INSTITUTO TECNOLÓGICO DE CELAYA

DEPARTAMENTO DE INGENIERÍA QUÍMICA

" IDENTIFICACIÓN Y CARACTERIZACIÓN DE SISTEMAS DINÁMICOS MEDIANTE REDES NEURONALES ARTIFICIALES "

P O R

RAÚL GONZÁLEZ GARCÍA

TESIS PRESENTADA AL DEPARTAMENTO DE INGENIERÍA QUÍMICA COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS

CELAYA, GTO. AGOSTO 1999

INSTITUTO TECNOLOGICO de celaya

I.Q./013/99

Asunto: Aprobación de Impresión
de Tesis de Doctorado.

Agosto 17 de 1999.

M.C. J. Guadalupe Michel Nava
Jefe del Depto. de Ingeniería Química.
P r e s e n t e

De acuerdo con la convocatoria hecha por esa Jefatura a su digno cargo, con el fin de determinar la aprobación de impresión del Proyecto de Tesis de Doctorado, titulado: ***"IDENTIFICACIÓN Y CARACTERIZACIÓN DE SISTEMAS DINÁMICOS MEDIANTE REDES NEURONALES ARTIFICIALES"*** presentado por él(la) **M.C. RAÚL GONZÁLEZ GARCÍA,** comunico a usted que esta Comisión Revisora de tesis ha tenido a bien aprobar la impresión de este trabajo.

Lo anterior, para los efectos consiguientes.

Atentamente,

Dr. Ramiro Rico Martínez
Presidente

Dr. Pedro A. Quintana Hernández
Secretario

Dr. Juan Francisco Javier Alvarado
Vocal

Dra. Gloria Ma. Martínez González
Suplente

Dr. Enrique Arce Medina
Revisor Externo

ssh*

Av. Tecnológico y A. García Cubas S/N
A.P. 57, C.P. 38010, Celaya, Gto.,
Correo Electrónico: lince@itc.mx
http:\\www.itc.mx
Conmutador 01 461 1 7575
Fax 01 461 1 7979

Dedicatoria

Quiero dedicar este trabajo de Tesis a DIOS por darme la oportunidad de llegar a este momento, y quien me ha iluminado en todo momento mi vida.

A esa gran mujer ... mi MADRE, por todo el amor y apoyo incondicionales que me ha brindado siempre y por sus ejemplos de lucha y tenacidad.

A mi PADRE que donde quiera que este se sentirá orgulloso de mi.

A mis hemanos por la confianza que han depositado en mi y por su apoyo incondicional.

A mi ESPOSA e HIJOS por darme más fuerzas cada día para seguir adelante.

Agradecimientos

Agradezco de todo corazón al Dr. Ramiro Rico Martínez por haberme brindado su amistad y por su participación como asesor del presente trabajo.

A los Doctores Pedro A. Quintana Hernández, Francisco Javier Alvarado, Gloria María Martínez González y al Dr. Enrique Arce Medina por haber participado amablemente en el comité de revisión del presente trabajo.

A todos mis compañeros y amigos.

A Lety por sus valiosos comentarios en la redacción de la presente.

A todo el grupo de secretarias del departamento de Ingeniería Química del ITC.

Doy gracias al CONACyT y al CONCYTEG por el apoyo económico brindado durante mis estudios de Doctorado.

Agradezco la ayuda económica proporcionada a mi asesor el Dr. Ramiro Rico Martínez por parte del CONACyT y la NSF para la realización del presente proyecto de investigación.

IDENTIFICACIÓN Y CARACTERIZACIÓN DE SISTEMAS DINÁMICOS MEDIANTE REDES NEURONALES ARTIFICIALES

Raúl González García

Doctorado en Ciencias en Ingeniería Química

Resumen

En el presente trabajo se utilizan identificadores, basados en redes neuronales artificiales, para caracterizar el comportamiento complejo que se presenta en sistemas dinámicos no-lineales continuos en el tiempo. La caracterización enfatiza la predicción de soluciones a largo plazo, para un entendimiento global del comportamiento presente en las observaciones.

En primer lugar se aplicaron dos diferentes metodologías de identificación para caracterizar un escenario de bifurcaciones en dos parámetros. Los datos para la identificación se tomaron de experimentos de la oxidación galvanostática de hidrógeno, y consistieron de mediciones de una sola variable de estado (potencial) a diferentes condiciones de corriente y concentración de iones cloro. Bajo esas condiciones, el sistema presenta comportamiento dinámico altamente no-lineal del tipo denominado oscilaciones de modo mixto. La primera de las metodologías, consistió en discretizar los datos experimentales para construir un modelo dinámico discreto unidimensional. La segunda, consistió en procesar las series originales a través de una red neuronal para extraer sus componentes principales no lineales. Con estos componentes principales

se construyó un modelo dinámico continuo tridimensional. En ambos casos, después de construir los modelos, se caracterizó el comportamiento dinámico en escenarios de bifurcaciones de dos parámetros.

También se desarrolló una metodología, basada en redes neuronales artificiales, para la identificación de sistemas con variaciones espacio-temporales. El identificador se obtuvo al acoplar una red neuronal artificial con esquemas de discretización espacial (diferencias finitas) para después incrustarlo en un esquema de integración numérica. Se presentó el desarrollo matemático del identificador y de sus derivadas para el caso del integrador de Runge-Kutta. Este procedimiento es la primera vez que se aplica con redes neuronales. La metodología fue probada con datos simulados de una ecuación diferencial parcial unidimensional y unidireccional, en condiciones donde se presenta comportamiento dinámico del tipo denominado ondas viajeras con multiplicidad de soluciones.

Dirigida por:

Dr. Ramiro Rico Martínez

Índice General

Resumen i

Índice de Figuras vi

Índice de Tablas x

Glosario de Abreviaturas xi

1 Introducción 1

1.1 Objetivo 3

1.2 Organización del trabajo 3

2 Antecedentes 5

2.1 Sistemas Dinámicos 5

2.1.1 Solución de un SD 6

2.1.2 Reducción de un sistema de tiempo continuo a uno de tiempo discreto 11

2.1.3 Caracterización de Sistemas: Diagramas de Bifurcaciones . . . 15

2.1.4 Análisis de OMMs 17

2.2 Redes Neuronales Artificiales 19
2.2.1 Reseña histórica 20
2.2.2 Funcionamiento de las RNAs 24
2.2.3 Arquitecturas 28
2.2.4 Entrenamiento 28

3 Teoría de Identificación con RNAs **31**
3.1 SDs discretos (mapas) 34
3.1.1 Construcción de la Función Error 34
3.1.2 Método Delta para el cálculo de derivadas en RNAs 36
3.2 SDs continuos (EDOs) 39
3.2.1 Construcción de la Función Error 41
3.2.2 Cálculo de derivadas 43
3.3 SDs espacio-temporales 44
3.3.1 Construcción de la Función Error 45
3.3.2 Cálculo de derivadas 49
3.4 Gradiente Conjugado 50
3.5 Validación Cruzada 52

4 Caracterización de OMMs **54**
4.1 Construcción de un modelo discreto 56
4.1.1 Procesamiento de datos experimentales (extracción de MPMs) 56
4.1.2 Modelo de identificación 58
4.1.3 Análisis de resultados 62
4.2 Construcción de un modelo continuo 68

4.2.1 Reconstrucción del espacio-fase (obtención de CPNLs) 68
4.2.2 Modelo de Identificación 74
4.2.3 Análisis de resultados 76

5 Caracterización de una EDP **82**
5.1 RNA para la identificación 83
5.2 Análisis de resultados 85

6 Discusión y Conclusiones **95**

Bibliografía **100**

A Oxidación galvanostática de H_2 **109**
A.1 Arreglo Experimental 110
A.2 Observaciones Experimentales 113

B Ecuación de Kuramoto-Sivashinsky **116**

C Integración y Derivación Numérica **121**
C.1 Integración Numérica 121
C.2 Diferencias Finitas 122

Índice de Figuras

2.1 Ilustración gráfica de Series de tiempo y Espacio-Fase. 7

2.2 Algunos comportamientos que se pueden encontrar en el estudio de sistemas dinámicos. 9

2.3 Comportamiento Oscilatorio. 12

2.4 Construcción de los Mapas de Poincaré. 14

2.5 Construcción de los Mapas de Próximo Mínimo. 14

2.6 Diagrama de Bifurcación en un parámetro. 16

2.7 Estructura básica de una red neuronal artificial. 25

2.8 Secuencia esquematizada de la activación de neuronas para producir su salida. 26

2.9 Dos funciones de activación comunes en RNAs. 27

2.10 Dos arquitecturas básicas de RNAs. 29

3.1 Esquema de RNA que puede ser usada en la identificación de mapas. 34

3.2 Esquema de RNA que puede ser usada en la identificación de EDOs. . 40

3.3 Macrored resultante al incrustar una RNA-EDO en el esquema de integración de Runge-Kutta. 42

3.4 Esquema de RNA que puede ser usada en la identificación de un proceso espacio-temporal unidimensional y unidireccional. 45
3.5 Discretización de la dimensión espacial. 46
3.6 RNA-EDP acoplada con un esquema de diferencias finitas. 47
3.7 Macrored resultante al incrustar una RNA-EDP en el esquema de integración de Runge-Kutta. 48

4.1 MPMs obtenidos a partir de series de tiempo experimentales de la oxidación de hidrógeno. 57
4.2 Estructura de la RNA que fue utilizada en la construcción de un MPM: $Em_{t+1} = F(Em_t; Cl, I)$. 59
4.3 Comparación de MPMs experimentales contra los predichos por la RNA-MPM. 64
4.4 Diagramas generales de bifurcaciones en dos parámetros (I contra Cl) predichos por la RNA-MPM. 65
4.5 Estructura de RNA utilizada para construir el espacio-fase en CPNLs. 70
4.6 Algunas series de tiempo experimentales representadas en el espacio-fase de los CPNLs. 73
4.7 Estructura de RNA utilizada para aproximar un conjunto de EDOs con datos de CPNLs. 75
4.8 Comparación de algunas series de tiempo experimentales contra las predichas (a largo plazo) por la RNA-EDO. 77
4.9 Diagrama general de bifurcaciones en un parámetro (I) predicho por la RNA-EDO. 78

4.10 Diagrama general de bifurcaciones en dos parámetros (I contra Cl) predicho por la RNA-EDO. 81

5.1 Estructura de RNA utilizada en la identificación del proceso espacio-temporal exhibido por la Ecuación de Kuramoto-Sivashinsky. 84

5.2 Comparación de los atractores reales contra los predichos por la RNA-KS. 86

5.3 Comparación de las series de tiempo reales contra las predichas por la RNA-KS. 87

5.4 Comparación de ondas viajeras reales contra las predichas por la RNA-KS en $t = \frac{1}{2}T$ y $t = T$. 89

5.5 Errores porcentuales a largo plazo (a través del tiempo) entre las series reales y las predichas por la RNA-KS. 91

5.6 Errores porcentuales a largo plazo entre las ondas reales y las predichas por la RNA-KS. 92

5.7 Errores porcentuales a corto plazo entre las series de tiempo reales y las predichas por la RNA-KS. 94

A.1 Esquema del arreglo experimental de la oxidación galvanostática de H_2 sobre electrodos de Pt. 111

A.2 Escenario de las corridas experimentales para una concentración de iones bismuto de 1×10^{-4} M. 112

A.3 Series de tiempo experimentales típicas de la oxidación de H_2 sobre electrodos de Pt, recolectadas a una concentración de iones bismuto de 1×10^{-4} M. 114

B.1 Soluciones a largo plazo de la EKS, proyectadas sobre dos de las mediciones del dominio espacial, para un valor del parámetro $\alpha = 15$. . . . 118

B.2 Evolución de las ondas viajeras de la EKS en $\alpha = 15$ para diferentes instantes en el tiempo. 119

Índice de Tablas

4.1 Información de los datos experimentales usados para construir el conjunto de entrenamiento de la RNA-MPM. 61

4.2 Información de los datos experimentales usados para construir el conjunto de validación de la RNA-MPM. 61

4.3 Información de los datos experimentales usados para construir el conjunto de entrenamiento de la RNA-CPNL. 71

Glosario de Abreviaturas

CPL ..... Componente Principal Lineal

CPNL ..... Componente Principal No-Lineal

EDO ..... Ecuación Diferencial Ordinaria

EDP ..... Ecuación Diferencial Parcial

EKS ..... Ecuación de Kuramoto-Sivashinsky

KS ..... Kuramoto-Sivashinsky

MP ..... Mapa de Poincaré

MPM ..... Mapa de Próximo Mínimo (o Máximo)

OMM ..... Oscilación de Modo Mixto

RNA ..... Red Neuronal Artificial

SD ..... Sistema Dinámico

SPA ..... Sistema de Parámetros Agrupados

SPD ..... Sistema de Parámetros Distribuidos

1. Introducción

Los sistemas físico-químicos que se encuentran lejos del equilibrio (termodinámico, mecánico, etc.) exhiben comportamientos que dependen de la evolución del tiempo. Este tipo de sistemas se denominan sistemas dinámicos. Con frecuencia se encuentra que dichos sistemas son altamente no-lineales por lo que pueden exhibir comportamientos muy complicados. Dichos comportamientos van desde estructuras regulares hasta aquellas que aparentemente carecen de regularidad en el dominio espacio-temporal. En teoría, todos los procesos reales involucran variaciones espacio-temporales (en mayor o menor escala). Existe gran variedad de procesos en la literatura de ingeniería química, que exhiben estas variaciones debido a su carácter inherentemente no-lineal. Como ejemplo se pueden mencionar los siguientes: reacciones electroquímicas en electrodos rotativos, procesos catalíticos, reactores tubulares, reactores de lecho empacado, operaciones de secado, etc.

Los procesos con variaciones espacio-temporales son conocidos como sistemas de parámetros distribuidos y pueden ser modelados por medio de ecuaciones diferenciales parciales, mientras que aquellos que sólo experimentan variaciones temporales se conocen como sistemas de parámetros agrupados y serán modelados con ecuaciones diferenciales ordinarias. Sin embargo, cuando la respuesta espacial es despreciable, las

aproximaciones tradicionales para la modelación de sistemas de parámetros distribuidos involucran simplificaciones que desatienden la respuesta espacial, centrándose sólo en la respuesta temporal, dando como resultado sistemas de parámetros agrupados.

En sistemas que presentan comportamiento dinámico suave o moderado, las aproximaciones basadas en principios básicos (balances de materia y energía, etc.) son generalmente exitosas en la reproducción de las observaciones experimentales. Sin embargo, muchos de los sistemas reales presentan comportamiento dinámico complejo aún cuando se permita despreciar la dependencia espacial. Estos comportamientos, en ocasiones, son manifestados en forma de oscilaciones de modo mixto, ondas viajeras o como comportamiento caótico. Las oscilaciones de modo mixto son oscilaciones alternantes (regulares) de pequeña y gran amplitud; las ondas viajeras se presentan como comportamiento oscilatorio espacio-temporal; y el comportamiento caótico es caracterizado por una aparente falta de regularidad en su respuesta dinámica. Para estos sistemas con dinámica compleja, la modelación basada en principios fundamentales puede resultar tediosa, complicada, ofrecer muy pobres capacidades predictivas y en la mayoría de las veces no se dispone de soluciones exactas. Es por eso que surge la necesidad de implementación de nuevas técnicas para la identificación de tales sistemas.

En ausencia de modelos confiables y tratables, basados en principios básicos, los modelos empíricos son los sustitutos ideales en la simulación. Dentro de este género, en los últimos años, han destacado los modelos basados en redes neuronales artificiales. Dichos modelos han llegado a ser una herramienta de gran valor para el procesamiento de series de tiempo y la identificación de sistemas, siendo poderosas herramientas para el entendimiento cualitativo del comportamiento dinámico y de la dependencia de los

parámetros de operación. Estas aproximaciones están limitadas en el sentido de que son modelos de caja negra para los cuales el conocimiento parcial del sistema físico fundamental no puede ser incorporado.

1.1 Objetivo

El objetivo general del trabajo es mostrar la capacidad que tienen los modelos empíricos, basados en redes neuronales artificiales, para capturar el comportamiento de largo plazo (atractores), tanto en el espacio-fase como en el espacio extendido de parámetros, en los sistemas que presentan comportamiento dinámico complejo.

El objetivo general es alcanzado con los siguientes objetivos particulares:

1. Caracterizar el escenario de bifurcaciones, en dos parámetros, presente en los datos experimentales de la oxidación galvanóstática de hidrógeno.

2. Caracterizar el escenario de soluciones múltiples, en el espacio-fase (no hay parámetros), de un proceso espacio-temporal simulado numéricamente.

1.2 Organización del trabajo

Se presentan tres alternativas para la identificación de los sistemas dinámicos. La primera de ellas es un modelo discreto para la identificación de un proceso continuo real de parámetros agrupados que presenta comportamiento del denominado oscilaciones de modo mixto. El modelo construido es apropiado para caracterizar el escenario de transiciones en dos parámetros del proceso real, aunque presenta algunas limitaciones. La segunda alternativa es un modelo continuo (conjunto de ecuaciones

diferenciales ordinarias) utilizado para la identificación del mismo sistema real de la alternativa anterior. Esta alternativa requiere procesar la informacion del sistema real mediante una red neuronal para extraer las principales características dinámicas por descomposición en componentes principales no-lineales, y después desarrollar una metodología de redes neuronales incrustadas en un esquema de integración numérica para lograr la identificación. Al aplicar estos dos procedimientos de caracterización independientes, al mismo proceso, puede dar una base para corroborar los resultados. La tercera de las alternativas es un modelo continuo con variaciones espacio-temporales. Este modelo es utilizado para la identificación de un proceso simulado por una ecuación diferencial parcial (la Ecuación de Kuramoto-Sivashinsky). Esta alternativa involucra desarrollar una metodología de redes neuronales acopladas con técnicas de discretización espacial y esquemas de integración numérica.

En el Capítulo 2 se presenta una breve introducción a los principios que sustentan el estudio de sistemas dinámicos y las redes neuronales artificiales. En el Capítulo 3 se desarrolla la teoría aplicada en la identificación de sistemas dinámicos con el uso de las redes neuronales artificiales. En el Capítulo 4 se muestran los resultados obtenidos en la identificación y caracterización de un sistema real que presenta comportamiento del denominado oscilaciones de modo mixto. La identificación se presenta en dos formas diferentes: como mapas de próximo mínimo y como un conjunto de ecuaciones diferenciales ordinarias. En el Capítulo 5 se presentan los resultados en la identificación de un sistema continuo simulado que presenta comportamiento espacio-temporal del tipo denominado ondas viajeras. Finalmente, en el Capítulo 6 se presentan algunas conclusiones y se ofrece una perspectiva del posible trabajo futuro relacionado con los tópicos descritos.

2. Antecedentes

2.1 Sistemas Dinámicos

Aquí se presentan algunas definiciones, en un sentido cualitativo, de la teoría de sistemas dinámicos. Demostraciones rigurosas de tales definiciones se pueden encontrar en la siguiente literatura [64, 15].

Un *sistema dinámico* (SD) puede ser definido como aquel en el cual una o varias de sus variables dependen, en su respuesta, de la evolución del tiempo. Dicho sistema puede ser representado de dos formas. La primera de ellas es como un *sistema dinámico discreto* o *mapa*:

$$\mathbf{U}_{t+1} = \mathbf{g}(\mathbf{U}_t; \boldsymbol{\mu}) \tag{2.1}$$

y la segunda forma es como un *sistema dinámico continuo* o conjunto de *Ecuaciones Diferenciales Ordinarias* (EDOs):

$$\frac{d\mathbf{U}}{dt} = \mathbf{f}(\mathbf{U}; \boldsymbol{\mu}) \tag{2.2}$$

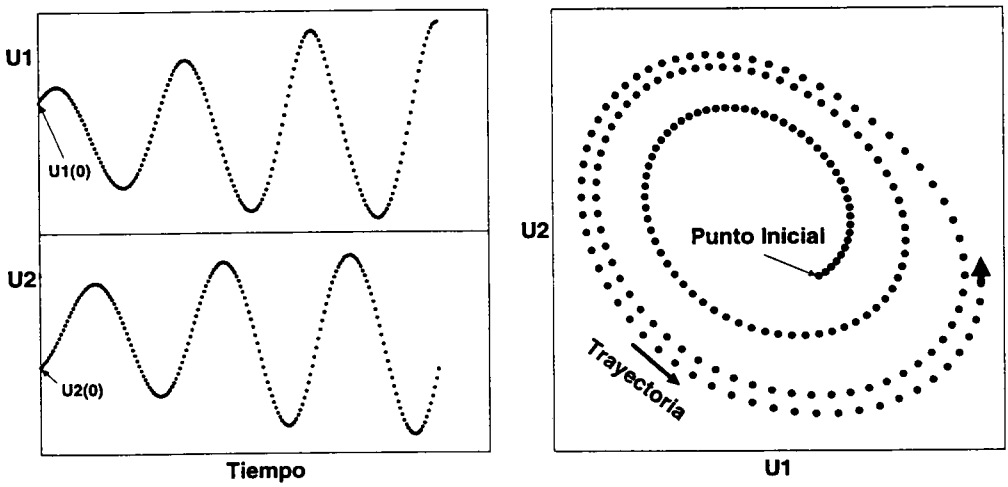
en estos sistemas $\mathbf{U}$ representa al vector de *variables dependientes o variables de estado*, $\boldsymbol{\mu}$ representa un vector de *parámetros* del sistema y t representa a la *variable*

independiente que en la Ecuación 2.1 se le conoce como *iteración* y en la Ecuación 2.2 se le da el nombre de *tiempo*. Las funciones del lado derecho [$\mathbf{g}(\cdot)$ y $\mathbf{f}(\cdot)$] son, en general, funciones vectoriales no-lineales de las variables involucradas. Si $\mathbf{g}(\cdot)$ y $\mathbf{f}(\cdot)$ no dependen explícitamente de t, a los SDs que representan se les conoce como *Sistemas Autónomos*.

2.1.1 Solución de un SD

Para resolver un sistema dinámico (discreto o continuo) se requiere de conocer la forma de la función del lado derecho [$\mathbf{g}(\cdot)$ o $\mathbf{f}(\cdot)$], un vector de valores de parámetros ($\boldsymbol{\mu}$) y un vector de valores iniciales de la variable de estado [$\mathbf{U}_{t=0} = \mathbf{U}_0$ o $\mathbf{U}(t=0) = \mathbf{U}(0)$].

Resolver el sistema de la Ecuación 2.1 involucra evaluar $\mathbf{g}(\cdot)$ y asignar el resultado a la variable $\mathbf{U}_{t+1}$ ($\mathbf{U}_1$); después el valor de $\mathbf{U}_1$ se toma para evaluar nuevamente la función $\mathbf{g}(\cdot)$ y el resultado se asigna a $\mathbf{U}_2$. Este procedimiento se repite cuantas veces sea necesario hasta completar un número determinado de vectores de estado ($\mathbf{U}_0, \mathbf{U}_1, \mathbf{U}_2, \ldots$). Sin embargo, resolver un sistema continuo (Ecuación 2.2) no resulta tan directo como el resolver un sistema discreto (Ecuación 2.1). La solución de los sistemas continuos involucra incrustarlos dentro de un esquema de integración. En los sistemas continuos simples (una o a lo sumo dos variables dependientes y función $\mathbf{f}(\cdot)$ lineal) se puede disponer de una solución analítica de la forma: $\mathbf{U} = \mathbf{U}(\mathbf{U}(0), \boldsymbol{\mu}, t)$, pero muchos de los sistemas continuos prácticos son en realidad multivariables y no-lineales, por lo cual requieren para su solución de técnicas de integración numérica. Las técnicas de integración numérica son, en general, una transformación de un sistema de tiempo continuo a un sistema de tiempo discreto, el cual puede ser iterado de la misma forma que se hace con el mapa. Dicha transformación es sólo un mecanismo



(a) Series de Tiempo (b) Espacio-Fase en dos dimensiones

Figura 2.1: *Ilustración gráfica de Series de tiempo y Espacio-Fase.*

para resolver el sistema continuo, la cual (aunque discreta) sigue preservando todas las propiedades dinámicas del sistema continuo.

Durante la resolución de un sistema dinámico, sus parámetros (μ) no cambian. Sin embargo, son de mucha importancia cuando se piensa en caracterizar los diferentes comportamientos del sistema (para esto dirigirse a la Sección 2.1.3).

El conjunto de valores que se generan al resolver las Ecuaciones 2.1 ó 2.2, generalmente se le conoce como *serie de tiempo*, debido a que se puede representar en una gráfica **U** contra t. Sin embargo, en el estudio de sistemas dinámicos es más conveniente representar estos datos en una gráfica que contenga únicamente variables de estado a la cual se le denomina diagrama (o retrato) *espacio-fase*. Los diagramas espacio-fase son una manera clara y compacta de representar la dinámica de un sistema, por lo cual serán los diagramas de elección en el presente trabajo. Estos conceptos son mostrados en la Figura 2.1.

La solución de un sistema dinámico, representada en un diagrama de espacio-fase, forma curvas que se denominan *trayectorias.* Si el sistema dinámico es iterado infinitamente (el tiempo tiende a infinito), la trayectoria formada puede llegar a converger hacia un punto, el cual se conoce como *punto fijo* o *estado estacionario.* Los puntos fijos de un sistema dinámico pueden calcularse a partir de las Ecuaciones 2.1 ó 2.2 resolviendo para $\mathbf{U}$ las siguientes ecuaciones algebraicas:

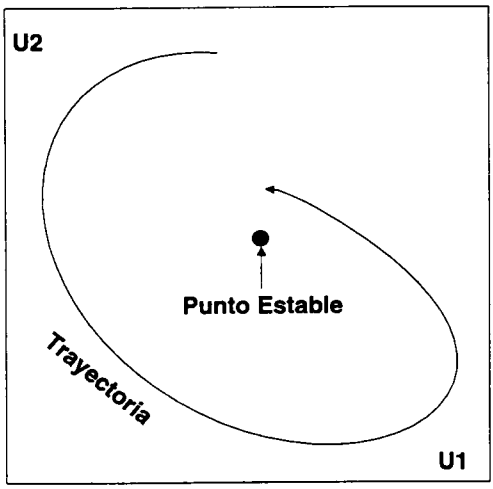
$$\widehat{\mathbf{U}} - \mathbf{g}(\widehat{\mathbf{U}}; \boldsymbol{\mu}) = 0 \tag{2.3}$$

ó

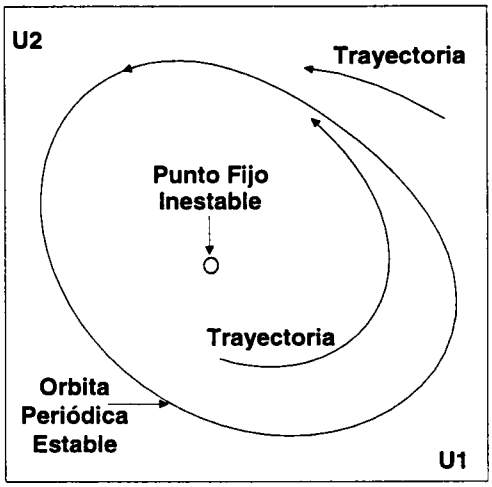
$$\mathbf{f}(\widehat{\mathbf{U}}; \boldsymbol{\mu}) = 0 \tag{2.4}$$

donde $\widehat{\mathbf{U}}$ representa el vector de cordenadas del punto fijo. Sin embargo, cuando un SD se itera infinitamente (hasta un número de iteraciones muy grande), no solamente se puede converger hacia puntos fijos, sino que también se puede converger a otras formas geométricas más complejas, por ejemplo trayectorias cerradas *(órbitas periódicas)* o *trayectorias caóticas.* Las soluciones en las que se involucra iterar el sistema hasta el infinito (suponiendo que existe el infinito), también se denominan soluciones a *largo plazo*, mientras que las soluciones en las que se involucra un lapso de tiempo corto (una pocas iteraciones) se les denomina soluciones a *corto plazo.* En la Figura 2.2 se presentan algunos de estos conceptos.

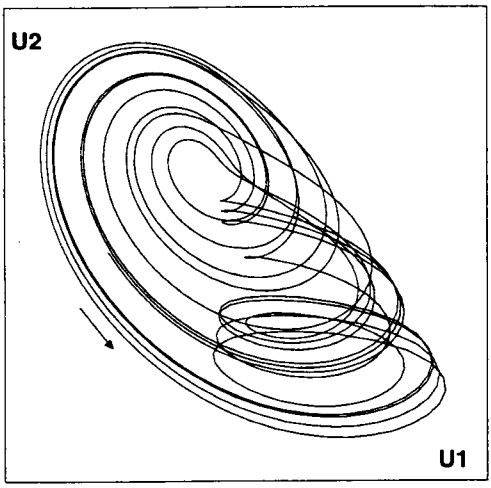
Las soluciones a largo plazo pueden ser clasificadas en dos categorias: las *estables* y las *inestables.* Un punto fijo es estable si el SD, comenzando cerca del punto, se itera hacia el *infinito positivo* y su trayectoria converge hacia dicho punto fijo. Por otro lado, el punto fijo es inestable si el SD, comenzando cerca del punto, se



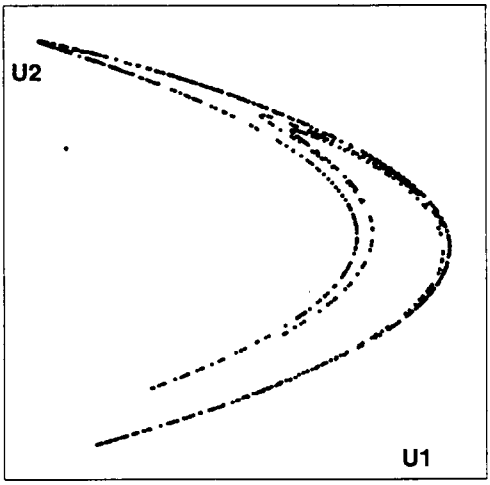
(a) Un Punto Fijo Estable.



(b) Una Orbita Periódica Estable y un Punto Fijo Inestable.



(c) Un Atractor Caótico de un Sistema Continuo.



(d) Un Atractor Caótico de un Sistema Discreto.

Figura 2.2: *Algunos comportamientos que se pueden encontrar en el estudio de sistemas dinámicos.*

itera hacia el *infinito positivo* y la trayectoria se aleja de dicho punto fijo[1]. Este mismo criterio puede ser aplicado a las órbitas periódicas para determinar su estabilidad. Metodologías más estrictas para la detección de estabilidad en soluciones a largo plazo se pueden encontrar en [64, 15]. En un SD pueden coexistir a la vez diferentes soluciones a largo plazo, es decir, combinaciones de puntos fijos (estables y/o inestables) con soluciones periódicas (estables y/o inestables) y soluciones *caóticas* (caracterizadas por su aparente falta de regularidad, Figuras 2.2c y d). Cabe mencionar que en los sistemas dinámicos reales sólo es posible detectar las soluciones estables, debido a que en dichos sistemas el tiempo siempre evoluciona hacia adelante. En el campo de SDs, a las soluciones estables a largo plazo también se les denomina *atractores.*

El comportamiento *oscilatorio* (Figura 2.3) es la representación del comportamiento periódico en diagramas de series de tiempo. Una manifestación especial de este comportamiento, son las llamadas *Oscilaciones de Modo Mixto* (OMMs). Las OMMs son formas de ondas que consisten de excursiones alternantes de pequeña y gran amplitud, y su comportamiento en el espacio-fase se manifiesta como órbitas de periodo múltiple (Figura 2.3). En teoría de sistemas, este comportamiento se ha clasificado como de alta complejidad. En sistemas de tiempo continuo el comportamiento de OMMs (ver la Figura 2.3) se presentará cuando el sistema sea no-lineal y la dimensión del espacio-fase sea al menos tres, además de que existan las condiciones necesarias para el comportamiento periódico. En el caso de los sistemas discretos (mapas), el comportamiento de OMMs se puede presentar para cualquier sistema (la dimensión puede ser uno) siempre y cuando sea no-lineal e *invertible* (*homeo-mórfico*). Los esta-

[1] Otra forma es iterar hacia el *infinito negativo* y si la trayectoria se acerca al punto fijo, entonces es inestable.

dos de OMMs se simbolizan L^S, es decir, L oscilaciones de amplitud grande seguidas por S oscilaciones de amplitud pequeña. Por ejemplo, en la Figura 2.3 se tienen las siguientes OMMs: (a) 1^0, (b) 1^1, (c) 1^2 y (d) $(1^1)^2$.

2.1.2 *Reducción de un sistema de tiempo continuo a uno de tiempo discreto*

La idea de reducir un sistema de tiempo continuo para ser estudiado como un sistema de tiempo discreto asociado, fue sugerida por Poincaré en 1899. La metodología consiste en transformar un sistema continuo en un sistema discreto, el cual preserva la dinámica fundamental del sistema continuo original, pero con todas las propiedades de un sistema discreto (no confundir con la integración de sistemas continuos del apartado anterior). Dicha transformación resulta muy útil en el estudio de sistemas continuos cuando las trayectorias son cercanas a órbitas periódicas o cuasiperiódicas (por ejemplo oscilaciones de modo mixto). La metodología permite la reducción de la dimensión del sistema, ya que su construcción involucra la eliminación de al menos una variable de estado; permite el estudio de la dinámica global de sistemas hasta de 4 variables; da claridad conceptual, ya que muchos detalles pueden ser aclarados fácilmente con el mapa asociado; y se obtiene una reducción substancial en el número de datos. Sin embargo, son incapaces de capturar el comportamiento asociado a la presencia de puntos fijos del sistema continuo, tampoco ofrecen detalles de las trayectorias entre puntos del mapeo, ni sobre la duración de los periodos.

