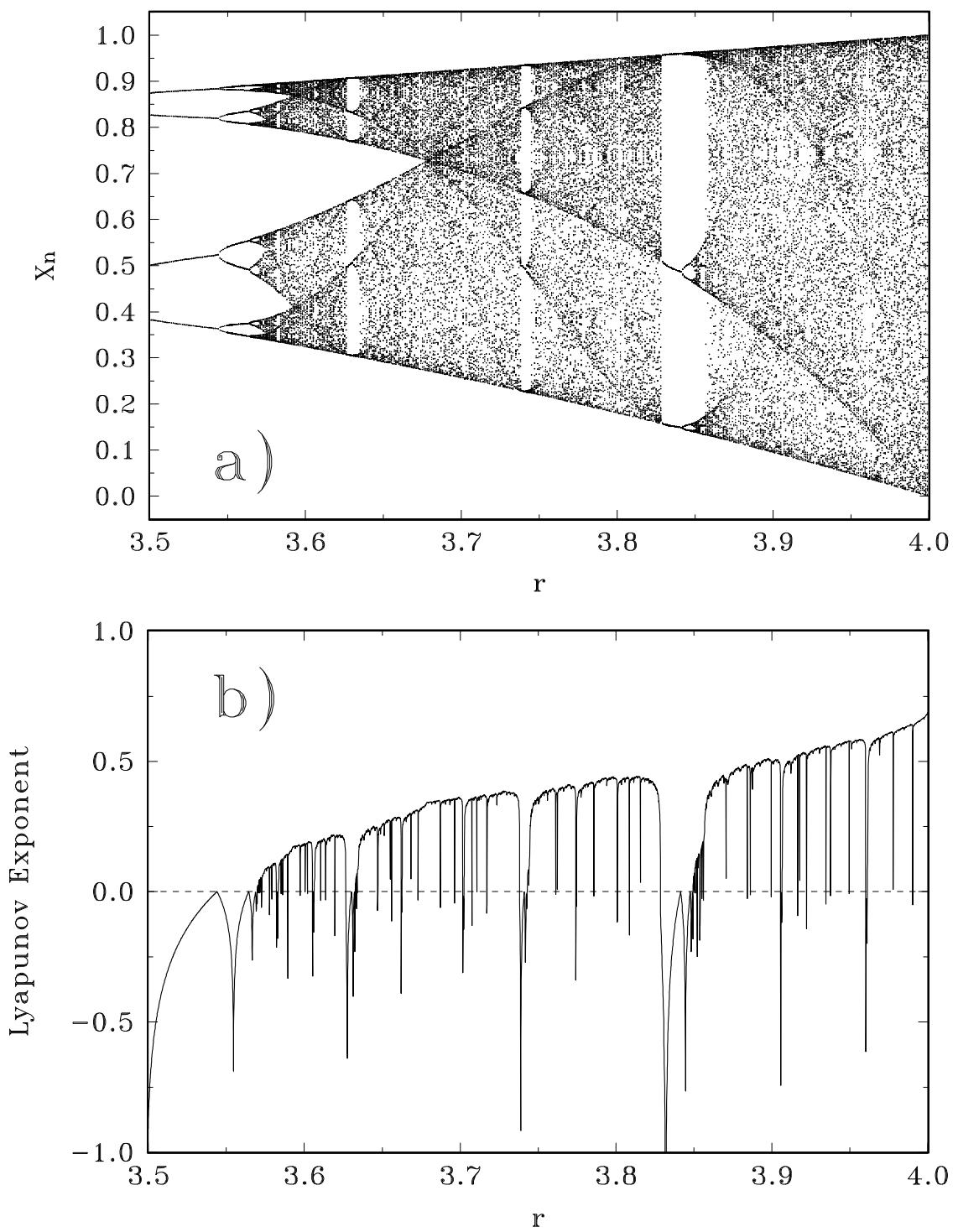


Figura 3.5: MAPA LOGÍSTICO: a) DIAGRAMA DE ÓRBITAS EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . b) EXPONENTE DE LYAPUNOV EN FUNCIÓN DEL PARÁMETRO DE CONTROL r .

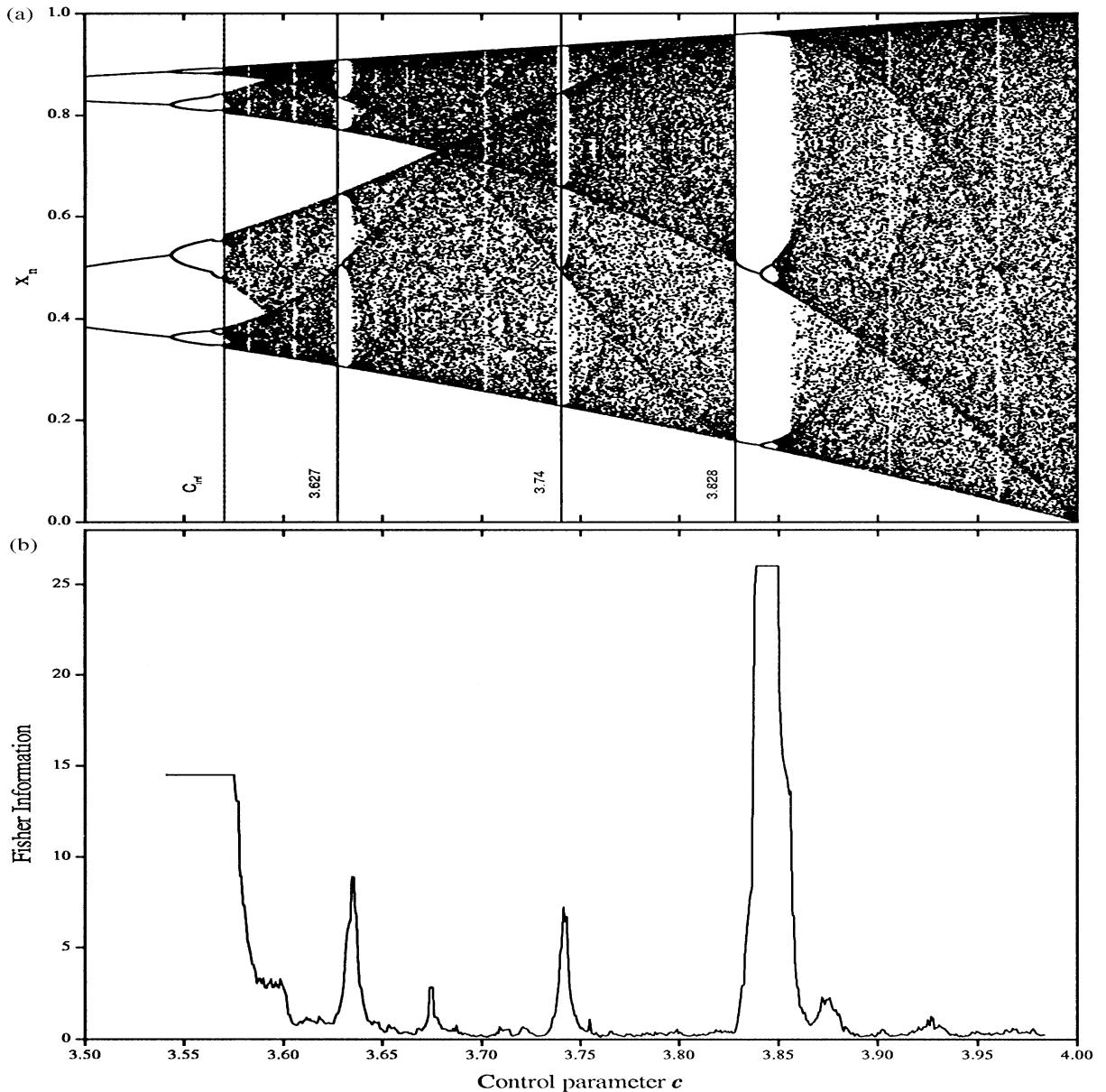


Figura 3.6: MAPA LOGÍSTICO: a) DIAGRAMA DE ÓRBITAS EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . b) MEDIDA DE INFORMACIÓN DE FISHER TAI EN FUNCIÓN DEL PARÁMETRO DE CONTROL. SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 10$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.

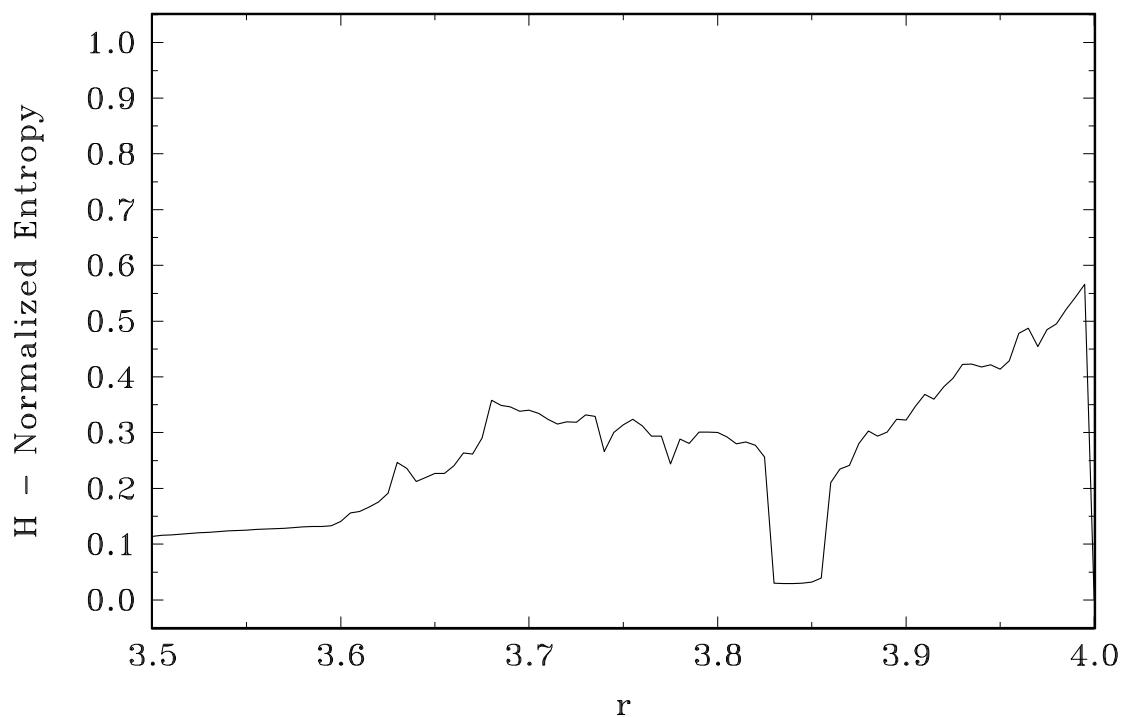


Figura 3.7: MAPA LOGÍSTICO: ENTROPIA WAVELET TOTAL, TWS, EN FUNCIÓN DEL PARÁMETRO DE CONTROL.

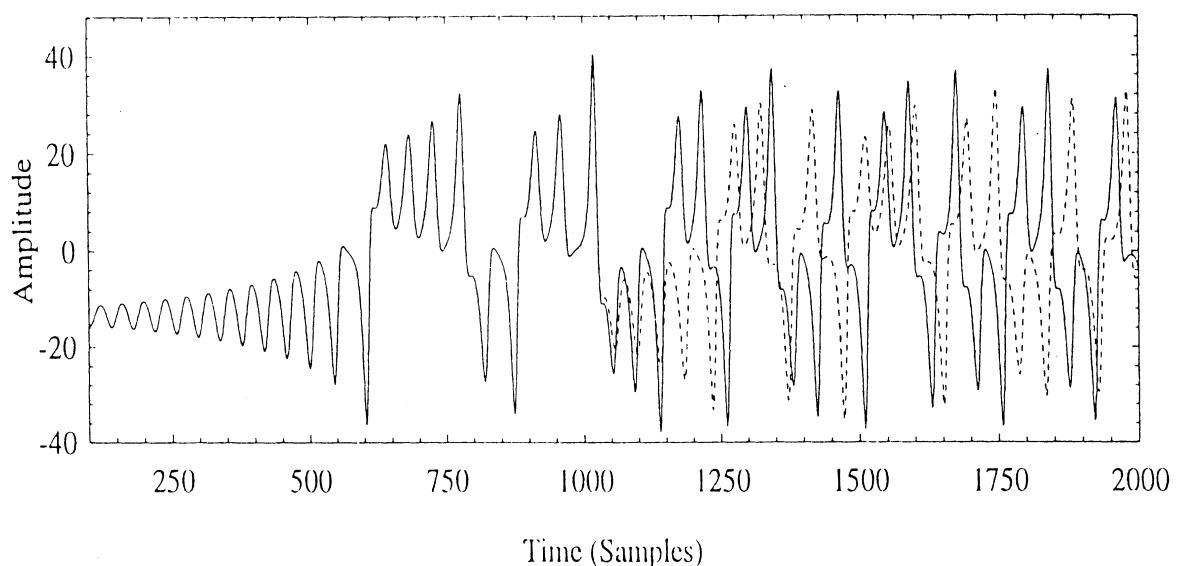


Figura 3.8: MODELO DE LORENZ: SEÑAL CORRESPONDIENTE A LA COORDENADA y , GENERADA CON PARÁMETRO FIJO (LÍNEA PUNTEADA), Y GENERADA CON PARÁMETRO VARIABLE (LINEA LLENA).

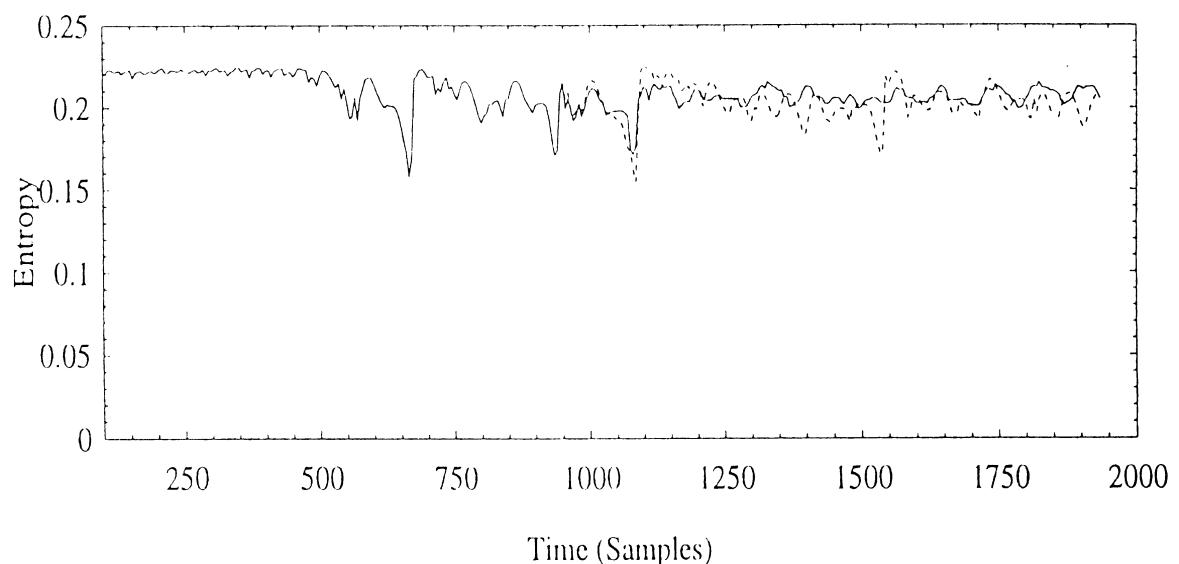


Figura 3.9: MODELO DE LORENZ: ENTROPIA DE SHANNON TAS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.

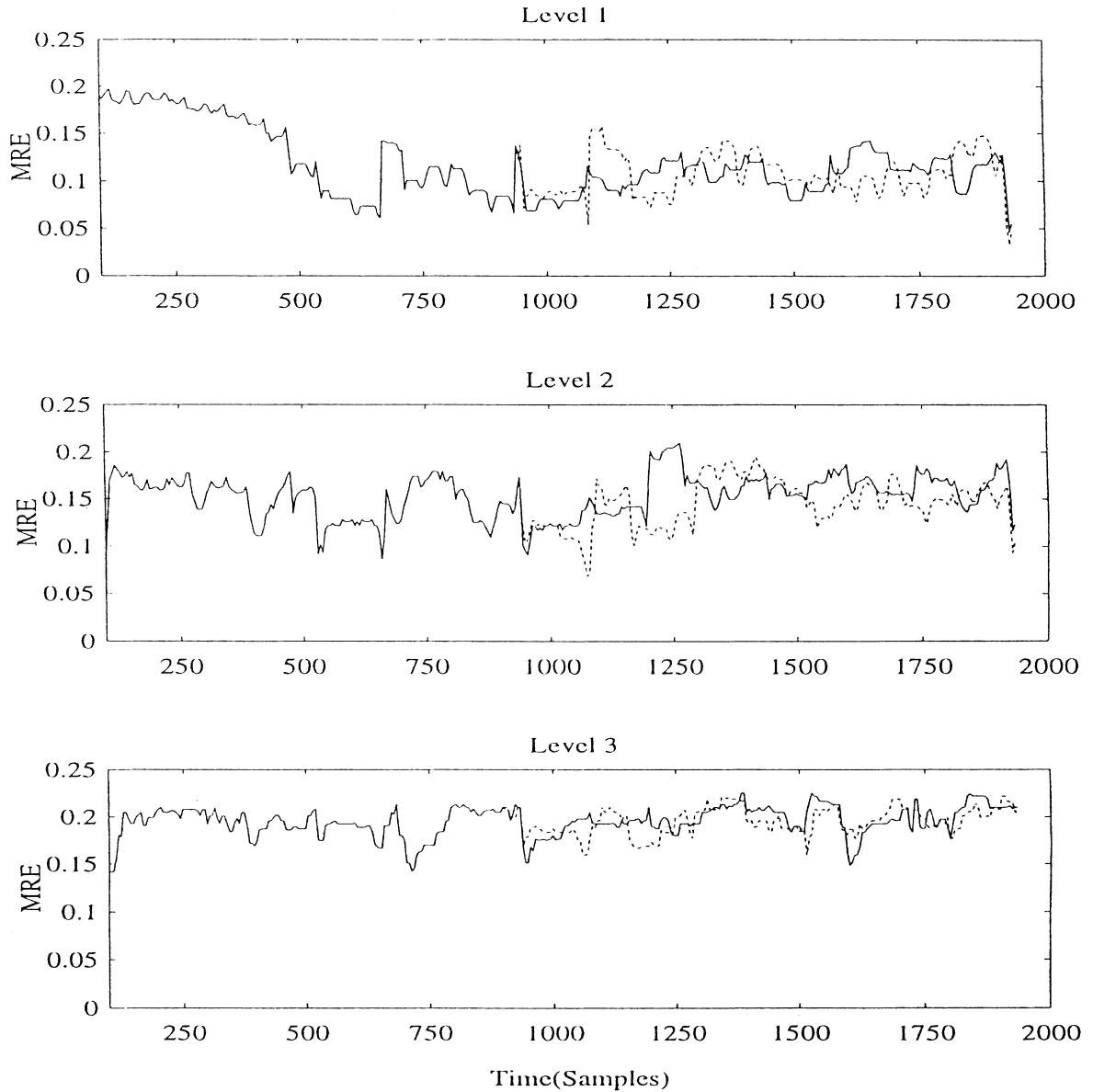


Figura 3.10: MODELO DE LORENZ: ENTROPIA MULTIESCALAS TMS DE LA SEÑAL GENERADA CON PARÁMETRO CONSTANTE (LÍNEA PUNTEADA) Y DE LA SEÑAL GENERADA CON PARÁMETRO VARIABLE (LÍNEA LLENA). LOS TRES PRIMEROS NIVELES DE RESOLUCIÓN. SE CONSIDERARON VENTANAS DE 128 MUESTRAS Y $N = 100$ SUBINTERVALOS EN LA PARTICIÓN DEL RANGO DE AMPLITUDES.

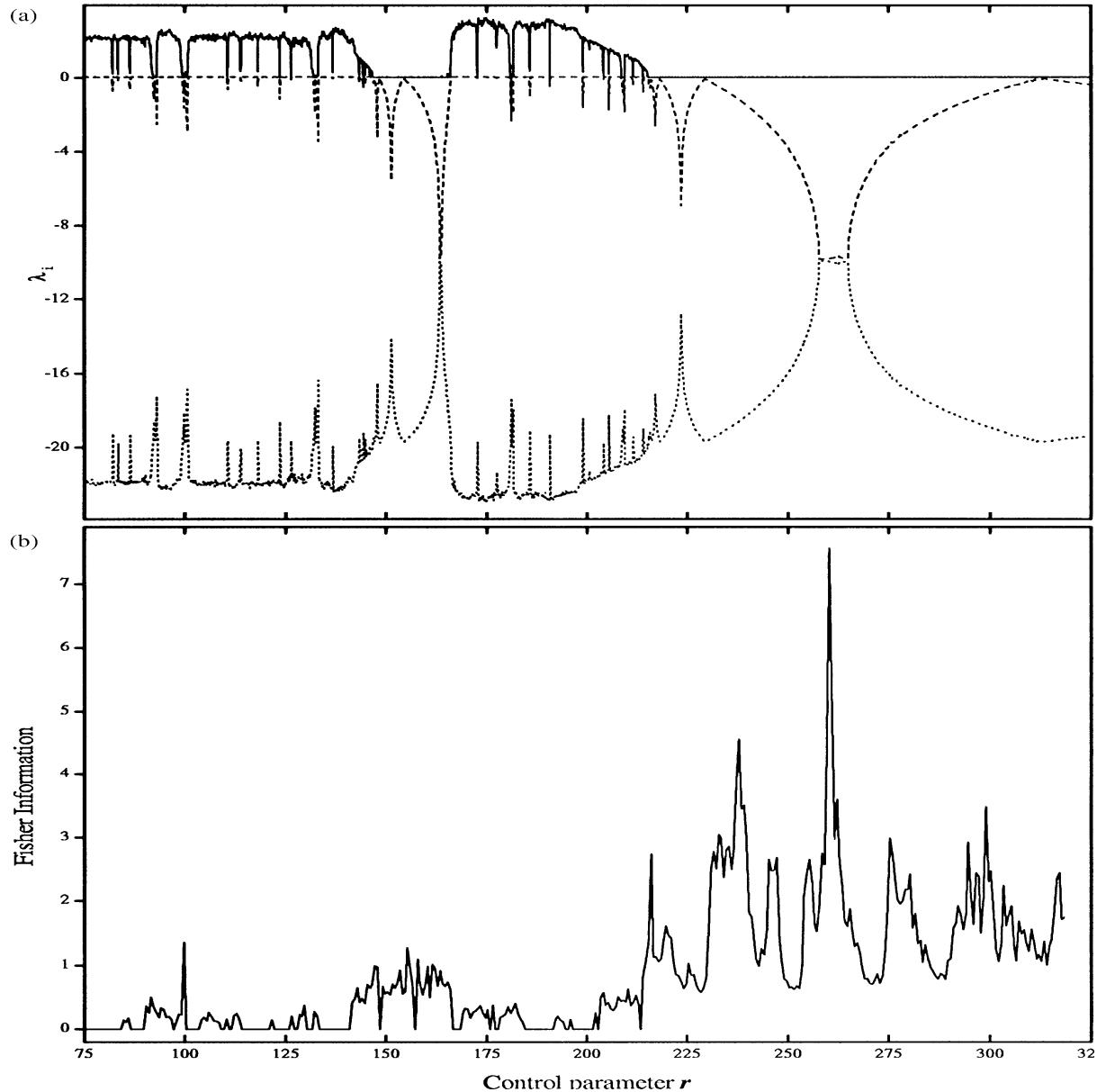


Figura 3.11: MODELO DE LORENZ: *a)* EXPONENTES DE LYAPUNOV EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . λ_1 (LÍNEA LLENA), λ_2 (LÍNEA SEGMENTADA) Y λ_3 (LÍNEA PUNTEADA). *b)* MEDIDA DE INFORMACION DE FISHER, TAI, EN FUNCIÓN DEL PARÁMETRO DE CONTROL. SE CONSIDERARON VENTANAS DE 512 MUESTRAS Y $N = 50$ SUBINTERVALOS EN EL RANGO DE LAS AMPLITUDES.

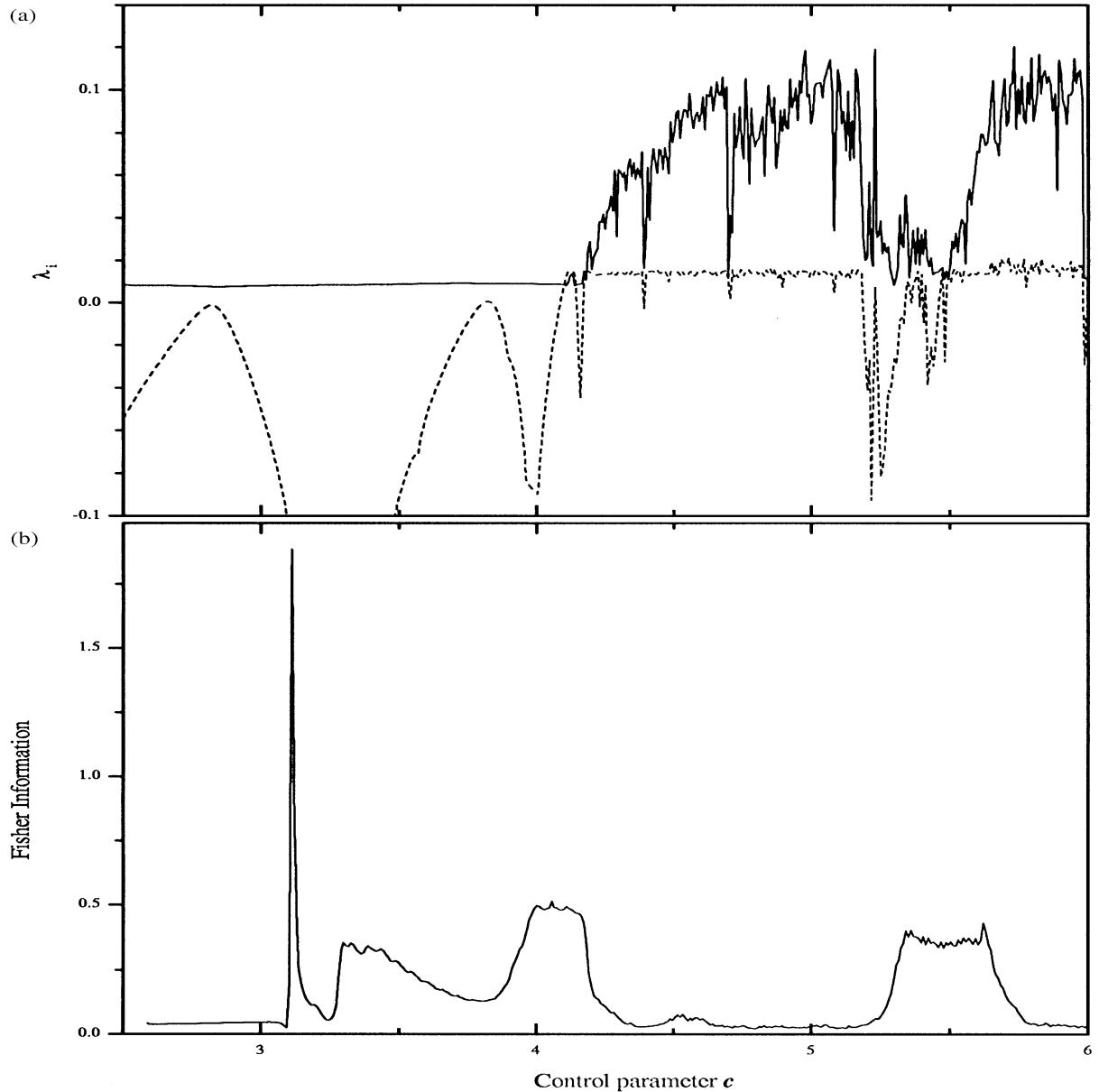


Figura 3.12: MODELO DE ROSSLER: *a)* EXPONENTES DE LYAPUNOV EN FUNCIÓN DEL PARÁMETRO DE CONTROL c . λ_1 (LÍNEA LLENA), λ_2 (LÍNEA SEGMENTADA) Y λ_3 (LÍNEA PUNTEADA). *b)* MEDIDA DE INFORMACION DE FISHER, TAI, EN FUNCIÓN DEL PARÁMETRO DE CONTROL. SE CONSIDERARON VENTANAS DE 1024 MUESTRAS Y $N = 10^4$ SUBINTERVALOS EN EL RANGO DE LAS AMPLITUDES.

Capítulo 4

Medidas de Complejidad (I):

Inhomogeneidad, desequilibrio y wavelets

4.1 Introducción

Si bien la materia organizada está presente en casi todos los aspectos del mundo natural, la disciplina científica que la estudia, carece de un método consistente que permita describir, cuantificar y detectar las diferentes clases de estructura que exhibe la naturaleza. La entropía estadística, como medida del “desorden”, es una cantidad aceptada y bien entendida en diversos contextos científicos. Sin embargo, es notable que medidas similares y complementarias de “estructura” o “complejidad”, han quedado fuera de la teoría corriente, estando las mismas muy poco desarrolladas.

Sin bien numerosas definiciones de “complejidad” han sido propuestas en los últimos 20 años, es escaso el consenso sobre el significado de este término. En la mayoría de los casos, las definiciones de complejidad introducidas son dependientes del contexto, como si tal dependencia fuera una propiedad inherente a la complejidad misma [123]. La subjetividad involucrada en esta contexto-dependencia y el malestar que esto provoca entre los científicos, es posiblemente la razón de la escasa coincidencia encontrada en el significado de este término.

Hasta el presente, la “complejidad” es manejada sobre la base del “caso por caso”: medidas de complejidad son desarrolladas típicamente para capturar las características significativas en un fenómeno específico. Es evidente que los actuales métodos no pueden ser extendidos para manejar las infinitas clases de organización que se encuentran en la naturaleza, especialmente aquellas producidas por procesos biológicos. Cuestiones relativamente simples, como por ejemplo: “¿qué grado de estructura o complejidad tiene un proceso que produce un dado nivel de desorden? ”, carecen aún de una respuesta adecuada.

Una nueva contribución en este sentido, será introducida en el presente capítulo. La complejidad de una serie temporal, será cuantificada a través de un método desarrollado en el marco del análisis multiresolución wavelet. Esta medida, si bien tendrá la ventaja de no ser “problema-específico”, será incapaz de cubrir, por sí sola, todos los aspectos que en años recientes han sido indicados bajo el término “complejidad”.

Este método resultará especialmente conveniente para el tratamiento digital de señales

no estacionarias y en particular para el análisis de la complejidad en series biológicas [108], como así también el tratamiento de sistemas físicos teóricos [70].

Una combinación del análisis wavelet multiresolución, por un lado, y la noción de complejidad estadística propuesta por Lopez Ruiz, Mancini y Calbet (LMC) en 1995 y recreada luego por Shiner, Davison y Landsberg (SDL), por el otro, será la base de la Complejidad Wavelet Total (TWC). En la sección 4.4.2, será presentado el método TWC, mientras que LMC y SDL serán desarrollados en las secciones 4.3.1 y 4.3.2, respectivamente.

Se mostrará en la sección 4.5.1 que TWC provee una nueva herramienta, útil para detectar y localizar cambios en la complejidad de series temporales provenientes de sistemas dinámicos no lineales. La noción de Complejidad Wavelet tiene las siguientes ventajas:

- a) Sobre la complejidad estadística, porque puede detectar cambios de complejidad en el comportamiento de una señal no estacionaria, gracias a las características de localización de la Transformada Wavelet.
- b) Sobre otras medidas de complejidad, porque el esfuerzo computacional que involucra es significativamente más bajo, lo que la hace efectivamente calculable en todos los casos.

4.2 Medidas de Complejidad

En los últimos 20 años, el estudio de sistemas complejos ha surgido como una disciplina reconocida en sí misma, aunque una buena definición de “sistema complejo” carece aún de una formulación adecuada. Intentos por formalizar el concepto de complejidad han llevado a profundizar en la teoría de la Información. Una buena revisión de los caminos que ha seguido el estudio de la complejidad, es presentada por Edmonds en Ref. [31].

Murray Gell-Mann, reconocido físico experto en “Sistemas Complejos”, ha discutido en numerosas conferencias y prestigiosas publicaciones, conceptos relacionados con la complejidad y el caos [44]. En base a sus fundamentadas apreciaciones, es fácil reconocer que una gran variedad de medidas serían necesarias para capturar todo lo que la intuición sugiere acerca de la *complejidad* o de su opuesto, la *simplicidad*.

- **Complejidad de Kolmogorov**

La primera y aun clásica medida de complejidad fue introducida independientemente por Kolmogorov (1965) [69], Chaitin (1966) [19] y Solomonoff (1964) [120], y es conocida como Complejidad de Información Algorítmica (AIC). Dada una particular Máquina de Turin universal, la complejidad de Kolmogorov de un string de caracteres (es decir, una descripción), es la longitud del menor programa corriendo en dicha máquina, que genere el string de caracteres dado. Por lo tanto, si una descripción admite una elevada compresión sin que haya pérdida de información, entonces es tomada como más simple que aquella que no puede admitir tal grado de compresión. Según esta definición, expresiones altamente ordenadas serán simples. Por el contrario, sucesiones aleatorias

tendrán máxima complejidad, ya que por definición de sucesión aleatoria, no puede haber un algoritmo generante más corto que aquel que simplemente “liste” toda la sucesión. Como puntualizó Gell-Mann [43], esto revela la inconveniencia de AIC como medida de complejidad.

AIC tiene más que ver con la cantidad de información de una sucesión que con su complejidad. Sucesiones con poca información estarán limitadas en su complejidad pero eso no quiere decir que todas las sucesiones con alto contenido de información sean complejas. Hay situaciones en las cuales existe mucha información, pero no está correlacionada, por lo tanto el “entero” es incomprimible y grande, aunque de cualquier manera “simple”. En efecto cuanto más correlaciones hay, más comprimible es la sucesión, y por tanto menor será AIC. Pero la correlación en sí misma, sugeriría un aumento de la complejidad y no una disminución. Gradualmente AIC fue llamada “aleatoriedad algorítmica” y carece de utilidad en la mayoría de las aplicaciones prácticas. Además AIC es incomputable en general, en el sentido que simplemente no hay algoritmo para calcularla.

Gell-Mann propone una medida, llamada Complejidad Efectiva (EC), que en armonía con la noción intuitiva de complejidad, se refiere a la descripción mas concisa, no ya de la entidad, sino del “conjunto de regularidades” de la entidad. Así una expresión totalmente aleatoria, al no tener regularidades, dará una Complejidad Efectiva casi nula.

• Complejidad y Teoría de Información

En general, la definición de medidas de complejidad, o bien continua en la dirección iniciada por la Complejidad de Kolmogorov [69, 19, 120] o sino adopta un punto de vista diferente basado en la Teoría de Información [23]. En este último caso se generan algoritmos tipo entropía, que aunque en principio son calculables, pueden ser muy difíciles de implementar y llevar a cabo efectivamente en sistemas experimentales. La “profundidad lógica” de Bennet es un ejemplo de la primer tendencia (se refiere al tiempo que tarda el menor programa de computadora). La “profundidad termodinámica” de Lloyd y Pagels [75], es un ejemplo del segundo grupo (se refiere a la entropía del conjunto de posibles trayectorias que conducen al corriente estado, es imposible de calcular).

La complejidad es a veces vista como un estado intermedio entre orden y desorden. Grassberger [50] consideró tres diagramas similares a los de la Fig. 4.1. La inmediata reacción es juzgar al primero y el último como simples y el del medio relativamente complejo. Esto se debe al hecho que la percepción del observador “filtra” la complejidad del diagrama de la derecha y lo interpreta como *representando* una situación sin *reglas*. O sea que el observador no juzga uniformemente. Podría suceder que el diagrama de la derecha represente una situación tan compleja que el observador no la reconozca. Se ve aquí la importancia del lenguaje de la representación. Esto es, si lo que se considera es la complejidad de las reglas que generan estos diagramas, entonces la intuición original es correcta. La confusión puede originarse en el hecho de que los diagramas no tienen un lenguaje propio, es el observador el que impone uno.

La principal barrera para arribar a un consenso respecto a la noción de complejidad, parece provenir de la falta de un lenguaje común en los distintos campos de la ciencia. Ciertamente, al ser la complejidad considerada en muy diversos campos de la ciencia, su significado adquirió cierta impresión o vaguedad. Existen muchos roles útiles para las

medidas de complejidad, aunque para garantizar su utilidad, es imperativo establecer claramente para qué se define una medida de complejidad, qué aspecto de la entidad será considerado y qué se intenta capturar.

Recientemente se ha introducido un nuevo tipo de medidas de complejidad, llamado “complejidad estadística”, que se refiere a cantidades que serán calculadas directamente desde la descripción probabilística del sistema.

- **Complejidad y entropía**

Frecuentemente el concepto de entropía ha sido usado como un punto de partida para estudiar complejidad. Pero, si bien es cierto que la entropía de un sistema está relacionada con la complejidad del mismo, también es cierto que por sí sola, no resulta una adecuada medida: “la aleatoriedad e impredicibilidad de un sistema falla para capturar sus patrones y estructura de correlación.”

Los algoritmos entrópicos tradicionales son, en general, una medida de la *incertezza* asociada a un proceso físico descripto por la distribución de probabilidades $P = \{p_i; i = 1, \dots, N\}$. Así, si la entropía es cero, podemos predecir con certeza *cuál* de las posibles salidas j realmente tendrá lugar. Nuestro conocimiento del proceso subyacente descripto por la distribución de probabilidades P es máximo. Por otro lado, nuestra ignorancia es máxima si la entropía es $\ln N$. Pero, estas dos circunstancias extremas de *i)* máximo conocimiento (*u* “orden perfecto”) y *ii)* máxima ignorancia (*o* “maxima aleatoriedad”), pueden coincidir en algún sentido y ser consideradas ambas como situaciones “triviales”. Es decir, situaciones que son extremos opuestos para las medidas entrópicas en general, serán coincidentes desde la óptica de la “no-trivialidad”. La medida de complejidad debería, por supuesto, asignar valor cero a estas situaciones consideradas “triviales” [5, 16, 77, 116, 117].

Pero además, la intuición supone que si un sistema es complejo debe contener cierta estructura o regularidad. Sin embargo, averiguar el grado de incertezza y aleatoriedad no es automáticamente equivalente a capturar la estructura de correlación presente, es decir las relaciones entre las componentes del sistema. Dicho de otra forma un aumento en la entropía no siempre está asociado con un aumento en la “no-trivialidad” o “complejidad”. Por ejemplo, una serie temporal con “ruido” tiene mayor entropía que la serie original, aunque el proceso generador de ruido destruye correlaciones y degrada el contenido de información de la señal original, es decir reduce la “complejidad” de la misma.

