

Universidad EAFIT
Facultad de Ciencias Básicas

Modelo de Fitzhugh-Nagumo y las redes neuronales

Tesis de Maestría presentada por

Jasmidt Vera Cuenca, cod. 200410027016
Yineth Medina Arce, cod. 200410015016

Director

Dr. Mauro Montealegre Cárdenas

Maestría en Matemáticas Aplicadas

Departamento de Matemáticas
Facultad de Ciencias Exactas
Universidad Surcolombiana

Universidad EAFIT
Medellín

2007

Universidad EAFIT
Facultad de Ciencias Básicas

Modelo de Fitzhugh-Nagumo y las redes neuronales

Tesis de Maestría presentada por

Jasmidt Vera Cuenca, cod. 200410027016
Yineth Medina Arce, cod. 200410015016

Director

Dr. Mauro Montealegre Cárdenas

Maestría en Matemáticas Aplicadas

Departamento de Matemáticas
Facultad de Ciencias Exactas
Universidad Surcolombiana

Universidad EAFIT
Medellín

2007

Agradecimientos

Agradezco a Dios por haberme permitido llegar a la culminación de este sueño tan anhelado, a la Universidad Eafit por brindarme la oportunidad de acceder a los conocimientos; al Doctor Mauro Montealegre Cárdenas por su paciencia, perseverancia y apoyo incondicional para el desarrollo y culminación de este proyecto, a mi amigo y profesor Gustavo Londoño Betancour por sus conocimientos y tiempo; a mi madre, a mi hija Natalia por su tiempo, tolerancia, comprensión y voz de aliento para no desfallecer en la firme decisión de alcanzar esta meta y a todos los que de una u otra forma han contribuido para la realización de este logro significativo para mi vida.

Jasmidt Vera Cuenca

Quiero expresar mis agradecimientos a la universidad Eafit por contribuir con mi engrandecimiento personal y profesional y que a la vez permitiera alcanzar uno de mis grandes sueños.

Agradezco especialmente la colaboración de:

Al doctor Mauro Montealegre Cárdenas, mi asesor de tesis quien con sus conocimientos paciencia y oportunas orientaciones contribuyó a la culminación de esta tesis.

Al profesor Gustavo Londoño por ofrecernos sus conocimientos y orientaciones durante el transcurso de la tesis.

Dedico esta tesis:

A mis dos hijos, Javier y Luis Ernesto que me regalaron su tiempo para que este logro fuera posible.

A mis familiares

A mis amigos (amigas).

Yineth Medina Arce

Índice general

Introducción	1
1. Definiciones relacionadas con los procesos dinámicos no lineales	3
1.1. Sistemas gradientes [PE]	3
1.1.1. Definición	3
1.1.2. Conjugación Topológica	4
1.2. Variedades invariantes	6
1.2.1. Conjuntos invariantes [GU,HO]	6
1.2.2. Teorema de las variedades invariantes bajo difeomorfismos [PE]	6
1.2.3. Variedades invariantes para flujos φ asociados a ecuaciones diferenciales	7
1.2.4. Variedades invariantes para órbitas periódicas hiperbólicas	7
1.2.5. Teorema de la variedad central	7
1.3. Formas normales	8
1.4. Bifurcación de Hopf	10
1.4.1. Bifurcación de Hopf para ecuaciones diferenciales	10
1.4.2. Bifurcación de Hopf para mapeos	10
1.4.3. Bifurcación de Bogdanov-Takens	11
1.5. Análisis de la forma normal en \mathbb{R}^2 en torno a un foco (para una bifurcación de Hopf)	11
1.6. Compactificación de Poincaré de un sistema dinámico	14
1.6.1. La proyección estereográfica	14
1.6.2. La proyección central	15

1.6.3. Observación	16
1.7. Metodos numéricos para resolver ecuaciones diferenciales . . .	16
1.7.1. Método de Euler [Ki,Che]	17
1.7.2. El método de Runge-Kutta	17
2. Modelo de Fitzhugh-Nagumo para el potencial de una membrana	19
2.1. Fundamentos	19
2.2. Estructura de una neurona	21
2.3. Excitabilidad de la membrana celular	22
2.4. Modelo de Fitzhugh Nagumo	24
2.4.1. Estados de equilibrio y ceroquinas	25
2.4.2. Estabilidad del potencial de reposo y linealización . . .	27
2.5. Análisis paramétrico	28
2.5.1. Existencia de un único estado de equilibrio	29
2.5.2. Caso autónomo	30
2.5.3. Estimulación de la célula con una corriente externa	32
2.5.4. Conclusión del Capítulo 2	34
3. Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3	37
3.1. Definiciones y conceptos básicos	37
3.1.1. Análisis de la estabilidad lineal	40
3.2. Simetrías del sistema	41
3.3. Espacio vectorial tangente generado por los vectores propios .	44
3.4. Reducción de la variedad central y forma normal	46
3.5. Análisis numérico: bifurcación y caos	48
3.6. Los coeficientes $R_{ij}^{(2,3)}$	49
3.7. Bifurcaciones de órbitas homoclínicas, heteroclínicas y periódicas	50
3.7.1. Introducción	50
3.7.2. Órbitas heteroclínicas	51
3.7.3. Las bifurcaciones heteroclínicas	51
3.8. Bifurcaciones Homoclínicas para el modelo de Fitzhugh-Nagumo con su perturbación singular, sistema lento-rápido. .	52
3.9. Conclusiones del capítulo tres	55
4. Las dinámicas de las redes neuronales no lineales	57
4.1. El modelo de Hopfield	57
4.1.1. Identificación de sistemas dinámicos mediante redes neuronales	61
4.2. El paradigma de la propagación hacia atrás	64

4.3. Modelo de aprendizaje de una neurona débilmente no lineal	69
4.3.1. Modelo matemático de una neurona y su proceso de aprendizaje	69
4.3.2. Conjugación topológica de cascada de gradientes	70
4.3.3. Comentario	71
4.3.4. Neuronas débilmente no lineales	72
4.3.5. Dinámica del proceso de aprendizaje de una neurona no lineal	75
4.4. El movimiento errante y el fenómeno co-operativo en una red neuronal: un modelo de memoria asociativa	78
4.4.1. Modelo	79
4.4.2. Memoria asociativa	81
Conclusiones	83
Bibliografía	85

Introducción

Este trabajo de tesis es motivado por el conocimiento de la dinámica nerviosa y las relaciones básicas con las redes neuronales artificiales, problema fundamental de la biofísica y la tecnología contemporánea.

Para ello nos soportamos en los trabajos de R. Fitzhugh [FH]; P. Murganandam y M. Lasklmanan [MU,MA]; Andrzej Bielecki [BiI] y [BiII]; M. Atencia, G. Joya y F. sandoval [AT,JO,SA], Carolina Barriga y otros [BA,CARR,ON]; entre otros.

El método usado en nuestra tesis es el análisis cualitativo de los sistemas dinámicos que surgen de la actividad de la membrana nerviosa, tal como lo presenta en sus trabajos de J. Gukhenheimer and P. J. Holmes [GU,HO], Morris F. Hirsch and Stephen Smale [MO,ST], Mauro Montealegre y otros [MON,LON,POL], y Lawrence Perko [PE]. Buscando aplicaciones básicas en los procesos de aprendizaje y de reconocimiento de parámetros, a partir de los estudios de [BA,CARR,ON] y [BiI].

Nuestro trabajo tiene las siguientes partes: Capítulo uno, una introducción general sobre sistemas dinámicos no lineales; en particular sobre: la sistemas gradientes, conjugación topológica, compactificación de Poincaré, variedades invariantes (estable, inestables y central) y las formas normales.

En el capítulo dos, estudiamos el modelo Fitzhugh-Nagumo reducido a dos dimensiones según la versión de Hodgkin-Huxley, análogo al modelo de Van Der Pol ; hacemos un estudio de la dinámica en torno de los equilibrios, en particular, los equilibrios no hiperbólicos como los de tipo de Hopf,

2 Introducción

cuya existencia tiene significado relevante en la actividad de la membrana nerviosa; las simulaciones necesarias son realizadas usando el software Integra.

En el capítulo tres estudiamos el modelo de Fitzhugh-Nagumo en la versión espacial espacio-temporal dada por una ecuación diferencial parcial; la cual por medio del cambio de variable del tipo de onda viajera conduce a una ecuación diferencial ordinaria en tres dimensiones; via variedad central llegamos de nuevo a un estudio cualitativo bidimensional confirmando los resultados obtenidos en el capítulo dos, ampliando así el conocimiento de la bifurcación de Hopf. Aquí las técnicas son más elaboradas pues se requiere escribir las ecuaciones en formas normales especiales, además de técnicas especiales para reconocer bifurcaciones homoclínicas y heteroclínicas en la versión S. N. Chow, B. Deng y D. Terman. Los estudios a este nivel aún son incipientes en la literatura conocida.

En el capítulo cuatro, presentamos en su primera parte un resumen de los conceptos básicos de las redes neuronales artificiales en sus versiones de Hopfield y Back propagation; resaltando los aspectos fundamentales de sus aprendizajes y capacidades de reconocimiento de parámetros. En la segunda sección de este último capítulo estudiamos con detalle el modelo de red neuronal debilmente no lineal y caótico; también un modelo de memoria asociativa, el cual tiene el comportamiento caótico. Estos dos últimos modelos aproximan el fenómeno biológico con los desarrollos tecnológicos dados por las redes neuronales.

CAPÍTULO 1

Definiciones relacionadas con los procesos dinámicos no lineales

1.1. Sistemas gradientes [PE]

1.1.1. Definición

Sea $\Omega \subseteq \mathbb{R}^n$ un conjunto abierto, y $v : \Omega \rightarrow \mathbb{R}$ una función potencial de clase C^2 , entonces el sistema gradiente tiene la siguiente forma

$$\dot{\vec{x}} = \frac{d\vec{x}(t)}{dt} = -\text{grad } v(\vec{x}(t)) \quad (1.1.1)$$

donde

$$\text{grad } v(\vec{x}(t)) = \left(\frac{\partial v(\vec{x}(t))}{\partial x_1}, \frac{\partial v(\vec{x}(t))}{\partial x_2}, \dots, \frac{\partial v(\vec{x}(t))}{\partial x_n} \right)^T,$$

$\text{grad } v : \Omega \rightarrow \mathbb{R}^n$ es el campo vectorial gradiente. El sistema (1.1.1) tiene las siguientes propiedades:

- i) si $\dot{v} = \frac{\partial v}{\partial t} : \Omega \rightarrow \mathbb{R}$ es la derivada de v a lo largo de las trayectorias de (1.1.1); esto es,

$$\dot{v}(x) = \left. \frac{dv(v(t))}{dt} \right|_{t=0}$$

por la regla de la cadena tenemos

$$\begin{aligned} \dot{v}(x) &= Dv(\vec{x}(t)) \cdot \dot{\vec{x}}(t) \\ &= \langle \text{grad } v(\vec{x}), -\text{grad } v(\vec{x}) \rangle \\ &= -|\text{grad } v(x)|^2 \end{aligned} \quad (1.1.2)$$

4 Definiciones relacionadas con los procesos dinámicos no lineales

- ii) Para $c \in \mathbb{R}$ constante, si $\frac{\partial v(\vec{x})}{\partial x_n} \neq 0$, para $\vec{x} \in B_\epsilon(\vec{U})$, por el teorema de la función implícita tenemos que

$$v(x_1, x_2, \dots, x_{n-1}, g(x_1, x_2, \dots, x_{n-1})) = c,$$

$g : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ es una superficie $n - 1$ dimensional y de clase C^1 . Entonces si $\text{grad } v(\vec{x}) \neq 0$, esto es \vec{x} es un punto regular de (1.1.1), las curvas de nivel de v atraviesan transversalmente las trayectorias de (1.1.1), como se muestra en la siguiente figura para el potencial: $v(x, y) = x^2(x - 1)^2 + y^2$

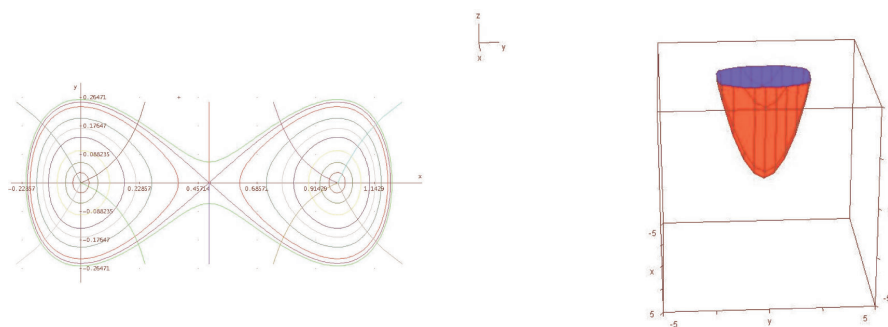


Figura 1.1.1 Sistema gradiente

- iii) La función $L : \Omega \rightarrow \mathbb{R}$, $L(\vec{x}) = v(\vec{x}) - v(\vec{x}_0)$, donde \vec{x}_0 es un punto de equilibrio de (1.1.1), esto es $\text{grad } v(\vec{x}_0) = 0$ es una función de Lyapunov; además los mínimos de v son equilibrios asintóticamente estables de (1.1.1).

1.1.2. Conjugación Topológica

- i) Sean dos dominios abiertos $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ para los campos vectoriales $f \in C^1(\Omega_1)$, $g \in C^1(\Omega_2)$ decimos que los dos sistemas

$$\begin{aligned} \dot{\vec{x}} &= g(\vec{x}) \\ \dot{\vec{y}} &= f(\vec{y}) \end{aligned} \tag{1.1.3}$$

son topológicamente conjugados si tienen la misma estructura dinámica y preservan la orientación dada por el tiempo, esto es, existe un homeomorfismo $H : \Omega_1 \rightarrow \Omega_2$ tal que

$$H_0 \phi_{t(\vec{x}, \tau)}(\vec{x}) = \psi_\tau \circ H(\vec{x}) \tag{1.1.4}$$

donde $t : \mathbb{R} \rightarrow \mathbb{R}$ para cada \vec{x} y además $\frac{\partial t}{\partial \tau} > 0$

Ejemplo 1.1.1. (Sobre el teorema de Hartman-Grobman [PE]) Si $\vec{0} \in \Omega \subseteq \mathbb{R}^n$ y $\vec{f} \in C^1(\Omega)$ con $\vec{f}(\vec{0}) = 0$ y además $A = Df(0)$, entonces para el flujo ϕ_t de $\dot{\vec{x}} = f(\vec{x})$ tenemos

$$H \circ \phi_t(\vec{x}) = e^{At} H(\vec{x})$$

donde H es un homeomorfismo de Ω en Ω ; $t \in I \subset \mathbb{R}$, I es un intervalo que contiene el origen del tiempo.

Ejemplo 1.1.2. Sea $\dot{\vec{x}} = \vec{f}(\vec{x})$ y para el cual construimos el nuevo campo vectorial $\dot{\vec{x}} = \frac{\vec{f}(\vec{x})}{1 + \|\vec{f}(\vec{x})\|}$, los cuales resultan ser topológicamente conjugados; sólo que el segundo es siempre un sistema dinámico porque su solución es

$$\varphi_t(x) = \vec{x}_0 + \int_0^t \frac{f(x(s))}{1 + \|f(x(s))\|} ds$$

con

$$\begin{aligned} f(\vec{x}_0) &= 0 \\ \varphi_t(\vec{x}) &: \Omega \times \mathbb{R} \longrightarrow \Omega \end{aligned}$$

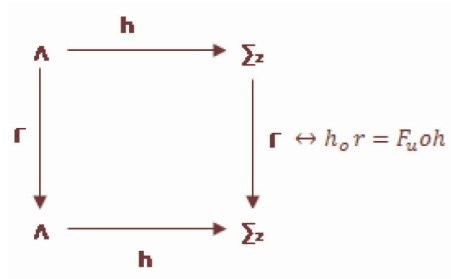
Ejemplo 1.1.3. Dos difeomorfismo $x_{n+1} = f(x_n)$ y $y_{n+1} = g(y_n)$ pueden ser topológicamente conjugados (y hasta diferencialmente conjugados), por ejemplo:

- i) $Q_c(x) = x^2 + c$ y $F_\mu(x) = \mu x(x - 1)$ son topológicamente conjugados via un homeomorfismo de la forma $h(x) = \alpha x + \beta$.
- ii) Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es un difeomorfismo con $Df(x) > 0, \forall x \in \mathbb{R}$; la ecuación diferencial es $\dot{x} = f(x) - x$ define un flujo $\varphi_t : \mathbb{R} \longrightarrow \mathbb{R}$, se puede mostrar que f y φ_1 son topológicamente conjugados.
- iii) Si dos orbitas periódicas de un sistema $\dot{\vec{x}} = f(\vec{x})$ son asintóticamente estables, entonces en vecindades de ellos se puede definir mapeos de Poincaré que resultan topológicamente conjugados.
- iv) Los mapeos $F(x) = 4x(x - 1)$ para $0 \leq x \leq 1$ y la aplicación tienda $G(x) = 2\left(\frac{1}{2} - \left|\frac{1}{2} - x\right|\right)$ son topológicamente conjugados, basta tomar el difeomorfismo $h(x) = 2\pi^{-1} \arcsen \sqrt{x}$. Se deduce que si $x_{n+1} = G(x_n)$ con $n = 0, 1, 2, \dots$ y $y_n = \sen^2\left(\frac{\pi x_n}{2}\right)$, entonces $y_{n+1} = F(y_n)$.
- v) Sea Λ el conjunto de cantor generado por la ecuación cuadrática $F_\mu(x) = \mu x(1 - x)$, para μ suficientemente grande, para la restricción siguiente

$$F_\mu|_\Lambda$$

existen dos intervalos I_0 y I_1 tal que $\Lambda \subseteq I_0 \cup I_1$; para $x \in \Lambda$ definimos el homeomorfismo $h(x) = s_0 s_1 s_2 \dots$ donde $s_J = 0$ si $F_\mu^J(x) \in I_0$, $s_J = 1$ si $F_\mu^J(x) \in I_1$, se puede mostrar que si $\mu > 2 + \sqrt{5}$ entonces

6 Definiciones relacionadas con los procesos dinámicos no lineales



donde

$$\sigma : \sum_2 \longrightarrow \sum_2 \quad (1.1.5)$$

es el mapeo shift.

1.2. Variedades invariantes

1.2.1. Conjuntos invariantes [GU,HO]

La órbita de un punto mediante un difeomorfismo $f(x_{n+1} = f(x_n))$, o mediante un flujo $\varphi(\varphi_t : x \longrightarrow \varphi_t(x))$, puede permanecer en una región especial del espacio de fases Ω . Un conjunto $\Lambda \subseteq \Omega$ es invariante para f (ó para φ) si $f^m(\Lambda) \subseteq \Lambda$ ó $(\varphi_t(\Lambda) \subseteq \Lambda); \forall m \in \mathbb{Z} (\forall t \in \mathbb{R})$. Es positivamente invariante si $m \in \mathbb{Z}^+$ ($t > 0$), o negativamente invariante si $m \in \mathbb{Z}^-$ ($t < 0$).

Ejemplo 1.2.1. El alfa limite de y , $\alpha(y)$ ó el omega limite de y , $w(y)$ son conjuntos invariantes, [PE].

1.2.2. Teorema de las variedades invariantes bajo difeomorfismos [PE]

Sea $f : \Omega \longrightarrow \mathbb{R}^n$ un difeomorfismo con un punto fijo $\vec{x}_0 \in \Omega$, esto es $f(\vec{x}_0) = \vec{x}_0$. Entonces existe una vecindad $N \subseteq \Omega$ con $\vec{x}_0 \in N$ que contiene las siguientes variedades estable y inestable:

$$W_{Loc}^s(\vec{x}_0) = \{\vec{x} \in N | f^m(\vec{x}) \longrightarrow \vec{x}_0 \text{ cuando } n \longrightarrow +\infty\}, \quad (1.2.1)$$

$$W_{Loc}^\mu(\vec{x}_0) = \{\vec{x} \in N | f^n(\vec{x}) \longrightarrow \vec{x}_0 \text{ cuando } n \longrightarrow -\infty\}, \quad (1.2.2)$$

\vec{x}_0 es hiperbólico si $Df(x_0)$ tiene todas las partes valores propios diferentes de (1.1.5).

Para \vec{x}_0 hiperbólico podemos extender las variedades locales invariantes anteriores a las variedades globales estables o inestables siguientes

$$W^s(\vec{x}_0) = \text{puntos de acumulación } \bigcup_{m \in \mathbb{Z}^+} f^{-m}(W_{Loc}^s(\vec{x}_0)), \quad (1.2.3)$$

$$W^u(\vec{x}_0) = \text{puntos de acumulación } \bigcup_{m \in \mathbb{Z}} f^m(W_{Loc}^u(\vec{x}_0)) \quad (1.2.4)$$

1.2.3. Variedades invariantes para flujos φ asociados a ecuaciones diferenciales

$\dot{\vec{x}} = f(\vec{x})$, con $f(\vec{x}_0) = 0$ equilibrio hiperbólico, tenemos las siguientes variedades invariantes locales:

$$W_{Loc}^s(\vec{x}_0) = \{x \in N \mid \varphi_t(x) \longrightarrow \vec{x}_0 \text{ cuando } t \rightarrow +\infty\},$$

$$W_{Loc}^u(\vec{x}_0) = \{x \in N \mid \varphi_t(x) \longrightarrow \vec{x}_0 \text{ cuando } t \rightarrow -\infty\};$$

con variedades globales:

$$W^s(\vec{x}_0) = \bigcup_{t>0} \varphi_{-t}(W_{Loc}^s(\vec{x}_0))$$

$$W^u(\vec{x}_0) = \bigcup_{t>0} \varphi_t(W_{Loc}^u(\vec{x}_0))$$

1.2.4. Variedades invariantes para órbitas periódicas hiperbólicas

Sea σ una órbita periódica de $\dot{\vec{x}} = f(\vec{x})$ se pueden definir $W^s(\sigma)$ y $W^u(\sigma)$, como se ilustra en la siguiente

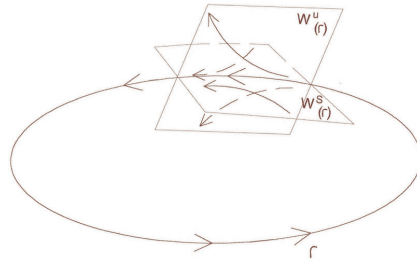


Figura 1.2.4 órbita periódica del tipo silla

1.2.5. Teorema de la variedad central

Si $f \in C^r$ es un campo vectorial, $f : \Omega \longrightarrow \mathbb{R}^n$, asociado a la ecuación diferencial $\dot{\vec{x}} = f(\vec{x}; \mu)$, \vec{x}_0 un punto de equilibrio no-hiperbólico ($f(\vec{x}_0) = 0$) con $A = Df(\vec{x}_0; \mu)$, el espectro σ de A se divide en tres partes: $\sigma_s, \sigma_\mu, \sigma_c$ con (podemos asumir $\vec{x}_0 = 0$):

$$Re(\lambda) = \begin{cases} < 0 & \text{si } \lambda \in \sigma_s \\ = 0 & \text{si } \lambda \in \sigma_c \\ > 0 & \text{si } \lambda \in \sigma_\mu \end{cases}$$

8 Definiciones relacionadas con los procesos dinámicos no lineales

con respectivos autoespacios E^s, E^u, E^c . Entonces también existe una variedad invariante W^c , la variedad central, tangente a E^c en \vec{x}_0 de clase C^{r-1} , la cual es el gráfico siguiente

$$W^c = \{(x_s, x_c, x_\mu) / x_c = h(x_s, x_\mu); h(0) = Dh(0) = 0\}$$

1.3. Formas normales

Sea la ecuación diferencial siguiente

$$\dot{x} = Ax + h_r(x) \quad \text{con } x \in \mathbb{R}^n \quad (1.3.1)$$

y el siguiente polinomio homogéneo $h_r(x) = h^2(x) + h^3(x) + \dots + h^r(x)$, $h^k(x)$ es de grado k , mediante el siguiente cambio de coordenadas próximas a la unidad,

$$\xi(y) = y + \xi^k(y), \quad y \in \mathbb{R}^n; \quad (1.3.2)$$

cuya derivada es,

$$D\xi^k(y) = I + \xi_y^k(y); \quad (1.3.3)$$

$\xi_y(y)$ con inversa

$$(\xi_y(y))^{-1} = I - \xi_y^k(y) + 0(|y|^{2k-2}) \quad (1.3.4)$$

Tenemos que para $y \in \Omega_k$, Ω_k una pequeña vecindad del origen en \mathbb{R}^n , tenemos que el sistema (1.3.1) se transforma en el siguiente sistema en la variable y :

$$\dot{y} = Ay + h^2(y) + h^{r-1}(y) + \{h^r(y) - [D\xi^k(y)Ay - A\xi^k(y)]\} + 0(|y|^{k+1}) \quad (1.3.5)$$

Ahora sea H_n^k el espacio vectorial de los campos de vectores de polinomios homogéneos (sin coeficientes independientes) con n variables de grado k . Definimos el operador de Lie L_A^K :

$$\begin{aligned} L_A^K &= H_n^k \longrightarrow H_n^k \\ (L_A^K \xi^k) &= \xi^k(y)Ay - AD\xi^k(y), \end{aligned} \quad (1.3.6)$$

L_A^K es denominado el colchete de Lie, y es tal que

$$H_n^k = R^k \oplus \mathcal{V}^k; \quad (1.3.7)$$

donde:

$$R^k = \text{imagen } (L_A^K)(H_n^k) \quad (1.3.8)$$

$$\mathfrak{v}^k = \ker (L_A^k)(H_n^k) \quad (1.3.9)$$

tomamos la base del espacio vectorial H_n^k dada por los polinomios $\{x^\alpha \cdot \exp_i\}_{i=1, \dots, n}$ donde

$$\begin{aligned} \alpha^k &= x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \quad \text{con} \quad k_1 + k_2 + \dots + k_n = k \\ e^i &= (0, 0, \dots, 1, 0, \dots, 0), 1 \quad \text{en la posición } i\text{.} \end{aligned} \quad (1.3.10)$$