Dentro de esta metodología se encuentran el *Mapa de Poincaré* (MP) y el *Mapa de Próximo Máximo/Mínimo* (MPM). La construcción del MP involucra un procedimiento sencillo y para explicarlo se supone que se tiene un sistema dinámico

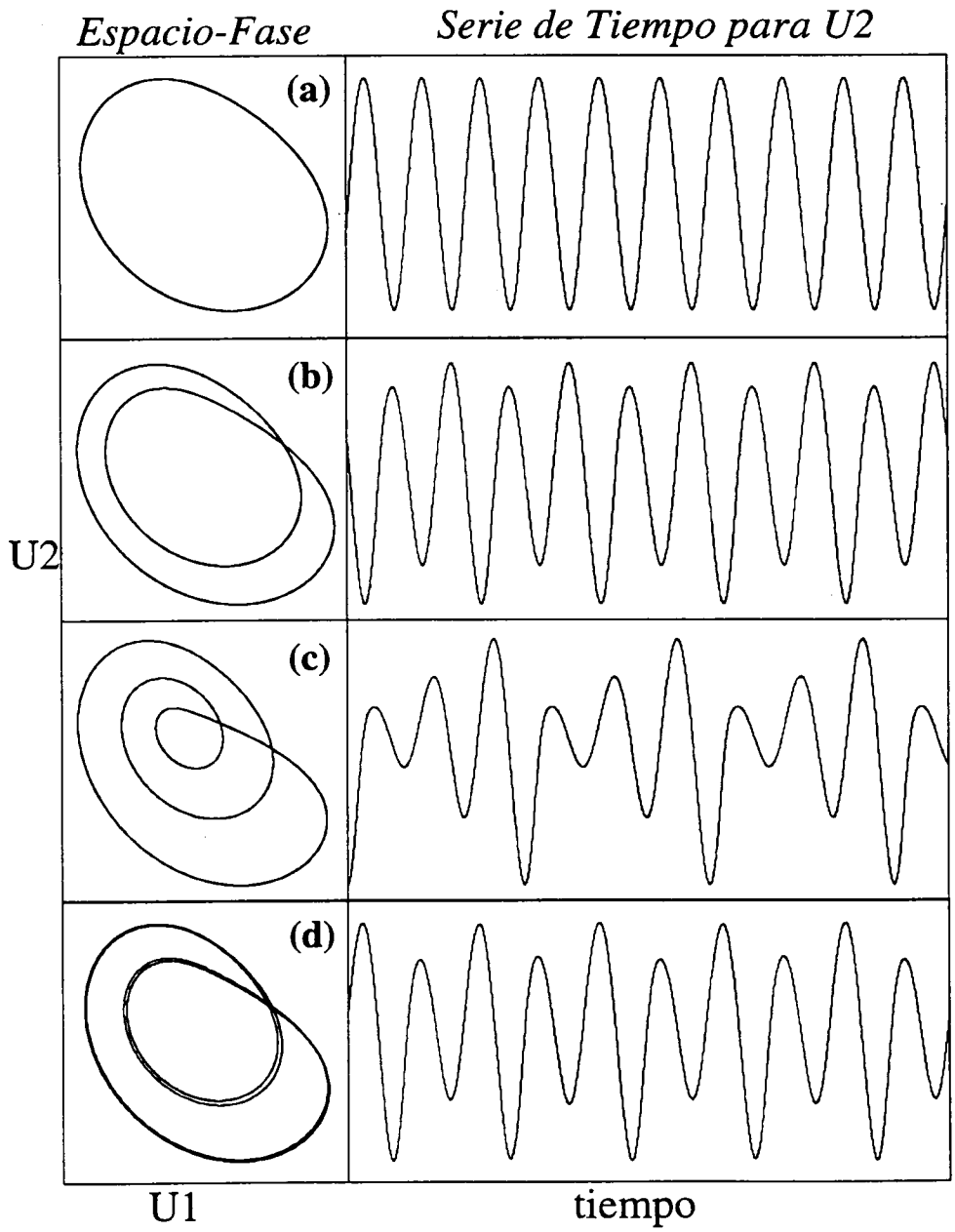


Figura 2.3: *Comportamiento Oscilatorio en el espacio-fase y en series de tiempo.*

tridimensional. En primer lugar se selecciona una superficie bidimensional (dentro del espacio-fase) que sea por todos lados transversal a las trayectorias cercanas a órbitas periódicas. Después se elige una dirección de mapeo, es decir, los puntos de cruce donde la trayectoria y el vector normal a la superficie cumplen con la condición: $\mathbf{f}(\mathbf{U}) \cdot \mathbf{n}(\mathbf{U}) > 0$. Por último se procede a recolectar los puntos del mapa que cumplen las condiciones anteriores. Este procedimiento se ejemplifica en la Figura 2.4.

Por su parte, la construcción del MPM también es simple, para ello se requiere escoger sólo una variable de estado de "interés" (en el caso experimental se trata de una variable medible), de la cual se obtienen los máximos (o mínimos) con respecto de su evolución dinámica a través del tiempo (un conjunto de datos discretos), por lo que el mapa será unidimensional. Una representación gráfica se puede obtener colocando el máximo (o mínimo) "actual" en el eje de las abscisas y el "próximo" máximo (o mínimo) en el eje de las ordenadas, de aquí el nombre de "mapas del próximo máximo (o mínimo)". Este procedimiento se ejemplifica en la Figura 2.5. En la práctica los MPMs son útiles cuando se desea estudiar un sistema real que tiene variables difíciles de medir o su medición resulta no costeable.

En cualquier caso (MP o MPM), la periodicidad de la órbita continua correspondiente puede determinarse a partir del número de puntos en el mapa. Es decir, un atractor constituido por un punto (en el MPM) representa una órbita de periodo uno semejante a la presentada en la Figura 2.3a, un atractor formado por dos puntos representa una órbita de periodo dos (semejante a la Figura 2.3b), etc. Los mapas asociados a órbitas *caóticas* aparecen estratificados, tienen *dimensión fractal* (no entera), y son referidos como *atractores extraños* [12, 40, 57] (un ejemplo de atractor extraño se puede observar en la Figura 2.2d).

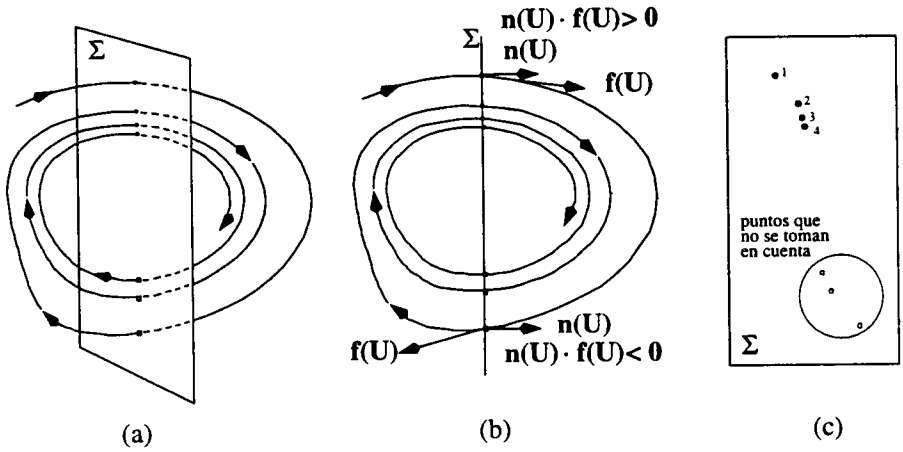


Figura 2.4: *Construcción de los Mapas de Poincaré. En (a) se escoge una sección transversa a la trayectoria; en (b) se elige una dirección de mapeo donde se cumpla:* $\mathbf{f(U) \cdot n(U) > 0}$*; y en (c) se recolectan los puntos del mapa.*

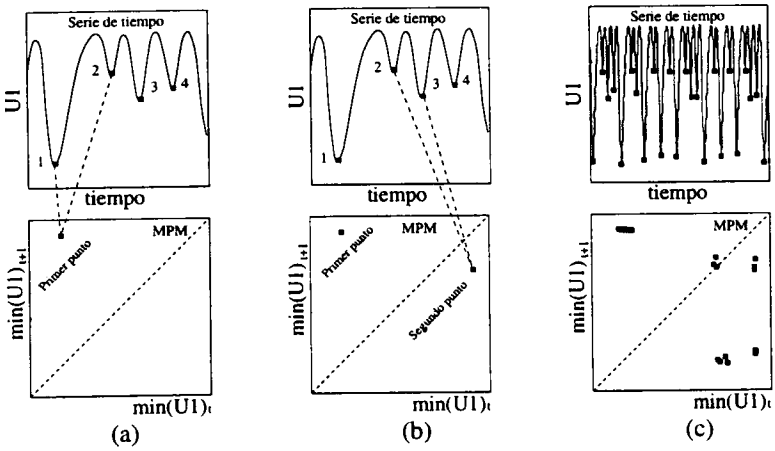


Figura 2.5: *Construcción de los Mapas de Próximo Mínimo. En (a) se toman dos valores mínimos consecutivos y se colocan como un punto del espacio* $min(U)_{proximo}$-$min(U)_{actual}$*; en (b) se toman los dos siguientes mínimos y se colocan como un segundo punto en el mapa de próximos mínimos; y en (c) se presenta una gráfica completa de un MPM.*

2.1.3 *Caracterización de Sistemas: Diagramas de Bifurcaciones*

En apartados anteriores se han descrito los tipos de comportamiento de los SDs cuando los parámetros de operación se mantienen fijos. Ahora se estudiará el efecto sobre las soluciones de dichos sistemas cuando los parámetros son modificados.

El término *bifurcación* se ha usado ampliamente para describir los cambios cualitativos significantes que ocurren sobre la estructura de las soluciones (a corto y largo plazo) en un sistema, cuando se modifican los parámetros de operación. Para su estudio, las bifurcaciones se han clasificado en dos grupos: locales y globales. La teoría de bifurcaciones locales se enfoca en los cambios estructurales que se generan en los puntos fijos de SDs o en situaciones donde el problema puede ser proyectado a esta forma (construcción de mapas a partir de EDOs). Mientras que la teoría de bifurcaciones globales se basa no sólo en los cambios estructurales de puntos fijos, sino también en todos los cambios que ocurren en una región extendida (con respecto de los puntos fijos) en el espacio-fase. En ambos casos, los cambios se presentan cuando los parámetros de operación son modificados. En el presente trabajo se tratará solamente con bifurcaciones locales.

Los diagramas de bifurcaciones (locales) son útiles para representar y entender de una manera clara y compacta la dinámica a largo plazo de un sistema. Cuando se gráfica la localización de los puntos fijos ($\widehat{\mathbf{U}}$) de un SD en función de un parámetro μ (en el espacio $\widehat{\mathbf{U}} - \mu$) se obtiene una gráfica como la mostrada en la Figura 2.6, la cual consta de varias líneas (sólidas, discontinua o punteadas) que se denominan *ramas de puntos fijos* o *ramas de órbitas periódicas.*

La Figura 2.6 muestra varios puntos (a, b, c, d, e y f) en los que el comportamiento es cualitativamente diferente (sobre las ramas) en ambos lados de dichos puntos, a

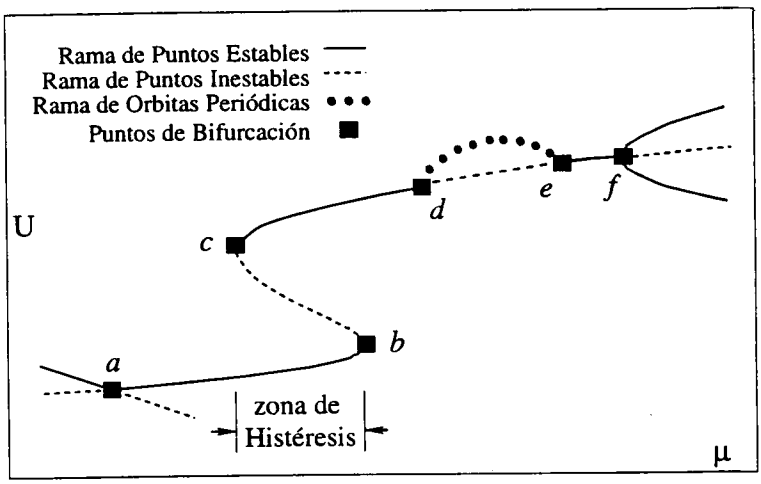


Figura 2.6: *Diagrama de Bifurcación en un parámetro. Las líneas sólidas representan ramas de puntos fijos estables; las líneas discontinuas, ramas de puntos fijos inestables; los círculos, ramas de órbitas periódicas; y los cuadros sombreados, puntos de bifurcación.*

ellos se les denomina *Puntos de Bifurcación* y la gráfica total se llama *Diagrama de Bifurcación de un Parámetro.* En esta figura sólo se presentan los puntos de bifurcación local más comunes en el estudio de sistemas dinámicos, los cuales se describen a continuación. El punto donde se cruzan dos ramas de puntos fijos y una de ellas pierde estabilidad mientras que la otra la gana (punto a), se le conoce como bifurcación *transcrítica*; el punto donde regresa una rama de puntos fijos perdiendo o ganando estabilidad (puntos b y c), se le conoce como bifurcación *nodo-punto de silla*; y el punto donde una rama de puntos fijos se expande en tres diferentes ramas (punto f), se le conoce como bifurcación de *horquilla* (sistemas continuos) o bifurcación de *doble periodo* (mapas). La región comprendida entre valores del parámetro de dos bifurcaciones nodo-punto de silla contiguas se le conoce como *región de histéresis.* En estas regiones puede observarse la coexistencia de diferentes estados estacionarios (multiplicidad de estados).

En sistemas dinámicos continuos con dos o más dimensiones, además de puntos fijos, se puede presentar el comportamiento periódico. Dicho comportamiento puede ser representado en los diagramas de bifurcación por un solo punto de la órbita (por ejemplo el valor máximo del estado $U1$), ya que es díficil representar la órbita completa. Los puntos de bifurcación d y e de la Figura 2.6 se conocen como *bifurcaciones de Hopf* y representa la transición de una rama de puntos fijos a una rama de soluciones periódicas. La rama de soluciones periódicas es representada por medio de círculos, y como en el caso de los puntos fijos aquí también se presentan las bifurcaciones de *doble periodo* y *nodo-punto de silla* (aunque no se muestran en la figura). Existen muchos otros tipos de bifurcaciones locales (de puntos fijos y órbitas periódicas), así como globales, que presentan estructuras más complicadas, los cuales estan fuera del alcance del presente trabajo, por lo que no serán discutidas aquí. Para una explicación de dichas bifurcaciones se recomienda la literatura citada anteriormente [64, 15].

Un diagrama de bifurcación de dos (o tres) parámetros es construido al graficar la localización de los puntos de bifurcación (tales como a, b, c, d, e y f del ejemplo anterior) en el espacio bidimensional (o tridimensional) de parámetros $\mu_1 - \mu_2$ $(-\mu_3)$. Esto sólo es posible cuando el sistema tiene más de un parámetro de operación.

2.1.4 Análisis de OMMs

El comportamiento de OMMs es un fenómeno muy complejo, que en la actualidad sigue siendo un tópico de investigación activo, ya que no se han encontrado los mecanismos principales que lo originan. Inicialmente, los estudios de este fenómeno se basaban principalmente en las observaciones experimentales (ver por ejemplo [21]). En recientes años ha aumentado el interés en el estudio de este fenómeno, lo que

ha generado que se propongan diferentes modelos matemáticos para explicarlo. Sin embargo, es difícil encontrar modelos, basados en principios básicos, que capturen cuantitativamente la dinámica experimental.

En el contexto de bifurcaciones, se sabe que las OMMs son el resultado de la interacción de bifurcaciones locales (doble periodo y nodo-punto de silla) con bifurcaciones globales (*intermitencia*). Una discusión de las bifurcaciones globales de intermitencia está más allá de los alcances de este trabajo (para una discusión de este tópico ver [64, 15]).

Maselco y Swinney [39] organizaron de una manera compacta la secuencia de observaciones en OMMs. Mostraron que el ordenamiento entre los diferentes estados de OMMs, en un diagrama de bifurcaciones, pueden ser arreglados de acuerdo a la *aritmética de Farey*[2]. Por ejemplo, si se tienen los estados de OMMs 1^1 y 1^0 (relación de Farey $\frac{1}{1} = 1$ y $\frac{1}{0} = \infty$) en valores del parámetro de operación μ_1 y μ_∞ respectivamente (donde $\mu_1 < \mu_\infty$), entonces se puede encontrar una OMMs con relación de Farey $\frac{1+1}{0+1} = 2$ (por ejemplo 2^1) en un valor del parámetro de operación μ_2 donde $\mu_1 < \mu_2 < \mu_\infty$. Barkley [4] investigó el fenómeno en diagramas de bifurcación de un parámetro con un modelo del sistema de Belousov-Zhabotinskii. Albahadily y col. [1] trabajaron con observaciones experimentales y encontraron que la transición a OMMs, inicia con "un estado de OMM muy cercano al estado 1^0". Posteriormente, Ringland y col. [48] estudiaron un mapa unidimensional con dos parámetros de operación y encontraron que los estados de OMMs "entran" por una bifurcación de nodo-punto de silla (via intermitencia I) y "salen" por una cascada de bifurcaciones de doble periodo. Además, asociaron el fenómeno con un escenario complejo de bifurcaciones en

[2]La suma Farey de los números naturales $\frac{L_1}{S_1} + \frac{L_2}{S_2}$ esta dada por $\frac{L_1+L_2}{S_1+S_2}$.

dos parámetros de operación.

Koper y Gaspard [28] propusieron un modelo de reacción en un electrodo rotativo, basado en principios básicos, que ayudó al entendimiento cualitativo de las transiciones observadas en la electrodisolución de cobre. Con este modelo se obtuvieron escenarios de bifurcación muy parecidos a los presentados por Ringland y col. [48], sin embargo, sólo existía concordancia cualitativa con las observaciones experimentales.

Las OMMs en la oxidación galvanostática de hidrogeno, fueron estudiadas por vez primera por Horányi y Visy [21], quienes propusieron un mecanismo intuitivo para explicar el fenómeno, describiendo los diferentes pasos durante una oscilación. Esta idea fue tomada por K. Krischer y col. [31, 33, 65], quienes encontraron que la presencia de aniones en la reacción es trascendental para la ocurrencia de las OMMs. Con estas ideas, ellos han propuesto un modelo matemático, basado en principios básicos, que predice la ocurrencia de OMMs. Dichas oscilaciones se presentan en valores de los parámetros de operación diferentes a los valores presentados experimentalmente.

Los modelos, basados en redes neuronales, también se han utilizado para estudiar el fenómeno de OMMs, sin embargo este tópico será tratado en la siguiente sección.

2.2 Redes Neuronales Artificiales

Las *Redes Neuronales Artificiales* (RNAs) son estructuras de modelo que surgieron como resultado del interés en la modelación de sistemas nerviosos [41]. Dichos modelos se han convertido en los últimos años en una herramienta de gran valor para el procesamiento de series de tiempo experimentales, reconocimiento de formas o patrones, codificación, clasificación, control, optimización, la identificación de sistemas no-lineales y muchas otras ramas del conocimiento. Las RNAs se han empleado en

la construcción de modelos dinámicos discretos (mapas comunes en la literatura de control de procesos) o continuos (conjunto de EDOs).

Las RNAs se pueden definir como una colección de unidades (*neuronas* o *nodos*) que se comunican entre sí por medio de conexiones (denominadas *axones*), tal como se ejemplifica en la Figura 2.7. De esta forma, las RNAs se agrupan en capas, de las cuales se distinguen tres tipos: capa de entrada, capas ocultas y capa de salida. La capa de entrada es donde se recibe la información; en las capas ocultas se procesa la información, y en la capa de salida se obtiene la respuesta de la red.

2.2.1 Reseña histórica

Las primeras aplicaciones teóricas sobre el cerebro y el pensamiento fueron dadas por algunos filósofos griegos, como Platón (427-347 a.C.) y Aristóteles (384-422 a.C.), y más tarde fueron mantenidas por Descartes (1596-1650) y los filósofos empiristas del siglo XVIII.

Sin embargo, no fue hasta 1936 cuando Alan Turing se interesó en estudiar el cerebro como una forma de ver el mundo de la computación. Para 1943, W. McCulloc y W. Pitts [41], a quienes se les atribuyen los fundamentos de las RNAs, lanzaron una teoría acerca de la forma de trabajar de las neuronas (modelaron una red simple mediante circuitos eléctricos). En 1949, D. Hebb [17] escribió un libro en el que se establece una conexión entre psicología y fisiología. Fue hasta 1957 cuando F. Rosenblatt [49] desarrolló la más antigua RNA (el Perceptrón). La más importante característica de esta red es su capacidad de generalizar, sin embargo, tenía la desventaja de no poder clasificar clases no-separables linealmente. En 1959, B. Widrow y M. Hoff [63] desarrollaron el modelo ADALINE (ADAptative LINear Elements). Esta

fue la primera RNA aplicada a un problema real (filtros adaptativos para eliminar ecos en las líneas telefónicas) y se ha usado comercialmente durante varias décadas. En 1967, S. Grossberg [14] (uno de los mayores investigadores de las RNAs) creó una red (Avalancha) consistente en elementos discretos con actividad que varía con el tiempo, para resolver actividades tales como reconocimiento continuo del habla y aprendizaje del movimiento de los brazos de un robot.

En 1969 surgieron numerosas críticas a las RNAs, las cuales paralizaron las investigaciones en este campo hasta 1982, por lo que muchas de las investigaciones dieron un giro hacia la inteligencia artificial, que prometía más por aquel entonces. M. Minsky y S. Papert [42] publicaron un libro (*Perceptrons*) en el que consideraban que la extensión a Perceptrones multinivel (el Perceptrón original sólo poseía una capa) era completamente estéril. Las limitaciones del Perceptrón eran importantes, sobre todo su incapacidad para resolver muchos problemas interesantes y sencillos. A pesar del rumbo que tomaron las investigaciones, durante esos años hubo algunos trabajos importantes en las investigaciones de RNAs, de los cuales se puede mencionar la tesis doctoral de P. Werbos [62]. En su trabajo doctoral, Werbos propone la regla DELTA para el cálculo de derivadas en el entrenamiento de RNAs. La regla DELTA pasó desapercibida, y fue hasta 1986 cuando fue redescubierta por D. Rumelhart y col. [50] (a quienes se les atribuye el mérito de este descubrimiento).

En 1982, coincidieron varios eventos que hicieron resurgir el interés por las investigaciones en RNAs, desde ese entonces las investigaciones en RNAs han crecido exponencialmente. En ese año, J. Hopfield [20] presentó su trabajo sobre RNAs en la Academia Nacional de Ciencias. En el trabajo, describe con claridad y rigor matemático una RNA a la que le ha dado su nombre, la cual es una variación del

Asociador Lineal (propuesto en 1977 por J. Anderson y col. [2]). Además, en 1982 se celebró la *U.S.-Japan Joint Conference on Cooperative/Competitive Neural Networks.* Siguiendo al trabajo de Hopfield, G. E. Hinton y col. [19] desarrollaron los modelos de RNAs llamados *Máquinas de Boltzman*, los cuales fueron usados por primera vez para resolver problemas de clasificar clases no-separables linealmente. De cualquier manera, el aprendizaje de las Máquinas de Boltzman es muy lento. En 1986, D. Rumelhart y col. [50] redescubrieron el mucho más rápido algoritmo *Backpropagation* (propagación hacia atrás) o regla DELTA. Desde entonces, este algoritmo, ha sido la elección para el entrenamiento en las investigaciones de RNAs.

A partir del resurgimiento de las RNAs, se han publicado muchos trabajos en los que se muestran sus diferentes capacidades como identificadores. Cybenko [7] realizó un estudio matemático riguroso y encontró que cualquier función arbitraria (y sus derivadas) puede ser aproximada usando RNAs con bastantes neuronas. Hornick y col. [22] y Irie y Miyake [24], en sus respectivos trabajos, demostraron que las RNAs tienen características de aproximadores universales. Lapedes y Farber [38] demostraron que una clase de RNAs pueden ser exitosamente usadas para construir modelos dinámicos con excelentes cualidades predictivas (principalmente para predicciones a *corto plazo*), aún en presencia de comportamiento caótico (determinístico). En un libro publicado por Weigend y Gershenfeld [61], destacan la capacidad de predicción de las RNAs en series de tiempo caóticas obtenidas de diferentes disciplinas de trabajo. También, se muestra que una red buena para predicciones a largo plazo, puede no serlo para predicciones a corto plazo, y viceversa.

En la identificación de sistemas dinámicos, continuos en el tiempo, han mostrado ser exitosas muchas arquitecturas de RNAs, al ser aplicadas en forma de modelos de

tiempo discreto. Por ejemplo, se pueden mencionar las RNAs de alimentación hacia adelante presentadas por Lapedes y Farber [38]; las de base radial presentadas por Wasserman [60]; o bien, las redes recurrentes presentadas por Su y McAvoy [54] y Tsung y Cottrell [58].

Sin embargo, las aproximaciones de tiempo continuo, para modelar sistemas continuos, no han gozado de tanta atención como su contraparte discreta. Chu y Shoureshi [6] usaron modelos continuos con técnicas de integración muy simples, para la identificación de sistemas de tiempo continuo. You y Nikolaou [67] explotaron las propiedades de las redes recurrentes para construir modelos continuos a partir de datos experimentales.

Rico-Marínez y col. [47, 45] han utilizado modelos, basados en RNAs, encaminados hacia la caracterización de transiciones (bifurcaciones) en un parámetro de operación. Ellos han utilizado tanto modelos continuos como discretos para tales fines. También han demostrado que las aproximaciones discretas tradicionales, para identificación de sistemas continuos, son inapropiadas para fines de caracterización.

Las investigaciones en identificación de sistemas espacio-temporales, en su mayoría, se han centrado en modelos polinomiales. En pocos trabajos se han utilizado las redes neuronales para identificar dichos sisistemas. Por ejemplo, K. Krischer y col. [34] construyeron un modelo para identificar una reacción catalítica con variaciones espacio-temporales (la oxidación de CO sobre Pt(110)). Sus observaciones consistieron de imágenes de la superficie catalítica a diferentes instantes en el tiempo, las cuales fueron digitalizadas. Ellos no usaron los datos digitalizados para modelar el sistema por ecuaciones diferenciales parciales. Procesaron los datos por medio de técnicas estadísticas (expansión de Karhunen-Loève) para reducir la dimensionalidad

del sistema a sólo 4 coeficientes. Dichos coeficientes, fueron usados para identificar el sistema de diferentes maneras, con modelos basados en RNAs. Recientemente, Lagaris y col. [36] presentaron un trabajo en el que utilizan las RNAs para resolver EDPs (se requiere conocer la EDP).

2.2.2 *Funcionamiento de las RNAs*

La construcción de una RNA implica escoger los siguientes elementos: una estructura de red (número de capas y nodos por capa); se debe proponer una *regla de activación* (cómo combinar las entradas de cada una de las neuronas para producir su salida); un *patrón de conectividad* (cómo se conectan las neuronas entre ellas y la dirección en la cual fluye la información), y una *regla de aprendizaje* o *entrenamiento* (cómo actualizar los parámetros). Para los problemas de identificación, la regla de aprendizaje es el resultado de la minimización del error entre los valores observados y predichos por las salidas de la red (mínimos cuadrados). Mientras que, la estructura, la regla de activación y los patrones de conectividad se basan en una selección empírica.

Para entender la forma en que trabajan las RNAs [37] se tomará como base la estructura mostrada en la Figura 2.7. Esta figura muestra una red típica de alimentación hacia adelante con dos capas ocultas. En esta arquitectura, las neuronas en alguna capa dada reciben como entradas las salidas de todas la neuronas en la capa de abajo. En el esquema, el resultado (salida) que arroja la j-ésima neurona en la i-ésima capa es denotada como V_{ij}, el número de neuronas en la capa i-ésima es denotado como n_i y el número de capas es m (para el caso de la Figura 2.7 se tiene $m = 4$).

El proceso para producir la salida de una neurona es esquematizado en la Figu-

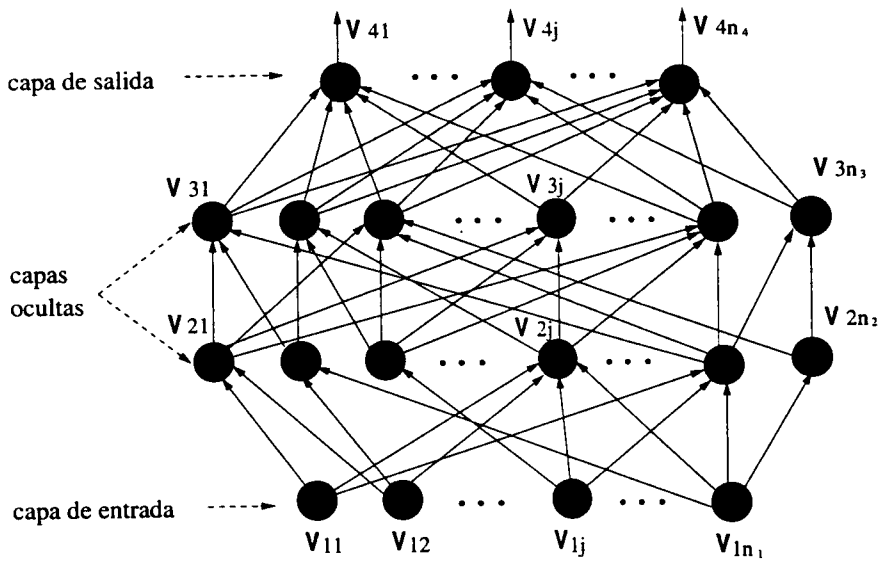


Figura 2.7: *Estructura básica de una red neuronal artificial. Las neuronas de alguna capa deben estar completamente interconectadas con las otras, aunque en la figura sólo se muestran unas cuantas conexiones.*

ra 2.8. Para los propósitos de cálculo, las conexiones entre dos neuronas son representadas matemáticamente por medio de los *pesos* **W**. De esta manera, W_{ijk} es el peso (conexión en forma gráfica) entre la j-ésima neurona de la i-ésima capa con la k-ésima neurona de la $(i-1)$-ésima capa. Para producir la salida de la neurona j en la capa i (V_{ij}), primero se hace la sumatoria de las salidas de cada neurona en la capa $i-1$ multiplicada por su correspondiente peso que conecta con la neurona ij, a esta sumatoria se le adiciona un parámetro Θ, llamado *umbral*, que va asociado con la neurona ij, el resultado de la sumatoria se define como Y_{ij} (Figura 2.8b nivel a).

$$Y_{ij} = \sum_{k=1}^{n_{i-1}} W_{ijk} V_{(i-1)k} + \Theta_{ij}$$

después el valor de Y_{ij} se introduce en una *función de activación* $g(Y_{ij})$ para producir

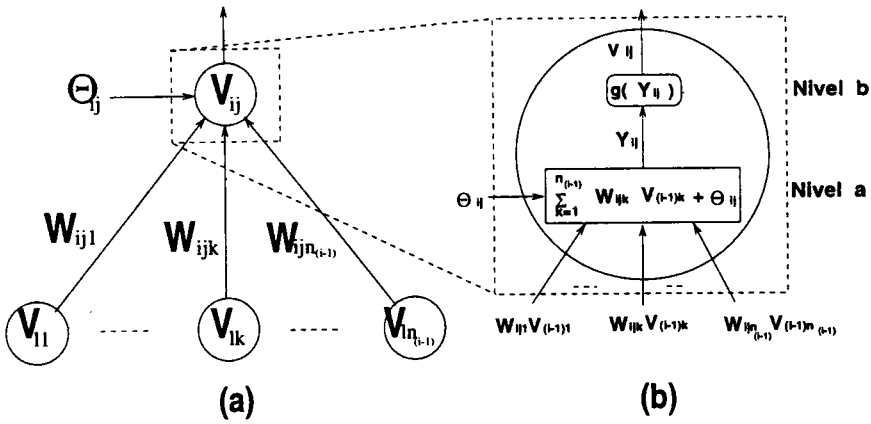


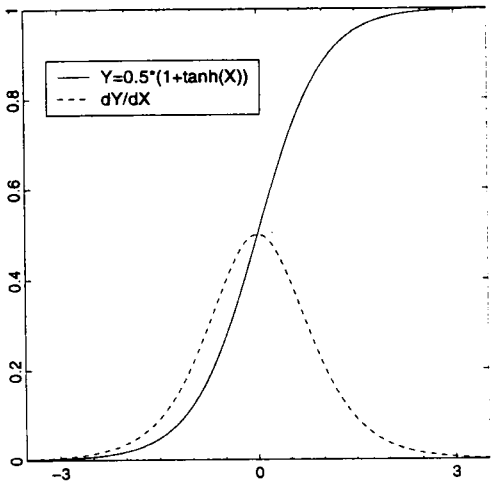
Figura 2.8: *Secuencia esquematizada de la activación de neuronas para producir su salida.*

la salida deseada V_{ij} (Figura 2.8b nivel b).