O sea que un algoritmo entrópico tradicional, no satisface en general las pautas mínimas requeridas a una medida de complejidad. Pero también es cierto que, aleatoriedad y correlación estructural no son aspectos totalmente independientes [5, 16, 77, 116, 117]. Efectivamente los extremos opuestos de *i)* orden perfecto y *ii)* máxima aleatoriedad no poseen estructura [58, 49, 25]. Entre estas dos instancias especiales existe un amplio rango de posibles grados de estructura física, grados que son reflejados en las características de la distribución de probabilidad subyacente. La medida de “complejidad” debería capturar adecuadamente estos grados de estructura, así como la entropía de Shannon captura aleatoriedad [5, 16, 77, 116, 117]. En síntesis, queda claro que las medidas de complejidad estadística, no sólo no se superponen con otras medidas de información, sino que son un útil complemento de ellas.

- **Aplicabilidad de las medidas de complejidad**

El ímpetu por desarrollar medidas de complejidad proviene en gran medida del campo de la biología. Recientemente, por ejemplo, se han generado discusiones acerca de si los sistemas biológicos evolucionan optimizando la complejidad, o dicho de otra forma, cómo evolucionarían los sistemas biológicos si la complejidad fuera maximizada.

En los últimos años ha crecido el interés por una medida de este tipo y su aplicación se ha diversificado. Así por ejemplo, ha despertado singular entusiasmo dentro del campo de la biología, cuantificar la “complejidad” de señales fisiológicas, tales como electrocardiogramas (ECG) o electroencefalogramas (EEG), y comparar esta medida para sistemas en salud y en enfermedad. La importancia de estas medidas radica en su capacidad tanto para evaluar modelos dinámicos en sistemas de control biológico, como para generar diagnósticos.

Una amplia clase de procesos de enfermedad, por ejemplo, degradan el contenido de información fisiológica y reducen capacidad adaptiva del individuo [22]. Al ser la información y la capacidad adaptiva, pautas de complejidad, se podría conjeturar que durante este tipo de procesos patológicos la complejidad biológica disminuye. Continuando en esta línea, podría pensarse que la pérdida de “complejidad” sea una característica general de dinámicas provenientes de organismos con alguna patología. La existencia de un “cuantificador” de “complejidad” calculable desde una señal experimental, sería entonces de gran utilidad, tanto para diagnóstico como para desarrollos formales en el campo de la biología.

4.3 Complejidad Estadística: dos Medidas Simples

4.3.1 Medida de complejidad LMC

Una medida de complejidad basada en una descripción probabilística del sistema físico, fue propuesta recientemente por Lopez-Ruiz, Mancini y Calbet (LMC) [77]. Esta medida incorpora las principales características sugeridas por la noción intuitiva de dicha magnitud. Investigando la noción de “complejidad” en física [95, 4], se ve que ésta comienza considerando tanto al cristal perfecto como al gas ideal aislado, como ejemplos típicos de modelos “simples” y por lo tanto como sistemas con “complejidad” nula.

Un cristal perfecto está completamente ordenado. La distribución de probabilidad de los estados accesibles está centrada alrededor de un estado preponderante. Una pequeña cantidad de “información” es suficiente para describirlo. Por otra parte, el gas ideal aislado es completamente desordenado. El sistema puede encontrarse en cualquiera de sus estados accesibles con la misma probabilidad. Almacena por lo tanto, una “información” máxima. Estos dos sistemas *simples* ocupan entonces extremos opuestos en las escalas de “orden” e “información”.

De la anterior observación se desprende que la complejidad no podrá caracterizarse sólo en términos de “orden” e “información”. LMC sugieren que una medida de complejidad debería tener en cuenta la distancia a la distribución equiprobable de los estados accesibles del sistema. Este “desequilibrio” daría una idea de la jerarquía probabilística del sistema. Así, el “desequilibrio” sería diferente de cero si hay estados privilegiados o

más probables. El cristal perfecto tendría entonces desequilibrio máximo, mientras en el gas ideal el desequilibrio sería nulo.

Por otra parte, LMC entienden también que la complejidad no debería ser medida únicamente, pues depende de la naturaleza de la descripción y de la escala de observación. Por lo tanto, LMC proponen una medida de complejidad basada en la descripción estadística del sistema que refleje el interjuego entre la “información” (S) y el “desequilibrio” (Q).

Sea $P = \{p_1, \dots, p_N\}$ la distribución de probabilidad reflejando el comportamiento del sistema. LMC adoptan para la información S y el desequilibrio Q , las siguientes definiciones:

- a) La cantidad de “información” almacenada en el sistema descripto por la distribución de probabilidad P , es medida mediante la entropía estadística de Shannon, S :

$$S(P) = - \sum_{i=1}^N p_i \ln p_i . \quad (4.1)$$

- b) El desequilibrio Q_E , calculado en base a la norma Euclídea, está dado por:

$$Q_E(P) = Q_E^0 \cdot \sum_{i=1}^N \left(p_i - \frac{1}{N} \right)^2 , \quad (4.2)$$

con $Q_E^0 = N/(N - 1)$ entonces $0 \leq Q_E \leq 1$.

A partir de estas definiciones, LMC proponen la siguiente medida de complejidad:

Definición 4.1:

La *complejidad* C_{LMC} de una distribución de probabilidad P , esta definida como:

$$C_{LMC}(P) = S(P) \cdot Q_E(P) . \quad (4.3)$$

Esta definición se ajusta a los argumentos intuitivos, siendo $C_{LMC} \sim 0$ tanto para los estados muy ordenados (homogéneos) como para los muy desordenados (heterogéneos), alcanzando un máximo en algún estado intermedio. Además, de acuerdo con lo que sugiere la intuición, la definición (4.3) está asociada a una *descripción particular* del sistema y por lo tanto a una *escala de observación*. Para cada escala de observación aparece un nuevo conjunto de estados accesibles, con su correspondiente distribución de probabilidad, que modifica el valor de la complejidad. Las leyes físicas de cada nivel de observación permitirán inferir la distribución de probabilidad de los nuevos estados accesibles, y por lo tanto diferentes valores para S , Q_E y C serán obtenidos.

La extensión de C_{LMC} al caso continuo es inmediata [77]. Si $p(x)$ es una distribución de probabilidad con soporte continuo y condición de normalización $\int p(x)dx = 1$, la complejidad C_{LMC} queda definida por:

$$C_{LMC}(p) = S(p) \cdot Q_{L_2}(p) = - \left(\int p(x) \ln p(x) dx \right) \cdot \left(\int p^2(x) dx \right) . \quad (4.4)$$

La definición de complejidad propuesta por LMC, ofrece un nuevo punto de vista, basado en la descripción estadística de los sistemas en una escala dada. En este esquema, el conocimiento de las leyes físicas que gobiernan la evolución dinámica en dicha escala es usado para encontrar los estados accesibles y su distribución de probabilidad. Este proceso podría inmediatamente indicar el valor de la complejidad, que esencialmente no es otra cosa más que el interjuego entre la información almacenada por el sistema y la *distancia a la equipartición* (medida de una jerarquía probabilística entre las partes observadas) de la distribución de probabilidad que describe el sistema considerado.

4.3.2 Medida de complejidad SDL

En Ref. [116], Shiner, Davison y Landsberg (SDL) introdujeron una medida de complejidad dependiente de dos parámetros, fácil de calcular y que se comporta como una cantidad termodinámica intensiva. Se expresa en términos de “desorden” y “orden”.

A menudo la entropía es tomada como una apropiada medida de desorden. Sin embargo, esto tácitamente supone que el tamaño del sistema, o sea el número de estados accesibles, no cambia. En efecto, si el número de estados del sistema aumenta entonces la entropía y por lo tanto el desorden del sistema también aumentará por la única razón de que aumentó el número de estados. Para evitar este problema, SDL han propuesto el “desorden” definido como:

$$\Delta(P) \equiv \frac{S(P)}{S_{\max}} , \quad (4.5)$$

donde S es la entropía de Shannon y S_{\max} es la máxima entropía posible para el sistema en estudio.

En el caso más simple de un sistema aislado, en el que no hay otra restricción más que la normalización de probabilidades, la variación de la funcional:

$$\delta \left[S(P) - \lambda \left(\sum_{j=1}^N p_j - 1 \right) \right] = 0 , \quad (4.6)$$

brinda como resultado, que la máxima entropía S_{\max} ocurre en la distribución equiprobable, $p_i = 1/N$, $1 \leq i \leq N$. Entonces, el valor de la máxima entropía posible es,

$$S_{\max} = \ln N . \quad (4.7)$$

Sin embargo en muchos casos este máximo no es alcanzable. Por ejemplo, si hay además otras restricciones, como por ejemplo, el número medio de partículas o la energía media libre, la máxima entropía será menor que la correspondiente a la distribución equiprobable, y es la apropiada máxima energía S_{\max} que debe ser usada en la definición del “desorden”. Cuando no hay otras restricciones y se maximiza la entropía, naturalmente se encuentra el máximo valor de la entropía de la ec.(4.7). Teniendo en cuenta estas consideraciones, SDL proponen la siguiente medida de complejidad:

Definición 4.2:

La *complejidad* C_{SDL} de una distribución de probabilidad P , esta definida como:

$$C_{SDL}^{\alpha,\beta}(P) = \Delta(P)^\alpha \cdot (1 - \Delta(P))^\beta , \quad (4.8)$$

donde α y β son parámetros reales.

La medida C_{SDL} satisface el criterio de máxima complejidad entre el orden y el desorden como ocurre con C_{LMC} y la parametrización con α y β permite ajustarla *ad hoc* de acuerdo a los casos específicos considerados.

En la literatura sobre complejidad existen una variedad de medidas comparadas y analizadas críticamente. En este sentido, Crutchfield, Feldman y Shalizi [27], así como también Binder y Perry [15], realizan severas críticas de la medida de SDL. Dicen que es “... sobreuniversal en el sentido que tiene la misma dependencia sobre el desorden para sistemas estructuralmente diferentes”. Esto deriva de una incorrecta interpretación de C_{SDL} , porque en realidad S_{max} depende del problema en estudio, y diferentes valores S_{max} son apropiados para responder diferentes cuestiones. Por lo tanto, si los sistemas son estructuralmente diferentes, S_{max} será seguramente diferente para cada sistema, lo que conducirá a un valor distinto para el “desorden” y por lo tanto para la complejidad C_{SDL} . Más aún, S_{max} , y por ende Δ y C_{SDL} , puede tomar distintos valores si se estudian cuestiones diferentes, aún en un mismo sistema dado.

4.3.3 Inhomogeneidad y desequilibrio

Habiéndose presentado dos medidas *simples* de complejidad, interesaría ahora rescatar las bondades de cada una de ellas, y proponer una nueva formulación que supere la performance de ambas. En este sentido, distinguiremos dos conceptos: *inhomogeneidad*, vinculado a SDL, y *desequilibrio*, introducido por LMC.

La inhomogeneidad, que será denotada H , está dada por $H = \Delta = S/S_{max}$. Es destacable que la “inhomogeneidad” H ha demostrado ser, por sí misma, una medida de utilidad en numerosas aplicaciones [97].

El desequilibrio Q , por otro lado, medirá ahora la distancia desde la distribución P considerada, a la distribución de probabilidad en la cual se alcanza S_{max} . Si llamamos $P_{max} = \{p_1^{max}, \dots, p_N^{max}\}$ a la probabilidad en la cual se alcanza S_{max} ,

$$Q(P) = Q^0 \cdot \sum_{i=1}^N (p_i - p_i^{max})^2 , \quad (4.9)$$

siendo Q^0 una constante de normalización. Teniendo en cuenta estas consideraciones, se definirá la complejidad C , como:

Definición 4.3:

Dada una distribución de probabilidad P , definimos la siguiente medida de *complejidad*, C , como:

$$C(P) = H(P) \cdot Q(P) , \quad (4.10)$$

donde $H(P) = S(P)/S_{max}$, siendo S la entropía de Shannon y S_{max} es el máximo valor que puede tomar la entropía para el sistema en estudio. El desequilibrio Q está definido por la ecuación (4.9), en la cual P_{max} representa a la distribución de probabilidad donde ocurre la máxima entropía S_{max} .

- Relación con C_{SDL} .

En el caso más simple, en el que la máxima entropía ocurra en la distribución equiprobable, C es una aproximación a C_{SDL} . Sea $\Omega = 1 - \Delta$. Se realizará una expansión por Taylor de Ω alrededor de la distribución equiprobable. Notando que

$$\Omega = 1 + \frac{\sum_{i=1}^N p_i \ln p_i}{\ln N}, \quad \Omega_0 = 0, \quad (4.11)$$

la expansión requerida será,

$$\begin{aligned} \Omega &= \Omega_0 + \sum_{i=1}^N \left(\frac{\partial \Omega}{\partial p_i} \right)_0 \left(p_i - \frac{1}{N} \right) + \frac{1}{2} \sum_{i=1}^N \left(\frac{\partial^2 \Omega}{\partial p_i^2} \right)_0 \left(p_i - \frac{1}{N} \right)^2 + \dots \\ &= 0 + \left(\frac{1}{\ln N} - 1 \right) \sum_{i=1}^N \left(p_i - \frac{1}{N} \right) + \frac{1}{2 \ln N} \sum_{i=1}^N \left(p_i - \frac{1}{N} \right)^2 + \dots \\ &= \frac{1}{2 \ln N} Q + \dots . \end{aligned} \quad (4.12)$$

Como el segundo término de la expansión se anula, se ve que el “desequilibrio” Q es, a menos de un factor dependiente de N , una expansión a segundo orden de Ω alrededor de la distribución equiprobable. Por lo tanto, la complejidad C satisface,

$$C = H \cdot Q = H \cdot \frac{2 \ln N}{N} \Omega \cong \frac{2 \ln N}{N} \cdot C_{SDL}, \quad (4.13)$$

en el caso más simple, en el que la máxima entropía ocurra en la distribución equiprobable.

4.4 Complejidad y Transformada Wavelet

4.4.1 Multi-escalas y complejidad

Dinámicas disipativas, quiebre de simetrías, transiciones de fase, bifurcaciones y formación de patrones, actuando sobre diferentes escalas espaciales y temporales, en diferentes niveles y sobre diferentes sustratos, son presumiblemente responsables de generar la amplia diversidad de estructuras observadas en el mundo natural. Cada uno de estos procesos tiene sus fundamentos, pero, ¿cuáles son los principios que definen y describen sus productos?

Un sistema físico en general, consta de diversas partes altamente interconectadas, la interacción de las cuales da lugar a un comportamiento complejo. Nuevas características emergen si se va de una escala a otra, y resulta interesante conocer los principios que rigen la aparición de estas nuevas propiedades. En sistemas complejos, se acepta que procesos que ocurren simultáneamente a diferentes escalas o niveles son todos importantes, y el comportamiento del sistema *entero* depende de una manera no trivial de sus unidades constitutivas.

Es conveniente entonces disponer de una nueva medida *cualitativa* que trate el sistema *entero*, es decir teniendo en cuenta sus múltiples escalas, porque las leyes que describen el comportamiento global son *cualitativamente* diferentes de aquellas que rigen sus unidades individuales.

Una serie discreta es en general, el resultado de mediciones realizadas a intervalos de tiempo (o tamaño de boxes) predefinidos y por lo tanto es una descripción del sistema en una escala determinada. Las clásicas medidas de complejidad están basadas en el análisis de una única escala, la de la señal original y no tienen en cuenta las fluctuaciones temporales complejas inherentes a los sistemas reales. Es decir que estas medidas únicamente caracterizan la estructura correspondiente a la menor escala, o sea la escala del proceso de muestreo. Los comportamientos multiescalas característicos de los sistemas naturales e ignorados por los algoritmos convencionales, inspiran una nueva medida *simple* y *calculable*, que proveerá un nuevo camino para caracterizar la complejidad.

4.4.2 Complejidad wavelet

Con la idea de definir una medida de complejidad que tenga en cuenta la inherente naturaleza multi-escala de los sistemas físicos, se introducirá la Complejidad Wavelet (WC). WC combina las propiedades de la medida de complejidad C , expuestas en la sección anterior, con las propiedades de localización tiempo-frecuencia de la transformada wavelet.

Suponiendo que la señal s a analizar tiene 2^M muestras, el número de coeficientes en el nivel j sera $K_j = 2^{(M+j)}$. Sea D_j , para $j = -1, \dots, -M$, denotando el conjunto de coeficientes wavelet del nivel de resolución j , como en la sección 3.4.1. Si ahora se consideran los coeficientes wavelet correspondientes a los M niveles de resolución, se tendrá el siguiente conjunto D^M :

$$D^M = \bigcup_{j=-1}^{-M} D_j = \{d_j(k) ; k = 1, \dots, K_j ; j = -1, \dots, -M\}. \quad (4.14)$$

Dado el conjunto de coeficientes D^M y de acuerdo a la definición 3.6, la distribución de energía multiescala P_{ME} asociada a D^M , estará dada por la ecuación (3.21). Es posible definir ahora la complejidad wavelet, del siguiente modo:

Definición 4.4:

Sea D^M el conjunto de coeficientes wavelet correspondientes a los M niveles de resolución de una serie temporal s de longitud 2^M . Se define la *complejidad wavelet* asociada a D^M , WC, como:

$$\text{WC}(D^M) \triangleq C(P_{ME}), \quad (4.15)$$

donde C es la medida de complejidad definida en la ec. (4.10).

4.4.3 Complejidad wavelet total (TWC)

El análisis multiresolución de la señal da lugar a una descomposición con localización temporal que permitirá estudiar la complejidad de la señal a través de ventanas

deslizantes, como fue hecho con la entropía wavelet total (TWS) en la sección 3.5.3. Suponiendo que la cantidad total de datos a analizar es 2^M , se considerarán sobre el conjunto de coeficientes D^M , dado por la ec. (4.14), ventanas *bidimensionales* deslizantes $W_{(i)}^L$ dependientes de un parámetro L (ver sección 3.5.3).

Se recordará que las ventanas deslizantes $W_{(i)}^L$ están definidas por:

$$\begin{aligned} W_{(i)}^L &= \{d_j(k); k = 1 + (i-1)2^{L+j}, \dots, i2^{L+j}; \\ &\quad j = -1, \dots, -L\} , \end{aligned} \quad (4.16)$$

para $i = 1, \dots, I_L$ donde $I_L = 2^{M-L}$ es el número de ventanas a considerar.

Las consideraciones anteriores permiten definir la evolución temporal de la complejidad wavelet de una señal s , TWC, como:

Definición 4.5:

Dada una señal s y un parámetro L , que determinará ventanas de ancho 2^L , se define la *complejidad wavelet total*, TWC_s^L , como:

$$\text{TWC}_s^L(i) \triangleq \text{WC}(W_{(i)}) = C(P_{\text{ME}}^{(i)}) . \quad (4.17)$$

donde $P_{\text{ME}}^{(i)}$ es la distribución de energía multiescala asociada a $W_{(i)}$.

Se puede observar en este momento, que a medida que i varía, la variable temporal $i2^L - 2(L-1)$, evoluciona. Por lo tanto la medida $\text{TWC}_s^L(i)$ representa la *evolución temporal de la complejidad* C , medida sobre el conjunto de *energías relativas* de las bandas de frecuencia o rangos de escala. La evolución temporal se mide, al igual que en el caso de la entropía TWS (ver sección 3.5.3), a través de ventanas bidimensionales deslizantes $W_{(i)}$ de tamaño L , definidas sobre el conjunto de coeficientes wavelet de un análisis multiresolución de $s(t)$.

Teniendo en cuenta que los coeficientes wavelet en cada nivel contienen los *detalles* de la señal en el correspondiente rango de frecuencias, la medida TWC proporciona información respecto de la complejidad en la interrelación de estos rangos de frecuencias. Por ejemplo, si la señal que analizamos tiene una frecuencia de muestreo de w_s , los coeficientes wavelet en el nivel j muestran los detalles de la señal en la banda de frecuencia $[2^j w_s, 2^{j+1} w_s]$. Por lo tanto la $\text{TWC}_s^L(i)$ dará información de la complejidad de la señal con respecto a la energía relativa de cada una de estas j -bandas y localizada en el espacio temporal dado por $W_{(i)}^L$. Si graficamos la TWS vs. el tiempo, considerando los puntos $(i2^L - 2(L-1), \text{TWC}_s^L(i))$, para $i = 1, 2, \dots, I_L$, este enfoque tiene las propiedades de localización de las wavelets, y por tanto, será útil para detectar cambios dinámicos en la complejidad de la señal.

4.5 Aplicaciones III: Medidas de Complejidad

Para ilustrar las propiedades de la medida de complejidad TWC propuesta, se considerará su aplicación al mapa logístico que ya fuera presentado en la sección 3.6.1.

Se analizará el comportamiento de la TWC ante la variación temporal continua del parámetro que gobierna la dinámica logística. Esta aplicación surge con un interés puramente ilustrativo y con la idea de facilitar la comparación de esta medida con otras presentes en la literatura pertinente.

Como se ha comentado en la presentación, las bondades de TWC se manifiestan principalmente al aplicarla en señales provenientes del mundo real. De este modo en el Capítulo 7 se utilizarán datos experimentales para ilustrar los conceptos discutidos en este trabajo. En efecto la capacidad en la detección de eventos transitorios, así como también la fácil implementación e interpretación de las técnicas desarrolladas, quedará clara mas adelante cuando se muestren aplicaciones realizadas en señales biomédicas [108], particularmente electroencefalogramas (EEG).

4.5.1 Mapa logístico

El mapa logístico constituye un ejemplo generalmente empleado para ilustrar nuevos conceptos en el tratamiento de sistemas dinámicos. Se trata de un sistema simple que presenta una rica variedad de comportamientos y según el valor del parámetro de control r el mapa puede mostrar un punto estable, duplicación de período o caos.

El mapa logístico es un sistema disipativo descripto por la ecuación en diferencia de primer orden:

$$x_{n+1} = r x_n (1 - x_n), \quad 0 \leq x_n \leq 1, \quad 1 < r \leq 4, \quad (4.18)$$

cuyo comportamiento dinámico está controlado por el parámetro r . Una minuciosa descripción de la dinámica logística se encuentra en la sección 3.6.1.

Siguiendo esta idea discutiremos aquí la aplicación de la medida de complejidad TWC_s^L (ver sec. 4.4.3), en el análisis del celebrado mapa logístico $F : x_n \rightarrow x_{n+1}$. La señal s que va a ser estudiada está definida como: $s_{r,n} \equiv x_n(r)$, donde $n = 1, \dots, M$ y el valor de r varía desde 3.5 hasta 4.0 con un paso $\Delta r = 0.001$. Esto es, para cada valor de r , con paso Δr , se generan M valores x_n de acuerdo a la ec. (4.18) y estos valores se acumulan formando la señal s . La evolución temporal de la señal s significa entonces, la evolución del parámetro de control r , y por lo tanto la aparición de una rica variedad de comportamientos. Esto hace que sea una señal útil para caracterizar la performance de nuestra medida TWC_s^L .

Se revisará primero el tratamiento del mapa logístico con las medidas de complejidad C_{LMC} y C_{SDL} definidas en la sección 4.3.1, lo que permitirá luego realizar una comparación. En la aplicación de las medidas C_{LMC} y C_{SDL} se seguirá la Ref. [77]. De este modo, para cada valor del parámetro r , se reducirá la dinámica del mapa logístico a una sucesión binaria (0 si $x \leq 1/2$; y 1 si $x > 1/2$), y se considerarán luego sucesiones binarias de longitud 12 como estados del sistema. Las probabilidades correspondientes son asignadas de acuerdo a la frecuencia de ocurrencia de cada estado luego de, como mínimo, 2^{22} iteraciones. El propósito es realizar un análisis general de la medida de complejidad en la región $3.5 \leq r \leq 4.0$, y efectuar una comparación entre la descripción provista por C_{LMC} y C_{SDL} con aquella que surge de la aplicación de TWC, cuyos resultados se muestran en la Fig. 4.2.

Al observar las Figs. 4.2 y 4.3, se ve que mientras C_{LMC} asigna complejidad 0 a la dinámica regular correspondiente a $3.5 < r < 3.56$, TWC acusa la existencia de un

período 4, apartándose ligeramente de la complejidad nula. Observamos en ambos casos un crecimiento de la complejidad a partir de $r = 3.56$, pero en C_{LMC} este crecimiento es abrupto, mientras que para la complejidad wavelet TWC resulta ser suave y gradual. C_{LMC} alcanza un máximo en la vecindad de $r > r_\infty \cong 3.57$, y TWC presenta un quiebre en la línea alrededor del mismo punto. Luego aparece una zona estocástica claramente visible en la Fig. 3.5.a. Mayores detalles son también detectados por ambas medidas, como por ejemplo las 3 más notables ventanas regulares, correspondientes a las órbitas de período 6, 5 y 3. Pero globalmente la medida básica C_{LMC} decrece, alcanzando valores casi nulos. Notar en cambio que la medida TWC crece globalmente en la región entre ventanas, y se notan picos descendentes que coinciden con las ventanas periódicas.

Por su parte, la aplicación de la medida C_{SDL} definida en la sección 4.3.2 se muestra en la Fig. 4.4. Los resultados obtenidos con el método de Shiner y colaboradores (C_{SDL}), resultan tener mayor coherencia con la dinámica logística, que los proporcionados por C_{LMC} . Pero aún así, se desprende de las figuras que algunos aspectos relevantes de la dinámica logística son descriptos más consistentemente por la nueva medida TWC. En efecto, valores particulares de r asociados a regímenes periódicos son bien identificados por ambas medidas, pero la complejidad basada en wavelets manifiesta un comportamiento global que tiene mayor coherencia con la física del sistema. Esto es TWC crece a medida que crece el grado de caoticidad, descendiendo sólo en los regímenes periódicos y finalmente cae tomando valores casi nulos cuando la caoticidad es muy grande.

4.6 Conclusiones

En este capítulo se presentó un método que permite cuantificar la complejidad de una señal: la complejidad wavelet total TWC. Este cuantificador surge de la combinación entre *a)* teoría de información (IT), por un lado y *b)* transformada wavelet (WT), por el otro.

Respecto al punto *a)* nos referimos aquí a dos *simples* medidas de información:

- La complejidad C_{LMC} presentada por Lopez Ruiz, Mancini y Calbet en 1995 [77].
- La complejidad C_{SDL} presentada por Shiner, Davison y Landsberg en 1999 [116].

Respecto al punto *b)*, la WT proporciona el marco teórico que posibilita el enfoque tiempo-frecuencia/escala para las medidas de información enunciadas en *a)*.

En los últimos años, varias medidas de complejidad han sido desarrolladas para comparar series temporales. Los principales parámetros de complejidad clásicos son *entropía*, *dimensión fractal* y *exponentes de Lyapunov*. Ellos están definidos en base a órbitas típicas de sistemas dinámicos. El problema conceptual básico es que estas definiciones no son válidas para una señal arbitraria (por ejemplo, *no estacionaria*), y entonces surgen problemas computacionales. Si bien se han desarrollado muchos algoritmos, ellos trabajan bien cuando la serie temporal es simulada a partir de un sistema dinámico de baja dimensión, pero la mayoría falla en presencia de ruido aditivo, por ejemplo. Para las series del mundo real, la eliminación de ruido mediante un filtrado clásico, requiere un cuidadoso preprocesamiento de los datos y ajuste de parámetros.

La medida simple de complejidad definida en este capítulo, en cambio, es calculada para cualquier tipo de serie temporal, sea estacionaria, caótica, ruidosa, etc. La noción

de complejidad a partir de una descripción probabilística del sistema, constituye en sí misma una propuesta interesante. Proporciona una medida que es al mismo tiempo fácil de calcular e intuitivamente coherente. Su presentación en el contexto del análisis wavelet, realizada en este capítulo, la convierte en una herramienta útil para cuantificar la complejidad multiescalas, que además tiene localización temporal. Efectivamente, si se tiene en cuenta que los coeficientes wavelet correspondientes a cada nivel de resolución contienen los *detalles* de la señal en el correspondiente rango de frecuencias, es fácil ver que la medida TWC proporciona información respecto del grado de complejidad presente en la interrelación de estas bandas de frecuencias.

El mencionado enfoque tiempo-frecuencia/escala en el cual se enmarca TWC, convierte a este método en uno particularmente apropiado para su aplicación en señales provenientes del mundo real. Las medidas convencionales para medir la complejidad de una señal, en general ignoran las múltiples escalas temporales. Esto hace que no sean confiables cuando son aplicadas, por ejemplo, a registros provenientes de sistemas biológicos [22]. Por el contrario, TWC contempla la naturaleza particular de estos sistemas y ha demostrado ser de utilidad en diversas aplicaciones experimentales [107, 108].

De ninguna forma se piensa que el procedimiento expuesto resuelva el problema de caracterizar la complejidad, que es un tema que está aún lejos de ser entendido. Sin embargo, se piensa que TWC puede dar luz en variados aspectos del conocimiento, al proponer un camino alternativo para estudiar las propiedades de un sistema físico que pueda ser descripto en términos de leyes de probabilidad.



Figura 4.1: DIAGRAMA DE GRASSBERGER: COMPLETAMENTE ORDENADO (DERECHA), CAOS (CENTRO) Y COMPLETAMENTE DESORDENADO (IZQUIERDA)

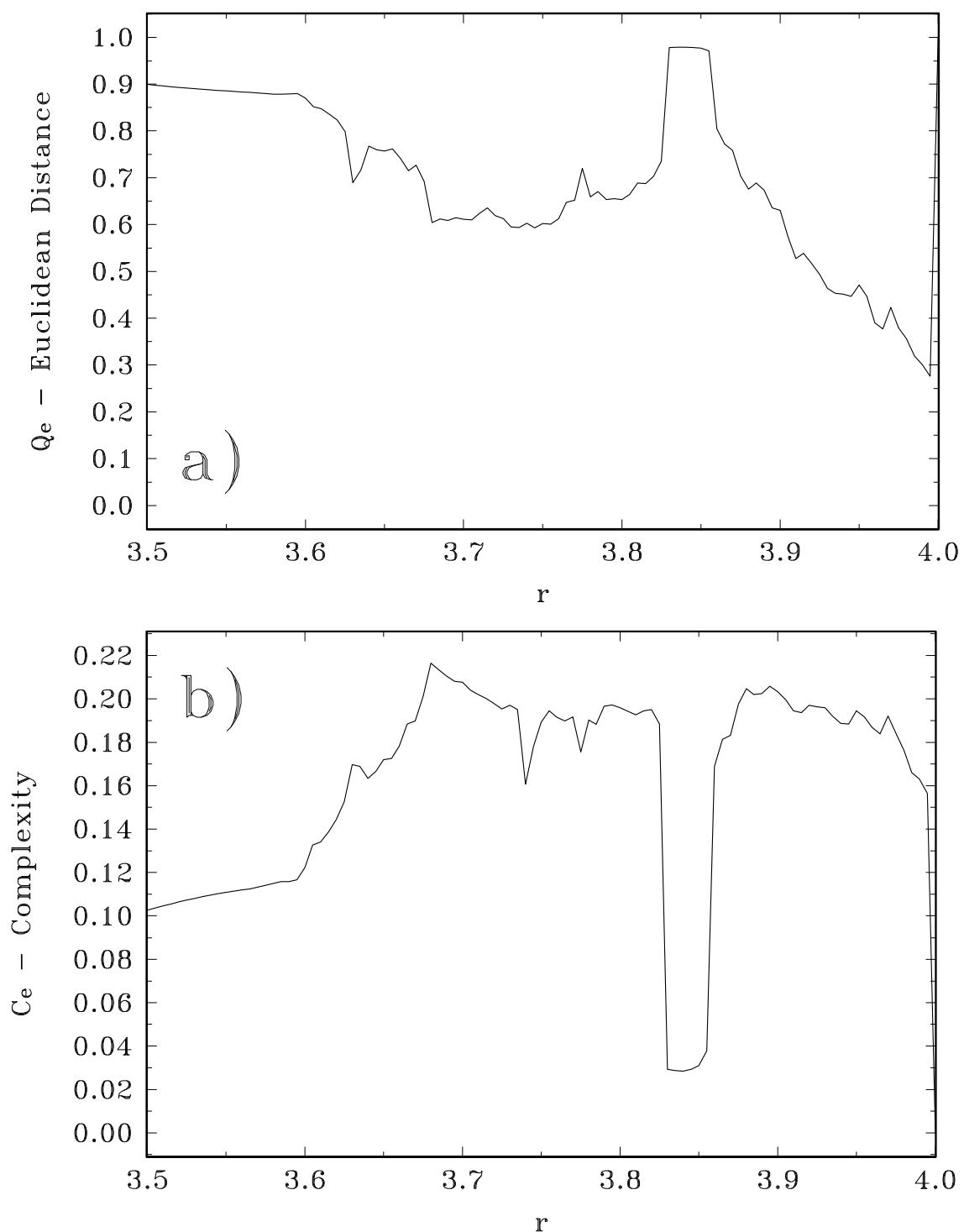


Figura 4.2: MAPA LOGÍSTICO (WAVELET SPLINE CUBICA): a) DESEQUILIBRIO EUCLÍDEO (Q_E) EN FUNCIÓN DEL PARAMETRO DE CONTROL r . b) COMPLEJIDAD TWC EN FUNCIÓN DEL PARAMETRO DE CONTROL r .

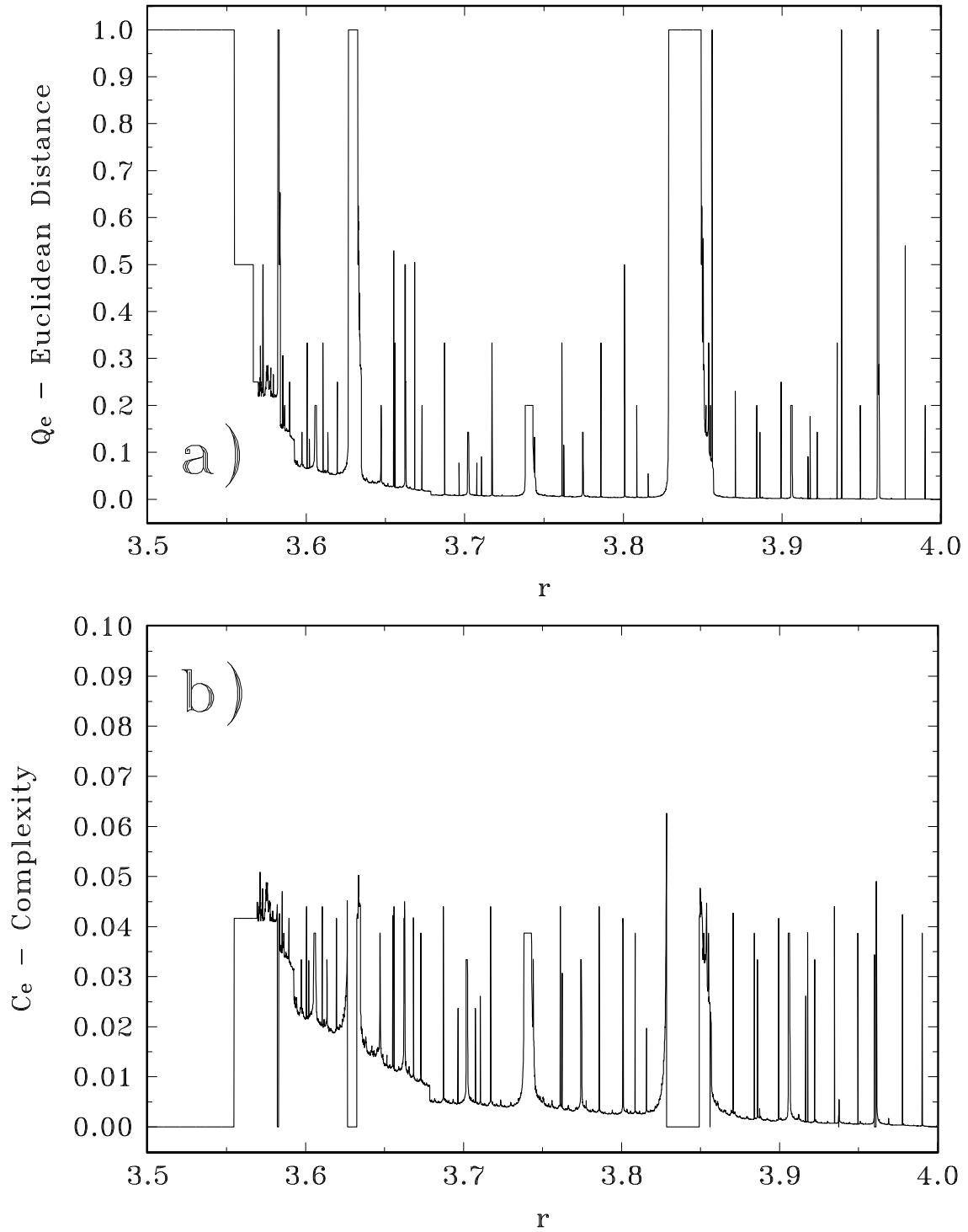


Figura 4.3: MAPA LOGÍSTICO (SUCECIÓN BINARIA): a) DESEQUILIBRIO EUCLÍDEO (Q_E) EN FUNCIÓN DEL PARÁMETRO DE CONTROL r . b) COMPLEJIDAD C_{LMC} EN FUNCIÓN DEL PARÁMETRO DE CONTROL r .

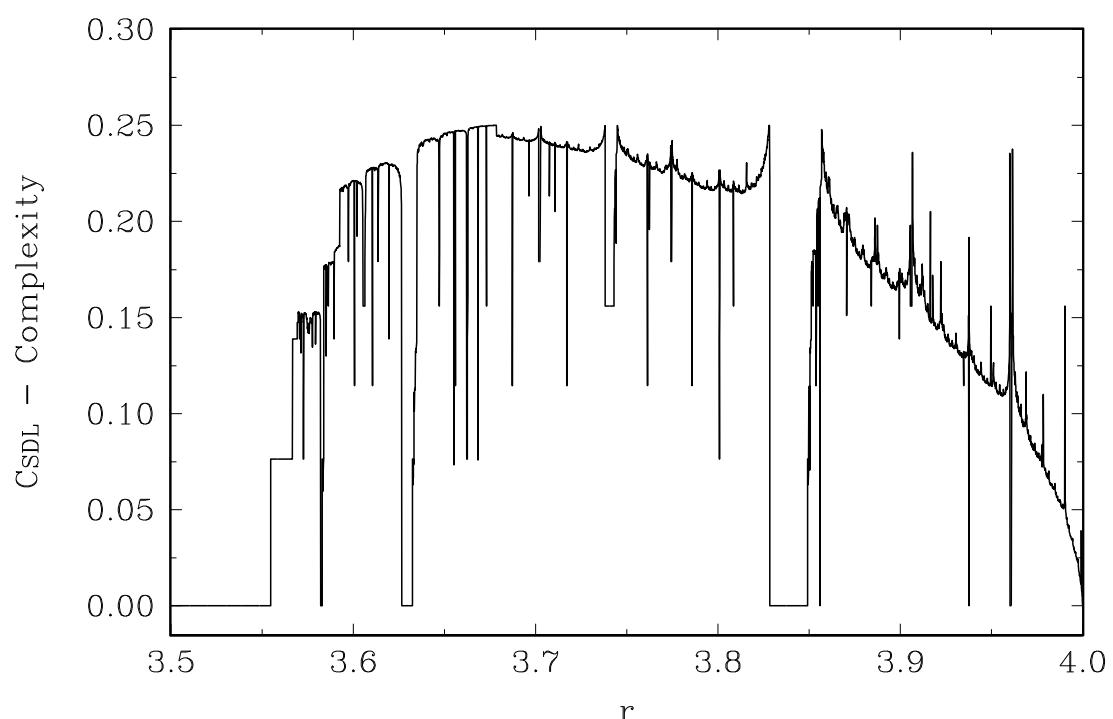


Figura 4.4: MAPA LOGÍSTICO (SUCESIÓN BINARIA): COMPLEJIDAD C_{SDL} EN FUNCIÓN DEL PARÁMETRO DE CONTROL r .