En este contexto tenemos el siguiente teorema y definiciones

Teorema 1.3.1. *Sea $\vec{X}(\vec{x})$ un campo de vectores diferenciable en \mathbb{R}^n , para la ecuación diferencial (1.3.1) con $X(0) = \vec{0}$, entonces existe una transformación polinomial en las nuevas coordenadas \vec{y} tal que este sistema se transforma en:*

$$\dot{y} = Ay + g^2(y) + \dots + g^r(y) + 0(|y|^{r+1}) \quad (1.3.11)$$

$$y \in \Omega_k, \quad \text{donde} \quad g^k \in \mathfrak{v}^k \quad \text{para} \quad k = 2, \dots, r \quad (1.3.12)$$

Definición 1.3.2. *Sea el vector de valores propios de (1.3.1) $\vec{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ entonces para $m \cdot \lambda = \lambda_1^{m_1} \cdot \lambda_2^{m_2} \dots \lambda_n^{m_n}$ obtenemos*

$$\wedge_{m,i} = (\vec{m} \cdot \vec{\lambda}) - \lambda_i, \quad (1.3.13)$$

Se puede probar que

$$L_A^K(x^\alpha \cdot e_j) = (\lambda \cdot \alpha - \lambda_j)x^\alpha e_j \quad (1.3.14)$$

Definición 1.3.3. *Una n -dupla de autovalores $\vec{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ es **resonante** de orden r , si*

$$\lambda_i = \sum_{j=1}^n m_j \lambda_j \quad (1.3.15)$$

Para $m_j \in \mathbb{N}$, con $\sum_{j=1}^n m_j = r$, y algún $j = 1, \dots, n$. Podemos notar que $h_r(x)$ de (1.3.1) puede desaparecer si escogemos ξ^k tal que

$$L_A^k(\xi^k(y)) = h_r(y), \quad (1.3.16)$$

entonces escogiendo la base canónica $\{x^\alpha \cdot e_i\}_{i=1, \dots, n}$ tenemos

$$\xi_r(x) = \sum_{m,i} \xi_{m,i} x^m e_i \quad (1.3.17)$$

Sustituyendo (1.3.17) en (1.3.16) vemos que si los autovalores no son resonantes de orden r , entonces

$$\xi_{m,i} = \frac{X_{m,i}}{(\vec{m} \cdot \vec{\lambda} - \lambda_i)} \quad (1.3.18)$$

donde $X_{m,i}$ con los coeficientes del campo vectorial (1.3.1)

10 Definiciones relacionadas con los procesos dinámicos no lineales

1.4. Bifurcación de Hopf

1.4.1. Bifurcación de Hopf para ecuaciones diferenciales

Consideremos un sistema que localmente tiene la siguiente forma

$$\begin{cases} \dot{x} = f(x, y) = f_{x\mu}\mu x - (w - f_{y\mu}\mu)y + \dots = \mu x - wy - (x^2 + y^2)x + \dots \\ \dot{y} = g(x, y) = (w + g_{x\mu}\mu)x + g_{y\mu}\mu y + \dots = wx + \mu y - (x^2 + y^2)y + \dots; \end{cases} \quad (1.4.1)$$

El cual en coordenadas polares es

$$\begin{cases} \dot{r} = \alpha(\mu) + R_e A(\mu)r^3 + 0(r^5) \\ \dot{\theta} = \beta(\mu) + I_m A(\mu)r^2 + 0(r^4) \end{cases}$$

o en términos de $z = x + iy$ y $\lambda(\mu) = \alpha(\mu) + i\beta(\mu)$ es equivalente $\dot{z} = \lambda(\mu)z + A(\mu)|z|^2z + 0(|z|^2)$. Resulta que (1.4.1) tiene un equilibrio en el origen, el cual es un foco y una órbita periódica de radio aproximado a $r = \sqrt{-\frac{\alpha(\mu)}{R_e A(\mu)}}$ y período aproximado a $\frac{2\pi}{\beta(\mu)}$. De la ecuación característica de este sistema en torno del origen, tenemos que sus valores propios

$$\lambda(\mu) = \alpha(\mu) \pm i\beta(\mu)$$

con

$$\begin{aligned} \alpha(\mu) &= \frac{1}{2}(f_{\mu x} + g_{\mu y})\mu + 0(\mu^2), \\ \beta(\mu) &= w(1 + \frac{\mu}{2w}(g_{\mu x} - f_{\mu y})) + 0(\mu^2), \end{aligned}$$

con la condición genérica para la bifurcación de Hopf siguiente

$$\frac{d}{d\mu}\alpha(\mu)|_{\mu=0} \neq 0$$

1.4.2. Bifurcación de Hopf para mapeos

$$F_\lambda \begin{pmatrix} x \\ y \end{pmatrix} = (\lambda + \beta(x^2 + y^2)) \begin{pmatrix} \cos(\theta_0 + \mu r^2), & -\sin(\theta_0 + \mu r^2) \\ \sin(\theta_0 + \mu r^2), & \cos(\theta_0 + \mu r^2) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1.4.2)$$

donde $\exp(ni\alpha) \neq 1$ para $n = 1, 2, 3, 4$. el cual tiene un punto fijo $(0, 0)$; este sistema se expresa en coordenadas polares como

$$\begin{cases} r_1 = \lambda r + \beta r^3 + 0(r^5) \\ \theta_1 = \theta + \alpha + \gamma r^2 + 0(r^4) \end{cases}, \quad (1.4.3)$$

1.5 Análisis de la forma normal en \mathbb{R}^2 en torno a un foco (para una bifurcación de Hopf) 11

de lo que se deduce la existencia de una órbita periódica en $r = \sqrt{\frac{(1-\lambda)}{\beta}}$, la estabilidad de la órbita depende del signo de β (la órbita de Hopf puede ser subcrítica o supercrítica). El sistema (1.4.3) en variable compleja tiene la siguiente expresión

$$z_{n+1} = 1 + \alpha\mu + 0(\mu^2) \exp(i\theta(\mu))z_n + A(\mu)|z_n|^2 z_n + 0(|z_n|^2)$$

con la condición generica $\frac{d|\lambda|}{d\mu}(0) = \alpha$, de $\lambda(\mu) = 1 + \alpha\mu + 0(\mu^2)$.

1.4.3. Bifurcación de Bogdanov-Takens

corresponde al sistema bidimensional de dos parámetros, ver [PE]

$$\begin{cases} \dot{x} = y \\ \dot{y} = \mu_1 + \mu_2 y + x^2 + xy \end{cases} \quad (1.4.4)$$

1.5. Análisis de la forma normal en \mathbb{R}^2 en torno a un foco (para una bifurcación de Hopf)

Seguiremos con el procedimiento de la forma normal dada por Crawford para el flujo en dos dimensiones

$$\dot{X} = V(X) \equiv V^{(1)}(X) + V^{(2)}(X) + \dots, \quad X \equiv (x, y) \in \mathbb{R}^2 \quad (1.5.1)$$

$$V^{(1)}(X) = \begin{pmatrix} \gamma x + w y \\ -w x + \gamma y \end{pmatrix} \quad (1.5.2)$$

Sea la transformación

$$X' = X + \phi^{(k)}(X) \quad (1.5.3)$$

cuya inversa es

$$X' = X - \phi^{(k)}(X) + 0(x^{2k+1}) \quad (1.5.4)$$

donde

$$\phi^k : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \quad (1.5.5)$$

es un mapeo polinómico homogéneo de grado k . Esta transformación en la ecuación (1.5.1) nos lleva a

$$\dot{X}' = V(X) - L^k[\phi^{(k)}(X)] + 0(x^{k+1}) \quad (1.5.6)$$

donde

$$L^k[\phi^{(k)}(X)] = DV^{(1)}(X)\phi^{(k)}(X) - D\phi^{(k)}(X)V^{(1)}(X). \quad (1.5.7)$$

12 Definiciones relacionadas con los procesos dinámicos no lineales

Para cambiar todos los términos de grado k , uno debe solucionar,

$$V^{(1)}(X) - L[\phi^{(k)}(X)] = 0 \quad (1.5.8)$$

o sea

$$\phi^{(k)}(X) = (L^k)^{-1}[V^{(1)}(X)]. \quad (1.5.9)$$

Aquí L^k puede ser visto como una matriz dimensionalmente finita y sera posible cambiar todos los términos de orden k si y solo si $\det L \neq 0$. Así $\phi^{(k)}(X)$ puede ser asumida como un vector en el espacio dimensional $2(k+1)$, a saber $\mathcal{K}^{(k)}(\mathbb{R}^2)$, esto es, $\phi^{(k)}(X) \in \mathcal{K}^{(k)}(\mathbb{R}^2)$ y L^k es una transformación lineal que actua sobre este espacio

$$L^k : \mathcal{K}^{(k)}(\mathbb{R}^2) \rightarrow L[\mathcal{K}^{(k)}(\mathbb{R}^2)] \quad (1.5.10)$$

$\mathcal{K}^{(k)}(\mathbb{R}^2) = H_2^k$ si el determinante $\det L^k = 0$, entonces L^k tiene uno de los valores propios cero. En la presencia de tal valor propio cero uno puede especificar un subespacio completo $\mathfrak{S}^{(k)}$, como en (1.3.7),

$$\mathcal{K}^{(k)}(\mathbb{R}^2) = L[\mathcal{K}^{(k)}(\mathbb{R}^2)] + \mathfrak{S}^{(k)} \quad (1.5.11)$$

Una vez $\mathfrak{S}^{(k)}$ es elegido, entonces los términos de orden k pueden ser escritos como

$$V^{(k)} = V_r^{(k)} + V_g^{(k)} \quad (1.5.12)$$

con

$$V^{(k)} \in L[\mathcal{K}^{(k)}(\mathbb{R}^2)] \quad (1.5.13)$$

y

$$V_g^{(k)} \in \mathfrak{S}^{(k)} \quad (1.5.14)$$

la componente $V_r^{(k)}$ desaparece mediante este cambio de coordenadas. Asumiremos la forma de $\phi^{(k)}$ como

$$\phi^{(k)}(x, y) = \begin{pmatrix} \phi_x^{(k)}(x, y) \\ \phi_y^{(k)}(x, y) \end{pmatrix} \quad (1.5.15)$$

donde

$$L^k[\phi^{(k)}(x, y)] = \begin{pmatrix} \gamma & w \\ -w & \gamma \end{pmatrix} \begin{pmatrix} \phi^{(k)}(x) \\ \phi^{(k)}(y) \end{pmatrix} - \begin{pmatrix} \frac{\partial \phi^k(x)}{\partial y} & \frac{\partial \phi^k(x)}{\partial y} \\ \frac{\partial \phi^k(y)}{\partial x} & \frac{\partial \phi^k(y)}{\partial y} \end{pmatrix} \times \begin{pmatrix} \gamma x & wy \\ -wx & \gamma y \end{pmatrix} \quad (1.5.16)$$

1.5 Análisis de la forma normal en \mathbb{R}^2 en torno a un foco (para una bifurcación de Hopf) 13

Para determinar los valores propios de L^k es mas conveniente usar coordenadas complejas (z, \bar{z}) . Entonces la ecuación (1.5.15) escrita en variable complejas como

$$L^k[\phi^{(k)}(z, \bar{z})] = \begin{pmatrix} \gamma - iw & 0 \\ 0 & \gamma + iw \end{pmatrix} \begin{pmatrix} \phi^{(k)}(z) \\ \phi^{(k)}(\bar{z}) \end{pmatrix} - \begin{pmatrix} \frac{\partial \phi^{(k)}(z)}{\partial z} & \frac{\partial \phi^{(k)}(z)}{\partial \bar{z}} \\ \frac{\partial \phi^{(k)}(\bar{z})}{\partial z} & \frac{\partial \phi^{(k)}(\bar{z})}{\partial \bar{z}} \end{pmatrix} \begin{pmatrix} (\gamma - iw)z \\ (\gamma + iw)\bar{z} \end{pmatrix} \quad (1.5.17)$$

La ecuación de arriba tiene vectores propios de la forma $(O(k), 0)$ y $(0, O(k))$. Supongamos el vector dado por

$$\xi_+^{(k,l)} = \begin{pmatrix} z^l \bar{z}^{k-l} \\ 0 \end{pmatrix}$$

y

$$\xi_-^{(k,l)} = \begin{pmatrix} 0 \\ z^l \bar{z}^{k-l} \end{pmatrix}, \quad l = 0, 1, 2, \dots, k \quad (1.5.18)$$

y sean los vectores propios de L^k y usando una base para $H^{(k)}(\mathbb{R}^2)$. Entonces los valores propios de L pueden obtenerse de la forma

$$L^k[\xi_{\pm}^{(k,l)}] = \lambda_{\pm}^{(k,l)} \xi_{\pm}^{(k,l)} \quad (1.5.19)$$

donde

$$\lambda_{\pm}^{(k,l)} = (1 - k)\gamma - iw(k - 2l \pm 1) \quad (1.5.20)$$

Para encontrar los términos de orden k que no se pueden quitar por la transformación uno tiene que encontrar la condición para valores propios cero (puesto que $\det L^k = 0$ implica que al menos L^k tiene un valor propio cero). Esto igualando a cero tanto la parte real como la imaginaria de (1.5.20) obtenemos:

$$(1 - k)\gamma = 0 \quad (1.5.21)$$

$$(k - 2l \pm 1)w = 0 \quad (1.5.22)$$

En la bifurcación de Hopf $\gamma = 0$ y $w \neq 0$ e implica que $k - 2l \pm 1 = 0$. Ya que $2l \pm 1$ es impar, para k par nunca se cumplira la ecuación (1.5.22) y para k impar, esto es $k = 2i \pm 1$ tenemos dos vectores propios:

$$\xi_+^{(k,(k+1)/2)} = \begin{pmatrix} z^l |z|^{k-l} \\ 0 \end{pmatrix}$$

14 Definiciones relacionadas con los procesos dinámicos no lineales

y

$$\xi_-^{(k,(k+1)/2)} = \begin{pmatrix} 0 \\ z^l |z|^{k-l} \end{pmatrix} \quad k = 3, 5, 7, \dots \quad (1.5.23)$$

Estos dos vectores forman una base para el subespacio complementario (1.5.11) $\mathfrak{S}^{(k)}$. Mediante la aplicación de la transformación de coordenadas sucesivamente uno puede obtener la forma normal de todo orden par cambiando los términos no lineales,

$$\begin{pmatrix} \dot{z} \\ \dot{\bar{z}} \end{pmatrix} = \begin{pmatrix} \gamma - iw & 0 \\ 0 & \gamma + iw \end{pmatrix} \begin{pmatrix} z \\ \bar{z} \end{pmatrix} + \sum_{j=1}^{\infty} \begin{pmatrix} \alpha_j z |z|^{2j} \\ \bar{\alpha}_j \bar{z} |z|^{2j} \end{pmatrix} \quad (1.5.24)$$

la cual puede ser escrita en coordenadas polares en la forma estandar

$$\dot{r} = r \left[\gamma + \sum_{j=1}^{\infty} a_j r^{2j} \right] \quad (1.5.25)$$

$$\dot{\theta} = w + \sum_{j=1}^{\infty} b_j r^{2j} \quad (1.5.26)$$

donde $a_j = \text{Re}(\alpha_j)$ y $b_j = -\text{Im}(\alpha_j)$.

1.6. Compactificación de Poincaré de un sistema dinámico

1.6.1. La proyección estereográfica

Se trata de proyectar la esfera $S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid \|(x, y, z)\| = 1\} \subseteq \mathbb{R}^3$ en el plano (x, y) ; ubicado en el polo norte de acuerdo con la proyección π :

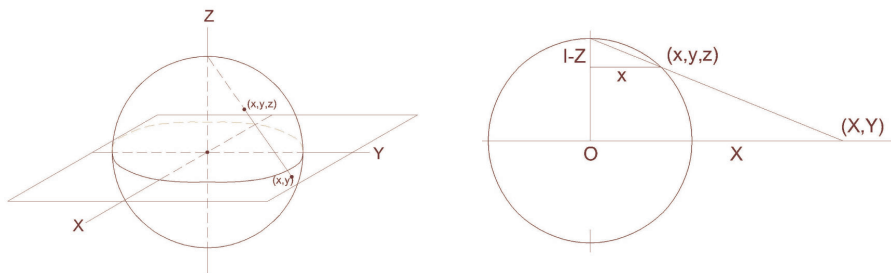


Figura 1.6.1 Proyección estereográfica

$$\pi = \begin{cases} x = \frac{X}{1-Z} \\ y = \frac{Y}{1-Z} \end{cases}$$

denominada proyección estereográfica; en este caso el polo norte $(0, 0, 1)$ es proyectado al infinito del plano, $\|(x, y)\| \rightarrow +\infty$. Los puntos dentro del círculo $x^2 + y^2 = 1$ corresponden a la semiesfera inferior; el origen corresponde al punto $(0, 0, -1)$; los puntos fuera del círculo $x^2 + y^2 = 1$ corresponden al hemisferio superior. Usando esta proyección podemos ver el sistema tridimensional

$$\dot{\vec{x}} = F(\vec{x}(t)) \iff \begin{cases} \dot{x} = -y + xz^2 \\ \dot{y} = x + yz^2 \\ \dot{z} = -z(x^2 + y^2) \end{cases}$$

Es tal que si $v(x, y, z) = x^2 + y^2 + z^2$ entonces $\dot{v} = 0$, esto es su flujo esta en la esfera S^2 ; el flujo en la vecindad del polo norte es repulsor, y esta cualidad en el infinito se puede calcular mediante su proyección en el plano.

1.6.2. La proyección central

De acuerdo a la gráfica los puntos $(X, Y, Z) \in S^2$ se proyectan al plano π ubicado tangente al polo norte con coordenadas (x, y) mediante el mapeo

$$C(X, Y, Z) = \left(\frac{X}{Z}, \frac{Y}{Z} \right) = (x, y)$$

$$C^{-1}(x, y) = \left(\frac{x}{\sqrt{1+x^2+y^2}}, \frac{y}{\sqrt{1+x^2+y^2}}, \frac{1}{\sqrt{1+x^2+y^2}} \right),$$

con mapeo inverso $C^{-1} : \pi \rightarrow S^2$,

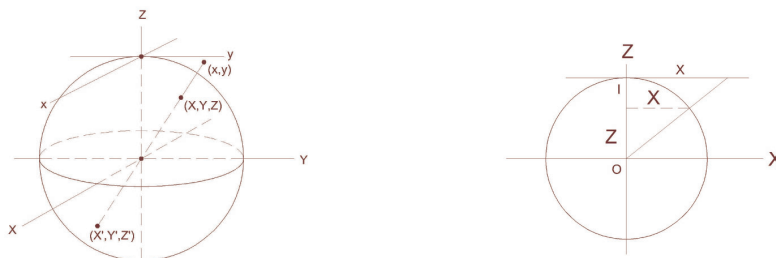


Figura 1.6.2. Proyección central

16 Definiciones relacionadas con los procesos dinámicos no lineales

En coordenadas locales de la esfera en un entorno de $x = 1$, tenemos $C_x(1, y, z) = (\frac{1}{z}, \frac{u}{z})$; en coordenadas locales de un entorno de $y = 1$, tenemos que $C_y(x, 1, z) = (\frac{u}{z}, \frac{1}{z})$. Los mapeos de C_x, C_y , y análogamente C_z se llaman mapeos de compactificación de Poincaré; C_x, C_y se usan para conocer el comportamiento del campo $\dot{\vec{x}} = f(\vec{x})$ en el infinito, el ecuador de la esfera. En el ecuador de S^2 pueden quedar nuevas singularidades y órbitas regulares correspondiendo al infinito del plano así:

$$\begin{cases} \frac{\partial x}{\partial t} = P(x, y) \\ \frac{\partial y}{\partial t} = Q(x, y) \end{cases} \implies \begin{cases} \frac{\partial u}{\partial t} = -uzP(\frac{1}{z}, \frac{u}{z}) + zQ(\frac{1}{z}, \frac{u}{z}) \\ \frac{\partial z}{\partial t} = -z^2P(\frac{1}{z}, \frac{u}{z}) \end{cases}$$

haciendo $d\tau = \frac{dt}{z^n}$, donde n es el mayor grado entre los polinomios P y Q , obtenemos

$$\begin{cases} \frac{\partial u}{\partial \tau} = P^*(u, z) \\ \frac{\partial z}{\partial \tau} = Q^*(u, z) \end{cases}$$

sistema el cual esta definido en $z = 0$, esto es en el infinito.

1.6.3. Observación

Como S^2 se puede expresar en coordenadas esféricas, $C : (\rho, \phi, \theta) \longrightarrow (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \theta)$, sus coordenadas son $\psi(x, y, z) = (\arccos(z), \arctan(\frac{y}{x})) = (\phi, \theta)$, donde se toma $\phi = \arccos(x^2 + y^2 + 1)^{-\frac{1}{2}}$. La inversa de $\psi_0 C^{-1}$ es el mapeo $P(\phi, \theta) = (\tan \phi \cos \theta, \tan \phi \sin \theta)$, y resulta (1.1.5) que el campo vectorial definido en el plano tangente π es dado por la siguiente expresión

$$\dot{y} = y(\phi, \theta) = (DP)^{-1}[f(P(\phi, \theta))]$$

1.7. Metodos numéricos para resolver ecuaciones diferenciales

Para la ecuación diferencial ordinaria

$$\frac{dx}{dt} = f(t, x)$$

con $a \leq t \leq b$, con condiciones iniciales $y(a) = \alpha$; tenemos por lo menos los siguientes métodos numéricos para resolver este tipo de ecuaciones diferenciales:

1.7.1. Método de Euler [Ki,Che]

Dividimos el intervalo $[a, b]$ en N puntos con tamaño del paso $h = \frac{b-a}{N}$, y hacemos $t_k = a + kh$ para cada $k = 0, 1, 2, \dots, N$ si $y(t)$ es la solución única de la ecuación diferencial, la expandimos usando el teorema de Taylor, de la siguiente forma:

$$y(t_{k+1}) = y(t_k) + (t_{k+1} - t_k) \cdot y'(t_k) + \frac{(t_{k+1} - t_k)^2}{2} \cdot y''(\xi_k)$$

para algún $\xi_k \in (t_k, t_{k+1})$. Esto es si $h = t_{k+1} - t_k$, entonces

$$w_{k+1} := y(t_{k+1}) = y(t_k) + hy'(t_k) + \frac{h^2}{2}y''(\xi_k),$$

donde podemos asumir que $O(h^2) = 0$

1.7.2. El método de Runge-Kutta

Da el siguiente algoritmo para el cálculo de pendientes:

$$\begin{aligned} w_{k+1} &= w_k + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), & i = 0, 1, 2, \dots, N - 1 \\ w_0 &= \alpha \\ k_1 &= hf(t_k, w_k) \\ k_2 &= hf\left(t_k + \frac{h}{2}, w_k + \frac{k_1}{2}\right) \\ k_3 &= hf\left(t_k + \frac{h}{2}, w_k + \frac{k_2}{2}\right) \\ k_4 &= hf(t_{k+1}, w_k + k_3). \end{aligned} \tag{1.7.1}$$

CAPÍTULO 2

Modelo de Fitzhugh-Nagumo para el potencial de una membrana

2.1. Fundamentos

El cerebro es un sistema complejo y la comprensión de la actividad cerebral, por su importancia y dificultad, constituye uno de los grandes retos de la ciencia moderna. No es posible prescindir del uso de modelos matemáticos para entender la funcionalidad de la mente, en términos de las bases fisicoquímicas de la fisiología del cerebro. La aproximación matemática al estudio del cerebro y del sistema nervioso en general contempla, entre otros aspectos, la construcción y el análisis de modelos de las unidades fundamentales que los constituyen: las células nerviosas o neuronas.

Es un hecho notable que los fenómenos eléctricos juegan un papel determinante en la fisiología de las células nerviosas. En el siglo XVII, a partir de los experimentos realizados por Luigi Galvani en la Universidad de Bolonia durante la década de 1780-90, se demostró que la aplicación de corrientes eléctricas provoca una acción neuromotora (contracciones) en las ancas de ranas. Después, a mediados del siglo XIX se constató que la acción nerviosa y muscular no se debe tanto a la circulación de corrientes eléctricas a lo largo de las fibras nerviosas, sino que es concomitante a variaciones del voltaje, y la consecuente circulación de corrientes eléctricas, a través de la membrana celular. En esta época, Carlo Matteucci y Emil du Bois-Reymond lograron medir estos flujos de cargas, tanto en músculos como en nervios,

20 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

en presencia de contracciones musculares o bajo la aplicación de pulsos eléctricos externos. Posteriormente, en 1868, Julius Bernstein -discípulo de du Bois-Reymond- logró medir el curso temporal de los cambios del voltaje de la membrana de una fibra nerviosa, dando así inicio a la electrofisiología moderna.

Un avance importante tuvo lugar a mediados del siglo pasado. Basados en una serie de experimentos, que fueron interrumpidos por la Segunda Guerra Mundial, los ingleses Alan Lloyd Hodgkin y Andrew Fielding Huxley [FH] estudiaron experimentalmente la dinámica del voltaje transmembranal en el axón (gigante) de una neurona de calamar, al ser aplicada una estimulación eléctrica externa. En 1952, a partir de estos estudios, dichos investigadores propusieron una teoría biofísica y un modelo con cuatro ecuaciones diferenciales no lineales, que explicó y reprodujo los resultados experimentales conocidos hasta el momento. Su estudio reveló que la onda del voltaje observada o espiga, que se propaga a lo largo de las fibras nerviosas, se debe a la permeabilidad iónica -selectiva y dependiente, a su vez, del mismo voltaje a través de la membrana de la célula nerviosa. Debido a estas investigaciones, Hodgkin y Huxley, recibieron en 1963 -junto con Sir John Carew Eccles- el premio Nobel de Medicina y Fisiología.

El modelo de Hodgkin y Huxley consta de cuatro ecuaciones diferenciales no lineales. Este tipo de ecuaciones tienen que ser estudiadas recurriendo a métodos cualitativos de análisis y simulaciones computacionales, basadas en la solución numérica del sistema. Para lograr una mejor comprensión del fenómeno se han construido modelos que simplifican la estructura matemática del sistema Hodgkin y Huxley, capturando exclusivamente la esencia dinámica de algunos de los procesos involucrados. Estos modelos simplificados proveen esquemas teóricos que permiten la comprensión de los fenómenos estudiados, a diferentes niveles de detalle.