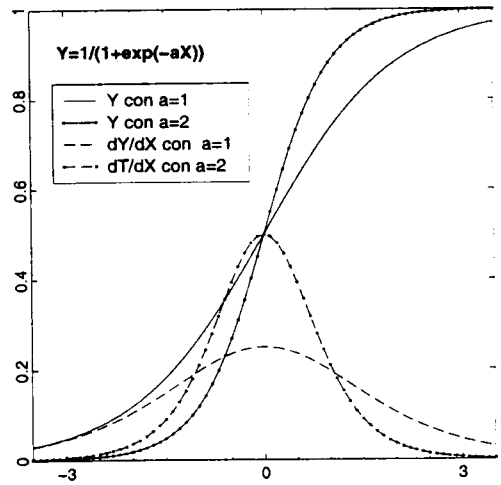
$$V_{ij} = g_{ij}(Y_{ij}) = g\left(\sum_{k=1}^{n_{i-1}} W_{ijk}V_{(i-1)k} + \Theta_{ij}\right)$$

En la nomenclatura presentada aquí, las entradas a la red coinciden exactamente con las salidas de las neuronas de la primera capa (V_{1j}).

Para la elección de la función de activación $g(Y)$, se ha demostrado [7] que las funciones discriminantes (monotónicas crecientes, por ejemplo sigmoidales), como las mostradas en la Figura 2.9, son apropiadas para aproximar una función no-lineal mediante la construcción de RNAs, aunque se puede usar cualquier otro tipo de funcionalidad. En los casos de identificación presentados aquí se usarán las funciones de activación de la Figura 2.9a para las capas ocultas y una función lineal o identidad ($V_{ij} = Y_{ij}$) para la capa de salida.



(a) función del tipo *tanh()*.



(b) función del tipo *exp()*.

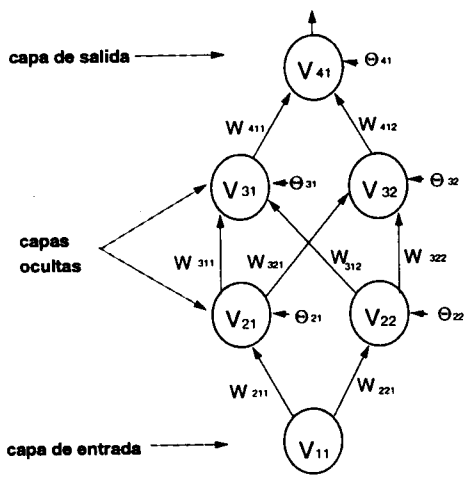
Figura 2.9: *Dos tipos de funciones de activación más usadas en RNAs. Estrictamente hablando, las funciones en (a) son un caso particular de (b) cuando* $a = 2$.

2.2.3 Arquitecturas

En lo que concierne a la arquitectura de las redes neuronales (Figura 2.10), se pueden clasificar en dos grupos: redes de alimentación hacia adelante y redes recurrentes. Dicha clasificación esta basada en la forma que se transmite la información. En las redes de alimentación hacia adelante (Figura 2.10a) la información fluye únicamente desde las capas de nivel inferior hacia las capas de nivel superior , es decir, la información sólo fuluye hacia adelante. Por otro lado, en las redes recurrentes (Figura 2.10b) la información fluye en cualquier dirección posible: hacia adelante, hacia atrás y lateralmente. Las redes recurrentes requieren un esfuerzo mayor en su evaluación y entrenamiento, comparado con las redes de alimentación hacia adelante (ver las funciones en la Figura 2.10). La arquitectura de las RNAs es determinada en gran parte por la aplicación que se pretende. Para la mayoría de los casos basta con una red de alimentación hacia adelante. En el presente trabajo se tratarán específicamente las redes de alimentación hacia adelante con dos capas de neuronas ocultas (las cuales se denominarán redes *estándar* [38]).

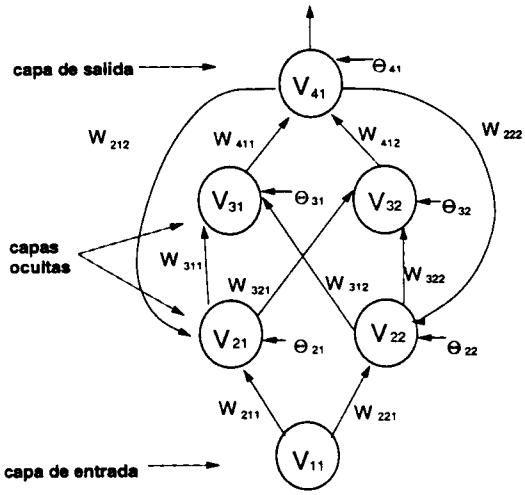
2.2.4 Entrenamiento

El entrenamiento de una RNA comprende la manera en que se actualizan (o cambian) sus parámetros para que sea capaz de reproducir ciertos eventos o situaciones. En principio, se pueden cambiar las propiedades del aproximador de red modificando el patrón de conectividad [50], sin embargo, esta alternativa es poco usada en la páctica. En lugar de eso se modifican las aproximaciones de la red cambiando el valor de sus parámetros (pesos y umbrales) por medio de la *experiencia*. Para los problemas de identificación, la regla de aprendizaje es el resultado de la minimización



$$V_{41} = W_{411}\, g(W_{311}\, g(W_{211}V_{11} + \Theta_{21}) + W_{312}\, g(W_{221}V_{11} + \Theta_{22}) + \Theta_{31}) + W_{412}\, g(W_{321}\, g(W_{211}V_{11} + \Theta_{21}) + W_{322}\, g(W_{221}V_{11} + \Theta_{22}) + \Theta_{32}) + \Theta_{41}$$

(a)



$$V_{41} = W_{411}\, g(W_{311}\, g(W_{211}V_{11} + W_{212}V_{41} + \Theta_{21}) + W_{312}\, g(W_{221}V_{11} + W_{222}V_{41} + \Theta_{22}) + \Theta_{31}) + W_{412}\, g(W_{321}\, g(W_{211}V_{11} + W_{212}V_{41} + \Theta_{21}) + W_{322}\, g(W_{221}V_{11} + W_{222}V_{41} + \Theta_{22}) + \Theta_{32}) + \Theta_{41}$$

(b)

Figura 2.10: *Dos arquitecturas básicas de RNAs. En (a) una red neuronal de alimentación hacia adelante y en (b) una red neuronal recurrente.*

del error entre los valores observados y predichos por las salidas de la red (mínimos cuadrados). Puesto que en el presente trabajo se aplican los madelos de RNA para la identificación de SD, el proceso de aprendizaje de estas redes se discutirá en el Capítulo 3.

3. Teoría de Identificación de SDs mediante RNAs

La teoría de SDs junto con las RNAs aplicadas de manera apropiada, en la identificación de sistemas complejos, llegan a ser excelentes herramientas en la exploración del comportamiento observado. En el presente capítulo se dan las bases para la identificación de los siguientes sistemas dinámicos:

$$\mathbf{U}_{t+1} = \mathbf{g}(\mathbf{U}_t; \boldsymbol{\mu}) \tag{3.1}$$

$$\frac{d\mathbf{U}}{dt} = \mathbf{f}(\mathbf{U}; \boldsymbol{\mu}) \tag{3.2}$$

$$\frac{\partial U}{\partial t} = F(U, \frac{\partial U}{\partial X}, \frac{\partial^2 U}{\partial X^2}, \cdots, \frac{\partial^\gamma U}{\partial X^\gamma}; \boldsymbol{\mu}) \tag{3.3}$$

donde las Ecuaciones 3.1 y 3.2 fueron definidas previamente en el Capítulo 2, y la Ecuación 3.3 es una *Ecuación Diferencial Parcial* (EDP) de comportamiento espacio-temporal, siendo la variable de estado (U) un escalar, X la dimensión espacial y γ el máximo orden de derivada espacial que puede aparecer en la ecuación.

Aquí se han escogido las RNAs como estructuras de modelo para la identificación de estos sistemas, en este sentido, la RNA hará las veces de la función $\mathbf{g}(\cdot)$, $\mathbf{f}(\cdot)$ o $F(\cdot)$, respectivamente. Las entradas a la RNA consistirán de todas las variables que se encuentran dentro de los paréntesis en las Ecuaciones 3.1, 3.2 ó 3.3, mientras que las salidas de dichas RNAs consistirán del vector de funciones $\mathbf{g}(\cdot)$, $\mathbf{f}(\cdot)$ o el escalar $F(\cdot)$, respectivamente. En lo que resta del trabajo, RNA-SD se referirá a un SD que es identificado mediante una RNA, y particularmente RNA-mapa, RNA-EDO o RNA-EDP se referirán a un SD como el de la Ecuación 3.1, 3.2 ó 3.3, respectivamente, que es identificado por medio de una RNA.

Independientemente del tipo de modelo que se construya, en primer término se asume que se dispone de un conjunto de p pares de entrada-salida

$$\begin{bmatrix} \mathbf{I}^1 & , & \mathbf{O}^1 \\ \mathbf{I}^2 & , & \mathbf{O}^2 \\ \mathbf{I}^3 & , & \mathbf{O}^3 \\ & \vdots & \\ \mathbf{I}^p & , & \mathbf{O}^p \end{bmatrix}$$

al cual se denominará *Conjunto de Entrenamiento.* En el contexto de SDs, este conjunto representa la evolución (a través del tiempo) del sistema dinámico que se desea identificar. El objetivo del proceso de *entrenamiento* es encontrar el conjunto de parámetros de la RNA-SD (pesos y umbrales en redes neuronales) que mejor ajusten la evolución dinámica de los datos experimentales. Es decir, buscar el conjunto de párametros adecuados para que la RNA-SD capture el mapa fundamental entre las entradas y salidas del conjunto de entrenamiento. Dicho proceso de entrenamiento es

logrado por minimización de la suma de los cuadrados de los errores (función error) entre los valores observados experimentalmente ($\mathbf{O}^t$) y las predicciones arrojadas por el modelo RNA-SD. Para el caso de identificación de SDs discretos (mapas), las predicciones arrojadas por el modelo RNA-mapa puden ser comparadas directamente con las observaciones experimentales y en los SDs continuos (EDOs o EDPs) esto sólo se logra a menos que se tengan mediciones directas de derivadas (ordinarias o parciales, respectivamente). En la mayoría de los casos de identificación, en lugar de mediciones directas de derivadas, se tienen mediciones atrasadas en el tiempo de una o más variables de estado, y en el caso de sistemas espacio-temporales (EDPs) las mediciones se tienen a diferentes posiciones en el espacio. Es por eso que se requiere dar algún procesamiento previo a las Ecuaciones 3.2 y 3.3 para poder ajustar los parámetros de la RNA (entrenamiento de la RNA).

La técnica de optimización utilizada en este trabajo, para el entrenamiento de RNAs, es la del *Gradiente Conjugado.* Esta técnica utiliza como información principal las derivadas de la función error con respecto a los parámetros de la red (pesos y umbrales).

La construcción de la función error y derivadas de un modelo RNA-SD además de la técnica del gradiente conjugado se presentan a continuación para el caso de RNAs de alimentación hacia adelante de 4 capas (2 capas ocultas) con función de activación sigmoidal en las capas ocultas y función de activación lineal (identidad) en la capa de salida. La teoría presentada no es exclusiva de este tipo de RNAs, sino que puede ser extrapolada a cualquier estructura de RNA, con los cambios pertinentes, mientras sea de alimentación hacia adelante.

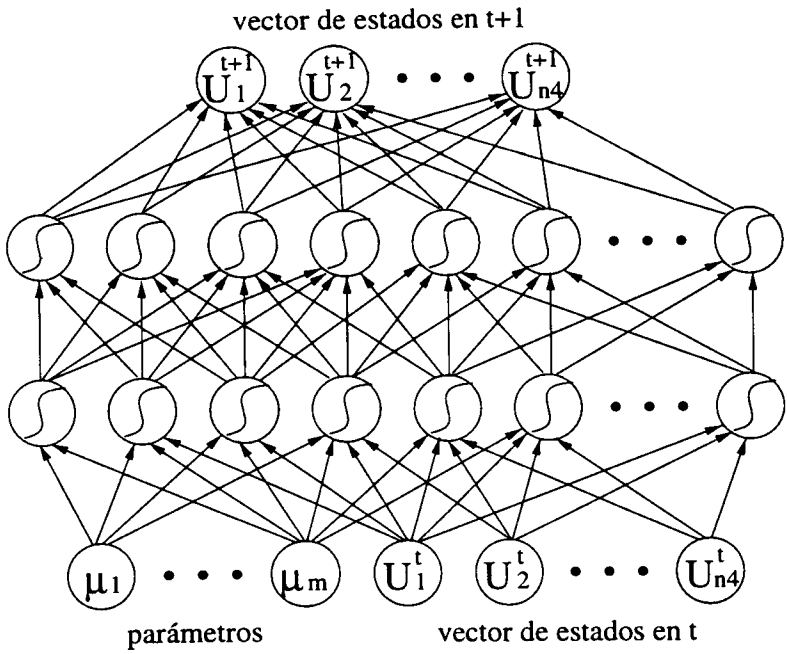


Figura 3.1: *Esquema de RNA que puede ser usada en la identificación de mapas.*

3.1 SDs discretos (mapas)

La Figura 3.1 muestra la estructura de una RNA representativa en la identificación de un sistema discreto como el de la Ecuación 3.1. En este esquema las entradas y salidas de la red son valores medibles directamente de los experimentos. La teoría que se presenta a continuación es también la base para cualquier sistema (dinámico o no) que pueda ser identificado mediante una RNA.

3.1.1 Construcción de la Función Error

Para obtener la función error, se requiere evaluar la red neuronal y comparar su salida contra los datos experimentales en cada punto del conjunto de entrenamiento. Esto se logra evaluando la RNA de la misma forma como se describe en el Capítulo 2, sólo que ahora la RNA es alimentada con el vector de entrada en el conjunto de entrenamiento

($\mathbf{V}_1^t = [\boldsymbol{\mu}, \mathbf{U}^t] = \mathbf{I}^t$). Téngase en cuenta que las entradas a la red son el vector de parámetros de operación y el vector de variables de estado. La función error (también conocida como *función de energía*) se construye comparando las salidas de la RNA evaluadas en el instante t, $\mathbf{V}_4^t(\mathbf{V}_1^t) = \mathbf{U}^{t+1}(\mathbf{U}^t)$, con las salidas de las observaciones experimentales en el mismo instante ($\mathbf{O}^t$). Para un punto cualquiera (t) del conjunto de entrenamiento se obtiene un error local:

$$E^t(\mathbf{W}, \Theta) = \frac{1}{2}\sum_{j=1}^{n_4}\left(O_j^t - V_{4j}^t(\mathbf{W}, \Theta)\right)^2$$

mientras que con el total de puntos contenidos en el conjunto de entrenamiento se obtiene el siguiente error global:

$$E(\mathbf{W}, \Theta) = \sum_{t=1}^{p} E^t(\mathbf{W}, \Theta) = \frac{1}{2}\sum_{t=1}^{p}\sum_{j=1}^{n_4}\left(O_j^t - V_{4j}^t(\mathbf{W}, \Theta)\right)^2 \tag{3.4}$$

donde V_{4j}^t es la j-ésima salida de la RNA (j-ésima variable de estado) evaluada con las entradas del t-ésimo punto del conjunto de entrenamiento y n_4 se refiere al número de nodos de la capa de salida (número de variables de estado). Aquí se utiliza la notación pertinente al contexto de RNAs ($\mathbf{V}_4^t$) en lugar de la notación de SDs ($\mathbf{U}^{t+1}$) ya que esta notación facilitará la obtención del cálculo de derivadas en el siguiente apartado. La sumatoria interna representa la contribución de cada variable de estado (salidas de la red) y la externa representa la contribución del total de pares de entrada-salida disponibles (conjunto de entrenamiento). De esta forma la función error depende directamenta de los parámetros de la RNA ($\mathbf{W}$ y Θ).

El entrenamiento se lleva a cabo en base a la minimización de la la función $E(\mathbf{W}, \Theta)$. El problema de minimización, así obtenido, es resuelto por medio del

método del gradiente conjugado que se describe en un apartado posterior.

3.1.2 Método Delta para el cálculo de derivadas en RNAs

El objetivo ahora es describir la forma de evaluar el vector de primeras derivadas de la función error $E(\mathbf{W}, \Theta)$ con respecto de los vectores $\mathbf{W}$ y Θ, es decir $\nabla E(\mathbf{W}, \Theta) = \left[\frac{dE}{dW_{ijk}}, \frac{dE}{d\Theta_{ij}}\right]$, el cual es necesario por el método del *Gradiente Conjugado* para el cálculo de las direcciones de búsqueda [18, 50, 63]. Tomando en consideración la Ecuación 3.4 se pueden obtener las siguientes expresiones de derivadas. Iniciando desde la capa más alta ($i = 4$ en este caso), se encuentran las siguientes expresiones de derivadas con respecto a cualquier peso y umbral que conectan a los nodos de la capa superior (W_{4jk} y Θ_{4j})

$$\begin{aligned}
\frac{dE(\mathbf{W},\Theta)}{dW_{4jk}} &= \frac{1}{2}\sum_{t=1}^{p}\sum_{i=1}^{n_4}\frac{d\left(O_i^t - V_{4i}^t\right)^2}{dW_{4jk}} \\
\frac{dE(\mathbf{W},\Theta)}{d\Theta_{4j}} &= \frac{1}{2}\sum_{t=1}^{p}\sum_{i=1}^{n_4}\frac{d\left(O_i^t - V_{4i}^t\right)^2}{d\Theta_{4j}}
\end{aligned}$$

derivando las expresiones del lado derecho

$$\begin{aligned}
\frac{dE}{dW_{4jk}} &= \frac{1}{2}\sum_{t=1}^{p}\sum_{i=1}^{n_4} 2\left(O_i^t - V_{4i}^t\right)\frac{-dV_{4i}^t}{dW_{4jk}} \\
\frac{dE}{d\Theta_{4j}} &= \frac{1}{2}\sum_{t=1}^{p}\sum_{i=1}^{n_4} 2\left(O_i^t - V_{4i}^t\right)\frac{-dV_{4i}^t}{d\Theta_{4j}}
\end{aligned}$$

como se esta derivando con respecto a los pesos y umbrales que conectan a algún nodo particular de la capa de salida, entonces el índice i debe coincidir con el índice j, por lo tanto desaparece la sumatoria interna y se tiene

$$\begin{aligned}
\frac{dE}{dW_{4jk}} &= -\sum_{t=1}^{p}\left(O_j^t - V_{4j}^t\right)\frac{dV_{4j}^t}{dW_{4jk}} \\
\frac{dE}{d\Theta_{4j}} &= -\sum_{t=1}^{p}\left(O_j^t - V_{4j}^t\right)\frac{dV_{4j}^t}{d\Theta_{4j}}
\end{aligned}$$

y ahora tomando en consideración la forma en que se evalúan las salidas de una RNA (Capítulo 2)

$$V_{4j} = \sum_{i=1}^{n_3} W_{4ji} V_{3i} + \Theta_{4j}$$

entonces se tendrán las expresiones buscadas para la última capa

$$\begin{aligned} \frac{dE}{dW_{4jk}} &= -\sum_{t=1}^{p} \left(O_j^t - V_{4j}^t\right) V_{3k}^t \\ \frac{dE}{d\Theta_{4j}} &= -\sum_{t=1}^{p} \left(O_j^t - V_{4j}^t\right) \end{aligned} \tag{3.5}$$

Si se define una nueva variable δ como:

$$\delta_{4j}^t \equiv O_j^t - V_{4j}^t \tag{3.6}$$

entonces se puede escribir

$$\begin{aligned} \frac{dE}{dW_{4jk}} &= -\sum_{t=1}^{p} \delta_{4j}^t V_{3k}^t \\ \frac{dE}{d\Theta_{4j}} &= -\sum_{t-1}^{p} \delta_{4j}^t \end{aligned} \tag{3.7}$$

donde la introducción de la nueva variable δ es la clave del cálculo de derivadas para las restantes capas.

Después de aplicar las ideas anteriores para el cálculo de las derivadas de $E(\mathbf{W}, \mathbf{\Theta})$ con respecto a los pesos y umbrales de las capas ocultas ($[W_{3jk}, \Theta_{3j}]$ y $[W_{2jk}, \Theta_{2j}]$) y llevando a cabo un poco de algebra (intercambios de sumatorias), se obtienen las siguientes expresiones en las capas ocultas ($i = 3,\ 2$)

$$\begin{aligned} \frac{dE}{dW_{ijk}} &= -\sum_{t=1}^{p} \delta_{ij}^t V_{(i-1)k}^t \\ \frac{dE}{d\Theta_{ij}} &= -\sum_{t-1}^{p} \delta_{ij}^t \end{aligned} \tag{3.8}$$

sólo que para estas capas δ esta definida como

$$\delta_{ij}^{t} \equiv g\prime(Y_{ij}^{t}) \sum_{h=1}^{n_{i+1}} W_{(i+1)hj}\delta_{(i+1)h}^{t} \tag{3.9}$$

donde $g\prime(Y_{ij}^{t})$ es la derivada de la función de activación, $g(Y_{ij}^{t})$, con respecto de Y_{ij}^{t}, es decir $\frac{dg(Y)}{dY}$. Las funciones de activación $g(Y)$ y sus derivadas $g\prime(Y_{ij}^{t})$ ya fueron mostradas anteriormente en la Figura 2.9.

En resumen, la evaluación del vector de derivadas de la función error (Ecuación 3.4) con respecto de los parámetros ($\mathbf{W}$ y Θ) de una RNA de alimentación hacia adelante con función de activación lineal en la capa de salida y cualquier función de activación en las capas ocultas, se obtiene con las siguientes expresiones:

$$\begin{aligned} \frac{dE}{dW_{ijk}} &= -\textstyle\sum_{t=1}^{p} \delta_{ij}^{t} V_{(i-1)k}^{t} \\ \frac{dE}{d\Theta_{ij}} &= -\textstyle\sum_{t-1}^{p} \delta_{ij}^{t} \end{aligned} \tag{3.10}$$

donde δ se debe evaluar de la siguiente forma

$$\delta_{ij}^{t} \equiv \begin{cases} O_{j}^{t} - V_{ij}^{t} & i = m = 4 \quad \text{capa de salida} \\ g\prime(Y_{ij}^{t}) \sum_{h=1}^{n_{i+1}} W_{(i+1)hj}\delta_{(i+1)h}^{t} & i = 3,\ 2 \quad \text{capas ocultas} \end{cases} \tag{3.11}$$

Obsérvese en la Ecuación 3.11 que la evaluación de las derivadas requiere de conocer la variable δ en la capa superior, por lo que las derivadas son propagadas desde la capa de salida hacia la capa de entrda. Este hecho hace que al cálculo de derivadas en RNAs como las mostradas en la Figura 3.1 se le denomine *Método de Propagación hacia Atrás*, y el hecho de utilizar la variable δ hace que al procedimiento se le conozca también como *Método Delta*. De esta manera, durante el entrenamiento,

la evaluación de la red involucra propagar la información hacia adelante y el cálculo de las derivadas involucra propagarla hacia atrás.

3.2 SDs continuos (EDOs)

La Figura 3.2 muestra la estructura de una RNA representativa de un sistema de EDOs como el de la Ecuación 3.2. Como lo muestra la figura, las salidas de la red son valores de derivadas temporales. De esta forma, la función error puede construirse de acuerdo a uno de los siguientes casos en el conjunto de entrenamiento:

1. El vector de salidas ($\mathbf{O}^l$) del conjunto de entrenamiento consta de mediciones directas de las derivadas temporales ($\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t}$).

2. El vector de salidas ($\mathbf{O}^l$) del conjunto de entrenamiento consta de mediciones atrasadas en el tiempo ($\mathbf{U}^{t+1}$) en lugar de derivadas temporales.

Para el primer caso, la construcción de la función error se hace de la misma forma como para el caso de los modelos discretos, pero este caso es muy aislado ya que en la práctica es difícil obtener mediciones confiables de las derivadas temporales. Para el segundo caso, la construcción de la función error involucra dos situaciones diferentes. Por un lado, se pueden transformar las mediciones discretas a derivadas temporales por medio de esquemas de diferencias finitas, y de esta forma la función error se construye directamente como en el caso de los modelos discretos. Sin embargo, esta no es la alternativa más viable, ya que el uso de diferencias finitas implica errores de redondeo que pueden dañar seriamente la calidad de las predicciones. Por otro lado, en lugar de transformar las mediciones discretas a derivadas temporales, se

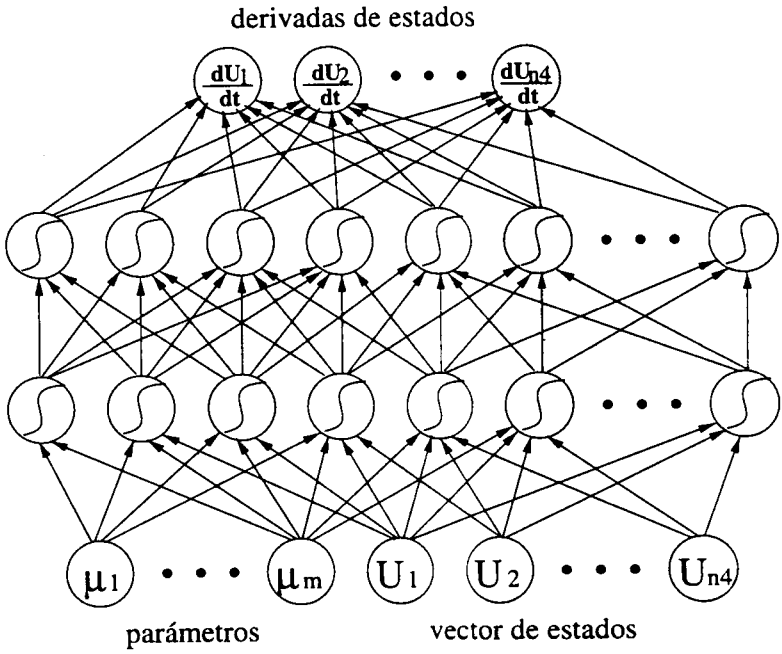


Figura 3.2: *Esquema de RNA que puede ser usada en la identificación de EDOs.*

incrusta la estructura de la RNA en un esquema de integración numérica que puede ser usado para predecir mediciones discretas, las cuales serán comparadas directamente con las mediciones experimentales. El uso de esta alternativa hace más tedioso el entrenamiento, pero se eliminan a un mínimo los errores por redondeo y no requiere de mediciones de derivadas temporales. A continuación se presenta la teoría aplicada a la alternativa de incrustar la RNA en un esquema de integración numérica. Por simplicidad, se utiliza el esquema de integración de Runge-Kutta, pero la metodología no se restringe solamente a este método, puede ser aplicada a cualquier otro método de integración.

3.2.1 Construcción de la Función Error

La transformación de las salidas de la RNA-EDO a una forma discreta puede hacerse a través de un método de integración numérica [8, 6] . Existen diferentes metodologías de integración de sistemas de EDOs, las cuales pueden ser explícitas (por ejemplo Runge-Kutta de cuarto orden) o implícitas (por ejemplo Euler implícito). Estas dos metodologías se explican en el Apéndice C.

Aquí se presenta la metodología sólo para el integrador explícito (Runge-Kutta). La incrustación de la RNA-EDO (Figura 3.2) en dicho integrador da como resultado una macrored como la mostrada en la Figura 3.3. Esta macrored representa la estructura de modelo que se desea identificar y representa un modelo continuo de caja gris, el cual tiene parámetros conocidos en la estructura de integración y parámetros desconocidos (que se desean encontrar) en la estructura de RNA. Ahora las salidas de esta macrored pueden ser comparadas directamente con las observaciones experimentales en cada punto del conjunto de entrenamiento.

La evaluación de la macrored en cada punto del conjunto de entrenamiento se explica a continuación. Primeramente se toman las entradas del conjunto de entrenamiento como entradas en la macrored ($[\boldsymbol{\mu}, \mathbf{U}^t] = \mathbf{I}^t$), despúes se evalúa la macrored y se comparan sus predicciones ($\mathbf{U}^{t+1}$) con el vector de salidas en el conjunto de entrenamiento ($\mathbf{O}^t$). Para un punto cualquiera (t) del conjunto de entrenamiento se obtiene el siguiente error local:

$$E^t(\mathbf{W}) = \frac{1}{2}\sum_{j=1}^{n}\left(O_j^t - U_j^{t+1}(\mathbf{W})\right)^2 \tag{3.12}$$

mientras que con el total de puntos contenidos en el conjunto de entrenamiento se

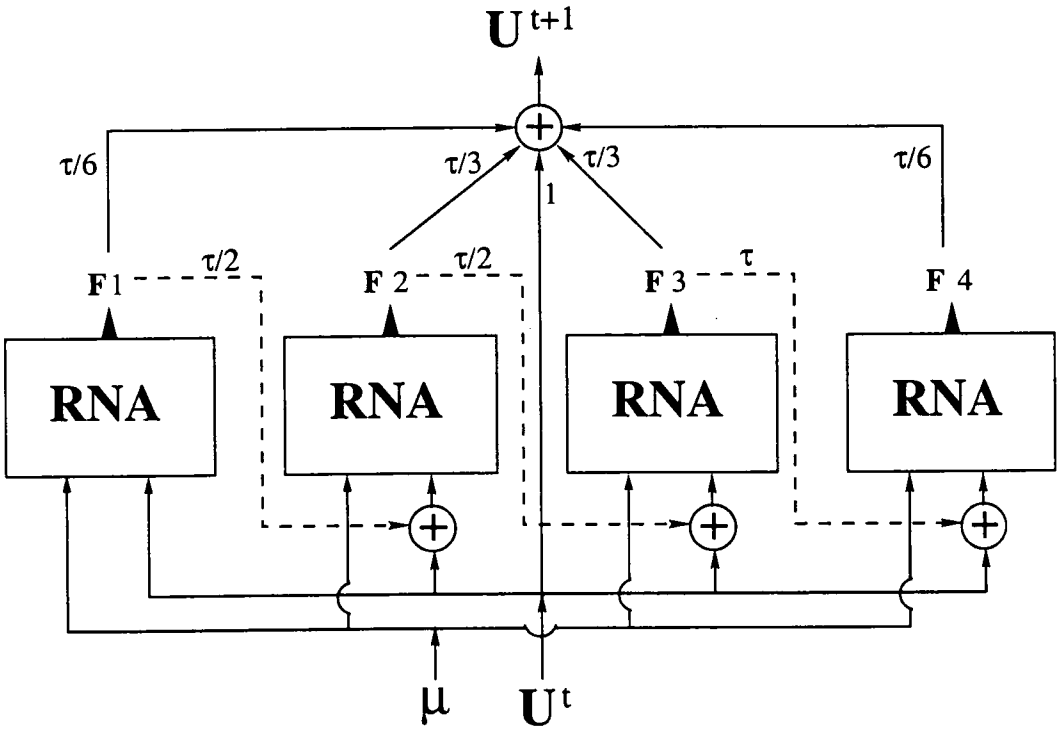


Figura 3.3: *Macrored resultante al incrustar una RNA-EDO en el esquema de integración de Runge-Kutta.*

obtiene el siguiente error global:

$$E(\mathbf{W}) = \sum_{t=1}^{p} E^t(\mathbf{W}) = \frac{1}{2}\sum_{t=1}^{p}\sum_{j=1}^{n}\left(O_j^t - U_j^{t+1}(\mathbf{W})\right)^2 \tag{3.13}$$

donde la sumatoria interna representa la contribución de cada variable de estado y la externa representa la contribución del total de mediciones que se tienen disponibles en el tiempo (tamaño del conjunto de entrenamiento). Por simplicidad y en lo que resta del trabajo, $\mathbf{W}$ representará al conjunto de parámetros ajustables de la RNA (pesos y umbrales). Esta notación facilitará la comprensión de las expresiones para el cálculo de derivadas.