Capítulo 5

Medidas de Complejidad (II):

La Noción de Distancia en la Formulación del Desequilibrio

5.1 Introducción

En el capítulo anterior se ha trabajado con medidas de complejidad estadística satisfaciendo una condición particular, que Crutchfield [27] ha llamado criterio de “one-hump”, en obvia referencia al requisito de tomar valores pequeños tanto en el orden perfecto como en la aleatoriedad total. Por otra parte, la Complejidad Wavelet Total (TWC), que fue introducida en la sección 4.4.3, se construye a partir de una medida de complejidad estadística *básica*, que fue llamada C (ver ec. (4.10)).

Con la idea de que, una mejor medida de complejidad estadística básica C , conducirá a una mejor medida de complejidad wavelet TWC, se plantea ahora el siguiente objetivo: en base a la medida básica C , definir otra que manteniendo el criterio de “one-hump” y la facilidad de su implementación, posea además mayor coherencia con la física subyacente. Para lograr este propósito, se dedicará el presente capítulo a un análisis crítico y a una posterior generalización, de la primaria medida de complejidad C (ver sección 4.3.3).

La medida C consiste básicamente en la multiplicación de un término que cuantifica la “inhomogeneidad” por otro que cuantifica el “desequilibrio”, esto es: $C = H \cdot Q$. La noción de “distancia estadística”, al ampliar el marco teórico que sustenta formulación de Q , permitirá una consistente generalización del mismo. En efecto, el físico W.K.Wootters presentó en 1981 la noción de “distancia estadística” entre estados cuánticos de un sistema, la cual es, por cierto, aplicable en cualquier espacio de probabilidad. La distancia de Wootters será tratada en la sección 5.3, mientras que su vinculación con la medida de información de Fisher [35] se analizará en una sección posterior. Con este nuevo enfoque de la información, se reformulará el desequilibrio Q y su asociada medida de complejidad [85].

Una medida de complejidad *útil* debería comportarse como una cantidad termodinámica de carácter “intensivo” [33], condición ésta no provista por el desequilibrio basado en la distancia de Wootters, ni por el basado en la distancia Euclídea. Buscando entonces la “intensividad” de Q , en la sección 5.4 se definirán dos tipos de divergencias basadas

en entropías, las divergencias relativas y las Jensen-divergencias [52, 13]. Asimismo se propondrá una extensión de la divergencia de Jensen.

Aplicaciones de estas nuevas medidas de complejidad a ejemplos clásicos de sistemas dinámicos, mostraron significativos avances: se observó que, si bien la C -complejidad (ver sección 4.3.3) detecta los cambios dinámicos debidos a las variaciones de los parámetros de control de los sistemas en estudio, la utilización del Wootters-desequilibrio incrementa notablemente la performance de la medida [85]. Las aplicaciones consideradas en la sección 5.5, pondrán de manifiesto esta situación. La divergencia de Jensen-Shannon por su parte, suma a las bondades mostradas por la medida de complejidad basada en Wootters, la propiedad de generar una medida que se comporta como una cantidad termodinámica intensiva [73].

5.2 Generalización del Desequilibrio

La medida de complejidad C , a partir de la cual se define la complejidad wavelet TWC (ver sección 4.4), está definida como:

$$C(P) = H(P) \cdot Q(P). \quad (5.1)$$

El desequilibrio Q , está dado por:

$$Q = \mathcal{D}(P, P_{max}) = \|P - P_{max}\|_E^2, \quad (5.2)$$

donde P_{max} es la *probabilidad que maximiza la entropía* (máximo desorden) y $\|\cdot\|_E$ simboliza la *distancia Euclídea*.

La propia definición del desequilibrio (5.2), sugiere algunos interrogantes:

- Por qué la distancia al *máximo desorden*?
- Por qué la distancia *euclídea*?

Una apertura en el enfoque, propuesta en el primer interrogante, podría conducir incluso a un paradigma diferente en el concepto de desequilibrio. Por ejemplo, R. Sole y B. Luque [119] propusieron en 1999 una simple generalización del desequilibrio basada en una medida más natural, como es la distancia a la *independencia*, definida por

$$Q^I = \sum_i \sum_j (p_{ij} - p_i \cdot p_j)^2, \quad (5.3)$$

donde p_{ij} representa la probabilidad conjunta para los eventos i y j . Q^I sería una medida de cuán lejos están los subsistemas de ser independientes (es decir de la identidad $p_{ij} = p_i \cdot p_j$). La complejidad asociada a este desequilibrio generalizado Q^I , satisface las condiciones de borde (mínima pauta exigida a una medida de complejidad): se anula tanto en el orden perfecto como en la aleatoriedad total. Además esta cantidad ha mostrado detectar consistentemente la complejidad cercana a puntos críticos en sistemas complejos, con un bajo esfuerzo computacional (ver ref. [119]), lo cual da validez a la generalización propuesta.

El segundo interrogante deja abierta una discusión sobre las condiciones particulares que requiere la distancia \mathcal{D} de la ecuación (5.2) para satisfacer el planteado concepto de desequilibrio. Desde el punto de vista formal, una distancia definida en \mathbb{R}^N es métrica si satisface las siguientes propiedades: *no-negatividad, identidad, simetría y desigualdad triangular*. Si bien las dos primeras condiciones son necesarias para cualquier medida de *discrepancia*, como es \mathcal{D} , las restantes no son esenciales. En efecto, \mathcal{D} es la distancia desde cada probabilidad P a la probabilidad P_{max} que permanece fija. O sea que \mathcal{D} es una distancia direccional, una discrepancia hacia un lado, por lo cual tanto la condición de simetría como la desigualdad triangular son irrelevantes en el presente contexto. Por otra parte, teniendo en cuenta que \mathcal{D} será aplicada sobre distribuciones de probabilidad, resultará natural su generalización en base a una “distancia estadística”.

5.3 Distancia Estadística de Wootters

En 1981, el físico W. K. Wootters [138] ha definido sobre la base de consideraciones puramente estadísticas, el concepto de distancia estadística entre dos distribuciones de probabilidad. Si se desea describir una situación desde un punto de vista estadístico (es decir, en términos de probabilidades), es necesario considerar un conjunto conteniendo un gran número de sistemas preparados idénticamente. La frecuencia de ocurrencia de un evento particular está dada por la fracción de sistemas en el conjunto, caracterizados por la ocurrencia del evento especificado. Ahora bien, la frecuencia de ocurrencia observada típicamente difiere de la probabilidad real. Debido a este error estadístico, es difícil distinguir (en un número finito de pruebas) entre dos preparaciones ligeramente diferentes de un mismo sistema.

En el contexto de la mecánica cuántica, donde Wootters ha aplicado estos conceptos, el problema adopta la siguiente forma: distinguir entre dos estados cuánticos, a partir de mediciones efectuadas sobre los mismos. En este sentido, se establece que dos preparaciones son “indistinguibles” en un número finito de pruebas si la diferencia entre sus probabilidades reales, es menor que el tamaño de una fluctuación típica. Esto sugiere una definición de distancia estadística entre dos estados dados, que solamente “cuente” los estados intermedios “distinguibles”.

El concepto de distancia estadística es independiente de la mecánica cuántica y puede ser definido en cualquier espacio de probabilidad. Sea entonces un experimento probabilístico con N salidas posibles. Las probabilidades $\{p_1, \dots, p_N\}$ de las N salidas varían sobre Ω , un espacio de dimensión $(N - 1)$ restringido por las condiciones de nonnegatividad y normalización que deben cumplir las probabilidades p_i . Igual que en el caso de estados cuánticos, la distancia estadística será proporcional al número de puntos “intermedios” distinguibles. Pero ahora es necesario aclarar qué se entiende por “intermedios”, lo cual requiere varios pasos adicionales. En efecto Wootters introduce las siguientes definiciones [138]:

- Se define la *longitud estadística* de una curva en un espacio de probabilidad, \mathcal{L} ,

como:

$$\mathcal{L} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \left[\text{número máximo de puntos mutuamente distinguibles (en } n \text{ pruebas), a lo largo de la curva} \right]. \quad (5.4)$$

El concepto de distinguibilidad deberá ser extendido ahora al espacio de probabilidades $(N - 1)$ -dimensional. En este caso, si llamamos $\{\xi_i, i = 1, \dots, N\}$ al conjunto de frecuencias reales (después de n pruebas) asociadas a las probabilidades $\{p_i, i = 1, \dots, N\}$, este conjunto está distribuido según una distribución multinomial, que cuando el número de pruebas n es suficientemente grande, puede ser aproximado por una Gaussiana [45]:

$$\rho(\xi) \propto \exp \left[-\frac{n}{2} \sum_{i=1}^N \frac{(\xi_i - p_i)^2}{p_i} \right]. \quad (5.5)$$

- La *región de incertezza* alrededor de un punto $P = (p_1, \dots, p_N)$ estará limitada a todos los puntos ξ tales que el exponente en la ec. (5.5) es menor en valor absoluto que $1/2$.
- Dos puntos P_1 y P_2 serán *distinguibles* en n pruebas, si sus regiones de incertezza no se superponen, o sea:

$$\frac{\sqrt{n}}{2} \left[\sum_{i=1}^N \frac{(p_i^{(1)} - p_i^{(2)})^2}{p_i^{(1)}} \right]^{1/2} > 1. \quad (5.6)$$

En este momento se está en condiciones definir la *distancia estadística* entre dos distribuciones de un espacio de probabilidad N -dimensional. Según Wootters la misma es la longitud estadística de la “menor curva (en el sentido de longitud estadística) que conecta ambos puntos”. A fin de encontrar una expresión explícita de la distancia estadística, se considerará una curva suave parametrizada por t que conecta los dos puntos. Dicha curva llamada $P(t) = (p_1(t), \dots, p_N(t)) \in \Omega$, con $0 \leq t \leq 1$, satisface que $P(0) = P_1$ y $P(1) = P_2$. Su longitud estadística, de acuerdo con el criterio de distinguibilidad dado por la ec. (5.6) y con la definición dada en la ec. (5.4), es:

$$\mathcal{L} = \frac{1}{2} \int_0^1 dt \left\{ \sum_{i=1}^N \frac{1}{p_i(t)} \left[\frac{dp_i(t)}{dt} \right]^2 \right\}^{1/2}, \quad (5.7)$$

que con el cambio de variables $x_i = p_i^{1/2}$ se convierte en

$$\mathcal{L} = \int_0^1 dt \left[\sum_{i=1}^N \left(\frac{dx_i}{dt} \right)^2 \right]^{1/2}. \quad (5.8)$$

La curva $P(t)$ deberá estar en el espacio de probabilidad, por lo tanto se cumplirá que: $1 = \sum_{i=1}^N p_i(t) = \sum_{i=1}^N x_i^2(t)$, lo cual implica que $x(t)$ estará sobre la esfera unidad en el x -espacio. Por lo tanto, la distancia estadística entre P_1 y P_2 será la menor distancia sobre

la esfera unidad entre los vectores $x^{(1)}$ y $x^{(2)}$, cuyas coordenadas son $x_i^{(j)} = \left(p_i^{(j)}\right)^{1/2}$, para $j = 1, 2$. Como la menor distancia es igual al ángulo entre los vectores unidad $x^{(1)}$ y $x^{(2)}$, se tiene que:

Definición 5.1:

Dadas dos distribuciones de probabilidad discretas P_1 y P_2 , se define la *distancia estadística* de Wootters, entre P_1 y P_2 , ($\mathcal{D}_W(P_1, P_2)$), como:

$$\mathcal{D}_W(P_1, P_2) = \cos^{-1} \left[\sum_{i=1}^N \left(p_i^{(1)} \right)^{1/2} \cdot \left(p_i^{(2)} \right)^{1/2} \right]. \quad (5.9)$$

Esta medida, como destaca Wootters, provee un natural camino para definir una distancia en un espacio de probabilidad. Al tener en cuenta las fluctuaciones estadísticas, sólo se consideran las probabilidades mutuamente distinguibles. Además \mathcal{D}_W se anula cuando las dos distribuciones de probabilidad coinciden, y obtiene su máximo valor ($\pi/2$) cuando los dos vectores P_1 y P_2 son ortogonales, es decir cuando a una salida con probabilidad positiva en una distribución, le corresponde probabilidad cero en la otra. Inspirados en el desarrollo que se acaba de exponer y que se debe a Wootters, se reformulará en la sección 5.3.2 el desequilibrio Q .

5.3.1 Relación con la medida de información de Fisher

Se presentará en esta sección la deducción de la distancia estadística de Fisher, en base a la presentación realizada por J. Hilgevoord y J. Uffink [57]. Sea X el conjunto de posibles salidas de un experimento probabilístico y $x \in X$ sus elementos. Aunque en la notación se supondrá X continuo, esta condición no es esencial. Dada una clase de distribuciones sobre X , se puede asignar únicamente a cada distribución un parámetro $\theta \in \Theta$, donde Θ es un espacio de parámetros arbitrario. Así, cada valor del parámetro θ especifica una distribución $P_\theta(x)$. Dado $\theta_0 \in \Theta$ y su distribución asociada $P_{\theta_0}(x)$, la correspondiente Información de Fisher I , introducida en el capítulo 2 (ver sección 2.3.2), está definida por:

$$I(\theta_0) = \int \frac{1}{P_\theta} \left(\frac{dP_\theta}{d\theta} \right)^2 dx \Big|_{\theta=\theta_0} = 4 \int \left(\frac{d\sqrt{P_\theta}}{d\theta} \right)^2 dx \Big|_{\theta=\theta_0}. \quad (5.10)$$

Ha sido observado por varios autores [57, 131] que la información de Fisher provee una geometría sobre el espacio de parámetros Θ . El objetivo de esta sección es introducir este punto de vista, que servirá para la definición de una distancia estadística.

Es útil considerar primero el caso especial de familias de traslación, donde las distribuciones son de la forma $P_\theta(x) = P(x - \theta)$, para $\theta \in \Theta$. En este caso la información $I(\theta_0) = I$ es independiente del valor de θ_0 (ver sección 2.3.2). Por lo tanto, $1/\sqrt{I}$ representa una región de incertezza cuyo tamaño es independiente del verdadero valor θ_0 . $1/\sqrt{I}$ puede ser considerado entonces, una *unidad de incertezza*. Si se “normaliza” la parametrización reemplazando θ por $s = \frac{1}{2}\sqrt{I}\theta$, es fácil ver que: si $|s_1 - s_2|$ es muy

pequeño ($s_i = \frac{1}{2}\sqrt{I}\theta_i$, $i = 1, 2$), sus asociadas distribuciones son difíciles de distinguir experimentalmente. O sea que la normalización del parámetro provee un camino para decir cuán lejos, en un sentido estadístico, están las dos distribuciones. Es decir se introduce una noción de “distancia estadística”.

Si la información I no es independiente de θ_0 , se puede aún realizar localmente esta normalización, obteniendo un elemento infinitesimal de distancia entre θ_0 y $\theta_0 + d\theta$:

$$ds = \frac{1}{2}\sqrt{I(\theta_0)} d\theta , \quad (5.11)$$

y la distancia entre θ_0 y θ_1 , es ahora:

$$s(\theta_0, \theta_1) = \frac{1}{2} \int_{\theta_0}^{\theta_1} \sqrt{I(\theta)} d\theta . \quad (5.12)$$

El significado geométrico de la información I es quizás más evidente si consideramos θ un parámetro multi-dimensional, $\vec{\theta} = (\theta_1, \dots, \theta_M)$. Una extensión directa de (5.10) es

$$\begin{aligned} I_{ij}(\vec{\theta}_0) &= \int \frac{1}{P_{\vec{\theta}}} \frac{\partial P_{\vec{\theta}}}{\partial \theta_i} \frac{\partial P_{\vec{\theta}}}{\partial \theta_j} dx \Big|_{\vec{\theta}=\vec{\theta}_0} \\ &= 4 \int \frac{\partial \sqrt{P_{\vec{\theta}}}}{\partial \theta_i} \frac{\partial \sqrt{P_{\vec{\theta}}}}{\partial \theta_j} dx \Big|_{\vec{\theta}=\vec{\theta}_0} \quad i, j = 1, \dots, M . \end{aligned} \quad (5.13)$$

La información de Fisher es ahora una matriz simétrica y definida positiva. Esto permitirá definir, a semejanza de la ec. (5.11) un elemento de distancia infinitesimal en el espacio de parámetros:

$$ds^2 = \frac{1}{4} \sum_{i,j=1}^M I_{ij}(\vec{\theta}) d\theta_i d\theta_j . \quad (5.14)$$

Desde el punto de vista geométrico, el caso anterior correspondiente a un parámetro unidimensional puede ser visto como un subconjunto, es decir como una curva en un espacio métrico de dimensión más alta. La distancia (5.12) mide entonces la longitud estadística a lo largo de esta curva. Pero este espacio de dimensión más alta permite hacer la distinción entre longitud y distancia, que no sería posible en el espacio unidimensional. La distancia entre dos puntos estará definida como la longitud de la menor curva que los conecta en el espacio de parámetros.

Esta idea es particularmente fructífera si se toma el espacio de parámetros Θ tal que mapee el conjunto de *todas* las distribuciones de probabilidad sobre X . Suponiendo que X contiene un número finito de posibles salidas, $X = \{x_1, \dots, x_N\}$, entonces basta con $(N - 1)$ parámetros independientes para mapear la totalidad de las distribuciones sobre X . El espacio de parámetros Θ es entonces un espacio de dimensión $N - 1$ y la matriz (5.13) adopta la siguiente forma:

$$I_{ij} = 4 \sum_{k=1}^N \frac{\partial \sqrt{P_{\vec{\theta}}(x_k)}}{\partial \theta_i} \frac{\partial \sqrt{P_{\vec{\theta}}(x_k)}}{\partial \theta_j} dx \quad i, j = 1, \dots, N - 1 . \quad (5.15)$$

Eligiendo un sistema de coordenadas apropiado, tal espacio puede ser representado como un subespacio curvado inmerso en un espacio euclídeo de dimensión mayor. En este caso tal representación se obtiene cuando se eligen las coordenadas $y_k(\vec{\theta}) = \sqrt{P_{\vec{\theta}}(x_k)}$. El elemento de distancia toma entonces la forma:

$$ds^2 = \frac{1}{4} \sum_{i,j=1}^{N-1} I_{ij} d\theta_i d\theta_j = \sum_{k=1}^N (dy_k)^2 , \quad (5.16)$$

que es el elemento de distancia usual en el espacio Euclídeo N -dimensional. Θ puede ser representado como una hipersuperficie $(N-1)$ -dimensional en este espacio de dimensión N , determinada por la condición de normalización de las distribuciones de probabilidad: $\sum_k (y_k)^2 = 1$, que representa la superficie de una hiperesfera unidad. La condición $y_k(\vec{\theta}) \geq 0$, para $k = 1, \dots, N$, restringe la superficie a un segmento convexo de la hiperesfera. Por lo tanto, la métrica definida por la información de Fisher da a Θ la geometría de la superficie de una esfera.

Puede considerarse ahora una curva regular unidimensional $\{P_{\theta}\}$ sobre la superficie de esta esfera conectando los puntos representando las distribuciones de probabilidad P_{θ_0} y P_{θ_1} . La menor de tales curvas es, claramente, un sector de un gran círculo y la longitud de esta curva será justamente la longitud de arco entre estos puntos. La distancia estadística entre dos distribuciones de probabilidad P_{θ_0} y P_{θ_1} , por lo tanto está dada por:

$$d_F(P_{\theta_0}, P_{\theta_1}) = \cos^{-1} \left[\sum_{k=1}^N \sqrt{P_{\theta_0}(x_k) \cdot P_{\theta_1}(x_k)} \right] , \quad (5.17)$$

que asimilando las notaciones coincide con la ecuación (5.9) obtenida por Wootters mediante un análisis diferente.

En efecto Wootters define la indistinguibilidad de dos distribuciones de probabilidad en base a las fluctuaciones estadísticas. O sea que analiza la dispersión de ambas distribuciones en función de las frecuencias de ocurrencia (ver ec. (5.5)). Mientras que la información de Fisher encierra en sí misma un significado geométrico. Este conduce de manera natural a la definición de un elemento infinitesimal de distancia (ver ec. (5.16)). En ambos casos utilizando un sistema de coordenadas apropiado, se arriba a la misma fórmula para la distancia estadística.

5.3.2 Desequilibrio según Wootters

La reformulación del desequilibrio en base a la distancia estadística de Wootters (ver ec. (5.9)), generalizará la definición dada en (5.2). Se conviene de aquí en más, que el sistema a estudiar no presenta otra restricción más que la correspondiente a la normalización de la probabilidad. Por lo tanto la distribución de probabilidad que maximiza la entropía será la distribución uniforme $P_e = (1/N, \dots, 1/N)$. Si llamamos Q_W al desequilibrio basado en Wootters, el mismo estará definido por:

Definición 5.2:

Dada una distribución de probabilidad discreta P , el *desequilibrio segun Wootters*, Q_W ,

está definido por:

$$\begin{aligned} Q_W(P) &= Q_0^{(W)} \cdot d_W(P, P_e) \\ &= Q_0^{(W)} \cdot \cos^{-1} \left\{ \sum_{i=1}^N [p(i)]^{1/2} \cdot \left[\frac{1}{N} \right]^{1/2} \right\}, \end{aligned} \quad (5.18)$$

donde $Q_0^{(W)} = 1 / \cos^{-1} \left\{ \left[\frac{1}{N} \right]^{1/2} \right\}$ y $0 \leq Q_W \leq 1$.

Si llamamos C_W a la correspondiente medida de complejidad, la misma estará definida por:

Definición 5.3:

La medida de *complejidad* basada en el desequilibrio según Wootters C_W , es:

$$C_W = H \cdot Q_W \quad (5.19)$$

donde H es la inhomogeneidad definida en la sección 4.3.3.

5.4 Divergencias

Es sabido que conceptos provenientes de teoría de la información intervienen de manera crucial en la formalización y resolución de problemas correspondientes al tratamiento de señales. Distancias y divergencias estadísticas, juegan un rol clave en predicción y estimación, como así también en la detección y trasmisión de señales. Así por ejemplo, la información de Kullback es subyacente a los criterios de selección de modelos. La información de Fisher proporciona precisión óptima de estimadores. En resumen, podemos decir que la cuantificación de la información en general es el corazón del análisis de señales moderno.

Las dos principales nociones de información concernientes a leyes de probabilidad están dadas por: *entropía* o medida de información y *divergencia* o ganancia de información. Esta sección estará consagrada a la introducción de las divergencias.

Siguiendo a M. Basseville [13], se distinguen dos clases de divergencias construidas a partir de *funcionales de entropías*. La primera clase contiene divergencias definidas como entropías relativas y la segunda divergencias definidas como diferencias de entropías. En este último caso su construcción reposa sobre la concavidad de la funcional de entropía considerada. Comenzaremos con una presentación formal de las entropías, definidas sobre espacios de medida más generales. Esta presentación sistemática y general, permitirá una introducción sencilla y natural del concepto de divergencia. Estas medidas serán posteriormente aplicadas al caso particular de espacios de probabilidad.

Sea X un espacio topológico, Ω una σ -álgebra de X y ν una medida definida sobre Ω . Sea además μ , una medida absolutamente continua con respecto a ν . Entonces, para todo $A \in \Omega$ se cumple que:

$$\mu(A) = \int_A \left(\frac{d\mu}{d\nu} \right) d\nu \quad (5.20)$$

donde $d\mu/d\nu \geq 0$ es la derivada de Radom-Nikodym. Se introducirán entropías integrales y no integrales, definidas con respecto a la medida de referencia ν .

Definición 5.4:

Sea (X, Ω, ν) un espacio de medida, $\nu(X) < \infty$ y μ una medida definida sobre Ω absolutamente continua respecto de ν .

- Una *entropía integral* está definida por:

$$S_{h,\nu}(\mu) \triangleq - \int_X h\left(\frac{d\mu}{d\nu}\right) d\nu , \quad (5.21)$$

donde h es una función (generalmente convexa) definida sobre $\mathbb{R}^+ = \{u \in \mathbb{R}, u \geq 0\}$.

- Una *entropía no integral* está definida por:

$$S_{\psi,\nu}(\mu) \triangleq \psi^{-1} \left[\int_X \frac{d\mu}{d\nu} \psi \left(-\ln \frac{d\mu}{d\nu} \right) d\nu \right] , \quad (5.22)$$

donde ψ es una función estrictamente monótona definida sobre \mathbb{R}^+ .

Ejemplos:

- 1) Las dos formas funcionales valen en particular para la *entropía de Shannon*:

$$S_{1,\nu} = - \int_X \frac{d\mu}{d\nu} \ln \frac{d\mu}{d\nu} d\nu , \quad (5.23)$$

que corresponde a la función convexa h definida en \mathbb{R}^+ por:

$$h(u) = h_1(u) = u \ln u . \quad (5.24)$$

- 2) El ejemplo típico de entropía integral es la *entropía de orden q de Havrda-Charvat-Tsallis* [55, 126]:

$$S_{q,\nu}^T(\mu) = \frac{1}{q-1} \int_X \left[\frac{d\mu}{d\nu} - \left(\frac{d\mu}{d\nu} \right)^q \right] d\nu , \quad (5.25)$$

que corresponde a la función h definida en \mathbb{R}^+ por:

$$h(u) = h_q(u) = \frac{1}{q-1} (u^q - u) \quad (5.26)$$

para $q \neq 1$. Utilizando el criterio de la derivada segunda, se ve que h_q es una función convexa si $q \in (0, 1) \cup (1, +\infty)$. En el límite para $q \rightarrow 1$ coincide con la entropía de Shannon:

$$\lim_{q \rightarrow 1} S_{q,\nu}^T(\mu) = S_{1,\nu}(\mu) . \quad (5.27)$$

3) El ejemplo típico de entropía no integral es la *entropía de orden q de Rényi* [102]:

$$S_{q,\nu}^R(\mu) = -\frac{1}{q-1} \ln \int_X \left[\frac{d\mu}{d\nu} \right]^q d\nu , \quad (5.28)$$

que corresponde a la función real ψ definida en \mathbb{R}^+ por:

$$\psi(u) = \psi_q(u) = e^{-(q-1)u} , \quad (5.29)$$

y también en el límite para $q \rightarrow 1$ coincide con la entropía de Shannon:

$$\lim_{q \rightarrow 1} S_{q,\nu}^R(\mu) = S_{1,\nu}(\mu) . \quad (5.30)$$

5.4.1 Entropías relativas

Las divergencias como entropías relativas fueron introducidas independientemente por Ali-Silvey [3] y Csiszar [28] en los años 1966 y 1967, respectivamente.

Sea (X, Ω) un espacio de medida y sean ν y μ dos medidas definidas sobre Ω . Supongamos que μ es absolutamente continua respecto de ν . Entonces, dada la entropía $S_\nu(\mu)$, se le asocia la divergencia o entropía relativa $I(\mu, \nu)$, de μ respecto de ν , definida como:

Definición 5.5:

La divergencia o entropía relativa (también llamada entropía condicional) está dada por:

$$I(\mu, \nu) \triangleq -S_\nu(\mu) . \quad (5.31)$$

Las entropías relativas, se clasifican en:

- **Entropía relativa integral**, si la entropía asociada es integral y será de la forma:

$$I_h(\mu, \nu) = \int_X h \left(\frac{d\mu}{d\nu} \right) d\nu . \quad (5.32)$$

- **Entropía relativa no integral**, si la entropía asociada es no integral y será de la forma:

$$I_\psi(\mu, \nu) = -\psi^{-1} \left[\int_X \frac{d\mu}{d\nu} \psi \left(-\ln \frac{d\mu}{d\nu} \right) d\nu \right] . \quad (5.33)$$

Se analizará el caso particular en el cual μ y ν son medidas de probabilidad. Sea entonces X la recta real, Ω los conjuntos de Borel en \mathbb{R} y consideremos las medidas de probabilidad P_1 y P_2 , con densidades que llamaremos $P_1(x)$ y $P_2(x)$.

Como P_1 es absolutamente continua respecto a P_2 , si $P_2(x) = 0$ entonces $P_1(x) = 0$, y $d\mu/d\nu = dP_1/dP_2 = P_1(x)/P_2(x)$. De este modo se puede expresar la entropía relativa $I(P_1, P_2)$ en términos de las densidades $P_1(x)$ y $P_2(x)$, del siguiente modo:

- Entropía relativa integral:

$$\begin{aligned} I_h(P_1, P_2) &= \int h\left(\frac{dP_1}{dP_2}\right) dP_2 \\ &= \int_{\mathbb{R}} h\left[\frac{P_1(x)}{P_2(x)}\right] P_2(x) dx . \end{aligned} \quad (5.34)$$

- Entropía relativa no integral:

$$\begin{aligned} I_\psi(P_1, P_2) &= -\psi^{-1} \left[\int \frac{dP_1}{dP_2} \psi(-\ln \frac{dP_1}{dP_2}) dP_2 \right] \\ &= -\psi^{-1} \left[\int_{\mathbb{R}} P_1(x) \psi(-\ln \frac{P_1(x)}{P_2(x)}) dx \right] . \end{aligned} \quad (5.35)$$

A modo de ilustración se presentarán a continuación, entropías relativas integrales y no-integrales, expresadas en términos de densidades:

Ejemplos:

- 1) La divergencia asociada a la entropía de Shannon (ver (5.23) y (5.24)) es llamada *divergencia o entropía relativa de Kullback* [72]. La entropía relativa de Kullback, de P_1 con respecto a P_2 resulta igual a:

$$K(P_1, P_2) = \int P_1(x) \ln \left[\frac{P_1(x)}{P_2(x)} \right] dx . \quad (5.36)$$

- 2) La divergencia asociada a la entropía de Havrda-Charvat-Tsallis (ver (5.25) y (5.26)) es llamada *entropía relativa de Havrda-Charvat-Tsallis*, (también conocida como q -Kullback). La q -Kullback entropía relativa de P_1 con respecto a P_2 resulta igual a:

$$K_q^T(P_1, P_2) = -\frac{1}{q-1} \int P_1(x) \left[1 - \left(\frac{P_1(x)}{P_2(x)} \right)^{(q-1)} \right] dx . \quad (5.37)$$

para $q \neq 1$.

- 3) La divergencia asociada a la entropía de Rényi (ver (5.28) y (5.29)) es llamada *divergencia o entropía relativa de Rényi*. La q -entropía relativa de Rényi, de P_1 con respecto a P_2 resulta igual a:

$$K_q^R(P_1, P_2) = \frac{1}{q-1} \ln \int [P_1(x)]^q [P_2(x)]^{(1-q)} dx . \quad (5.38)$$

5.4.2 Divergencias de Jensen

En esta sección se introducirá otra forma de construir divergencias a partir de entropías. Remarquemos que la diferencia de entropías: $S(P_2) - S(P_1)$ no define en general una ganancia de información (o divergencia) satisfactoria, por la simple razón que no es siempre positiva. Un camino posible para evitar esta dificultad es recurrir a la diferencia de Jensen.

Diferencia de Jensen: La *diferencia de Jensen* entre dos distribuciones de probabilidad es por definición la diferencia entre la entropía de la mezcla (lineal) y la media (aritmética) de las dos entropías individuales.

Definición 5.6:

Para $0 \leq \beta \leq 1$, se llama *divergencia de Jensen* asociada a una entropía S , a la expresión J_S^β definida por:

$$J_S^\beta(P_1, P_2) = S(\beta P_1 + (1 - \beta) P_2) - \beta S(P_1) - (1 - \beta) S(P_2). \quad (5.39)$$

La positividad de esta cantidad depende de la concavidad de S . La divergencia de Jensen-Shannon se puede escribir también en función de entropías relativas. Si se denota S_1 a la entropía de Shannon:

$$J_{S_1}^\beta = S_1(\beta P_1 + (1 - \beta) P_2) - \beta S_1(P_1) - (1 - \beta) S_1(P_2) \quad (5.40)$$

$$= \beta K(P_1, \beta P_1 + (1 - \beta) P_2) + (1 - \beta) K(P_2, \beta P_1 + (1 - \beta) P_2), \quad (5.41)$$

donde K es la entropía relativa (o divergencia) de Kullback definida en (5.36). Es decir que la divergencia de Jensen-Shannon es la media (aritmética) de las Kullback-*divergencias* entre cada ley y su mezcla (lineal).

Como S_1 es una función cóncava $J_{S_1}^\beta$ es positiva y $J_{S_1}^\beta(P_1, P_2) = 0$ si y sólo si $P_1 = P_2$. Además es una función simétrica de sus argumentos P_1 y P_2 . Si se calcula la divergencia de Jensen (5.39) asociada a una entropía que no posea propiedades definidas de concavidad, puede perderse la positividad de la medida. Por ejemplo, en el caso de la entropía de Renyi $S = S_q^R$ que sólo es cóncava definida para $0 < q < 1$, no podría definirse la Jensen-Renyi según (5.39) para todo valor q , pues no podría garantizarse su positividad.

Pero la relación entre las ecuaciones (5.40) y (5.41), permite imaginar una generalización de las divergencias de Jensen (5.39) cambiando la entropía relativa en (5.41). Se podría realizar esta generalización reemplazando la entropía relativa de Kullback K por la entropía relativa de Havrda-Charvat-Tsallis K_q^T (ver ec. (5.37)). Se obtiene así la siguiente expresión:

$$\mathcal{J}_{S_q^T}^\beta = \beta K_q^T(P_1, \beta P_1 + (1 - \beta) P_2) + (1 - \beta) K_q^T(P_2, \beta P_1 + (1 - \beta) P_2). \quad (5.42)$$

Las propiedades mencionadas para el caso $J_{S_1}^\beta$ se mantienen. Es directo verificar que $\mathcal{J}_{S_q^T}^\beta \geq 0$ para $q > 0$ y $\mathcal{J}_{S_q^T}^\beta = 0$ si y sólo si $P_1 = P_2$.

Es importante destacar que (5.42) no coincide con la medida que surge de reemplazar S por S_q^T en la expresión (5.39). Efectivamente, usando la desigualdad de Jensen se puede demostrar que, para $q > 1$:

$$\mathcal{J}_{S_q^T}^\beta(P_1, P_2) > J_{S_q}^\beta(P_1, P_2) . \quad (5.43)$$

O sea que la identidad equivalente a la que existe entre (5.41) y (5.40) no se satisface en este caso. Por lo tanto (5.42) define una nueva medida basada en entropías relativas, que generaliza a (5.39).

Similarmente se podría reemplazar la información de Kullback K en (5.41) por la divergencia asociada a la entropía no integral de Rényi $K = K_q^R$. Se obtiene así la siguiente expresión:

$$\mathcal{J}_{S_q^R}^\beta = \beta K_q^R(P_1, \beta P_1 + (1 - \beta)P_2) + (1 - \beta) K_q^R(P_2, \beta P_1 + (1 - \beta)P_2) . \quad (5.44)$$

5.4.3 Desequilibrio según Jensen

En esta sección se reformulará el desequilibrio en base a la divergencia de Jensen asociada a la entropía de Shannon, según la ec. (5.40). Si se llama Q_{JS} al nuevo desequilibrio, el mismo estará definido por:

Definición 5.7:

Dada una distribución de probabilidad discreta P , el *desequilibrio según Jensen-Shannon*, Q_{JS} , está definido por:

$$\begin{aligned} Q_{JS}(P) &= Q_0^{(JS)} \cdot J_S^\beta(P, P_e) \\ &= Q_0^{(JS)} \cdot [S(\beta P + (1 - \beta)P_e) - \beta S(P) - (1 - \beta) S(P_e)] , \end{aligned} \quad (5.45)$$

donde $Q_0^{(JS)} = -[\beta + (1 - \beta)/N] \ln[\beta + (1 - \beta)/N] - (N - 1)((1 - \beta)/N) \ln[(1 - \beta)/N] - (1 - \beta) \ln N$ es un factor de normalización.

Si se llama C_{JS} a la correspondiente medida de complejidad, la misma estará definida por:

Definición 5.8:

La medida de *complejidad* basada en el desequilibrio según Jensen-Shannon, es:

$$C_{JS} = H \cdot Q_{JS} , \quad (5.46)$$

donde H es la inhomogeneidad definida en la sección 4.3.3.

Nótese que Q_{JS} está definida ahora en términos de entropías. Por lo tanto ésta es una cantidad termodinámica extensiva. Luego, la medida de complejidad C_{JS} resulta ser una cantidad termodinámica intensiva [52, 73].

5.5 Aplicaciones IV: Medidas C_W y C_{JS}

El mapa logístico es un sistema disipativo, ecológicamente motivado, descripto por la ecuación en diferencia de primer orden:

$$x_{n+1} = r x_n (1 - x_n), \quad 0 \leq x_n \leq 1, \quad 1 < r \leq 4, \quad (5.47)$$

cuyo comportamiento dinámico está controlado por el parámetro r . Una descripción de la dinámica asociada al mapa logístico se encontrará en la sección 3.6.1.

El mapa logístico constituye un clásico ejemplo generalmente empleado para ilustrar nuevos conceptos en el tratamiento de sistemas dinámicos. Es un simple y bien estudiado sistema que presenta una rica variedad de comportamientos. Dependiendo del valor del parámetro r el mapa puede mostrar un punto estable, duplicación de período o caos. Siguiendo esta idea se discutirá aquí la aplicación de las medidas de complejidad C_W (ver sección (5.3)) y C_{JS} (ver sección (5.4.2)), en el análisis de este mapa. Igual que en la sección 4.5 del capítulo 4, también ahora se considerará una representación binaria del sistema.