El trabajo de Richard FitzHugh, en el laboratorio de Biofísica de los National Institutes of Health, USA, constituyó un gran avance en esta dirección. Basado en el trabajo previo de Balthazar van der Pol, propuso una simplificación considerable del modelo de Hodgkin y Huxley. Su modelo consta de dos ecuaciones diferenciales de primer orden, una lineal y otra cúbica. Siendo susceptible de un análisis bastante completo, este sistema permite una comprensión cualitativa del fenómeno de excitabilidad, desde el punto de vista de la dinámica matemática y constituye un modelo clásico de la neurofisiología. Otros fenómenos importantes, como la producción de ráfagas periódicas de impulsos nerviosos (bursting) no pueden ser explicados con el modelo de FitzHugh y requieren modelos de mayor complejidad

matemática [BA,CARR,ON].

Simultánea e independientemente del trabajo de FitzHugh, el investigador japonés Jin-ichi Nagumo propuso como análogo neuronal, un circuito eléctrico no lineal, gobernado por un sistema de dos ecuaciones también semejantes a las de van der Pol.

Actualmente, al análogo simplificado propuesto por estos autores, se le conoce como modelo de FitzHugh-Nagumo. La importancia actual de este modelo trasciende el ámbito de la biofísica y la neurofisiología, siendo de interés para los profesionales de otras ramas de las ciencias que necesitan comprender la constelación de fenómenos no lineales, que son concomitantes al fenómeno de excitabilidad. Particularmente, el modelo es de interés para ingenieros y científicos interesados en el estudio de sistemas de integración, disparo y oscilaciones de relajación que, típicamente se producen en éstos.

2.2. Estructura de una neurona

Las células nerviosas varían en su forma y tamaño pero prototípicamente están constituidas por tres partes principales: El cuerpo celular o soma, las dendritas y el axón (ver figura 2.1). El soma contiene al núcleo de la célula y por lo tanto es poseedor del material genético de la neurona. Aquí ocurren los mecanismos bioquímicos sintetizadores de enzimas y demás proteínas necesarias para mantenerlas vivas. Las dendritas y el axón son filamentos con múltiples ramificaciones que le permiten a la célula nerviosa recibir o transmitir señales a otras células.

Tradicionalmente, se cree que en el soma se lleva a cabo un proceso de integración de las señales provenientes de las dendritas que terminan con el envío de una respuesta concordante hacia otras células receptoras, a través de una larga fibra que es el axón, la interacción entre las células se produce a través de conexiones llamadas *sinapsis*. Estas se clasifican en dos tipos: las sinapsis químicas y las eléctricas. En la sinapsis químicas existe un espacio (intersináptico) que separa las neuronas de ambas células y por lo tanto es necesaria la presencia de agentes químicos (neurotransmisor) que se difunden a través del espacio intersináptico para posibilitar el restablecimiento de la señal eléctrica. Por su parte, en las eléctricas existe una conexión directa que permite el paso de corrientes eléctricas de una célula a otra, sin la mediación de neurotransmisores.

22 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

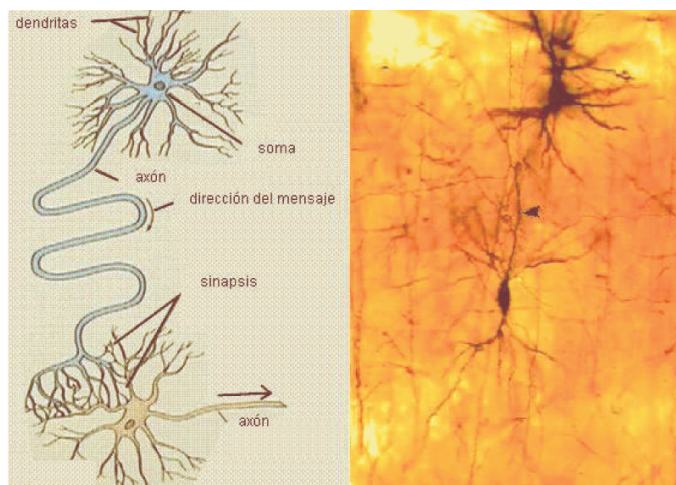


Figura 2.1 Esquema de una neurona típica (izquierda). Microfotografía de un par de neuronas de la corteza cerebral de una rata de 19 días de vida (derecha).

En los circuitos cerebrales, una sola neurona sólo puede recibir entre 10,000 y 20,000 sinapsis de cientos de neuronas, y mandar proyecciones a cientos y a veces miles de neuronas blanco. La funcionalidad del cerebro emerge de la suma de todas las interacciones que tienen lugar entre las células constituyentes. El sustrato de esta funcionalidad tiene una complejidad extraordinaria. Un cerebro cuenta con varios cientos de miles de millones de neuronas, que individualmente pueden llegar a tener conexiones con varias centenas de otras células nerviosas.

2.3. Excitabilidad de la membrana celular

Las neuronas están cubiertas por una membrana compuesta de complejos proteínicos y lípidos orientados que forman estructuras encargadas del transporte de solutos y electrolitos, entre el interior y el exterior de la célula nerviosa. Entre éstas están los llamados canales iónicos, que sirven para regular las corrientes iónicas transmembranales responsables de originar la formación de los trenes de espigas de voltaje, que son señales nerviosas. Los primeros estudios experimentales que se llevaron a cabo, fueron facilitados por el descubrimiento del axón gigante de una de las células que controlan la acción neuromotora del calamar. Cuando, en esta preparación clásica se coloca un par de electrodos, uno en el interior y otro en el exterior del axón, se mide un voltaje de aproximadamente $-70mV$, llamado *voltaje o potencial de membrana en reposo*. Al aplicar un pequeño y breve pulso de corriente eléctrica despolarizante, se advierte un incremento en el voltaje

a través de la membrana celular, que decae asintóticamente hasta que se alcanza otra vez el potencial de reposo. Si la amplitud del pulso de corriente aplicado es suficientemente grande, lo cual corresponde a llevar el potencial de la membrana por encima de un *umbral* de aproximadamente $-55mV$, se observa un aumento desproporcionado del voltaje, hasta que éste alcanza un valor máximo cercano a $+30mV$, para luego decrecer más allá del potencial del reposo, hasta alcanzar su valor máximo a los $-80mV$. En una última fase del proceso, lentamente, se recupera el voltaje de reposo: $-70mV$. El fenómeno anterior es llamado *impulso nervioso*, *potencial de acción* o *espiga de voltaje* y dura alrededor de dos milisegundos (ver figura 2.2).

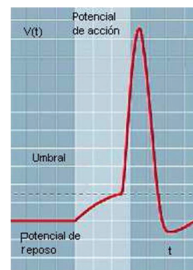


Figura 2.2 Curso temporal del voltaje eléctrico a través de la membrana celular, durante un potencial de acción.

Experimentalmente, se comprueba que la amplitud máxima del potencial de acción ($+30mV$ en la preparación clásica) es independiente de la magnitud del estímulo eléctrico aplicado, siempre y cuando éste rebese el valor del *umbral* ($-55mV$). Este hecho es conocido como la ley del “todo o nada”. Al aplicar secuencialmente breves estímulos (supraumbrales) suficientemente espaciados en el tiempo, la membrana responde cada vez produciendo idénticos potenciales de acción. En cambio, si el lapso entre estímulos se va reduciendo, se observa que, para intervalos de tiempo de entre $0,25ms$ y $0,5ms$, es imposible excitar a la membrana por segunda vez. Este lapso crítico es llamado *período refractario*.

Bajo ciertas condiciones, si en vez de aplicar un breve pulso de corriente, se le aplica a la célula una corriente constante de magnitud I , ésta responde con un tren periódico de disparos de potenciales de acción, que tiene una frecuencia creciente con el valor de la intensidad de la corriente aplicada (ver figura 2.3). Esto quiere decir que la intensidad del estímulo se codifica en términos de la frecuencia de las espigas.

Existe un amplio rango de frecuencias que se observan fisiológicamente, pero como los períodos refractarios son generalmente de menos de un

24 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

milisegundo, las frecuencias máximas resultan ser inferiores a 1000 espigas por segundo.

Es importante destacar que las neuronas cerebrales son mucho más complejas que los axones del calamar. La diversidad de canales iónicos que tienen les permite generar trenes de potenciales de acción con diferentes estructuras temporales; unas disparan en ráfagas, otras “adaptan” su disparo (esto es, cambian su frecuencia con el tiempo), otras más disparan periódicamente todo el tiempo, etc. Si se quisiera modelar esta riqueza dinámica con ecuaciones del tipo de Hodgkin y Huxley, cada neurona requería de un sistema de entre 10 y 20 ecuaciones diferenciales no lineales (para tomar en cuenta los diferentes tipos de canales iónicos involucrados). Por lo tanto, es realmente notable que los aspectos más básicos de la dinámica de las células nerviosas puedan comprenderse con el simple modelo de FitzHugh-Nagumo.

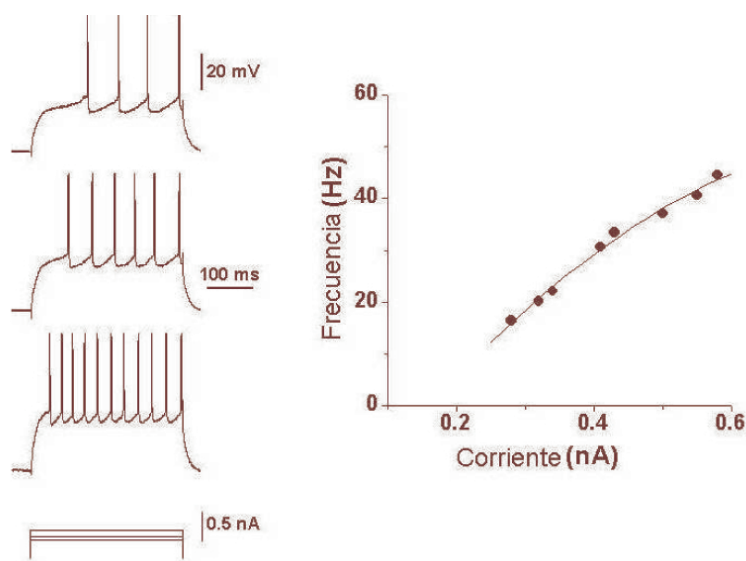


Figura 2.3 Respuesta de una célula del neuroestriado del cerebro de una rata al aplicarse una corriente sostenida durante un intervalo de tiempo de aproximadamente medio segundo.

2.4. Modelo de Fitzhugh Nagumo

Las características más sobresalientes del potencial de acción y su dinámica, fueron modeladas satisfactoriamente por Hodgkin y Huxley, con su sistema de cuatro ecuaciones diferenciales no lineales. A pesar de que todos los elementos allí involucrados, tienen un claro significado biofísico, sin embargo,

la complejidad matemática dificulta enormemente su análisis. Con el afán de comprender la esencia de la dinámica del fenómeno de excitabilidad, FitzHugh construyó un sistema más simple (dos ecuaciones diferenciales), basado en la famosa ecuación de Van Der Pol [GU,HO]. Se trata de un sistema que a pesar de ser no lineal, es susceptible de un análisis cualitativo bastante completo por ser bidimensional. El sistema se representa generalmente en la forma siguiente:

$$\frac{dv}{dt} = V(v, w) = I - v(v - a)(v - 1) - w \quad (2.4.1)$$

$$\frac{dw}{dt} = W(v, w) = b(v - gw) \quad (2.4.2)$$

siendo $I, g \geq 0, b > 0$ y $0 < a < 1$ los parámetros. No se trata de un modelo que describa detalladamente la realidad biofísica de las células nerviosas sino que constituye más bien lo que se conoce como un “*análogo*” que imita la dinámica neuronal. Para poder hacer la interpretación de la dinámica del sistema FitzHugh-Nagumo en términos biofísicos, se considera la variable de estado v como el voltaje a través de la membrana y se considera que el parámetro I representa la corriente externa aplicada a la célula nerviosa. Siguiendo a FitzHugh, conviene pensar la variable de estado w , como una variable de recuperación del sistema sin significado biofísico específico. El sistema FitzHugh-Nagumo constituye hoy en día el modelo estándar del fenómeno de excitación.

2.4.1. Estados de equilibrio y ceroclinas

Al resolver, analítica o numéricamente, el sistema de ecuaciones diferenciales que gobiernan la dinámica de la neurona, encontramos la evolución temporal de las variables de estado del sistema, $(v(t), w(t))$, a partir de alguna condición inicial (v_0, w_0) , predeterminada. El espacio donde habitan las variables de estado (v, w) , es llamado *espacio de fases o espacio de estados* y en él podemos visualizar la evolución del sistema (la neurona), graficando las órbitas de la ecuación diferencial (es decir las imágenes de las soluciones de la ecuación diferencial). Los puntos del espacio de estados de un sistema de ecuaciones diferenciales, para los cuales las derivada respecto al tiempo de las variables de estado involucradas se anulan simultáneamente, constituyen los *estados de equilibrio o estados estáticos* del sistema.

Igualando a cero las ecuaciones (2.4.1,2.4.2) del sistema de FitzHugh-Nagumo se encuentran las ecuaciones de los estados de equilibrio:

$$0 = I - v(v - a)(v - 1) - w$$

26 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

$$0 = b(v - gw)$$

Por lo tanto

$$\begin{aligned} w &= I - v(v - a)(v - 1) \\ w &= g^{-1}v \end{aligned} \tag{2.4.3}$$

Las curvas que determinan cada una de estas ecuaciones por separado en el espacio estado (v, w) , son llamadas curvas ceroclinas del sistema. Éstas son el lugar geométrico de los puntos del espacio de estados en los que el campo vectorial asociado al sistema es, respectivamente, vertical u horizontal. La primera de estas ecuaciones corresponde a los estados en los cuales no hay cambio instantáneo en v (se obtuvo al hacer $\frac{dv}{dt} = 0$), y la segunda, a los estados para los cuales no hay cambio instantáneo en w (se obtuvo al hacer $\frac{dw}{dt} = 0$).

Cada una de las ecuaciones del sistema (2.4.3) representa una curva en el espacio de estados (v, w) del sistema FitzHugh-Nagumo. A la primera ecuación le corresponde la gráfica de un polinomio cúbico y a la segunda le corresponde la gráfica de una recta que pasa por el origen. Los puntos en donde estas curvas (la cúbica y la recta) se intersecan, son los puntos de equilibrio del sistema -no se producen cambios, ni en v ni en w . Como estas curvas pueden tener hasta tres intersecciones, el número máximo de equilibrios del sistema FitzHugh-Nagumo es tres. Ejemplos de cuando se tienen uno, dos o tres puntos fijos se muestran en la figura 2.4.

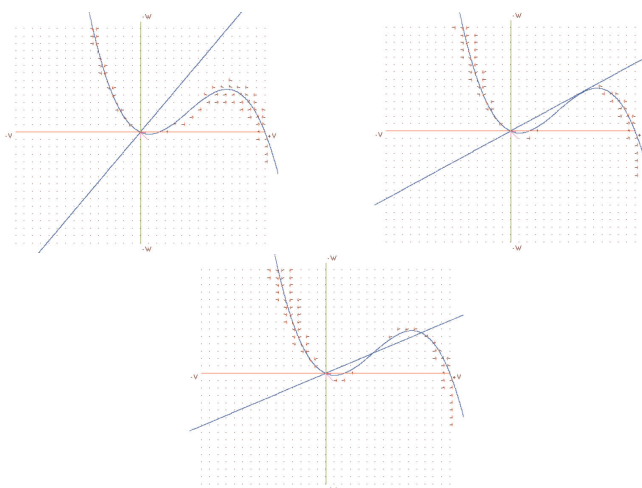


Figura 2.4.1 Se grafican las dos ceroclinas del sistema de FitzHugh-Nagumo. Dependiendo del valor de los parámetros se pueden obtener uno, dos o tres cruces. Aquí se muestran tres ejemplos: si $a = 0,15$, $b = 0,01$, $g = 2,5$, $I = 0$ se tiene un solo punto fijo; si $a = 0,15$, $b = 0,01$, $g = 5,45$, $I = 0$ se tienen dos puntos fijos; si $a = 0,15$, $b = 0,01$, $g = 7,0$, $I = 0$ se tienen tres puntos fijos.

2.4.2. Estabilidad del potencial de reposo y linealización

Para analizar la dinámica de un sistema de ecuaciones diferenciales, es necesario estudiar sus estados de equilibrio, y la estabilidad de los mismos. Dos sistemas que tengan un número distinto de estados de equilibrio, o el mismo número, pero con diferente estabilidad, tendrá un comportamiento cualitativamente diferente. Los estados de equilibrio de un sistema dinámico pueden ser inestables, estables o asintóticamente estables. Si se trata de un estado de equilibrio inestable, una pequeña perturbación puede hacer que el sistema evolucione a estados muy lejanos de él y se pierda el equilibrio; si se trata de un equilibrio estable, las soluciones vecinas no se alejarán nunca de él; y si el equilibrio es asintóticamente estable, después de cualquier pequeña perturbación, éste tenderá a restablecerse automáticamente.

Es conocido que, bajo ciertas condiciones que se puedan verificar computacionalmente, la dinámica local, alrededor de un estado de equilibrio de un sistema no lineal, puede aproximarse bastante bien, utilizando sólo la parte lineal del campo vectorial. A continuación se recordará la idea de este procedimiento y el teorema que establece la relación entre el sistema no lineal y su linealización, así como las condiciones de su validez.

Si se tiene un sistema no lineal en dos dimensiones, dado por las funciones continuamente diferenciables $X_i(x_1, x_2)$, $i = 1, 2, \dots$ en una vecindad de un punto entonces, utilizando la expansión de Taylor se tiene:

$$X_i(x_1, x_2) = X_i(\xi, \eta) + (x_1 - \xi) \frac{\partial X_i}{\partial x_1}(\xi, \eta) + (x_2 - \eta) \frac{\partial X_i}{\partial x_2}(\xi, \eta) + R_i(x_1, x_2)$$

donde $R_i(x_1, x_2)$ satisface $\lim_{r \rightarrow 0} \left[\frac{R_i(x_1, x_2)}{r} \right] = 0$ con $r = [(x_1 - \xi)^2 + (x_2 - \eta)^2]^{\frac{1}{2}}$.

Si (ξ, η) es un punto fijo del sistema de ecuaciones diferenciales $\dot{x} = X(x)$, con $x = (x_1, x_2)$ y $X = (X_1, X_2)$, entonces $X_i(\xi, \eta) = 0$. Utilizando este hecho y la expansión de Taylor se encuentra que:

$$\dot{x}_1 = (x_1 - \xi) \frac{\partial X_1}{\partial x_1}(\xi, \eta) + (x_2 - \eta) \frac{\partial X_1}{\partial x_2}(\xi, \eta) + R_1(x_1, x_2)$$

$$\dot{x}_2 = (x_2 - \eta) \frac{\partial X_2}{\partial x_1}(\xi, \eta) + (x_2 - \eta) \frac{\partial X_2}{\partial x_2}(\xi, \eta) + R_2(x_1, x_2)$$

y la parte lineal de este sistema, utilizando las nuevas coordenadas $y_1 = x_1 - \xi$ y $y_2 = x_2 - \eta$, se puede escribir como:

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial X_1}{\partial x_1} & \frac{\partial X_1}{\partial x_2} \\ \frac{\partial X_2}{\partial x_1} & \frac{\partial X_2}{\partial x_2} \end{pmatrix}_{(x_1, x_2) = (\xi, \eta)} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

28 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

A esta matriz de 2×2 se le conoce como la matriz de linealización en el punto (ξ, η) . El siguiente teorema establece la relación entre un sistema no lineal (en el plano) y el sistema lineal que se obtiene utilizando la matriz de linealización.

Teorema 2.4.1. *Sea $\dot{x} = X(x)$ un sistema no lineal, con $x = (x_1, x_2)$, tal que su matriz de linealización $A = DX((\xi, \eta))$ en el estado de equilibrio (ξ, η) sea no singular ($\det A \neq 0$). Entonces, si el sistema linealizado no es un centro (es decir los valores propios de A no son imaginarios), la dinámica del sistema y la de linealización son cualitativamente equivalentes en una vecindad de $x = (\xi, \eta)$.*

La matriz A de linealización del sistema FitzHugh-Nagumo alrededor del punto $(\xi, \eta) = (v_0, w_0)$ es:

$$\begin{bmatrix} \frac{\partial V}{\partial v} & \frac{\partial V}{\partial w} \\ \frac{\partial W}{\partial v} & \frac{\partial W}{\partial w} \end{bmatrix}_{(v,w)=(v_0,w_0)} = \begin{bmatrix} -3v_0^2 + 2(a+1)v_0 - a & -1 \\ b & -bg \end{bmatrix}$$

y el sistema linealizado alrededor del equilibrio es:

$$\begin{aligned} \dot{v} &= (-3v_0^2 + 2(a+1)v_0 - a)v - w \\ \dot{w} &= b(v - gw) \end{aligned} \tag{2.4.4}$$

La estabilidad de los puntos fijos está relacionada con los valores propios de la matriz de este sistema lineal. Cuando la parte real de todos los valores propios es negativa, el equilibrio es asintóticamente estable; cuando al menos uno es cero, es estable; cuando algún valor propio es positivo, el sistema lineal es inestable.

2.5. Análisis paramétrico

Los resultados experimentales indican que las neuronas presentan un solo estado de equilibrio correspondiente al potencial de reposo de la membrana. Para analizar esta situación en el modelo matemático, consideraremos que los parámetros del sistema FitzHugh-Nagumo tienen valores que garantizan la existencia de un único punto estático y entonces examinaremos dos casos: autónomo, cuando la estimulación externa es cero ($I = 0$) y el forzado, cuando a la célula se le aplica, externamente, una corriente continua ($I = \text{constante}$).

2.5.1. Existencia de un único estado de equilibrio

Como hemos observado anteriormente, dependiendo de los valores de sus parámetros, el sistema FitzHugh-Nagumo puede tener uno, dos o tres estados de equilibrio. Ahora queremos averiguar bajo qué condiciones paramétricas se puede asegurar la existencia de un único punto de equilibrio. Para esto observamos primero que el parámetro b no es relevante en este análisis: las coordenadas de los puntos de equilibrio del sistema no dependen de b , ya que éste no figura en las ecuaciones de la ceroclinas. Conviene hacer dos observaciones geométricas en el plano de fases del sistema FitzHugh-Nagumo. Si se mantienen fijos los parámetros a y g , la modificación de los valores del parámetro I tiene como consecuencia la traslación de la ceroclina cúbica, en la dirección del eje w ; si se mantienen fijos los parámetros I y a , el modificar el parámetro g tiene como efecto un cambio en el valor de la pendiente de la ceroclina recta. A partir de estas consideraciones, no es difícil llegar a la siguiente observación geométrica.

Fijemos el parámetro a en el intervalo $(0, 1)$ y denotemos por c a la pendiente de la ceroclina cúbica en su punto de inflexión. Si la pendiente $(\frac{1}{g})$ de la ceroclina recta es mayor o igual a c , entonces para todo valor del parámetro I , el sistema FitzHugh-Nagumo tiene un único punto de equilibrio.

Calculemos el valor de la pendiente de la ceroclina cúbica, $w = I - v(v - a)(v - 1)$, en su punto de inflexión:

$$\frac{d^2w}{dv^2} = -6v + 2(a + 1)$$

y

$$\frac{d^2w}{dv^2} = 0 \iff v = \frac{a + 1}{3}$$

Entonces en $v = \frac{a+1}{3}$ la ceroclina cúbica tiene el punto de inflexión. La pendiente que tiene la ceroclina en este valor de v se calcula fácilmente:

$$\left. \frac{dw}{dv} \right|_{v=\frac{a+1}{3}} = \frac{a^2 - a + 1}{3} = c$$

obtenemos así el siguiente resultado .

Proposición 2.5.1. *Para calcular selección de parámetros, (a, I, b, g) del sistema FitzHugh-Nagumo. Si se cumple la condición*

$$\frac{1}{g} \geq \frac{a^2 - a + 1}{3} = c \tag{2.5.1}$$

entonces hay un único equilibrio.

30 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

2.5.2. Caso autónomo

Analizamos primero el caso en el que $I = 0$, lo cual significa que a la membrana neuronal no se le aplica corriente alguna. Bajo estas condiciones, los experimentos fisiológicos revelan que el potencial de reposo se comporta como un atractor. Si el potencial de la membrana es perturbado con un pulso de corriente (ver figura 2.6), éste espontáneamente se recupera regresando a su valor inicial (potencial de reposo). Cuando $I = 0$ y se cumple la desigualdad (2.5.1), en el modelo matemático el origen (v_0, w_0) es el único estado de equilibrio del sistema FitzHugh-Nagumo debido a que las ceroclinas de este sistema se intersecan en el origen del espacio.

Averigüemos ahora si este (único) estado de equilibrio del modelo es estable, como debe corresponder a la realidad biológica observada. Para ésto aplicaremos el procedimiento de linealización y buscaremos condiciones sobre los parámetros, de manera que la parte real de los valores propios del sistema linealizado sea menor que cero. De acuerdo con la ecuación (2.4.4), el sistema linealizado alrededor del $(0, 0)$ sería en este caso

$$\begin{aligned}\dot{v} &= -av - w \\ \dot{w} &= b(v - gw)\end{aligned}$$

Los valores propios del sistema resultan entonces:

$$\lambda_{1,2} = \frac{1}{2} \left[(-a - bg) \pm \sqrt{(-a - bg)^2 - 4(abg + b)} \right]$$

ó

$$\lambda_{1,2} = - \left(\frac{a + bg}{2} \right) \pm \frac{\sqrt{(a - bg)^2 - 4b}}{2}$$

Para llevar a cabo el análisis conviene separar dos casos: cuando los valores propios son reales y cuando éstos no son reales.

Caso real. Este se produce cuando $(a - bg)^2 \geq 4b$. Para que los valores propios sean menores que cero, se tiene que cumplir que $a + bg > \sqrt{(a - bg)^2 - 4b}$, lo cual equivale a que $ga > -1$; ésto se cumple siempre, pues, desde el planteamiento de las hipótesis de modelo FitzHugh-Nagumo, hemos supuesto que $a, b > 0$ y $g \geq 0$.

Caso no real. Si $(a - bg)^2 < 4b$, la parte real de los valores propios es menor que cero si $a + bg > 0$. Esto también se cumplirá siempre, ya que $a, b > 0$ y $g \geq 0$. Así llegamos a la conclusión que expresa la siguiente proposición.