3.2.2 Cálculo de derivadas

Teniendo en cuenta la función error (Ecuación 3.13) y la forma en que se evalúa la macrored (Figura 3.3), se pueden calcular las derivadas. La derivada de la función de energía con respecto a cualquier parámetro ajustable de la RNA (W) se obtiene de la siguiente manera:

$$\frac{dE}{dW} = \frac{1}{2}\sum_{t=1}^{p}\sum_{j=1}^{n}\frac{d\left(O_j^t - U_j^{t+1}\right)^2}{dW} = -\sum_{t=1}^{p}\sum_{i=0}^{n}\left(O_i^t - U_i^{t+1}\right)\frac{dU_i^{t+1}}{dW} \tag{3.14}$$

donde el factor $\frac{dU_i^{t+1}}{dW}$ es representativo de la salida de la macrored y esta influenciado por el número de evaluaciones de la EDO dentro del esquema de integración numérica. Este término se evalúa de la manera siguiente:

$$\frac{dU_i^{t+1}}{dW} = \frac{\tau}{6}\left(\frac{dF1_i^t}{dW} + 2\frac{dF2_i^t}{dW} + 2\frac{dF3_i^t}{dW} + \frac{dF4_i^t}{dW}\right) \tag{3.15}$$

donde las variables Fr_i^t ($r = 1, 2, 3$ ó 4) representan el i-ésimo valor de salida en la RNA-EDO que fue avaluada con el vector de entrada $\mathbf{I}^t$; cada término dentro del paréntesis del lado derecho representa la contribución de la evaluación de la RNA-EDO en el esquema de integración. Estas contribuciones pueden ser evaluadas de la siguiente manera (por comodidad y claridad se omiten los superíndices que se refieren al número de evaluación en el conjunto de entrenamiento)

$$\frac{dFr_i}{dW} = \frac{\partial Fr_i}{\partial W} + Hr\sum_{j=1}^{n}\frac{\partial Fr_i}{\partial \widehat{U}_j}\frac{\partial F(r-1)_j}{\partial W} \tag{3.16}$$

donde

$$Hr = \begin{cases} 0 & \text{para} \quad r = 1 \\ \frac{\tau}{2} & \text{para} \quad r = 2, 3 \\ \tau & \text{para} \quad r = 4 \end{cases} \tag{3.17}$$

además los términos $\frac{\partial Fr_i}{\partial W}$ y $\frac{\partial F(r-1)_h}{\partial W}$ implican el cálculo de la derivada de salida de la RNA-EDO con respecto a los parámetros ajustables (aplicación de la regla Delta), y el término $\frac{\partial Fr_i}{\partial \widehat{U}_j}$ implica las derivadas de la salida de la RNA-EDO con respecto a sus entradas (semejante a la regla Delta). La variable $\widehat{U}$ viene a ser una entrada modificada a la RNA que depende del número de evaluación de la EDO en el esquema de integración. Para aplicar este algoritmo en el cálculo de derivadas de la función error con respecto a cualquier parámetro ajustable sólo es necesario evaluar las derivadas $\frac{\partial Fr_i}{\partial W}$ y $\frac{\partial Fr_i}{\partial \widehat{U}_j}$ que vienen siendo la aplicación de la regla Delta en términos de RNAs.

3.3 SDs espacio-temporales (EDPs)

En la sección previa se ha desarrollado una metodología de RNAs para la construcción de modelos de EDOs. Aquí se propone extender dicha metodología a sistemas que incluyen variaciones espacio-temporales (EDPs) basadas en RNAs acopladas con técnicas de discretización espacial, frecuentemente usadas en métodos de aproximación para la solución de EDPs. La teoría aplicada aquí es para sistemas unidireccionales (una variable espacial) y unidimensionales (una variable de estado), pero puede ser extendida a sistemas multidireccionales y multidimensionales. La Figura 3.4 muestra la estructura de una RNA representativa de un sistema dinámico espacio-temporal como el de la Ecuación 3.3. Las entradas a la red constituyen el vector de parámetros del sistema ($\boldsymbol{\mu}$) y las derivadas espaciales, mientras que la predicción de la RNA es

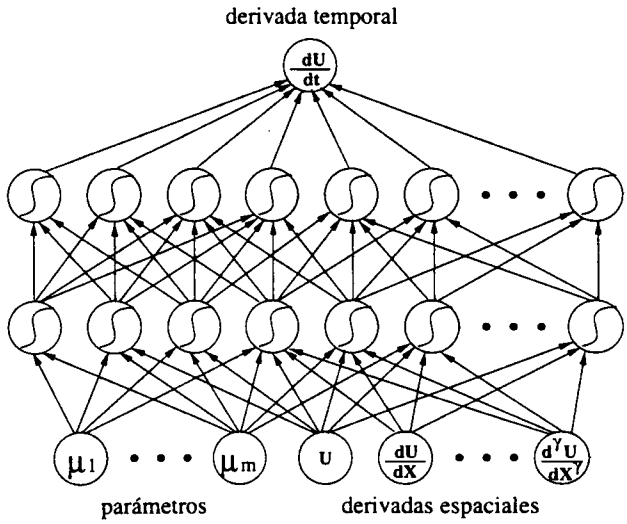


Figura 3.4: *Esquema de RNA que puede ser usada en la identificación de un proceso espacio-temporal unidimensional y unidireccional.*

la derivada temporal. En este caso se supone que las mediciones en el conjunto de entrenamiento se han recolectado a diferentes posiciones en la dirección espacial y a diferentes instantes en el tiempo.

3.3.1 Construcción de la Función Error

Si se supone que se cuenta con un conjunto de entrenamiento compuesto de mediciones atrasadas en el tiempo a diferentes posiciones en la dirección espacial y, si se desea eliminar al máximo los errores por redondeo, entonces la construcción de la función error implica utilizar técnicas de integración numérica para EDPs. La solución numérica de EDPs, tales como la Ecuación 3.3, con frecuencia se hace en dos pasos (método de líneas): 1) se transforma la EDP a un sistema de EDOs por medio de aproximaciones de diferencias finitas en las derivadas espaciales y, 2) se integra el sistema de EDOs resultante por medio de alguna técnica de integración numérica.

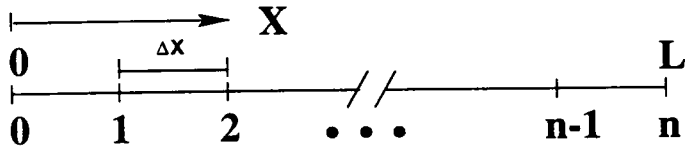


Figura 3.5: *Discretización de la dimensión espacial.*

Para transformar la EDP a un sistema de EDOs, se subdivide el intervalo espacial de interés en "n" sub-intervalos separados por una distancia $\Delta X = \frac{L}{n}$, como lo muestra la Figura 3.5. De esta manera, las derivadas espaciales se pueden aproximar usando diferencias centrales con algún orden de exactitud deseado (ver Apéndice C) por expansión de series de Taylor. Para un orden de error $O(\Delta X^4)$ se tienen las siguientes primeras cuatro derivadas (si se desea, se puede utilizar otro orden de error):

$$\begin{aligned}
\left(\frac{dU}{dX}\right)_i &\approx \frac{U_{i-2}-8U_{i-1}+8U_{i+1}-U_{i+2}}{12\,\Delta X} \\
\left(\frac{d^2U}{dX^2}\right)_i &\approx \frac{-U_{i-2}+16U_{i-1}-30U_i+16U_{i+1}-U_{i+2}}{12\,\Delta X^2} \\
\left(\frac{d^3U}{dX^3}\right)_i &\approx \frac{U_{i-3}-8U_{i-2}+13U_{i-1}-13U_{i+1}+8U_{i+2}-U_{i+3}}{8\,\Delta X^3} \\
\left(\frac{d^4U}{dX^4}\right)_i &\approx \frac{-U_{i-3}+12U_{i-2}-39U_{i-1}+56U_i-39U_{i+1}+12U_{i+2}-U_{i+3}}{6\,\Delta X^4}
\end{aligned} \tag{3.18}$$

Con la introducción de estas aproximaciones en las derivadas espaciales de la Ecuación 3.3, se obtiene un conjunto equivalente de $n+1$ EDOs sobre la variable de estado evaluada en cada subdivisión del dominio espacial:

$$\frac{dU_i}{dt} = F(U_{i-\kappa}, \cdots, U_i, \cdots, U_{i+\kappa}; \boldsymbol{\mu}) \qquad i = 0, \cdots, n \tag{3.19}$$

donde κ es el máximo número de puntos, alrededor del punto de interés i, necesarios para poder representar la derivada espacial de mayor orden (γ) con un determinado error de aproximación. Para el caso de la Ecuación 3.18, κ toma el valor de 3. En

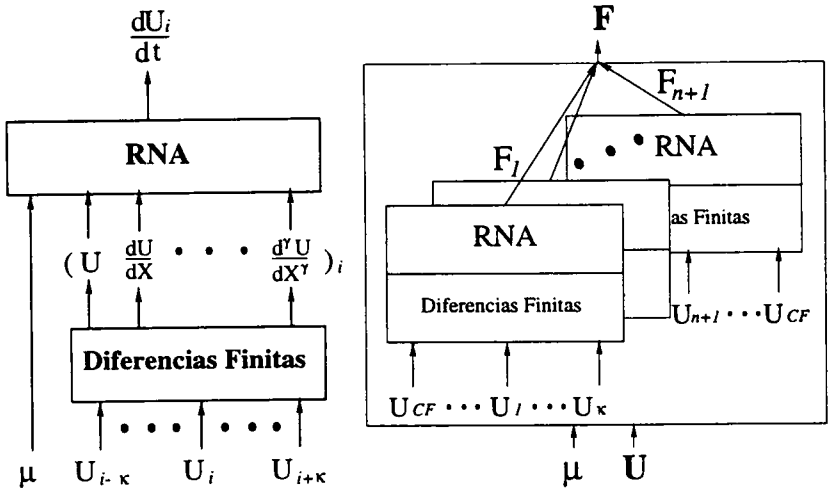


Figura 3.6: *RNA-EDP acoplada con un esquema de diferencias finitas. En (a) se evalúa la derivada temporal en un punto del espacio, y en (b) se evalúa la derivada temporal en todo el espacio.*

las primeras y en las últimas κ ecuaciones es donde entran en juego las condiciones de frontera. Este proceso es esquematizado en la Figura 3.6.

De tal forma que la solución de este sistema de EDOs (y por consiguiente la solución de la EDP original) se puede obtener de la misma manera como se resuelve la Ecuación 3.2. Entonces, al incrustar la RNA y el esquema de diferencias finitas dentro del esquema de integración de Runge-Kutta, se obtendrá una macrored como la mostrada en la Figura 3.7. Esta macrored representa la estructura de modelo que se desea identificar y representa un modelo continuo de caja gris, el cual tiene parámetros conocidos en la estructura de integración y en la estructura de diferencias finitas, mientras que la estructura de RNA-EDP contiene los parámetros ajustables.

La función error con este procedimiento se presenta de forma semejante que con

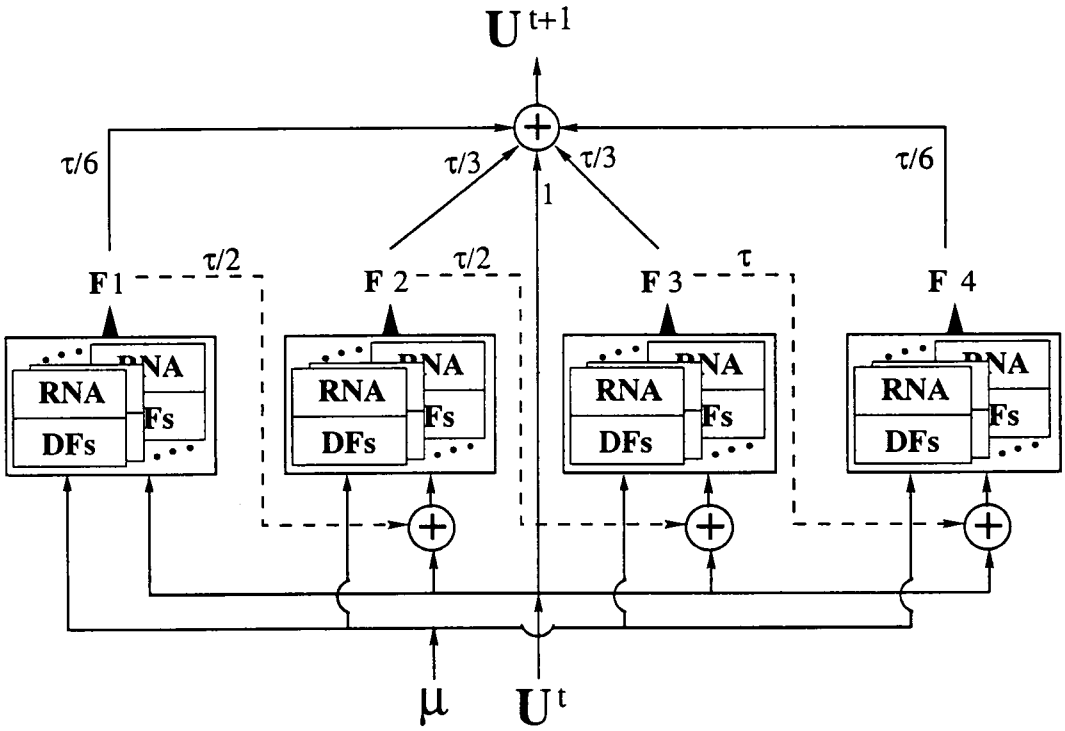


Figura 3.7: *Macrored resultante al incrustar una RNA-EDP en el esquema de integración de Runge-Kutta.*

la RNA-EDO y se muestra a continuación:

$$E(\mathbf{W}) = \frac{1}{2}\sum_{t=1}^{p}\sum_{j=0}^{n}\left(O_j^t - U_j^{t+1}\right)^2 \tag{3.20}$$

donde la sumatoria interna representa la contribución de cada variable de estado (discretizaciones en el espacio) y la externa representa la contribución del total de mediciones que se tienen disponibles en el tiempo (tamaño del conjunto de entrenamiento).

3.3.2 Cálculo de derivadas

La derivada de la función error con respecto a cualquier parámetro (W) de la RNA-EDP es muy semejante a aquella que se presenta en la RNA-EDO. Estas derivadas se pueden calcular de la siguiente manera:

$$\frac{dE}{dW} = -\sum_{t=1}^{p}\sum_{i=0}^{n}\left(O_i^t - U_i^{t+1}\right)\frac{dU_i^{t+1}}{dW} \tag{3.21}$$

donde el factor $\frac{dU_i^{t+1}}{dW}$ esta influenciado por el número de evaluaciones del conjunto de EDOs dentro del esquema de integración y pude ser evaluado como:

$$\frac{dU_i^{t+1}}{dW} = \frac{\tau}{6}\left(\frac{dF1_i^t}{dW} + 2\frac{dF2_i^t}{dW} + 2\frac{dF3_i^t}{dW} + \frac{dF4_i^t}{dW}\right) \tag{3.22}$$

donde cada término dentro del paréntesis del lado derecho representa la contribución de la evaluación del conjunto de EDOs (Ecuación 3.19) en el esquema de integración. Estas contribuciones pueden ser evaluadas como sigue (por comodidad y claridad se omiten los superíndices):

$$\frac{dFr_i}{dW} = \frac{\partial Fr_i}{\partial W} + Hr\sum_{h=i-\kappa}^{i+\kappa}\sum_{j=0}^{\gamma}\frac{\partial Fr_i}{\partial U_{X^j}}\frac{\partial U_{X^j}}{\partial \widehat{U}_h}\frac{\partial F(r-1)_h}{\partial W} \tag{3.23}$$

donde r indica el número de evaluación del conjunto de EDOs en el esquema de integración (toma valores de 1, 2, 3 ó 4), además:

$$Hr = \begin{cases} 0 & \text{para} \quad r = 1 \\ \frac{\tau}{2} & \text{para} \quad r = 2,3 \\ \tau & \text{para} \quad r = 4 \end{cases}$$

En la Ecuación 3.23 los términos $\frac{\partial F_{\tau_i}}{\partial W}$ y $\frac{\partial F(\tau-1)_h}{\partial W}$ implican el cálculo de la derivada de la salida de la RNA-EDP con respecto a los parámetros de la red (aplicación de la regla Delta); el término $\frac{\partial F_{\tau_i}}{\partial U_{Xj}}$ implica las derivadas de la salida de la RNA-EDP con respecto a las derivadas espaciales de entrada (semejante a la regla Delta); el término $\frac{\partial U_{Xj}}{\partial \widehat{U}_h}$ implica las derivadas de las salidas del esquema de diferencias finitas con respecto a las entradas a dicho esquema (matriz de coeficientes de diferencias finitas). La variable $\widehat{U}$ viene a ser una entrada modificada en el esquema de diferencias finitas. De esta forma, sólo es necesario evaluar las derivadas $\frac{\partial F_{\tau_i}}{\partial W}$ y $\frac{\partial F_{\tau_i}}{\partial U_{Xj}}$ para aplicar este algoritmo.

3.4 Método del Gradiente Conjugado

Cuando se tiene una función cualquiera $f = f(\mathbf{U})$ con $\mathbf{U} \in R^n$ y $f \in R^1 \mapsto R^n$, y se desea encontrar el valor de $\mathbf{U} = \widehat{\mathbf{U}}$ que minimice (maximice) el valor de la función $f(\cdot)$, entonces se encuentra ante un problema de optimización multivariable no restringida [13].

Existen diversos métodos de optimización para resolver este problema, la mayoría de los cuales son métodos iterativos de búsqueda. En estos métodos se utiliza la siguiente expresión (mapa) para bucar el vector de parámetros $\widehat{\mathbf{U}}$ que minimice la función objetivo:

$$\mathbf{U}_{t+1} = \mathbf{U}_t + \eta_t \mathbf{P}_t$$

donde $\mathbf{P}$ es un vector que indica la dirección en la cual se hace la búsqueda y η es un escalar que indica el tamaño del paso (*velocidad de entrenamiento* en la literatura de redes neuronales) en dicha dirección. La asignación de un valor a η_t puede hacerse

de una manera arbitraria o puede obtenerse mediante algún método de *optimización unidimensional* [13] (tal como *sección dorada*). Sin embargo, la asignación de un vector de direcciones $\mathbf{P}_t$, representa un esfuerzo mayor. Esta asignación es el principal punto que caracteriza los diferentes *Métodos de Optimización.* En el método del Gradiente Conjugado (utilizado en el presente trabajo) se usa la siguiente expresión para el cálculo de ese vector:

$$\mathbf{P}_t = -\nabla f_t + \beta_t \mathbf{P}_{t-1}$$

donde $\nabla f_t \equiv \nabla f(\mathbf{U}_t)$ es el vector de primeras derivadas (jacobiano) de la función objetivo en la iteración t, β_t es un parámetro que busca garantizar que el vector de direcciones actuales $\mathbf{P}_t$ sea conjugado a todos los vectores de direcciones anteriores $\mathbf{P}_{t-1}, \mathbf{P}_{t-2}, \cdots, \mathbf{P}_1$. La mayoría de los métodos del gradiente conjugado que actualmente se encuentran en la literatura de optimización difieren en la forma de calcular el parámetro β_t, la forma más utilizada (que se usa en el actual trabajo) es:

$$\beta_t = \frac{\nabla f_t \cdot \nabla f_t}{\nabla f_{t-1} \cdot \nabla f_{t-1}} = \frac{\sum_{i=1}^{n} \left(\frac{\partial f}{\partial U_i}\right)_t^2}{\sum_{i=1}^{n} \left(\frac{\partial f}{\partial U_i}\right)_{t-1}^2}$$

donde ∇f_t y ∇f_{t-1} son los respectivos jacobianos de la función f en la iteración t y $t-1$.

En conclusión, el vector de direcciones de búsqueda $\mathbf{P}_t$ queda determinado únicamente con la información de la iteración anterior $(t-1)$ y la información de la iteración actual (t). En la iteración inicial (1) no se conoce la información de la iteración 0, por lo que $\mathbf{P}_t$ se calcula como: $\mathbf{P}_1 = -\nabla f(\mathbf{U}_1)$, y en las subsecuentes iteraciones se

calcula utilizando β. El método se reinicia cada n iteraciones para evitar el deterioro por errores de aproximación, esto significa que el vector de direcciones $\mathbf{P}_t$ se calcula, cada n iteraciones, de la misma forma que en la primera iteración.

El vector $\nabla f(\mathbf{U})$ es la información principal usada en el cálculo de $\mathbf{P}_t$. El cálculo de este vector en el entrenamiento de RNA-SDs ya se ha mostrado en los apartados anteriores . En este caso $\nabla f(\mathbf{U}) = \nabla E(\mathbf{W})$.

3.5 Validación Cruzada

Una cuestión común en el ajuste de funciones es la que concierne al momento en que se debe detener el ajuste para que se generalicen las características del sistema ajustado. En esta sección se describe la técnica de validación cruzada que sirve como criterio de paro en el ajuste de funciones.

Cuando un método iterativo es utilizado para ajustar los parámetros de un modelo a partir de datos experimentales, entonces es necesario saber el momento en que se debe detener el ajuste. El criterio utilizado para detener el proceso de entrenamiento debe tomar en cuenta, por ejemplo, la calidad con que se reproducen las salidas, la cantidad en que se elimina el ruido, si existe o no sobre-entrenamiento (sobreajuste), etc. Este proceso se conoce como *validación del modelo*, y puede llevarse a cabo (en el entrenamiento de RNAs) por las siguientes técnicas: validación cruzada, depuramiento de la red o aún intentando otras estructuras de modelo (RNA) [26, 44, 51].

En los modelos ajustados aquí, se utilizará el método de *Validación Cruzada.* Este método ayuda a estimar la habilidad del modelo, durante el ajuste, para generalizar las predicciones. El método consiste en dividir los datos experimentales en dos conjuntos, uno de ellos es el *conjunto de entrenamiento* (mencionado anteriormente) y el otro

es llamado *conjunto de prueba* o *validación.* El conjunto de entrenamiento es usado para modificar los parámetros del modelo (pesos y umbrales en RNAs) con la finalidad de disminuir el valor absoluto de la función error (esto es garantizado por cualquier método de optimización), mientras que el conjunto de validación es usado para estimar la habilidad que tiene el modelo para generalizar las predicciones.

Cuando se construye una gráfica de la función error (calculada a partir del conjunto de prueba) respecto del número de iteraciones, se observa que es irregular, es decir, tiene muchos picos. El comportamiento de este error (aunque irregular), por lo general, empieza disminuyendo hasta alcanzar un mínimo (global) para después aumentar. Cuando el error de validación alcanza un mínimo global, se dice que la RNA tiene propiedades de *generalización*, es decir, es capaz de reproducir adecuadamente cualquier dato del conjunto de entrenamiento y fuera de éste. Por otro lado, cuando el error de validación comienza a aumentar, se dice que la RNA comienza a *sobreentrenarse* (*sobreajustarse*), es decir, el modelo empieza a capturar características particulares del conjunto de entrenamiento y no será capaz de reproducir adecuadamente cualquier dato fuera del conjunto de entrenamiento. De esta manera, el mejor momento para detener el proceso de entrenamiento es cuando se haya alcanzado un mínimo global en el error de validación. Por lo cual, el error de validación sirve como criterio para detener el entrenamiento en el momento en que se ha alcanzado un mínimo global.

4. Caracterización de OMMs

Aquí se presentan los resultados obtenidos en la identificación, y posterior caracterización, de observaciones experimentales del proceso de oxidación galvanostática de hidrógeno en presencia de iones cloro. Las observaciones experimentales se describen en el Apéndice A. Dichas observaciones se tienen presentes en forma de series de tiempo de una sola variable de estado (el potencial, E) en un escenario de dos parámetros de operación (concentración de iones cloro, Cl, y la intensidad de corriente, I). El comportamiento presentado en estas observaciones es del denominado *Oscilaciones de Modo Mixto* (OMMs).

Los esfuerzos en la identificación de este sistema con modelos basados en principios básicos (balances de materia, de energía, etc.) es un tópico de investigación activo, aunque todavía no se han obtenido resultados satisfactorios. Por el momento los modelos (y mecanismos) propuestos sólo predicen algunas ocurrencias de OMMs en valores de parámetros muy diferentes a los que se observan experimentalmente [21, 66, 33]. Aquí se pretende caracterizar la oxidación de hidrógeno con modelos basados en *Redes Neuronales Artificiales* (RNAs), los cuales puedan ayudar a complementar la exploración experimental y guiar hacia el entendimiento de los mecanismos generadores del comportamiento de OMMs.

En la primera parte del capítulo se presenta la identificación con un modelo dinámico discreto (mapa) y en la segunda parte se presenta con un modelo continuo de *Ecuaciones Diferenciales Ordinarias* (EDOs). Al aplicar dos procedimientos de caracterización independientes, se tendrán bases para corroborar los resultados. Independientemente del tipo de modelo (discreto o continuo), la caracterización de las transiciones e inestabilidades implícitas en las observaciones experimentales es lograda con el uso de modelos dinámicos enfocados en la predicción de soluciones a largo plazo. En este sentido, la predicción de soluciones a largo plazo de series de tiempo juega un papel muy importante en la exitosa caracterización del comportamiento observado. Es decir, aunque las soluciones a largo plazo del sistema real y el modelo pueden diferir punto por punto (de hecho son inexactas para series de tiempo caóticas), el modelo ajustado todavía se puede considerar exitoso si los atractores predichos se presentan en una zona del espacio-fase cercana a los atractores reales. De esta forma, el atractor del modelo aproxima con muy buena exactitud ciertas propiedades estadísticas del atractor real (amplitud, espectro de potencia y dimensión entre otras) y además, el punto clave de este trabajo, puede ayudar a entender la naturaleza cualitativa de las transiciones (bifurcaciones) como una función de los parámetros de operación.

El cálculo de soluciones a partir de modelos requiere del uso de herramientas analíticas y computacionales de la teoría de sistemas dinámicos. En el presente trabajo se hace uso de paquetes de software para el cálculo de las soluciones de los modelos RNA-SDs construidos. Esto significa que los modelos RNA-SDs son acoplados con cada paquete de software utilizado. En particular se utilizan los paquetes DSTOOL [16] y SCIGMA [56] para analizar el comportamiento dinámico en el espacio-fase, y AU-

TO [10] para el cálculo de los diagramas de bifurcaciones (de uno y dos parámetros).

4.1 Construcción de un modelo discreto

Aquí se aplica una metodología para la construcción de modelos discretos (mapas) a partir de mediciones experimentales de una sola variable de estado en un sistema continuo. La metodología va acompañada de dos pasos: un procesamiento previo de los datos experimentales y la construcción del modelo dinámico.

4.1.1 Procesamiento de datos experimentales (extracción de MPMs)

Cuando se desea reproducir, por medio de modelos discretos, la dinámica a largo plazo y la secuencia de bifurcaciones en sistemas experimentales con dinámica compleja, las aproximaciones usuales basadas en mediciones atrasadas [55], generalmente son inapropiadas para tal efecto. Esto se debe a que el sistema experimental (aunque muestreado en forma discreta) presenta carácter continuo, mientras que el modelo construido es de carácter discreto [47, 46]. En tales casos, las representaciones discretas (tales como MPs o MPMs), extraídas de sistemas continuos a partir de series de tiempo experimentales, pueden ser usadas para dichos fines. En esta sección se utiliza la discretización por medio de MPMs ya que con ellos se puede obtener una caracterización exitosa de un gran número de transiciones y por su simplicidad son un punto de inicio en los esfuerzos de caracterización de escenarios de transiciones complejas, como el que se presenta en un régimen de OMMs.

La extracción de MPMs (como fue explicado en el Capítulo 2) se aplicó a todas las series de tiempo experimentales disponibles de la oxidación galvanostática de hidrógeno (Apéndice A). Algunos de los MPMs obtenidos se muestran en la Figura

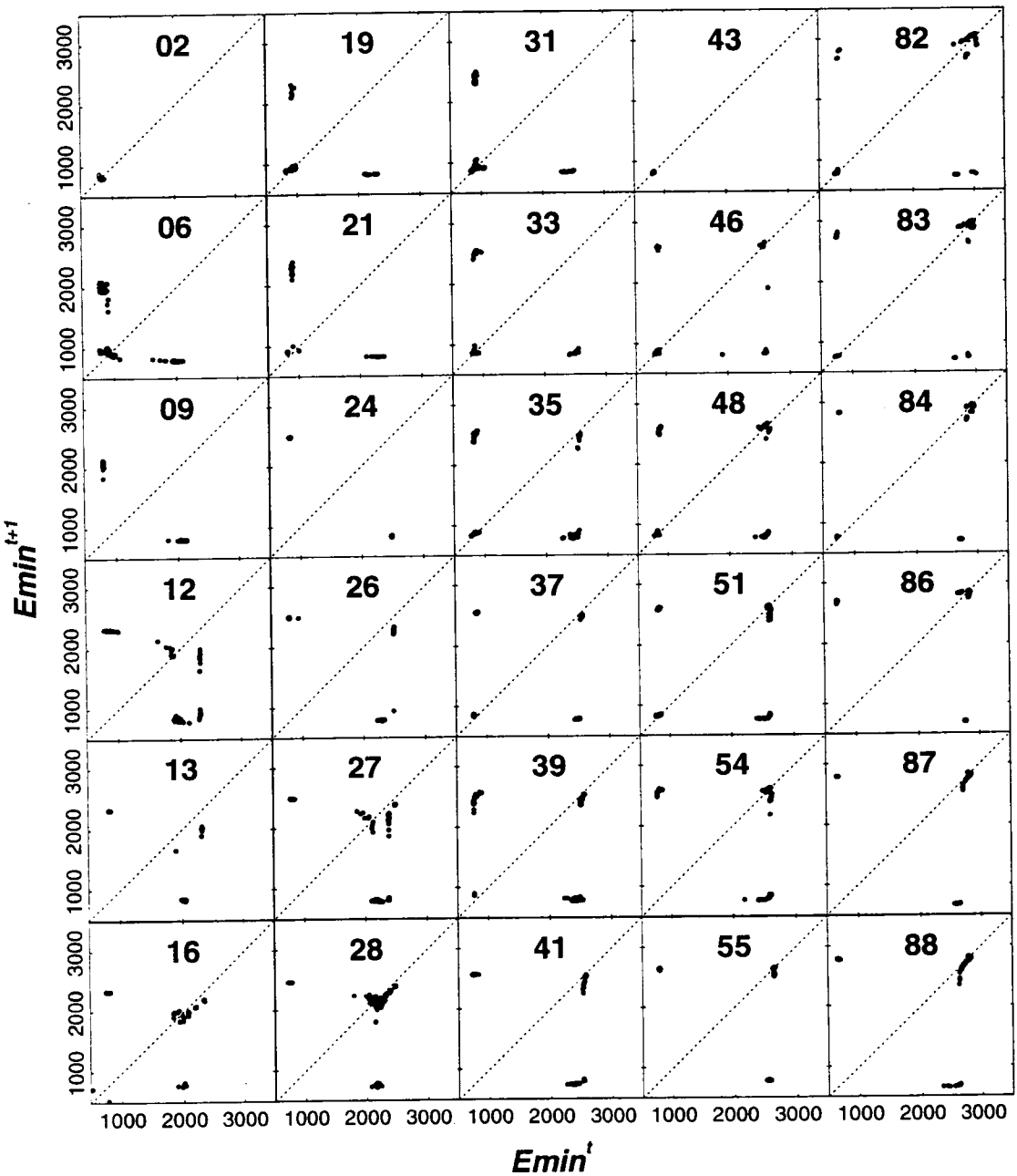


Figura 4.1: *MPMs obtenidos a partir de series de tiempo experimentales de la oxidación de hidrógeno. Ver en el Apéndice A la nota sobre las unidades del potencial.*

4.1. Si se unen todos los puntos de cualquier MPM, se tendrían curvas de dos extremos (parecidas a la letra N), las cuales son más notables en los mapas caóticos. Dichas curvas son características del comportamiento de OMMs representado en MPMs [48, 28].

4.1.2 Modelo de identificación

Para los fines de identificación de OMMs con los MPMs disponibles, se construye un modelo dinámico discreto unidimensional de la forma:

$$Em_{t+1} = F(Em_t; Cl, I) \tag{4.1}$$

donde Em_t y Em_{t+1} son dos valores mínimos sucesivos de potencial en una serie de tiempo experimental, Cl e I son los valores de los parámetros de operación y representan la concentración de iones cloro ($[Cl^-]$) y la intensidad de corriente aplicada, respectivamente. La función del lado derecho, $F(Em_t; Cl, I)$, representa una estructura de modelo no-lineal y será aproximada por un mapa basado en RNAs, al cual se le denominará RNA-MPM. El entrenamiento de esta RNA se lleva a cabo de la forma como se explica en la Sección 3.1.

Los resultados presentados aquí, fueron obtenidos con un identificador que tiene la configuración de RNA que se presenta en la Figura 4.2. Dicho identificador consistío de una RNA estándar de alimentación hacia adelante con 4 capas (dos capas ocultas). La capa de entrada consistió de una neurona para la variable de estado (Em_t) y dos neuronas para los parámetros de operación del sistema (Cl e I). En la capa de salida se tuvo una neurona con función de activación lineal, para predecir el valor futuro de

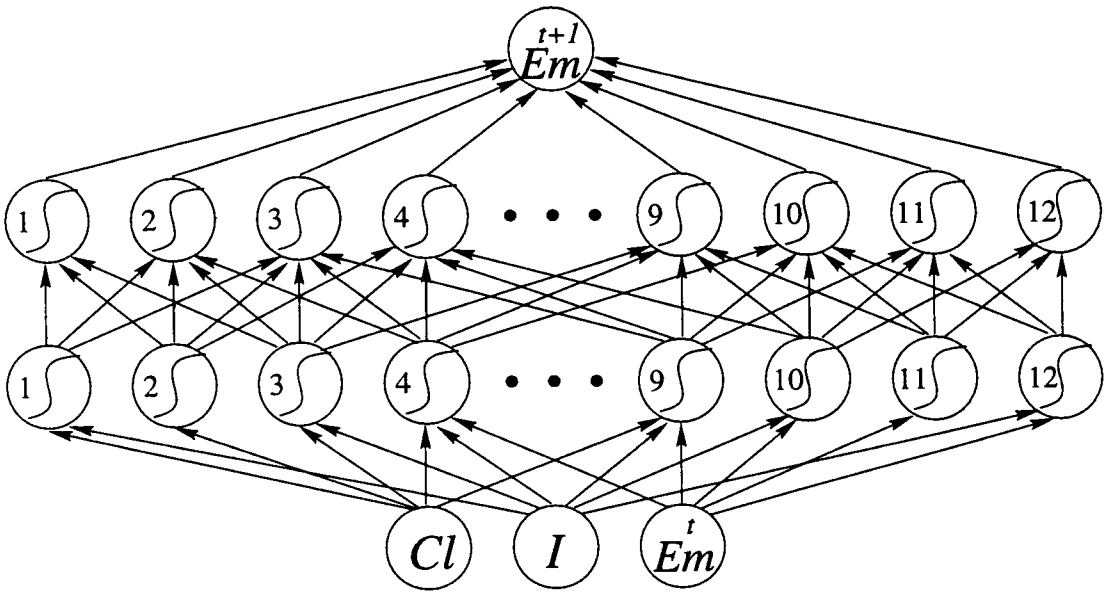


Figura 4.2: *Estructura de la RNA que fue utilizada en la construcción de un MPM:* $Em_{t+1} = F(Em_t; Cl, I)$. *La red debe estar completamente conectada hacia adelante, aunque en la figura sólo se muestran unas cuantas conexiones (las siguientes RNAs se presentan de igual manera).*

potencial mínimo (Em_{t+1}). Cada capa oculta consistió de 12 neuronas con función de activación no-lineal del tipo sigmoidal.