La señal s que va a ser estudiada con las nuevas medidas de complejidad C_W y C_{JS} , está definida como: $s_{r,n} \equiv x_n(r)$, donde $n = 1, \dots, M$ y el valor de r varía desde 3.5 hasta 4.0 con un paso $\Delta r = 0.001$. Esto es, para cada valor de r , con paso Δr , se generarán M valores x_n de acuerdo a la ec. (5.47) y estos valores se acumulan formando la señal s . La evolución temporal de la señal s significa entonces, la evolución del parámetro de control r , y por lo tanto la aparición de una rica variedad de comportamientos. Esto hace que sea una señal útil para caracterizar la performance de las medidas de complejidad generalizada C_W y C_{JS} .

Se seguirá la Ref. [77], y para cada valor del parámetro r , se reducirá la dinámica del mapa logístico a una sucesión binaria (0 si $x \leq 1/2$; y 1 si $x > 1/2$), considerándose luego sucesiones binarias de longitud 12 como estados del sistema. Las probabilidades correspondientes son asignadas de acuerdo a la frecuencia de ocurrencia de cada estado luego de, como mínimo, 2^{22} iteraciones. La idea ahora es analizar las descripciones de la señal logística provistas por ambas medidas de complejidad estadística generalizadas, C_W y C_{JS} , lo que nos permitirá calificar sus respectivos comportamientos.

En particular la complejidad C_W , basada en la distancia estadística de Wootters, y la C_{JS} que utiliza la divergencia de Jensen-Shannon, serán comparadas con la medida simple de complejidad C (ver ec. (4.10)). En este caso $S_{max} = \ln N$, por lo tanto $C \equiv C_{LMC}$. Se prestará especial atención a la capacidad demostrada para captar los cambios en la complejidad asociados a la variación (temporal) del parámetro r .

En la Fig. 4.3.a se muestra la evolución temporal (variación continua con el parámetro r) del desequilibrio Q_E y la complejidad original $C \equiv C_{LMC}$ cuando r varía en $3.5 < r < 4.0$. La Fig. 5.1 presenta también el desequilibrio y la complejidad, pero ahora según Wootters. En Fig. 5.1.a se muestra Q_W en función del parámetro r . Mientras que 5.1.b se refiere a C_W .

Se observa en ambos casos (Euclídea y Wootters), un abrupto crecimiento de la complejidad alrededor de $r = 3.56$, con un máximo en la vecindad de $r > r_\infty \approx 3.57$. Pasado este punto las dos medidas se comportan de manera muy diferente. Si bien ambas casi se anulan en las grandes ventanas de estabilidad, en las regiones entre-ventanas la diferencia es notable.

Considérese por ejemplo, $r \in [3.58, 3.62]$ en la Fig. 4.3. Los numerosos picos indican un crecimiento local de la complejidad. Si se compara con el diagrama de órbitas (ver Fig. 3.5), todo indica que los picos coinciden con las ventanas periódicas. La conclusión es inevitable. La original medida de complejidad $C \equiv C_{LMC}$ le asigna a las ventanas periódicas una mayor complejidad que a la circundante zona caótica. Notar en cambio que la medida de Wootters crece en la región entre ventanas y rápidamente cae en las ventanas periódicas.

Claramente C_W se comporta en concordancia con el bien conocido esquema de bifurcaciones del mapa logístico. No sucede lo mismo con la C -complejidad. El árbitro que puede juzgar la controversia entre las dos medidas es el comportamiento del exponente de Lyapunov. El valor de λ y por lo tanto el grado de caoticidad asociado, crece con r , alcanzando un máximo en $r = 4$. Una sensible medida de complejidad estadística debería acompañar tal crecimiento. En otras palabras, una razonable medida de complejidad estadística debería tomar muy pequeños valores en $r = r_\infty$ y crecer junto con el grado de caoticidad hasta que la caoticidad sea tan grande que la complejidad caiga nuevamente. Cuando prevalece totalmente el caos, la complejidad estadística debería anularse. Nuestra medida modificada según Wootters se comporta precisamente de esta manera, como uno claramente lo ve en Fig. 5.1.

La Fig. 5.2 muestra la evolución temporal de la complejidad de Jensen-Shannon. Puede apreciarse claramente el hecho que la C_{JS} puede distinguir entre diferentes periodicidades, capacidad que es natural pedirle a una adecuada medida de complejidad (ver Ref. [33]).

La Fig. 5.3 amplifica una zona interesante comprendida entre $3.845 \leq r \leq 3.859$ para mostrar con claridad algunos detalles. Nuevamente, la medida de C -complejidad es: *i)* incapaz de distinguir diferentes periodicidades, *ii)* asigna mayor complejidad a regiones periódicas que a su vecinas zonas caóticas. La C_{JS} , por su parte, se comporta de una manera muy diferente (y aparentemente mas correcta). Ella diferencia entre distintos grados de periodicidad (ver el diagrama de bifurcaciones en la Fig. 3.5).

5.6 Conclusiones

La noción de complejidad estadística desarrollada por López-Ruiz, Mancini y Calbet C_{LMC} [77] constituye un importante paso hacia la construcción de una familia de medidas que son fáciles de calcular e intuitivamente coherentes. Ella exhibe no obstante algunas dificultades que han sido claramente puntualizadas por varios autores, como por ejemplo, en Refs. [33, 5].

Profundizando entonces la dirección iniciada por LMC y continuada luego por Shiner, Davison y Landsberg, hemos revisado la medida de complejidad estadística C (ver ecuación 4.10), reemplazando su definición por una que incorpora, en lugar del desequilibrio Euclídeo $Q = Q_E$, el inducido por la distancia estadística de Wootters [138], Q_W . Aplicaciones de esta nueva medida de complejidad estadística, C_W , a la señal generada con el mapa logístico, muestran que se logran importantes mejoras en el comportamiento de la medida de complejidad estadística. En efecto, al variar el parámetro de control r , C_W se manifiesta de manera compatible al exponente de Lyapunov, mientras que la medida original C presenta dificultades.

En cuanto a la medida de divergencia de Jensen-Shannon (JS), tres problemas básicos son solucionados, si se usa el desequilibrio Q_{JS} inducido por la divergencia de JS, a saber:

- . C_{JS} se comporta como una cantidad termodinámica “intensiva”,
- . C_{JS} distingue diferentes grados de periodicidad, y
- . C_{JS} describe consistentemente la dinámica del sistema.

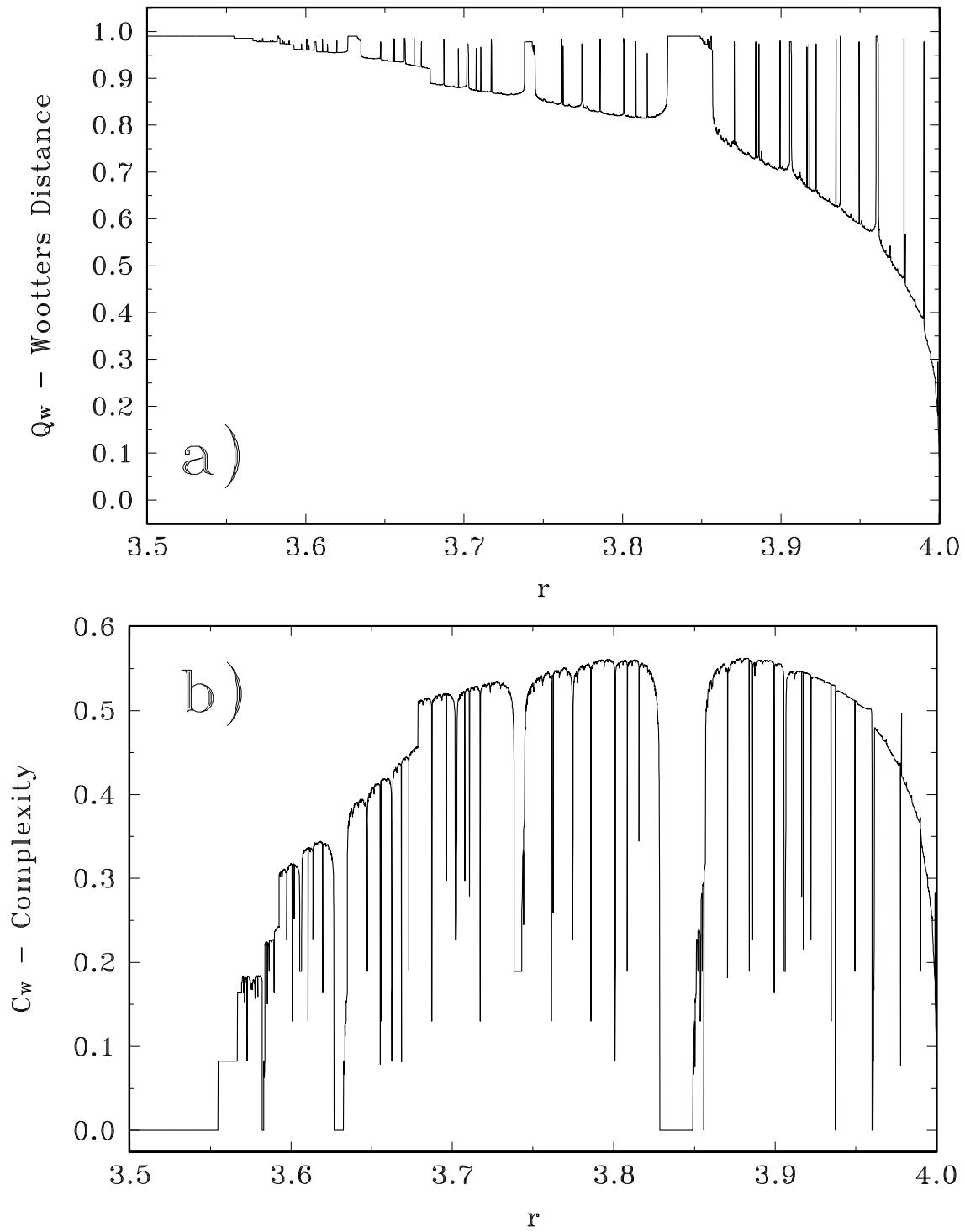


Figura 5.1: MAPA LOGÍSTICO (SUCESIÓN BINARIA): a) DESEQUILIBRIO DE WOOTTERS (Q_W) Y b) MEDIDA DE COMPLEJIDAD C_W EN FUNCIÓN DEL PARÁMETRO DE CONTROL r .

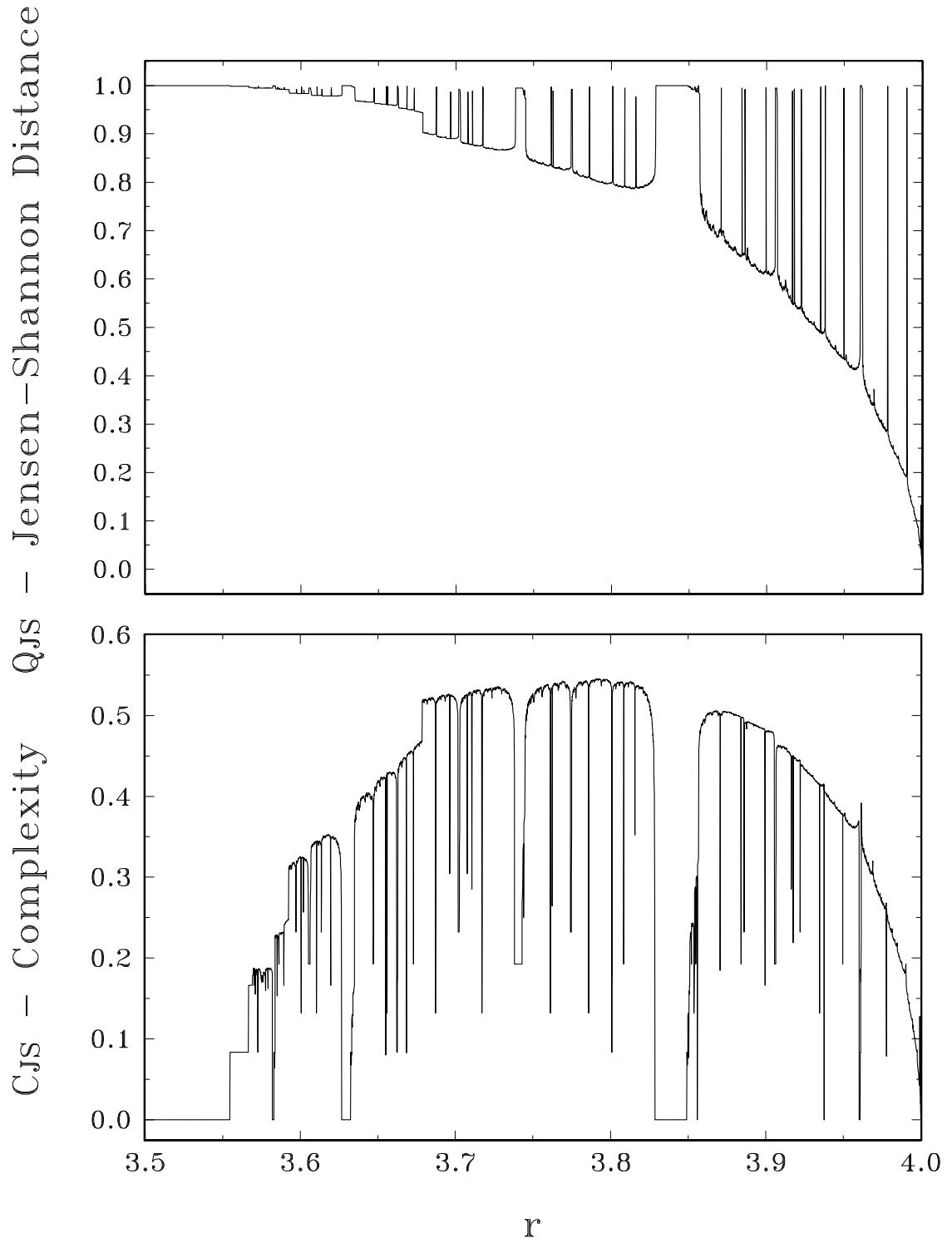


Figura 5.2: MAPA LOGÍSTICO (SUCESIÓN BINARIA): *a)* DESEQUILIBRIO DE JENSEN-SHANNON (Q_{JS}) Y *b)* MEDIDA DE COMPLEJIDAD C_{JS} EN FUNCIÓN DEL PARÁMETRO DE CONTROL r .

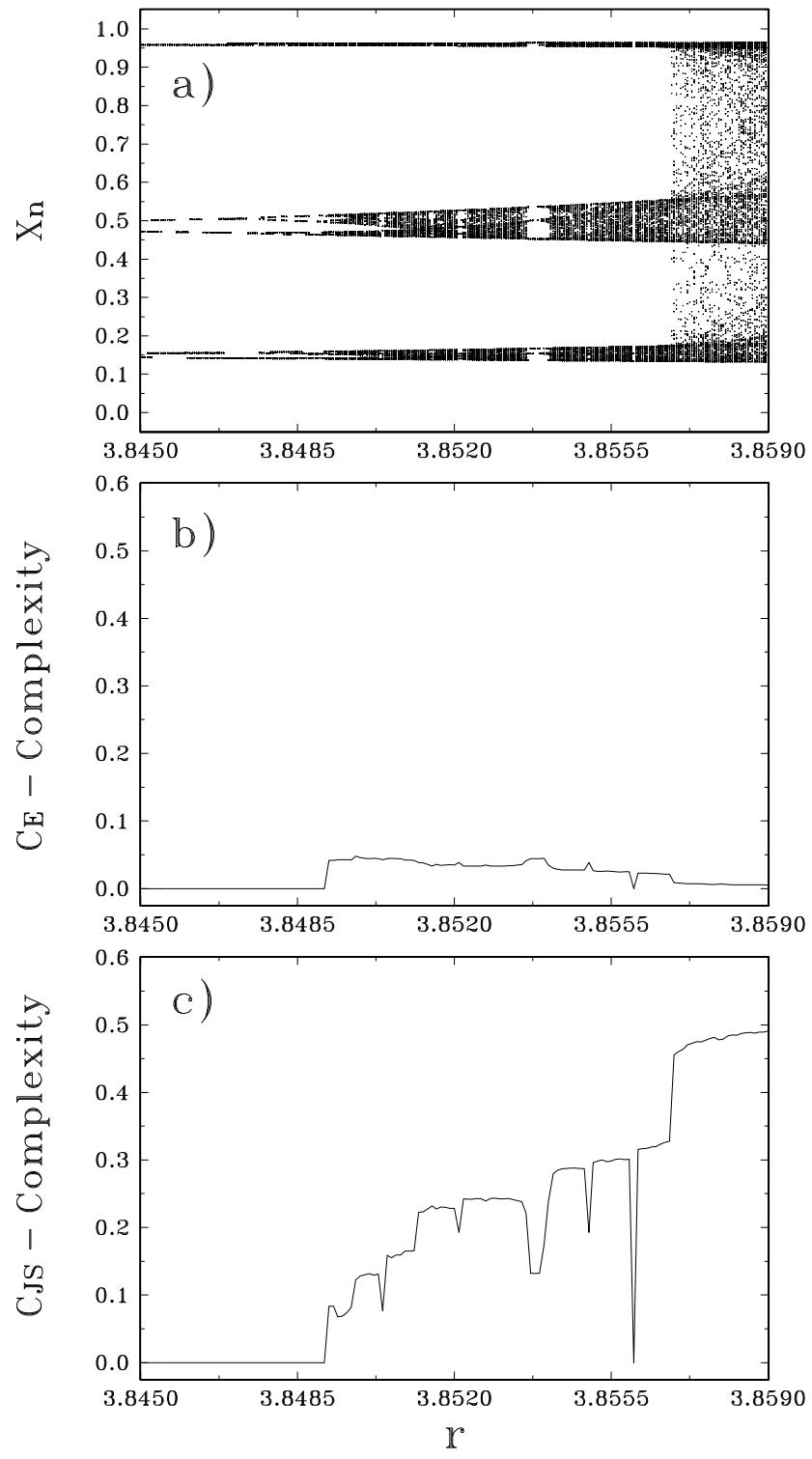


Figura 5.3: Grafico de cantidades de interes en funcion del parametro r restringido al rango asociado a la ventana de periodo 3. a) diagrama de orbitas, b) medida de complejidad C_{LMC} , c) medida de complejidad C_{JS} .

Capítulo 6

Medidas de Complejidad (III):

Una Familia de Medidas

6.1 Introducción

En el capítulo 5 se trató la noción de distancia con relación a su rol en la formulación del desequilibrio Q . En base a los desarrollos que allí se realizaron, se generalizarán en este capítulo los dos factores que definen la medida de complejidad C original (ver ec. (4.10)): la “inhomogeneidad” H y el “desequilibrio” Q . Esto dará lugar a una “familia de medidas de complejidad estadística generalizada”, cuyos integrantes serán llamados $C_{\nu,q}^{(\kappa)}$, para adecuados valores de los parámetros.

En la sección 6.3 se realizará un estudio analítico y geométrico de aspectos relevantes de esta familia de medidas. En particular se tratará de caracterizar los valores extremos de las medidas de complejidad estadística generalizada, realizando este análisis en referencia al gráfico “complejidad versus la entropía normalizada”. Estos diagramas “complejidad-entropía” o C vs H , utilizados inicialmente por Crutchfield y colaboradores [26], resultan ser muy elocuentes ya que al usar coordenadas pertenecientes al ámbito de teoría de información, proporcionan un enfoque que es explícitamente independiente de los parámetros del sistema. Esto permite que sistemas muy diferentes puedan compararse directamente en términos de la información intrínseca del proceso.

En la subsección 6.3.3 se presentará un algoritmo general que permite en forma sistemática obtener cotas para los valores de la complejidad en el diagrama C vs H , cuando C pertenece a la “familia de medidas de complejidad estadística generalizada”.

Por último, los resultados obtenidos al aplicar medidas de complejidad al mapa logístico, que ya fueron presentados en la sección 5.5, serán ahora graficados en diagramas complejidad-entropía. Si junto a dichos resultados se grafican también las correspondientes cotas de la complejidad en función de H (calculadas con el referido algoritmo), se podrán deducir nuevas conclusiones que serán expuestas en la sección 6.4.

6.2 Medidas Generalizadas

6.2.1 Inhomogeneidad basada en entropías de orden q .

Dada una distribución de probabilidades discreta $P = (p_1, \dots, p_N)$, su inhomogeneidad $H(P)$ en términos de la entropía de Shannon S (ver sección 4.3.3), está dada por:

$$H(P) = S(P) / S_{max} , \quad (6.1)$$

donde S_{max} es el máximo valor que puede tomar la entropía. Supondremos por el momento, que el sistema no presenta otra restricción más que la correspondiente a la normalización de las probabilidades, y por lo tanto la entropía es maximizada por la probabilidad uniforme $P_e = (1/N, \dots, 1/N)$. Con lo cual $S_{max} = S(P_e) = \ln N$.

Generalizando la clásica entropía de Shannon, se ha realizado en el capítulo 5 una presentación formal de las *entropías* definidas sobre espacios de medida generales [83, 109]. Estas fueron introducidas a través de formas funcionales, clasificándolas en entropías integrales y no integrales, como puede verse en la sección 5.4. Entropías paramétricas de orden q [102], aparecen en ambas clases como ejemplos típicos. Así, la entropía *de orden q de Havrda-Charvat-Tsallis*, ec. (5.25), es un clásico exponente de entropía integral, mientras que la entropía *de orden q de Rényi*, ec. (5.28), es un ejemplo de entropía no integral. Es interesante notar que en el límite para $q \rightarrow 1$, ambas medidas convergen a la entropía de Shannon.

Las expresiones de las entropías de orden q , vistas en la sección 5.4 para medidas generales, adoptan en el caso de probabilidades discretas las siguientes expresiones:

- a) La entropía de orden q de Havdra-Charvat-Tsallis, $S_q^{(T)}$, está dada por:

$$S_q^{(T)}(P) = \frac{1}{q-1} \sum_{i=1}^N (p_i - p_i^q) . \quad (6.2)$$

- b) La entropía de orden q de Rényi, $S_q^{(R)}$, está dada por:

$$S_q^{(R)}(P) = \frac{1}{1-q} \ln \left[\sum_{i=1}^N p_i^q \right] . \quad (6.3)$$

La inhomogeneidad H basada en las anteriores q -entropías, se presenta ahora de acuerdo a las siguientes expresiones:

- La *inhomogeneidad de orden q de Havdra-Charvat-Tsallis*, $H_q^{(T)}$, es:

$$H_q^{(T)}(P) = S_q^{(T)}(P) / S_{max}^{(T)} , \quad (6.4)$$

con

$$S_{max}^{(T)} = S_q^{(T)}(P_e) = \frac{1 - N^{1-q}}{q-1} , \quad (6.5)$$

donde $q \in (0, 1) \cup (1, \infty)$.

- La *inhomogeneidad de orden q de Rényi*, $H_q^{(R)}$, es:

$$H_q^{(R)}(P) = S_q^{(R)}(P) / S_{max}^{(R)}, \quad (6.6)$$

con

$$S_{max}^{(R)} = S_q^{(R)}(P_e) = \ln N, \quad (6.7)$$

donde $q \in \mathbb{R}$ y $q \neq 1$.

Ambas inhomogeneidades se reducen, claramente, a la inhomogeneidad H original de la ec. (6.1), en el límite $q \rightarrow 1$:

$$\lim_{q \rightarrow 1} H_q^{(T)}(P) = \lim_{q \rightarrow 1} H_q^{(R)}(P) = H_1(P) = H(P). \quad (6.8)$$

6.2.2 Desequilibrio basado en distancias generalizadas

El desequilibrio Q de una distribución de probabilidad discreta P , ha sido presentado en el capítulo 4. Su definición en términos de una distancia \mathcal{D} , está dada por:

$$Q(P) = Q_0 \cdot \mathcal{D}(P, P_e), \quad (6.9)$$

donde Q_0 es una constante de normalización (ver sección 4.3.3).

Q representa una medida normalizada de la “distancia” entre P y P_e . Podría adoptarse en (6.9) un punto de vista diferente y considerar la *ganancia de información que surge de reemplazar P_e por P* . Las “divergencias”, al ser medidas de información condicionales, de una probabilidad respecto de otra, satisfacerían este enfoque. En efecto, medidas de divergencia (entropías condicionales o diferencias de entropías), brindan una natural opción en el concepto utilizado para medir desequilibrio.

Se han considerado dos clases de divergencias (ver sección 5.4): entropías relativas y divergencias de Jensen. Como es sabido, también en este contexto las medidas paramétricas de orden q tienen su representación. Discutiremos a continuación la extensión del desequilibrio en base a ambas divergencias.

Entropías relativas de orden q

Sean P_1 y P_2 dos distribuciones de probabilidad discretas, denotadas por: $P_j = (p_1^{(j)}, \dots, p_N^{(j)})$, para $j = 1, 2$.

La entropía relativa de P_1 con respecto a P_2 asociada a la medida de Shannon, es la *entropía relativa de Kullback* (ver ec. (5.36)). En el caso discreto se presenta como:

$$K(P_1, P_2) = \sum_{i=1}^N p_i^{(1)} \cdot \ln \left[\frac{p_i^{(1)}}{p_i^{(2)}} \right]. \quad (6.10)$$

Por otra parte, entropías condicionales asociadas a entropías paramétricas han sido vistas en la sección 5.4.1 para el caso continuo. Las desarrollaremos a continuación para el caso discreto:

a) La entropía relativa de orden q de Havdra-Charvat-Tsallis, $K_q^{(T)}$, está dada por:

$$K_q^{(T)}(P_1, P_2) = \frac{1}{q-1} \sum_{i=1}^N \left(p_i^{(1)} \right)^q \left[\left(p_i^{(2)} \right)^{1-q} - \left(p_i^{(1)} \right)^{1-q} \right]. \quad (6.11)$$

para $q \neq 1$.

b) La entropía de orden q de Rényi, $K_q^{(R)}$, está dada por:

$$K_q^{(R)}(P_1, P_2) = \frac{1}{q-1} \ln \left[\sum_{i=1}^N \left(p_i^{(1)} \right)^q \left(p_i^{(2)} \right)^{1-q} \right]. \quad (6.12)$$

Se cumple que:

$$\lim_{q \rightarrow 1} K_q^{(T)}(P_1, P_2) = \lim_{q \rightarrow 1} K_q^{(R)}(P_1, P_2) = K_1(P_1, P_2) = K(P_1, P_2). \quad (6.13)$$

Observación 1: Simbolicemos con K_q una entropía relativa, ya sea del tipo Kullback (K_1), Tsallis o Renyi. Se puede comprobar mediante el cálculo directo que $K_q(P, P_e)$ se reduce a:

$$K_q(P, P_e) = k_q [S_q(P_e) - S_q(P)], \quad (6.14)$$

donde S_q es la correspondiente entropía asociada y k_q es un factor constante ($k_q = 1$ en el caso Kullback y Renyi, $k_q = N^{1-q}$ en el caso Tsallis).

Teniendo en cuenta la *Observación 1*:

- El *desequilibrio de orden q basado en entropías relativas*, Q_q , se reduce a:

$$Q_q = k_q [1 - H_q] \quad (6.15)$$

donde H_q es la correspondiente inhomogeneidad y k_q es el factor constante definido en la ec. (6.14).

Divergencias de Jensen de orden q

El desequilibrio (6.9), puede ser también extendido en base a las divergencias de Jensen de orden q presentadas en la sección 5.4.2, a saber:

- El *desequilibrio de orden q de Jensen-Tsallis*, $Q_{\mathcal{J}_q}^{(T)}$, es:

$$Q_{\mathcal{J}_q}^{(T)}(P) = Q_0^{\mathcal{J}T} \cdot \mathcal{J}_{K_q^{(T)}}(P, P_e), \quad (6.16)$$

siendo $Q_0^{\mathcal{J}T}$ un factor de normalización. O sea que $Q_0^{\mathcal{J}T}$ es igual al inverso del máximo valor posible de $\mathcal{J}_{K_q^{(T)}}$, es decir aquel que se obtiene cuando una componente de P es igual a 1 y las demás obviamente 0.

- El desequilibrio de orden q de Jensen-Rényi, $Q_{\mathcal{J}_q}^{(R)}$, es:

$$Q_{\mathcal{J}_q}^{(R)}(P) = Q_0^{\mathcal{J}R} \cdot \mathcal{J}_{K_q^{(R)}}(P, P_e), \quad (6.17)$$

siendo $Q_0^{\mathcal{J}R}$ un factor de normalización. O sea que $Q_0^{\mathcal{J}R}$ es igual al inverso del máximo valor posible de $\mathcal{J}_{K_q^{(R)}}$, es decir aquel que se obtiene cuando una componente de P es igual a 1 y las demás obviamente 0.

6.2.3 Complejidad generalizada

En base a la forma funcional planteada $C = H \cdot Q$, se definirá a continuación una familia de medidas de complejidad que contemple las extensiones planteadas para la inhomogeneidad y el desequilibrio. El desequilibrio en particular, puede concebirse como ya se ha visto, a través de: *a*) una distancia (“longitud comprendida”), o *b*) una divergencia (“información ganada”).

En el caso *a*), podemos a su vez plantear:

- distancia métrica.
- distancia estadística.

En el caso *b*), la divergencia por su parte puede ser:

- entropía relativa generalizada.
- divergencia de Jensen generalizada.

Llamaremos $C_{\nu,q}^{(\kappa)}$ a la familia generalizada de medidas de complejidad estadística, según la siguiente notación:

$$C_{\nu,q}^{(\kappa)}[P] = Q_\nu[P] \cdot H_q^{(\kappa)}[P]. \quad (6.18)$$

donde $\kappa = S, R$ y T indica la entropía de Shannon, Rényi o Tsallis con el correspondiente parámetro q . Notar que para el caso de Shannon tenemos $q = 1$ y en consecuencia este parámetro será omitido para simplificar la notación. $\nu = E, W, K, K_q, J$ y \mathcal{J}_q indican que el desequilibrio es evaluado con las medidas: Euclídea, Wootters, Kullback, Kullback extendida de orden q , Jensen o Jensen extendida de orden q respectivamente. Notar que cuando $\kappa = S$ ($q = 1$) y $\nu = E$ se tiene la definición original de la complejidad estadística C .

Observacion 2: Llámemos Q_q al desequilibrio formulado de acuerdo a las pautas del caso *b*). Como se observó ya en la ecuación (6.2.3), $Q_q = k_q[1 - H_q]$, donde H_q es la inhomogeneidad asociada y k_q un factor constante (ver ec. (6.2.3)). Lo cual deriva en una complejidad que, a menos de un factor constante, tiene la siguiente forma $C_q = H_q(1 - H_q)$, en coincidencia con la medida C_{SDL} (ver sección 4.3.2), definida por Shiner, Davison y Landsberg [116].

Han sido formuladas numerosas críticas en relación a C_{SDL} , motivadas en gran parte por ser C_{SDL} una función *sólo* de la entropía del sistema [33]. No parece ser particularmente útil que todos los sistemas con la misma entropía tengan la misma complejidad estadística.

Sin embargo, es importante destacar que las medidas de la familia $C_{\nu,q}^{(\kappa)}$ no se comportan en general como una función *simple* de la entropía, en el sentido que para un dado valor de $H^{(\kappa)}$ y un conjunto de sistemas con entropía igual a $H^{(\kappa)}$, existe un amplio rango de valores para la complejidad de dichos sistemas, variando entre un valor mínimo C_{min} y un valor máximo C_{max} [5, 16]. Por lo tanto, evaluar la medida de complejidad $C_{\nu,q}^{(\kappa)}$ proporcionará, en general, información adicional que tendrá en cuenta otras peculiaridades de la distribución de probabilidad del sistema. De este modo resultará interesante estudiar las características particulares de esta familia de medidas lo que conducirá finalmente a la obtención de las mencionadas cotas C_{min} y C_{max} .

6.3 Análisis de la Familia de Medidas

Investigaremos en esta sección los valores extremos que alcanzan las medidas de complejidad generalizada $C_{\nu,q}^{(\kappa)}$ sobre el espacio de probabilidad $\Omega = \{p_i : \sum p_i = 1, p_i \geq 0\}$. El procedimiento a seguir es independiente de la medida elegida tanto para la inhomogeneidad $H_q^{(\kappa)}$, como para el desequilibrio Q_ν .

Un miembro particular de esta familia, $C_{E,1}^{(S)} \equiv C_{LMC}$, ha sido ampliamente explorado por Anteneodo y Plastino [5]. Mediante simulaciones de Monte Carlo los mencionados autores encontraron cuál es la distribución de probabilidad discreta que maximiza la complejidad C_{LMC} . Por su parte también Lopez Ruiz, Mancini y Calbet [16] mostraron que la medida de complejidad por ellos definida C_{LMC} , permanece dentro de ciertas cotas C_{max} y C_{min} , si se la grafica en el plano *complejidad versus entropía*.

El objetivo propuesto en esta sección es elaborar un procedimiento *sistemático* que permita obtener cotas para los valores de $C_{\nu,q}^{(\kappa)}$ cuando $H_q^{(\kappa)}$ permanece constante. El análisis incluirá por supuesto como una instancia particular, el desarrollo realizado por Anteneodo y Plastino [5].

Para simplificar la notación, en lo sucesivo omitiremos los índices κ, ν, q excepto cuando ellos sean necesarios. Es decir que $C(P) = H(P) \cdot Q(P)$ simbolizará la complejidad estadística generalizada (6.18).

6.3.1 Caracterización analítica de los extremos

Extendemos aquí el formalismo propuesto por Anteneodo y Plastino en [5], para mostrar que si $H_q^{(\kappa)}$ permanece constante, las configuraciones $P = (p_1, \dots, p_N)$ que optimizan $C_{\nu,q}^{(\kappa)}$ admiten a lo sumo dos valores distintos de las probabilidades p_i . Trabajaremos en el plano *complejidad versus entropía*, al que denotaremos $(H \times C)$.

Notemos que, la inhomogeneidad $H \equiv H_q^{(\kappa)}$ y el desequilibrio $Q \equiv Q_\nu$ referidos en la notación (6.18), son funciones definidas en el espacio de probabilidad Ω , que responden

a la siguiente forma funcional:

$$H(P) = \phi \left(\sum_{i=1}^N f(p_i) \right) , \quad Q(P) = \psi \left(\sum_{i=1}^N g(p_i) \right) , \quad (6.19)$$

donde $\phi(u)$ y $\psi(u)$ son funciones continuas estrictamente monótonas. Además las segundas derivadas de las funciones $f(u)$ y $g(u)$ satisfacen:

$$f''(u) = A_1 u^{\alpha_1} , \quad g''(u) = A_2 u^{\alpha_2} , \quad (6.20)$$

con $A_j, \alpha_j \in \mathbb{R}$ para $j = 1, 2$.

El siguiente teorema, fundamental en la deducción de posteriores resultados, caracterizará las configuraciones extremas de C .

Teorema 6.1:

Las distribuciones de probabilidad P que son extremos de C para H constante, responden a la siguiente configuración:

$$P = \begin{cases} 0 & \text{para } 1 \leq i \leq m \\ p & \text{para } m+1 \leq i \leq m+n \\ (1-p \cdot n) / (N-m-n) & \text{para } m+n+1 \leq i \leq N \end{cases} \quad (6.21)$$

con $n \in \mathbb{N}$.

Demostración:

Como $C = H \cdot Q$, encontrar extremos de C para H constante, es equivalente a encontrar extremos de Q . Los extremos de Q estarán sujetos a dos restricciones:

- a) normalización de las probabilidades: $\sum_{i=1}^N p_i = 1$.
- b) valor constante de la entropía: $H(P) = h$.

El método de multiplicadores de Lagrange conduce a la siguiente variación:

$$\delta \left(Q(P) - \lambda_1 \left[\left(\sum_{i=1}^N p_i \right) - 1 \right] - \lambda_2 [H(P) - h] \right) = 0 , \quad (6.22)$$

donde λ_1 y λ_2 son multiplicadores de Lagrange.

Teniendo en cuenta (6.19) y (6.20), se calculan $\partial Q / \partial p_i$ y $\partial H / \partial p_i$, para $i = 1, \dots, N$ y reemplazando en (6.22), se obtiene:

$$\phi' \left(\sum_{j=1}^N f(p_j) \right) \cdot f'(p_i) - \lambda_1 - \lambda_2 \cdot \psi' \left(\sum_{j=1}^N g(p_j) \right) \cdot g'(p_i) = 0 , \quad (6.23)$$

para $i = 1, \dots, N$.

Llamando $F(p_i)$ a la ec. (6.23) y reescribiendo sus términos, se tiene:

$$F(p_i) = f'(p_i) - a \cdot g'(p_i) - b = 0 \quad \text{para } i = 1, \dots, N , \quad (6.24)$$

donde

$$\begin{aligned} a &= \lambda_2 \psi' \left(\sum_{j=1}^N g(p_j) \right) / \phi' \left(\sum_{j=1}^N f(p_j) \right), \\ b &= \lambda_1 / \phi' \left(\sum_{j=1}^N f(p_j) \right). \end{aligned} \quad (6.25)$$

Notar que a y b están bien definidas y son no nulas debido a la propiedad de monotonía de las funciones ϕ y ψ . Además a y b son iguales para todos los p_i , $i = 1, \dots, N$, por lo tanto en (6.24) pueden ser pensadas como constantes. Si se calcula la función derivada del miembro izquierdo de la ec. (6.24), resulta:

$$F'(u) = f''(u) - a \cdot g''(u), \quad (6.26)$$

y es fácil ver que $F'(u)$ tendrá a lo sumo una raíz real positiva. En efecto, teniendo en cuenta la requerida forma funcional para $f''(u)$ y $g''(u)$ (ver (6.20)), $F'(u) = 0$ implica una ecuación de la forma $u^\alpha = \beta$ con $\alpha, \beta \in \mathbb{R}$, que tiene a lo sumo una raíz real positiva. Por lo tanto recurriendo al teorema del valor medio, inferimos que (6.24) no puede tener más de dos raíces reales positivas. Lo cual implica que los p_i no nulos admiten a lo sumo dos valores distintos, como expresa la ec. (6.21).

Ahora bien, el requerimiento de positividad de las probabilidades ($p_i \geq 0$) determina un subconjunto \mathcal{A} de la variedad $(N - 2)$ -dimensional definida por las restricciones de normalización y entropía constante. El método de multiplicadores de Lagrange brinda los extremos en el interior de \mathcal{A} . Siendo \mathcal{A} un dominio acotado, los extremos podrían estar también en el contorno. Supongamos entonces que el extremo P está en el contorno de \mathcal{A} , entonces al menos una coordenada p_i se anula. Supongamos que P tiene m coordenadas nulas. Con un razonamiento similar al anterior pero ahora sobre el espacio de dimensión $N - m$, se deduce que las $(N - m)$ coordenadas no nulas de P admiten a lo sumo dos valores distintos. Es decir que los $P \in \Omega$ que son extremos de C sujetos a H constante, son de la forma (6.21). Con lo cual se completa la prueba. ■

En el caso particular de $C_{E,q=1}^{(S)} \equiv C_{LMC}$, Anteneodo y Plastino [5] mostraron mediante argumentos analíticos, que el extremo de C_{LMC} se obtiene en configuraciones P de la forma (6.21), y luego usando procedimientos numéricos concluyeron que en el valor de n en la ec. (6.21), debe ser $n = 1$. Nos interesa investigar si este resultado $n = 1$, continúa siendo válido al extenderlo al caso general $C_{\nu,q}^{(\kappa)}$.