Proposición 2.5.2. *Si en el sistema de FitzHugh-Nagumo existe un solo punto fijo y no hay estimulación externa (es decir cuando $I = 0$), entonces para todo valor de los parámetros $a, b > 0$ y $g \geq 0$, el estado de equilibrio $(v_0, w_0) = (0, 0)$ es asintóticamente estable.*

En referencia a la figura (2.5), conviene notar que las constantes $a = 0,15$, $b = 0,01$ y $g = 0,25$ verifican la condición (2.5.1) y en consecuencia el sistema tiene un único estado de equilibrio. Por la proposición anterior este estado de equilibrio $(v_0, w_0) = (0, 0)$ es un atractor. allí se calculan y muestran las ceroclinas del sistema FitzHugh-Nagumo y algunas órbitas en el espacio de fases, así como los correspondientes cursos temporales del voltaje, cuando la corriente aplicada es cero. Se aprecia claramente que los cursos temporales que produce el modelo son cualitativamente iguales a los potenciales de acción que se observan en el experimento (comparar con la figura 2.2). Se muestra además, que el experimento de perturbación del potencial de reposo de la neurona, queda bien modelado por las ecuaciones del sistema FitzHugh-Nagumo (escogiendo soluciones con condiciones iniciales de la forma $(v_0, w_0) = (0, v_u \pm \varepsilon)$ siendo v_u el potencial umbral y ε una pequeña perturbación). Observamos también en esta figura que al aplicar a la neurona una serie de perturbaciones, que tengan el efecto de incrementar progresivamente el valor del voltaje de la membrana $v_1 < v_2 < v_3 < \dots$, se obtiene primero una respuesta pasiva (lineal) de la membrana, cuyo voltaje no excede a la amplitud del la perturbación inicial. Después, cuando la perturbación cruza el valor umbral, se obtiene una respuesta activa (no lineal), caracterizada por un incremento desproporcionado del voltaje $v(t)$, hasta alcanzar un valor máximo que es mucho mayor que el valor de la perturbación inicial. Finalmente, se observa también que la diferencia entre los cursos temporales de perturbaciones que superan el umbral es muy poco significativa, dando lugar a potenciales de acción que son prácticamente indistinguibles.

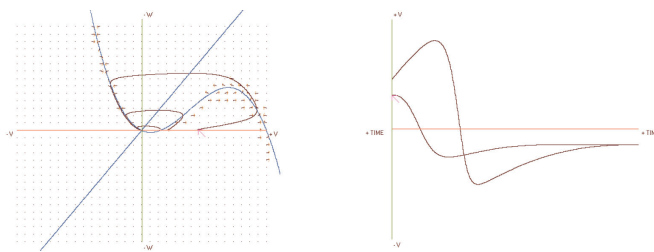


Figura 2.5.2 Fénomeno de excitabilidad. Izquierda: Órbitas el espacio de estados. Derecha: Correspondientes cursos temporales del potencial eléctrico. Los parámetros utilizados en ambos casos son $I = 0$, $a = 0,15$, $b = 0,01$ y $g = 2,5$.

32 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

2.5.3. Estimulación de la célula con una corriente externa

Investigaremos ahora la forma en que la membrana responde cuando se le aplica una corriente I constante en el tiempo (lo que en la jerga de los ingenieros electrónicos se conoce como una corriente continua; ver figura 2.6). Físicamente es de esperar que exista un valor crítico de I (la corriente aplicada) a partir de la cual la dinámica eléctrica de la membrana celular exhiba trenes periódicos de potenciales de acción, cuya frecuencia varíe o simplemente crezca al aumentar la intensidad de la corriente, lo cual se registra en las ecuaciones 2.4.1 y 2.4.2.

Veamos qué predice el modelo de FitzHugh-Nagumo en estas condiciones. Aumentar el valor de la corriente I , traslada la gráfica de la ceroquina cúbica verticalmente en el sentido positivo del eje w del plano de fases, lo cual tiene como efecto incrementar el valor de la componente v del estado de equilibrio del sistema (ver figura 2.10). Es importante estudiar la variación de la estabilidad de este equilibrio cuando ocurre tal cambio.

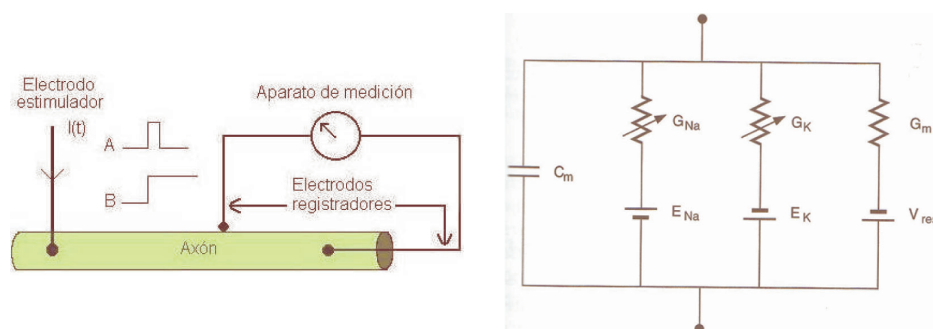


Figura 2.5.3 a) Estimulación del axón con una corriente $I(t)$. En el caso A se aplica un pulso de corriente y en el caso B una corriente continua.

Puede demostrarse que si los parámetros del sistema FitzHugh-Nagumo satisfacen la condición

$$bg \geq c \quad (2.5.2)$$

el equilibrio es asintóticamente estable para todo valor de I . En este caso, el incrementar la corriente tiene sólo el efecto de aumentar el valor de la componente v del equilibrio (es decir el valor del potencial de reposo), como lo ilustra la figura 2.7; obsérvese que la función cúbica se desplaza verticalmente con respecto a las ecuaciones 2.4.1 y 2.4.2.

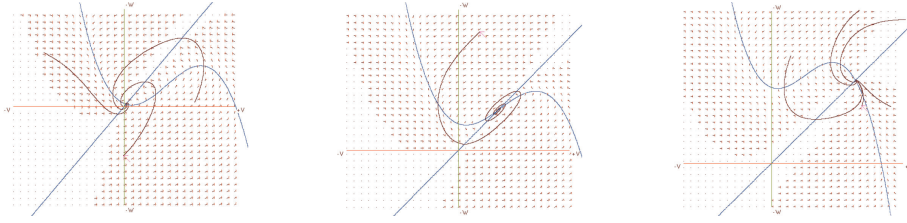


Figura 2.5.3 b) Configuración paramétrica ($a = 0,15$, $b = 0,14$ y $g = 2,5$) para la cual el punto de equilibrio resulta ser un atractor para todo valor de I ; los valores de la corriente para cada una de las gráficas son $I = 0,01$ (arriba a la izquierda), $I = 0,095$ (arriba a la derecha) e $I = 0,35$ (abajo).

Cuando se tiene una configuración paramétrica del sistema FitzHugh-Nagumo con $bg < c$ puede demostrarse la existencia de dos valores de la corriente aplicada $I_1(g, b)$ y $I_2(g, b)$ para los cuales la traza de la matriz de linealización del sistema en el punto de equilibrio se anula y tales que al traspasarlos, variando el valor de I , la traza cambia de signo. Como el signo de la traza de la matriz de linealización da el signo de la parte real de los valores propios del sistema linealizado, asociado a esta transición se produce un cambio en la estabilidad del equilibrio. Este hecho se ilustra en la figura 2.8

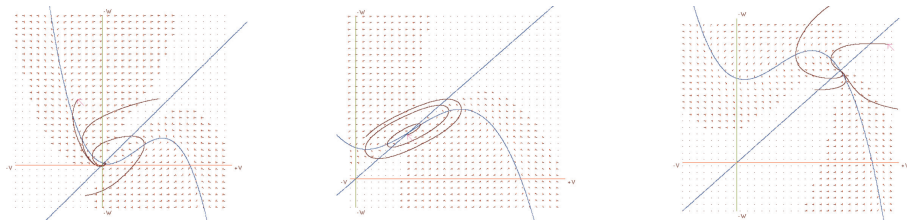


Figura 2.5.3 c) Configuración paramétrica ($a = 0,15$, $b = 0,08$ y $g = 2,5$) para la cual el estado de equilibrio pierde la estabilidad al incrementar la intensidad de la corriente aplicada. Obsérvese que un incremento adicional de la corriente aplicada puede volver a estabilizar el voltaje de reposo. Los valores de la corriente, para cada una de las gráficas son $I = 0,01$ (arriba a la izquierda), $I = 0,095$ (arriba a la derecha) e $I = 0,35$ (abajo).

La situación anterior configura lo que se conoce como escenario de la *bifurcación Andronov-Hopf*. Obsérvese que, asociado al proceso de inestabilización del estado de equilibrio, aparece un ciclo límite estable (es decir una órbita periódica atractora). En la figura 2.9 se muestra una secuencia de ciclos límite en la que se van disminuyendo los valores del parámetro b . El límite $b \rightarrow 0$ es lo que se conoce en lenguaje matemático como el límite singular del sistema de ecuaciones diferenciales. Es cuando $b \ll 1$ que el modelo FitzHugh-Nagumo reproduce cualitativamente la dinámica del

34 Modelo de Fitzhugh-Nagumo para el potencial de una membrana

potencial de acción de una neurona. En este caso se obtienen trayectorias en el espacio de estados casi horizontales, debido a que la variable w empieza a variar muy lentamente respecto a la variable v (los valores propios de \dot{w} se hacen mucho más pequeños que los de \dot{v}). Al respecto, ver 2.4.1; nótese que cerca de la ceroclina cúbica los valores de \dot{v} y \dot{w} pueden tener el mismo orden de magnitud. En la figura 2.10 se muestra cómo al aumentar los valores de la corriente I se produce la transición, del régimen excitable al régimen oscilatorio del sistema FitzHugh-Nagumo. En la figura 2.10 se muestra como, de acuerdo a lo observado experimentalmente (comparar con la figura 1.3), la frecuencia de disparo de los trenes de potenciales de acción aumenta con la intensidad de la corriente.

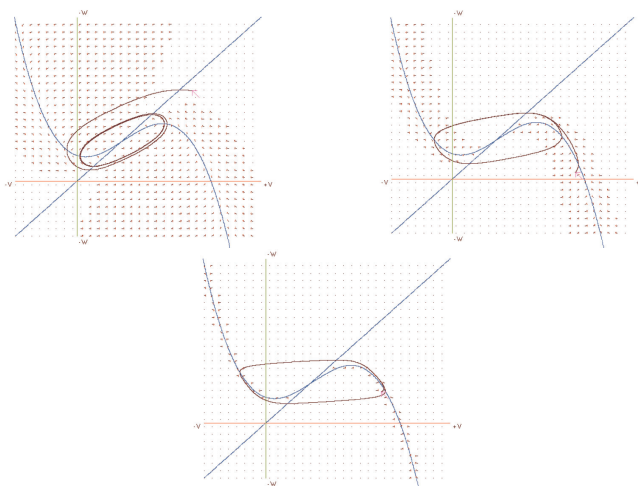


Figura 2.5.3 d) Ciclos límite estables asociados a los estados de equilibrio inestables. En todos los casos se utilizó $a = 0,15$, $g = 2,5$ e $I = 0,095$; los valores de b son $b = 0,08$ (arriba a la izquierda), $b = 0,03$ (arriba a la derecha) y $b = 0,01$ (abajo).

2.5.4. Conclusión del Capítulo 2

La dinámica es una característica fundamental de la fisiología de una neurona y del sistema nervioso en general. Esta dinámica es no lineal y su complejidad es tal que no puede ser comprendida sin el uso de una teoría matemática basada en ecuaciones diferenciales. El trabajo pionero realizado por Hodgkin y Huxley, a mediados del siglo pasado, constituyó una contribución fundamental en esta dirección. Dada la gran complejidad de los sistemas estudiados, es también muy importante construir modelos simplificados que permitan aislar y comprender la esencia de la dinámica de los fenómenos involucrados. El modelo de FitzHugh-Nagumo provee

un escenario de complejidad mínima para entender el fenómeno de la excitabilidad en un contexto geométrico. Por esta razón, la aparición de este modelo marca una nueva etapa en la historia de la neurofisiología, en la que el análisis geométrico de órbitas en el espacio de fases se vuelve fundamental para lograr una comprensión visual de la mecánica del impulso nervioso.

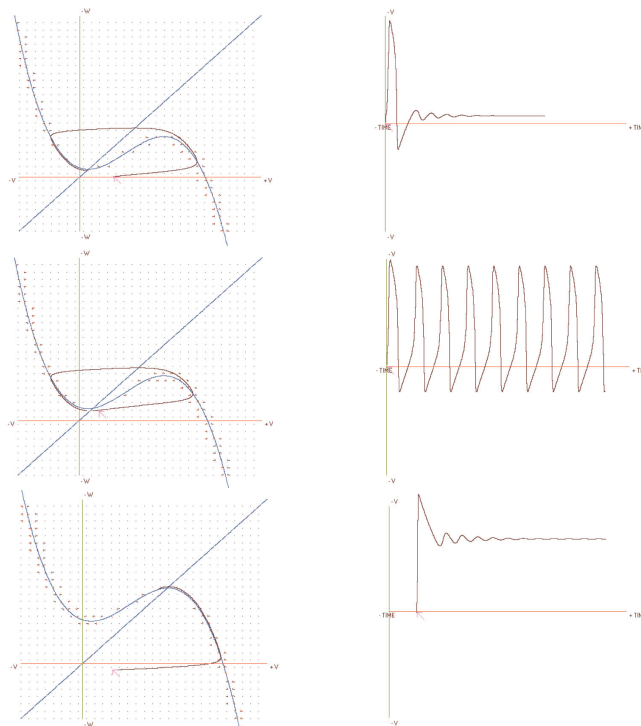


Figura 2.5.4 Transición del régimen excitable al régimen oscilatorio del potencial eléctrico de una neurona al aumentar la intensidad de la corriente. Obsérvese cómo el voltaje del estado de equilibrio aumenta con la corriente. En todas las gráficas se utilizó $a = 0,15$, $b = 0,01$ y $g = 2,5$; de arriba a abajo, los valores del parámetro I , tanto en el retrato de fase (izquierda) como en el curso temporal (derecha), son respectivamente, 0,035; 0,05 y 0,16.

CAPÍTULO 3

Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

3.1. Definiciones y conceptos básicos

Un modelo matemático bien conocido que describe la propagación de pulsos o voltajes a lo largo de la membrana de las células es el formado por las ecuaciones

$$V_{xx} - V_t = F(v) + R - I \quad (3.1.1)$$

$$R_t = c(V + a - bR) \quad (3.1.2)$$

el cual fue sugerido por FitzHugh-Nagumo (FHN). Estas son ecuaciones diferenciales parciales para el pulso de voltaje $V(x, t)$ a lo largo de la fibra nerviosa y para la variable recuperadora $R(x, t)$ donde a, b son constantes positivas, c corresponde al factor de temperatura e I es la corriente externa inyectada. Las ecuaciones (3.1.1) y (3.1.2) tiene aplicaciones en electrofisiología y también en la genética de poblaciones. Cuando la segunda variación espacial de $V(x, t)$, a saber V_{xx} , es insignificante entonces las ecuaciones (3.1.1) y (3.1.2) se reducen a la ecuación del oscilador de BONHOEFER-VAN DER POL (BVP).

Algunas investigaciones sobre las dinámicas caóticas en presencia de corrientes externas constantes y periódicas han sido hechas por RA-JASEKAR LAKSHMANAN [MU,MA]. Estos sistemas bidimensionales describen la cinética local. Además de lo dicho anteriormente, existen varios estudios en la literatura sobre la bifurcación y el caos de las diferentes

38 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

formas del modelo neuronal. Por otra parte, al buscar soluciones en forma de ondas viajeras

$$V = V(x - ut) = V(\xi) \quad (3.1.3)$$

$$R = R(x - ut) = R(\xi) \quad (3.1.4)$$

donde u es la velocidad de la **onda viajera**, las ecuaciones (3.1.1) y (3.1.2) se reducen al siguiente sistema de ecuaciones diferenciales ordinarias.

$$\ddot{V} = F(v) + R - uv - I \quad (3.1.5)$$

$$\dot{R} = \frac{c}{u}(v + a - bR) \quad (3.1.6)$$

el cual puede reformularse como:

$$\dot{v} = U$$

$$\dot{U} = f(v) + R - uU - I$$

$$\dot{R} = \frac{c}{u}(v + a - bR)$$

Aquí la diferenciación es con respecto a ξ , e I es una corriente externamente inyectada, la cual podemos suponer constante. En nuestro análisis estamos suponiendo que tenemos un axón de longitud infinita. Es suficiente para tratar la no linealidad en $F(v)$ en forma adecuada, en particular se escogerá la forma cúbica,

$$F(V) = -V + \frac{V^3}{3} \quad (3.1.7)$$

(Comparar con la correspondiente en 2.4.1.) En los últimos años algunos estudios se han llevado a cabo usando el método de perturbación en ausencia de corriente externa. El comportamiento caótico ha sido estudiado en modelos neuronales similares. Por ejemplo Kahlert y Rössler, estudiaron el comportamiento caótico de manera diferente del modelo de FHN en el modelo de Rinzel-Keller. Muy recientemente Rajasekar y Lakshmanan han investigado numéricamente y señalando la existencia del movimiento límite de carácter cíclico que bifurca en una sucesión de duplicidad de período para este movimiento caótico y la eliminación del caos usando mecanismos diferentes, con y sin, la inyección de corriente constante y periódica para ciertos parámetros escogidos específicamente.

Queremos señalar que existen numerosos estudios sobre la bifurcación de Hopf en las ecuaciones diferenciales parciales (3.1.1) y (3.1.2). Sin embargo en este artículo no consideraremos todo el sistema de ecuaciones parciales sino sólo las ecuaciones diferenciales ordinarias (3.1.3) y (3.1.4).

Nosotros aplicaremos la teoría de bifurcaciones basada en la reducción de la variedad central y el análisis de la forma normal para que una bifurcación de Hopf ocurra en las ecuaciones (3.1.3) y (3.1.4), las cuales modelan la conducción nerviosa con y sin la inyección de corrientes externas constantes ($I = A_0$ e $I = 0$, respectivamente). Notamos primero que las ecuaciones (3.1.3) y (3.1.4) poseen dos puntos de equilibrio correspondientes $(V, \dot{V}, R) = (\alpha, 0, \beta)$, donde α y β satisfacen las ecuaciones

$$\alpha^3 + 3\alpha \left(\frac{a}{b} - 1 \right) + 3 \left(\frac{a}{b} - A_0 \right) = 0 \quad (3.1.8)$$

y

$$\beta = \frac{\alpha + a}{b} \quad (3.1.9)$$

Ahora, por conveniencia, haremos la transformación lineal del sistema de ecuaciones ordinarias

$$V = x + \alpha, \quad \dot{V} = y, \quad R = z + \beta \quad (3.1.10)$$

obteniendo

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ \alpha^2 - 1 & -u & 1 \\ -\frac{c}{u} & 0 & \frac{bc}{u} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha x^2 + \frac{x^3}{3} \\ 0 \end{pmatrix} \quad (3.1.11)$$

Es fácil ver que la transformación (3.1.10) solamente cambia el punto de equilibrio de $(\alpha, 0, \beta)$ a el origen $(0,0,0)$ [PE]. Nuestro primer propósito es encontrar la forma normal para que la bifurcación de Hopf ocurra y analizar la naturaleza de ésta para varios parámetros elegidos, al igual que entender la dinámica caótica de (3.1.11). Para ello en la sección 2.2 llevaremos a cabo el análisis de (3.1.11) alrededor de puntos fijos. Consideraremos un parámetro específico, a saber $\frac{a}{b} = A_0$, para el cuál el sistema posee simetría por reflexión y veremos que varias bifurcaciones de interés ocurrirán en este sistema simétrico.

En la sección 2.3 hemos clasificado el espacio vectorial lineal asociado con el sistema linealizado de (3.1.11) y su variedad central, esto es por, una reducción de dimensionalidad local la cual es invariante.

Llevaremos a cabo el análisis de la forma normal para poner el sistema reducido en una forma más simple, por medio de una transformación de coordenadas no lineales próxima a la identidad. En las secciones 2.4 y 2.5 verificaremos la existencia de ciclos de límite estable integrando el sistema.

3.1.1. Análisis de la estabilidad lineal

A continuación consideraremos, sin perder generalidad el sistema (3.1.11) correspondiente a la ecuación FHN, el cual tiene tres puntos de equilibrio, obtenidos haciendo $\dot{x} = \dot{y} = \dot{z} = 0$.

$S^0 = (0, 0, 0)$; y

$$S^\pm = \left(\frac{-3\alpha^2 \pm \sqrt{12 - \frac{12}{b} - 3\alpha^2}}{2}, 0, \frac{-3\alpha^2 \pm \sqrt{12 - \frac{12}{b} - 3\alpha^2}}{2b} \right) \quad (3.1.12)$$

Los puntos de equilibrio S^\pm existen únicamente si $b > 1$ y $\alpha < \sqrt{1 - \frac{1}{b}}$. Sin embargo el parámetro b siempre será tomado como menor o igual que 1, debido a una interpretación fisiológica. Esto significa que el punto de equilibrio S^0 es el que aparece en el caso fisiológico.

En lo sucesivo consideraremos este caso. Para analizar la estabilidad de S^0 , linealizaremos la ecuación (3.1.11) alrededor de S^0 . Su matriz jacobiana es

$$J(S^0) = \begin{pmatrix} 0 & 1 & 0 \\ \alpha^2 - 1 & -u & 1 \\ -\frac{c}{u} & 0 & \frac{bc}{u} \end{pmatrix} \quad (3.1.13)$$

La estabilidad es determinada por los valores propios de $J(S^0)$ los cuales son las raíces de su ecuación característica. Esta es

$$P(\lambda) = \lambda^3 + p_1\lambda^2 + p_2\lambda + p_3 = 0 \quad (3.1.14)$$

donde

$$\begin{aligned} p_1 &= u - \frac{bc}{u} \\ p_2 &= 1 - \alpha^2 - bc \\ p_3 &= \frac{c}{u}[(\alpha^2 - 1)b + 1] \end{aligned}$$

El punto de equilibrio S^0 es asintóticamente estable si y sólo si todos los valores propios tienen parte real negativa. La condición necesaria y suficiente para que los valores propios tengan parte real negativa es:

$$p_1 > 0, \quad p_3 > 0 \quad y \quad p_1p_2 - p_3 > 0.$$

Así para un valor dado de (a, b, c, μ) ($a > 0, b > 0, c > 0, \mu > 0$) y de A_0 , las condiciones se convierten en

$$u > bc \quad (3.1.15)$$

$$(1 - \alpha^2)b + 1 > 0 \quad (3.1.16)$$

$$u(1 - \alpha^2 - bc) + (bc - c) > 0 \quad (3.1.17)$$

La estabilidad se puede perder cuando un par de valores propios complejo cruza el eje imaginario, siendo aún el tercero negativo. En la bifurcación de Hopf $J(S^0)$ se tienen valores propios $\pm iw_0$ y $\lambda_t \neq 0$. Esta situación corresponde a $p_1 p_2 = p_3 p_2 = w_0^2 > 0$, en la ecuación (3.1.14), es decir las condiciones son

$$u^2(1 - \alpha^2 - bc) + b^2 c^2 - c = 0 \quad (3.1.18)$$

$$w_0^2 = 1 - \alpha^2 - bc \quad (3.1.19)$$

$$\lambda_1 = -u + \frac{bc}{u} < 0 \quad (3.1.20)$$

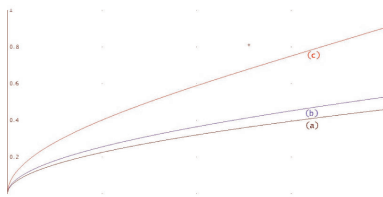
De (3.1.18) para valores fijos de los parámetros a, b, c y A_0 , la condición umbral para que la bifurcación ocurra es

$$u_H = \sqrt{\frac{c - b^2 c^2}{1 - \alpha^2 - bc}} \quad (3.1.21)$$

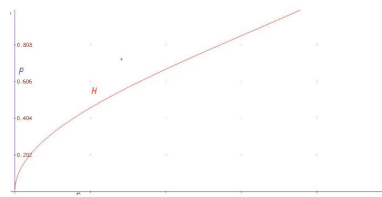
la figura 3.1(a), muestra la curva de la bifurcación de Hopf en el plano $c - u$, para varios valores de A_0 y b fijo [ver ecuación 3.1.8].

3.2. Simetrías del sistema

Es interesante considerar las **propiedades de estabilidad**, para el parámetro $\frac{a}{b} = A_0$ en la ecuación (3.1.11), la cual admite ciertas **propiedades de simetría**. El valor específico elegido naturalmente corresponde a la situación en la cual el valor de la razón de dos parámetros a y b representan varios parámetros asociados con las variables de activación e inactivación de la ecuación original Hodgkin-Huxley, de la cual la ecuación FHN ha sido derivada y balanceada por la inyección externa de la corriente A_0 .

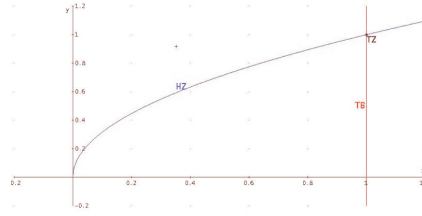


(a)



(b)

42 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3



(c)

FIGURA 3.1. Las curvas para la bifurcación de Hopf en el plano $c - u$ para varios valores de A_0

Para este parámetro elegido a se convierte en cero, el cual puede ser reducido de la ecuación (3.1.11, 3.1.8), la ecuación **FHN** se convierte en

$$\dot{x} = y \quad (3.2.1)$$

$$\dot{y} = -x + \frac{x^3}{3} - uy + z \quad (3.2.2)$$

$$\dot{z} = -\frac{c}{u}(x - bz) \quad (3.2.3)$$

La ecuación (3.1.11, 3.1.13) es ahora invariante bajo la acción del grupo simétrico \mathcal{Z}_2

$$(x, y, z) \xrightarrow{\mathcal{Z}_2} (-x, -y, -z)$$

Entonces, el polinomio característico correspondiente a S^0 para los valores propios del sistema linealizado es

$$\lambda^3 + p_1\lambda^2 + p_2\lambda + p_3 = 0 \quad (3.2.4)$$

donde ahora

$$p_1 = u - \frac{bc}{u}, \quad p_2 = 1 - bc \quad y \quad p_3 = \left(\frac{c}{u}\right)(1 - b).$$

De (3.1.11, 3.1.13) resultan dos casos diferentes, a saber:

a) $0 < b < 1$

b) $b = 1$

Cada uno de estos casos conducen a varias bifurcaciones correspondientes a estados estables y soluciones periódicas.