Antes de elegir la configuración final, se intentaron diferentes estructuras de RNAs. Todas estas configuraciones fueron de RNAs de alimentación hacia adelante de cuatro capas (dos capas ocultas con el mismo número de neuronas cada una); función de activación no lineal del tipo sigmoidal; una neurona en la capa de salida; y tres neuronas en la capa de entrada. La diferencia entre las todas las configuraciones intentadas, fue el número de neuronas en las capas ocultas. Se intentaron configuracines con 4, 6, 8, 9, 10, 11 y 12 neuronas. Las RNAs con pocas neuronas ocultas (4, 6 y 8 neuronas) presentaron pobres capacidades predictivas, en lo referente al

escenario de bifurcaciones. Mientras que, las RNAs con mayor número de neuronas ocultas presentaron mejores capacidades predictivas, y el tiempo de entrenamiento fue mayor. De esta manera, se escogió una RNA con 12 neuronas ocultas, con la cual se obtuvieron las predicciones de mayor calidad. Al tiempo de entrenamiento se le dió poco peso, puesto que todas las arquitecturas intentadas presentan tiempos de entrenamiento relativamente pequeños.

El proceso de entrenamiento de la RNA-MPM se llevó a cabo de la forma como se explicó en la Sección 3.1. El número de vectores usados en el conjunto de entrenamiento se eligió tomando en consideración un criterio basado en el compromiso que existe entre la calidad de las predicciones a largo plazo y el tiempo de entrenamiento, lo cual condujo a probar varios conjuntos de entrenamiento. De esta manera, el conjunto de entrenamiento final (Tabla 4.1) fue construido con datos de 41 diferentes series de tiempo (ya discretizadas como MPMs) de un total de 61 series disponibles (ver la Figura A.2); y el conjunto de prueba (Tabla 4.2) fue construido con 20 diferentes series de tiempo. Cada vector en el conjunto de entrenamiento (prueba) consistió de 4 elementos (3 de entradas y uno de salida). Estos datos fueron escaladas en el rango $[-0.9, +0.9]$ debido a que la función de activación de la RNA requiere de pesos infinitamente grandes para alcanzar valores de -1.0 ó $+1.0$.

La elección de las series de tiempo utilizadas, en el conjunto de entrenamiento, y el número de vectores tomados desde cada serie elegida, se llevó a cabo tomando los siguientes criterios. Estos criterios se enumeran en orden de prioridad.

1. Abarcar todo el espacio de parámetros.

2. Tomar un número igual de series por cada corte de concentración de iones cloro.

Tabla 4.1: *Información de los datos experimentales usados para construir el conjunto de entrenamiento de la RNA-MPM. Consta de 2100 puntos.*

Cl=0.004M		Cl=0.005M		Cl=0.006M		Cl=0.0065M		Cl=0.010M	
Exp	pts.	Exp.	pts.	Exp.	pts.	Exp.	pts.	Exp.	pts.
01	25	17	25	29	19	43	25	80	19
03	40	18	75	30	60	44	75	81	50
04	75	20	50	33	50	48	50	82	59
08	65	22	50	35	45	50	60	84	78
10	25	24	25	36	65	52	60	86	70
12	70	26	40	37	30	54	50	87	70
13	40	27	75	40	40	55	45	88	75
14	55	28	75	41	40	56	66		
16	42			42	70				

Tabla 4.2: *Información de los datos experimentales usados para construir el conjunto de validación de la RNA-MPM. Consta de 450 puntos.*

Cl=0.004M		Cl=0.005M		Cl=0.006M		Cl=0.0065M		Cl=0.010M	
Exp.	pts.	Exp.	pts.	Exp.	pts.	Exp.	pts.	Exp.	pts.
00	25	18	15	32	25	44	15	81	15
06	25	21	25	34	25	46	25	83	25
09	25	23	25	39	25	51	25	85	25
15	65	25	25	41	15	53	25	88	15

3. Tomar las series de tiempo que presentaron OMMs bien definidas. Por ejemplo, las OMMs 1^0, 1^1, $2^1 1^1$, 1^3, etc.

4. Tomar las series de tiempo con menos ruido. Por ejemplo, las series de tiempo 09 y 10 representan OMMs 1^1 en el corte $Cl = 0.004$ M, de las cuales se eligió la serie 10 que tiene menos ruido.

5. Incluir al menos dos periodos en cada serie.

6. Incluir algunas ocurrencias caóticas.

El proceso de entrenamiento se inició asignando los parámetros de la RNA aleatoriamente en un rango de $[-0.2, +0.2]$. Se entrenaron varias RNAs con diferentes parámetros de inicio. La que mejores resultados presentó, requirió de aproximadamente 20 ciclos completos del gradiente conjugado (cada ciclo consistió de 217 pasos o iteraciones) para alcanzar convergencia en el entrenamiento. El error (estandarizado) al final del entrenamiento fue de 0.3137. Este error normalizado se consigue al aplicar la siguiente transformación a la función error (Ecuación 3.4):

$$e = \sqrt{\frac{2E}{n \times p}} \tag{4.2}$$

donde n y p son el número de variables de estado y número de datos en el conjunto de entrenamiento, respectivamente.

4.1.3 Análisis de resultados

La Figura 4.3 muestra una comparación cualitativa, mediante retratos en el espacio-fase del MPM, de los atractores periódicos experimentales y los atractores predichos

por la RNA-MPM. En la figura se muestran las predicciones a largo plazo una vez que se ha convergido en los atractores. La RNA-MPM predice exitosamente la mayoría de los atractores periódicos de OMMs observados experimentalmente. El número y posición de los puntos (periodos) en el MPM son predichos razonablemente, así como el ordenamiento de las OMMs y el rango de los valores de parámetros (Cl e I) en el cual se observaron originalmente cada uno de ellos. La RNA-MPM es capaz de predecir la presencia de una secuencia de Farey, la cual se describe más adelante, así como los detalles adicionales del escenario de transiciones en la vecindad del rango de valores de parámetros de interés.

La Figura 4.4 resume un escenario de bifurcaciones en dos parámetros (I contra Cl), el cual es predicho por la RNA-MPM. En la figura se muestran los estados principales de OMMs (1^0, 1^1, 0^1) que predice la RNA-MPM. Estos estados de OMMs están ordenados de acuerdo a la aritmética de Farey. Las regiones que se marcan como "Transición" representan la predicción de transiciones entre los estados principales de OMMs (notados en la figura) e involucran bifurcaciones globales (se sabe que son del tipo intermitencia I [15]) y locales que dan lugar a la presencia de otros regímenes de OMMs. Las bifurcaciones globales no pueden ser estudiadas con este tipo de modelos, ya que el modelo no representa toda la dinámica global en el espacio-fase del sistema original. Las bifurcaciones globales sólo pueden ser estudiadas con un modelo dinámico continuo.

Los estados principales de OMMs y las regiones marcadas como "Transición" están delimitados por líneas continuas o discontinuas. Las líneas continuas representan la aparición o desaparición de los estados principales de OMMs por medio de bifurcaciones de doble periodo y las líneas discontinuas representan bifurcaciones de

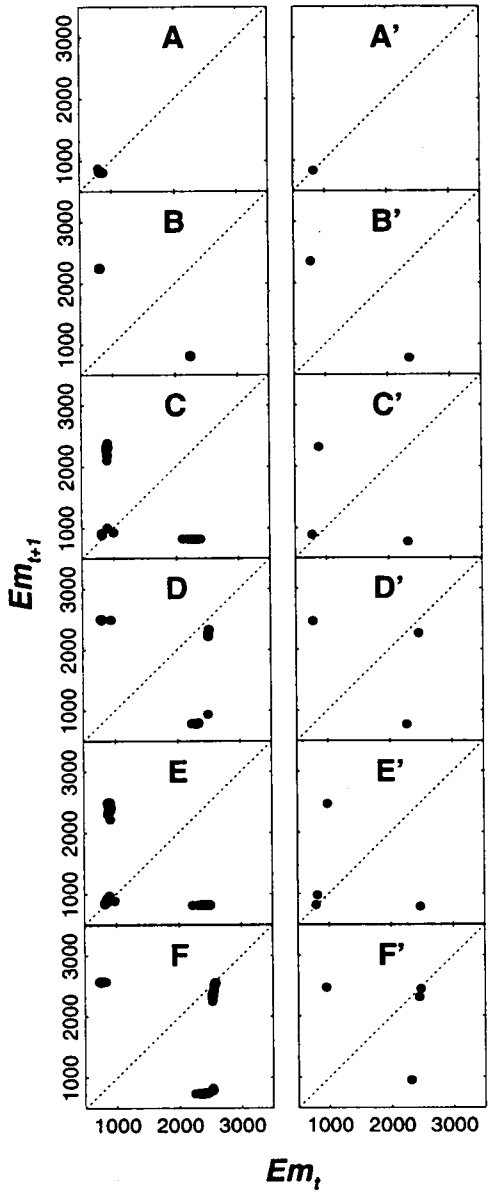


Figura 4.3: *Comparación de MPMs experimentales contra los predichos por la RNA-MPM. En la columna izquierda se muestran atractores experimentales y en la derecha, las predicciones de la RNA-MPM. El valor de los parametros es: A(Cl = 0.004 M, I = 1.3 mA), B(0.004, 1.11), C(0.005, 1.5), D(0.005, 1.21), E(0.006, 1.71), F(0.006, 1.32), A'(0.004, 1.3), B'(0.004, 1.11), C'(0.005, 1.475), D'(0.005, 1.16), E'(0.006, 1.6482) y F'(0.006, 1.135).*

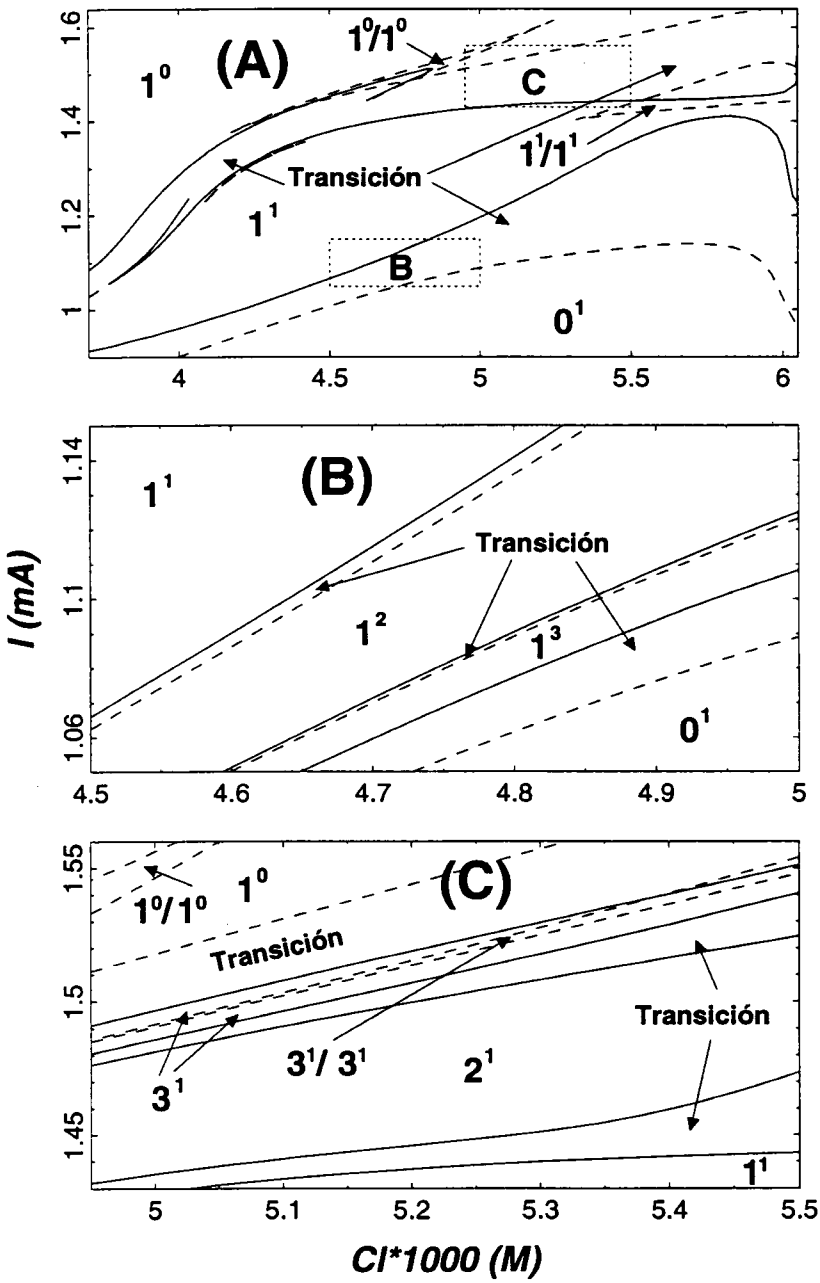


Figura 4.4: *Diagramas generales de bifurcaciones en dos parámetros (I contra Cl) predichos por la RNA-MPM. En (A) se presenta el diagrama en todo el espacio de parámetros de interés, mientras que en (B) y (C) se presenta en más detalle algunas regiones de (A).*

nodo-punto de silla. Los números elevados a una potencia y separados por una diagonal representan soluciones múltiples de OMMs, es decir, que se puede llegar a una u otra solución (pero no a ambas) dependiendo de las condiciones iniciales. Estos son algunos de los aspectos que no son fácilmente discernibles a partir de los datos experimentales.

En la Figura 4.4A se observa que a valores altos de I se tienen sólo OMMs 1^0, es decir, el sistema exhibe oscilaciones simples de gran amplitud; a valores bajos de I se tienen sólo OMMs 0^1, es decir, el sistema exhibe oscilaciones simples de pequeña amplitud; y en la región intermedia se tienen OMMs 1^1, es decir, el sistema exhibe una oscilación de amplitud grande alternada con una oscilación de amplitud pequeña. Este primer resultado concuerda satisfactoriamente con aquellos presentados en la literatura para OMMs [48, 29].

En la Figura 4.4B se presenta en forma más detallada el comportamiento de la zona superior de transición de la Figura 4.4A. Al variar I desde valores altos hasta valores bajos se presentan diferentes regiones de estados principales de OMMs, ordenándose de acuerdo a la aritmética de Farey (1^1, 1^2, 1^3, ... ,0^1). Entre dos de estos estados principales de OMMs se presentan regiones de transición que inician con una cascada de bifurcaciones de doble periodo, pasando por otros estados secundarios de OMMs y terminan en una bifurcación de nodo-punto de silla (en concordancia con los escenarios sugeridos por Ringland y col. [48] y Koper y Gaspard [29]). La región de transición localizada en la parte más baja de la figura presenta otros estados principales de OMMs del tipo 1^n ($n > 3$) separados entre sí por regiones de transición.

En la Figura 4.4C se presenta en forma más detallada el comportamiento de la zona inferior de transición de la Figura 4.4A. Aquí, los estados principales aparecen

(en valores bajos de I) y desaparecen (en valores altos de I) por medio de bifurcaciones de doble periodo. Este comportamiento es debido a que dentro de la región del estado principal de OMM (por ejemplo la OMM 3^1) se encuentra una zona de multiplicidad de soluciones o histéresis (como la $3^1/3^1$ presentada dentro del estado principal 3^1) que complica la determinación del ordenamiento de las transiciones. Sin embargo, los estados principales y secundarios de OMMs que se muestran en este diagrama siguen apareciendo ordenados de acuerdo a una secuencia de Farey.

La construcción de modelos empírico (basados en RNAs) como el presentado aquí, permite la caracterización del comportamiento de sistemas dinámicos no-lineales. Los méritos de este tipo de metodologías han sido ilustrados para un sistema electroquímico típico en un régimen con soluciones del tipo de oscilaciones de modo mixto. La concordancia entre el comportamiento dinámico experimental y el predicho por la red es cualitativamente muy buena en una región amplia del espacio de dos parámetros (I contra Cl), aunque sólo existe concordancia cuantitativa para algunas regiones de este espacio.

Con las representaciones discretas (como las usadas aquí), sólo es posible estudiar bifurcaciones locales en los sistemas que presentan comportamiento periódico complejo. Esta manera de atacar el problema no permite estudiar las bifurcaciones de estados estacionarios ni las bifurcaciones globales. Una descripción completa del comportamiento dinámico (incluyendo bifurcaciones de estados estacionarios y estructuras en el espacio fase) sólo se puede capturar con un modelo dinámico continuo (modelo de ecuaciones diferenciales ordinarias) como el que se presenta a continuación.

4.2 Construcción de un modelo continuo

Aquí se aplica una metodología para la construcción de modelos continuos (conjunto de EDOs) a partir de mediciones experimentales de una sola variable de estado en un sistema continuo. La metodología se presenta en dos etapas: un procesamiento previo de los datos y la construcción del modelo. La primera etapa (procesamiento) es equivalente a la reconstrucción del comportamiento del sistema en el espacio-fase [55, 43]. La segunda etapa resulta al suponer que las series de tiempo obtenidas de la reconstrucción del espacio-fase, son el resultado de la integración numérica de un conjunto de EDOs que describen el comportamiento dinámico observado experimentalmente. De esta manera el lado derecho del conjunto de EDOs puede ser aproximado usando RNAs, tal como se explicó en el Capítulo 3.

4.2.1 Reconstrucción del espacio-fase (obtención de CPNLs)

Para reconstruir el comportamiento del sistema en el espacio-fase, partiendo de las series de tiempo de una sola variable de estado, se hace uso de la técnica denominada análisis de *Componentes Principales No-Lineales* (CPNLs) [9, 30, 52, 59]. Esta técnica es similar a la técnica estándar de *Componentes Principales Lineales* (CPLs) en procesamiento de señales propuesta por Broomhead y King [5]. Para el análisis de CPLs se dispone de una teoría establecida y bien desarrollada, mientras que el análisis de CPNLs esta basado, principalmente, en reglas heurísticas y propiedades muy generales (aproximadores universales) de RNAs. Por ello, la manera en que el análisis de CPNLs relaciona el espacio-fase original no esta bien entendida, sin embargo, la calidad de la reconstrucción es muy superior a la obtenida con el uso de CPLs.

El procedimiento usual, en el análisis de CPNLs, involucra entrenar una RNA de

alimentación hacia adelante para aprender el *mapa identidad*, es decir, que aproxime sus propias entradas (*auto-asociación*). Dicha red consiste de 5 capas, de las cuales, la capa intermedia (conocida como *cuello de botella*) tiene función de activación lineal y un número de nueronas menor al número de neuronas en la capa de entrda (y salida). De esta manera, cuando se entrena la red, el número de neuronas en la capa cuello de botella representa el número de CPNLs significantes, mientras que la capa de salida constituye los datos filtrados.

Los resultados presentados aquí, fueron obtenidos con un identificador que tiene la configuración de RNA que se presenta en la Figura 4.5. A esta RNA se le denominará RNA-CPNL. La elección de este identificador se hizo probando diferentes estructuras de RNA (número de neuronas en cada capa), todas ellas de alimentación hacia adelante con 5 capas (tres capas ocultas). Las capas de entrada y salida consistieron de un vector de 20 elementos y el objetivo del entrenamiento de esta red fue aproximar el mapa identidad (auto-asociación). Es decir, los vectores de entrada/salida fueron construidos con el mismo segmento de 20 mediciones atrasadas (19 intervalos de tiempo) de la serie de tiempo, lo cual representó un segmento de 0.79167 segundos (ver el Apéndice A para los detalles de la velocidad de muestreo). Este segmento de 20 elementos se escogió de una manera empírica. Las capas ocultas de la RNA final estuvieron formadas de la siguiente manera: la primera (capa inferior) sirvió para transformar los segmentos de la serie de tiempo en CPNLs (codificador) y consistió de 10 neuronas con función de activación no-lineal del tipo sigmoidal; la segunda (capa intermedia o cuello de botella) consistió de 3 neuronas con función de activación lineal y representó el espacio-fase de CPNLs; la tercera (capa superior) sirvió para transformar los CPNLs en segmentos de series de tiempo (decodificador) y consistió

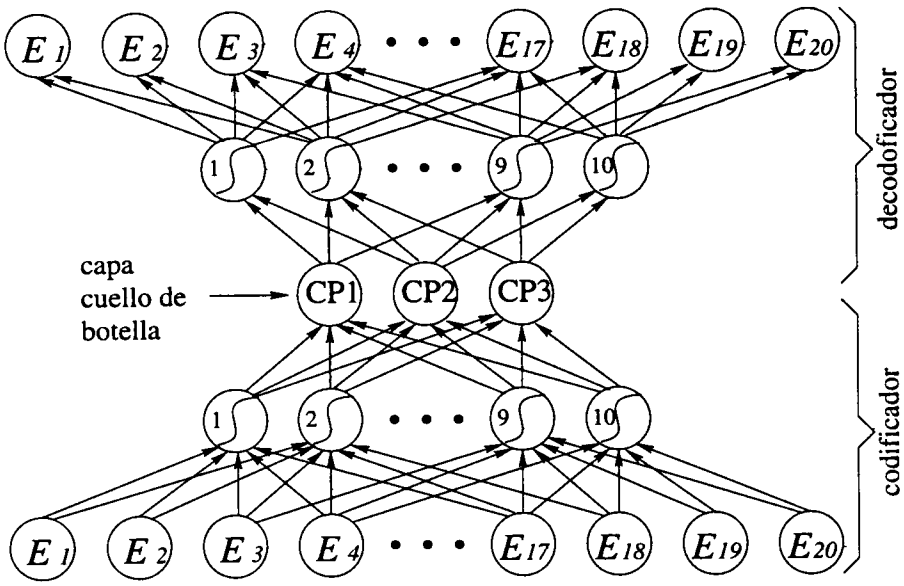


Figura 4.5: *Estructura de RNA utilizada para construir el espacio-fase en CPNLs.*

de 10 neuronas con función de activación no-lineal del tipo sigmoidal.

Los valores de salida de las neuronas en la capa cuello de botella fueron los CPNLs buscados (obtenidos una vez que se alcanzó convergencia en el entrenamiento de la red). El número de neuronas en dicha capa (tres) se escogió tomando en cuenta que se necesita al menos un espacio-fase tridimensional para poder reproducir el comportamiento de OMMs. Se construyeron series de tiempo en CPNLs "deslizando", el vector de entrada (y salida) de la RNA-CPNLs, 3 intervalos de muestreo por vez. El deslizamiento estuvo basado en criterios empíricos. Esto representa un intervalo de tiempo de 0.125 segundos entre puntos de la serie de tiempo de los CPNLs. Es decir, los CPNLs pueden suponerse como el resultado de la integración de un conjunto de EDOs con un incremento de tiempo de 0.125 segundos.

El proceso de entrenamiento de la RNA-CPNL se llevó a cabo de la forma como se explicó en la Sección 3.1, para el entrenamiento de mapas. El número de vectores

Tabla 4.3: *Información de los datos experimentales usados para construir el conjunto de entrenamiento de la RNA-CPNL. Consta de 2225 puntos.*

$Cl = 0.004$ M		0.005 M		0.006 M	
Exp.	pts.	Exp.	pts.	Exp.	pts.
01	100	17	75	29	75
09	100	20	200	32	250
10	100	21	150	34	150
13	150	24	100	40	130
		26	120	41	150
		27	175	42	200

usados en el conjunto de entrenamiento se eligió tomando en consideración un criterio basado en el compromiso que existe entre la calidad (eliminación de ruido y forma cualitativa de los atractores) de los atractores reconstruidos y el tiempo de entrenamiento. Además se tomaron en cuenta la calidad de las predicciones obtenidas con la RNA-MPM del apartado anterior. Estos criterios condujeron a probar varios conjuntos de entrenamiento. De esta manera, el conjunto de entrenamiento final (Tabla 4.1) fue construido con datos de 16 diferentes series de tiempo de un total de 61 series disponibles (ver la Figura A.2). El conjunto de prueba se construyó con sólo 3 series de tiempo, debido a que este conjunto fue usado sólo para dar un estimado (no fue definitivo) del punto en el que se debería detener el entrenamiento. Cada vector en el conjunto de entrenamiento (prueba) consistió de 20 elementos (ya que el vector de entradas es exactamente igual al de salidas). Estos datos fueron escalados en el rango $[-0.9, +0.9]$ debido a que la función de transferencia nodal de la RNA requiere de pesos infinitamente grandes para alcanzar valores de -1.0 ó $+1.0$.

La elección de las series de tiempo utilizadas, en el conjunto de entrenamiento, y el número de vectores tomados de cada serie elegida, se llevó a cabo tomando los siguientes criterios. Estos criterios se enumeran en orden de prioridad.

1. Tomar solamente series de tiempo pertenecientes a los cortes de cloro 0.004, 0.005 y 0.006 M. En estos cortes se obtienen resultados de mayor calidad (basado en la predicciones del modelo RNA-MPM), ya que se tienen series experimentales mejor definidas.

2. Tomar las series de tiempo que presentarón OMMs bien definidas. Por ejemplo, las OMMs 1^0, 1^1, $2^1 1^1$, 1^3, etc.

3. Tomar las series de tiempo con menos ruido. Por ejemplo, las series de tiempo 09 y 10 representan OMMs 1^1 en el corte $Cl = 0.004$ M, de las cuales se eligió la serie 10 que tiene menos ruido.

4. Tomar un número igual de periodos.

El proceso de entrenamiento se inició asignando los parámetros de la RNA aleatoriamente en un rango de $[-0.2, +0.2]$. Se entrenaron varias RNAs con diferentes parámetros de inicio. La que mejores resultados presentó, requirió de 50 ciclos completos del gradiente conjugado (cada ciclo consistió de 503 pasos o iteraciones) para alcanzar convergencia en el entrenamiento. El error (estandarizado) al final del entrenamiento fue de 0.008375 (calculado con las Ecuaciones3.4 y 4.2).

La RNA-CPNL, ya entrenada, se utilizó para extraer las series de tiempo en CPNLs. Para esto se utilizó solamente la parte inferior de la RNA-CPNLs (codificador). En la Figura 4.6 se representan algunos resultados obtenidos al extraer los CPNLs a partir de las series de tiempo experimentales. Los resultados obtenidos quedaron escalados en un rango $[-0.9, +0.9]$. Las gráficas se presentan en el espacio-fase del CPNL1 contra el CPNL2.

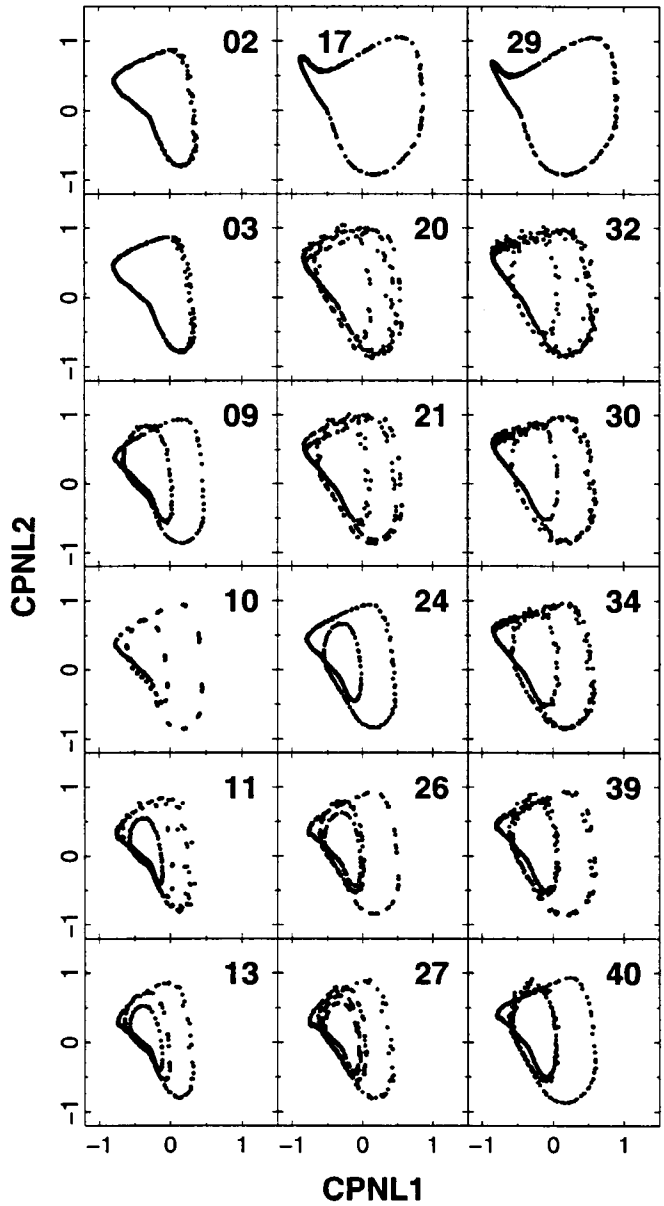


Figura 4.6: *Algunas series de tiempo experimentales representadas en el espacio-fase de los CPNLs. La columna de la izquierda son atractores para $Cl = 0.004$ M; la columna central, para $Cl = 0.005$ M; y la derecha, para $Cl = 0.006$ M. Los valores de I son como sigue:* 02($I = 1.3$); 03(1.26); 09(1.16); 10(1.11); 11(0.91); 13(0.871); 17(2.5); 20(1.52); 21(1.5); 24(1.32); 26(1.21); 27(1.07); 29(2.5); 32(1.73); 30(1.71); 34(1.65); 39(1.49); 40(1.44)

4.2.2 Modelo de Identificación

Con las series de tiempo de los CPNLs se procede a construir un modelo dinámico continuo tridimensional de la forma:

$$\begin{aligned} \frac{d\mathrm{CP1}}{dt} &= f_1(\mathbf{CP}; Cl, I) \\ \frac{d\mathrm{CP2}}{dt} &= f_2(\mathbf{CP}; Cl, I) \\ \frac{d\mathrm{CP3}}{dt} &= f_3(\mathbf{CP}; Cl, I) \end{aligned} \tag{4.3}$$

donde **CP** representa el vector de CPNLs (CP1, CP2 *y* CP3), Cl e I son nuevamente los valores de los parámetros de operación. El vector de funciones del lado derecho, $\mathbf{f}(\mathbf{CP}; Cl, I)$ representa una estructura de modelo no-lineal, la cual es aproximada con un modelo continuo basado en RNAs, al que se le denominará RNA-EDO.

Los resultados presentados aquí, fueron obtenidos con un identificador que tiene la configuración de RNA que se presenta en la Figura 4.7. Dicho identificador consistió de una RNA de alimentación hacia adelante con 4 capas (dos capas ocultas). Conteniendo 5 neuronas en la capa de entrada: 2 para los parámetros de operación y 3 para el vector de CPNLs. En la capa de salida se tuvieron 3 neuronas con función de activación lineal, para predecir la derivada temporal de cada CPNL. Cada capa oculta consistió de 10 neuronas con función de activación no-lineal del tipo sigmoidal.