6.3.2 Formalismo geométrico

Para analizar propiedades geométricas del espacio de probabilidad $\Omega \subset \mathbb{R}^N$, se introducirán previamente algunas definiciones.

Definiciones generales

Definición 6.1:

Un conjunto $\{a_1, a_2, \dots, a_N\}$ de vectores de \mathbb{R}^N es *geométricamente independiente* si

para cualquier conjunto de escalares $t_i \in \mathbb{R}$, se cumple que:

$$\text{Si } \sum_{i=1}^N t_i = 0 \wedge \sum_{i=1}^N t_i a_i = 0 \quad \text{entonces} \quad \{t_i = 0, \text{ para } i = 1, \dots, N\}. \quad (6.27)$$

Definición 6.2:

Sea $\{a_1, a_2, \dots, a_N\}$ un conjunto de vectores geométricamente independiente de \mathbb{R}^N . Se llama $(N-1)$ -simplex expandido por a_1, a_2, \dots, a_N al siguiente conjunto, Δ^{N-1} :

$$\Delta^{N-1} = \left\{ x \in \mathbb{R}^N : x = \sum_{i=1}^N t_i a_i \quad \text{con} \quad \sum_{i=1}^N t_i = 1 \quad \text{y} \quad t_i \geq 0, \forall i \right\}. \quad (6.28)$$

En un sentido informal, un $(N-1)$ -simplex es una generalización $(N-1)$ dimensional de un triángulo en \mathbb{R}^2 [71]. Cada a_i es un *vértice* y los lados son los segmentos que unen cada par de vértices. Se llama *cara* del simplex al subsimplex expandido por un subconjunto propio de $\{a_1, a_2, \dots, a_N\}$, y *centroide o baricentro* al vector

$$\mu_{N-1} = \sum_{i=1}^N \frac{a_i}{N}. \quad (6.29)$$

Por otra parte, los vectores P de Ω expresados en la base canónica de \mathbb{R}^N , adoptan la siguiente forma:

$$P \in \Omega \iff P = \sum_{i=1}^N p_i e_i, \quad (6.30)$$

con $\sum_{i=1}^N p_i = 1$ and $p_i \geq 0$. Por lo tanto, Ω es el $(N-1)$ -simplex expandido por $\{e_1, e_2, \dots, e_N\}$:

$$\Omega = \Delta^{N_1} = [e_1, \dots, e_N], \quad (6.31)$$

donde llamamos $N_1 = N - 1$ y $[e_1, \dots, e_N]$ es el conjunto de combinaciones convexas de los N vectores geométricamente independientes e_i en \mathbb{R}^N .

Subdivisión baricéntrica

Sea φ un mapeo biyectivo $\varphi : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$ y sea $b_\varphi^j \in \Delta^{N_1}$ el punto

$$b_\varphi^j \triangleq \text{baricentro de } [e_{\varphi(j)}, \dots, e_{\varphi(N)}] = \sum_{i=\varphi(j)}^{\varphi(N)} e_{\varphi(i)} / (N - j + 1), \quad (6.32)$$

para $j = 1, \dots, N$. El punto b_φ^j se llama vértice de orden j por ser el *centroide* de un subsimplex de dimensión $(N-j)$. Por ejemplo, b_φ^1 es el centroide del N_1 -simplex *entero*. Mientras que b_φ^N es el centroide de *un punto*.

Se llamará $\Delta_\varphi^{N_1} \subset \Delta^{N_1}$ al N_1 -simplex siguiente:

$$\Delta_\varphi^{N_1} = [b_\varphi^1, \dots, b_\varphi^N]. \quad (6.33)$$

Definición 6.3:

Se llama *subdivisión baricéntrica* del N_1 -simplex, al siguiente conjunto de subsimplices:

$$\mathcal{SB} \triangleq \{\Delta_\varphi^{N_1} = [b_\varphi^1, \dots, b_\varphi^N], \varphi \in \mathcal{P}_N\}, \quad (6.34)$$

donde $N_1 = N - 1$ y \mathcal{P}_N representa el conjunto de permutaciones de $\{1, \dots, N\}$.

La complejidad en Ω

La siguiente proposición permitirá investigar los valores que toma la complejidad C sobre Ω , considerando solamente los valores que toma sobre un subconjunto propio de Ω .

Proposición 6.1:

C toma los mismos valores sobre cada uno de los $N!$ símplices de la subdivisión baricéntrica.

Demostración:

La medida C no depende del orden de los p_i , por lo tanto

$$C(p_{\varphi(1)}, \dots, p_{\varphi(N)}) \equiv C(p_{\varphi'(1)}, \dots, p_{\varphi'(N)}), \quad (6.35)$$

para cualquier permutación φ y φ' definidas sobre $\{1, \dots, N\}$. ■

Se realizará entonces el estudio de la complejidad C considerando solamente los valores que toma sobre uno de los $N!$ símplices de la subdivisión baricéntrica. En particular se considerará el simplex correspondiente a la permutación *identidad*, que será llamado $\Delta_I^{N_1} = [b_I^1, \dots, b_I^N]$.

En la Fig. 6.1, se muestra Ω para $N = 4$. En este caso Ω es un 3-simplex incluido en un hiperplano de dimensión 3 determinado por la restricción de normalización ($\sum_{i=1}^N p_i = 1$). Está generado por los vectores canónicos de \mathbb{R}^4 , como se ve en la Fig. 6.1.a. Se muestra también el baricentro de Ω , $\mu_3 = (1/4, 1/4, 1/4, 1/4)$, que en el contexto que tratamos, está asociado a la probabilidad del equilibrio P_e . En Fig. 6.1.b se ha representado Δ_I^3 , el sub-simplex correspondiente a la permutación identidad I , mostrando también sus $N = 4$ vértices b_I^j , y sus lados que se llamarán $L_{j,k}$. Puede observarse en la Fig. 6.1.b que algunos lados unen vértices consecutivos mientras que otros no. Esto, aunque obvio, es muy importante ya que determinará diferencias sustanciales a la hora de caracterizar los extremos, como será analizado en la sección siguiente.

6.3.3 Caracterización geométrica de los extremos

Para cada $P \in \Omega$ se representará el par $(H(P), C(P))$ en el plano $(H \times C)$ y se llamará C_{max} y C_{min} a las curvas determinadas por los valores extremos de C en función de H .

El objetivo de esta sección es encontrar un procedimiento que permita calcular estas cotas y sea válido para cualquier miembro de la familia. Para simplificar, en adelante se omitirá el subíndice I en la notación de los vértices del simplex $\Delta_I^{N_1} = [b_I^1, \dots, b_I^{N_1}]$.

Proposición 6.2:

Los vectores P de la forma (6.21) constituyen los *lados* de $\Delta_I^{N_1}$.

Demostración:

Sea $L_{j,k}$, con $1 \leq j < k \leq N$, el *lado* que une b^j con b^k (ver por ejemplo la Fig. 6.1.b). $L_{j,k}$ puede definirse paramétricamente mediante: $\mathcal{L}_{j,k}^* : [0, 1] \rightarrow L_{j,k} \subset \Omega$, donde:

$$\mathcal{L}_{j,k}^*(\lambda) = \lambda b^j + (1 - \lambda) b^k . \quad (6.36)$$

Si se expresan los vértices en la base canónica, se tiene

$$\begin{aligned} \mathcal{L}_{j,k}^*(\lambda) &= \left(\frac{\lambda}{N-j+1} \right) \sum_{i=j}^{k-1} e_i \\ &+ \left(\frac{\lambda}{N-j+1} + \frac{1-\lambda}{N-k+1} \right) \sum_{i=k}^N e_i . \end{aligned} \quad (6.37)$$

Llamando $\rho = \lambda/(N-j+1)$, la ecuación anterior define ahora una función $\mathcal{L}_{j,k} : [0, 1/(N-j+1)] \rightarrow L_{j,k} \subset \Omega$, dada por:

$$\mathcal{L}_{j,k}(\rho) = \mathcal{L}_{j,k}^*(\rho(N-j+1)) = \rho \sum_{i=j}^{k-1} e_i + \left(\frac{1-\rho(k-j)}{N-k+1} \right) \sum_{i=k}^N e_i , \quad (6.38)$$

donde $j < k$. Luego si un vector P pertenece al *lado* $L_{j,k}$, será $P = \mathcal{L}_{j,k}(\rho)$ para algún valor $0 \leq \rho \leq 1/(N-j+1)$. De lo que se desprende que P tendrá la forma (6.21), con $m = j-1$ y $n = k-j$. Inversamente, si P es de la forma (6.21), fácilmente se ve que pertenece a un $L_{j,k}$. ■

Corolario 6.1

Los extremos de C para H constante, se alcanzan en los *lados* de $\Delta_I^{N_1}$.

Puede notarse que el *lado* $L_{1,N}$ conecta los vértices no consecutivos (excluyendo el caso $N=2$) b^1 y b^N . El vértice b^1 es el baricentro del simplex entero, y se asocia a la “aleatoriedad total”, $H(b^1) = 1$. Por su parte, b^N , es el baricentro de un punto, y está asociado a la “certeza”, $H(b^N) = 0$. Es decir que $L_{1,N}$ vincula dos vértices con valores extremos de entropía. Esto sugiere la existencia de una relación de orden en los vértices. Para analizar la existencia de este orden, se considerará la entropía H restringida al lado $L_{j,k}$. Teniendo en cuenta (6.38) y suponiendo $j < k$, dicha restricción puede expresarse en función de ρ , como $H(\mathcal{L}_{j,k}(\rho))$, para $0 \leq \rho \leq 1/(N-j+1)$. Se satisface la siguiente proposición:

Proposición 6.3:

Para $1 \leq j < k \leq N$, sea $\mathcal{L}_{j,k}(\rho)$ la definición paramétrica (6.38) del lado $L_{j,k}$. Se cumple que $H(\mathcal{L}_{j,k}(\rho))$ es una función *estrictamente creciente* de ρ , para $0 \leq \rho \leq 1/(N-j+1)$.

Demostración:

Considerando la definición de la entropía H , se llega rápidamente a que la derivada de $H(\mathcal{L}_{j,k}(\rho))$ respecto de ρ es positiva si $0 \leq \rho \leq 1/(N-j+1)$. ■

Corolario 6.2

Sea la entropía H , considerada como una función $H : \Omega \rightarrow [0, 1]$. La función entropía H restringida a un lado $L_{j,k}$, con $j < k$, es una función inyectiva.

Corolario 6.3

Los valores de la función entropía H en los vértices de $\Delta_I^{N_1}$, forman una *partición* del intervalo $[0, 1]$.

$$0 = H(b^N) < H(b^{(N-1)}) < \dots < H(b^2) < H(b^1) = 1 \quad (6.39)$$

Para un conjunto arbitrario $\mathcal{X} \subset \Omega$, denotemos $H[\mathcal{X}]$ a la imagen por H de \mathcal{X} . Con esta notación, se ve que $H[L_{1,N}] = [0, 1]$. Mientras que si $j < k$, $H[L_{j,k}] = [H(b^k), H(b^j)] \subset [0, 1]$. Si además se tiene en cuenta la relación de orden establecida en los corolarios anteriores, se deduce que:

Corolario 6.4

$\{H[L_{i,i+1}], i = N_1, \dots, 1\}$ es una sucesión de conjuntos disjuntos cuya unión es el intervalo $[0, 1]$.

Para cada valor $h \in [0, 1]$, se llamará \mathcal{E}_h al conjunto $\mathcal{E}_h = \{P \in \Delta_I^{N_1} : H(P) = h\}$.

Proposición 6.4:

- a) Dado el simplex $\Delta_I^{N_1}$ y $h \in [0, 1]$, existe un único punto $P_h^{(1)}$ de $\mathcal{B}^{(1)} \equiv L_{1,N}$ y un único punto $P_h^{(N_1)}$ de $\mathcal{B}^{(N_1)} \equiv \bigcup_{i=1}^{N_1} L_{i,i+1}$ en el conjunto \mathcal{E}_h .
- b) Dada una cara bidimensional $[b^i, b^j, b^k]$ del simplex $\Delta_I^{N_1}$, con $i < j < k$, y $h \in [H(b^k), H(b^i)]$, existe un único punto $P_h^{(ik)}$ de $L_{i,k}$ y un único punto $P_h^{(ijk)}$ de $L_{i,j} \cup L_{j,k}$ en el conjunto \mathcal{E}_h .

Demostración:

Tanto a) como b) se comprueban fácilmente aplicando el Corolario 6.4 y la propiedad de inyectividad dada por el Corolario 6.2. ■

Puede notarse que, $\mathcal{B}^{(1)}$ representa al *lado* que une los vértices extremos, mientras que $\mathcal{B}^{(N_1)}$ es la *unión de los lados* que conectan vértices consecutivos.

En la Figs. 6.2.a y 6.2.b se muestran para $N = 6$ los valores que adopta la complejidad $C_E^{(S)}(P)$ y $C_W^{(S)}(P)$, respectivamente, en función de $H^{(S)}(P)$, cuando P varía sobre $\mathcal{B}^{(1)}$ o sobre $\mathcal{B}^{(N_1)}$. Observando la ec. (6.38) se ve que si $P \in \mathcal{B}^{(1)}$, entonces satisface la ec. (6.21)

con $n = 1$, o sea está en el interior de Ω ($m = 0$), tiene *una sola* coordenada con valor p y las demás valen $(1 - p)/(N - 1)$. Por otra parte, si $P \in \mathcal{B}^{(N_1)}$, entonces P está en algún lado $L_{i,i+1}$ y por lo tanto tiene $m = i - 1$ coordenadas nulas, $n = 1$ coordenadas con valor p y las restantes $N - m - 1$ con valor $(1 - p)/(N - 1)$. En particular si P está en $L_{1,2}$, corresponde al interior de Ω .

En las Figs. 6.3.a y 6.3.b se muestran valores de $C_E^{(S)}(P)$ y $C_W^{(S)}(P)$, respectivamente, en función de $H^{(S)}(P)$, obtenidos por simulación numérica del espacio de probabilidades Ω . Si se comparan con las Figs. 6.2.a y 6.2.b, respectivamente se ve que las cotas extremas de la complejidad se corresponden con valores sobre $\mathcal{B}^{(1)}$ o sobre $\mathcal{B}^{(N_1)}$. En particular para $C_E^{(S)}$ la máxima complejidad C_{max} está sobre $\mathcal{B}^{(1)}$ lo cual confirma (satisface eq. (6.21) para $n = 1$) lo probado numéricamente por Anteneodo y Plastino. Para $C_W^{(S)}(P)$ en la Fig. 6.2.b se ve que C_{max} está sobre $\mathcal{B}^{(N_1)}$ y C_{min} sobre $\mathcal{B}^{(1)}$.

Usando la notación de la Proposición 6.4, se supondrá que C satisface la siguiente condición de regularidad:

C satisface *una* de las dos condiciones siguientes (ver Fig. 6.4):

- I) Dada una cara arbitraria $[b^i, b^j, b^k]$, con $i < j < k$, se cumple que $C(P_h^{(ik)}) \geq C(P_h^{(ijk)})$, para todo $h \in [H(b^k), H(b^i)]$.
- II) Dada una cara arbitraria $[b^i, b^j, b^k]$, con $i < j < k$, se cumple que $C(P_h^{(ik)}) \leq C(P_h^{(ijk)})$, para todo $h \in [H(b^k), H(b^i)]$.

Para cada $h \in [0, 1]$ se denotará:

$$P_h^{\max} = \max\{C(P), P \in \mathcal{E}_h\} \quad \text{y} \quad P_h^{\min} = \min\{C(P), P \in \mathcal{E}_h\}. \quad (6.40)$$

Proposición 6.5:

Si C satisface la condición de regularidad I se cumple:

$$P_h^{\max} = P_h^{(1)} \quad \text{y} \quad P_h^{\min} = P_h^{(N_1)} \quad (\text{ver Proposición 6.4}). \quad (6.41)$$

Si por el contrario, C satisface II, entonces:

$$P_h^{\min} = P_h^{(1)} \quad \text{y} \quad P_h^{\max} = P_h^{(N_1)}. \quad (6.42)$$

Demostración:

Se supondrá que C satisface la condición I. De acuerdo al Corolario 6.3 los valores de H en los vértices forman una partición de $[0, 1]$. Por lo tanto, existe j tal que $H(b^{j+1}) \leq h \leq H(b^j)$. Sean i, k subíndices tales que $i < j < k$. Se considerarán $P_h^{(ik)}$, $P_h^{(ijk)}$, $P_h^{(1)}$ y $P_h^{(N_1)}$ como en la Proposición 6.4.

Aplicando la condición de regularidad I primero a la cara bidimensional $[b^1, b^j, b^k]$ y luego a $[b^j, b^k, b^N]$, se comprueba que: $C(P_h^{(1)}) \geq C(P_h^{(ik)})$. Por otro lado, como $H(b^{j+1}) \leq h \leq H(b^j)$ y $P_h^{(N_1)} \in \mathcal{B}^{(N_1)}$, entonces $P_h^{(N_1)}$ está en $L_{j,j+1}$. Aplicando la condición I, ahora primero a la cara bidimensional $[b^i, b^j, b^{j+1}]$ y luego a $[b^j, b^{j+1}, b^k]$, se comprueba que: $C(P_h^{(N_1)}) \leq C(P_h^{(ik)})$.

Es decir que para cualquier i, k tal que $i < j < j + 1 < k$ se cumple que $C(P_h^{(N_1)}) \leq C(P_h^{(ik)}) \leq C(P_h^{(1)})$. Pero como se sabe (Corolario 6.1) que los extremos de la medida de complejidad C para H constante se alcanzan en los lados del simplex, queda demostrado que:

$$C(P_h^{(N_1)}) \leq C(P) \leq C(P_h^{(1)}) \quad \forall P \in \mathcal{E}_h \quad (6.43)$$

En otras palabras, si C satisface la condición I , $P_h^{\max} = P_h^{(1)}$ y $P_h^{\min} = P_h^{(N_1)}$. De manera similar se demuestra el caso en el cual C satisface la condición II . ■

Si se llama:

- $\mathcal{G}^{(1)}$ a la aplicación:

$$\mathcal{G}^{(1)} : [0, 1] \rightarrow [0, 1], \text{ tal que } \mathcal{G}^{(1)}(h) = C(P_h^{(1)}),$$

- $\mathcal{G}^{(N_1)}$ a la aplicación:

$$\mathcal{G}^{(N_1)} : [0, 1] \rightarrow [0, 1], \text{ tal que } \mathcal{G}^{(N_1)}(h) = C(P_h^{(N_1)}),$$

donde $P_h^{(1)}$ y $P_h^{(N_1)}$ están definidos en la Proposición 6.4, se cumplirá el siguiente Teorema.

Teorema 6.2:

Sea C una medida de complejidad estadística generalizada que satisface la condición de regularidad enunciada previamente. Si C_{\max} y C_{\min} son las cotas máxima y mínima de la complejidad en el plano $(H \times C)$, se cumple que:

- Si C satisface la condición I entonces $C_{\max} \equiv \mathcal{G}^{(1)}$ y $C_{\min} \equiv \mathcal{G}^{(N_1)}$.
- Si C satisface la condición II entonces $C_{\max} \equiv \mathcal{G}^{(N_1)}$ y $C_{\min} \equiv \mathcal{G}^{(1)}$.

Demostración:

Se obtiene aplicando la Proposición 6.5 para cada $h \in [0, 1]$. ■

Ahora bien, para todo $h \in [0, 1]$ se tiene que $P_h^{(1)} \in L_{1,N}$, entonces la curva $\mathcal{G}^{(1)}$ se obtiene graficando $(H(P), C(P))$ para $P \in L_{1,N}$. Por otro lado, $P_h^{(N_1)} \in \bigcup_{i=1}^{N_1} L_{i,i+1}$, entonces la curva $\mathcal{G}^{(N_1)}$ se obtiene graficando $(H(P), C(P))$ para P recorriendo $L_{i,i+1}$, con $i = 1, \dots, N_1$. Es decir que $\mathcal{G}^{(N_1)}$ está definida a trozos. Cada uno de estos trozos, corresponde con un lado $L_{i,i+1}$, que une los vértices consecutivos $b^{(i)}$ y $b^{(i+1)}$. Se deduce fácilmente de la ec. (6.38), que las configuraciones correspondientes a los lados que unen vértices consecutivos, adoptan la forma (6.21) con $n = 1$. Por lo tanto $\mathcal{G}^{(N_1)}$ se obtiene en configuraciones de la forma (6.21) con $n = 1$.

Evaluación de C_{\max} y C_{\min}

Se supondrá que C satisface la condición de regularidad I . En este caso $C_{\max} \equiv \mathcal{G}^{(1)}$ y $C_{\min} \equiv \mathcal{G}^{(N_1)}$. Se comenzará con el cálculo de C_{\max} .

Como $\mathcal{G}^{(1)}(\mathbf{h}) = C(P_{\mathbf{h}}^1)$ y $P_{\mathbf{h}}^1 \in L_{1,N}$, entonces para evaluar C_{max} , P debe variar sobre $L_{1,N}$. La curva se obtiene graficando $\{(H(P), C(P))\}$ con $P \in L_{1,N}$. Las configuraciones P quedan definidas por la función $\mathcal{L}_{1,N} : [0, 1/N] \rightarrow L_{1,N}$:

$$\mathcal{L}_{1,N}(\rho) = \rho \sum_{i=1}^{N-1} e_i + (1 - \rho(N-1)) e_N . \quad (6.44)$$

P varía desde la certeza al equilibrio, a través de configuraciones que pertenecen al interior de Ω ($m = 0$), con una única probabilidad $p_i = p$ ($n = 1$) y las demás son todas iguales a $(1-p)/(N-1)$.

Por su parte, $C_{min} \equiv \mathcal{G}^{(N_1)}$. Como $\mathcal{G}^{(N_1)}(H) = C(P_{\mathbf{h}}^{N_1})$ y $P_{\mathbf{h}}^{N_1} \in \bigcup_{i=1}^{N_1} L_{i,i+1}$, entonces para evaluar C_{min} , P debe ser variado sobre $\bigcup_{i=1}^{N_1} L_{i,i+1}$. La curva se obtiene graficando $\{(H(P), C(P))\}$ con $P \in \bigcup_{i=1}^{N_1} L_{i,i+1}$. Las configuraciones quedan definidas por la función $\mathcal{L}_{i,i+1} : [0, 1/(N-i+1)] \rightarrow L_{i,i+1}$:

$$\mathcal{L}_{i,i+1}(\rho) = \rho e_i + \left(\frac{1-\rho}{N-i} \right) \sum_{l=i+1}^N e_l , \quad (6.45)$$

con $i = 1, \dots, N-1$. P varía desde la certeza al equilibrio, a través de configuraciones que pertenecen a lados consecutivos. O sea que la mínima complejidad se obtiene en configuraciones P que estarán en $L_{1,2}$ del interior de Ω ($m = 0$), o en el contorno. Estas configuraciones asociadas a C_{min} son tales que si pertenecen a lado $L_{i,i+1}$, tienen $m = i-1$ coordenadas nulas y sus probabilidades no nulas tienen un único valor $p_i = p$ ($n = 1$) y los demás son todos iguales a $(1-p)/(N-i)$.

6.4 Aplicaciones V

Aplicaciones al mapa logístico permitirán mostrar la utilidad de la representación “complejidad *versus* entropía normalizada” (C vs. H), y por ende, de las cotas C_{max} y C_{min} . El mapa logístico ya ha sido presentado en la sección 3.6.1, donde se expuso una minuciosa descripción del mismo. Se trata de un sistema disipativo, ecológicamente motivado, descripto por la ecuación en diferencia de primer orden:

$$x_{n+1} = r x_n (1 - x_n) , \quad 0 \leq x_n \leq 1 , \quad 1 < r \leq 4 , \quad (6.46)$$

cuyo comportamiento dinámico está controlado por el parámetro r . Se ha discutido en el capítulo anterior la aplicación de las medidas de complejidad C_W (ver sec. (5.3)) y C_{JS} (ver sec. (5.4.2)), en el análisis del mapa logístico considerando una representación binaria del sistema, cuando r varia entre 3.5 y 4.0. Se mostrarán aquí los resultados obtenidos con ambas medidas de complejidad pero representados en un diagrama “complejidad *versus* entropía normalizada” (C vs. H).

Si se observa el gráfico correspondiente a la entropía normalizada H en función del parámetro de control r dado en la Fig. 6.5.c., puede notarse que para configuraciones correspondientes a ventanas periódicas, la entropía H satisface $H < H^* \sim 0.3$, mientras que para aquellas asociadas a comportamientos caóticos se cumple $H > H^*$. En la

Fig. 6.6.a. se presenta el exponente de Lyapunov en el diagrama λ vs. H , observándose en coincidencia que $H < H^* \sim 0.3$ se corresponde con $\lambda < 0$ (comportamiento periódico), mientras que $H > H^*$ se asocia con $\lambda > 0$ (comportamiento caótico).

Las Figs. 6.6.b. y 6.6c. presentan el diagrama complejidad-entropía para las medidas de complejidad C_E y C_{JS} , respectivamente. Los valores alcanzados por estos cuantificadores cuando r varía entre 3.5 y 4.0, son graficadas paramétricamente como una función de H . Este diagrama brinda acceso directo a las propiedades de información del sistema independientemente del valor del parámetro de control r . Para el caso de la medida basada en la distancia Euclídea, la Fig. 6.6.b. muestra por un lado, que C_E adopta para la mayor parte de las configuraciones valores muy próximos a la cota mínima de complejidad, C_{min} . Además, C_E alcanza valores más grandes en la región de mayor orden ($H < H^* \sim 0.3$), asociada a comportamientos periódicos, mientras que en la zona intermedia, correspondiente a comportamientos complejos, toma valores pequeños. Finalmente tiende a 0 en la aleatoriedad total. Dado que C_E es introducida como una medida de estructura, este comportamiento no parece ser razonable. Efectivamente, si se observa la Fig. 6.6.c se notará enseguida que esta dificultad desaparece al usar la medida de complejidad basada en la divergencia de Jensen-Shannon. C_{JS} presenta un comportamiento opuesto, alcanzando valores máximos en las proximidades de $H \sim 0.7$ asociada a un comportamiento caótico del mapa logístico (ver Fig. 6.5). Además, al variar r entre 3.4 y 4.0 el mapa logístico presenta comportamientos de muy variada complejidad, que se traducen en el amplio rango de valores que alcanza C_{JS} . Por otra parte analizando la región de mayor orden $H < H^* \sim 0.3$, asociada a ventanas con diferentes periodicidades, se ve que C_{JS} tiene la capacidad de distinguirlas asignándole complejidades crecientes.

6.5 Conclusiones

Anteneodo y Plastino [5] mostraron usando argumentos analíticos y numéricos que la distribución de probabilidad que maximiza la medida de complejidad de Lopez-Ruiz, Mancini y Calbet, C_{LMC} , es de la forma dada en la ec. (6.21). En este capítulo, mediante argumentos analíticos y geométricos se ha extendido este resultado. Para las medidas de complejidad estadística generalizada que satisfagan una condición de regularidad (ver sección 6.3.3), se mostró que ellas alcanzan sus valores extremos en configuraciones de la forma (6.21), con $n = 1$.

El estudio detallado de los valores que alcanzan las medidas de complejidad, llevado a cabo en el plano C vs. H , permitió mostrar la existencia de cotas de la medida C en función de H , llamadas C_{max} y C_{min} . Las características geométricas del espacio de probabilidad y las propiedades particulares de las medidas de complejidad estadística, facilitaron el desarrollo de un procedimiento sistemático para la evaluación de las cotas, válido para cualquier medida de la familia.

Estos resultados revisten gran interés, ya que el uso de diagramas C vs. H ofrece la posibilidad de estudiar los sistemas desde un punto de vista explícitamente independiente de los parámetros de control. Por ejemplo, si al investigar el comportamiento de un sistema, se grafica paramétricamente la medida de complejidad del mismo en función de H , junto con las cotas C_{max} y C_{min} de la medida, se obtiene una visión que permite analizar

el sistema directamente en términos de la intrínseca información del proceso. Este enfoque también brinda la posibilidad de vincular la complejidad del sistema en estudio, con los valores extremos de la medida. Finalmente, mediante las cotas pueden también caracterizarse algunos aspectos de cada medida de complejidad en sí misma, como por ejemplo la existencia de un valor máximo absoluto y la forma de la configuración P para la cual dicha medida de complejidad es máxima.

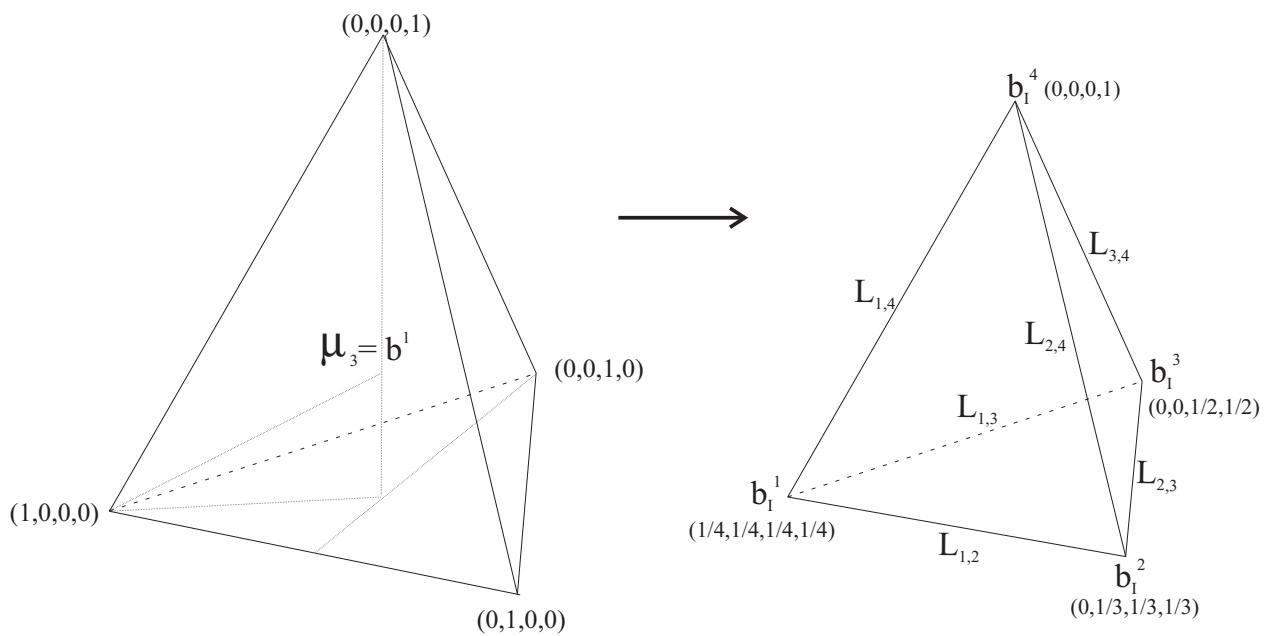


Figura 6.1: ESPACIO DE PROBABILIDADES Ω PARA $N = 4$: a) $\Omega \equiv \Delta^2$ EN EL HIPERPLANO DE DIMENSIÓN 3. LAS LÍNEAS PUNTEADAS TRAZAN LA SUBDIVISIÓN BARICÉTRICA Y μ_3 BARICENTRO DE Ω . b) SUB-SIMPLEX Δ_4^3 .

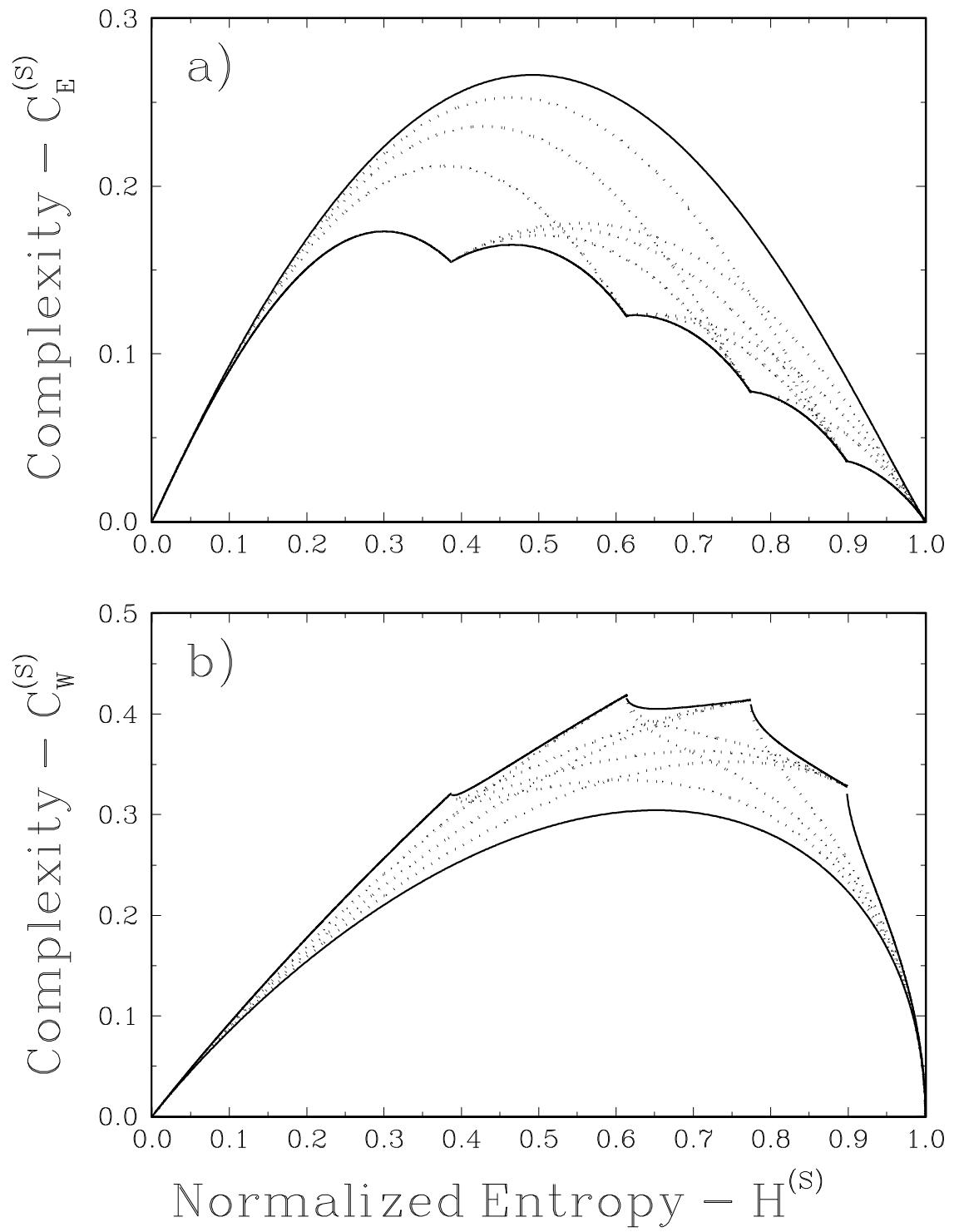


Figura 6.2: GRÁFICO DE LA MEDIDA DE COMPLEJIDAD EN FUNCIÓN DE $H^{(S)}(P)$, CUANDO P VARÍA SOBRE $\mathcal{B}^{(1)}$ O SOBRE $\mathcal{B}^{(N_1)}$, PARA $N = 6$. a) $C_E^{(S)}(P)$ Y b) $C_W^{(S)}(P)$.

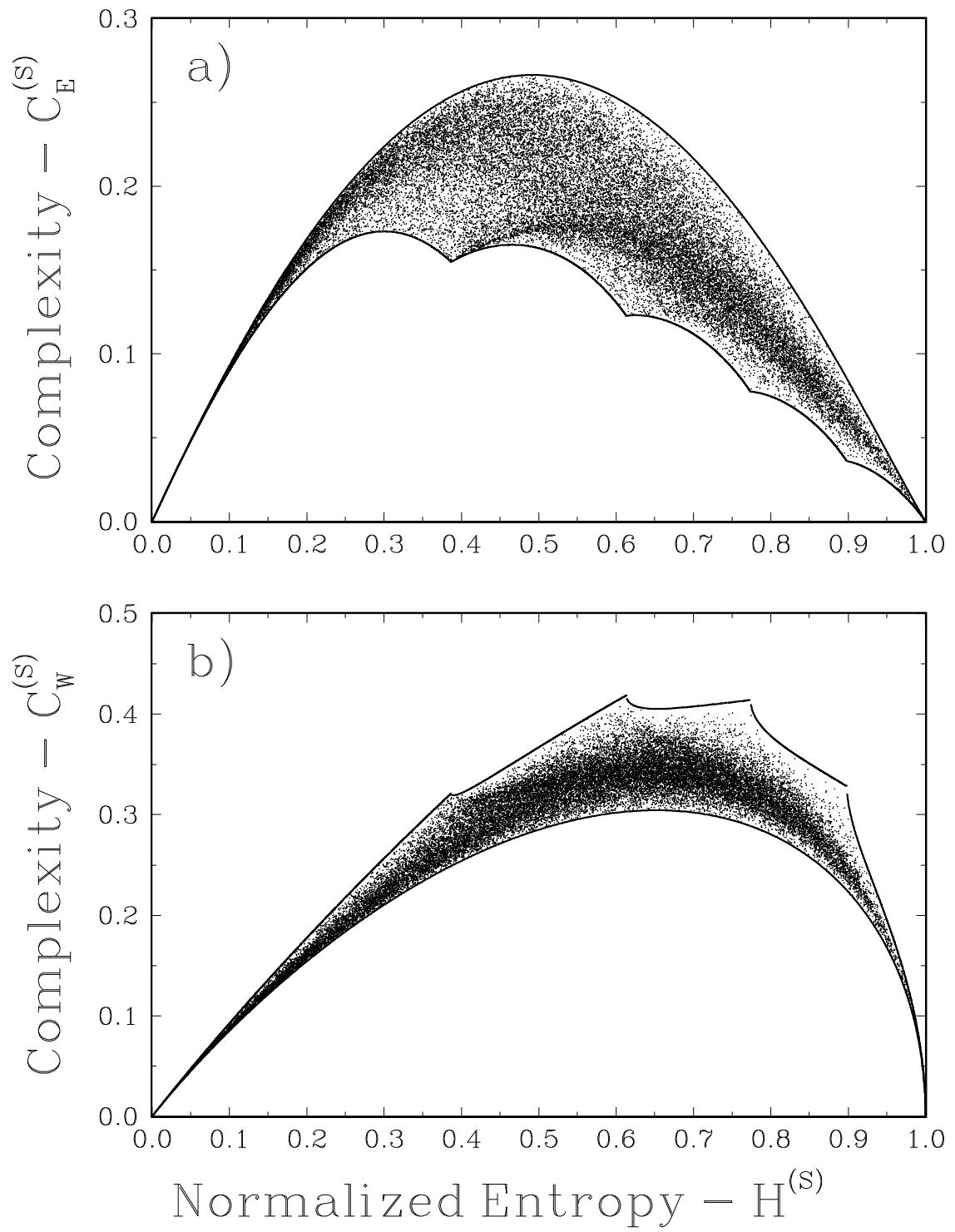


Figura 6.3: SIMULACIÓN NUMÉRICA DE LA MEDIDA DE COMPLEJIDAD EN FUNCIÓN DE $H^{(S)}(P)$, CUANDO P VARÍA EN Ω , PARA $N = 6$. a) $C_E^{(S)}(P)$ Y b) $C_W^{(S)}(P)$.