Caso (a): $0 < b < 1$

i) **Bifurcación Orquilla:**

Cuando $p_1 \neq 0$, $p_2 \neq 0$ y $p_3 = 0$, el sistema (3.1.11, 3.1.13) tiene un valor propio simple cero mientras que los otros dos valores propios tienen parte real distinta de cero. Esto conduce a una bifurcación orquilla, la cual ocurre sobre la línea $c = 0$, en el plano $c - u$. Esto se muestra en la figura 3,1b) por la línea p para $b = 0,5$.

ii) **Bifurcación de Hopf:**

Cuando $p_1 p_2 = p_3$ y $p_2 = w_s^2 \neq 0$, en la ecuación 3.2.1; este sistema tiene un par de valores propios imaginarios puros ($\lambda_1 = \bar{\lambda}_2 = iw_s$) y un valor propio real $\lambda_3 = p_1 = \frac{bc}{u-1}$. Este corresponde a una bifurcación de Hopf que se ilustra en el plano $c - u$, exactamente sobre la curva $u_H = \sqrt{\frac{c-b^2c^2}{1-\alpha^2-bc}}$, ver figura (1b) cuando H para $b = 0,5$.

Caso (b): $b = 1$

i) **Bifurcación Takens-Bogdanov:**

Cuando $p_2 = p_3 = 0$ y $p_1 \neq 0$ el sistema (3.1.11, 3.1.13) tiene dos valores propios cero y el tercero distinto de cero ($\lambda_1 = \lambda_2 = 0; \lambda_3 \neq 0$). Para esta elección obviamente $1 - bc = 0$, $1 - b = 0$ y $\lambda_3 = \frac{bc}{u-1}$. Correspondientemente tenemos $c = 1,0$. Así para $u \neq 0$ el sistema (3.1.11, 3.1.13) tiene dos bifurcaciones Taken-Bogdanov sobre la línea $c = 1,0$. Esto es representado en la figura (1c) y la línea TB.

ii) **Bifurcación de Hopf-Cero:**

Cuando $p_1 = p_2 = 0$ y $p_3 = w^2 > 0$ el sistema (3.1.11, 3.1.13) tiene un par de valores propios imaginarios puros y un valor propio cero ($\lambda_2 = \bar{\lambda}_1 = iw_s$). Esta situación corresponde a una bifurcación Hopf-Cero o bifurcación orquilla a lo largo de la curva $u_H = \sqrt{c}$. En la figura (1c) se muestra por la curva H_z .

iii) **Triple valor propio Cero:**

Cuando $p_1 = p_2 = p_3 = 0$ el sistema (3.1.11,3.1.13) tiene un valor propio triple. Esta situación se muestra en la figura (1c). Sin embargo para $\frac{a}{c} \neq A_0$, $\alpha \neq 0$, el término cuadrático ax^2 , aparece en la ecuación (3.1.11) y la simetría Z_2 se pierde. Debido a esto uno no puede analizar sencillamente las diferentes clases de bifurcación como tal.

3.3. Espacio vectorial tangente generado por los vectores propios

Ahora nos permitiremos considerar los valores propios $(\lambda_1, \lambda_2, \lambda_3)$ de la matrix Jacobiana $J(S^0)$ (ver ecuación (3.1.13)) y (3.1.11). De los valores propios de $J(S^0)$ uno puede fácilmente encontrar los vectores propios, los cuales forman un espacio vectorial lineal. Es más conveniente transformar el sistema original (3.1.11) a, una forma estándar por medio de una transformación lineal, la cual transforma el jacobiano en bloques de forma diagonal. Los vectores correspondientes a los valores propios λ_j , ($j = 1, 2, 3...$) están dados por:

$$v_j = \begin{pmatrix} \frac{bc}{u} - \lambda_j \\ \lambda_j \left(\frac{bc}{u} - \lambda_j \right) \\ \frac{c}{u} \end{pmatrix} \quad (3.3.1)$$

Para el valor propio real λ_1 , el vector propio es real y para el par conjugado $\lambda_2 = \bar{\lambda}_3 = \gamma + iw$ tenemos

$$v_{23} = v_r \pm iv_i \quad (3.3.2)$$

donde:

$$v_j = \begin{pmatrix} \frac{bc}{u} - \gamma \\ w^2 - \gamma^2 + \frac{bc}{u}\gamma \\ -\frac{c}{u} \end{pmatrix}, \quad v_i = \begin{pmatrix} -w \\ w \left(\frac{bc}{u} - 2\gamma \right) \\ 0 \end{pmatrix} \quad (3.3.3)$$

La transformación lineal que lleva el jacobiano del problema lineal a un bloque de Jordan es:

$$v_j = S u_j \quad , j = 1, 2, 3, \dots \quad (3.3.4)$$

así que : $SJ = JS$

$$J(S^0) = \begin{pmatrix} y & w & 0 \\ -w & y & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} = S^{-1} J(S^0) S \quad (3.3.5)$$

donde

$$S = \begin{pmatrix} \frac{bc}{u} - \gamma & -w & \frac{bc}{u} - \lambda_1 \\ w^2 - \gamma^2 + \frac{bc}{u}\gamma & w \left(\frac{bc}{u} - 2\gamma \right) & \lambda_1 \left(\frac{bc}{u} - \lambda_1 \right) \\ \frac{c}{u} & 0 & \frac{c}{u} \end{pmatrix} \quad (3.3.6)$$

y

$$S^{-1} = \frac{1}{|s|} \begin{pmatrix} \frac{cw}{u} \left(\frac{bc}{u} - 2\gamma \right) & \frac{cw}{u} & w \left(\lambda_1^2 + \frac{b^2c^2}{u^2} + 2\gamma \left(\frac{bc}{u} - \lambda_1 \right) \right) \\ \frac{c}{u} \left(\gamma^2 - w^2 - \lambda_1^2 + \frac{bc}{u} (\lambda_1 - \gamma) \right) & \frac{c}{u} (\lambda_1 - \gamma) & \left(\frac{bc}{u} - \lambda_1 \right) \left(w^2 - \gamma^2 + \gamma \lambda_1 - \frac{bc}{u} (\lambda_1 - \gamma) \right) \\ 0 & 0 & \lambda_1 \end{pmatrix} \quad (3.3.7)$$

$$|S| = \frac{cw}{u} (\lambda_1^2 + \gamma^2 + 2w^2 - 2\gamma \lambda_1) \quad (3.3.8)$$

Considerando toda la ecuación no lineal (3.1.11), la transformación lineal de las variables $(x, y, z) \rightarrow (x', y', z') = S^{-1}(x, y, z)$ lleva el sistema (3.1.11) a la forma normal

$$\begin{pmatrix} \dot{x}' \\ \dot{y}' \\ \dot{z}' \end{pmatrix} = \begin{pmatrix} \gamma & w & 0 \\ -w & \gamma & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} R_1(x', y', z') \\ R_2(x', y', z') \\ R_3(x', y', z') \end{pmatrix} \quad (3.3.9)$$

donde

$$\begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} = S^{-1} \begin{pmatrix} 0 \\ ax^2 + \frac{x^3}{3} \\ 0 \end{pmatrix} \quad (3.3.10)$$

con (x, y, z) expresadas en términos de (x', y', z') . La forma de $R_j(x', y', z')$, $j = 1, 2, 3$ es

$$R_j(x', y', z') = R_j^{(2)}(x', y', z') + R_j^{(3)}(x', y', z') \quad (3.3.11)$$

donde

$$R_j^{(2)} = R_{j1}^{(2)} x'^2 + R_{j2}^{(2)} y'^2 + R_{j3}^{(2)} x' y' + R_{j4}^{(2)} x' z' + R_{j5}^{(2)} y' z' + R_{j6}^{(2)} z'^2 \quad (3.3.12)$$

y

$$R_j^{(3)} = R_{j1}^{(3)} x'^3 + R_{j2}^{(3)} y'^3 + R_{j3}^{(3)} x' y'^2 + R_{j4}^{(3)} x'^2 y' + \dots + R_{j10}^{(3)} z'^3 \quad (3.3.13)$$

Los coeficientes $R_{jk}^{(2)}$ y $R_{jk}^{(3)}$ pueden ser fácilmente calculados y ellos están dados en el apéndice A. En el límite $\gamma = 0$, fácilmente observamos de (3.3.9) que la parte linealizada corresponde a una suma directa de un subespacio central y un subespacio estable. Considerando el efecto de todo lo no lineal de la ecuación (3.3.9) uno puede encontrar una variedad central, la cual coincide con el subespacio central del sistema linealizado en el origen (punto de equilibrio) y en este caso es posible restringir la dinámica de la ecuación (3.3.9) localmente a tal variedad central.

3.4. Reducción de la variedad central y forma normal

Del atractor local y las propiedades de invarianza de la variedad central es razonable representar la coordenada z' como una función de x' y y' para (x', y') suficientemente pequeños. Cerca al origen representamos la variedad central por una función $h : \mathfrak{R}^2 \longrightarrow \mathfrak{R}^2$, describiendo la coordenada z' de la variedad central, esto es, $z' = h(x', y')$. Así se puede mostrar satisfactoriamente la relación como: $z' = \frac{\partial h}{\partial t} = \left(\frac{\partial h}{\partial x'} \right) \dot{x}' + \left(\frac{\partial h}{\partial y'} \right) \dot{y}'$.

$\frac{\partial h}{\partial x'} = [yx' + wy' + R_1(x', y', z')] + \frac{\partial h}{\partial y'}[-wx' + yy' + R_2(x', y', z')]$ esto es igual a

$$\lambda_1 h(x', y') + R_3(x', y', z') \quad (3.4.1)$$

Con

$$h(0, 0) = \frac{\partial h}{\partial x'}(0, 0) = \frac{\partial h}{\partial y'}(0, 0) = 0 \quad (3.4.2)$$

una solución asintótica para $h(x', y')$ cerca a $(x', y') = 0$ tiene la forma $h(x', y') = h_1 x'^2 + h_2 y'^2 + h_3 x' y' + \dots$ y el tercero y el término de orden mayor, el coeficiente, h_1 , h_2 y h_3 pueden ser evaluados usando la ecuación (3.4.1). Una evaluación sencilla produce

$$h_3 = \frac{2w(R_{32}^{(2)} - R_{31}^{(2)}) + (2y - \lambda_1)R_{33}^{(2)}}{(2w)^2 + (2y - \lambda_1)^2} \quad (3.4.3)$$

$$h_1 = \frac{wh_3 + R_{31}^{(2)}}{2y - \lambda_1} \quad (3.4.4)$$

$$h_2 = \frac{-wh_3 + R_{32}^{(2)}}{2y - \lambda_1} \quad (3.4.5)$$

esta aproximación reduce el sistema (3.3.9) a un modelo de dos dimensiones en la variedad central, y por su invarianza tenemos

$$\begin{aligned} \dot{x}' = & yx' + wy' + R_{11}^{(2)} x'^2 + R_{12}^{(2)} y'^2 + R_{13}^{(2)} (x' y' + h_1 R_{14}^{(2)}) + R_{15}^{(3)} x'^3 + (h_2 R_{15}^{(2)} + R_{12}^{(2)}) y'^3 \\ & + (h_2 R_{14}^{(2)} + h_3 R_{15}^{(2)} + R_{12}^{(2)}) x' y'^2 + (h_1 R_{15}^{(2)} + h_3 R_{14}^{(2)} + R_{14}^{(2)}) x'^2 y' \end{aligned} \quad (3.4.6)$$

$$\begin{aligned} \dot{y}' = & -wx' + yy' + R_{21}^{(2)} x'^2 + R_{22}^{(2)} y'^2 + R_{23}^{(2)} x' y' + (h_1 R_{24}^{(2)} + R_{21}^{(3)} x'^3) + (h_2 R_{25}^{(2)} + R_{22}^{(3)}) y'^3 \\ & + (h_2 R_{24}^{(2)} + h_3 R_{25}^{(2)} + R_{23}^{(3)}) x' y'^2 + (h_1 R_{25}^{(2)} + h_3 R_{24}^{(2)} + R_{24}^{(3)}) x'^2 y' \end{aligned} \quad (3.4.7)$$

Donde los términos de grado cuarto y más altos han sido ignorados. Ahora, uno puede siempre especificar el flujo en la variedad central en la forma

$$(\dot{x}', \dot{y}') = V(x', y') = V^{(1)}(x', y') + V^{(2)}(x', y') + \dots \quad (3.4.8)$$

donde $V^{(i)}, i = 1, 2, 3, \dots$ que corresponden al grado de los términos x', y' . Para la identificación de la bifurcación de Hopf es necesario poner el sistema reducido (3.4.8) a una forma sencilla, la así llamada **forma normal**. Esto puede ser hecho por medio de cambios de las coordenadas no lineales cerca a la identidad, considerando la transformación de coordenadas transpuesta siguiente

$$(x, y)^T = (x', y')^T + \phi^{(k)}(x', y') \quad (3.4.9)$$

con el inverso :

$$(x', y')^T = (x, y)^T - \phi^{(k)}(x, y) + 0^{(2k-1)} \quad (3.4.10)$$

donde $\phi^{(k)}(x, y)$ es un polinomio homogéneo de grado k . Después de emplear esta transformación y después de algunas manipulaciones uno puede obtener la forma normal (en coordenadas polares). En términos de la representación $z = x + iy = re^{i\theta}$, $\bar{z} = x - iy = re^{-i\theta}$ obtenemos

$$\dot{r} = yr + Re(\alpha_1)r^3 \quad (3.4.11)$$

$$\dot{\theta} = w - Im(\alpha_1)r^2 \quad (3.4.12)$$

(Omitiendo los términos de orden cuatro y mayores) se obtiene:

$$\begin{aligned} \alpha_1 = & R_+^{(3,2)} + R_+^{(2,1)}(\phi_+^{(2,3)} - \phi_-^{(2,1)}) + \phi_+^{(2,3)}(R_-^{(2,1)} - R_+^{(2,2)}) + 2(R_-^{(2,2)}\phi_+^{(2,0)} \\ & - R_-^{(2,0)}\phi_-^{(2,2)}) - \gamma(\phi_+^{(2,1)}\phi_-^{(2,1)} + 2\phi_+^{(2,0)}\phi_-^{(2,2)} + 3\phi_+^{(2,1)}\phi_+^{(2,2)}) \\ & + iw(\phi_+^{(2,1)}\phi_+^{(2,2)} - \phi_+^{(2,1)}\phi_-^{(2,1)} - 6\phi_+^{(2,0)}\phi_-^{(2,2)}) \end{aligned} \quad (3.4.13)$$

con:

$$\begin{aligned} R_+^{(3,2)} = & \frac{1}{8}[3(h_1R_{14}^{(2)} + R_{11}^{(3)} + h_2R_{25}^{(2)} + R_{22}^{(3)}) + h_2R_{14}^{(2)} + h_3R_{15}^{(2)} + R_{25}^3 + \\ & h_1R_{25}^{(2)} + h_3R_{24}^{(2)} + R_{24}^{(3)}] + \frac{1}{8}[3(h_1R_{24}^{(2)} + R_{21}^{(3)} - h_2R_{15}^{(2)} - R_{12}^{(3)}) + \\ & h_2R_{24}^{(2)} + h_3R_{25}^{(2)} + R_{23}^{(3)} - h_3R_{14}^{(2)} - R_{14}^{(3)}] \end{aligned} \quad (3.4.14)$$

$$R_+^{(2,0)} = \frac{1}{4}(R_{11}^{(2)} - R_{12}^{(2)} - R_{23}^{(2)}) + \frac{i}{4}(R_{13}^{(2)} + R_{21}^{(2)} - R_{22}^{(2)})$$

$$R_+^{(2,0)} = \bar{R}_-^{(2,2)} \quad (3.4.15)$$

48 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

Figura 2. Forma normal numéricamente computalizada, coeficientes $a_1 = Re(\alpha_1)$ y críticamente ($\gamma = 0$) en la ecuación (3.4.11) el argumento contra u y A_0 .

$$R_+^{(2,1)} = \frac{1}{2}(R_{11}^{(2)} + R_{12}^{(2)}) + \frac{i}{2}(R_{21}^{(2)} + R_{22}^{(2)})$$

$$R_+^{(2,1)} = \bar{R}_-^{(2,1)} \quad (3.4.16)$$

$$R_-^{(2,0)} = \frac{1}{4}(R_{11}^{(2)} - R_{12}^{(2)} + R_{23}^{(2)}) + \frac{i}{4}(R_{21}^{(2)} - R_{13}^{(2)} - R_{22}^{(2)})$$

$$R_-^{(2,0)} = \bar{R}_+^{(2,2)} \quad (3.4.17)$$

y

$$\phi_{\pm}^{(2,l)} = \frac{R_{\pm}^{(2,l)}}{\lambda_{\pm}^{2,l}} \quad l = 0, 1, 2, 3, \dots \quad (3.4.18)$$

donde

$$\lambda_{\pm}^{(2,l)} = -\gamma + iw(3 \pm 2l) \quad (3.4.19)$$

la naturaleza del coeficiente en el crítico ($\gamma = 0$) como una función del parámetro de bifurcación (A_0 o u) puede ser fácilmente examinado numericamente. Como por ejemplo primero consideremos el caso de fuerza libre con ($A_0 = 0$) con otros parámetros fijos en $b = 0,5$, $c = 0,1$ y u variando hacia abajo de 1.25. Podemos encontrar una bifurcación de Hopf super crítica en $u = 1,09$. $Re(\alpha_1)$ toma un valor negativo en $u = 0,9$. Similarmente fijando $u = 0,9$ y los otros parámetros como $b = 0,5$ y $c = 0,1$ encontramos que ocurre una bifurcación de Hopf supercrítica en $A_0 = 2,163$ y $A_0 = -0,163$; donde $Re(\alpha_1)$ toma valores negativos.

3.5. Análisis numérico: bifurcación y caos

Hemos verificado la existencia del ciclo limite del sistema (3.1.1)-(3.1.2) por integración numérica usando el algoritmo de Runge-Kuta de cuarto grado para el parámetro elegido determinado para el análisis de la forma normal.

El análisis muestra que existe un estado de oscilación periódico en ausencia de cualquier corriente externa para valor del parámetro $u < 1,099$ mientras que los otros parámetros están fijos en $a = 0,6$, $b = 0,5$, $c = 0,1$. Similarmente en la presencia de corrientes constantes externas ellos existen tales oscilaciones para $A_0 > 2,163$ y $A_0 < -0,163$ para parámetros fijos en $a = 0,5$, $b = 0,5$, $c = 0,1$ y $u = 0,9$.

Después del análisis numérico uno encuentra que el sistema (3.1.1)-(3.1.2) admite movimiento caótico en ausencia completa de fuerzas externas, asignadas bifurcaciones de duplicación de período, cuando u esta decreciendo desde 1,09 hasta abajo. Existen varios atractores periodicos (periodos T, 2T y 4T) y el atractor caotico para los valores de parámetros $a = 0,6$, $b = 0,5$, $c = 0,1$ en la región $0,7 < u < 0,09$ y en ausencia de fuerzas externas. En presencia de corrientes externas constantes este sistema admite bifurcaciones de duplicación de período que conducen a un movimiento caotico en el dominio del parametro $A_0 \in (2,182; 2,224)$ y $A_0 \in (-0,182; -2,224)$ con los otros parametros permaneciendo fijos en $a = 0,5$, $b = 0,5$, $c = 0,1$ y $u = 0,75$

En la figura 3,1 hemos presentado los resultados obtenidos del análisis numerico. En la figura (3,1a) la curva umbral limita la bifurcación de Hopf y de bajo de la curva está la región de puntos fijos estables.

3.6. Los coeficientes $R_{ij}^{(2,3)}$

Los coeficientes $R_{ij}^{(2)}$ y $R_{ij}^{(3)}$ en las ecuaciones (1.3.13)estan dadas por:

$$A1) R_{11}^{(2)} = -R_{31}^{(2)} = \frac{1}{|s|} \frac{c}{u} \alpha w \left(\frac{bc}{u} - \gamma \right)^2$$

$$A2) R_{12}^{(2)} = -R_{32}^{(2)} = \frac{1}{|s|} \frac{c}{u} \alpha w^3$$

$$A3) R_{13}^{(2)} = -R_{33}^{(2)} = -\frac{1}{|s|} \frac{c}{u} 2aw^2 \left(\frac{bc}{u} - \gamma \right)$$

$$A4) R_{14}^{(2)} = -R_{34}^{(2)} = -\frac{1}{|s|} \frac{c}{u} 2\alpha \left(\frac{bc}{u} - \gamma \right) \left(\frac{bc}{u} - \lambda_1 \right)$$

$$A5) R_{15}^{(2)} = -R_{35}^{(2)} = -\frac{1}{|s|} \frac{c}{u} \alpha w^2 \left(\frac{bc}{u} - \lambda_1 \right)$$

$$A6) R_{21}^{(2)} = -\frac{1}{|s|} \frac{c}{u} \alpha (\lambda_1 - \gamma) \left(\frac{bc}{u} - \gamma \right)^2$$

$$A6) R_{21}^{(2)} = -\frac{1}{|s|} \frac{c}{u} \alpha (\lambda_1 - \gamma) \left(\frac{bc}{u} - \gamma \right)^2$$

$$A7) R_{22}^{(2)} = \frac{1}{|s|} \frac{c}{u} \alpha (\lambda_1 - \gamma) w^2$$

$$A8) R_{23}^{(2)} = -\frac{1}{|s|} \frac{c}{u} \alpha (\lambda_1 - \gamma) w \left(\frac{bc}{u} - \lambda_1 \right)$$

$$A9) R_{24}^{(2)} = -\frac{1}{|s|} \frac{c}{u} 2\alpha (\lambda_1 - \gamma) \left(\frac{bc}{u} - \gamma \right) \left(\frac{bc}{u} - \lambda_1 \right)$$

50 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

$$\text{A10)} \quad R_{25}^{(2)} = -\frac{1}{|s|} \frac{c}{u} \alpha (\lambda_1 - \gamma) w \left(\frac{bc}{u} - \lambda_1 \right)$$

$$\text{A11)} \quad R_{11}^{(3)} = -R_{31}^{(3)} = \frac{1}{|s|} \frac{c}{u} \frac{w}{3} \left(\frac{bc}{u} - \gamma \right)^3$$

$$\text{A12)} \quad R_{12}^{(3)} = -R_{32}^{(3)} = \frac{1}{|s|} \frac{c}{u} \frac{w^4}{3}$$

$$\text{A13)} \quad R_{13}^{(3)} = -R_{33}^{(2)} = -\frac{1}{|s|} \frac{c}{u} w^3 \left(\frac{bc}{u} - \gamma \right)$$

$$\text{A14)} \quad R_{14}^{(3)} = -R_{34}^{(3)} = -\frac{1}{|s|} \frac{c}{u} w^2 \left(\frac{bc}{u} - \gamma \right)^2$$

$$\text{A15)} \quad R_{21}^{(3)} = -\frac{1}{|s|} \frac{c}{u} \frac{1}{3} (\lambda_1 - \gamma) \left(\frac{bc}{u} - \gamma \right)^3$$

$$\text{A16)} \quad R_{22}^{(3)} = -\frac{1}{|s|} \frac{c}{u} \frac{w^3}{3} (\lambda_1 - \gamma)^3$$

$$\text{A17)} \quad R_{23}^{(3)} = -\frac{1}{|s|} \frac{c}{u} w^2 (\lambda_1 - \gamma) \left(\frac{bc}{u} - \gamma \right)$$

$$\text{A18)} \quad R_{24}^{(3)} = -\frac{1}{|s|} \frac{c}{u} w (\lambda_1 - \gamma) \left(\frac{bc}{u} - \gamma \right)^2$$

3.7. Bifurcaciones de órbitas homoclínicas, heteroclínicas y periódicas

3.7.1. Introducción

A seguir estudiamos las condiciones para la existencia de órbitas Homoclínicas u órbitas periódicas [CH,DE,TE] que surgen de un par de orbitas Heteroclínicas que los equilibrios A y B , para una familia de ecuaciones diferenciales de la forma.

$$\dot{x} = f(x, \alpha), \quad x \in \mathbb{R}^N, \quad \alpha \in \mathbb{R}^2 \quad (3.7.1)$$

donde asumimos que $A, B \in \mathbb{R}^N$ son equilibrios hiperbólicos de (3.7.1) para un valor del parámetro α . Por una solución de (3.7.1) es Heteroclínica desde el equilibrio A hasta el equilibrio B , la que notamos $\Gamma(t)$ solución de (3.7.1) y satisface.

$$\lim_{t \rightarrow -\infty} \Gamma(t) = a \quad \lim_{t \rightarrow -\infty} \Gamma(t) = b$$

Una solución Homoclínica para el equilibrio A es una solución de (3.7.1) que satisface la siguiente aproximación asintótica

$$\lim_{|t| \rightarrow \infty} \Gamma(t) = A$$

en [CH,DE,TE] se muestra que una bifurcación Homoclínica ocurre para un valor específico α_0 .

3.7.2. Órbitas heteroclínicas

Entonces hay dos curvas α_{aa} y α_{bb} en el espacio de parámetros los cuales nacen en α_0 las cuales corresponden a soluciones Homoclínicas. La curva α_{aa} es tangente a α_{ab} en α_0 y α_{bb} es tangente en α_0 . También consideramos la existencia de soluciones periódicas de (3.7.1). Probaremos que las curvas α_{aa} y α_{bb} forman la frontera de un sector Λ constituido por soluciones periódica de (3.7.1). Por ejemplo las ecuaciones reacción-difusión de la forma siguiente:

$$V_t = DV_{xx} + F(V, \lambda) \quad (3.7.2)$$

donde $V \in \mathbb{R}^N$, D es una matriz diagonal no negativa, y $\lambda \in \mathbb{R}$ es un parámetro. Una solución homoclínica es una onda viajera de (3.7.2) es una solución de la forma $V(x, t) = V(z)$, $z = x + ut$; corresponden a soluciones de (3.7.2) que parecen tener forma y velocidad constante [CO].