Antes de elegir la configuración final, se intentaron diferentes estructuras de RNAs. Todas estas configuraciones fueron: de alimentación hacia adelante de cuatro capas; con el mismo número de neuronas en cada capa oculta; función de activación no-lineal del tipo sigmoidal; tres neuronas en la capa de salida; y cinco neuronas en la capa de entrada. La diferencia entre todas las configuraciones intentadas, fue el número de neuronas en las capas ocultas. Se intentaron configuraciones con 8, 9, 10, 11 y

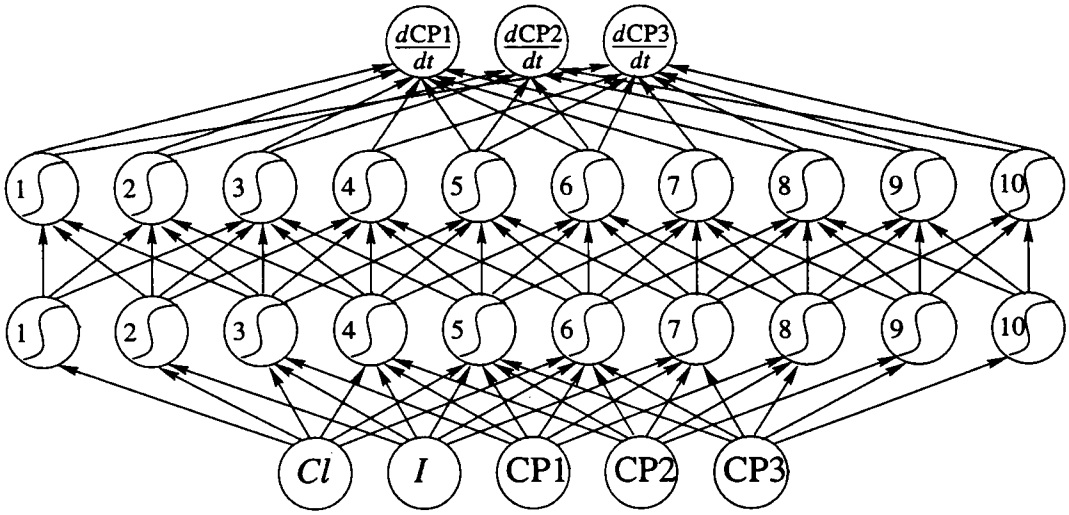


Figura 4.7: *Estructura de RNA utilizada para aproximar un conjunto de EDOs con datos de CPNLs.*

12 neuronas. Las RNAs con mayor número de neuronas ocultas presentaron mejores capacidades predictivas a largo plazo, pero el tiempo de entrenamiento fue mayor, y viceversa. De esta manera, se escogió una RNA con 10 neuronas ocultas, para mantener un equilibrio entre la calidad de las predicciones a largo plazo y el tiempo de entrenamiento.

El entrenamiento de la RNA-EDO se llevó acabo de la misma forma como se explicó en la Sección 3.2. Los conjunto de entrenamiento y prueba se construyeron exactamente con las mismas series de tiempo usadas en el entrenamiento de la red de CPNLs (ver la Tabla 4.3), sólo que ahora las series de tiempo fueron en CPNLs. Cada vector en el conjunto de entrenamiento (prueba) consistió de 8 elementos (5 entradas y 3 salidas). El intervalo de tiempo que se utilizó en la integración de la RNA-EDO fue de 0.125 segundos. No fue necesario escalar las series de tiempo en CPNLs, ya que se tenían presentes en el rango $[-0.9, +0.9]$.

El proceso de entrenamiento se inició asignandole parámetros a la RNA-EDO

en forma aleatoria en un rango de $[-0.2, +0.2]$. Se entrenaron varias RNAs con diferentes parámetros de inicio. La que mejores resultados presentó, necesitó de 100 ciclos completos del gradiente conjugado (cada ciclo con 203 pasos o iteraciones) para alcanzar convergencia en el entrenamiento. El error (estandarizado) al final del entrenamiento fue de 0.01327 (calculado con las Ecuaciones3.13 y 4.2).

4.2.3 Análisis de resultados

En la Figura 4.8 se presenta la comparación de algunas series de tiempo originales contra las predichas a largo plazo por la RNA-EDO. Las series de tiempo predichas se obtuvieron de la siguiente manera: la RNA-EDO se integra a largo plazo de donde se obtienen series de tiempo en CPNLs; después cada punto de estas series en CPNLs se procesa a través de la parte superior de la RNA-CNPLs (decodificador) para generar un punto en la serie de tiempo del potencial (se desechan 19 de los 20 puntos generados).

Las gráficas de la figura muestran que la RNA-EDO ha capturado adecuadamente la forma cualitativa de las oscilaciones, así como la amplitud y frecuencia de éstas. Estos resultados dan la confianza de que la RNA-EDO ha capturado, adecuadamente, también la secuencia de transiciones e inestabilidades que se presentan en los datos experimentales.

En la Figura 4.9 se presenta un diagrama de bifurcaciones en un parámetro (la intensidad de corriente) predicho por el modelo RNA-EDO correspondiente a un valor de $Cl = 0.004$ M. El modelo predice que para valores bajos de I se tiene solamente un punto fijo estable, el cual pierde su estabilidad por medio de una bifurcación de de Hopf subcrítica (punto A) en un valor de $I = 0.6473$ mA. El estado estacionario

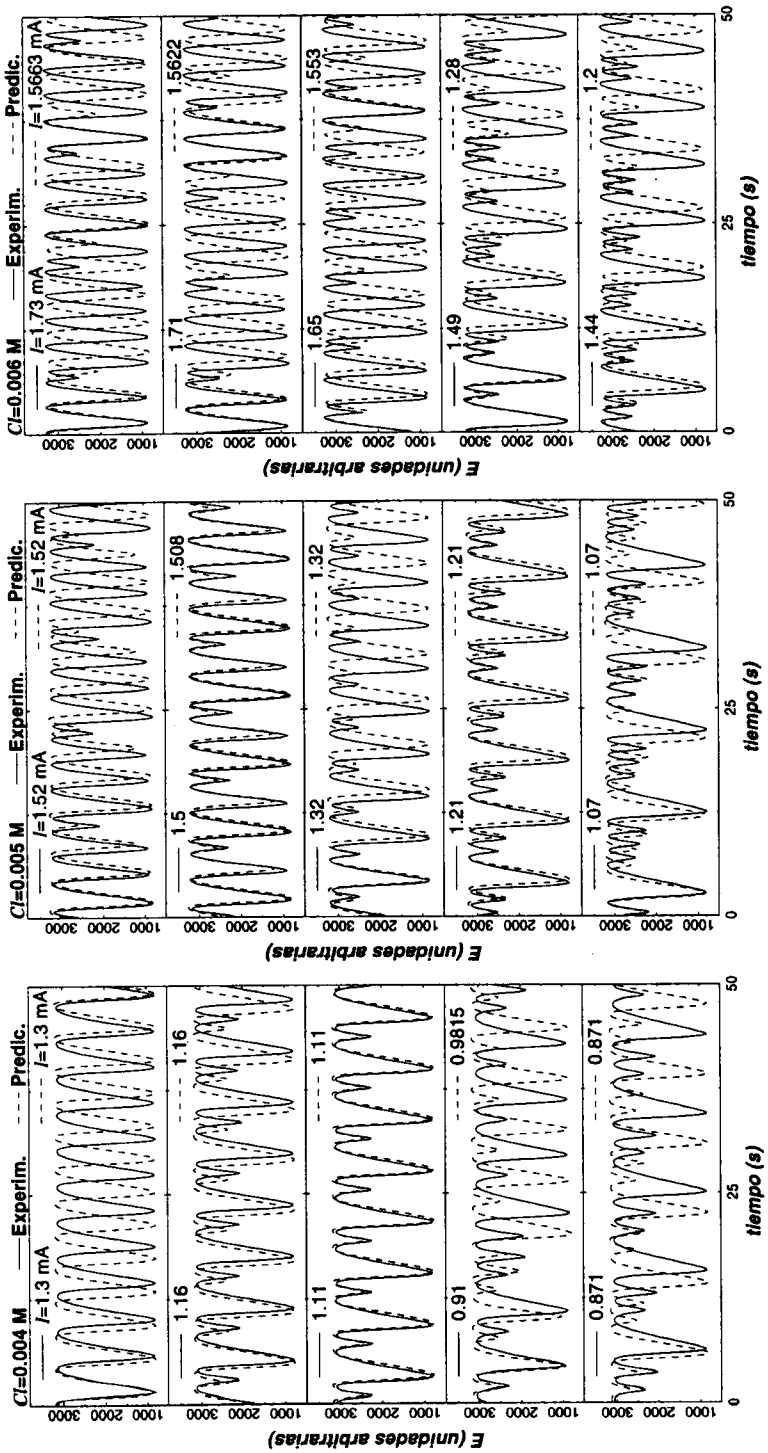


Figura 4.8: *Comparación de algunas series de tiempo experimentales contra las predichas (a largo plazo) por la RNA-EDO. Las líneas continuas representan las series originales y las discontinuas representan las predicciones de la RNA-EDO. En cada gráfica se presentan los valores de los parámetros usados.*

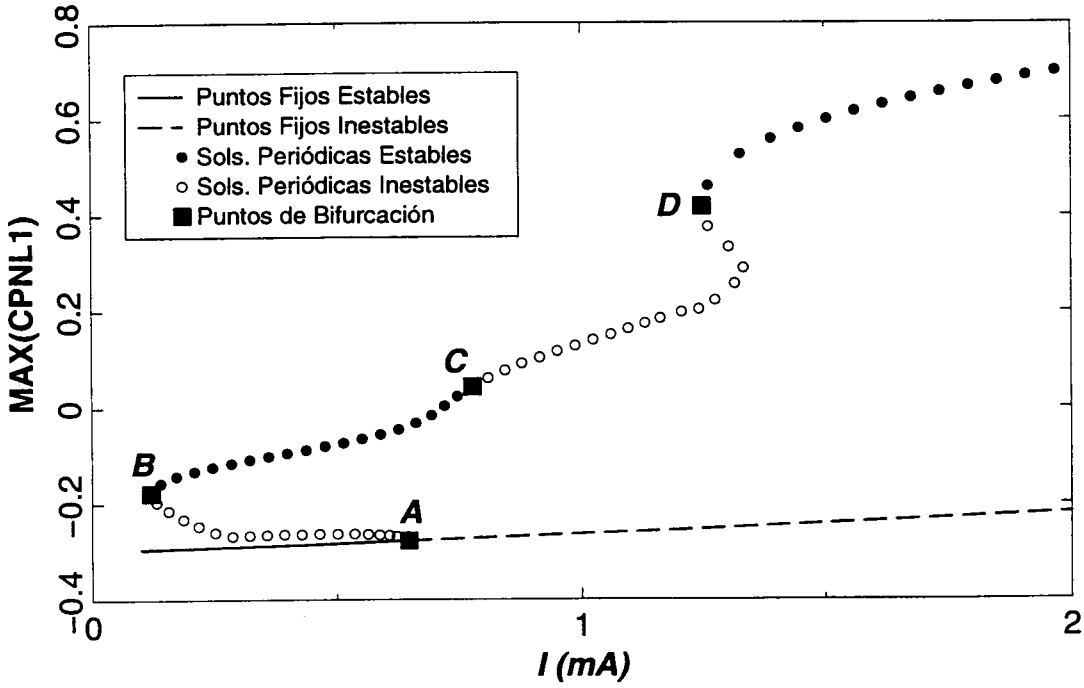


Figura 4.9: *Diagrama general de bifurcaciones en un parámetro (I) predicho por la RNA-EDO.*

y la solución periódica no se encuentran en los datos experimentales. Las solución periódica, que se abre en la bifurcación de Hopf, es inestable y de amplitud pequeña, la cual gana estabilidad por medio de una bifurcación de nodo-punto de silla (punto B) a un valor de $I = 0.1210$ mA. La solución periódica estable (de amplitud pequeña) se pierde a un valor de $I = 0.7783$ mA por medio de una cascada de bifurcaciones de doble periodo (punto C) que da lugar a un escenario de OMMs. Al aumentar el valor de I también crece la amplitud de la oscilación periódica inestable la cual encuentra una bifurcación de nodo-punto de silla pero se mantiene inestable. Después vuelve a encontrar otra bifurcación de nodo-punto de silla (punto D) a un valor de $I = 1.2484$ mA en donde gana estabilidad y continúa aumentando de amplitud mientras aumenta el valor del parámetro de operación. La solución periódica estable de amplitud grande que se encuentra en valores de I mayores de 1.2484 mA corresponde a las series de tiempo de amplitud grande que se encuentran en los datos experimentales, es decir la OMM 1^0. Mientras que la solución periódica estable de amplitud pequeña que se encuentra a valores de I menores de 0.7783 mA corresponde a la OMM 0^1, de la cual no se tienen datos experimentales. La zona comprendida entre los puntos C y D es donde se encuentran todo el escenario de OMMs $(\ldots, 1^3, 1^2, 1^1, 2^1, 3^1, \ldots)$. Esta zona esta acorde con las observaciones experimentales. La secuencia de bifurcaciones que se presenta en la figura es para un valor de $Cl = 0.004$ M, pero este comportamiento también se observa para cualquier valor de Cl dentro del rango de interés (0.004 a 0.006 M). Como puede observarse, el modelo RNA-EDO predice una conexión entre la OMM 0^1 y la 1^0 a través de una rama de soluciones periódicas inestables (línea de círculos entre los puntos C y D). Por el momento se está ante un nuevo descubrimiento en el estudio de OMMs, ya que dicha conexión no se ha reportado en la literatura.

Sin embargo, esta afirmación no es definitiva pues se presume que la conexión es el resultado de un cambio de soluciones, por parte del software utilizado, en una región de bifurcaciones globales. Para corroborar este resultado, es necesario hacer un estudio meticuloso de la dinámica global del sistema a lo largo de la rama de soluciones periódicas inestables, lo cual no es nada trivial. De cualquier manera, se seguirá trabajando en este rubro.

En la Figura 4.10 se presenta un diagrama general de bifurcaciones en dos parámetros (intensidad de corriente contra concentración de iones cloro). En la gráfica solamente se muestran tres estados principales de OMMs (0^1, 1^1 y 1^0). La línea de bifurcación inferior representa la continuación (en dos parámetros) del punto de bifurcación de doble periodo que se encuentra etiquetado como "C" en la Figura 4.9, mientras que la línea de bifurcación superior representa la continuación del punto de bifurcación nodo-punto de silla que se encuentra etiquetado como "D" en la Figura 4.9. Las líneas intermedias son la continuación (en dos parámetros) de las bifurcaciones que limitan la OMM 1^1. Como puede apreciarse, la OMM 1^1 comienza en valores bajos de I por medio de una bifurcación de nodo-punto de silla y termina en valores altos de I por medio de una cascada de bifurcaciones de doble periodo. Los otros estados de OMMs ($\ldots, 1^2, 2^1, 3^1, \ldots$) se encuentran en las zonas sombreadas de la figura. Dichos estados de OMMs, aunque no se muestran en la figura, presentan un comportamiento similar al del estado de OMM 1^1. Estos resultados concuerdan satisfactoriamente con los presentados en la literatura, los cuales son caracterizados como pertenecientes a una secuencia Farey de OMMs.

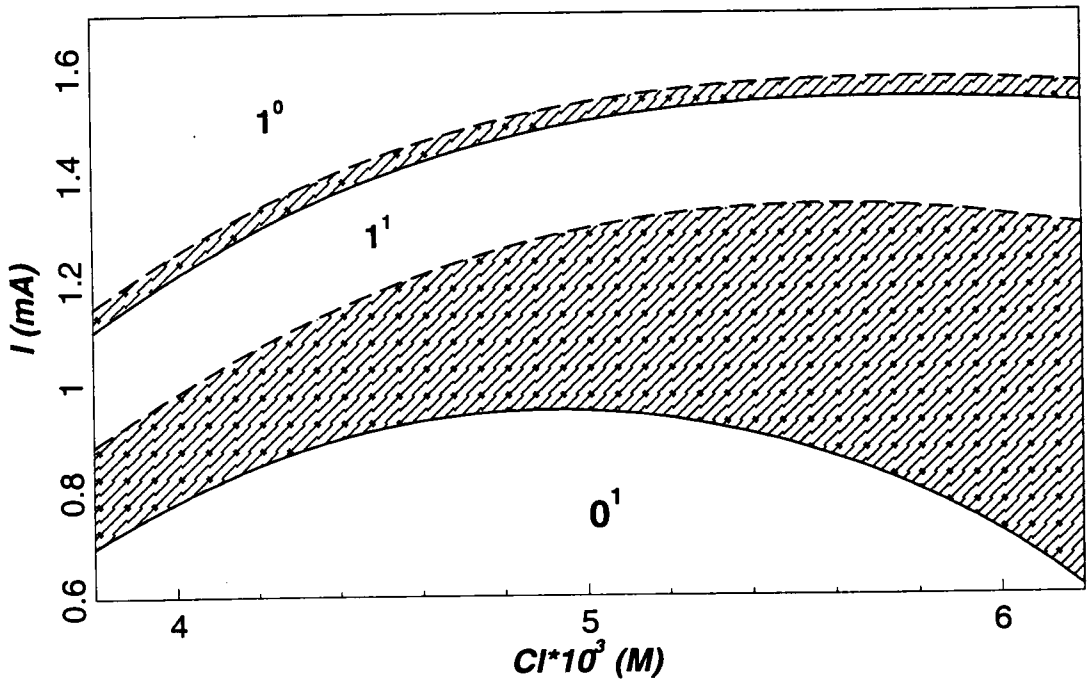


Figura 4.10: *Diagrama general de bifurcaciones en dos parámetros (I contra Cl) predicho por la RNA-EDO.*

5. Caracterización de un proceso espacio-temporal

El ejemplo presentado aquí pretende demostrar los méritos que tienen los modelos basados en RNAs para capturar la dinámica de largo plazo en sistemas con variaciones espacio-temporales de las denominadas ondas viajeras. Sin embargo, cuando la EDP se conoce, las RNAs también pueden ser usadas para aproximar numéricamente sus soluciones [36].

Para ejemplificar la aplicabilidad del algoritmo presentado en la Sección 3.3, considérese la Ecuación de Kuramoto-Sivashinsky (Apéndice B). Los datos para la ilustración han sido muestreados a partir de la solución de esta EDP en un valor del parámetro $\alpha = 15$, evaluados en 21 posiciones equidistantes del rango espacial de interés. Para este valor de α la solución de la Ecuación de Kuramoto-Sivashinsky presenta los dos diferentes atractores que se muestran en la Figura B.1.

5.1 RNA para la identificación

El modelo dinámico, utilizado para la identificación del comportamiento espacio-temporal de la Ecuación de Kuramoto-Sivashinsky, tiene la siguiente forma:

$$\frac{\partial U}{\partial t} = F\left(U, \frac{\partial U}{\partial t}, \frac{\partial^2 U}{\partial t^2}, \frac{\partial^3 U}{\partial t^3}, \frac{\partial^4 U}{\partial t^4}\right) \tag{5.1}$$

donde la función del lado derecho, $F(\cdot)$, representa una estructura de modelo no-lineal. Esta estructura de modelo es aproximada por una RNA, la cual se denominará RNA-KS, en relación a la Ecuación de Kuramoto-Sivashinsky.

Los resultados presentados aquí, fueron obtenidos con un identificador que tiene la configuración de RNA que se presenta en la Figura 5.1. Dicho identificador consistió de una RNA de alimentación hacia adelante con 4 capas. En la capa de entrada se tuvieron 5 neuronas pertenecientes a las derivadas espaciales de la variable de estado (el parámetro α no aparece aquí ya que se mantuvo constante). En la capa de salida se tuvo una neurona con función de activación lineal, para predecir la derivada temporal de la variable de estado. Cada capa oculta consistió de 8 neuronas con función de activación no-lineal del tipo sigmoidal.

Antes de elegir la configuración final, se intentaron diferentes estructuras de RNAs. Todas estas configuraciones fueron de RNAs de alimentación hacia adelante de cuatro capas, con el mismo número de neuronas en cada capa oculta; función de activación no-lineal del tipo sigmoidal; una neurona en la capa de salida (para $\frac{dU}{dt}$); y cinco neuronas en la capa de entrada. En este caso se conoce la forma de la ecuación fundamental y se decidió incluir todas las derivadas direccionales, pero esto no queda restringido sólo a cinco (se pudo haber utilizado un número mayor). La diferencia

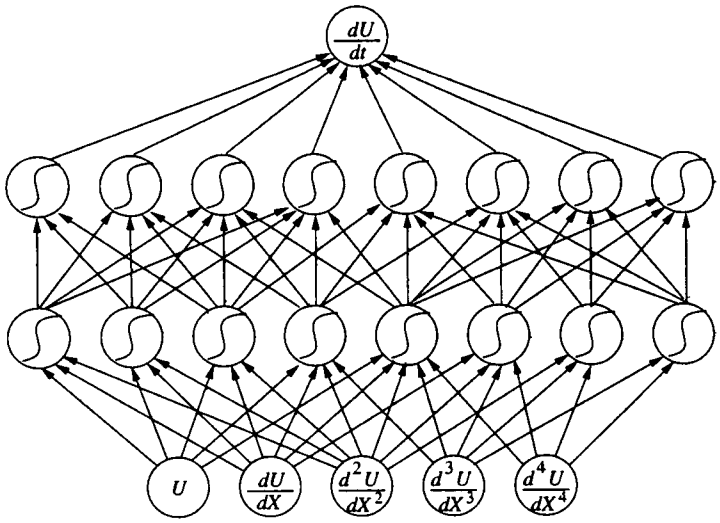


Figura 5.1: *Estructura de RNA utilizada en la identificación del proceso espacio-temporal exhibido por la Ecuación de Kuramoto-Sivashinsky.*

entre todas las configuraciones intentadas, fue el número de neuronas en las capas ocultas. Se intentaron configuraciones con 4, 6, 7, 8, 9, 10, 11 y 12 neuronas. Las RNAs con pocas neuronas ocultas (4, 6 y hasta 7 neuronas) presentaron pobres capacidades predictivas a largo plazo y el tiempo de entrenamiento fue pequeño. Mientras que, las RNAs con demasiadas neuronas ocultas presentaron buenas capacidades predictivas, pero el tiempo de entrenamiento fue muy grande. De esta manera, se escogió una RNA con 8 neuronas ocultas, para mantener un equilibrio entre la calidad de las predicciones a largo plazo y el tiempo de entrenamiento.

El entrenamiento de la RNA-KS se llevó acabo de la misma forma como se explicó en la Sección 3.3. El conjunto de entrenamiento fue construido tomando en consideración un criterio basado en el compromiso que existe entre la calidad de las predicciones a largo plazo y el tiempo de entrenamiento. De esta forma, el conjunto de entrenamiento final fue construido con 600 mediciones a diferentes instantes en el

tiempo. Cada medición representó un vector de 42 puntos: 21 puntos en el vector de entrada y 21 puntos en el vector de salida. El intervalo de muestreo entre los vectores de entrada y salida del conjunto de entrenamiento fue de 0.0001 unidades de tiempo. Esto significa, que 0.0001 unidades de tiempo fue el intervalo de tiempo utilizado en la integración. Entre dos mediciones consecutivas del conjunto de entrenamiento se usaron diferentes intervalos de tiempo (no necesariamente 0.0001 unidades). Estas mediciones se obtuvieron de trayectorias cercanas a los dos atractores que presenta la Ecuación KS en $\alpha = 15.0$ (se tomaron 300 mediciones por cada atractor). Las entradas de la RNA-KS fueron discretizadas utilizando diferencias finitas con un orden de aproximación $\mathbf{O}(\Delta X^4)$ (ver el Apéndice C). Lo cual permitió utilizar los datos del conjunto de entrenamiento (y prueba) para evaluar la salida de la RNA-KS.

El proceso de entrenamiento se inició asignando los parámetros de la RNA aleatoriamente en un rango de $[-0.2, +0.2]$. Se entrenaron varias RNAs con diferentes parámetros de inicio. La que mejores resultados presentó, necesitó de 13 ciclos completos del gradiente conjugado (cada ciclo con 129 pasos o iteraciones) para alcanzar convergencia en el entrenamiento. El error (estandarizado) al final del entrenamiento fue de 4.97×10^{-5} (calculado con las Ecuaciones3.20 y 4.2).

5.2 Análisis de resultados

Las solución a largo plazo que se obtiene con el modelo RNA-KS, al igual que la EKS, presenta dos atractores. Dichos atractores se denominarán A y B. En lo que resta del capítulo, las figuras (A) serán relativas al atractor A y las figuras (B) serán relativas al atractor B. Recordando que una solución a corto plazo involucra integrar un sistema una sola iteración, y largo plazo involucra integrar muchas iteraciones con

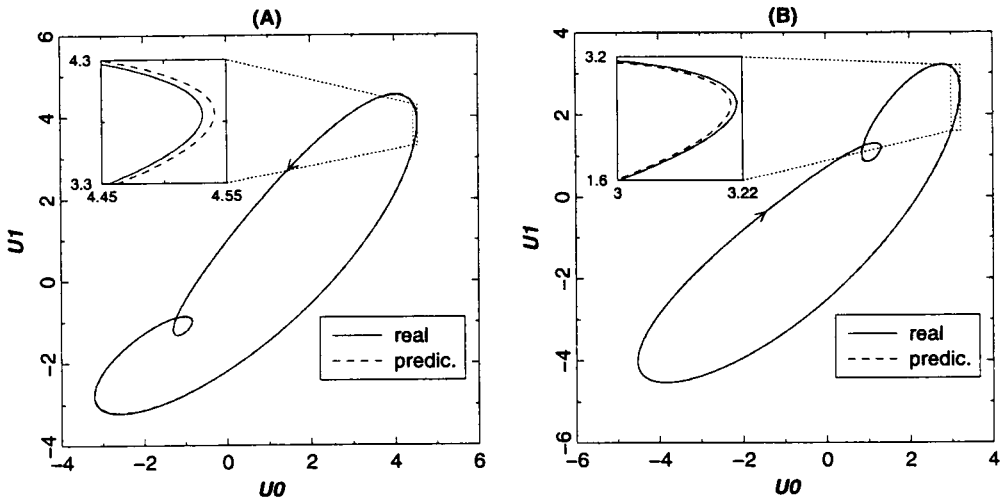


Figura 5.2: *Comparación de soluciones a largo plazo reales contra las predicciones hechas por la RNA-KS. Los estados $U0$ y $U1$ representan el valor de U en el nodo 0 y 1 respectivamente.*

sólo presentarle un punto inicial (ver la Sección 2.1).

En la Figura 5.2 se presenta una comparación (en el espacio-fase) de los atractores reales contra las predicciones obtenidas con el modelo RNA-KS. Los atractores presentados son bien definidos, es decir, son las soluciones a largo plazo obtenidas al integrar (la Ecuación de Kuramoto-Sivashinsky y la RNA-KS) durante un tiempo muy grande hasta obtener convergencia. La concordancia cualitativa entre los atractores predichos y los reales es muy buena, ya que el atractor predicho logra sobreponerse con el atractor real. Dentro de la misma gráfica se presenta otra gráfica que muestra el detalle de una región de la curva, para dar una idea del grado en que se sobreponen estos atractores. De esta manera el modelo RNA-KS logra capturar en forma extraordinaria la amplitud de los dos atractores (A y B).

En la Figura 5.3 se hace una comparación en series de tiempo para el primer nodo ($U0$). Al igual que en la figura anterior, las series de tiempo también son soluciones

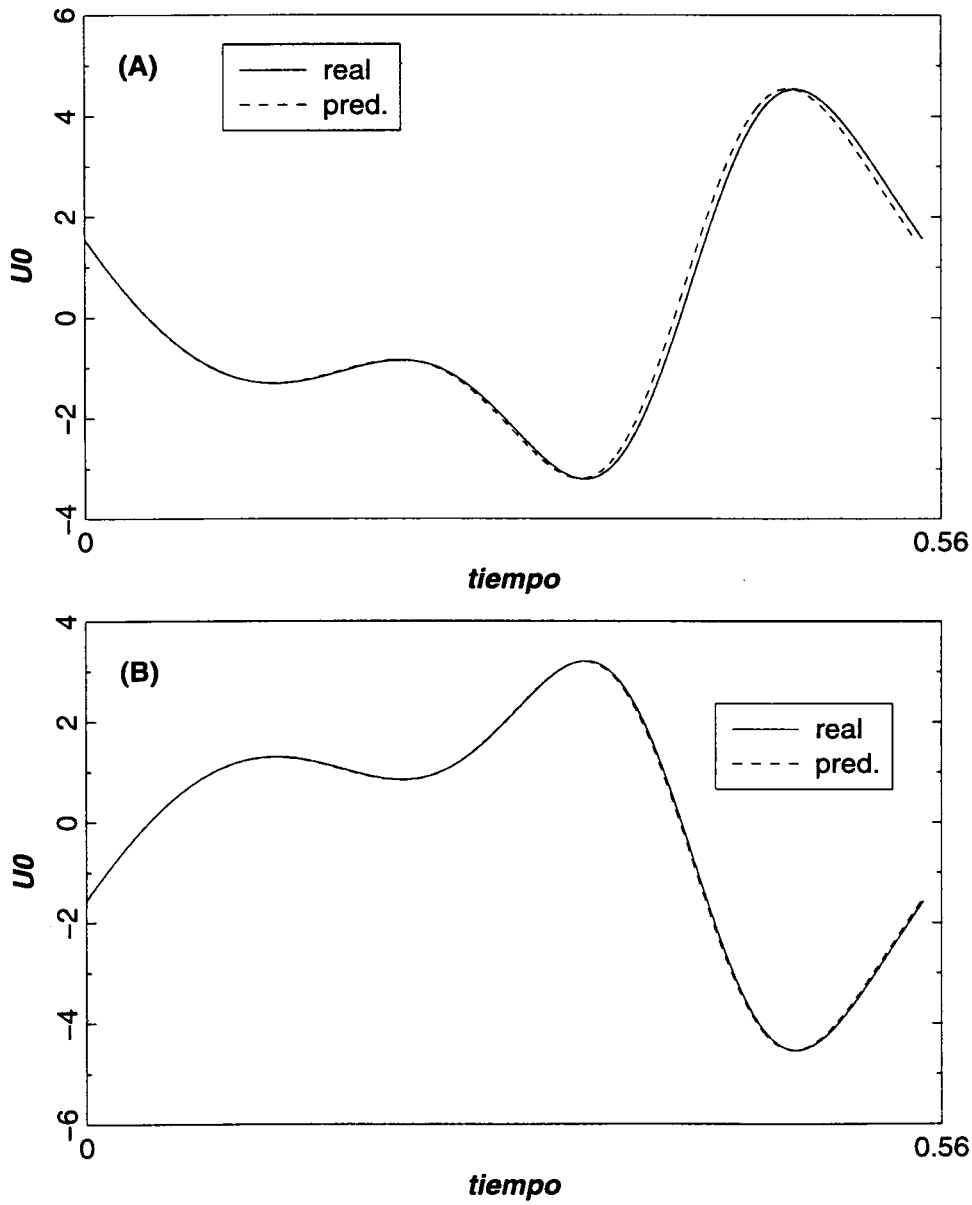


Figura 5.3: *Comparación de las series de tiempo reales contra las predichas por la RNA-KS.*

a largo plazo bien definidas. Las series de tiempo predichas por el modelo RNA-KS son más rápidas que las series de tiempo reales. Es decir, el tiempo de periodo de las series de tiempo predichas (0.5431 unidades de tiempo en el atractor A y 0.5469 unidades de tiempo en el atractor B) es más pequeño que el de las series reales (0.5483 unidades de tiempo en ambos atractores). Sin embargo, la diferencia de los tiempos de periodo apenas representan un 0.95% del tiempo real en el atractor A, y un 0.26% en el atractor B. Tomando en cuenta que sólo se usaron unas cuantas mediciones para construir el conjunto de estrenamiento, el modelo RNA-KS ha sido capaz de predecir satisfactoriamente el tiempo de periodo para los dos atractores.

En la Figura 5.4 se presenta la comparación de las ondas viajeras (espacio $U - X$) reales contra las predichas por el modelo RNA-KS. La comparación es presentada en los siguientes instantes del tiempo: $t = \frac{1}{2}T$ y $t = T$. Donde T es el tiempo de periodo del atractor real y representa 0.5483 unidades de tiempo real (5483 iteraciones en la integración). Las ondas reales pertenecen a diferentes posiciones de los atractores que se presentan en la Figura 5.2. Mientras que, las ondas predichas no pertenecen a un atractor, pero si a una trayectoria muy cercana a dicho atractor. Dicha trayectoria se determinó utilizando como punto inicial un punto sobre el atractor real, del cual se estimó toda la trayectoria. De cualquier manera, las predicciones siguen siendo a largo plazo, ya que si se toma en cuenta el incremento de tiempo (0.0001 unidades de tiempo) en la integración de las ecuaciones (EKS y la RNA-KS), se requirieron de 2742 iteraciones para alcanzar la onda del medio periodo ($t = \frac{1}{2}T$) y 5483 iteraciones para alcanzar la onda en el periodo completo ($t = T$). Nuevamente, las predicciones de la RNA-KS son de muy buena calidad.