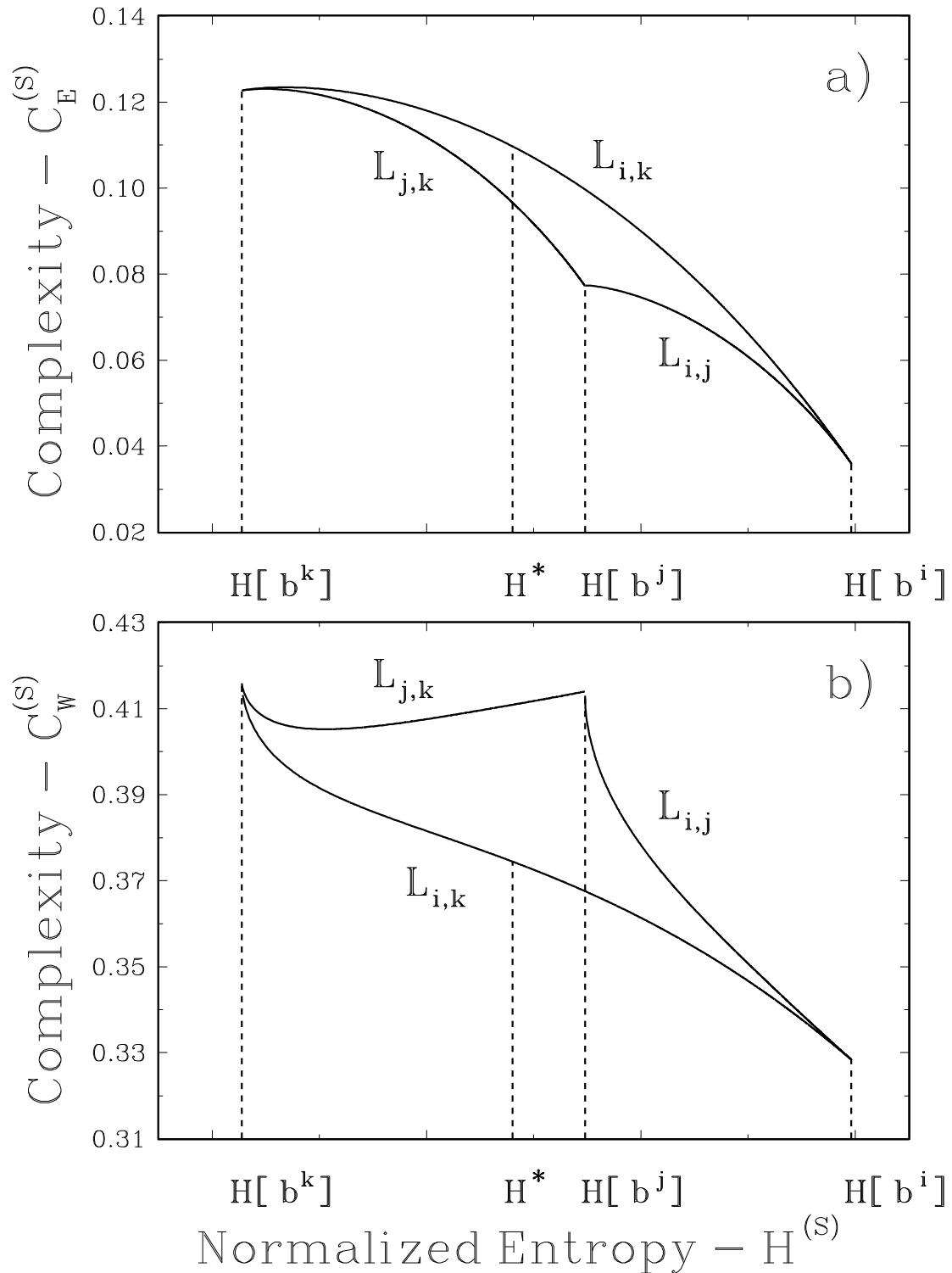


Figura 6.4: CARA BIDIMENSIONAL: GRÁFICO DE LA MEDIDA DE COMPLEJIDAD EN FUNCIÓN DE $H^{(S)}(P)$, CUANDO P VARÍA SOBRE L_{ik} O SOBRE $L_{ij} \cup L_{jk}$. a) $C_E^{(S)}(P)$ Y b) $C_W^{(S)}(P)$.

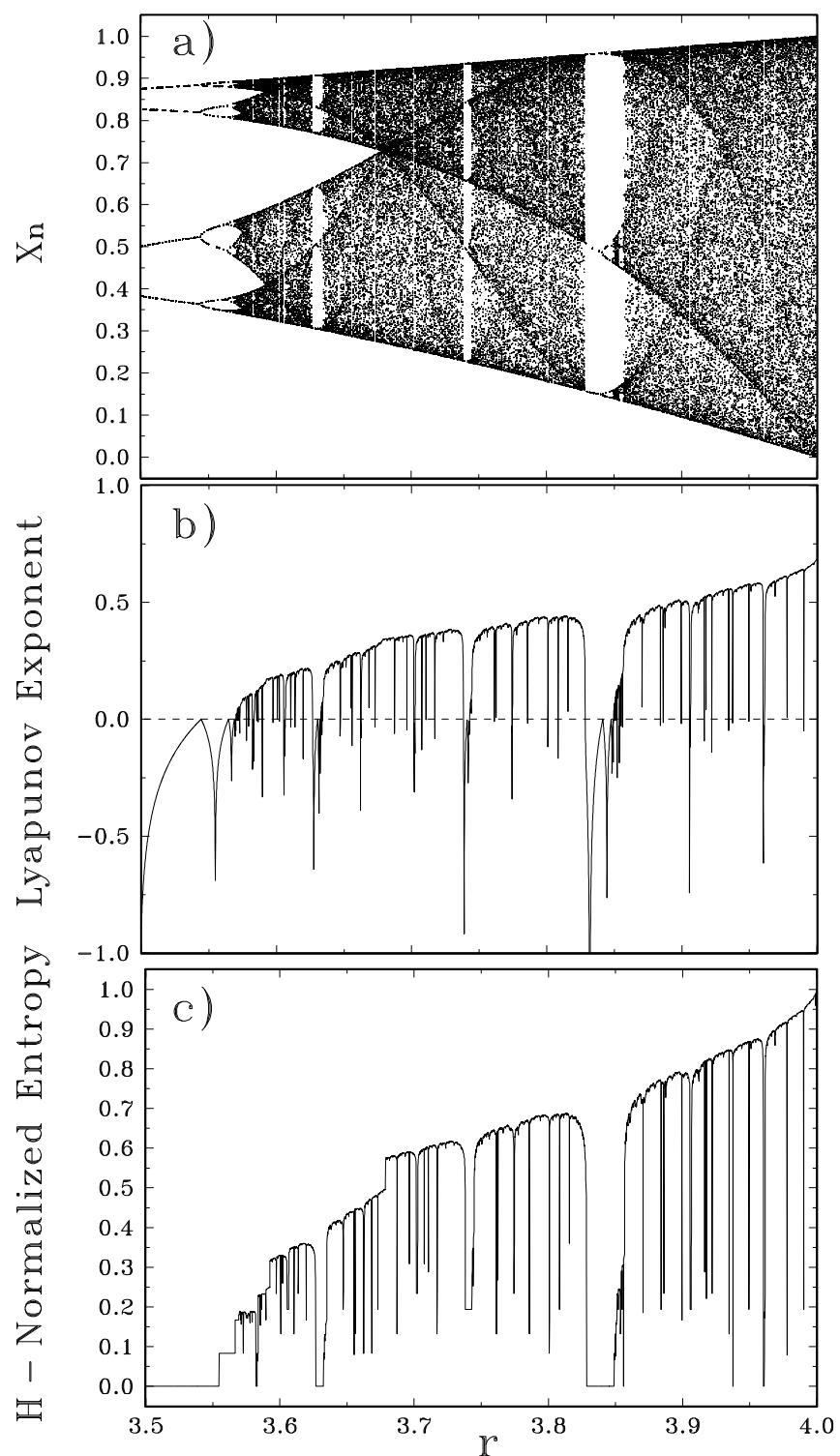


Figura 6.5: a) DIAGRAMA DE ÓRBITAS b) LYAPUNOV EXPONENT Y c) ENTROPÍA NORMALIZADA H PARA EL MAPA LOGÍSTICO EN FUNCIÓN DEL PARÁMETRO r .

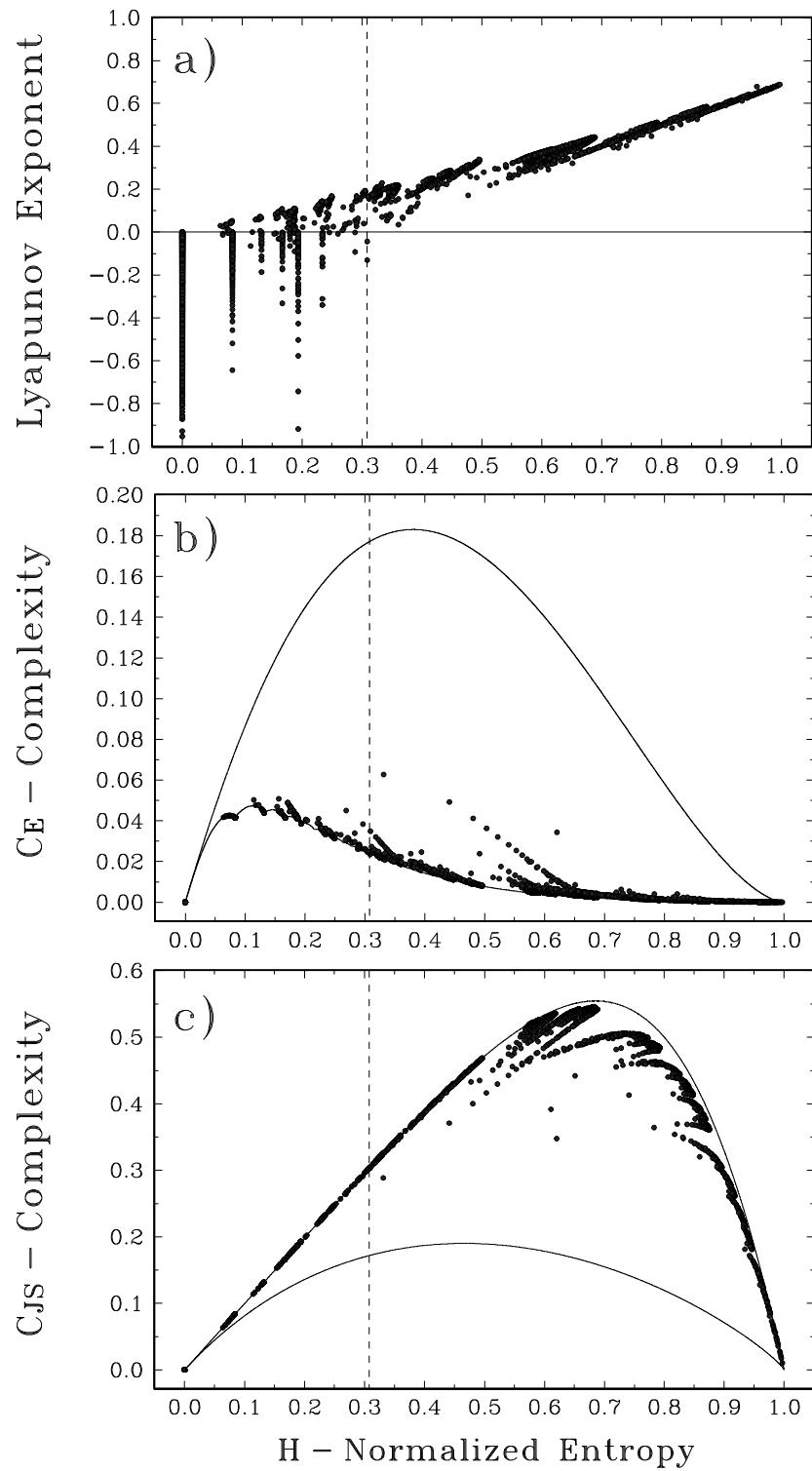


Figura 6.6: CANTIDADES DE INTERÉS SON GRAFICADAS EN FUNCIÓN DE LA ENTROPIA NORMALIZADA H : a) EXPONENTE DE LYAPUNOV λ . NOTAR QUE LAS VENTANAS PERIÓDICAS SE DISTINGUEN CLARAMENTE PARA $H^* \leq 0.3$. b) COMPLEJIDAD C_E Y c) COMPLEJIDAD C_{JS} . EN b) Y c) SE GRAFICARON TAMBIÉN LAS COTAS C_{max} Y C_{min} (LÍNEA CONTINUA). LA LÍNEA VERTICAL PUNTEADA SEÑALA EL VALOR H^* PARA EL FINAL DE LA ZONA DE PERIÓDICIDAD.

Capítulo 7

Aplicaciones a señales EEG

7.1 Introducción

Las medidas de entropía y complejidad introducidas en capítulos anteriores, serán aplicadas al análisis de señales electroencefalográficas (EEG), prestando especial atención a la capacidad que demuestren estas técnicas para el estudio de *epilepsia*. Luego de presentar los mínimos conceptos de neurofisiología necesarios para la exposición, se describirán los resultados obtenidos al aplicar los métodos TWS y TWC definidos en las secciones (3.5.3) y (4.4.3) respectivamente.

A pesar de ser muy amplia la bibliografía y muy diversos los trabajos dedicados al desarrollo de métodos para el análisis de señales de EEG, existen puntos fundamentales que aún hoy son controvertidos. En efecto, debido a la elevada complejidad de estas señales, las técnicas tradicionales fracasan en algunos tópicos, que resultan por lo tanto candidatos ideales para ser estudiados con nuevos métodos cuantitativos.

Si se analizan las técnicas introducidas previamente en este trabajo, podrá notarse que la medida de complejidad wavelet total TWC (4.4.3) admite ser implementada, con cualquier integrante de la familia de medidas de complejidad estadística generalizada desarrollada en el capítulo 6. En las aplicaciones que ilustrarán este capítulo, su implementación estará basada en dos medidas simples de complejidad (C_{LMC} y C_{SDL}) definidas en la sección 4.3, así como también en sus generalizaciones basadas en la distancia de Wootters C_W y la divergencia de Jensen-Shannon C_{JS} . De este modo, se generarán medidas de complejidad wavelet total que serán llamadas: TWC_{LMC} , TWC_{SDL} , TWC_W , TWC_{JS} , respectivamente. Se analizará el comportamiento de cada una de ellas en relación a las series electroencefalográficas.

7.2 Electroencefalograma (EEG)

Los primeros registros de la actividad eléctrica cerebral fueron reportados por Caton en 1875 en cerebros expuestos de conejos y monos, pero recién en 1929 Hans Berger [14] realizó la primer medición de actividad eléctrica cerebral en humanos. Distintos patrones morfológicos característicos del EEG han sido correlacionados con funciones, disfunciones y enfermedades del sistema nervioso central. De este modo los EEGs emergen como una de las más importantes técnicas de diagnóstico en neurofisiología.

Un registro EEG puede ser definido, sintéticamente, como la actividad eléctrica media del cerebro en distintos sitios de la cabeza. Más específicamente, es la suma del flujo de corriente *extracelular* de un extenso grupo de neuronas. Para detalles sobre la generación de EEG a partir de la acción de potenciales de neuronas, se sugieren los trabajos de Steriade y colaboradores [124], Lopes da Silva [76], Speckermann y Elger [121], Pedley y Traub [96] y Başar [10].

Registros EEG son obtenidos por medio de electrodos de alta conductividad ubicados en diferentes sitios de la cabeza [92]. Estas mediciones son principalmente realizadas sobre la superficie de la cabeza (*EEG scalp*), de manera no invasiva. Aunque para el estudio y un mejor diagnóstico de ciertas patologías se utilizan electrodos especiales ubicados en el cerebro mediante una intervención quirúrgica (*EEG intracraneano*).

7.2.1 Análisis de señales EEG

El análisis tradicional de la actividad eléctrica cerebral a través de registros electroencefalográficos (EEG), se basa principalmente en una inspección visual de la señal. Si bien es indiscutible la utilidad de esta técnica, su naturaleza subjetiva impide un protocolo sistemático. Para sortear esta dificultad, en los últimos años se han desarrollado técnicas cuantitativas que introducen medidas objetivas, capaces de reflejar no solamente la actividad cerebral en sí misma, sino también información sobre la dinámica cerebral asociada. De este modo, distintos métodos del análisis lineal y no lineal han sido aplicados a las señales EEG [34, 125, 122, 98, 6, 7].

Los registros EEG, al igual que la mayoría de las señales biológicas revelan una dinámica subyacente no lineal [8, 88, 60, 61, 134, 135, 46, 47]. Algunos autores han demostrado que la variabilidad de las señales EEG no representa ruido, sino la presencia de un atractor [8, 11, 88]. Una completa revisión del estado del arte en este área puede consultarse en [32] y en el volumen especial [9].

También desde el ángulo de la neurofisiología se ha entendido que las señales EEG provienen de sistemas altamente no lineales [34, 103, 101]. Numerosos bucles de retroalimentación han sido encontrados en cada uno de los niveles jerárquicos del sistema nervioso central. Las neuronas, de manera individual, también aparecen como elementos altamente no lineales. Una combinación de medidas espectrales y no lineales son entonces perspectivas promisorias para emprender el análisis de los EEGs [34].

Con respecto a las técnicas no lineales, ellas analizan *i*) la evolución temporal de la complejidad dinámica del EEG (asociada con la medida de la dimensión de correlación D_2 , [18, 74]) y *ii*) el grado de caoticidad (en términos del mayor exponente de Lyapunov, Λ_{max} [61, 59]). Mediante un análisis de este tipo, puede detectarse por ejemplo, una transición de un comportamiento muy complejo a uno más simple, en el comienzo o aún antes, de una crisis epiléptica [74, 61, 59, 99]. Sin embargo, a pesar de la obvia relevancia fisiológica de estos hallazgos, un requerimiento básico para la aplicación de herramientas métricas de la dinámica no lineal en datos experimentales es que la señal correspondiente sea *estacionaria*. Este no es el caso de las señales de EEG (los EEG solo pueden ser considerados señales estacionarias para períodos de tiempo muy cortos, típicamente para intervalos de algunos segundos). Además, para la evaluación de D_2 y Λ_{max} se requieren registros con un número elevado de datos significativos (sin sobremuestreo), lo cual corresponderá a registros de tiempos largos. Más aún, la evaluación de los

invariantes métricos mencionados requieren el cálculo de parámetros adicionales que se degradan rápidamente en presencia de ruido aditivo. Dificultando así la implementación e interpretación de los resultados obtenidos mediante estas técnicas.

Por otra parte, cambios dinámicos en la actividad eléctrica cerebral, vinculados al procesamiento de información, se manifiestan en tres campos bien diferenciados: *i*) tiempo, *ii*) frecuencia y *iii*) espacio. Así es que los correspondientes estudios requieren técnicas capaces de describir la variación cualitativa y cuantitativa de algún aspecto de la señal, tanto en tiempo como en frecuencia. Las técnicas basadas en el análisis wavelet introducidas en los capítulos anteriores, al satisfacer estos requerimientos, se convierten en herramientas promisorias. Se mostrará en este capítulo su utilidad en el análisis de los EEGs.

Debe destacarse también que en presencia de ruido, los métodos provenientes de la dinámica no lineal ya no son aplicables. En cambio los nuevos métodos basados en el análisis wavelet presentan ventajas: los segmentos de señal a analizar no requieren un registro elevado de datos y el costo computacional de los algoritmos es reducido, como ya fue señalado en capítulos anteriores. En el caso en que el ruido esté mayormente concentrado solo en algún rango de frecuencia, su contribución en la evaluación de los cuantificadores wavelets puede ser eliminada muy fácilmente.

7.2.2 EEG y epilepsia

La epilepsia es una entidad médica caracterizada por crisis recurrentes, que afecta alrededor del 1 % de la población y cuya etiología ha avanzado muy lentamente. Una importante aplicación de los EEGs es el diagnóstico de la epilepsia. La abrupta aparición de puntas de muy alta amplitud (*spikes*), ritmicidades anormales, enlentamiento (*slowing*) del registro y otros varios paroxismos son en general indicios de epilepsia, que ayudan a identificar, clasificar y localizar las crisis.

A fines del siglo XIX, los neurólogos ingleses J. Hughlings y J. Jackson [63], formularon la hipótesis de que las complejas manifestaciones clínicas observadas durante una crisis eran producto de una excesiva descarga de los agregados neuronales. Esta hipótesis fue confirmada por Matsumoto y Ajmone-Marsan en 1964 [86], quienes describieron las correlaciones celulares de las manifestaciones ictales (crisis epiléptica) e inter-ictales (período entre crisis).

Actualmente se define a la *Epilepsia* como un desorden neurológico caracterizado por crisis o cambios bruscos en la actividad de ciertas neuronas que interfieren con el funcionamiento normal de toda la red neuronal, provocando disturbios en las actividades sensorial y motora como así también posibles pérdidas de conciencia. Este grupo de neuronas que ve alterado su funcionamiento puede o no estar circunscripto a una zona bien localizada del cerebro (foco epileptogénico). Durante el crisis epiléptica la red neuronal presenta oscilaciones típicas que normalmente se propagan involucrando otras zonas del cerebro.

Desde el punto de vista electrofisiológico pueden distinguirse dos tipos de signos bioeléctricos: la actividad ictal y la inter-ictal. La punta (*spike*) inter-ictal es el sello de la actividad epileptiforme y representa un estadio epiléptico intermedio con un alto grado de sincronía [65]. En el electroencefalograma humano las puntas (*spikes*) epilépticas inter-ictales aparecen como breves transitorios (30 - 150 ms) cuyas amplitudes usual-

mente oscilan de 80 a 250 μV . La actividad ictal está caracterizada por un súbito cambio de la frecuencia de la señal EEG, dando paso a un nuevo tipo de ritmo que dominará el trazado [89]. Este tipo de actividad está correlacionada con el comportamiento ictal característico (sintomatología clínica). Se supone que el factor principal o al menos uno de los que detiene la crisis después de algunos segundos o minutos, es el fenómeno de fatiga neuronal relacionado con el agotamiento momentáneo de neurotransmisores en las terminales sinápticas debido a las numerosas descargas durante la crisis epiléptica. En la actualidad los métodos para controlar las crisis epilépticas son del tipo farmacológico y en algunos casos de epilepsias refractarias se recurre al tratamiento quirúrgico.

La clasificación de las crisis epilépticas es una tarea muy difícil. A continuación se presentará una versión simplificada:

1. **Epilepsia Focal:** involucra una zona específica de la corteza, comúnmente la porción límbica del encéfalo como el hipocampo, la amígdala o la corteza temporal. En este caso se habla de foco epileptogénico, de modo que se lo puede identificar y con frecuencia se utiliza la remoción quirúrgica para evitar crisis futuras. En general este tipo de *epilepsia* es consecuencia de alguna lesión orgánica localizada (por ejemplo un tumor que comprime un área del encéfalo) o anomalía funcional.
2. **Epilepsia Generalizada:** se caracteriza por descargas neuronales extremas en todas las áreas del encéfalo; se producen descargas sincrónizadas de alto voltaje en toda la corteza, incluso en el tálamo y el tallo encefálico.
 - *Epilepsia Petit Mal:* se caracteriza por 3 a 30 s de inconciencia a veces con contracciones musculares (pero sin actividad tónico-clónica), seguidas por la recuperación de la conciencia y reanudación de las actividades previas. A veces éste es el preámbulo de una crisis tipo *Grand Mal*. El patrón de ondas es un complejo punta-onda que puede registrarse sobre la mayor parte o toda la corteza cerebral.
 - *Epilepsia Grand Mal:* Consiste en una repentina pérdida de conciencia con descargas en la médula espinal que producen convulsiones tónico-clónicas. Una crisis tónico-clónica (*o Grand Mal*) dura normalmente alrededor de 40 - 90 s y está caracterizado por violentas contracciones musculares. Una fase inicial tónica con espasmos masivos (es decir, tensión muscular extrema sin movimiento) es reemplazada unos segundos más tarde por la fase clónica con movimientos flexores violentos y características contracciones rítmicas del cuerpo entero hasta el final del crisis [92]

En el caso de epilepsia del tipo Grand Mal, la inspección visual de registros de superficie es en general muy dificultosa, porque la señal se encuentra oscurecida por la presencia de artificios musculares. En la mayoría de los casos el análisis queda confinado a la interpretación de las anormalidades eléctricas que o bien preceden o siguen a la actividad tónico-clónica. Por lo tanto, ya que las crisis de este tipo son difíciles de estudiar con las técnicas tradicionales, será muy interesante analizarlos con las técnicas propuestas en los capítulos previos [92].

Análisis de la actividad cerebral durante crisis tónico-clónica (TC), sólo han sido hechas en circunstancias especiales, tales como en pacientes tratados con *curare* (un inhibidor de la respuesta muscular) [41, 42] o eliminando la actividad de alta frecuencia muscular con el uso de filtros tradicionales [48]. Gastaut y Broughton [41] describieron un patrón de frecuencias durante una crisis epiléptica TC de pacientes sometidos a relajación muscular por aplicación de *curare* y respiración artificial. Luego de un corto período, entre 1 s y 3 s, caracterizado por una fase de desincronización, Gastaut y Broughton encontraron un “ritmo de reclutamiento epiléptico” de alrededor de 10 Hz [42] con un rápido aumento de la amplitud dominando la señal de EEG. Hacia la finalización de la crisis existe un progresivo aumento de las bajas frecuencias asociado con la fase clónica.

En la Fig. 7.1 se presenta una señal de EEG scalp correspondiente a una crisis epiléptica tónico-clónica registrado con un electrogordo colocado en la región central derecha (canal C4). Por medio del análisis clínico y la inspección visual, el equipo médico determina las distintas etapas de la crisis. En este registro la fase pre-ictal está caracterizada por una señal con una amplitud máxima de 50 μV . La crisis epiléptica comienza a los 80 s con una descarga de ondas lentas superpuestas con actividad rápida de menor amplitud. Esta descarga dura aproximadamente 9 s con una amplitud media de 100 μV . Luego, la crisis comienza a desarrollarse plenamente, resultando más complicado el análisis de la señal debido a la presencia de gran cantidad de artificios provenientes de la actividad muscular. A pesar de las dificultades técnicas es posible establecer el inicio de la fase clónica a los 125 s aproximadamente y el final de la crisis cercano a los 155 s, donde hay una caída abrupta en la amplitud de la señal.

7.3 Separación de Señales y Análisis Wavelet

En las crisis epilépticas tónico-clónica, como ya se dijo, el análisis cuantitativo resulta especialmente difícil por la presencia de artificios derivados de la actividad muscular, los que suelen alcanzar amplitudes muy altas [92]. El filtrado tradicional del ruido contaminante (filtros pasa-banda basados en la Transformada de Fourier), tiene algunas desventajas [90, 91, 125] debido que al filtrar las frecuencias relacionadas con los artificios provenientes de la actividad electromiográfica, también se ve afectada la morfología de las frecuencias remanentes y por ende el proceso de filtrado altera el valor de los invariantes métricos no lineales.

Para evitar estas limitaciones, se utilizará un método de separación de señales basado en wavelets ortogonales. Cuando el ruido (contaminante de la señal) está presente sólo en bandas de frecuencia específicas, un proceso de filtrado o separación de señal puede ser eficientemente implementado. En el caso de las crisis tónico-clónica, por ejemplo, se quiere eliminar el ruido correspondiente a las bandas de alta-frecuencia ya que la actividad muscular está concentrada en un rango de frecuencias mayores que 20 Hz.

Usando la ecuación de descomposición wavelet (2.44) hasta la escala j_0 , se obtiene una versión suavizada de la señal original recurriendo a:

$$\tilde{s}(t) = \sum_{j=-N_J}^{-j_0} \sum_k d_j(k) \psi_{j,k}(t) = \sum_{j=-N_J}^{-j_0} r_j(t) . \quad (7.1)$$

$\tilde{s}(t)$ tiene menor contenido de altas frecuencias comparada con la señal obtenida si se incluye en la suma el nivel $j_0 + 1$. Pero a su vez tendrá la mitad de datos. La relación de muestreo original puede obtenerse con una interpolación spline cúbica.

Una razón fundamental para usar el método de separación de señales basado en wavelets ortogonales, es que permite analizar la señal “sin ruido” sin alterar la interpretación de la dinámica asociada. Mallat [79] ha mostrado que para eliminar el ruido de una señal unidimensional (suponiendo ruido aditivo), es más conveniente un método basado en wavelets ortogonales que los basados en Fourier, porque la morfología de las bandas de frecuencias no filtradas no resultan afectadas, y por lo tanto la dinámica asociada a ellas no cambia [80, 30, 111].

El análisis espectral de las series de EEG es realizado clásicamente estudiando diferentes bandas de frecuencia. Intensidades absolutas y relativas de estas bandas son analizadas y correlacionadas con diferentes patologías. Se definirán en este trabajo seis bandas de frecuencia para un análisis wavelet apropiado dentro del esquema multiresolución a ser utilizado. Se denotarán estos niveles de bandas de resolución por B_j ($|j| = 1, \dots, 6$). Sus límites temporales, resolución temporal, así como su correspondencia con las tradicionales bandas de frecuencia EEG, están dadas en la Tabla 7.1.

Se usará en este trabajo un esquema multiresolución basado en funciones wavelet spline cubicas como wavelet “madres”. En la Fig. 7.2 se muestra la señal EEG original sin la contribución de las bandas de frecuencia B_1 y B_2 , asociadas a los niveles de resolución wavelet $j = -1$ y -2 , respectivamente. Estos contienen artificios de alta frecuencia relacionados con la actividad muscular, que oscurecen el EEG. Aunque la actividad cerebral de alta frecuencia es también eliminada con este procedimiento, su contribución durante la etapa ictal es de muy poca magnitud en relación a la contribución de las medias y bajas frecuencias [106, 105]. Una vez que los artificios de las altas frecuencias fueron eliminados, se puede analizar la evolución temporal de los cuantificadores basados en wavelet, sobre la señal “remanente” o “ limpia”. Se usarán en dicha aplicación ventanas temporales de 2.5 s, o sea $L = 256$ datos.

Resumiendo, el filtrado wavelet con spline cúbicas ortogonales resulta un buen método para separar en forma precisa el ruido (actividad muscular) y luego recuperar la señal “limpiada” sumando las componentes residuales correspondientes a las bandas de frecuencia de interés previa interpolación de las mismas en orden de obtener la misma frecuencia de registro. De esta manera, el método permite la eliminación de bandas de frecuencias no deseadas que ocultan otra información o efectos desconocidos. Dada la ortogonalidad de las funciones wavelet empleadas, se puede asegurar que, sólo las bandas de frecuencias previamente seleccionadas son las efectivamente removidas.

7.3.1 Resultados

- **Energía Multi-escala**

La Fig. 7.3 muestra la evolución temporal para las energías multi-escala wavelet (ver sección 3.5.1) sin contribuciones electromiográficas representadas por las bandas de frecuencia B_1 y B_2 . Este cuantificador provee información dinámica acerca del grado de importancia de cada banda. Así se ve que la fase pre-ictal está caracterizada por una preponderancia de los ritmos lentos (pre-ictal: $[B_5 + B_6] \sim 50\%$). La crisis comienza

a los 80 s con una descarga de ondas lentas sobreimpuestas a una actividad rápida de bajo voltaje. Esta descarga dura aproximadamente 8 s y produce una marcada actividad en las bandas de frecuencia B_5 y B_6 (banda Delta), que alcanza el 80% de la energía wavelet.

Comenzando a los 90 s, la actividad de baja frecuencia, representada por B_5 y B_6 , decrece abruptamente a valores relativos más bajos que el 10%, mientras que las otras bandas de frecuencia (bandas Theta y Alpha) se vuelven más importantes. Se observa también en Fig. 7.3 que el comienzo de la fase clónica está correlacionado con un aumento en la actividad de la banda de frecuencia B_4 .

Luego de 140 s, cuando las descargas clónicas comienzan a ser intermitentes, la actividad en B_5 crece nuevamente hacia el final de la crisis, cuando la actividad de B_6 también aumenta de muy rápidamente y la actividad de ambas bandas es claramente dominante. Las bandas de frecuencia B_5 y B_6 (banda delta) mantienen esta predominancia durante toda la etapa post-ictal.

Se concluye a partir de este ejemplo que la crisis está dominado por las bandas de frecuencia B_3 and B_4 (ritmos Alpha y Theta, 12.8 - 3.2 Hz), con un correspondiente decrecimiento abrupto de la actividad en las bandas de baja frecuencia B_5 y B_6 (ritmo Delta, 3.2 - 0.8 Hz) [106, 105]. Claramente este comportamiento puede ser asociado al “ritmo de reclutamiento epiléptico” descripto por Gastaut y Broughton [41, 42].

• Entropía Wavelet Total (TWS)

En la Fig. 7.4 la línea punteada representa la evolución temporal de la entropía wavelet normalizada (se incluyeron todas las bandas de frecuencia), mientras que la línea continua corresponde a los resultados que ignoran la contribución de la actividad muscular representada por las bandas de alta frecuencia (B_1 y B_2).

Es interesante observar el comportamiento de la TWS durante los primeros 10 s desde el comienzo de la crisis. Se ve que en este intervalo de tiempo la TWS exhibe valores crecientes cuando se incluyen todas las bandas, al comparar con los valores medios correspondientes a la etapa pre-ictal. Si no se incluyen las bandas de frecuencia B_1 and B_2 (bandas que reflejan principalmente la actividad muscular), claramente se observa una tendencia decreciente en los valores de la TWS inmediatamente luego del inicio de la crisis. Es claro que el valor medio de la TWS en este intervalo de tiempo es menor que aquel correspondiente a la etapa pre-ictal. Por lo tanto el comportamiento de la TWS después del comienzo de la crisis es compatible con un aumento en el grado de desorden del sistema, inducido por una actividad de alta frecuencia. Las actividades superpuestas de media y baja frecuencia, sin embargo, son responsables del comportamiento más ordenado de la señal “remanente”.

El comportamiento de TWS luego de 90 s (en ambos casos, con y sin inclusión de las bandas de alta frecuencia) es indicativo del hecho que el sistema exhibe una tendencia a ser más “ordenado”. Esta tendencia es mejor apreciada sin actividad muscular. Más aún, se puede notar que la TWS adopta un valor mínimo alrededor de 125 s, en coincidencia con el comienzo de la etapa clónica. El pico observado en la TWS en el entorno de los 145 s podría asociarse a la finalización del ritmo de reclutamiento epileptico. Pasado este punto la TWS muestra valores crecientes hasta los 155 s, que marca el final de la crisis. Se ve que TWS muestra valores casi constantes en la etapa post-ictal, similares a los obtenidos en la etapa pre-ictal.

Dos máximos relativos son observados entorno de los 145 s y 155 s. Estos tiempos pueden asociarse al *i*) final del ritmo de reclutamiento epiléptico y *ii*) final de la crisis, respectivamente. Los cambios producidos en la señal de EEG alrededor de los 125 s (transición de la etapa tónica a la clónica) son el resultado de un mecanismo enteramente diferente de aquel que produce las variaciones a los 145 s y 155 s. Como ha sido mencionado previamente estos cambios estarían asociados a la “fatiga” neuronal, y por ende a los mecanismos inhibitorios que determinan el final de la crisis.

Sintetizando, se puede asociar un más robusto grado de orden en la actividad del EEG durante la fase ictal que durante las etapas pre- y post-ictal, compatible con un proceso de sincronización en la actividad cerebral. Este comportamiento puede pensarse como inducido por un hipotético foco epiléptico que genera el ritmo de reclutamiento observado.

- **Complejidad Wavelet Total (TWC)**

En la Fig. 7.5 se muestran los resultados obtenidos al aplicar la complejidad wavelet total basada en la medida de Lopez Ruiz y coautores (TWC_{LMC}) y también la basada en la medida de Landsberg y coautores (TWC_{SDL}). En esta figura mostramos los resultados de *i*) la señal EEG original *ii*) la señal sin la contribución de las bandas de frecuencia B_1 y B_2 (asociadas principalmente a la actividad muscular). La región correspondiente al “ritmo de reclutamiento epiléptico” (90 – 145 s) tiene un aspecto diferente según estén o no todas las bandas de frecuencia. Podría pensarse que la actividad muscular destruye la complejidad subyacente. Esto si bien es esperable, muestra cuán eficientes son las medidas de complejidad propuestas. La nueva característica que enfatizaremos aquí es que la zona 90 – 145 s posee un alto grado de orden (baja entropía) *junto* con valores más altos de la complejidad (en particular $TWC_{SDL} \sim 1$). Evidentemente *coexisten* orden y complejidad. La caída de la complejidad hacia los 145 s es indicativa del final de la fase de reclutamiento. La caída más pronunciada a los 155 s señala la finalización de la crisis.

La Fig. 7.6.a muestra la evolución temporal de la serie EEG sin las contribuciones de B_1 y B_2 en el plano (H, C) ($C \equiv TWC_{(LMC)}$). Las cotas de máxima y mínima complejidad C_{max} y C_{min} , respectivamente, son graficadas en función de la entropía H , para $N_J = 4$ niveles de resolución wavelet. Notar que los valores correspondientes a las etapas pre-ictal y post-ictal, están concentradas en la zona de entropía relativamente alta ($H > 0.8$) y baja complejidad, mientras que un comportamiento diferente puede asociarse a la etapa ictal.

La Fig. 7.6.b muestra un detalle del plano (H, C) correspondientes al intervalo 90 – 145 s, asociados al ritmo de reclutamiento epiléptico. Se ve que la complejidad LMC es mayor en la etapa tónica 90 – 125 s que en la clónica 125 – 140 s. Más aún, nítidamente se aprecia el hecho que la “trayectoria” EEG permanece cercana a C_{max} durante la etapa tónica, mientras que después de un corto tiempo de transición (120 – 125 s), tiende a aproximarse a la complejidad mínima C_{min} , a medida que la trayectoria evoluciona hacia el final del ritmo de reclutamiento (145 s).