Estamos interesados en solución ondas viajeras de (3.7.2) que conecta dos puntos de equilibrio de (3.7.2); esto es asumimos que $A, B \in \mathbb{R}^N$ satisfacen $F(A, \lambda) = F(B, \lambda) = 0$ para todo λ . Aquí cero 0 es el origen en \mathbb{R}^N . Consideramos soluciones de ondas viajeras de (3.7.2) que satisfacen:

$$\lim_{z \rightarrow -\infty} V(z) = A \quad y \quad \lim_{z \rightarrow +\infty} V(z) = B$$

Observe que una solución onda viajera satisface el siguiente sistema de ecuaciones diferenciales ordinarias:

$$DV'' - uV' + F(V, \lambda) = 0,$$

y si hacemos $x = V$, $V' = y$, entonces esta ecuación es equivalente a el sistema de primer orden siguiente,

$$\dot{x} = y, \quad \dot{y} = \theta y - F(x, \lambda) \quad (3.7.3)$$

junto con las condiciones de frontera

$$\lim_{z \rightarrow -\infty} (V(z), \dot{V}(z)) = (A, 0) \quad y \quad \lim_{z \rightarrow +\infty} (V(z), \dot{V}(z)) = (B, 0)$$

3.7.3. Las bifurcaciones heteroclínicas

El problema de probar la existencia de una onda viajera de (3.7.2) se reduce a encontrar una solución homoclínica ó heteroclínica de (3.7.3). Note que la velocidad u es un parámetro por tanto depende de dos parámetros y es un caso especial de (3.7.1). Un sistema dado podría tener, para un valor

52 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

dado de parámetros frentes de ondas, pulsos, pulsos múltiples y soluciones periódicas. Si un sistema dado tiene muchas soluciones del tipo ondas viajeras, entonces la existencia de algunas de ellas (quizás, frentes de ondas viajeras) podría ser fácilmente probados, mientras la existencia de otras (quizás, pulsos) podría ser más difícil de probar; ondas más complicadas podrían presentarse como bifurcaciones de las ondas simples.

Las pruebas de estos resultados están basadas en la idea de Sil'nikov sobre el mapeo de Poincaré sobre cierta sección transversal para la órbita homoclínica ó heteroclínica. Con las variables de Sil'nikov estos mapeos de Poincaré reducen el problema a familias biparamétricas de ecuaciones transcendentales. La unicidad de órbitas periódicas y Homoclínicas sigue los argumentos del teorema de la función Implícita; la existencia de órbitas homoclínicas ó heteroclínicas se derivan de ciertas ecuaciones de bifurcación que surgen en los mapeos de Poincaré.

3.8. Bifurcaciones Homoclínicas para el modelo de Fitzhugh-Nagumo con su perturbación singular, sistema lento-rápido.

Consideramos las Ecuaciones de Fitzhugh - Nagumo

$$v_t = v_{xx} + f(v) - W \quad , \quad R_t = \varepsilon(v - \gamma W) \quad (3.8.1)$$

En (3.8.1) ε y γ son constantes positivas con $0 < \varepsilon \ll 1$. Por $f(v)$, tomamos

$$f(v) = v(1-v)(v-a), \quad 0 < a < \frac{1}{2},$$

una solución onda viajera de (3.8.1), (v, R) esta limitada, de las soluciones no constantes de la forma

$$(V(x, t), W(x, t)) = (v(z), W(z)), \quad \text{donde hacemos } z = x+ut, y \quad u = cte \quad (3.8.2)$$

sustituyendo (3.8.2) en (3.8.1), (v, W) satisface el siguiente sistema de Ecuaciones diferenciales ordinarias; donde $v = x, x' = y, z = w$; para obtener el siguiente sistema:

$$x' = y, \quad y' = ux - f(x) + y, \quad z' = \varepsilon(x - \gamma z) - u \quad (3.8.3)$$

Para este sistema γ es suficientemente grande, entonces existen 3 puntos de equilibrio. En lo que sigue llamamos $\varphi = (\varphi_1, \varphi_2, \varphi_3)$ y $\varepsilon = (\vartheta_1, 0, \vartheta_2)$ son los puntos de equilibrio no triviales $0 = (0, 0, 0)$ es el otro punto de Equilibrio.

3.8 Bifurcaciones Homoclínicas para el modelo de Fitzhugh-Nagumo con su perturbación singular, sistema lento-rápido. 53

Por un pulso indicaremos una solución $\Gamma_c = \Gamma_c(z)$ de (3.8.3) que satisface

$$\lim_{|z| \rightarrow +\infty} \Gamma_c(z) = 0$$

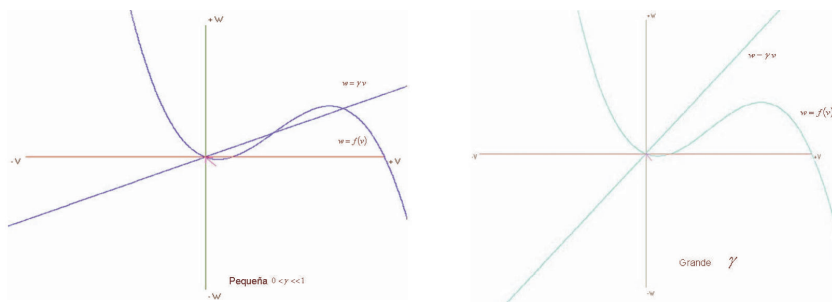


Figura 3.8 a) Modelo lento-rápido

Por un pulso ϑ , Γ_ϑ indicamos una solución homoclínica que satisface la siguiente aproximación asintótica

$$\lim_{|z| \rightarrow +\infty} \Gamma_\vartheta(z) = \vartheta$$

Por un frente de onda Γ_F indicamos una solución que satisface

$$\lim_{z \rightarrow +\infty} \Gamma_F(z) = \vartheta \quad y \quad \lim_{z \rightarrow -\infty} \Gamma_F(z) = 0$$

Por onda de retorno heteroclínica Γ_B indicamos una solución que satisface

$$\lim_{z \rightarrow +\infty} \Gamma_B(z) = 0 \quad y \quad \lim_{z \rightarrow -\infty} \Gamma_B(z) = \vartheta$$

Los parámetros relevantes son γ y la velocidad de onda u . Se han mostrado la existencia de constantes $0 < \gamma_1 < \gamma_2 < \gamma_3$ para las que tenemos lo siguiente:

- 1) si $\gamma_1 < \gamma$ y $0 < \varepsilon \ll 1$, entonces existe un frente de onda Γ_F para algún u , es decir existe una curva de bifurcación $\theta_F(\gamma)$
- 2) si $\gamma_1 < \gamma < \gamma_3$ y $0 < \varepsilon \ll 1$, entonces existe una onda retorno Γ_B para algún u , es decir existe una curva de bifurcación $\theta_B(\gamma)$;
- 3) En el límite $\varepsilon \rightarrow 0$, diagrama de bifurcación con estas curvas diferenciales $\theta_F(\gamma)$ y $\theta_B(\gamma)$ son aproximadamente como se muestra en la figura (3.8 b).

54 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

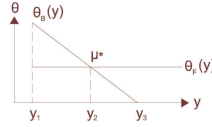


Figura 3.8 b) Diagrama de bifurcación sobre las soluciones globales del sistema FitzHugh-Nagumo

Estas dos curvas se cortan precisamente en $\gamma = \gamma_2$ en el límite $\varepsilon \rightarrow 0$ y para toda ε pequeña, pero no cero, las curvas $\theta_B(\gamma)$ y $\theta_F(\gamma)$ corresponden a las mismas características cualitativas en el espacio de fases como se muestra en la figura (3.8 c). Para γ cerca a γ_1 , $\theta_B(\gamma) > \theta_F(\gamma)$, mientras para γ cerca a γ_3 , $\theta_B(\gamma) < \theta_F(\gamma)$. Por tanto, debe un punto en el espacio de parámetros existir $\mu^* = (\gamma^*(\varepsilon), u^*(\varepsilon))$ donde las dos curvas se cruzan. No es difícil demostrar que cuando $0 < \varepsilon \ll 1$, las linealizaciones en 0 y ϑ tienen dos valores propios negativos y un valor propio positivo. Uno de los valores propios negativos es cero en el límite $\varepsilon \rightarrow 0$, mientras que los otros dos permanecen uniformemente lejos de cero cuando $\varepsilon \rightarrow 0$. Entonces hay cuatro posibilidades para la localización de el sector Λ , que corresponden en el espacio de fases a las órbitas periódicas; sin embargo; Λ es determinado por los siguientes resultados tomados del teorema (2.1.) de [CH,DE,TE]:(1) La velocidad del frente excede la velocidad del pulso; (2) cuando ambos existen, corresponde frente de ondas y a su retorno; existe unicamente un pulso si la velocidad de onda de retorno excede la velocidad del frente de ondas. La afirmación correspondiente también incluye a los pulsos ε . (Ver fig 3.8 a), b) y c)).

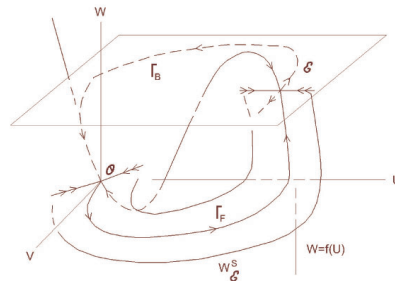


Figura 3.8 c) Retrato de fases correspondientes a soluciones globales (heteroclínicas o homoclínicas)

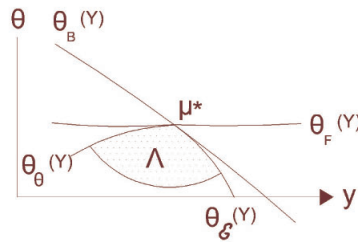


Figura 3.8 d) Diagrama de bifurcación sobre la región Ω correspondiente a soluciones n -homoclínicas y n -periódicas

La consecuencia importante de esta sección es que cerca de la solución heteroclínica, existe una cantidad infinita de ondas periódicas viajando a velocidades siempre mayores que la velocidad del pulso.

3.9. Conclusiones del capítulo tres

En la primera parte del presente artículo hemos estudiado la estabilidad lineal de la ecuación de **FitzHugh-Nagumo** tanto en la presencia como en la ausencia de corrientes externas constante. También del análisis lineal hemos encontrado el parámetro que eleva los pulsos periódicos es la presencia o ausencia de corrientes externas constantes. Para un valor de parámetro particular escogido tenemos un sistema simétrico el cual nos señala la existencia de una bifurcación codimensional dos de Takens-Bogdanov.

En la segunda parte hemos aplicado el método de la teoría de bifurcación basado en la reducción de la variedad central y el análisis de su forma normal para la ecuación de **FitzHugh-Nagumo** y así probar la existencia de pulsos periódicos estables. Finalmente hemos obtenido la región en el espacio de parámetros donde existen los equilibrios estables.

La existencia de oscilaciones caóticas y periódicas en el modelo neuronal de **FitzHugh-Nagumo**, en ausencia o presencia de cualquier estímulo externo corresponden la existencia de trayectorias periódicas del potencial de la membrana. La naturaleza de las oscilaciones en ausencia de corrientes externas es determinada por la velocidad de las ondas viajeras y esto representa la propagación del tren de ondas en la neurona; la velocidad de la onda

56 Estudio cualitativo del modelo FitzHugh-Nagumo espacio-temporal en \mathbb{R}^3

viajera, es fijada como parámetro del sistema. En la presencia de estímulo externo constante, la activación de los pulso puede ser determinado por los parámetros internos del sistema y la magnitud del estímulo externo aplicado.

Cualquier otro fenómeno ondulatorio está sujeto al control de la velocidad u . Así cuando el sistema (3.1.1)-(3.1.2) admite soluciones periódicas, estas corresponden a la propagación de pulsos espacialmente periódicos con oscilaciones periódicas temporales. Similarmente las soluciones caóticas y, las cuales surgen de la existencia de soluciones homoclínicas ó heteroclínicas corresponden también al tipo de caos espacio temporal.

CAPÍTULO 4

Las dinámicas de las redes neuronales no lineales

4.1. El modelo de Hopfield

El modelo básico de una red neuronal no lineal es el siguiente [Cr] y usa el modelo formal de una neurona dado por Mc Culloch and Pitts en 1943; se puede ver como un modelo con memoria asociativa no lineal o un modelo con memoria direccionada

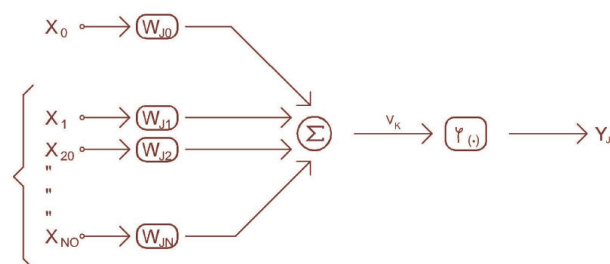


Figura 4.1 a) Modelo de Hopfield

con X_0 fijado como ± 1 , es el nivel de activación de cada neurona (on ó off) $w_{j0} = \theta_j$ es el valor del umbral valores que satisfacen

$$u_j = \sum_{i=1}^N w_{ji}x_{ij}, \quad y_j = \varphi(u_j - \theta_j) = \varphi(v_j)$$

58 Las dinámicas de las redes neuronales no lineales

la más elemental φ para de terminar el estado de cada neurona es dado por

$$s_j = \begin{cases} +1 & \text{si } v_j > 0 \\ 1 & \text{si } v_j < 0 \end{cases}$$

esto es por $s_j = \text{sign}(v_j)$. donde φ es la función de activación; la matriz de pesos signaticos es

$$W = \frac{1}{N} \sum_{j=1}^p X_{ij} X_{ij}^T - \frac{p}{N} I$$

I matriz identica. La cual se acostumbra escojer como $\varphi(v_j) = \frac{1}{1+\exp(-v_j)}$ ó también como $\varphi(v_j) = \frac{1-\exp(g_i v)}{1+\exp(-g_i v)}$ con $g_i = \left. \frac{\partial \varphi_i}{\partial v} \right|_{v=0}$; la ganancia de la neurona para esta última se tiene $v = \varphi^{-1}(x) = -\frac{1}{g_i} \ln\left(\frac{1-x}{1+x}\right)$, la evolución de los estados es dada por

$$x_j(n+1) = \text{sign}\left[\sum_{i=1}^N w_{ji} x_i(n)\right],$$

y el proceso se de tiene en los puntos fijos:

$$y = x_j(n) = x_j(n+1)$$

Para tiempos continuos en el lugar de los pesos sinapticos w_{ji} tenemos una respuesta al impulso estandar $h_{ji}(t) = w_{ji} h_0(t)$ para todos i, j ; asumiendo $h_0(t) = \frac{1}{\tau} \exp(-\frac{t}{\tau})$, τ un tiempo constante, este filtro es un circuito RC simple, como el siguiente modelo:

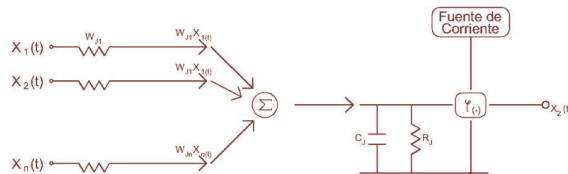


Figura 4.1 b) Modelo RC de Hopfield

Entonces la corriente que ingresa al nodo $\varphi(\cdot)$ es

$$\sum_{i=1}^N w_{ji}x_i(t) + I_j$$

y la corriente que egresa del mismo nodo es

$$\frac{v_j(t)}{R_j} + C_2 \frac{dv_j(t)}{dt}$$

por la ley de Kirchoff obtenemos el siguiente sistema acoplado de ecuaciones diferenciales:

$$C_j \frac{dv_j(t)}{dt} + \frac{v_j(t)}{R_j} = \sum_{i=1}^N w_{ji}x_i(t) + I_j \quad (4.1.1)$$

donde $C_j \frac{dv_j(t)}{dt}$ da la capacidad de agregar memoria dinámica a la neurona; tomando el potencial de activación φ , el cual generalmente es la función logística

$$\varphi(v_i) = \frac{1}{1 + \exp(-v_j)}$$

$j = 1, 2, 3, \dots, N$ (4.1.1) también se puede escribir como

$$\frac{dv_j(t)}{dt} = -v_j(t) + \sum_i w_{ji}\varphi(v_i(t)) + I_j, \quad j = 1, 2, 3, \dots, N \quad (4.1.2)$$

La más común si φ_j es el mismo para todo j , y además los pesos y la corriente externa también se normalizan con respecto a R_j . Entonces w_{ij} forma la matriz de pesos sináptica e I_j la matriz de bias.

El modelo modelo (4.1.2) es básicamente el siguiente sistema acoplado de ecuaciones diferenciales:

$$\frac{dx_j(t)}{dt} = -x_j(t) + \varphi\left(\sum_i w_{ji}v_i(t)\right) + k_j, \quad j = 1, 2, 3, \dots, N \quad (4.1.3)$$

usamos la transformación $v_k(t) = \sum_j w_{kj}x_j(t)$ y $I_k = \sum_j w_{kj}K_j, K_j$ donde constituye el vector de bias. Para que los modelos discretos que generen procesos de aprendizaje es necesario que los sistemas acoplados tengan equilibrios atractores. Además poseen retroalimentación, como se ve en los siguientes diagramas de bloques.

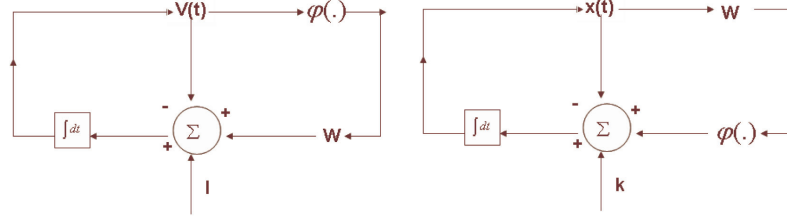


Figura 4.1 c) Sistema dinámico de Hopfield

En 1984 [Ho] propuso la siguiente función de energía de Lyapunov la cual se usa para el estudio de la estabilidad orbital de (4.1.3) (estable o asintóticamente estable):

$$E(\vec{x}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ji} x_i x_j + \sum_{j=1}^N \frac{1}{R_j} \int_0^{x_j} \varphi_j^{-1}(x) dx - \sum_{j=1}^N I_j x_j \quad (4.1.4)$$

cuya derivada con respecto al tiempo es:

$$\frac{dE}{dt} = - \sum_{j=1}^N \left(\sum_{i=1}^N w_{ji} x_i - \frac{v_j}{R_j} + I_j \right) \frac{dx_j}{dt} \quad (4.1.5)$$

De (4.1.2) tenemos la siguientes propiedades:

$$\begin{aligned} \frac{dE}{dt} &= - \sum_{j=1}^N C_j \left(\frac{dv_j}{dt} \right) \frac{dx_j}{dt} \\ \frac{dE}{dt} &= - \sum_{j=1}^N C_j \left[\frac{d}{dt} \varphi_j^{-1}(x_j) \right] \frac{dx_j}{dt} \\ \frac{dE}{dt} &= - \sum_{j=1}^N C_j \left(\frac{dx_j}{dt} \right)^2 \left[\frac{d}{dx_j} \varphi_j^{-1}(x_j) \right] \leq 0 \quad \text{para } x_j \neq 0; \quad \forall_j; \end{aligned} \quad (4.1.6)$$

Si

$$x = \varphi_i(v) = \frac{1 - \exp(-g_i v)}{1 + \exp(-g_i v)}; \quad g_i = \left. \frac{d\varphi}{dv} \right|_{v=0}$$

entonces

$$v = \varphi_i^{-1}(x) = -\frac{1}{g_i} \ln \left(\frac{1-x}{1+x} \right)$$

En resumen tenemos el siguiente teorema de Hopfield para las redes neuronales

Teorema 4.1.1. *la función de energía E para una red Hopfield consistente de N neuronas es una función monótona decreciente de los estados de las red $\{x_j|_{j=1,2,3,\dots,N}\}$.*

Hasta ahora el modelo de Hopfield es el siguiente:

$$\begin{aligned} x_i &= \varphi(u_i) \\ \frac{du_i}{dt} &= E(u_i, x, w, I) \\ E(x) &= E(u_i, x, w, I) \end{aligned} \quad (4.1.7)$$

con

$$\frac{du_i}{dt} = -\frac{\partial E}{\partial x_i} \quad (4.1.8)$$

en (4.1.8) se tiene un sistema gradiente. El modelo (4.1.7) se usa con eficiencia para resolver problemas de optimización que se pueden ubicar en un hipercubo, como por ejemplo: El problema de viajante comercio (TPS; Traveling Saleman Problem), la coloración de mapas con k colores (kCP, k color problem), partición de un grafo (GBP, Graph Bipartition problem) ver [AT,JO,SA].

4.1.1. Identificación de sistemas dinámicos mediante redes neuronales

Para un sistema dinámico

$$y = \frac{dx}{dt} = g(x, P, \theta) \quad (4.1.9)$$

lo expresamos como operador lineal de sus parámetros

$$y = \frac{dx}{dt} = A(x, P)(\theta_n + \theta) \quad (4.1.10)$$

con estimación $\hat{\theta}(t)$ y predicción $e(t)$ son las siguientes:

$$\begin{aligned} \tilde{\theta}(t) &= \theta - \hat{\theta} \\ e(t) &= y - f(x, P, \tilde{\theta}), \quad \theta \text{ desconocido} \end{aligned} \quad (4.1.11)$$

esto es, $e(\tilde{\theta}) = A(\theta_n + \theta) - A(\theta_n + \tilde{\theta}) = A(\theta - \tilde{\theta}) = A\tilde{\theta}$. Obtenemos el nuevo sistema dinámico [AT,JO,SA] siguiente,

$$\frac{d\tilde{\theta}}{dt} = -k\nabla(e^t.e) \quad (4.1.12)$$

62 Las dinámicas de las redes neuronales no lineales

sistema que minimiza la norma $\|e\|^2 = e^T \cdot e$, k resulta ser la ganancia del estimador. Alternativamente en el modelo de Hopfield, en la versión de Abe, se fuerza la equivalencia entre la función de Lyapunov $E(w)$ con la función objetivo $\frac{1}{2}\|e\|^2$, resultando

$$E = \frac{1}{2}e^T \cdot e = \frac{1}{2}\tilde{\theta}^T A^T A \theta + \theta^T (A^T A \theta_n - A^T y) + \frac{1}{2}\|y - A \theta_n\|^2,$$

cuyos estados s representan en cada instante la estimación propuesta $\tilde{\theta}(t)$, con pesos w y bias I siguientes:

$$w = -A^T A, \quad I = A^T A \theta_n - A^T y$$

Ejemplo 4.1.1. para un brazo rigido de robot, con P un par aplicado por un motor

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{g}{l} \text{sen} x_1 - \frac{v}{ml^2} x_2 + \frac{1}{ml^2} P, \end{aligned}$$

realizamos las transformaciones

$$y = \frac{dx_2}{dt}$$

$$\theta_n + \theta = \left(-\frac{g}{l}, -\frac{v}{ml^2}, \frac{1}{ml^2} \right)^T$$

$$A = (\text{sen} x_1, x_2, P)$$

mediante el sistema $y = \frac{dx}{dt} = A(\theta_n + \theta)$ por el método del gradiente o el modelo de Hopfield, podemos encontrar las curvas de estimación, para $\theta_1 = -\frac{g}{l}$

Ejemplo 4.1.2. Para el sistema de Lorenz

$$\dot{x} = -\beta x_1 + x_2 x_3$$

$$\dot{x}_2 = w(x_3 - x_2)$$

$$\dot{x}_3 = -x_1 x_2 + \rho x_2 - x_3$$

donde seleccionaremos $w = 10$, $\beta = \frac{8}{3}$ y $\rho = 28$, para obtener los estados en tiempo discreto usamos el sistema en diferencias $\dot{x} = Ax$ donde:

$$x = [x_1, x_2, x_3]^T, \quad A = \begin{bmatrix} -\beta & 0 & x_2 \\ 0 & -\sigma & \sigma \\ -x_2 & \rho & -1 \end{bmatrix}, \quad s_1 = Ax_k$$

$$s_2 = A(x_k + s_1), \quad s_3 = A\left(x_k + \frac{s_1 + s_2}{4}\right); \quad \text{si } \left|\frac{s_1 - 2s_3 + s_2}{3}\right| \leq \frac{|x_k|}{1000}$$

$$\vee \left|\frac{s_1 - 2s_3 + s_2}{3}\right| < 1, \quad \text{entonces } x_{k+1} = x_k + \frac{s_1 + 4s_3 + s_2}{6}, \quad k = 0, 1, 2, \dots$$

se trata de un sistema no lineal discreto sin entradas de control de la forma $x(k+1) = f(x(k))$ con una red neuronal multicapa

$$B\hat{x}(k+1) = A\hat{x}(k) + w_k\sigma[x(k)],$$

con

$$w_1(0) = \begin{bmatrix} \text{randon}(0, 1) \\ 1 \\ 1 \end{bmatrix}$$

son los pesos de las capas ocultas, $\sigma = \tan h(\cdot)$, $\beta = 4$; actualizando los pesos mediante la relación

$$w_1(k+1) = w_1(k) - \frac{n}{1 + \|\sigma\|^2} e(k)\sigma^T,$$

$$e(k) = \hat{x}(k) - x(k);$$

$n \leq 1$, para un aprendizaje estable. Así obtenemos los valores mas adecuados del comportamiento de la red neuronal monocapa con respecto a la señal del sistema

Ejemplo 4.1.3. En general para un sistema con entradas de control $n(t)$

$$\frac{dx(t)}{dt} = \phi(x(t), u(t))$$

$$y(t) = \psi(x(t))$$

o en tiempo discreto

$$x(k+1) = \phi(x(k), u(k))$$

$$y(k) = \psi(x(k))$$

se tiene una red neuronal en tiempo discreto como la siguiente

$$\beta\hat{x}(k+1) = A\hat{x}(k) + \sigma[w_{1,t}(k)\hat{x}(k)] + \phi[w_{2,t}(k)\hat{x}(k)]u(k),$$

donde $\hat{x}(k) \in \mathfrak{R}^n$ representa el estado interno de la red neuronal; $A \in \mathfrak{R}^{n \times n}$ es una matriz estable; $w_{1,t}$ y $w_{2,t}$ son matrices en $\mathfrak{R}^{n \times n}$ que corresponden a los pesos de las redes neuronales; $\sigma(\cdot) \in \mathfrak{R}^n$ es una función monóticamente creciente; $\phi = \text{diagonal}(\phi_i(\hat{x}_j))r$ donde $\sigma_i(\cdot)$ y $\phi_i(\cdot)$ son funciones sigmoidales del tipo

$$\sigma_i(x_i) = \frac{a_i}{1 + e^{-b_i x_i}} - c_i$$

4.2. El paradigma de la propagación hacia atrás

Como ejemplo, tenemos el problema de la OR exclusiva, cuya red básica requiere de otra capa de neuronas ocultas para su solución como se muestra en las siguientes gráficas (en la segunda se colocan valores de umbral).