Debido a que en las figuras anteriores es difícil detectar las diferencias entre los

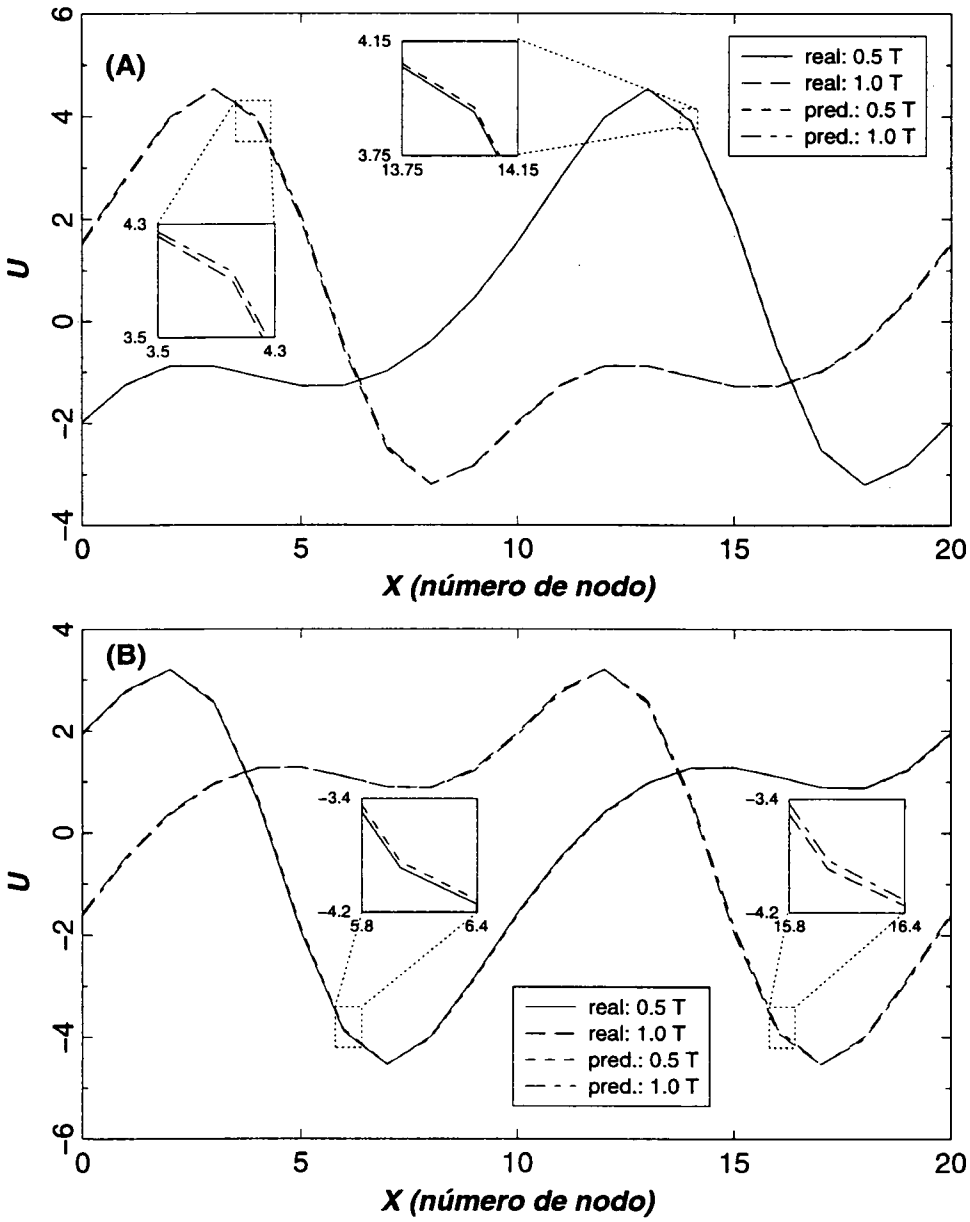


Figura 5.4: *Comparación de ondas viajeras reales contra las predichas por la RNA-KS en $t = \frac{1}{2}T$ y $t = T$. La dirección X se representa por el número de nodo.*

atractores reales y predichos, en las figuras restantes se compara la propagación de errores porcentuales con respecto a la máxima amplitud del atractor (real). Los errores son determinados con la siguiente fórmula:

$$error = \frac{U_{real} - U_{pred.}}{(U_{max} - U_{min})_{real}} \times 100 \tag{5.2}$$

donde U_{max} y U_{min} son los valores máximo y mínimo, respectivamente, que puede tomar la variable de estado real U. En el Atractor A se tienen los valores: $U_{max} = 4.5312$ y $U_{min} = -3.2109$; mientras que en el Atractor B sus valores son: $U_{max} = 3.2109$ y $U_{min} = -4.5312$. De cualquier manera, para los efectos del cálculo de errores, la diferencia es la misma: $U_{max} - U_{min} = 4.5312 - (-3.2109) = 3.2109 - (-4.5312) = 7.7421$, lo cual permite hacer la comparación en una misma base, que representa la amplitud de la onda. La comparación pudo hacerse con respecto a cualquier otra base, sin embargo, se tomó la amplitud, ya que representa una propiedad importante en predicción de soluciones periódicas a largo plazo.

En la Figura 5.5 se presenta la propagación de errores de los dos primeros nodos ($U0$ y $U1$) a través del tiempo. Las series de tiempo utilizadas para calcular el error fueron determinadas de la misma forma que en la figura anterior (la serie real es un atractor definido y la serie predicha se calcula con un punto inicial sobre el atractor real). Los errores comienzan con valores bajos y conforme transcurre el tiempo sus valores tienden a crecer. Esto se debe a que el tiempo de periodo de los atractores reales es diferente al de los atractores predichos (ver la Figura 5.3). Sin embargo, el error apenas sobrepasa el 1%(en base a un ciclo de la onda).

En la Figura 5.6 se muestra la forma en que cambia el error con respecto a la

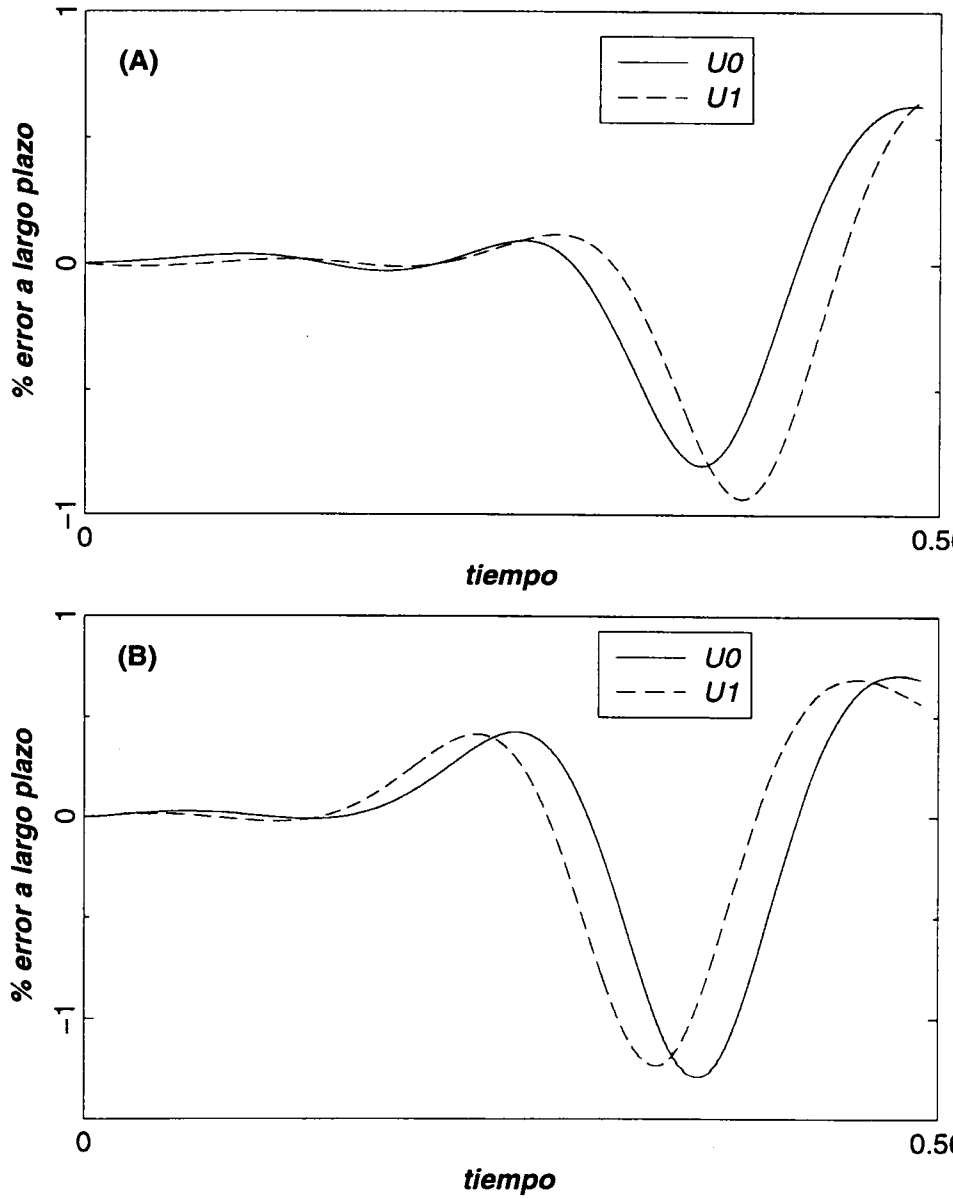


Figura 5.5: *Errores porcentuales a largo plazo (a través del tiempo) entre las series reales y las predichas por la RNA-KS.*

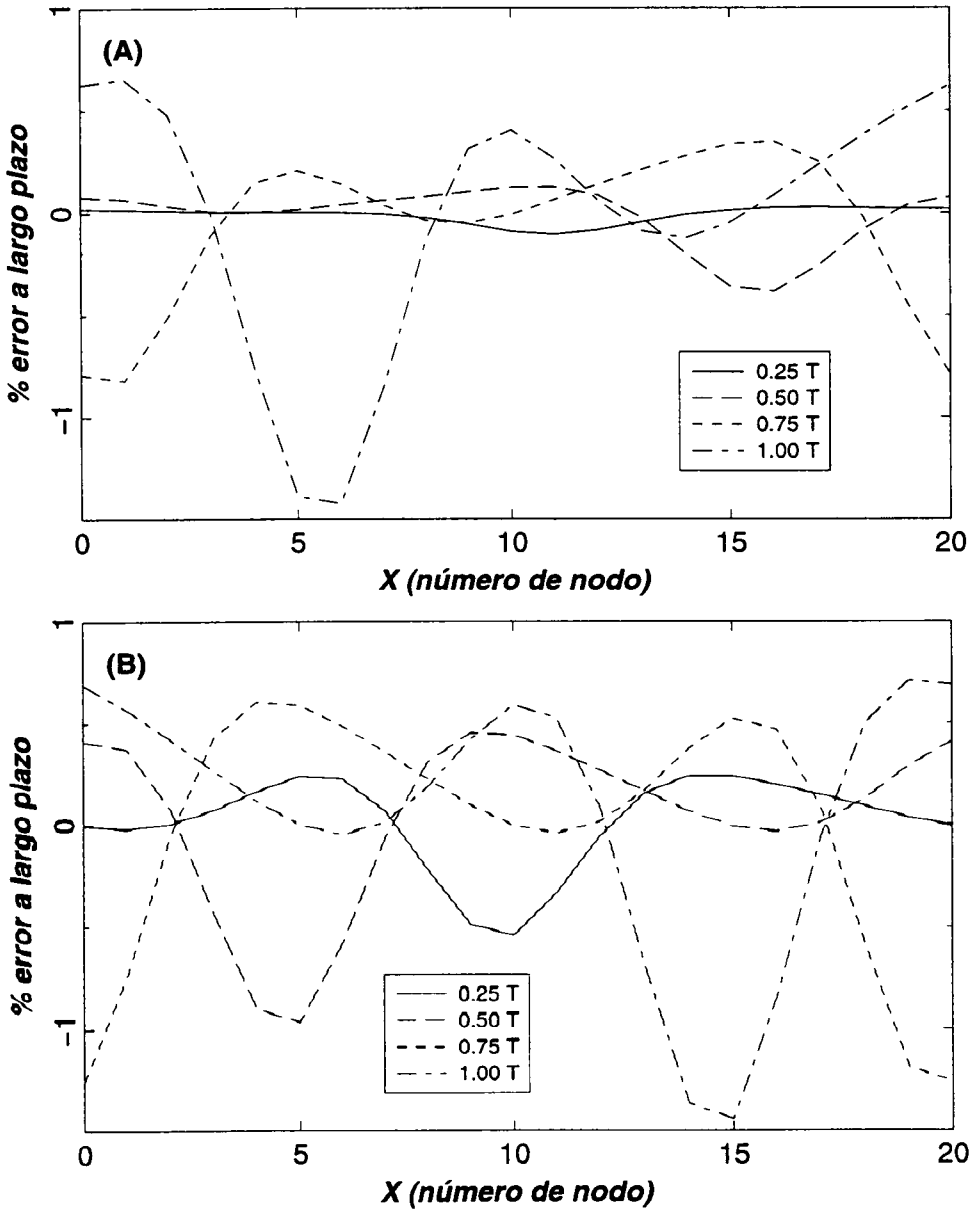


Figura 5.6: *Errores porcentuales a largo plazo entre las ondas reales y las predichas por la RNA-KS.*

dimension espacial (X) en diferentes instantes del tiempo ($t = \frac{1}{4}T$, $t = \frac{1}{2}T$, $t = \frac{3}{4}T$ y $t = T$). Las series de tiempo utilizadas para calcular el error fueron determinadas de la misma forma que en la figura anterior (la serie real es un atractor definido y la serie estimada se calcula con un punto inicial sobre el atractor real). Esta figura pretende mostrar los errores (no apreciables) de la Figura 5.4. Las dos figuras mostradas anteriormente no rebasan el 1.5% de error, lo cual es muy alentador sabiendo que se trata de soluciones a largo plazo.

Por último, en la Figura 5.7 se presenta una comparación de errores porcentuales a corto plazo. Es decir, la RNA-KS sólo predice el valor U un incremento de tiempo hacia adelante. Esto significa que a la RNA-KS se le alimentó con puntos pertenecientes a un atractor real y, ésta estimó el siguiente punto sobre dicho atractor. En estas figura el error no rebasa más del 0.0015 % lo cual representa un excelente resultado.

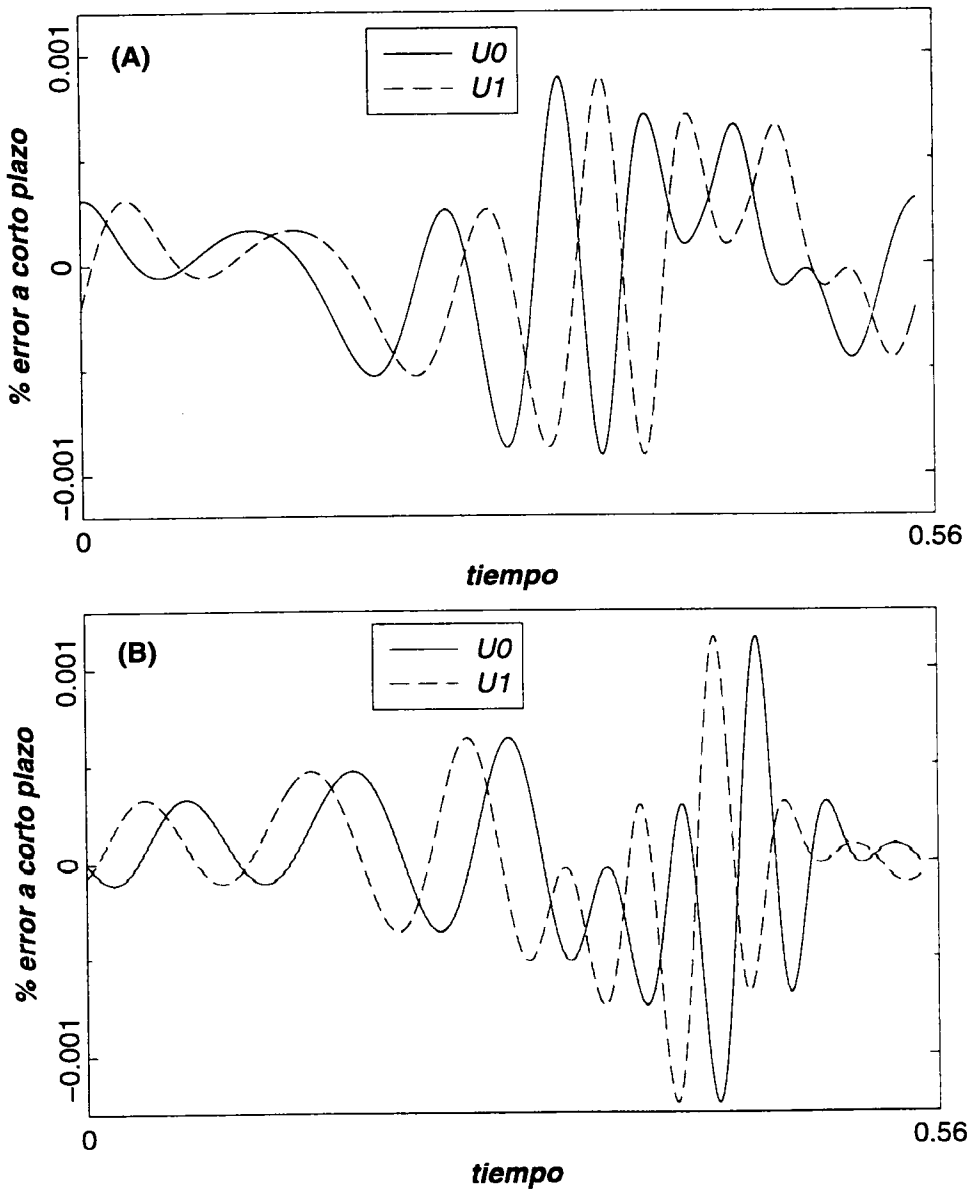


Figura 5.7: *Errores porcentuales a corto plazo entre las series de tiempo reales y las predichas por la RNA-KS.*

6. Discusión y Conclusiones

Aquí se utilizaron los modelos de RNAs en la identificación de sistemas dinámicos continuos (sistemas de parámetros agrupados y de parámetros distribuidos) que presentan comportamientos complejos. Estos modelos, posteriormente, fueron utilizados para caracterizar, de manera exitosa, el comportamiento dinámico de los sistemas originales. La caracterización estuvo basada, principalmente, en predicciones a largo plazo que ayudaron a entender de manera clara y compacta las transiciones (bifurcaciones) e inestabilidades del proceso real. Sin embargo, los modelos obtenidos no sólo pueden ser usados en la caracterización de los sistemas, sino que también, pueden ser utilizados en el diseño preliminar de equipo y sistemas de control; pueden ayudar en la optimización del proceso, o pueden asistir en el diseño de nuevos experimentos (para mejorar la caracterización o explorar regiones difíciles de experimentar como son los escenarios de multiplicidad). De cualquier forma, la metodología de identificación de sistemas dinámicos ejemplificada aquí, no sustituye los esfuerzos de identificación de sistemas basada en principios fundamentales.

En el Capítulo 4 se aplicaron los algoritmos presentados en las Secciones 3.1 y 3.2 en la identificación de los datos experimentales de un proceso continuo que presenta comportamiento complejo de OMMs. El procesamiento de los datos sirvió como

un punto de arranque para: determinar la dimensión del modelo ajustado, filtrar el ruido contenido en los datos experimentales y reducir sustancialmente los datos. Es importante recalcar que el adecuado procesamiento de los datos experimentales es determinante en el éxito de las aproximaciones. Los resultados obtenidos con los dos modelos (MPMs y EDOs) son apropiados para caracterizar el escenario de transiciones presentes en los datos experimentales. Las predicciones obtenidas con estos modelos son muy parecidas, en el sentido de la secuencia de bifurcaciones locales. La diferencia es que el modelo discreto de MPMs no es capaz de predecir la presencia de estados estacionarios, mientras que el modelo continuo de EDOs sí lo hace, además de poder predecir la forma como empiezan a generarse las OMMs a partir de los estados estacionarios. Como trabajo futuro se propone explorar, de manera más meticulosa, la dinámica (local y global) del modelo continuo ya obtenido. Con esto se pretende conocer más sobre los fenómenos generadores del comportamiento de OMMs, para posteriormente utilizarlos como guía en los modelos basados en principios básicos, y para la realización de experimentos adicionales que permitan corroborar algunas de las predicciones que en la base de datos actual no se observan (como lo es la predicción de una bifurcación de Hopf subcrítica a valores bajos de la intensidad de corriente).

En el Capítulo 5 se le dió una aplicación práctica al algoritmo basado en RNAs para la identificación de sistemas espacio-temporales que fue propuesto en la Sección 3.3. La base de la implementación fue el uso de diferencias finitas acopladas con la RNA e incrustadas en esquemas de integración numérica (método de líneas), para crear la estructura del identificador, lo cual permitió la formulación del entrenamiento de la red basada en mediciones equidistantes de la variable de estado. La aplicación fue ejemplificada con datos obtenidos de la integración de la EKS sin tomar en cuenta

la dependencia del parámetro de operación (α). Es decir, los datos fueron obtenidos sólo para un valor del parámetro de operación donde la EKS presenta comportamiento complicado: ondas viajeras y multiplicidad de soluciones. Los resultados obtenidos con esta implementación fueron de excelente calidad para los diferentes atractores que presenta el sistema original (EKS). Una extensión natural de este trabajo es la caracterización del escenario de transiciones cuando el parámetro de operación es modificado. Para esto es necesario entrenar nuevas RNAs, en las que se incluya el parámetro de operación en la entrada de la RNA. Esto permitirá predecir los diferentes tipos de bifurcaciones presentes en el sistema y dará idea de las capacidades predictivas de las RNAs en escenarios de bifurcaciones en sistemas espacio-temporales. Sin embargo, este trabajo requiere de grandes cantidades de datos en el proceso de identificación, lo cual se transforma en excesivos tiempos de ejecución en el proceso de entrenamiento. Esta desventaja puede eliminarse implementando otras alternativas de integración (por ejemplo, métodos espectrales); implementación de algoritmos en paralelo, o simplemente utilizar mejores equipos de cómputo.

En cualquier caso, de los presentados aquí, la metodología de identificación esta basada en predicciones a corto plazo, mientras que los resultados en la caracterización (secuencia de bifurcaciones) de los sistemas esta basada únicamente en predicciones a largo plazo. En trabajos futuros se buscará implementar nuevos métodos de identificación que tomen en cuenta las predicciones a largo plazo, la forma estructural de los atractores o explorando las propiedades dinámicas de RNAs recurrentes. De esta manera se busca agilizar los procesos de entrenamiento y predecir de manera todavía más clara y exacta los escenarios de bifurcaciones presentes en las observaciones experimentales.

En el proceso de entrenamiento de RNAs, los tiempos de ejecución del entrenamiento son demasiado grandes, debido a la estructura misma de la RNA (contiene un número muy grande de parámetros ajustables). Una alternativa, para reducir al máximo estos tiempos, es la implementación en paralelo de los algoritmos de entrenamiento de las RNAs que simulan el comportamiento dinámico de sistemas. En trabajos futuros se pretende explorar esta alternativa ya que representa una de las más promisorias posibilidades para el entrenamiento de RNAs hacia implementaciones en línea y en tiempo real.

A continuación se presentan las conclusiones más importantes del presente trabajo.

- Los modelos de RNAs son capaces de capturar la dinámica de largo plazo en sistemas que presentan comportamiento dinámico complejo.

- Gran parte del éxito de las aproximaciones depende de un adecuado preprocesamiento de los datos experimentales.

- El éxito de las predicciones a largo plazo, aún cuando el modelo se entrene basado en predicciones a corto plazo, se logra asignándole a la RNA, durante el entrenamiento, datos pertenecientes a trayectorias cercanas a los atractores.

- La caracterización se puede llevar a cabo aún cuando se le presente a la RNA, durante el entrenamiento, información parcial de la dinámica del proceso.

- Los dos modelos presentados para la caracterización de las OMMs predicen un estado 0^1 que no se encuentra en las observaciones experimentales.

- El modelo continuo, para la caracterización de OMMs, es capaz de predecir un

estado estacionario en valores bajos de intensidad de corriente, el cual desaparece por medio de una bifurcación de Hopf subcrítica.

- El modelo continuo, para la caracterización de OMMs, predice que los estados de OMMs 0^1 y 1^0 estan conectados (Figura 4.9) por medio de una órbita periódica inestable.

- El modelo discreto, para la caracterización de OMMs, predice algunas regiones de "histéresis", mientras que el modelo continuo no lo hace.

Bibliografía

[1] F. Albahadily, J. Ringland, y M. Schell. Mixed-mode oscillations in an electrochemical system. I. A Farey sequence which does not occur on a torus. *J. Chem. Phys*, 90:813–821, 1989.

[2] J. Anderson, J. Silverstein, S. Ritz, y R. Jones. Distinctive features, categorical perception and probability learning: some applications on a neural model. *Psychological Review*, 84:413–451, 1977.

[3] M. Baer, M. Falcke, y M. Eiswirth. Dispersion relation and spiral rotation in an excitable surface reaction. *Physica A*, 188:78–88, 1992.

[4] D. Barkley. Near-critical behavior for one-parameter families of circle maps. *Phys. Lett. A*, 129:219–222, 1988.

[5] D. S. Broomhead y G. P. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217–236, 1986.

[6] S. R. Chu y R. Shoureshi. A neural network approach for identification of continuous-time nonlinear dynamic systems. *Proceedings of the 1991 American Control Conference*, 1:1–5, 1991.

[7] G. Cybenko. Approximation by superposition of a sigmoidal function. *Math. Control Signals Systems*, 42:303–314, 1989.

[8] H. Dai y N. K. Sinha. Robust identification of continuous-time multivariable systems using numerical integration methods. *Control and Dynamic Systems*, 56:51–91, 1993.

[9] D. DeMers. Dimensionality reduction for non-linear time series. *Neural and Stochastic Methods in Image and Signal Processing*, SPIE 1766:200–211, 1992.

[10] E. J. Doedel y J. P. Kernévez. AUTO: A program for continuation and bifurcation problems in ordinary differential equations. *Applied Mathematics Report, California Institute of Technology*, 1–226, 1986.

[11] G. Ertl. Oscillatory catalytic reactions at single-crystal surfaces. *Adv. Catalysis*, 37:213–277, 1990.

[12] H. Froehling, J. P. Crutchfield, D. Farmer, N. H. Packard, y R. Shaw. On determining the dimension of chaotic flows. *Physica D*, 3:605–617, 1981.

[13] B. S. Gottfried y J. Weisman. *Introduction to Optimization Theory*. Prentice Hall, Englewood Cliffs, NJ, 1973.

[14] S. Grossberg. Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity. *Proceedings of the National Academy of Sciences*, 59:368–372, 1968.

[15] J. Guckenheimer y P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, Heidelberg, 1983.

[16] J. Guckenheimer, M. R. Myers, F. J. Wicklin, y P. A. Worfolk. DSTOOL: A Dynamical System Toolkit with an Interactive Graphical Interface. *Center For Applied Mathematics Cornell University, Ithaca NY*, 1991.

[17] D. O. Hebb. *The Organization of Behavior.* John Wiley & Sons , New York, 1949.

[18] R. Hecht-Nielsen. Theory of the backpropagation neural network. *Proceedings of the 1989 International Joint Conference on Neural Networks, Washington, DC*, I:593–608, 1989.

[19] G. E. Hinton, D. H. Ackley, y T. J. Sejnowski. Boltzmann machines: Constraint satisfaction networks that learn. *Carnegie-Mellon University, Technical Report CMU-CS-84-119*, 1984.

[20] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA*, 79:2554–2558, 1982.

[21] G. Horányi y C.Visy. Potential oscillations in the course of galvanostatic oxidation of hydrogen at platinum electrode in the presence of electrosorbing cations. *Journal of Electroanalytical Chemistry*, 103:353, 1979.

[22] K. Hornick, M. Stinchcombe, y H. White. Multi-layer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1988.

[23] J. S. Il'yashenko. Global analysis of the phase portrait for the Kuramoto-Sivashinsky equation. *Technical Report 665, IMA Preprint*, 1990.

[24] B. Irie y S. Miyake. Capabilities of three layer perceptrons. *Proceedings of the 1988 International Joint Conference on Neural Networks, San Diego, CA*, 1:641–648, 1988.

[25] S. Jakubith, H. H. Rotermund, W. Engel, S. A. von Oertzen, y G. Ertl. Spatiotemporal concentration patterns in a surface reaction: Propagating and standing waves, rotating spirals and turbulence. *Phys. Rev. Letters*, 65:3013–3016, 1990.

[26] E. D. Karnin. A simple procedure for pruning back-propagation trained neural networks. *IEEE Trans. Neural Networks*, 1:239–242, 1990.

[27] T. Kodera, T. Yamazaki, M. Masuda, y R. Ohnishi. Limit cycle in electrochemical oscillation – Potential oscillation during anodic oxidation of H_2. *Electrochimica Acta*, 33:537–540, 1988.

[28] M. Koper y P. Gaspard. Mixed-mode and chaotic oscillations in a simple model of an electrochemical oscillator. *J. Phys. Chem.*, 95:4945–4947, 1991.

[29] M. Koper y P. Gaspard. The modelling of mixed-mode and chaotic oscillations in electrochemical systems. *J. Chem. Phys.*, 96:7797–7813, 1992.

[30] M. A. Kramer y J. A. Leonard. Diagnosis using backpropagation neural networks– Analysis and criticism. *Computers chem. Engng*, 14:1323–1338, 1990.

[31] K. Krischer, M. Lübke, M. Eiswirth, W. Wolf, J. L. Hudson, y G. Ertl. A hierarchy of transitions to mixed mode oscillations in an electrochemical reaction. *Physica D*, 62:123–133, 1993.

[32] K. Krischer, M. Lübke, W. Wolf, M. Eiswirth, y G. Ertl. Chaos and interior crisis in an electrochemical reaction. *Ber. Bunsenges. Phys. Chem.*, 95:820–823, 1991.

[33] K. Krischer, M. Lübke, W. Wolf, M. Eiswirth, y G. Ertl. Oscillatory dynamics of the electrochemical oxidation of H_2 in the presence of Cu^{2+}: Structure sensitivity and the rule of anions. *Electrochimica Acta*, 40:69–81, 1995.

[34] K. Krischer, R. Rico-Martínez, I. G. Kevrekidis, H. H. Rotermund, G. Ertl, y J. L. Hudson. Model identification of a spatiotemporally varying catalytic reaction. *AIChE J.*, 39:89–98, 1993.

[35] Y. Kuramoto y T. Tsuzuki. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Prog. Theor. Phys.*, 55(2):356–369, Feb. 1976.

[36] I. E. Lagaris, A. Likas, y D. I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *Preprint*, 1997.

[37] A. S. Lapedes y R. M. Farber. How neural networks work. *In: Neural Information Processing Systems (D.Z. Anderson, ed.), AIP Press*, 422–456, 1987.

[38] A. S. Lapedes y R. M. Farber. Nonlinear signal processing using neural networks: Prediction and system modeling. *Los Alamos Report LA-UR-87-2662*, 1987.

[39] J. Maselko y H. L. Swinney. Complex periodic oscillations and Farey arithmetic in the Belousov-Zhabotinskii reaction. *J. Chem. Phys.*, 85:6430–6441, 1986.

[40] G. Mayer-Kress. *Dimensions and Entropies in Chaotic Systems.* Springer Verlag, 1986.

[41] W. S. McCulloch y W. A. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematics and Biophysics*, 5:115–133, 1943.

[42] M. Minsky y S. Papert. *Perceptrons: An Introduction to Computational Geometry.* MIT Press, Cambridge, MA, 1969.

[43] N. H. Packard, J. P. Crutchfield, J. D. Farmer, y R. S. Shaw. Geometry from a time series. *Phys. Rev. Letters*, 45:712–716, 1980.

[44] R. Reed. Pruning algorithms – A survey. *IEEE Trans. Neural Networks*, 4:740–747, 1993.

[45] R. Rico-Martínez y I. G. Kevrekidis. Continuous-time modeling of nonlinear systems: A neural network-based approach. *Proceedings of the 1993 IEEE International Conference on Neural Networks*, III:1522–1525, 1993.

[46] R. Rico-Martínez, I. G. Kevrekidis, M. C. Kube, y J. L. Hudson. Discrete- vs. continuous-time nonlinear signal processing: Attractors, transitions and parallel implementation issues. *Proceedings of the 1993 American Control Conference*, 2:1475–1479, 1993.

[47] R. Rico-Martínez, K. Krischer, M. C. Kube I. G. Kevrekidis, y J. L. Hudson. Discrete- vs. continuous-time nonlinear signal processing of Cu electrodissolution data. *Chem. Eng. Comm.*, 118:25–48, 1992.

[48] J. Ringland, N. Issa, y M. Schell. From U sequence to Farey sequence: A unification of one-parameter scenarios. *Phys. Rev. A*, 41:4223–4235, 1990.

[49] F. Rosenblatt. The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386–408, 1958.

[50] D. Rumelhart y J. McClelland ed. *Parallel Distributed Processing: Explorations in microstructure of cognition. Vol. 1. Foundations.* MIT Press, Cambridge, MA, 1986.

[51] S. Saarinen, R. Bramley, y G. Cybenko. Ill-conditioning in neural network training problems. *SIAM J. Sci. Stat. Comput.*, 14:693–714, 1993.

[52] E. Saund. Dimensionality-reduction using connectionist networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11:305–314, 1989.

[53] G. I. Sivashinsky. Nonlinear analysis of hidrodynamics instability in laminar flames-I. Derivation of basic equations. *Acta Astronaut.*, 4:1177–1206, 1977.

[54] H-T. Su y T. J. McAvoy. Identification of chemical processes using recurrent networks. *Proceedings of the 1991 American Control Conference*, 2314–2319, 1991.

[55] F. Takens. *Detecting strange attractors in turbulence.* In: Dynamical Systems and Turbulence, (D. A. Rand y L. S. Young, eds.), Lect. Notes in Math., Springer, Heidelberg, 1981.

[56] M. A. Taylor, M. S. Jolly, y I. G. Kevrekidis. SCIGMA: Technical Report. *Dept. of Chem. Eng., Princeton University, Princeton, NJ*, 1990.