La evolución temporal para TWC_W es mostrada en la Fig. 7.7. La línea punteada y la sólida respectivamente, representan los valores obtenidos cuando las bandas de frecuencia B_1 y B_2 son incluídas o no en la evaluación de TWC_W . La complejidad

según Wootters TWC_W exhibe valores casi constantes, similares para la fase pre-ictal y post-ictal, cuando el análisis incluye las bandas B_1 y B_2 . No incluyendo estas dos bandas en el análisis los resultados son muy diferentes. En la etapa ictal los valores de TWC_W son mayores que en la pre-ictal. Se observa que TWC_W tiene una caída a los 120 s en correspondencia con la transición tónico-clónica. Esto se asocia a una marcada disminución de la entropía TWS, opuesto a lo que pasa tanto a los *i*) 140 s (final del ritmo de reclutamiento epiléptico), o *ii*) 158 s (finalización de la crisis), donde los valores de TWS crecen. Puede suponerse que un régimen dinámico diferente se tiene a los 120 s, el cual no coincide con el observado a los 140 s y a los 158 s.

En Fig. 7.8 se grafica la evolución temporal de la serie EEG en el plano (H, C_W). La máxima y la mínima complejidad C_W^{max} y C_W^{min} son graficadas en función de H , para el número de niveles de resolución usado en la evaluación de los cuantificadores ($N = 4$) [108, 84, 110]. Los valores correspondientes a las etapas pre- y post-ictal se concentran en la zona de entropía relativamente alta ($H > 0.8$) y baja complejidad. Un comportamiento muy diferente se asocia a la etapa ictal. Más aún, en la Fig. 7.8.b se puede distinguir claramente entre la etapa tónica y la clónica. La Fig. 7.8.b muestra datos correspondientes al intervalo 97.5 – 140 s, asociados con el ritmo de reclutamiento epiléptico. La trayectoria EEG permanece cerca de C_W^{max} durante la etapa tónica, mientras tiende a aproximar la complejidad mínima C_W^{min} cuando la trayectoria evoluciona hacia el final del ritmo de reclutamiento.

Las Figs. 7.9 y 7.10 muestran resultados similares a los anteriores obtenidos en este caso con la medida de complejidad estadística generalizada basada en la divergencia de Jensen-Shannon. En síntesis, se ha mostrado aquí que la fase de reclutamiento, además de presentar un elevado nivel de orden (ver Ref. [105]) exhibe valores altos de la medida de complejidad estadística generalizada TWC. Esto ha sido verificado en todos los casos cuando se usó la complejidad basada en la medida de *a*) LMC, *b*) SDL, *c*) Wootters o *d*) Jensen-Shannon. Al graficar la evolución temporal de la crisis epiléptica en el plano (H, C) las etapas tónica y clónica son claramente diferenciables, siendo esta distinción una de las principales contribuciones del presente análisis.

Resultados similares a los presentados previamente fueron obtenidos en 20 señales de EEG correspondientes a crisis tónico-clónica, registrados con identico protocolo y en la derivación C4. Estos registros corresponden a 8 pacientes diferentes [106, 105]. En las Figs. 7.11.a y 7.11.b se presentan los valores medios y los correspondientes errores standar de TWS y TWC_W , para las etapas pre-ictal e ictal, de las 20 señales analizadas. La medida TWS toma valores mas bajos en la etapa ictal que en la pre-ictal. En particular, si en su evaluación están incluídas todas las bandas de frecuencia (ver Fig. 7.11.a), la diferencia entre los valores medios es negativa en el 50 % de los casos ($\Delta H = H^{(ictal)} - H^{(pre-ictal)} < 0, p = 0.6737$, usando T-test de una-muestra). Pero el número de diferencias negativas aumenta hasta el 95 % ($\Delta H < 0, p = 0.0001$) si las bandas de frecuencia B_1 y B_2 son excluidas (ver Fig. 7.11.b). Inversamente, la complejidad TWC_W toma valores mas altos en la etapa ictal que en la pre-ictal. En particular, si en la evaluación están incluídas todas las bandas de frecuencia (ver Fig. 7.12.a), la diferencia entre los valores medios es positiva en el 25 % de los casos ($\Delta C_W = C_W^{(ictal)} - C_W^{(pre-ictal)} > 0, p = 0.9513$, usando T-test de una-muestra). Pero el número de diferencias positivas aumenta hasta el 85 % ($\Delta C_W > 0, p = 0.0002$) si

las bandas de frecuencia B_1 y B_2 son excluidas (ver Fig. 7.12.b). Por lo tanto, podría concluirse que la actividad eléctrica del cerebro durante una crisis epiléptica TC tiene asociado un comportamiento que si bien es más ordenado presenta elevada complejidad.

Un aspecto crítico de los resultados presentados es la posible distorsión debido a la propagación espacial de la crisis, ya que solo se analizaron datos registrados por el electrodo $C4$ y el área epileptogénica (foco epiléptico) de las crisis estaban principalmente localizadas en ubicaciones del lóbulo temporal (1 paciente con foco en el lóbulo temporal derecho; 3 pacientes con foco en el lóbulo temporal izquierdo; 2 pacientes con foco bitemporal; 2 pacientes sin foco claramente localizado [105]). Para solucionar este problema TWS y TWC_W fueron aplicadas también a registros provenientes de las derivaciones $T3$ y $T4$, obteniendo resultados similares a los obtenidos a partir del electrodo $C4$. Aún pensando que estos electrodos adicionales presentan mas artificios comparados con $C4$, el “ritmo de reclutamiento epiléptico” fue todavía observado, así como también el decrecimiento de la entropía junto con el aumento de complejidad, durante la etapa ictal. Al contar con registros de EEG provistos por electrodos localizados en diferentes sitios, podría suponerse que un análisis multicanal brindaría resultados más realistas. El problema del presente conjunto de datos es la inhomogeneidad en la localización del área epileptogénica, lo cual invalida el análisis multicanal usando un test ANOVA.

7.4 Conclusiones

En el presente capítulo se ha presentado una aplicación de las técnicas previamente introducidas, al análisis de señales electrencefalográficas (EEG) correspondientes a pacientes con epilepsia del tipo *Grand Mal*. El objetivo principal es mostrar la utilidad de las nuevas técnicas tanto en el diagnóstico de la epilepsia como en la caracterización dinámica del modelo subyacente a la correspondiente señal EEG. En este sentido, se ha introducido como aporte original al análisis de EEG, la noción de *complejidad wavelet*, que proporcionó un interesante enfoque a la hora de analizar crisis epilépticas y en particular el proceso de transición de fase tónico a la fase clónica.

Una ventaja importante de la entropía y la complejidad wavelet en relación a otros métodos de análisis de señales, es que los métodos aquí propuestos permiten realizar el análisis usando algoritmos simples que se aplican sobre segmentos de la señal de corta duración, lo que los hace particularmente convenientes para el estudio de crisis epilépticas. Efectivamente, la posibilidad de usar ventanas temporales de pocos segundos (2.5s) permite que las técnicas propuestas exhiban su capacidad discriminatoria al detectar, por ejemplo, tanto el inicio y como el final de la crisis. Pero sin duda el resultado más interesante que surge de esta aplicación, es la clara detección de la transición de fase tónico a la fase clónica que se produce al usar la medida de complejidad wavelet. Es importante resaltar que dicha transición es casi imperceptible al realizar una inspección visual del EEG.

Por otra parte, se ha visto también que las nuevas técnicas basadas en wavelets describen con precisión el ritmo de reclutamiento epiléptico reportado por Gastaut y Fischer-Williams [42] y Gastaut y Broughton [41] para crisis epilépticas tónico-clónica. Se mostró además que durante la fase de reclutamiento la medida de complejidad wavelet alcanza elevados valores, lo que sugeriría la presencia de un estado cerebral cuya dinámica

subyacente posee orden con máxima complejidad. Sumando estos resultados al hecho de que una dinámica caótica (provista por la evaluación de la evolución temporal de la caoticidad en función del tiempo [106]) puede ser asociada a la totalidad de este tipo de registros EEG, puede concluirse que esto es una evidencia de un proceso de auto-organización [12, 53]).

Luego puede conjeturarse que en el caso de las crisis epilépticas tónico-clónica, el foco epiléptico induce un estado cerebral auto-organizado caracterizado por orden y alta complejidad [84, 110]. En síntesis el uso de los cuantificadores propuestos basados en los métodos tiempo-frecuencia pueden contribuir no sólo al análisis de señales de EEG, sino también a una mejor comprensión de la dinámica cerebral asociada.

Tabla 7.1: INTERVALOS DE FRECUENCIA (EN Hz) ASOCIADOS CON LOS NIVELES DE RESOLUCIÓN WAVELET j Y TIEMPO DE RESOLUCIÓN ΔT (EN s) PARA UNA FRECUENCIA DE MUESTREO $\omega_s = 102.4\ Hz$. LAS BANDAS DE FRECUENCIA EEG TRADICIONALES CORRESPONDEN A LAS SIGUIENTES INTERVALOS DE FRECUENCIAS: δ (0.5-3.5 Hz); θ (3.5-7.5 Hz); α (7.5-12.5 Hz); β (12.5-30.0 Hz); γ (30.0-70.0 Hz).

Notation	Wavelet Band				EEG Band
	ω_{min}	ω_{max}	j	ΔT	
B_1	25.6	51.2	-1	0.0195	β, γ
B_2	12.8	25.6	-2	0.0391	β
B_3	6.4	12.8	-3	0.0781	θ, α
B_4	3.2	6.4	-4	0.1562	θ
B_5	1.6	3.2	-5	0.3125	δ
B_6	0.8	1.6	-6	0.6250	δ

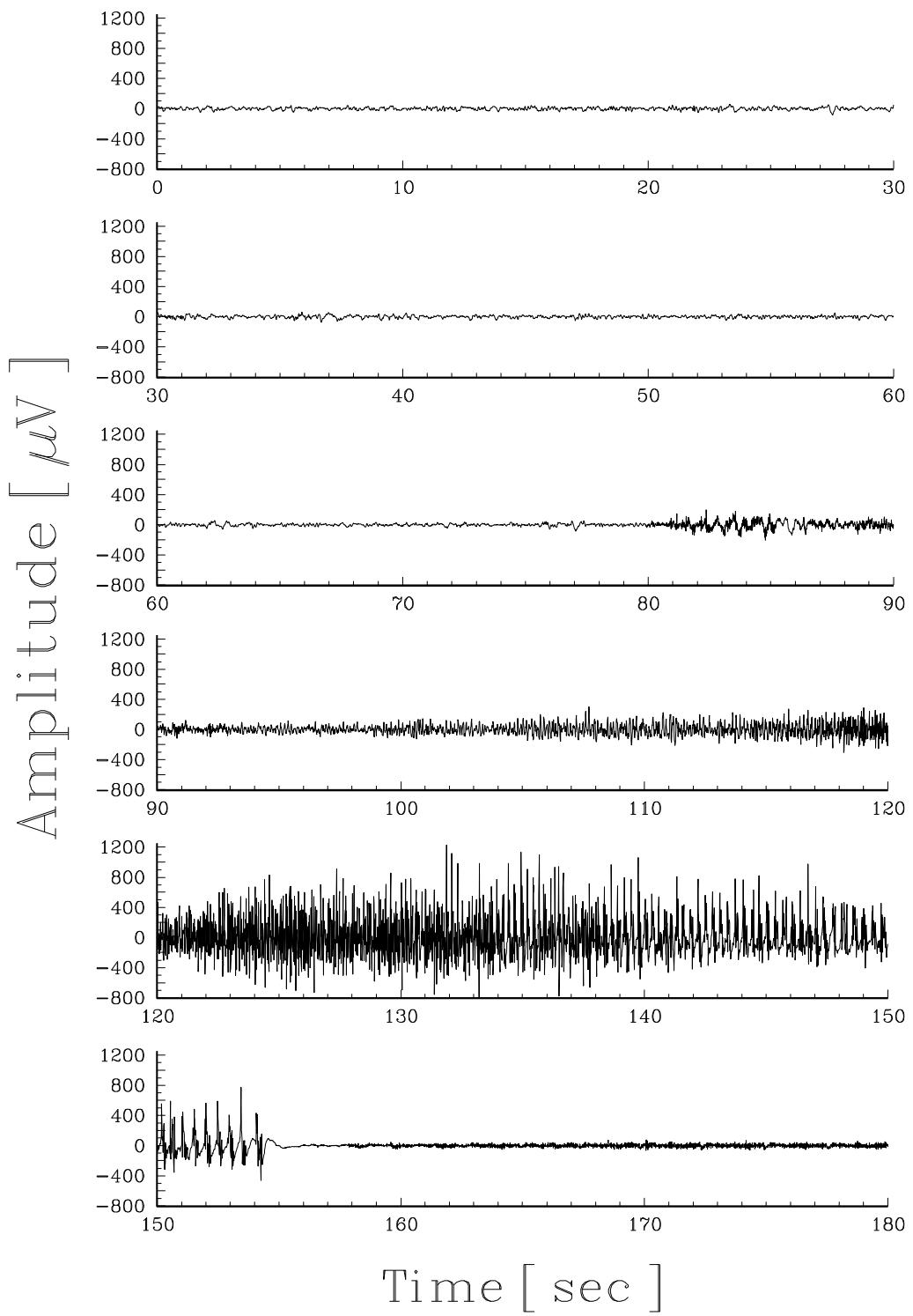


Figura 7.1: SEÑAL EEG (*scalp*) CORRESPONDIENTE A UNA CRISIS EPILEPTICA TÓNICO-CLÓNICA, REGISTRADO EN LA UBICACIÓN CENTRAL DERECHA C4. LA CRISIS COMIENZA A LOS 80 s Y LA FASE CLÓNICA A LOS 125 s. LA CRISIS FINALIZA A LOS 155 s.

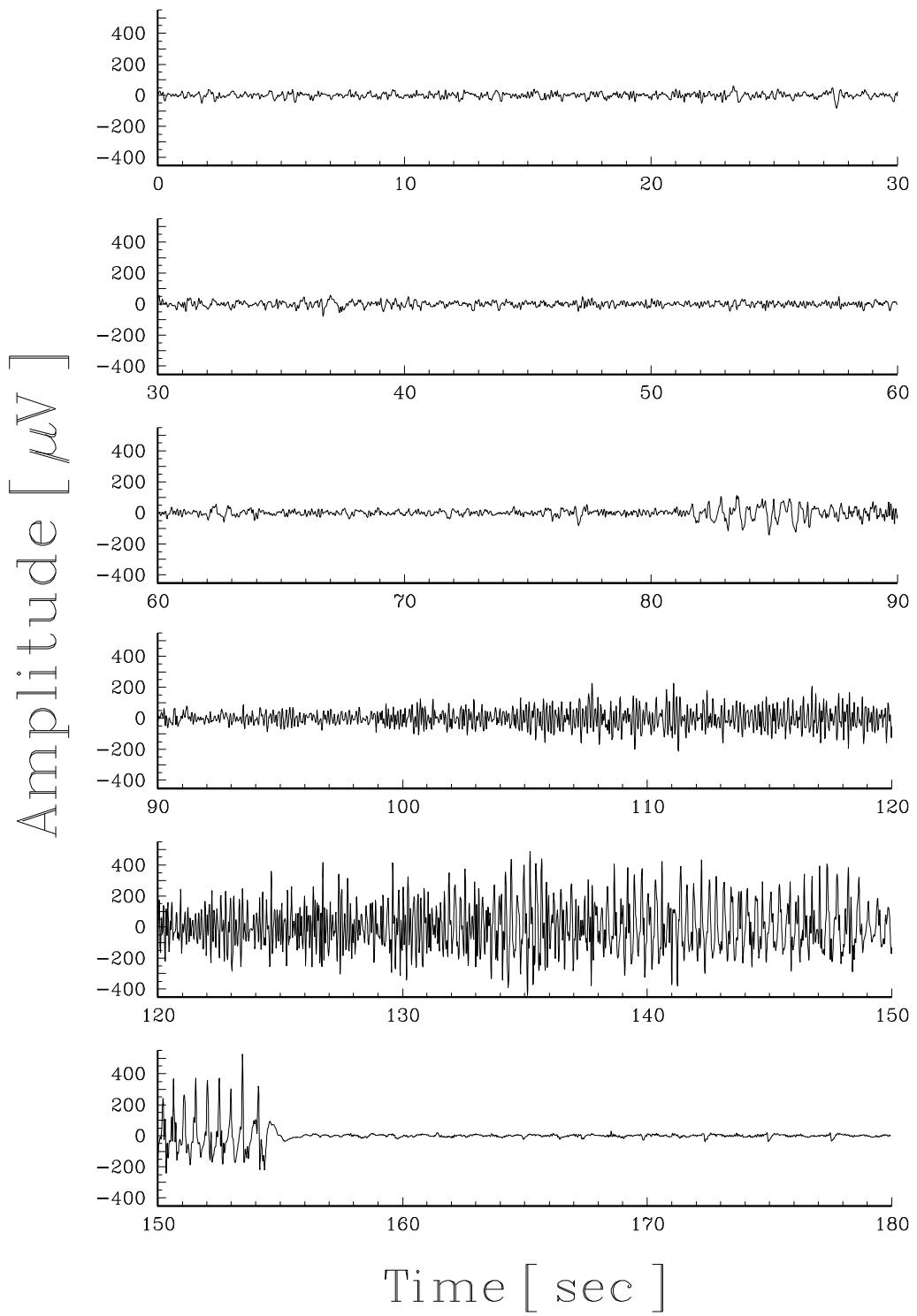


Figura 7.2: SEÑAL EEG LIBRE DE ARTIFICIOS (RUIDO). SIMILAR A LA FIG. 7.1 SIN LA CONTRIBUCIÓN DE LA ACTIVIDAD MUSCULAR REPRESENTADA POR BANDAS DE FRECUENCIA B_1 Y B_2 , VER TABLA. 7.1.

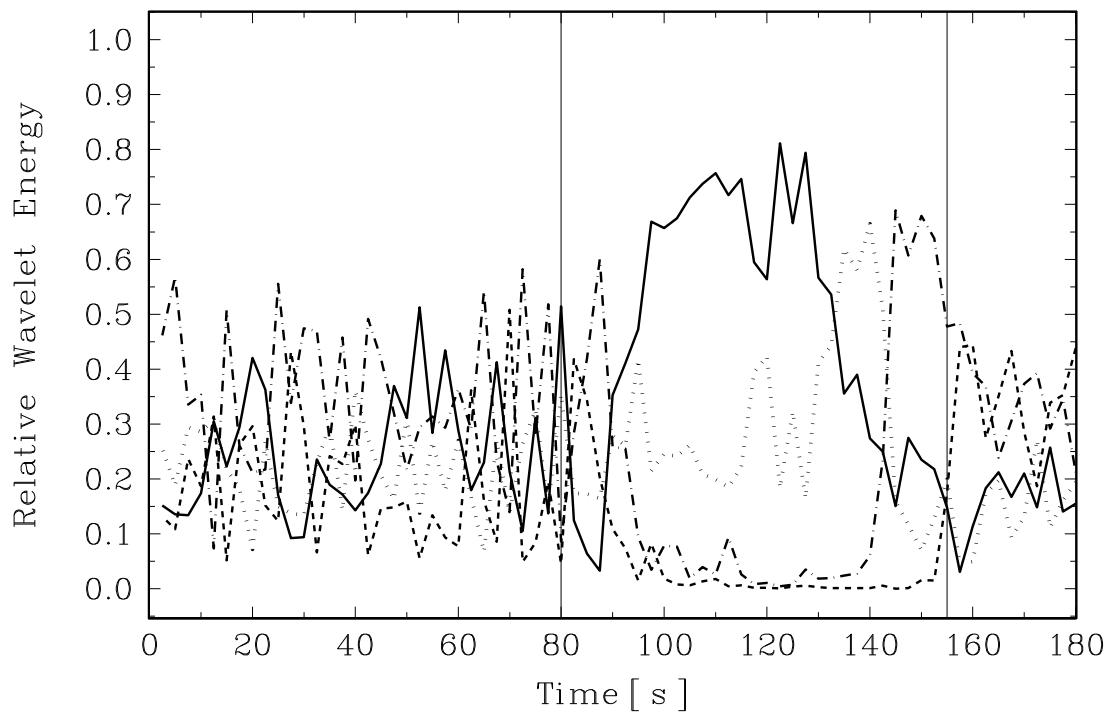


Figura 7.3: EVOLUCIÓN TEMPORAL DE LA ENERGÍA MULTIESCALA WAVELET CORRESPONDIENTE A LA SEÑAL EEG SIN ACTIVIDAD MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2) PARA LAS BANDAS DE FRECUENCIA B_3 (LÍNEA SÓLIDA), B_4 (LÍNEA PUNTEADA), B_5 (LÍNEA PUNTO-GUIÓN), B_6 (LÍNEA GUIÓN). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILEPTICA.

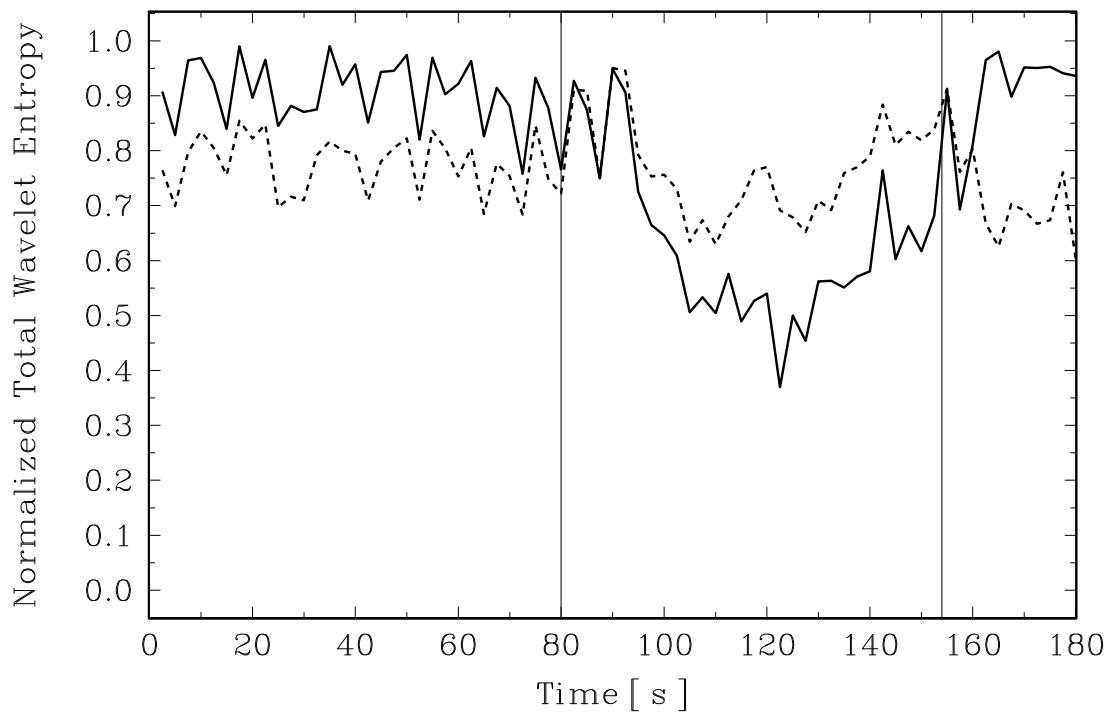


Figura 7.4: EVOLUCIÓN TEMPORAL DE LA ENTROPÍA WAVELET TOTAL (TWS). LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DEL CRISIS EPILÉPTICA.

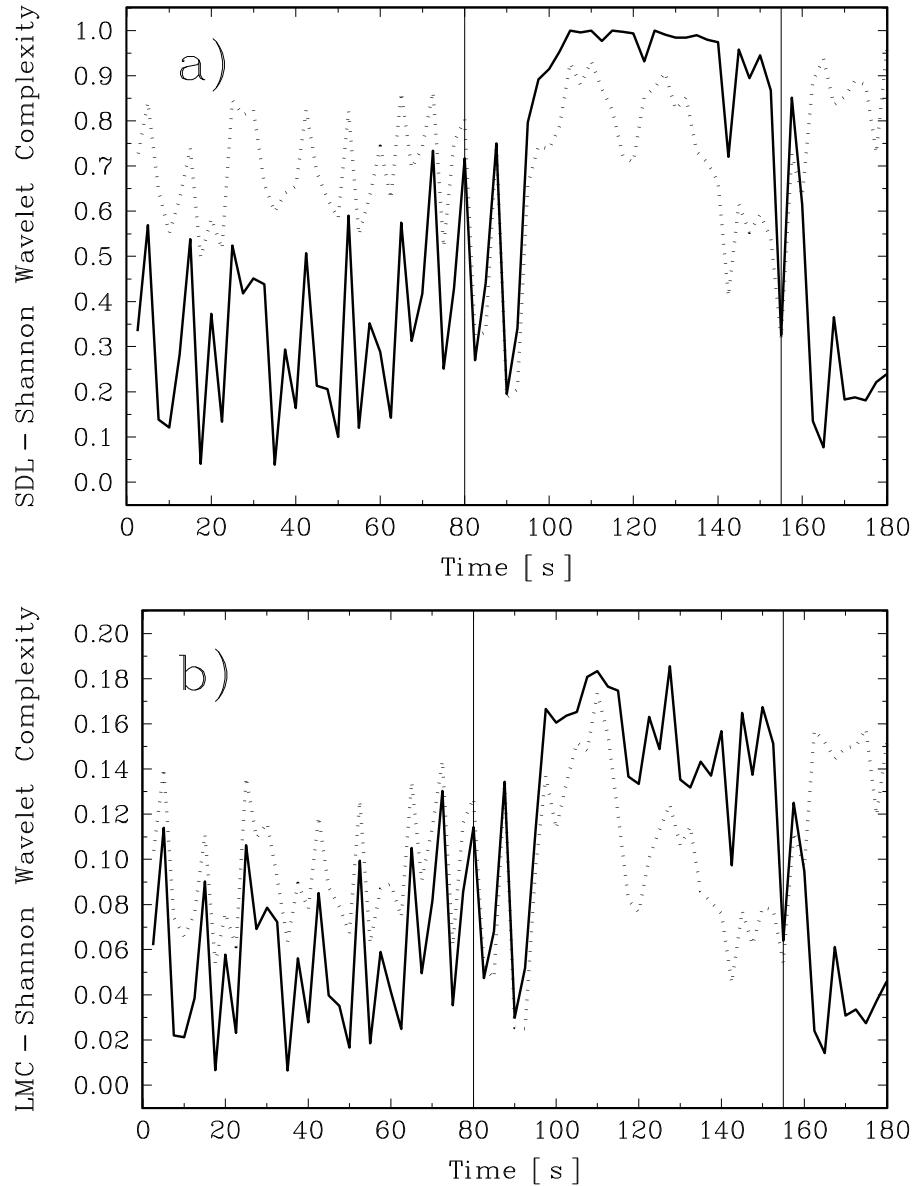


Figura 7.5: a) EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_{SDL} . LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTRÓPIA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). b) IGUAL AL ANTERIOR PARA LA EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_{LMC} . LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.

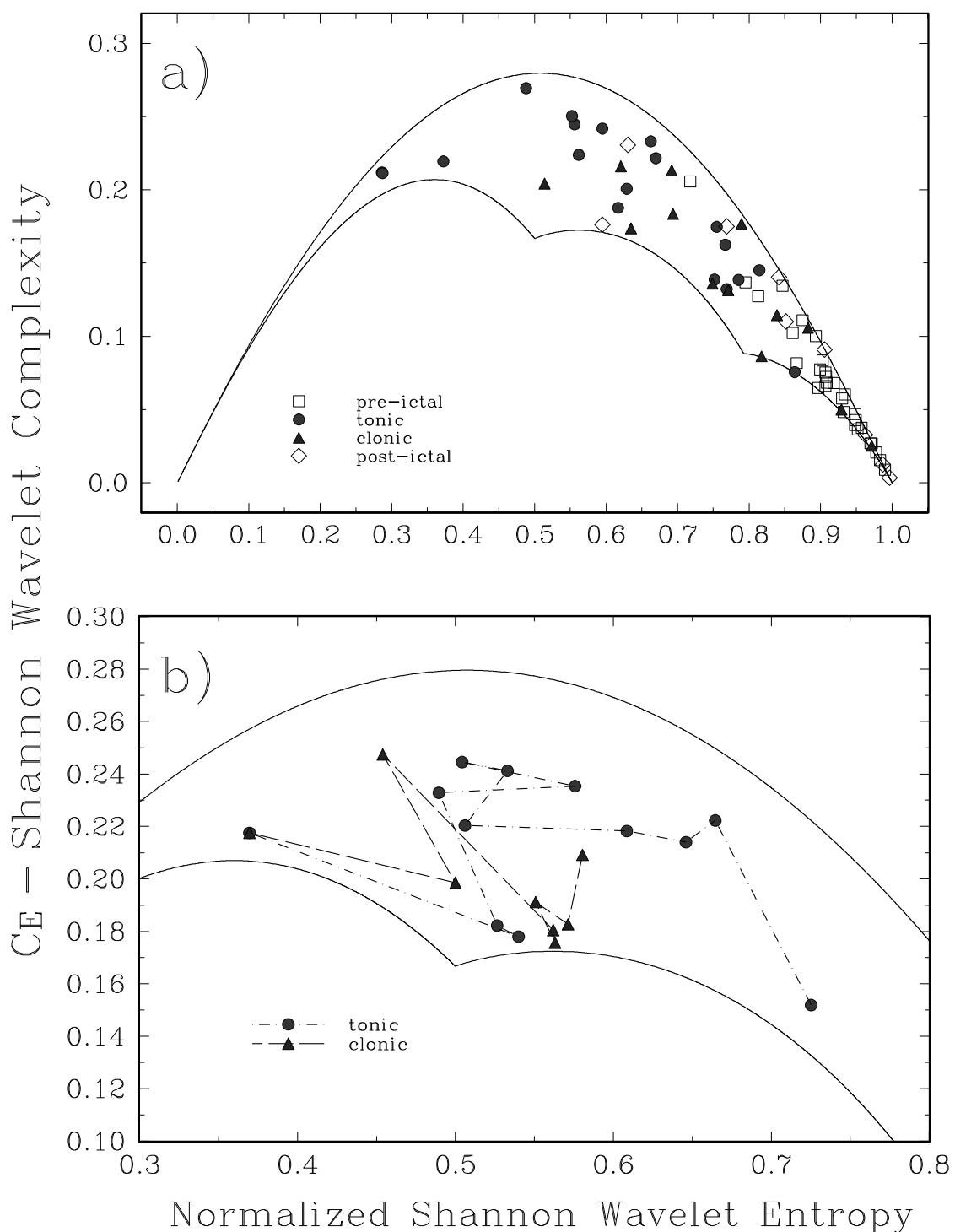


Figura 7.6: a) DIAGRAMA “COMPLEJIDAD *vs* ENTROPÍA NORMALIZADA” CORRESPONDIENTE A TWC_{LMC} PARA LA SEÑAL EEG LIBRE DE ARTIFICIOS (FIG. 7.2). CON LÍNEA CONTINUA SE HA REPRESENTADO LA COMPLEJIDAD MÁXIMA C_{max} Y MÍNIMA C_{min} EN FUNCIÓN DE LA ENTROPÍA WAVELET SHANNON NORMALIZADA. b) EVOLUCIÓN EN ESTE ESPACIO, PARA LAS FASES TÓNICA Y CLÓNICA. DETALLE CORRESPONDIENTE AL INTERVALO ASOCIADO CON EL RITMO DE RECLUTAMIENTO EPILÉPTICO.

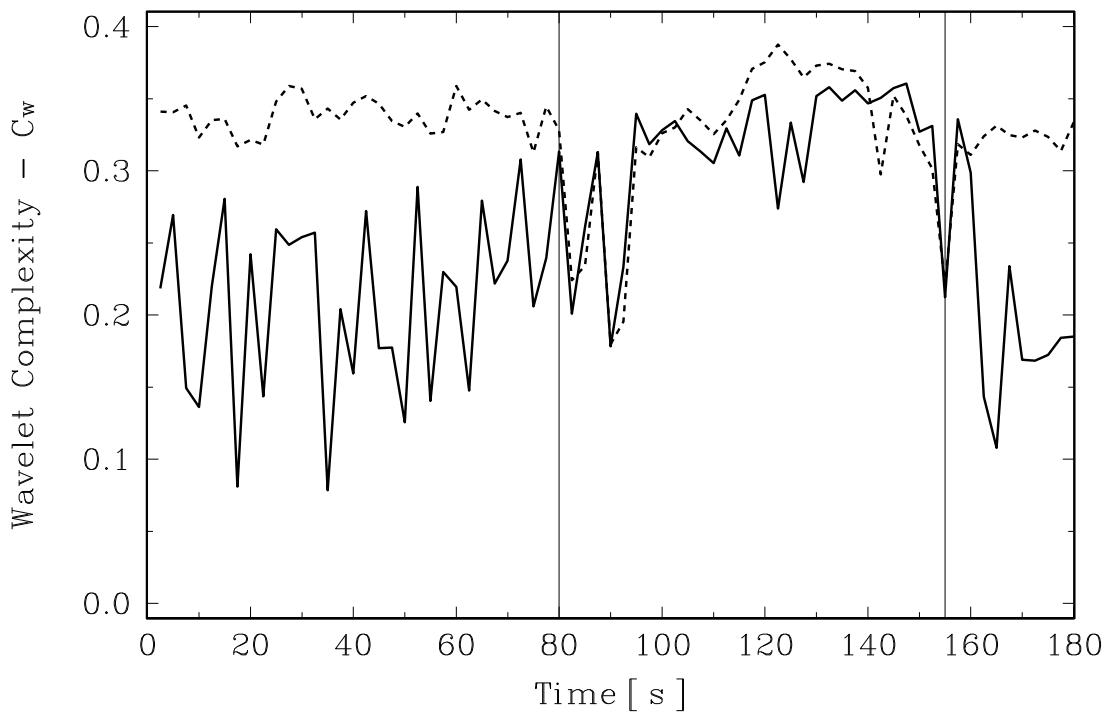


Figura 7.7: EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_W . LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.

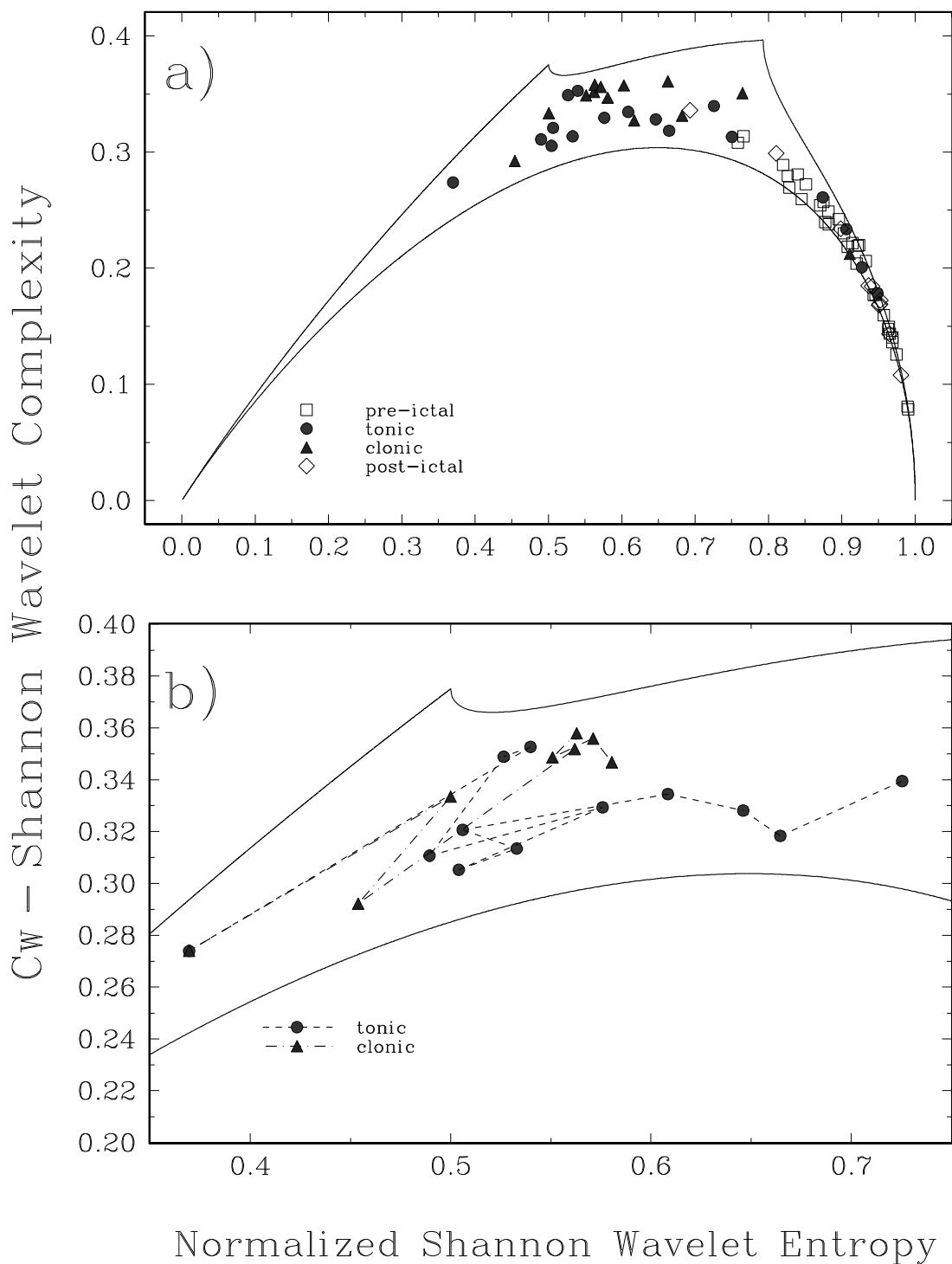


Figura 7.8: a) DIAGRAMA “COMPLEJIDAD *vs* ENTROPÍA NORMALIZADA” CORRESPONDIENTE A TWC_W PARA LA SEÑAL EEG LIBRE DE ARTIFICIOS (FIG. 7.2). CON LÍNEA CONTINUA SE HA REPRESENTADO LA COMPLEJIDAD MÁXIMA C_{max} Y MÍNIMA C_{min} EN FUNCIÓN DE LA ENTROPÍA WAVELET SHANNON NORMALIZADA. b) EVOLUCIÓN EN ESTE ESPACIO, PARA LAS FASES TÓNICA Y CLÓNICA. DETALLE CORRESPONDIENTE AL INTERVALO ASOCIADO CON EL RITMO DE RECLUTAMIENTO EPILÉPTICO.