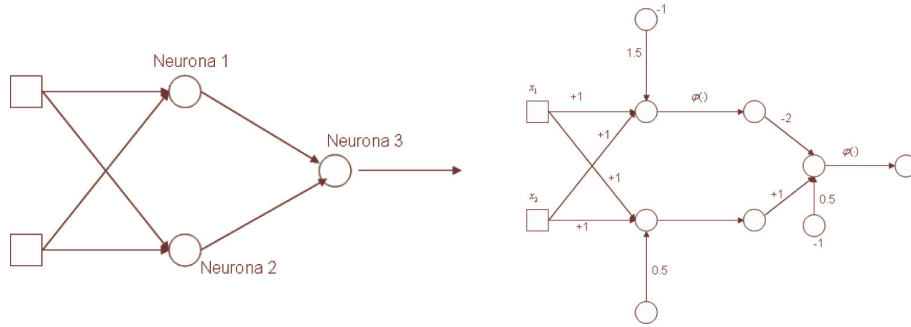


Figura 4.2 a) Modelo de propagación hacia atrás

En este modelo consideramos el error con signo de la neurona de salida “j” en el n-iterado definido como

$$e_j(n) = d_j(n) - y_j(n) = d_j(n) - \varphi_j^0\left(\sum_j w_{kj}^0 L_{pj} + \theta_k^0\right) \quad (4.2.1)$$

donde $d_j(n)$ es la salida deseada en la neurona j para el estado n ,

$$i_{pj} = \varphi_{pj}\left(\sum_{i=1}^N w_{pi}^h + \theta_j^h\right)$$

ϕ_j^h es el umbral para la neurona oculta j (hidden,) la media suma de errores cuadráticos simultáneos en todos los estados $S \subseteq \mathbb{N}$ es

$$\xi(n) = \frac{1}{2} \sum_{j \in S} e_j^2(n) \quad (4.2.2)$$

con error cuadrático promedio

$$\xi_{prom}(N) = \frac{1}{N} \sum_{n=1}^N \xi(n), \quad (4.2.3)$$

esta última es llamada la función de costo dependiente de los parámetros libres (pesos sinápticos y valores de umbral), como se representa en la siguiente red:

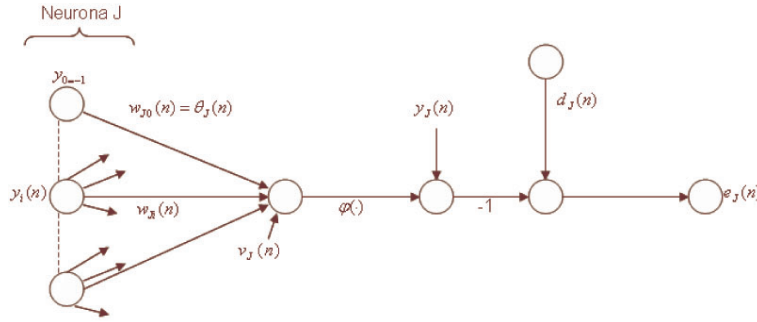


Figura 4.2 b) Propagación hacia atrás con la función costo dependiente de parámetros

De $y_j = \varphi_j(v_j(n))$ con $v_j(n) = \sum_{i=1}^N w_{ji}(n) \cdot y_i(n)$ obtenemos el siguiente gradiente o factor de sensibilidad:

$$\frac{\partial \xi(n)}{\partial w_{ji}(n)} = \frac{\partial \xi(n) \cdot \partial e_j(n) \cdot \partial y_j(n) \cdot \partial v_j(n)}{\partial e_j(n) \cdot \partial y_j(n) \cdot \partial v_j(n) \cdot \partial w_{ji}(n)} \quad (4.2.4)$$

$$\frac{\partial \xi(n)}{\partial w_{ji}(n)} = e_j(n)(-1) \cdot \varphi'_j(v_j(n)) \cdot y_i(n)$$

para razón de aprendizaje definida en las correcciones de los pesos así:

$$\Delta w_{ji}(n) = -\eta \frac{\partial \xi(n)}{\partial w_{ji}(n)} \quad (4.2.5)$$

que en términos de $\delta_j(n)$ se escribe como $\Delta w_{ji} = \eta \delta_j(n) \cdot y_i(n)$ y se concluye que:

$$\delta_j(n) = e_j(n) \cdot \varphi'_j(v_j(n)) \quad (4.2.6)$$

para muchas neuronas N tomamos los siguientes términos bias:

$$I_j = \begin{cases} \xi_j & \text{para la neurona de salida } j \\ 0 & \text{para todo valor;} \end{cases} \quad (4.2.7)$$

y para las respuestas en los estados finales de reposo $x_j(\infty)$ considera las siguientes señales errores con signo:

$$e_j = \begin{cases} d_j - x_j(\infty) & \text{para la neurona de salida } j \\ 0 & \text{para todo valor;} \end{cases} \quad (4.2.8)$$

el objetivo es encontrar un algoritmo que minimice la siguiente función

$$\xi = \frac{1}{2} \sum_k e_k^2$$

considerando los cambios en los pesos sinápticos w_{ij} por el siguiente método de pasos descendentes

$$\Delta w_{ij} = -\eta \frac{\partial \xi}{\partial w_{ij}} \quad (4.2.9)$$

donde η es una constante de proporcionalidad, la cual debe ser controlada; de tal manera que $x_j(\infty)$ y $\{\xi_j\}$ constituyan un atractor para los valores de salida deseados $\{d_j\}$. Esto implica la construcción de una red adjunta con señales contrarias, la cual funciona para el ajuste de los pesos para el tiempo futuro “ $t + 1$ ” así:

$$w_{ji}^h(t + 1) = w_{ji}(t) + \eta \delta_{pj}^h x_i(t)$$

η es la velocidad de aprendizaje y proviene del sistema gradiente descendente como se ilustra en la segunda de las siguientes redes:

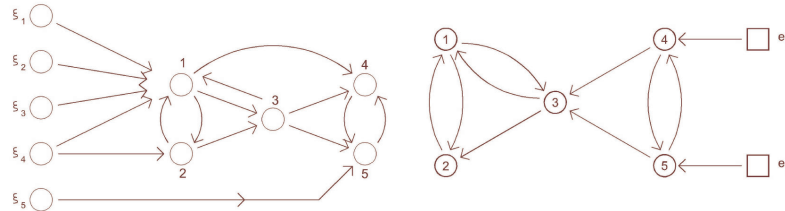


Figura 4.2 c) La segunda gráfica corresponde a la red adjunta (Dual) del modelo de propagación hacia atrás

Más exactamente la computación hacia atrás ajusta los pesos sinápticos en la capa l de acuerdo a la siguiente regla

$$w_{ji}^l(n + 1) = w_{ji}^l(n) + \alpha [w_{ji}^l(n) - w_{ji}^l(n + 1) - w_{ji}^l(n - 1)] + \eta \delta_j^l(n) y_i^{l-1}(n),$$

donde α es la constante de momento; y

$$\delta_j^L(n) = \exp_j^L y_j^L [1 - y_j^L],$$

para la neurona j en la capa de salida L ;

$$\delta_j^l(n) = y_j^l(n) [1 - y_j^l(n)] \sum_k \delta_k^{l+1}(n) w_{kj}^{l+1}(n)$$

para la neurona j en la capa oculta l . En forma continua la evolución de la red en el espacio de pesos sinápticos y umbrales (4.2.9) se relaciona con la ecuación de (4.1.1).

$$\frac{dv_j(t)}{dt} = -v_j(t) + \sum_i w_{ij}x_i(t) + I_j, \quad j = 1, 2, 3, \dots, N \quad (4.2.10)$$

donde $x_i = \varphi(v_i)$, el vector $x(t)$ esta en el de espacio de fases; (4.2.10) es la ecuación de propagación hacia adelante. En resumen, la relación entre las dos redes anteriores es la siguiente

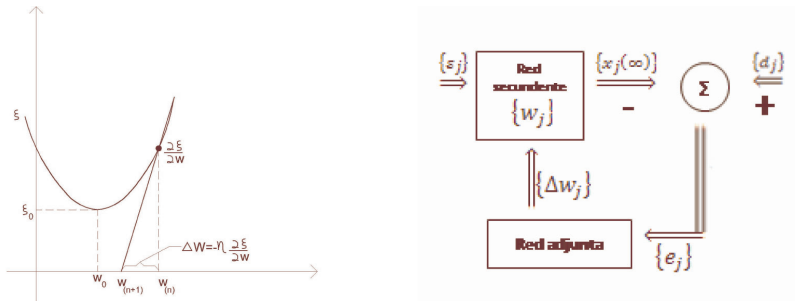


Figura 4.2 d) Resumen de sistema dinámico del modelo de propagación hacia atrás.

De la (4.2.8) y (4.2.9) tenemos

$$\frac{\partial \xi}{\partial w_{ij}} = - \sum_k e_k \frac{\partial x_k(\infty)}{\partial w_{ij}} \quad (4.2.11)$$

lo que conduce a

$$\Delta w_{ij} = \eta \sum_k e_k \frac{\partial x_k(\infty)}{\partial w_{ij}} \quad (4.2.12)$$

En lo que sigue estudiamos los puntos de equilibrio del sistema de propagación hacia atrás:

$$\frac{dy_j(t)}{dt} = -y_j(t) + \sum_i \varphi'(v_i(\infty))w_{ij}y_i(\infty) + e_j, \quad j = 1, 2, 3, \dots, N \quad (4.2.13)$$

Usando el siguiente procedimiento para el cálculo de $\frac{\partial x_k(\infty)}{\partial w_{rs}}$:

- a) El valor del potencial activación de la neurona j en el tiempo $t = \infty$ es

$$v_j(\infty) = \sum_i w_{ij}x_i(\infty) + I_j \quad (4.2.14)$$

y teniendo en cuenta que

$$\frac{\partial x_j(\infty)}{\partial w_{rs}} = \frac{\partial x_j(\infty)}{\partial v_j(\infty)} \cdot \frac{\partial v_j(\infty)}{\partial w_{rs}},$$

y de $x_j(\infty) = \varphi(v_j(\infty))$ obtenemos

$$\frac{\partial x_j(\infty)}{\partial v_j(\infty)} = \varphi'(v_j(\infty))$$

b) De (4.2.14) tenemos

$$\frac{\partial v_j(\infty)}{\partial w_{rs}} = \sum_i \left(\frac{\partial w_{ji}}{\partial w_{rs}} x_i(\infty) + w_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}} \right),$$

con

$$\frac{\partial w_{ij}}{\partial w_{rs}} = \delta_{jr} \Upsilon_{is}$$

donde δ_{jr} son los respectivos deltas de Kronocker, entonces

$$\frac{\partial v_j(\infty)}{\partial w_{rs}} = \Upsilon_{jr} x_s(\infty) + \sum_i w_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}}$$

c) De (4.2.12) usando el hecho de que

$$\sum_r L_{rj} \sum_k e_k (L_{kr}^{-1}) = e_j \quad (4.2.15)$$

y además

$$y_r(\infty) = \sum_k e_k (L_{kr}^{-1});$$

obtenemos

$$\sum_r \{ \delta_{ij} - \varphi'(v_r(\infty)) w_{rj} \} y_r(\infty) = e_j,$$

lo que conduce a la siguiente expresión

$$y_j(\infty) - \sum_r \varphi'(v_r(\infty)) w_{rj} y_r(\infty) = e_j,$$

esto es

$$0 = y_j(\infty) + \sum_i \varphi'(v_i(\infty)) w_{ij} y_i(\infty) + e_j, \quad j = 1, 2, 3, \dots, N$$

que corresponde a los puntos de equilibrio de (4.2.13); desafortunadamente no existe un método para garantizar la convergencia hacia un punto fijo de la propagación hacia atrás.

4.3. Modelo de aprendizaje de una neurona débilmente no lineal

En los últimos años las redes neuronales artificiales han sido estudiadas muy intensamente. Hay muchos artículos que describen las aplicaciones de las redes neuronales en la solución de problemas de la teoría de control, robótica, reconocimiento de voz, reconocimiento de patrones, comprensión de datos, sistemas expertos y muchos otros. En problemas del proceso de aprendizaje de las redes neuronales multicapas, los modelos matemáticos son una herramienta conveniente para la investigación. De otro lado recientemente algunos artículos han sido dedicados al estudio de las propiedades cualitativas obtenidas por el método de discretización de la variable temporal.

La pregunta básica es, ¿Si las propiedades cualitativas de los sistemas dinámicos de tiempo-continuo son preservadas bajo la discretización?. Para responder esta pregunta usamos varios conceptos sobre sistemas dinámicos diferenciables: Las propiedades de estabilidad y atracción. La estructura que surge del equilibrio punto silla, sus variedades invariantes, los promedios invariantes de la topología algebraica pueden ser mencionados como ejemplos. También han sido estudiadas aplicaciones numéricas, así como la conjugación topológica local y global.

El análisis de una neurona débilmente no lineal está basado en el teorema de *Fečkan*, mientras que el caso de una neurona no lineal el resultado del teorema de Bielecki [Bi].

4.3.1. Modelo matemático de una neurona y su proceso de aprendizaje

Una neurona artificial es una unidad provista de algunos pesos de entrada y una salida. Así, podemos decir que una neurona es un operador F el cual es un mapeo

$$F : \mathbb{R}^k \longrightarrow \mathbb{R} \ni (\mathbf{w}, \mathbf{x}) \longrightarrow F(\mathbf{w}, \mathbf{x}) = \varphi(\mathbf{w} \cdot \mathbf{x}) \in \mathbb{R} \quad (4.3.1)$$

donde φ , la llamada función de activación de una neurona, es un mapeo $\varphi : \mathbb{R} \ni \beta \longrightarrow \varphi(\beta) \in \mathbb{R}$; donde w corresponde a los pesos siguientes

$$\boxed{w_1 \quad w_2 \quad \dots \quad w_k} \quad (4.3.2)$$

Una red neuronal artificial es un sistema de neuronas, las cuales están

conectadas de tal manera que la señal de salida de una neurona es la entrada de otra neurona. Muy frecuentemente las redes multicapas son consideradas; los pesos de la neurona son determinados durante el proceso de aprendizaje. Así una neurona entrenada es un mapeo

$$F^* := F(\mathbf{w}, \cdot) : \mathbb{R}^k \longrightarrow \mathbb{R} \quad (4.3.3)$$

Hay varios métodos de aprendizaje para las redes neuronales artificiales, muchos de ellos son procesos iterativos. Observamos que el método de gradiente descendente lleva a la variación de la sinápsis de la forma siguiente

$$\mathbf{w}^{(t+1)} = \mathbf{w}^p - h \cdot \text{grad}E(\mathbf{w}) \quad (4.3.4)$$

donde $w = [w_1, w_2, \dots, w_k]$ es un vector de pesos de una neurona, iterativamente generado por el método de Euler para la ecuación diferencial

$$w = -\text{grad}E(w) \quad (4.3.5)$$

Si consideramos una neurona, una función de desviación llamada también función criterio, que representa el papel del potencial E en la ecuación gradiente (4.3.5). Con frecuencia se usa el criterio del mínimo cuadrado para definir la función criterio de cada neurona, que esta dado por la fórmula

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N [\varphi(\beta^n) - z^{(n)}]^2 \quad (4.3.6)$$

donde f es una función de activación de una neurona, $\beta^{(n)} = x_1^{(n)}w_1 + \dots + x_k^{(n)}w_k$ y el vector $\mathbf{x}^{(n)} = [x_1^{(n)}, \dots, x_k^{(n)}]$ es una señal de entrada. El número $z^{(n)}$ es una respuesta deseada de la neurona si un vector $\mathbf{x}^{(n)}$ es la señal de entrada y n es el número de vectores de entrada usados en el proceso de aprendizaje. La sucesión finita $((\mathbf{x}^{(1)}, \mathbf{z}^{(1)}), \dots, (\mathbf{x}^{(N)}, \mathbf{z}^{(N)}))$ es llamada secuencia de aprendizaje.

4.3.2. Conjugación topológica de cascada de gradientes

Aplicaremos dos teoremas al análisis del proceso de aprendizaje. Los teoremas considerados son aplicados a la ecuación diferencial y al método de Euler. El h-mapeo relacionado con el tiempo inducido por las soluciones es comparado con la cascada obtenida por el método de Euler.

El resultado obtenido por Bielecki que es un sistema dinámico en una

esfera bidimensional S^2 bajo alguna suposición natural, es correctamente reproducido por el método de Euler para un paso de tiempo suficientemente pequeño. Esto significa que el h-mapeo relacionado con el tiempo del sistema dinámico inducido esta topologicamente globalmente conjugado con el sistema dinámico discreto obtenido por el método de Euler, el cual puede ser sintetizado como sigue:

Teorema 4.3.1. *Sea*

$$\phi : S^2 \times \mathbb{R} \longrightarrow \mathbb{R}^2 \quad (4.3.7)$$

Un sistema dinámico que tiene un número finito de singularidades, todas hiperbólicas, y generadas por la ecuación diferencial

$$\dot{X} = -\text{grad}E(X) \quad (4.3.8)$$

donde $E \in C^2(S^2, \mathbb{R})$. Supongase que además, el sistema dinámico ϕ no tiene conexiones silla-silla. Además, asuma que $\phi_h : S^2 \longrightarrow S^2$ es una discretización del sistema ϕ , es decir $\phi_h(x) = \phi(x, h)$; mientras que un mapeo $\psi_h : S^2 \longrightarrow S^2$ es generado por el método de Euler para la ecuación (4.3.8). Entonces, para $h > 0$ suficientemente pequeño, y existe un homeomorfismo $\alpha = \alpha_h : S^2 \longrightarrow S^2$ el cual ϕ_h y ψ_h , esto es:

$$\psi_h \circ \alpha = \alpha \circ \phi_h \quad (4.3.9)$$

La prueba del teorema se presenta en [BiI].

4.3.3. Comentario

- 1) Como es conocido, el axioma A y la fuerte condición de transversalidad son equivalentes a la estabilidad estructural de un sistema dinámico (4.3.8) [PA,ME]. Por otro lado, para sistemas dinámicos gradientes, el axioma A implica que el sistema tiene unicamente un número finito de singularidades, todas hiperbólicas; mientras que la condición de transversalidad fuerte implícita que el sistema gradiente no tiene conexión sillan-sillas. Así que la estabilidad estructural del sistema dinámico considerado (S^2, ϕ) implica la suposición del teorema (4.3.1). Además, el conjunto de sistemas dinámicos estructuralmente estables es abierto y denso en el espacio de sistemas dinámicos gradientes. (ver [PA,ME], Pág 116).
- 2) Un sistema dinámico generado por la ecuación (4.3.8) tiene únicamente un número finito de singularidades, todas hiperbólicas, y sin conexiones sillan-silla; este sistema es llamado un sistema gradiente Morse-smale.

En el teorema *Fečkan* se considera la relación entre un flujo lineal debilmente perturbado con la discretización siguiente

Teorema 4.3.2. *Sea (ϕ, \mathbb{R}^m) el flujo generado por $\dot{X} = Ax + g(x)$ donde $A \in C(\mathbb{R}^m)$ no tiene valores propios en el eje imaginario, $g \in C^1(\mathbb{R}^m, \mathbb{R}^m)$, $g(0) = 0$ para cada $x \in \mathbb{R}^n$ y un b suficientemente pequeño, con las siguientes desigualdades:*

$$\sup |g(x)| < \infty \quad y \quad |Dg(x)| \leq b.$$

Entonces para un conjunto compacto $K \subset \mathbb{R}^m$ allí existe un número $h_0 > 0$ y una C° conjugación

$$H_h : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad h \in (0, h_0)$$

tal que sobre el conjunto K se cumple la siguiente conjugación:

$$\phi(\cdot, h) \circ H_h(\cdot) = H_h(\cdot) \circ \psi_h(\cdot)$$

donde ψ_h es el mapeo dada por el método de Euler

$$\psi_h(x) = x + h \cdot Ax + h \cdot g(x)$$

4.3.4. Neuronas débilmente no lineales

Usaremos los teoremas de Grobman-Hartman y *Fečkan* para el análisis del proceso de iteraciones, el cual es usado para las series de neuronas debilmente no lineales.

Definición 4.3.3. *Una neurona es llamada debilmente no lineal si su función de activación es de la forma*

$$\varphi : \mathbb{R}^n \ni \beta \longrightarrow \varphi(\beta) = \beta + g(\beta)$$

Donde g satisface la suposición, considerada en el teorema (4.3.2) para la función g y además, para cada β la segunda derivada $|D^2 g(\beta)| < c$, $|\beta g'(\beta)| < d_1$, $|\beta g''(\beta)| < d_2$, donde las constantes c, d_1, d_2 son suficientemente pequeñas.

Parece que las neuronas debilmente no lineales no han sido suficientemente estudiadas, porque el proceso de aprendizaje de una neurona debilmente no lineal tiene la misma dinámica que el sistema dinámico lineal. El siguiente teorema nos permite mostrar esto.

Teorema 4.3.4. *El método de aprendizaje del gradiente descendiente de una neurona débilmente no lineal es modelado por una ecuación diferencial de tipo (4.3.8).*

Demostración. : Sea una neurona que tiene una M componentes de entrada y una secuencia de aprendizaje dada por $(\mathbf{x}^{(1)}, z^{(1)}), (\mathbf{x}^{(2)}, z^{(2)}), \dots, (\mathbf{x}^{(N)}, z^{(N)})$. Entonces la excitación total de una neurona en el n-ésimo paso del proceso de aprendizaje esta dado por:

$$\beta^{(n)} = \sum_{m=1}^M x_m^{(n)} w_m^{(n)}$$

asumiendo que una neurona es débilmente no lineal, si la función de activación es de la forma

$$\varphi(\beta^{(n)}) = \beta^{(n)} + g(\beta^{(n)})$$

mientras que la función criterio de media cuadrática esta dada por

$$E(w_1, w_2, \dots, w_M) = \frac{1}{2} \sum_{n=1}^N [\beta^{(n)} + g(\beta^{(n)}) - z^{(n)}]^2$$

podemos realizar el cálculo de la k-ésima componente del gradiente de la función costo:

$$\frac{\partial E}{\partial w_k} = \sum_{n=1}^N [\beta^{(n)} + g(\beta^{(n)}) - z^{(n)}] [x_k^{(n)} + g'(\beta^{(n)}) \cdot x_k^{(n)}]$$

esto es,

$$\frac{\partial E}{\partial w_k} = \sum_{n=1}^N x_k^{(n)} \beta^{(n)} + \sum_{n=1}^N x_k^{(n)} [g(\beta^{(n)}) + g'(\beta^{(n)}) f(\beta^{(n)}) - z^{(n)} g'(\beta^{(n)}) - z^{(n)}]$$

Así, el gradiente de la ecuación diferencial que modela el proceso de aprendizaje es la siguiente

$$\dot{\mathbf{w}} = -grad E(\mathbf{w})$$

puede ser escrita en la siguiente forma:

$$\dot{\mathbf{w}} = -(A(\mathbf{w}) + \tilde{g}(\mathbf{w}))$$

Asumamos que la entrada de la neurona tiene M-componentes y la secuencia de aprendizaje consiste de N-elementos. La matrix A es entonces:

$$A = \begin{pmatrix} \bar{x}_1 \circ \bar{x}_1 & \cdots & \bar{x}_1 \circ \bar{x}_M \\ \vdots & \vdots & \vdots \\ \bar{x}_M \circ \bar{x}_1 & \cdots & \bar{x}_M \circ \bar{x}_M \end{pmatrix} \quad (4.3.10)$$

Donde \bar{x}_M es un vector con N-componentes. Las señales dadas en la m-ésima entrada de la neurona son sus N-componentes para $m = 1, 2, \dots, M$. La k-ésima componente del vector $\tilde{g}(w)$ es de la forma:

$$\tilde{g}(w) = \sum_{n=1}^N x_k^{(n)} [g(\beta^{(n)}) + g'(\beta^{(n)})f(\beta^{(n)}) - z^{(n)} g'(\beta^{(n)}) - z^{(n)}]$$

Así, si el mapeo de g satisface las suposiciones específicas (4.3.1) del comentario(4.3.3), entonces el mapeo \tilde{g} satisface las hipótesis del teorema (4.3.1). Esto completa la prueba. Las propiedades de la función g indicada en el comentario (4.3.3) implica también la suposición del teorema Grobman-Hartman. Esto significa que se tiene el siguiente teorema: \square

Teorema 4.3.5. *El flujo generado por la ecuación*

$$\dot{\mathbf{w}} = -(A\mathbf{w} + \tilde{g}(\mathbf{w}))$$

Es globalmente topológicamente conjugado con el flujo generado por la discretización h-mapeo del tiempo generado por la ecuación siguiente

$$\dot{\mathbf{w}} = -A\mathbf{w}.$$

Además, la cascada generada por la discretización h-mapeo del tiempo es globalmente topológicamente conjugado con el flujo generado por el sistema no-lineal siguiente:

$$\dot{\mathbf{w}} = -(A\mathbf{w} + \tilde{g}(\mathbf{w})) \tag{4.3.11}$$

En una bola grande, el h-mapeo es conjugado con la cascada generada por el método de Euler para la ecuación (4.3.11).

Notamos que el número del rango de valores en el cual se puede representar en un computador esta acotado. Además, en una célula neuronal, los neurotransmisores son liberados en pequeñas cantidades, en forma de paquetes alrededor de 10^{-17} por impulso, el impulso de entrada no puede ser muy grande porque la célula puede ser destruida. Así, tanto en las redes neuronales artificiales como en las biológicas, los valores absolutos de los vectores \mathbf{w} y \mathbf{x} estan acotadas, y por lo tanto en las modelaciones numéricas de neuronas, tanto biológicas como artificiales, necesitamos únicamente considerar los vectores frontera \mathbf{x} y \mathbf{w} . Así, la consideración topológicamente del conjugado en una bola grande es adecuada para el análisis de un proceso de aprendizaje. El proceso de aprendizaje se implementara en un computador

de acuerdo al método de Euler.

El teorema (4.3.5) garantiza que su dinámica es la misma que la dinámica de una cascada lineal sólo si la matrix A no tiene valores propios sobre el eje imaginario. Esto implica, por ejemplo, la estabilidad asintótica de un proceso de aprendizaje. Sin embargo, se hace necesario estimar el valor de h_0 para aplicar los resultados de nuestros análisis a redes neuronales que estan implementadas en un computador.