[57] J. Theiler. Estimating the fractal dimension of chaotic time series. *The Lincoln Laboratory Journal*, 3:63–86, 1990.

[58] F-S. Tsung y G.W Cottrell. Learning temporal signals in phase space. *International Symposium on Nonlinear Theory and Its Applications*, Hawaii, December, 1993.

[59] S. Usui, S. Nakauchi, y M. Nakano. Reconstruction of Munsell color space by a five-layered neural network. *Proceedings of the 1990 IEEE International Joint Conference in Neural Networks, San Diego, CA*, 2:515–520, 1990.

[60] P. D. Wasserman. *Advanced Methods in Neural Computing.* Van Nostrand Reinhold, New York, 1993.

[61] A. S. Weigend y N. A. Gershenfeld. *Time Series Prediction: Forecasting the Future and Understanding the Past.* Addison-Wesley, 1993.

[62] P. Werbos. *Beyond Regression: New Tools for Prediction and Analysis of Behavioral Sciences.* Ph.D. Thesis, Harvard University, 1974.

[63] B. Widrow y M. E. Hoff. Adaptive switching circuits. *IREWESCON Convention Record*, 4:96–104, 1960.

[64] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, volume 2 of *Texts in Applied Mathematics.* Springer Verlag, 1990.

[65] W. Wolf, K. Krischer, M. Lübke, M. Eiswirth, y G. Ertl. Modelling oscillations in galvanostatic H_2 oxidation at Pt in presence of metal ions. *Journal of Electroanalytical Chemistry*, 385:85–93, 1995.

[66] T. Yamazaki, T. Kodera, R. Ohnishi, y M. Masuda. Potential oscillation during anodic oxidation of hydrogen at a platinum electrode – II. Kinetic analysis. *Electrochimica Acta*, 35:431–435, 1990.

[67] Y. You y M. Nikolaou. Dynamic process modeling with recurrent neural networks. *AIChE Journal*, 39:1654–1667, 1993.

A. Oxidación galvanostática de hidrógeno

En este apéndice se presentan datos experimentales de mediciones de potencial (E) del proceso de oxidación de hidrógeno en la presencia de iones bismuto y cloro bajo condiciones galvanostáticas sobre una superficie de platino como catalizador. Los experimentos fueron realizados por W. Wolf, M. Lübke y M. Eiswirth del Fritz-Haber-Institut der Max-Planck-Gesellschaft en Berlín, Alemania, a quienes se les agradece su colaboración en este proyecto. Información adicional acerca del comportamiento experimental en la oxidación de hidrógeno sobre platino, puede ser encontrada en [31, 65, 33].

La reacción química de interés es:

$$2H_2 + O_2 \rightarrow 2H_2O + 2e^-$$

El comportamiento oscilatorio en el proceso de oxidación de H_2 en electrodos de platino fue estudiado por vez primera por Horányi y Visy [21]. Ellos encontraron que esta reacción exhibe oscilaciones bajo condiciones galvanostáticas (pero no potenciostáticas) en la presencia de diferente electroadsorción de iones metálicos y fueron

los primeros en proponer un mecanismo intuitivo para describir las oscilaciones. Este mecanismo toma en cuenta la presencia de iones metálicos, dejando a un lado la presencia de los no metálicos. Modelos basados en la idea de Horányi y Visy [21] se han construido, a base de principios básicos, para predecir comportamiento oscilatorio [27, 66]. Recientemente Krischer y col. [31, 33] han mostrado que el efecto de los iones no-metálicos juega un papel esencial en la ocurrencia de las oscilaciones, logrando así un modelo matemático (muy simplificado), basado en principios fundamentales, que predice comportamiento de OMMs, aunque un modelo matemático completamente satisfactorio todavía no ha sido propuesto.

A.1 Arreglo Experimental

El reactor donde se llevó a cabo la reacción fue una celda electroquímica convencional que consistió en un arreglo de 3 electrodos colocado en un matraz de 3 bocas (Figura A.1). Un alambre de platino rotatorio en el interior de un cilindro central funcionó como *electrodo de trabajo*; en un cilindro lateral se colocó una vara de platino de 25 cm^2 de área superficial que funcionó como *contra-electrodo*, y en el otro cilindro se colocó un *electrodo de referencia* compuesto de Hg/Hg_2SO_4. La velocidad de rotación del electrodo de trabajo y la temperatura de operación se mantuvieron constantes durante las corridas experimentales.

El electrólito consistió de una solución 1 M de $HClO_4$ a la cual se le adicionaron pequeñas concentraciones de iones bismuto y cloro. Durante las mediciones el H_2 fue burbujeado continuamente a través de la solución agitada, hasta lograr saturarla con éste. La deposición repetida y la disolución subsecuente de bismuto mejoró la reproducibilidad de las oscilaciones [25, 32]. El potencial fue medido entre el electrodo de

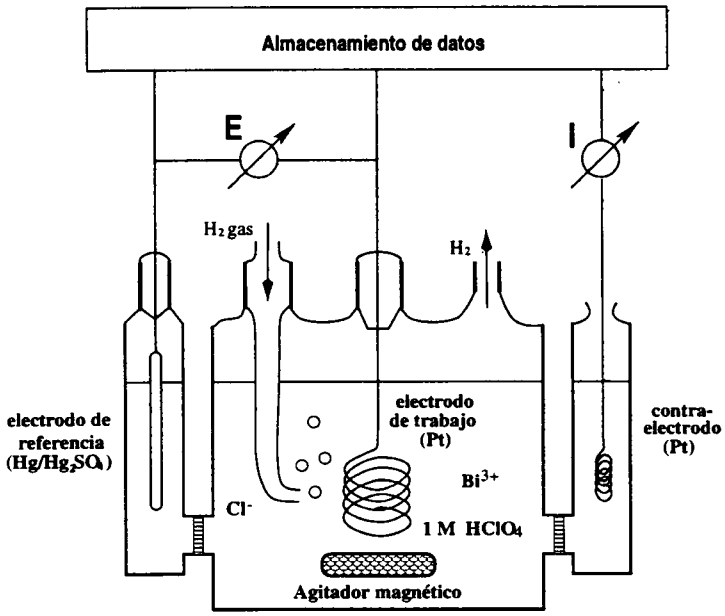


Figura A.1: *Esquema del arreglo experimental de la oxidación galvanostática de H_2 sobre electrodos de Pt.*

trabajo y el electrodo de referencia con una frecuencia de 24 Hz (0.0417 segundos entre mediciones) y sus señales fueron digitalizadas y almacenadas en una microcomputadora para el análisis de los datos. Detalles adicionales acerca de la organización y preparación de los electrodos se encuentran en [3, 11, 32, 31].

Bajo estas condiciones la variable de estado que se midió fue el potencial (E) en mV. Además, se tienen los siguientes parámetros de operación: la intensidad de corriente aplicada (I) en mA; la concentración de iones bismuto ($[Bi^{2+}]$) en mol/l, y la concentración de iones cloro ($[Cl^-]$) en mol/l.

Las observaciones estudiadas en el presente trabajo involucran el escenario de transiciones observadas cuando sólo dos parámetros de operación (la corriente aplicada y la concentración de iones cloro) fueron modificados, mientras que la concentración de iones bismuto fue mantenida constante en un valor $1 \times 10^{-4}\,M$. En la Figura A.2

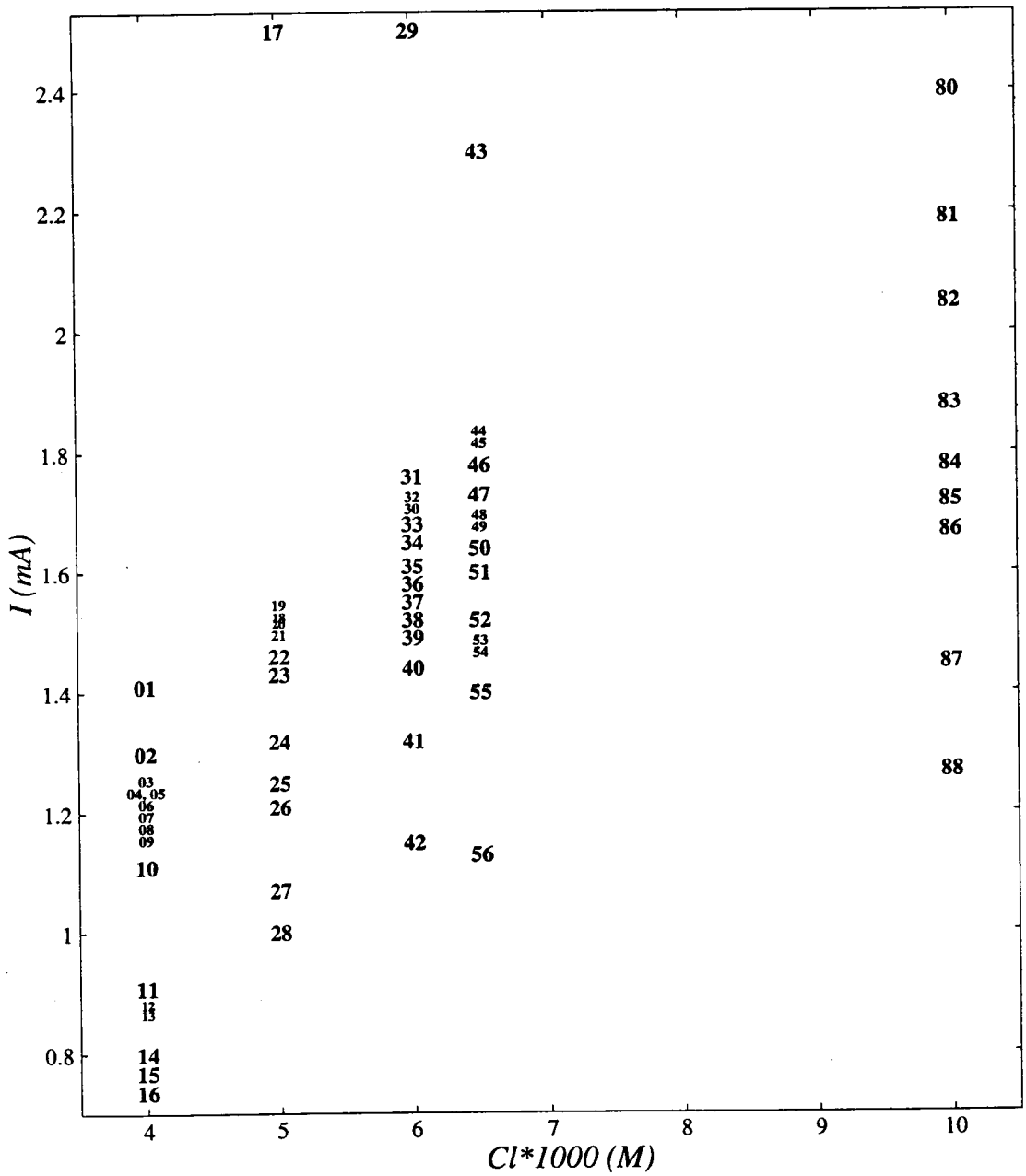


Figura A.2: *Escenario de las corridas experimentales para una concentración de iones bismuto de* $1 \times 10^{-4}\,M$. *Cl representa la concentración de iones cloro,* $[Cl^-]$, *en mol/l e I representa la intensidad de corriente aplicada en mA.*

se presenta un escenario esperimental en los dos parámetros de operación utilizados. En la Figura A.2 sólo se presentan los experimentos que fueron usados en el proceso de caracterización con las RNAs. En este escenario faltan los experimentos etiquetados desde el número 57 al 79 que no se usaron en la identificación. En el escenario mostrado, el rango en la concentración de iones Cl^- va desde 4×10^{-3} hasta $1 \times 10^{-2}\,M$ (rango muy estrecho), sin embargo, se observa una gran variedad de OMMs y comportamiento caótico.

A.2 Observaciones Experimentales

En la Figura A.3 se presenta series de tiempo experimentales típicas del comportamiento oscilatorio que presenta la oxidación galvanostática de H_2 sobre electrodos de Pt. En esta figura, cada columna de gráficas representa experimentos para un valor particular de la concentración de iones cloro (corresponde a un corte perpendicular en la Figura A.2). Así, por ejemplo, la columna central presenta segmentos de 6 series de tiempo recolectadas a diferentes valores de la corriente aplicada para una concentración de iones cloro de $5 \times 10^{-3}\,M$ y corresponden a los experimentos señalados como 17, 21, 24, 26, 27 y 28 en la Figura A.2. El comportamiento presentado por esta columna de gráficas se describe a continuación.

Para el valor de la corriente más alto ($I = 2.5\,mA$) el sistema exhibe un atractor periódico compuesto de oscilaciones de amplitud grande, todas ellas del mismo tamaño (una oscilación de modo mixto del tipo 1^0). Al disminuir la corriente a un valor de $1.5\,mA$, se observa un atractor periódico compuesto de dos oscilaciones de amplitud grande seguidas de una oscilación de amplitud pequeña (una OMM del tipo 2^1). Al seguir bajando la corriente hasta $1.32\,mA$ se observa un atractor periódico

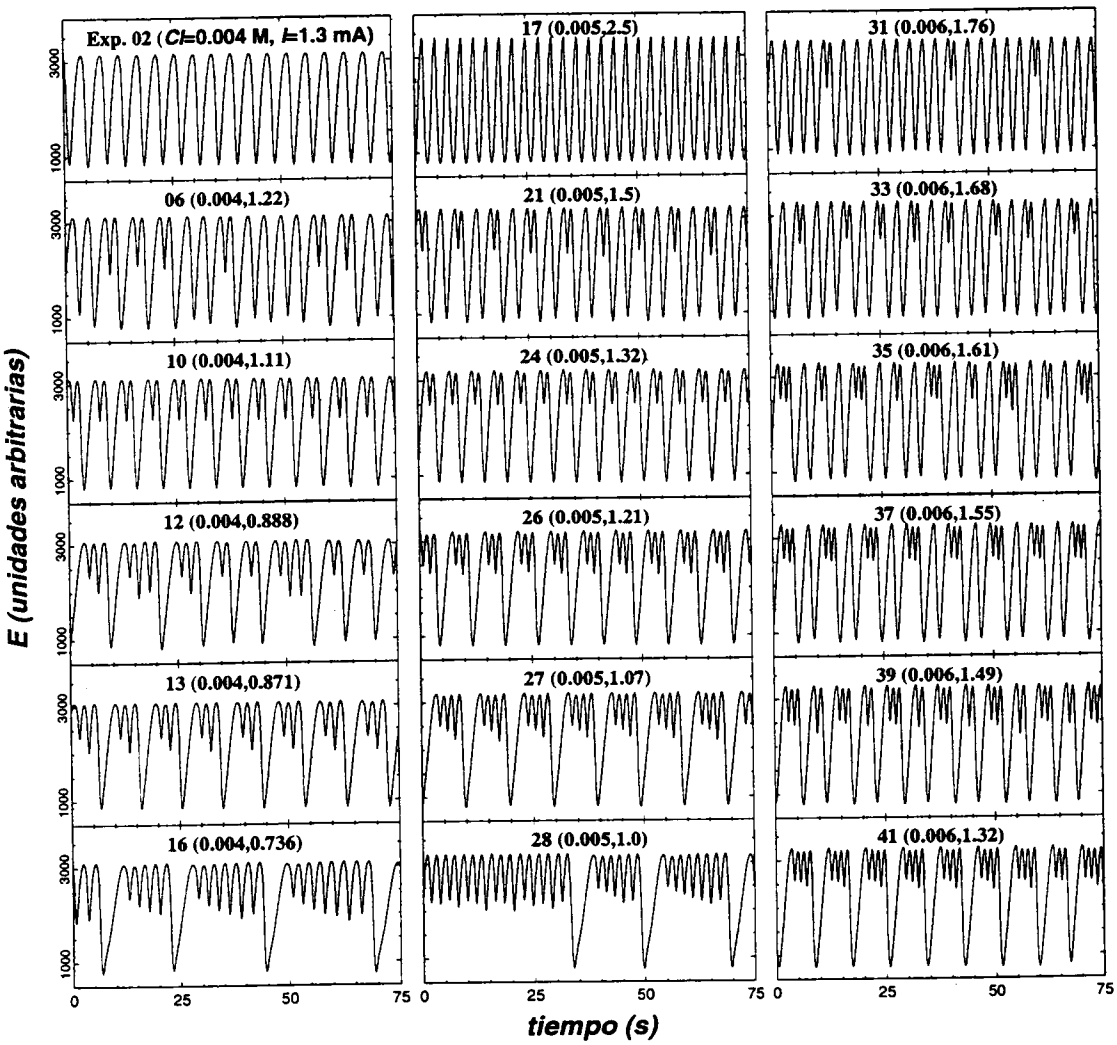


Figura A.3: *Series de tiempo experimentales típicas de la oxidación de H_2 sobre electrodos de Pt, recolectadas a una concentración de iones bismuto de $1 \times 10^{-4}\,M$. Por conveniencia los resultados se reportan en unidades arbitrarias de potencial (E) y para convertir a mV se aplica la siguiente fórmula: $e(mV) = (E - 533)/3$.*

compuesto de una oscilación de amplitud grande seguida de una oscilación de amplitud pequeña (una OMM del tipo 1^1). Si se sigue disminuyendo la corriente hasta $1.21\,mA$ se encuentra un atractor periódico compuesto de una oscilación de amplitud grande seguida por dos oscilaciones de amplitud pequeña (una OMM del tipo 1^2). Posteriormente, al disminuir hasta $1.07\,mA$ se presenta otro atractor periódico compuesto de una oscilación de amplitud grande seguida de tres oscilaciones de amplitud pequeña (una OMM del tipo 1^3). Finalmente, para el valor más bajo de corriente ($1.0\,mA$) se observa un atractor caótico donde predominan las oscilaciones de amplitud pequeña. El experimento 28 es el último para esta secuencia de experimentos, pero se puede estimar que al seguir disminuyendo la corriente, se tendrán solamente oscilaciones de amplitud pequeña (una OMM del tipo 0^1).

Observese que la secuencia de transiciones presentada, al variar la corriente desde valores altos hasta valores bajos, es 1^0, 2^1, 1^1, 1^2, 1^3 y 0^1 (que no se tiene experimentalmente). Entre los estados 1^0 y 2^1 se encuentran otros estados principales de OMMs, los cuales presentan varias oscilaciones de amplitud grande seguidas de una oscilación de amplitud pequeña, y entre los estados 1^3 y 0^1 se encuentran otros estados principales de OMMs que presentan una oscilación de amplitud grande seguida de varias oscilaciones de amplitud pequeña. Este mismo comportamiento es presentado por las columas del lado izquierdo y derecho de la Figura A.3.

B. Ecuación de Kuramoto-Sivashinsky

La *Ecuación de Kuramoto-Sivashinsky* (EKS) describe la modelación, en forma simplificada, de flujo de líquidos en películas delgadas. Esta ecuación surge del análisis de perturbaciones de soluciones regulares cerca de inestabilidades para una amplia variedad de escenarios físicos. A la ecuación se le da este nombre en honor a sus descubridores, ya que fue obtenida independientemente por Kuramoto en 1976 [35] y por Sivashinsky en 1977 [53]. Dicha ecuación, en su versión diferencial y adimensional, tiene la siguiente forma:

$$\frac{\partial U}{\partial t} + 4\frac{\partial^4 U}{\partial X^4} + \alpha\left(U\frac{\partial U}{\partial X} + \frac{\partial^2 U}{\partial X^2}\right) = 0$$

donde U representa la variable de estado y α representa el parámetro de operación. En reacciones con difusión, U representa una perturbación del frente plano de la flama inestable propagándose en la mezcla combustible-oxígeno y α es proporcional al número de Lewis (difusividad térmica/difusividad másica). Mientras que en la interfase de películas delgadas de líquido fluyendo hacia abajo de un plano inclinado,

U representa la altura de la interfase y α puede pensarse como el recíproco de la tensión superficial. En los últimos años, la EKS ha sido el punto de una amplia investigación debido a su rico contenido dinámico manifestado en función del parámetro de operación (tal como soluciones periódicas, homoclínicas y caóticas).

En el presente trabajo se ha escogido la EKS debido a que es una de las *Ecuaciones Diferenciales Parciales* (EDPs) no-lineales más simples que tiene un significado físico, además de su amplio contenido dinámico. Las soluciones de la EKS dependen en gran medida de las condiciones de frontera que se tengan. Aquí se utilizaron condiciones de frontera periódicas en el rango $X \in [0, 2\pi]$, ya que con dichas condiciones siempre se garantiza una solución "suave" de la EKS [23].

Para resolver la EKS se utilizó el "método de líneas" tal como fue explicado en el Capítulo 3 para el caso de RNAs. La integración se realizó discretizando la dimensión espacial (X) en 20 intervalos (21 nodos), y para este intervalo de X en particular se utilizó un paso de integración igual a 0.0001 unidades de tiempo (el integrador se vuelve inestable cuando se aumenta un orden de magnitud en el paso de integración). El parámetro de operación (α) se mantuvo constante en un valor igual a 15, ya que con este valor se obtienen, dependiendo de las condiciones iniciales, dos diferentes atractores de ondas viajeras. Estos atractores se muestra en la Figura B.1 en el espacio-fase bi-dimensional de U en el nodo 0 ($U0$) contra U en el nodo 1 ($U1$). El número de pasos (o iteraciones) necesarios para completar un ciclo (en cualquier atractor) son 5483, es decir que cada atractor consta de 5484 puntos y el tiempo del periodo (T) es de 0.5483 unidades de tiempo. Al primer atractor se le denominará *Atractor A* y al segundo *Atractor B*. En lo que resta, las figuras *(A)* serán relativas al *Atractor A* y las figuras *(B)* serán relativas al *Atractor B*.

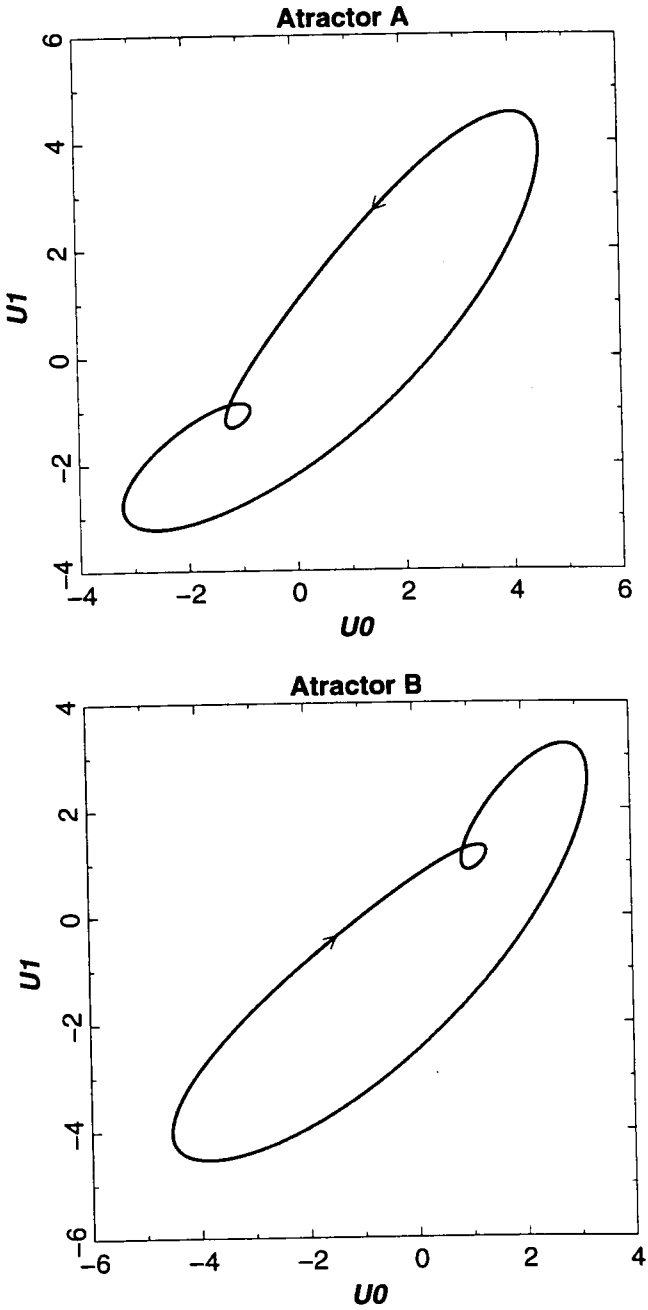


Figura B.1: *Soluciones a largo plazo de la EKS, proyectadas sobre dos de las mediciones del dominio espacial, para un valor del parámetro* $\alpha = 15$. *Los estados* $U0$ *y* $U1$ *representan el valor de* U *en el nodo* 0 *y* 1 *respectivamente.*

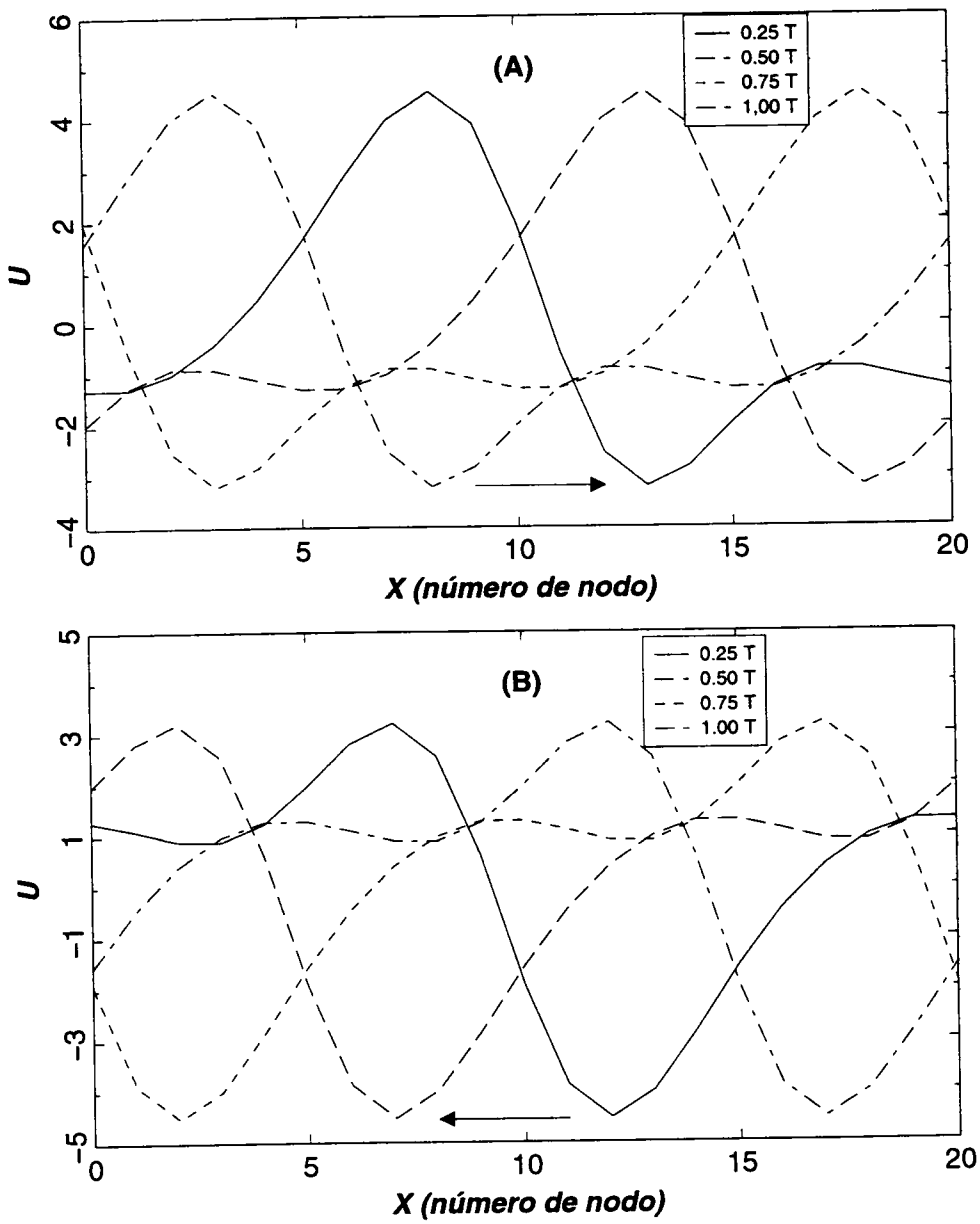


Figura B.2: *Evolución de las ondas viajeras de la EKS en $\alpha = 15$ para diferentes instantes en el tiempo. La dirección X se representa por el número de nodo.*

En la Figura B.2 se muestra la evolución de las onda viajera (en el espacio $X-U$) a diferentes instantes de tiempo. La dirección X representa el número de nodo, en lugar de su valor original que esta en el rango $[0, 2\pi]$. Las curvas representadas en la Figura B.2 estan tomadas en los intantes $t = \frac{1}{4}T$ (T es el tiempo del periodo), $t = \frac{1}{2}T$, $t = \frac{3}{4}T$ y $t = 0 = T$. En esta figura se puede observar la dirección de movimiento de la onda viajera. Por ejemplo, la onda viajera del atractor A se mueve, a través del tiempo, de izquierda a derecha, mientras que la onda viajera del atractor B se mueve de derecha a izquierda.

C. Integración y Derivación Numérica

C.1 Integración Numérica

Cuando se tiene un conjunto de EDOs :

$$\frac{d\mathbf{U}}{dt} = \mathbf{F}(\mathbf{U}; \boldsymbol{\mu}) \tag{C.1}$$

Se pueden utilizar métodos explícitos o implícitos para integrar el sistema en algún intervalo de tiempo τ.

Runge-Kutta de 4º orden

Los integradores explícitos más ampliamente usados pertenecen a la familia de integradores de Runge-Kutta. En particular Runge-Kutta de 4º orden es uno de los más robustos. El método esta dado por la siguiente fórmula:

$$\mathbf{U}_k = \mathbf{U}_{k-1} + \frac{\tau}{6}(\mathbf{F}1_{k-1} + 2\,\mathbf{F}2_{k-1} + 2\,\mathbf{F}3_{k-1} + \mathbf{F}4_{k-1}) \tag{C.2}$$

donde:

$$\begin{aligned}
\mathbf{F1}_{k-1} &= \mathbf{F}(\mathbf{U}_{k-1};\boldsymbol{\mu}) \\
\mathbf{F2}_{k-1} &= \mathbf{F}(\mathbf{U}_{k-1}+\tfrac{\tau}{2}\mathbf{F1}_{k-1};\boldsymbol{\mu}) \\
\mathbf{F3}_{k-1} &= \mathbf{F}(\mathbf{U}_{k-1}+\tfrac{\tau}{2}\mathbf{F2}_{k-1};\boldsymbol{\mu}) \\
\mathbf{F4}_{k-1} &= \mathbf{F}(\mathbf{U}_{k-1}+\tau\mathbf{F3}_{k-1};\boldsymbol{\mu})
\end{aligned}$$

Euler Implícito

El método de Euler implícito es uno de los integrdores más ampliamente usados en la familia de los integradores impícitos. El método viene dado por la siguiente fórmula:

$$\mathbf{U}_{k+1} = \mathbf{U}_k + \frac{\tau}{2}\left[\mathbf{F}(\mathbf{U}_k;\boldsymbol{\mu}) + \mathbf{F}(\mathbf{U}_{k+1};\boldsymbol{\mu})\right]$$

C.2 Diferencias Finitas

Aquí se presentan las primeras cuatro derivadas aproximadas por medio de diferencias finitas.

Segundo Orden

$$\begin{aligned}
\left(\frac{dU}{dX}\right)_i &\approx \frac{-U_{i-1}+U_{i+1}}{2\,\Delta X} \\
\left(\frac{d^2U}{dX^2}\right)_i &\approx \frac{U_{i-1}-2U_i+U_{i+1}}{\Delta X^2} \\
\left(\frac{d^3U}{dX^3}\right)_i &\approx \frac{-U_{i-2}+2U_{i-1}-2U_{i+1}+U_{i+2}}{2\,\Delta X^3} \\
\left(\frac{d^4U}{dX^4}\right)_i &\approx \frac{U_{i-2}-4U_{i-1}+6U_i-4U_{i+1}+U_{i+2}}{\Delta X^4}
\end{aligned} \tag{C.3}$$

Cuarto Orden

$$
\begin{aligned}
\left(\frac{dU}{dX}\right)_i &\approx \frac{U_{i-2}-8U_{i-1}+8U_{i+1}-U_{i+2}}{12\,\Delta X} \\
\left(\frac{d^2U}{dX^2}\right)_i &\approx \frac{-U_{i-2}+16U_{i-1}-30U_i+16U_{i+1}-U_{i+2}}{12\,\Delta X^2} \\
\left(\frac{d^3U}{dX^3}\right)_i &\approx \frac{U_{i-3}-8U_{i-2}+13U_{i-1}-13U_{i+1}+8U_{i+2}-U_{i+3}}{8\,\Delta X^3} \\
\left(\frac{d^4U}{dX^4}\right)_i &\approx \frac{-U_{i-3}+12U_{i-2}-39U_{i-1}+56U_i-39U_{i+1}+12U_{i+2}-U_{i+3}}{6\,\Delta X^4}
\end{aligned}
\tag{C.4}
$$