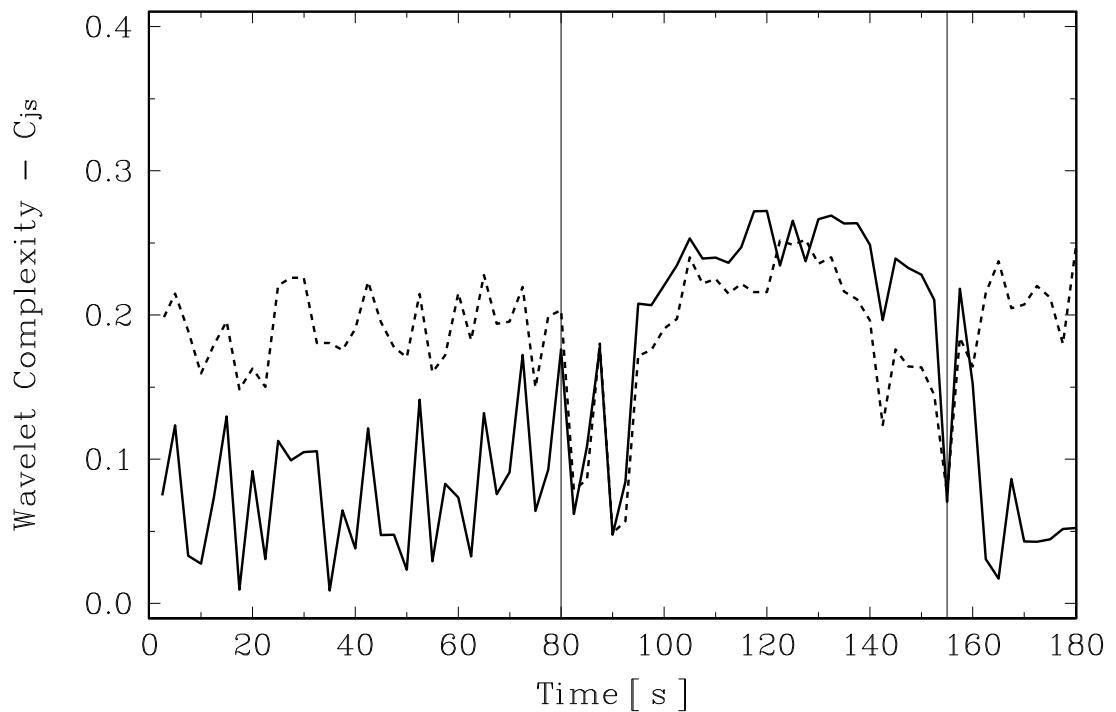


Figura 7.9: EVOLUCIÓN TEMPORAL DE LA COMPLEJIDAD WAVELET TOTAL TWC_{JS} . LA LÍNEA PUNTEADA Y LA SÓLIDA REPRESENTAN, RESPECTIVAMENTE, LA EVOLUCIÓN TEMPORAL DE LA ENTROPÍA PARA LA SEÑAL EEG ORIGINAL (FIG. 7.2) Y SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS, FIG. 7.2). LAS LÍNEAS VERTICALES REPRESENTAN EL COMIENZO Y EL FINAL DE LA CRISIS EPILÉPTICA.

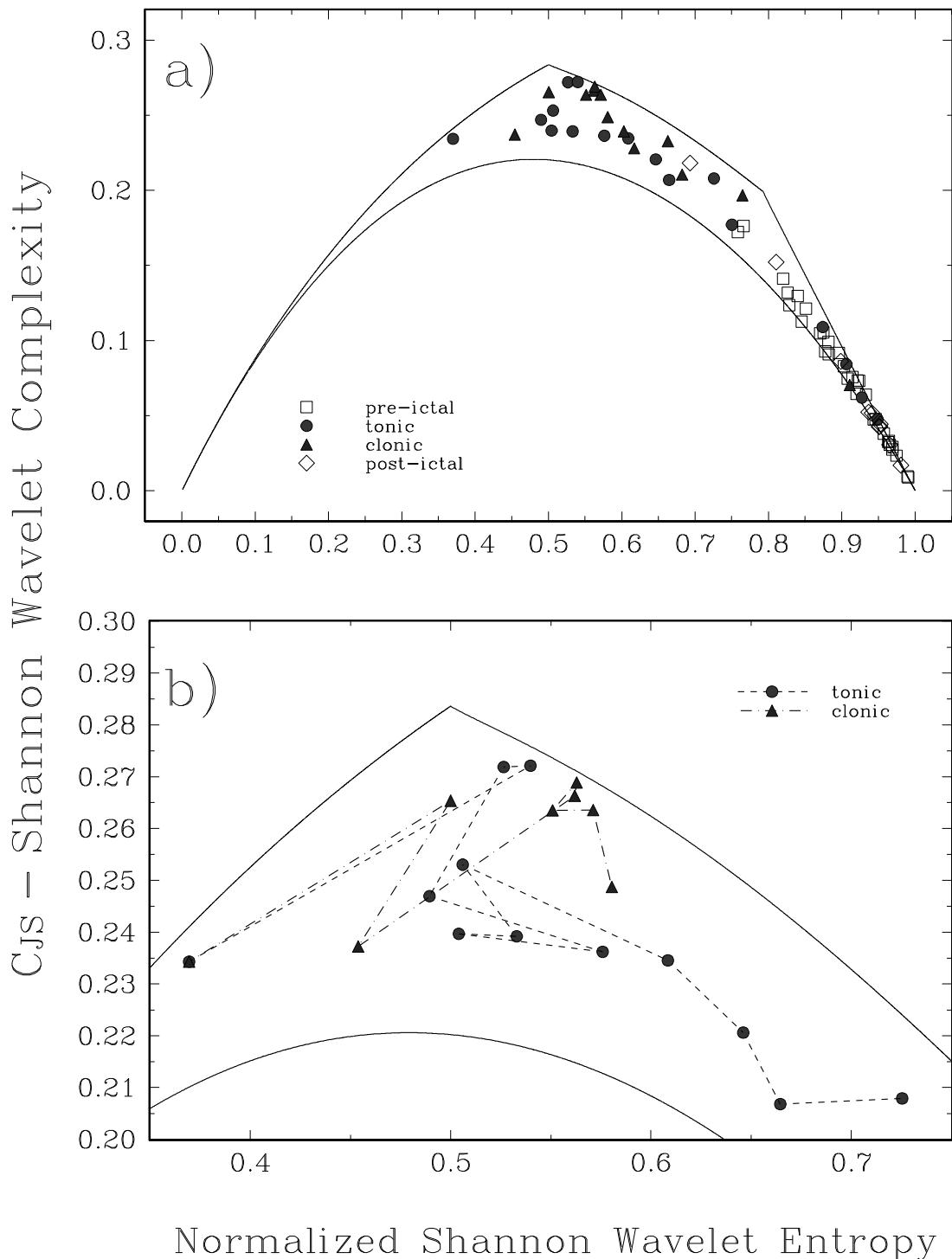


Figura 7.10: *a)* DIAGRAMA “COMPLEJIDAD vs ENTROPÍA NORMALIZADA” CORRESPONDIENTE A TWC_{JS} PARA LA SEÑAL EEG LIBRE DE ARTIFICIOS (FIG. 7.2). CON LÍNEA CONTINUA SE HA REPRESENTADO LA COMPLEJIDAD MÁXIMA C_{max} Y MÍNIMA C_{min} EN FUNCIÓN DE LA ENTROPÍA WAVELET SHANNON NORMALIZADA. *b)* EVOLUCIÓN EN ESTE ESPACIO, PARA LAS FASES TÓNICA Y CLÓNICA. DETALLE CORRESPONDIENTE AL INTERVALO ASOCIADO CON EL RITMO DE RECLUTAMIENTO EPILÉPTICO.

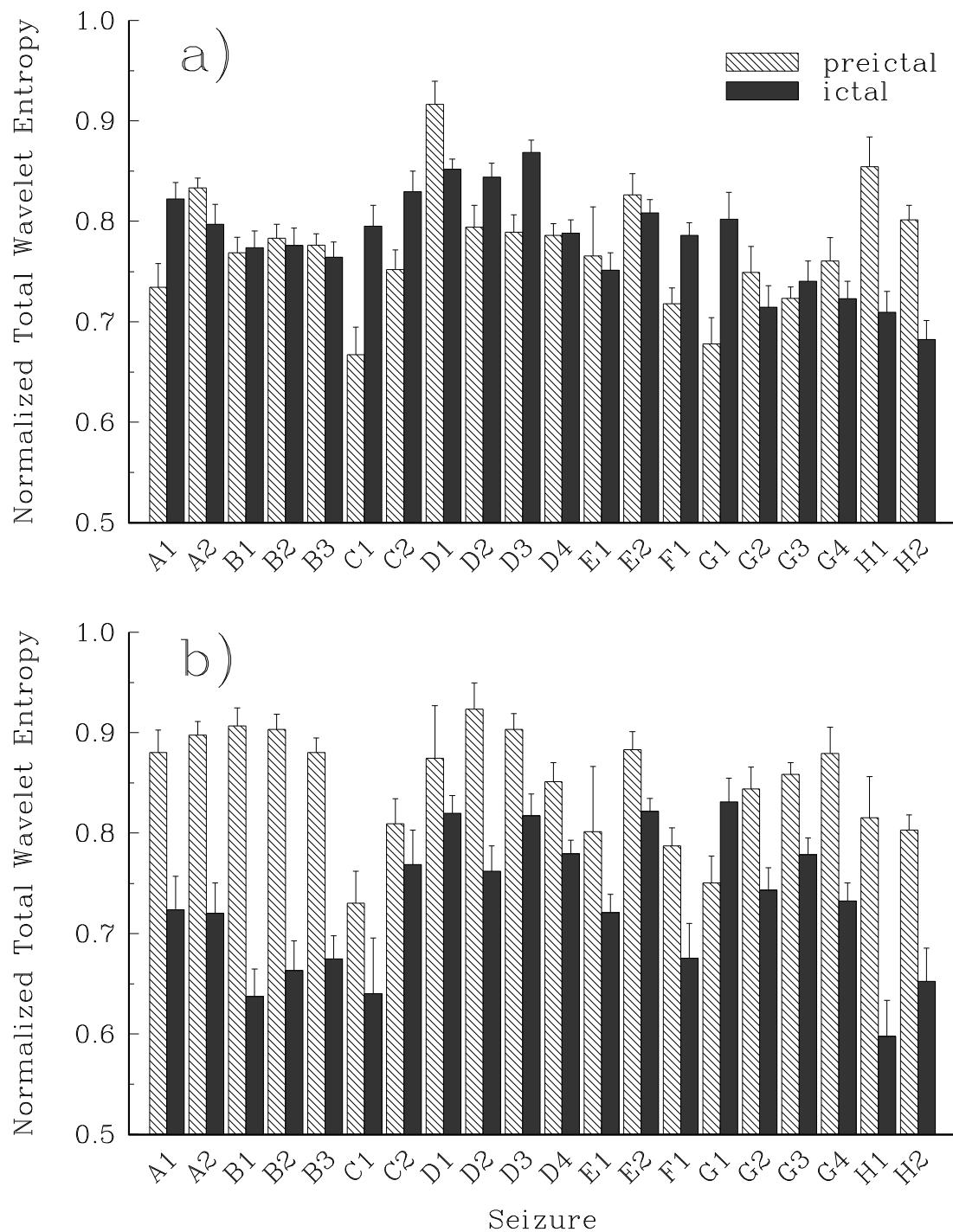


Figura 7.11: TWS ($media \pm SE$) EVALUADO PARA LAS 20 CRISIS EPILEPTICAS TÓNICO-CLÓNICAS PARA LOS PERÍODOS PRE-ICTAL E ICTAL. a) SEÑAL EEG ORIGINAL. b) SEÑAL EEG SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS). CADA PACIENTE ES REPRESENTADO POR UNA LETRA Y EL NÚMERO A SU LADO IDENTIFICA LA CRISIS CORRESPONDIENTE.

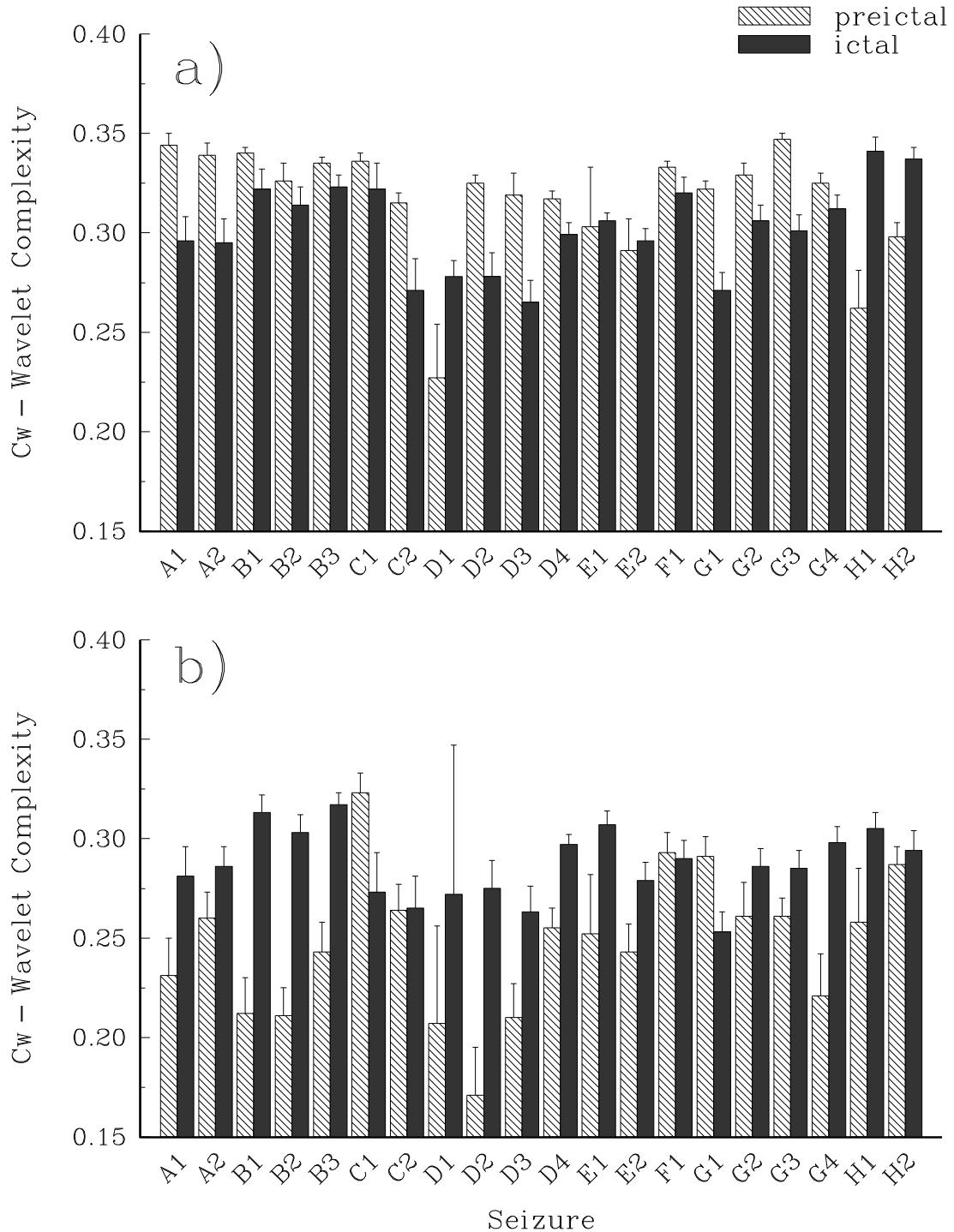


Figura 7.12: $TWC_W(\text{media} \pm SE)$ VALUADO PARA LAS 20 CRISIS EPILEPTICAS TÓNICO-CLÓNICAS PARA LOS PERÍODOS PRE-ICTAL E ICTAL. a) SEÑAL EEG ORIGINAL. b) SEÑAL EEG SIN CONTRIBUCIÓN MUSCULAR (LAS CONTRIBUCIONES DE LAS BANDAS B_1 Y B_2 SON CONSIDERADAS NULAS). CADA PACIENTE ES REPRESENTADO POR UNA LETRA Y EL NÚMERO A SU LADO IDENTIFICA LA CRISIS CORRESPONDIENTE.

Capítulo 8

Resumen y consideraciones finales

La idea que orientó la realización de esta tesis fue el desarrollo de métodos con capacidad de detección y cuantificación de las variaciones en las dinámicas subyacentes a series temporales no estacionarias. De esta forma, se pueden caracterizar aspectos específicos de los fenómenos físicos asociados a dichas series temporales. En particular los desarrollos fueron orientados hacia medidas que cuantifiquen el grado de “estructura” o “complejidad” de una señal, mediante algoritmos de fácil implementación y cálculo.

La noción de complejidad a partir de una descripción probabilística del sistema de interés constituye en sí misma una propuesta interesante. Esta medida de la complejidad, *complejidad estadística* C , es al mismo tiempo fácil de calcular e intuitivamente coherente, y su extensión en el contexto del análisis wavelet, TWC, adquiere las propiedades de localización que la convierten en una herramienta útil para cuantificar la complejidad multiescalas.

Con la idea de que una mejor medida de complejidad estadística básica C , conducirá a una mejor medida de complejidad wavelet TWC, se generalizó el desequilibrio Q en base a la noción de distancia estadística de Wootters y también mediante la divergencia de Jensen-Shannon. Aplicaciones de estas nuevas medidas de complejidad estadística, C_W y C_{JS} , a la señal generada con el mapa logístico, mostraron importantes mejoras en el comportamiento de este cuantificador.

En efecto, cuando el parámetro de control r del mapa logístico es variado con continuidad a través del tiempo, existen considerables cambios en la dinámica subyacente. La medida de complejidad C_W mostró ser sensitiva a estos cambios, manifiestándose además de manera compatible al exponente de Lyapunov, lo que no ocurre con la medida original C . Por otra parte, una medida de complejidad debería ser capaz de distinguir entre estructuras de diferentes periodicidades, lo cual es satisfecho por C_W , como quedó demostrado en el ejemplo correspondiente al mapa logístico. Sin embargo, C_W es una cantidad termodinámicamente ni “intensiva” ni “extensiva”.

En cuanto a la medida de complejidad asociada a la divergencia de Jensen-Shannon C_{JS} , también detecta con claridad los cambios dinámicos manteniendo coherencia con la física subyacente. Lo interesante en este caso, es que adiciona la propiedad de comportarse como una cantidad termodinámica intensiva, lo cual tratándose de una medida de complejidad resulta muy ventajoso.

El estudio detallado de los valores que alcanzan las medidas de complejidad estadística, llevado a cabo en el plano *complejidad vs. entropia*, $C \times H$, permitió mostrar

la existencia de cotas para la medida C en función de H , llamadas C_{max} y C_{min} . Estos resultados revisten gran interés, ya que el uso de diagramas $C \times H$ ofrece la posibilidad de estudiar los sistemas desde un punto de vista explícitamente independiente de los parámetros de control. Se obtiene así una visión que permite analizar al sistema directamente en términos de la información intrínseca del proceso. Este enfoque también brinda la posibilidad de vincular la complejidad del sistema en estudio, con los valores extremos de la medida.

Las medidas de complejidad estadística generalizada basadas en wavelets, fueron aplicadas al análisis de señales electrencefalográficas (EEG) correspondientes a pacientes con epilepsia generalizada de tipo tónico-clónica. El objetivo principal fue mostrar la utilidad de las nuevas técnicas tanto en el diagnóstico de la epilepsia como en la caracterización dinámica del modelo subyacente a la correspondiente señal EEG. Una ventaja importante de la entropía y la complejidad wavelet es la simplicidad de su implementación y su aplicabilidad sobre segmentos de la señal de corta duración.

El aporte original al análisis de señales EEG, es la noción de complejidad wavelet, que proporcionó un interesante enfoque a la hora de analizar este tipo de crisis epilépticas y en particular el proceso de transición de la fase tónico a la fase clónica. En base a los resultados obtenidos, puede conjeturarse que en el caso de las crisis epilépticas tónico-clónica, el foco epiléptico induce un estado cerebral auto-organizado caracterizado por orden y alta complejidad. En síntesis el uso de los cuantificadores propuestos basados en los métodos tiempo-frecuencia pueden contribuir no sólo al análisis de señales de EEG, sino también a una mejor comprensión de la dinámica cerebral asociada.

Agradecimientos

Quisiera agradecer:

A mi Director, Dr. Angel L. Plastino, por su comprensión y estímulo permanente.
A mi co-Director y amigo, Dr. Osvaldo A. Rosso, por el entusiasmo que me trasmitió en todo momento y la fundamental ayuda que me brindó durante la realización de este trabajo.

A todos mis compañeros del Grupo de Física y muy especialmente a mis amigas, Graciela y Elsa, por su apoyo incondicional y constante.

A la Secretaría del Instituto de Física, por su colaboración en diversas etapas de este trabajo.

Y por último a mi familia, Carlos, Mariana, Andrés y Paula, por ser ellos los que impulsaron, sostuvieron y lograron, la realización de esta tesis.

Por los datos EEG, quisiera agradecer al Dr. Adrian Rabinowicz, Departamento de Neurología y Programa de Epilepsia, del Instituto de Investigaciones Neurológicas Raúl Carrera (FLENI), Buenos Aires, Argentina, por los registros EEG con electrodos de superficie, utilizados en este trabajo.

La Plata, 30 de enero de 2004.

Bibliografía

- [1] H. D. I. Abarbanel. *Analysis of observed chaotic data*. Springer-Verlag, New York, 1996.
- [2] A. Aldroubi and M. Unser. *Wavelet in Medicine and Biology*. CRC Press, Boca Raton, 1996.
- [3] S. M. Ali and M. Silvey. *Jal. Royal Stat. Soc. B* 28 (1966) 131.
- [4] P. W. Anderson. *Physics Today*, July 1991.
- [5] C. Anteneodo and A. R. Plastino. *Phys. Lett. A* 223 (1997) 348.
- [6] A. P. Anokhin, N. Birbaumer, W. Lutzenberger, A. Nikolaev and F. Vogel. *Electr. and Clin. Neurophysiol.* 63 (1996) 99.
- [7] J. E. Arle and R. H. Simon. *Electr. and Clin. Neurophysiol.* 75 (1990) 296.
- [8] A. Babloyantz. Evidence of chaotic dynamics of brain activity during the sleep cycle. In Meyer-Kress (Ed). *Dimensions and Entropies in Chaotic Systems*, pp. 241-245. Springer Verlag, Berlin, 1986.
- [9] A. Babloyantz and A. Destexhe, *Chaos in neural networks. Proceedings of the IEEE 1st International Conference on Neural Networks*. M. Caudill and C. Butler, Vol. 4, pp. 31–40, San Diego (1987).
- [10] E. Başar. *EEG-Brain dynamics. Relation between EEG and brain evoked potentials*. Amsterdam, Elsevier, 1980.
- [11] E. Başar. *Chaos in Brain Function*. Springer Verlag, Berlin, 1990.
- [12] E. Başar, H. Flohr, H. Haken, A. J. Mandell (Eds). *Sinergetics of the Brain*. Springer, Berlin, 1983.
- [13] M. Basseville. *Information: entropies, divergences et moyennes*. Institut de Recherche en Informatique et Systèmes Aleatoires (IRISA), INRIA, France, 1996.
- [14] H. Berger. *Arch. Psychiat. Nervenkr.* 87 (1929) 527.
- [15] P. Binder and N. Perry. *Phys. Rev. E* 62 (2000) 2998 .
- [16] X. Calbet and R. López-Ruiz. *Phys. Rev. E* 63 (2001) 066116.

- [17] A. Capurro, L. Diambra, D. Lorenzo, O. Marcadar, M. T. Martín, A. Plastino, E. Rofman, M. E. Torres and J. Velluti. INRIA, Rapport de Recherche, 3184 , 1997.
- [18] M. C Casdagli, L. D. Iasemedis, R. S. Savit, R. L. Gilmore, S. Roper and J. C. Sackellares. Electr. and Clin. Neurophysiol. 102 (1997) 98.
- [19] G. Chaitin, J. Assoc. Comput. Mach. 13 (1966) 547.
- [20] C. K. Chui, *An Introduction to wavelets*. Academic Press, Boston, 1992.
- [21] R. R. Coifman, *Adapted multiresolution analysis, computation, signal processing and operator theory*. Proceeding of the International Congress of Mathematicians, Qioto, Japan, (1990) 879.
- [22] M. Costa, A. Goldberg and C. Peng Phys. Rev. Lett. 89 (2002).
- [23] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [24] J. C. Crepeau and L. K. Isaacson. J. Non-Equilib. Thermodyn. 15 (1990) 115.
- [25] J. P. Crutchfield. Phys. Rev. Lett. 63 (1989) 105.
- [26] J. P. Crutchfield and D. P. Feldman. Phys. Rev. E 55 (1997) 54.
- [27] J. Crutchfield, D. Feldman and C. Shalizi. Phys. Rev. E 62 (2000) 2996.
- [28] I. Csiszar. Studia Scientiarum Mathematicarum Hungarica 2 (1967) 299.
- [29] C. E. D'Atellis and Fernandez-Berdaguer, Eds. *Wavelet theory and harmonic analysis in applied sciences*. Birkhauser, 1997.
- [30] I. Daubechies. *Ten Lectures on wavelets*. SIAM, Philadelphia, 1992.
- [31] B. Edmonds. PhD Thesis, University of Manchester, 1999.
<http://www.cpm.mmu.ac.uk>
- [32] Th. Elbert, W. Ray, Z. Kwalik, J. Skinner, K. Graf, and N. Birbaumer. Physiological Reviews 74 (1994) 1.
- [33] D. P. Feldman and J .P. Crutchfield. Phys. Lett. A 238 (1998) 244.
- [34] J. Fell, J. Roeschke, K. Mann and C. Schaeffner. Electr. and Clin. Neurophys. 98 (1996) 401.
- [35] R. A. Fisher. Proc. Camb. Soc. 22 (1925) 700.
- [36] B. R. Frieden. Founds. Phys. 21 (1991) 757.
- [37] B. R. Frieden. *Physics from Fisher Information*. Cambridge University Press, Cambridge, England, 1998.
- [38] B. R. Frieden and B. H. Soffer. Phys. Rev. E 52 (1995) 2274.

- [39] J. Froyland and K. H. Alfsen. Phys. Rev. A 29 (1984) 2928.
- [40] D. Gabor J. Inst. Elec. Eng. 93 (1946) 429.
- [41] H. Gastaut and R. Broughton. *Epileptic Seizures*. C. C. Thomas (Ed). Springfield, Illinois, 1972.
- [42] H. Gastaut and M. Fischer-Williams. In: J. Field, H. W. Mgoun, V. E. Hall (Eds). *Hambook of Physiology*, Vol. 1, p.329. Williams and Wilkins, Baltimore, 1959.
- [43] M. Gell-Mann. *The Quark and the Jaguar: Adventures in the simple and the Complex*. Freeman, 1994.
- [44] M. Gell-Mann. Complexity 1 (1995) 16.
- [45] B. V. Gnedenko. *The Theory of Probability*. Chelsea, New York, 1962.
- [46] A. Goldberger, D. Rigney, J. Mietus, E. Antman, and S. Greenwald. Experientia 44 (1988) 983.
- [47] A. Goldberger, D. Rigney, and B. West. *Chaos and fractals in physiology*. Scientific American, 262 (1990) 43
- [48] J. Gotman, J. R. Ives and P. Gloor, Electroenceph. Clin. Neurophysiol. 52 (1981) 626.
- [49] P. Grassberger. Int. J. Theor. Phys. 25 (1986) 907.
- [50] P. Grassberger. Helvetica Physica Acta 62 (1989) 498.
- [51] P. Grassberger, Th. Schreiber and C. Schaffrath. Int. Jour. of Bifurcation and Chaos 1 (1991) 521.
- [52] I. Grosse, P. Bernaola-Galvan, P. Crapena, R. Román-Roldán, J. Oliver, H. E. Stanley. Phys. Rev. E 65 (2002) 41905.
- [53] H. Haken. *Information and Self-Organization: A Macroscopic Approach to Complex Systems*. Springer, Berlin, 2000.
- [54] R. V. L. Hartley. Bell Sys. Tech. J. (1928) 535.
- [55] J. Havrda and F. Charvat. Kybernetika. 3 (1967) 30.
- [56] M. Hénon. Commun. Math. Phys. 50 (1676) 69.
- [57] J. Hilgevoord and J. Uffink. Fund. of Phys. 21 (1991) 323.
- [58] B. A. Huberman and T. Hogg. Physics D 22 (1986) 376.
- [59] L. D. Iasemidis and J. C. Sackellares. *The evolution with time of spatial distribution of the largest Lyapunov exponent on the human epileptic cortex*. In: Measuring chaos in human brain. D. Duke and W. Pritchards (Eds). World Scientific, Singapore, pp. 49–89, 1991.

- [60] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. Williams. *Modelling of ECoG in temporal lobe epilepsy*. 25th Ann. Rocky Mountain Bioing. Symposium, pp. 1201-1203, 1988.
- [61] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri and W. J. Williams. Brain Topography 2 (1990) 187.
- [62] A. Jackson. *Perspectives of Nonlinear Dynamics*, Vols. 1 and 2. Cambridge University Press, Cambridge, England, 1991.
- [63] J. Jackson and J. Hughlings. Br. Med. J. 1 (1890) 765.
- [64] E. T. Jaynes. *Statistical Physics*. W. K. Ford, Ed. Benjamin, New York, 1963.
- [65] D. Johnston and T. H. Brown. *Mechanisms of neural burst generation*. In P. A. Schwartzkroin and H. V. Wheal (Eds). *Electrophysiology of epilepsy*. Academic Press, London, P. 277, 1984.
- [66] H. Kantz and T. Schereiber, *Nonlinear time series analysis*. Cambridge University Press, Cambridge, 1997.
- [67] A. Katz. *Principles of statistical mechanics*. Freeman, San Francisco, 1967.
- [68] J. Klein, *Statistical visions in time: a history of time series analysis*. Cambridge, England: Cambridge University Press, 1997.
- [69] A. Kolmogorov. Problems of Information Transmission 1 (1965) 1.
- [70] A. Kowalski, M. T. Martin, A. Plastino, A. Proto and O. A. Rosso. Physics Letters A, 311 (2003) 180.
- [71] P. Kronheimer. Lecture Notes in Pure and Appl. Math. 272 (2000).
- [72] S. Kullback and R. A. Leibler, Ann. Math. Stat. 22 (1951) 79.
- [73] P. W. Lamberti, M. T. Martín, A. Plastino and O. A. Rosso, Physica A 334 (2003) 119.
- [74] K. Lehnertz and C. E. Elger. Phys. Rev. Lett. 80 (1998) 5019
- [75] S. Lloyd and H. Pagels. Ann. Phys. 213 (1988) 188.
- [76] F. H. Lopes da Silva. Electr. Clin. Neurophysiol. 79 (1991) 81.
- [77] R. Lopez-Ruiz, H. L. Mancini and X. Calbet. Phys. Lett. A 209 (1995) 321.
- [78] E. Lorenz. J. Atmos. Sci. 20 (1963) 130.
- [79] S. Mallat. *Multiresolution representations and wavelets*. Grasp. Lab. 153, Univ. of Pennsylvania, Philadelphia, 1998.
- [80] S. Mallat, *A Wavelet tour of signal processing*. Academic Press, second edition, San Diego, 1999.

- [81] M. T. Martín, F. Pennini and A. Plastino. Physics Letters A 256 (1999) 173.
- [82] M. T. Martín, J. Perez and A. Plastino. Physica A 291 (2001) 523.
- [83] M. T. Martín, A. R. Plastino, A. Plastino. Physica A 275 (2000) 262.
- [84] M. T. Martín, A. Plastino, O. A. Rosso. *Self-organization in neural matter*. Proceedings of Workshop on Condensed Matter Theories, Sidney, Australia, 2001. Vol. 17 Condensed Matter Theories, Nova Science Publishers, in press.
- [85] M. T. Martín, A. Plastino and O. A. Rosso, Phys. Lett. A, 126 (2003) 311.
- [86] H. Matsumoto and C. Ajmone-Marsan. Exp. Neurol. 9 (1964) 305 .
- [87] Y. Meyer. *Ondelettes* , Hermann Ed., Paris, France, 1990.
- [88] Y. Meyer. *Les Ondelettes. Algorithmes et applicattions*. Armand Colin, 1992.
- [89] O. Michel, R. Baraniuk and P. Flandrin. Time-frequency based distance and divergence measures. In: *IEEE Int. Time-Freq. and Time-Scale Symp.* Philadelphia, PA, 1994.
- [90] F. Mitschake, Phys. Rev. A 41 (1990) 1169.
- [91] F. Mitschake, M. Moller and W. Large. Phys. Rev. A 37 (1988) 4518.
- [92] E. Niedermeyer and F. H. Lopes da Silva (Eds). *Electroencephalography, Basic Principles, Clinical Applications, and Related Field*. Urban and Schwarzenberg, Baltimore, 1987.
- [93] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge, England, 1993.
- [94] E. Ott, T. Sauer and J. A. Yorke. *Coping with chaos*. John Wiley & Sons Inc., New York, 1994.
- [95] G. Parisi. Phys. World 6 (1993) 9.
- [96] T. A. Pedley, R. D. Traub. *Physiological basis on EEG*. In: D. D. Dalv and T. A. Pedley (Eds). Raven Press Ltd.,New York, 2nd ed., p. 107, 1990.
- [97] R. Piasecki, M. T. Martín and A. Plastino. Physica A (2001), in press.
- [98] J. P. Pijn, J. van Nerven, A. Noest and F. H. Lopes da Silva. Electr. and Clin. Neurophysiol. 79 (1991) 371.
- [99] J. P. Pijn, D. Velis, M. van der Heiden, J. de Goede, C. van Veelen and F. H. Lopes da Silva. Brain Topography, 9 (1997) 249.
- [100] G. E. Powell and I. C. Percival. J. Phys. A: Math. Gen. 12 (1979) 2053.
- [101] W. S. Pritchard, D. W.Duke and K. Krieble. Psycophysiology 32 (1995) 486.

- [102] A. Renyi. *On measures of entropy and information*. In: Fourth Berkeley Symposium on Math. Stat. and Prob., Vol. 1, p. 547, 1970.
- [103] S. A. Rombouts, R. W. Keunen and C. J. Stam. Phys. Lett. A 202 (1995) 352.
- [104] O. E. Rossler. Ann. New York Acad. Sci. 316 (1979) 376.
- [105] O. A. Rosso, S. Blanco and A. Rabinowicz. Signal Processing 86 (2003) 1275.
- [106] O. A. Rosso and M. L. Mairal. Physica A 312 (2002) 469.
- [107] O. A. Rosso, M. T. Martín, and A. Plastino. Physica A 313 (2002) 587 .
- [108] O. A. Rosso, M. T. Martín, and A. Plastino. Physica A 320 (2003) 497.
- [109] O. A. Rosso, M. T. Martín, A. Plastino. *Generalized information measures and the analysis of brain electrical signals*. Nonextensive Entropy- Interdisciplinary applications. Edited by: M. Gell-Mann and C. Tsallis. Oxford University Press, Oxford, 2003, in press.
- [110] O. A. Rosso, M. T. Martín, A. Plastino. *A measure of self-organization in neural activity*. Instabilities and Non-equilibrium Structures IX Edited by: O. Descalzi, J. Martinez Mardones, S. Ricca and E. Tirapegui. Kluwer Academic Publishes, Dordrech , 2004, in press.
- [111] V. J. Samar, A. Bopardikar, R. Rao and K. Swartz. Brain and Language 66 (1999) 7.
- [112] H. G. Schuster. *Deterministics Chaos: An introduction*. Physic Verlag, Weinheim, 1988.
- [113] E. Serrano. Tesis Doctoral, FCEyN, Universidad de Buenos Aires, 1996.
- [114] E. P. Serrano. Escuela de Ciencia y Tecnología. Univ. Nac. de General San Martin.
- [115] C. E. Shannon, Bell System Technol. J. 27 (1948) 379.
- [116] J. S. Shiner, M. Davison and P. T. Landsberg. Phys. Rev. E 59 (1999) 1459.
- [117] J. S. Shiner, M. Davison and P. T. Landsberg. Phys. Rev. E 62 (2000) 3000 .
- [118] R. Silver. *Physics and probability. Essays in honor of Edwin T. Jaynes*. Edited by W. T. Grandy, Jr. and P. W. Milonni. Cambridge University Press, New York, 1993.
- [119] R. Sole and B. Luque. *Statistical measures of complexity for Strongly Interacting Systems*. arXiv: adap-org/9909002 v1, 1999.
- [120] R. J. Solomonoff. Inform. Contr. (1964) 1, 224.
- [121] E. J. Speckermann and C. E. Elger. *Introduction to the neurophysiological basis of the EEG and DC potentials*. In: E. Niedermeyer and F. Lopes da Siva (Eds). Williams and Wilkins, Baltimore, 3rd ed., p. 15, 1993.

- [122] C. J. Stam, T. C. A. M. van Woerkon and W. S. Pritchard. Electr. and Clin. Neurophysiol. 99 (1996) 214.
- [123] R. K. Standish. arXiv:nlin.AO/0101006 v1 2 Jan 2001.
- [124] M. Steriade, P. Gloor, R. R. Llinas, F. Lopes da Silva and M. M. Mesulam. Electr. and Clin. Neurophysiol. 76 (1990) 481.
- [125] J. Theiler and P. E. Rapp. Electr. and Clin. Neurophysiol. 98 (1996) 213.
- [126] C. Tsallis. J. Stat.Phys. 52 (1988) 479
- [127] M. E. Torres. Tesis Doctoral, FCEIyA, Universidad Nacional de Rosario, 1998.
- [128] M. E. Torres, L. Gamero and E. D'Atellis. *A multiresolution entropy approach to detect epileptic form activity in the EEG*. In: I. Pitas (Ed), IEEE Workshop on nonlinear signal and image processing, v.II, p.791, 1995.
- [129] M. E. Torres, L. Gamero and E. D'Atellis. Latin American Applied Research, 53 (1995) 57.
- [130] M. E. Torres, L. Gamero and E. D'Atellis. Technical Report 2812, INRIA. Rapport de recherche, 1996.
- [131] J. Uffink. Tesis Doctoral, Universidad de Utrecht, 1990.
- [132] M. Unser, A. Aldrubi and M. Eden. IEEE Trans. on Signal Processing 41 (1993) 821.
- [133] J. Ville. Cables et Transmission 2A (1948).
- [134] P. Walters. *Ergodic theory: Introductory lectures*. Lecture Notes in Mathematics 458. Springer Verlag, New York, 1975.
- [135] B. J. West (Ed). *Fractal physiology and chaos in medicine*. Nonlinear Phenomena in Life Science Vol. I. World Scientific, Singapore, 1993.
- [136] E. P. Wigner. Phys. Rev. 40 (1932) 749.
- [137] A. Wolf, J .B. Swift and H. L. Swinney. Physica D 16 (1985) 285.
- [138] W. K. Wootters. Phys. Rev. D 23 (1981) 357.