4.3.5. Dinámica del proceso de aprendizaje de una neurona no lineal

Consideremos la red neuronal artificial unicapa constituidas de neuronas no lineales . Asumiremos que la entrada de cada neurona tiene dos componentes y además se afirma que la función de activación de cada neurona es un mapeo acotado de clase dos $C^2(\mathbb{R}, \mathbb{R})$, con primera y segunda derivadas también acotadas. En la práctica se utilizan más tipos de funciones de activación por ejemplo, las funciones bipolar y unipolar; otras funciones radiales satisface las suposiciones especificadas. Puesto que las neuronas aprende independientemente en una red unicapa, podemos considerar el proceso de aprendizaje únicamente para solo una neurona. Usando el método del gradiente descendiente bajo las hipótesis el teorema (4.3.1) tenemos la estabilidad asintótica del proceso de aprendizaje.

Consideremos la esfera

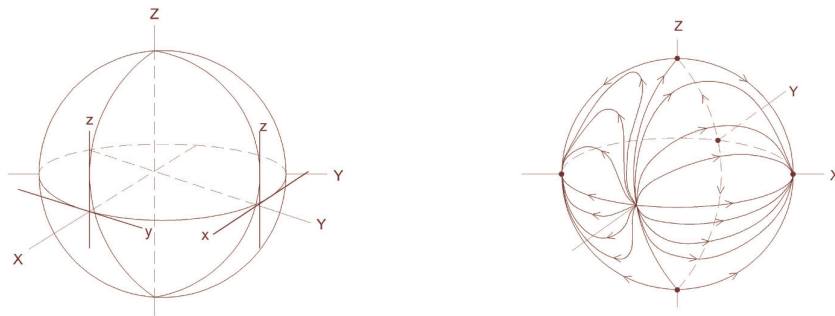


Figura 4.3 Compactificación de Poincaré y retrato de fases en el infinito.

Por medio de proyección estereográfica el plano \mathbb{R}^2 se compactifica en S^2 , como se explico en la sección 1.6.1. El teorema puede ser usado para el análisis de procesos de aprendizaje de una neurona no lineal con dos componentes de entrada suministrando el potencial E en la ecuación gradiente (4.3.5) puede completarse en el polo norte de la esfera de tal manera que se obtiene una función de clase $C^2(S^2, \mathbb{R})$ y no presenta puntos fijos adicionales no

hiperbólicos.

La función criterio dado por (4.3.6) tiene varios límites si una imagen inversa $\pi^{-1}(\mathbf{w})$ de un vector \mathbf{w} (donde π es la proyección estereográfica) converge al polo norte \mathbf{w} , es decir $|\mathbf{w}|$ converge al infinito. En consecuencia puede converger al infinito de tal modo que el producto escalar $\mathbf{x} \circ \mathbf{w}$ permanece como una constante arbitraria. No obstante, es posible modificar la función criterio de tal modo que en una cierta bola $B((0,0), r)$ el potencial sea modificado, y así será posible completarlo de manera conveniente tomando el radio tan grande como se necesaria.

Entonces como observamos en la sección anterior, para el análisis del proceso de aprendizaje considerar la conjugación topológica sobre un balón suficientemente grande. En efecto podemos modificar el potencial E usando la función escalar siguiente:

$$g(\mathbf{w}) = \begin{cases} e^{-(r-a)^4} & \text{para } r \geq a \\ 1 & \text{para } r \in [0, a) \end{cases} \quad (4.3.12)$$

donde $r := |\mathbf{w}|^2 = w_1^2 + w_2^2$ y a es una constante positiva, esta función g es de clase $C^2(\mathbb{R}^2, \mathbb{R})$. Haciendo la constante a suficientemente grande, definimos el siguiente mapeo E^*

$$E^*(\mathbf{w}) := g(\mathbf{w}) \cdot E(\mathbf{w}) \quad (4.3.13)$$

entonces la primera y la segunda derivada parcial de la función E^* son de la siguiente forma:

$$\begin{aligned} \frac{\partial E^*(\mathbf{w})}{\partial w_i} &= E(\mathbf{w}) \cdot \frac{\partial g(\mathbf{w})}{\partial w_i} + g(\mathbf{w}) \cdot \frac{\partial E(\mathbf{w})}{\partial w_i}; \\ \frac{\partial E^*(\mathbf{w})}{\partial w_i} &= E(\mathbf{w}) \cdot \frac{\partial g(\mathbf{w})}{\partial w_i} + g(\mathbf{w}) \cdot \sum_{n=1}^N [(f(\beta^{(n)}) - z^{(n)}) \cdot \frac{df}{d\beta^{(n)}} \cdot \frac{\partial \beta^{(n)}}{\partial w_i}] \\ \frac{\partial E^*(\mathbf{w})}{\partial w_i} &= E(\mathbf{w}) \cdot \frac{\partial g(\mathbf{w})}{\partial w_i} + g(\mathbf{w}) \cdot \sum_{n=1}^N [x_i^{(n)} \cdot \frac{df}{d\beta^{(n)}} \cdot (f(\beta^{(n)}) - z^{(n)})]; \end{aligned}$$

luego,

$$\frac{\partial^2 E^*(\mathbf{w})}{\partial w_j \partial w_i} = E(\mathbf{w}) \cdot \frac{\partial^2 g(\mathbf{w})}{\partial w_j \partial w_i} + \frac{\partial E(\mathbf{w})}{\partial w_j} \cdot \frac{\partial g(\mathbf{w})}{\partial w_i} + g(\mathbf{w}) \frac{\partial^2 E(\mathbf{w})}{\partial w_j \partial w_i} + \frac{\partial g(\mathbf{w})}{\partial w_j} \cdot \frac{\partial E(\mathbf{w})}{\partial w_i}$$

donde

$$\frac{\partial^2 E(\mathbf{w})}{\partial w_j \partial w_i} = \frac{\partial}{\partial x_j} \sum_{n=1}^N [x_i^{(n)} \cdot (f(\beta^{(n)}) - z^{(n)})],$$

esto es

$$\frac{\partial^2 E(\mathbf{w})}{\partial w_j \partial w_i} = \sum_{n=1}^N x_i^{(n)} x_j^{(n)} \cdot \left[\frac{\partial^2 f}{d(\beta^{(n)})^2} \cdot (f(\beta^{(n)}) - z^{(n)}) + \frac{df}{d\beta^{(n)}} \cdot \frac{df}{d\beta^{(n)}} \right]$$

La forma de estas derivadas y las suposiciones especificadas al comienzo de esta sección implican que si nos aproximamos al polo norte de la esfera S^2 , no sólo la primera derivada sino también la segunda derivada convergen a cero. Esto significa que si el potencial E^* se completa en el polo norte entonces un punto fijo no hiperbólico adicional podría aparecer. Sin embargo las soluciones de la ecuación

$$\dot{\mathbf{w}} = -\text{grad } E^*(\mathbf{w}) \quad (4.3.14)$$

son transversales el círculo $K((0,0), r = 2a)$ en dirección hacia su interior, es decir, el producto escalar $-\text{grad } E^*(\mathbf{w}) \circ \mathbf{w}$ es negativo para $|\mathbf{w}| = \sqrt{2a}$. En efecto tenemos

$$-\text{grad } E^*(\mathbf{w}) \cdot \mathbf{w} = \sum_{i=1}^2 [E(W) \cdot \frac{\partial g(\mathbf{w})}{\partial w_i} \cdot w_i + g(\mathbf{w}) \cdot \sum_{n=1}^N [x_i^{(n)} \cdot \frac{df}{d\beta^{(n)}} \cdot (f(\beta^{(n)}) - z^{(n)})] w_i]$$

puesto que para $r = 2a$,

$$\frac{\partial g(\mathbf{w})}{\partial w_i} := g(\mathbf{w}) = \begin{cases} -8 \cdot (r - a)^3 w_i \cdot e^{-(r-a)^4} & \text{para } r > a \\ 0 & \text{para } r \in [0, a) \end{cases} \quad (4.3.15)$$

obtenemos

$$-\text{grad } E^*(\mathbf{w}) = \sum_{i=1}^2 e^{-a^4} \cdot [-8a^3 \cdot w_i^2 \cdot E(\mathbf{w}) + \sum_{n=1}^N [x_i^{(n)} \cdot \frac{df}{d\beta^{(n)}} \cdot (f(\beta^{(n)}) - z^{(n)})] w_i]$$

Tanto el potencial E y la primera derivada están acotadas para a grande, el primer término en la suma es asintóticamente igual a $a^3 e^{-a^4}$, en tanto que en la segunda es igual a e^{-a^4} . Así para a suficiente grande, el valor del primer término es negativo y su valor absoluto es más grande que el valor absoluto de la segunda componente de la suma.

Via la proyección estereográfica el potencial $v(w) = -r^2$ en $|w| \Rightarrow \sqrt{3a}$ tiene su máximo en el polo norte, entonces el sistema dinámico

$$\dot{w} = -\text{grad } v(w)$$

tiene un punto de equilibrio hiperbólico en este polo. Finalmente para el potencial

$$\tilde{E}(w) = \begin{cases} E^*(w) & \text{si } w \leq 2\sqrt{a} \\ v(w) & \text{si } w > 3\sqrt{a} \end{cases}$$

hace el sistema (4.3.5) totalmente hiperbólico.

4.4. El movimiento errante y el fenómeno co-operativo en una red neuronal: un modelo de memoria asociativa

Un nuevo modelo para memoria asociativa con el tiempo discreto, en el cual el caos es tomado en cuenta y además el estado interior de las neuronas dependen de la historia pasada. La red Neuronal por sí misma puede buscar los mínimos de energía consecutivamente usando el movimiento errante y el fenómeno co-operativo. Se mostrará que el modelo descrito en esta sección tiene gran utilidad en problemas de memoria asociativa. Uno de los propósitos de estudiar las redes neuronales es entender la información procesada en el cerebro. Si conocemos el mecanismo de procesamiento de la información, nosotros podemos construir la red neuronal, que es útil para el procesar de información complicada.

El primer modelo de una neurona fue creado por Mac Culloch and Pitts en 1943. Este modelo es simple pero poderoso, específicamente, el rendimiento de las i neuronas o_i toma uno (el estado encendido) o cero (el estado no encendido) según si la suma de los pesos de sus entradas para otras neuronas están arriba o por abajo del umbral θ_i , tenemos

$$o_i(n+1) = F[u_i(n+1)] \quad , \quad o_i(n+1) = F \left[\sum_{j=1}^N w_{ij} o_j(n) - \theta_i \right] , n = 1, 2, 3, \dots \quad (4.4.1)$$

donde N es el número de neuronas en la red neuronal y w_{ij} son los pesos sinápticos que conectan las i neuronas con las j . El tiempo es tomado en pasos discretos. La función ganancia F es la función paso de Heaviside o la función sigmoideal.

La red neuronal es usada muchas veces en memorias asociativas y problemas de optimización (Traveling Salesman Problem T.S.P). Las redes neuronales Hopfield y Tank 1985 mutuamente conectadas tienen energía $E(n)$ definida.

$$E(n) = -\frac{1}{2} \sum_{i \neq j} \sum w_{ij} o_i(n) o_j(n) - \sum_i I_i o_i(n) + \sum_i \theta_i o_i(n) \quad (4.4.2)$$

donde I_i y θ_i son las entradas externas y el umbral respectivamente. Si la red tiene conexión simétrica, esto es $w_{ij} = w_{ji}$ con cero para conexión consigo misma $w_{ii} = 0$, y la función ganancia incrementa monótonicamente

4.4 El movimiento errante y el fenómeno co-operativo en una red neuronal: un modelo de memoria asociativa 79

la variación $\Delta E(n)$ de $E(n)$ debido a la variación $\Delta o_i(n)$ de $o_i(n)$ es

$$\Delta E(n) = - \left(\sum_{j \neq i} w_{ij} o_j(n) + I_i - \theta_i \right) \quad (4.4.3)$$

En su modelo neuronal Hopfield ha demostrado en la sección (4.1) que la energía total de la red decrece monótonicamente debido a $\Delta E(n)$ es siempre negativa. Si consideramos un mínimo de energía como una memoria asociativa, como en el caso de las rutas correctas en TSP, la red neuronal Hopfield puede buscar los mínimos. Al mismo tiempo, sin embargo, el modelo tiene la dificultad que el sistema no puede escapar de la mínima. Significa cuando empezamos el estado inicial que el sistema recupera solo una de las memorias, si algunas de las memorias estan guardadas en el sistema. Para quitar la dificultad algunos métodos son propuestos, por ejemplo, el Método estocástico ó el Método caótico. En el primer método, intengamos salir del minimo adicionando ruido Gaussiano. En el segundo, el estado del sistema sale del mínimo usando el movimiento errante, lo cual conduce naturalmente al caos. La salida de una neurona es determinado según si los osciladores caóticos acoplados están sincronizados entre si o no. En esta sección estudiamos el modelo que se ha obtenido modificando el Modelo de redes de Hopfield con una conexión negativa de retroalimentación; este modelo es descrito como un mapeo acoplado globalmente. El parámetro de bifurcación de su mapeo es determinado por las neuronas circundantes.

4.4.1. Modelo

Es bien conocido que después de activarse, la neurona tiene que esperar por un periodo, llamado periodo refractorio, antes que pueda activarse de nuevo. Por consiguiente esta claro que la historia anterior del estado interno de la neurona tiene gran influencia en la conducta de la neurona. Representamos el estado interno de la neurona i ($1 \leq i \leq N$) en la red neuronal por dos variables x_i^n y u_i^n en n pasos, los cuales son llamados la variable básica y la variable de la memoria. Usando la variable u_i^n podemos estimar si el estado del sistema se establece en uno de los mínimos o no. La variable básica x_i^n cambia caóticamente, lo cual se usa para escapar del mínimo spurious haciendo uso del Movimiento Caótico errante. La variable básica x_i^n se genera usando el mapeo $f(x, a) : x_i^n = f(x_i^{n-1}, a_i^n)$, donde a_i^n es un parámetro de bifurcación en el paso n , emplearemos el mapeo logístico siguiente:

$$\begin{aligned} x_i^n &= f(x_i^{n-1}, a_i^n) \\ &= 2a_i^n(1 + x_i^{n-1})(1 - x_i^{n-1}) - 1 \end{aligned} \quad (4.4.4)$$

La serie de tiempo generada por el mapa logístico f depende mucho del parámetro de bifurcación a_i^n , entonces es importante conocer como elegir a_i^n . La memoria variable u_i^n es la cantidad que describe la historia anterior, para la información de entrada de la neurona i . Luego la salida o_i^n de la neurona i definido en términos de u_i^n y x_i^n es:

$$o_i^n = \text{sgn}(u_i^n)|x_i^n| \quad (4.4.5)$$

La variable de memoria u_i^n se define de la siguiente manera:

$$\begin{aligned} u_i^n &= \frac{u_i^{n-1}}{\alpha} + \sum_j w_{ij}o_j^n + I_i - \theta_i \\ u_i^n &= \sum_{m=0}^{n-1} \alpha^{-m} \left(\sum_j w_{ij}o_j^{n-m-1} + I_i - \theta_i \right) \end{aligned} \quad (4.4.6)$$

donde α es el parámetro de decaimiento. Como es claro en la ecuación (4.4.6), u_i^n involucra la memoria pasada. Si α es suficientemente larga y $|x_i^n| = 1$, podemos no tener en cuenta la memoria pasada y nuestro modelo coincide con el modelo convencional propuesto por McCulloch - Pitt como Ecuación (4.4.1). El parámetro de la bifurcación a_i^n del mapeo $f(x, a)$ definido en términos de la variable de la memoria variable u_i^n esta dada por:

$$a_i^n = A_1 + \frac{A_2}{1 + \exp(-A_3 u_i^n)} \quad (0 \leq a_i^n) \leq 1 \quad (4.4.7)$$

donde A_1 , A_2 , y A_3 son parámetros. Si el sistema entra en uno de los mínimos de energía, la salida o_i^n permanece constante y así u_i^n incrementara o decrecerá monótonicamente, lo que es claro en la ecuación (4.4.6). Por eso introducimos el umbral θ con respecto a $|u_i^n|$ para escapar del mínimo. Si $|u_i^n| \geq \theta$ la magnitud y el signo de u_i^n en n pasos se actualiza de acuerdo a:

$$u_i^n = -\text{sgn}(u_i^n)\beta(\theta, N) \quad (4.4.8)$$

donde $\beta(\theta, N)$ es un parámetro. Esta actualización significa el cambio de signo de la salida o_i^n en la ecuación (4.4.5). En la ecuación (4.4.6) este cambio de signo es equivalente a reemplazar w_{ij} por $-w_{ij}$. Esto puede implicar que parte de la función de energía en la ecuación (4.4.2) se modifique. Así el sistema puede escapar del mínimo, entre mas grande sea N el tamaño de la red mas grande debe ser β , por eso la magnitud del aumento β debe ser una función del tamaño de la red Neuronal N y del Umbral θ .

4.4 El movimiento errante y el fenómeno co-operativo en una red neuronal: un modelo de memoria asociativa 81

Como se discutió arriba, la red neuronal trabaja como sigue. Inicialmente los valores de u_i^n , o_i^n y x_i^n se dan aleatoriamente, la variable de la memoria u_i^{n+1} en el siguiente paso se calcula con la ecuación (4.4.6) usando o_i^n y u_i^n . Sustituyéndolas en la ecuación (4.4.7) obtenemos el parámetro de bifurcación a_i^{n+1} . La variable básica x_i^{n+1} se determina con la ecuación (4.4.4) con el parámetro de bifurcación a_i^{n+1} . De la ecuación (4.4.5) obtenemos la salida de una neurona. Si el valor absoluto de u_i^n es mas grande que el umbral θ , la magnitud y el signo de u_i^n se actualiza de acuerdo a la ecuación (4.4.8). Repitiendo este proceso, podemos calcular la dinámica de la red neuronal.

4.4.2. Memoria asociativa

Si la red neuronal tiene 4×4 neuronas y tres patrones, lo que se muestra se muestran en la Figura 4.4.2. El peso Sináptico w_{ij} es determinado de acuerdo a la regla Hebbian siguiente,

$$w_{ij} = \sum_s (2\xi_i^s - 1)(2\xi_j^s - 1) \quad (4.4.9)$$

con $w_{ii} = 0$, donde ξ_i denota los s-vectores patrones guardados y ξ_i^s toma 1 y 0 para blanco y negro, respectivamente. Para simplificar ponemos $I_i = 0$, $\theta_i = 0$ y $A_3 = 2,5$. Nuestro modelo es diferente del Hopfield, sin embargo usaremos su definición de energía E de la ecuación (4.4.2). Para $I_i = 0$ y $\theta_i = 0$, la energía E^n en el paso n , es

$$E^n = -\frac{1}{2} \sum_{i \neq j} w_{ij} o_i^n o_j^n. \quad (4.4.10)$$

Puesto que la salida de nuestra red es analógica, nosotros ponemos $o_i^n = 1$ para $o_i^n > 0$ y $o_i^n = 0$ para $o_i^n < 0$, en el cálculo de la energía (4.4.10) y podemos pintar de cada neurona en la figura 4.4.2 de negro o blanco según si o_i^n es positiva o negativa respectivamente.

La red toma la mínima energía -20 cuando la red recupera uno de los patrones guardados e invierte patrones. Si la red neuronal recupera un patrón falso, se toman el más alto nivel de energía.

Podemos ajustar el rango de a_i^n cambiando A_1 y A_2 , tomamos inicialmente $\alpha = 1$, $\theta = 15$, $\beta = 3$ para $A_1 = 0,9539$ y $A_2 = 0,0032$ ($0,9539 \leq a_i^n \leq 0,9571$), el cual corresponde a una región de caos intermitente, localizada en una ventana de periodo 3 en el diagrama de bifurcación

del mapeo logístico. Por ejemplo el sistema puede encontrar la ruta mas corta en la solución del problema T.S.P. (problema del comerciante viajero)

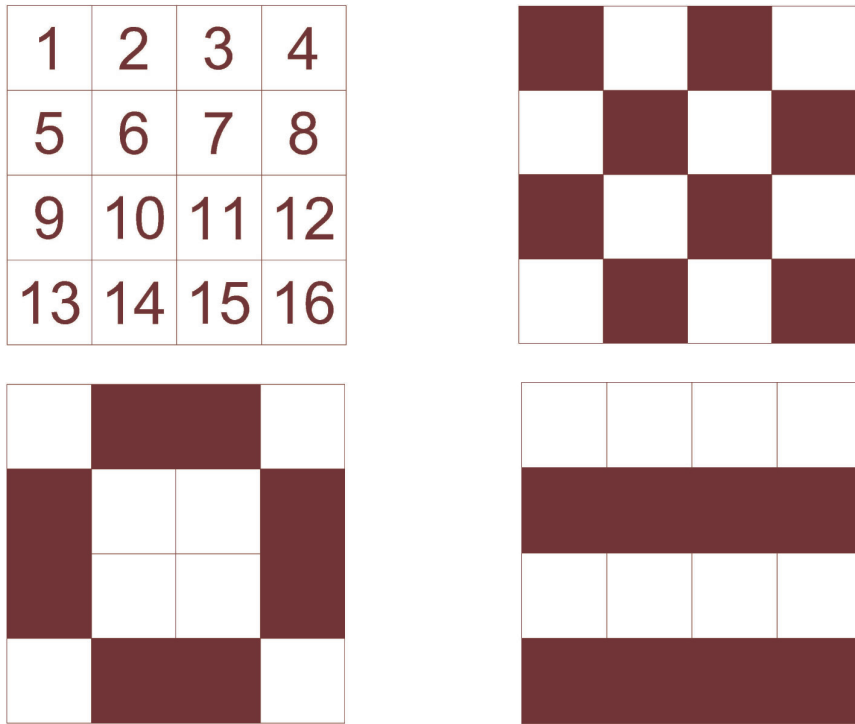


Figura 4.4.2 Ejemplo de la dinámica del modelo de Hopfield con memoria asociativa

Conclusiones

- 1) Se conocieron, analizaron y simularon las pequeñas oscilaciones o autooscilaciones dinámicamente próximas a los estados de reposo; las cuales corresponden a las bifurcaciones de Hopf. Fenómeno detectado en el modelo de Hodgkin-Huxley y posteriormente en el modelo de FitzHugh-Nagumo. Este hecho es relevante para conocer el comportamiento de la dinámica neuronal.
- 2) En coordenadas apropiadas, en formas normales algebraicamente convenientes y mediante reducción a la variedad central, se descubre también grandes oscilaciones; las cuales se explican por las conexiones heteroclínicas entre equilibrios, dicho de otra manera, mediante soluciones solitónicas. Esto explica la inestabilidad estructural al modelo de FitzHugh-Nagumo.
- 3) Estudiamos una primera aproximación entre redes neuronales artificiales, en sus paradigmas de Hopfield y propagación hacia atrás, con la representación no lineal en contextos compactos de la actividad neuronal; esto a través de sistemas dinámicos del tipo gradiente y el concepto de memoria asociativa.

Bibliografía

- [MU,MA] P. Muruganandam y M. Lasklmanan [1997]; Bifurcación Analysis of the Travelling Waveform of FitzHugh-Nagumo nerve conduction model equation; Chaos, Vol 7, N° 3, 476-487.
- [FH] R. FitzHugh [1961]; Impulses and Phisiological states in theoretical models of nerve membrane, Biophysic journal 1, 445.
- [GU,HO] J. Gukhenheimer and P.J Holmes [1983]; Nonlinear oscillations, Dynamical Systems, and Bifurcations of vector fields, Applied Mathematical Science Series (springer, New York) vol. 42.
- [BiI] Andrzej Bielecki [2001]; Dynamical properties of learning process of weakly nonlinear and nonlinear neurons, nonlinear Analysis: Real world Applications 2, 249-258.
- [BiII] Andrzej Bielecki, D. Jablonski, M. Kedzierski [2004]; Properties and Applications of weakly nonlinear Neurons, Elsevier Journal of Computational and Applied mathematics, 93-106.
- [MA,SH] Makoto Makishima y Toshihiro Shimizu [1998]; Wandering Motion and Co-operative Phenomena in Chaotic Neural Network, International Journal of Bifurcations and Chaos, vol N° 5, 891-898.
- [Cr] Alejandro Cruz Sandoval [2003]; Tesis de Maestria de Instituto Politecnico Nacional de Mexico, Ciudad de Mexico D.F.
- [AT,JO,SA] M. Atencia, G Joya y F. Sandoval [2004]; Optimización con sistemas Neuronales de Hopfield. Aplicación a la identificación

- paramétrica de sistemas Dinámicos; en Optimización Inteligente, Universidad de Malaga, España.
- [CH,DE,TE] S. N. Chow, B Deng and D. Terman [1990]; The Bifurcation of Homoclinic and Periódic Orbits From Two Heteroclinic orbits, Siam Math Journal, vol 21 N° 1, 205-220.
- [BA,CARR,ON] Carolina Barriga M., Humberto Carrillo Calvet y Fernando Ongay L. [2003]; El Modelo de Fitzhuhg-Nagumo para el Potencial Eléctrico de una Neurona, Aportes matemáticos series comunicaciones 32, 31-32, Mexico D.F.
- [MO,ST] Morris F. Hirsch and Stephen Smale [1998]; Differential equations, Dynamical Systems, and linear Algebra, Academic Press, New York.
- [PE] Lawrence Perko [1996]; Differential equations and Dynamical Systems, Springer-Verlag.
- [MON,LON,POL] Mauro Montealegre C., Gustavo Londoño B. y Luis Arturo Polania Q.[2001]; Los Fundamentos de los Sistemas Dinámicos, Universidad Surcolombiana.
- [CO] Conley[1975]; Traveling wave solutions of nonlinear differential equations, in structural stability, the theory of catastrophes and applications in Scienes; A Dold and B. Eckmann; eds, Lecture notes in mathematics, springer-verlag, Berlin.
- [Ki,Che] David Kincaid y Ward Cheney, Análisis numérico las matemáticas del cálculo científico, [1994]; Addison wesley.
- [Ho] Hopfield, J.J en “Neurons with graded response have collective computational properties like those of two-state neurons”, Proceedings of the National Academy of Sciences of de U.S.A 81, 3088-3092.1984
- [PA,ME] J. Palis and W. de Melo; Geometric Theory of dynamica systems; Springer [1998].
- [Si] Simon Haykin; Neuronal Networks; [1994] Macmillan.