

00382

2  
REJ



**UNIVERSIDAD NACIONAL  
AUTONOMA DE MEXICO**

**FACULTAD DE CIENCIAS  
División de Estudios de Posgrado**

**DINAMICAS EXTENDIDAS EN REDES  
BIOLOGICAS**

**FALLA DE ORIGEN**

**T E S I S**

Que para obtener el grado Académico de

**DOCTORA EN CIENCIAS**

**(FISICA)**

**p r e s e n t a**

**ANGELICA LEONOR GELOVER SANTIAGO**

**Directores de Tesis: Dr. Jorge Simón Lomnitz Adler  
Dr. Gustavo Carlos Martínez Mekler**



Universidad Nacional  
Autónoma de México



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

2030 AN 2121 AY22  
A3761870 AJ 30 81A2

**Dedico este trabajo a mis padres**

**Y a mis cinco hermanos**

**Por todo el amor y el apoyo que me han brindado  
siempre**

**A la memoria de Jorge Lomnitz Adler**

Quiero expresar mi más sincero agradecimiento a todas aquellas personas que de una u otra forma me han brindado su ayuda durante todos estos años. En primer lugar, gracias a mis asesores: al Dr. Gustavo Martínez Mekler y al desaparecido Dr. Jorge Lomnitz Adler, por dirigir este trabajo, por sus enseñanzas, porque siempre me animaron a seguir adelante y sobre todo por su paciencia. Al Dr. Germinial Cocho Gil porque aún sin ser mi asesor me brindo siempre todo su apoyo. A los demás integrantes del Jurado: Dr. José Luis del Río, Dr. Francisco Lara, Dr. François Leyvraz, Dr. Rafael Pérez y Dr. Xavier Soberón, gracias por leer este trabajo y porque con sus valiosos comentarios contribuyeron de manera muy importante al mejoramiento del mismo. Al Prof. Ricardo Lima porque siempre he contado con sus consejos y sobre todo con su amistad. Al Dr. José Luis Rius por brindarme la oportunidad de trabajar con él en un programa tan importante, el PACIME. Al Dr. Rafael Pérez por permitirme ocupar su espacio. A los doctores Franco Bagnoli, Raúl Herrera, Pedro Miramontes, Raúl Rechtman, Víctor Romero, Alejandro Salcido, Rubén Santamaría, Manuel Torres y Gerardo Vázquez por la sincera amistad con que me han distinguido. A Verónica por la gran amistad que nos une. A Paty porque siempre he podido contar con ella. A mis amigos de tanto tiempo Chelo, Alejandro, José Antonio, Mario, Pedro y Oliverio por las cosas que hemos compartido juntos. A Jeremy Ahouse por brindarme su amistad aún estando tan lejos. A mis amigos del CONACYT: Ale, Anita, Avilio, Bernardo, Jesús y José Ramón por todo lo que han hecho por mi. A mis padres por el amor sin fronteras que me han dado a lo largo de toda mi vida y por la libertad que me dieron para decidir mi futuro. A mis hermanos Silvia, Guille, Jaime, Adriana e Israel, por el ilimitado cariño que nos ha unido siempre. A *mis pequeños* José Abel, Carlos Alberto, Erika Mariana y Angel Eduardo porque sus alegrías también son mías.

A los que he hecho patente mi agradecimiento, y a los que ahora pude olvidar, sinceramente: ¡Gracias!.

**Finalmente quiero agradecer de manera muy especial la hospitalidad que me brindó el IFUNAM durante tantos años, la ayuda de las secretarias, del personal de cómputo y de los dibujantes. Así mismo agradezco a la DGAPA-UNAM por la beca con la que me distinguió.**

# Contenido

<b>1</b>	<b>Introducción</b>	<b>1</b>
<b>2</b>	<b>Bases teóricas</b>	<b>7</b>
2.1	Sistemas extendidos . . . . .	7
2.1.1	Redes de mapeos acoplados . . . . .	8
2.2	Dinámica de poblaciones . . . . .	10
<b>3</b>	<b>Fenomenología</b>	<b>15</b>
3.1	ADN y ARN, las moléculas de la vida . . . . .	15
3.2	Síntesis de proteínas y el código genético . . . . .	18
3.2.1	Dinámica ribosomal . . . . .	22
3.3	Fisicoquímica de los ácidos nucleicos . . . . .	26
<b>4</b>	<b>El sistema inmune y el virus del SIDA</b>	<b>33</b>
4.1	El sistema inmune . . . . .	33
4.2	Virus y retrovirus . . . . .	37
4.3	El retrovirus del SIDA . . . . .	38
4.3.1	La visión tradicional . . . . .	38
4.3.2	La visión actual . . . . .	41
<b>5</b>	<b>El modelo</b>	<b>45</b>
5.1	El espacio de secuencias . . . . .	46
5.2	Construcción del modelo . . . . .	48
<b>6</b>	<b>Dinámica del modelo</b>	<b>55</b>
6.1	Secuencias que codifican un solo gen . . . . .	55
6.1.1	Acoplamiento de corto alcance . . . . .	55

6.1.2	Acoplamiento de largo alcance . . . . .	62
6.2	Extensión a $K$ genes . . . . .	65
<b>7</b>	<b>Bifurcaciones</b> . . . . .	<b>73</b>
7.1	Mapeos unidimensionales . . . . .	73
7.2	Mapeos multidimensionales . . . . .	78
7.3	Cascada de bifurcaciones . . . . .	80
<b>8</b>	<b>Modelos tipo campo medio</b> . . . . .	<b>87</b>
8.1	Descripción de la dinámica . . . . .	87
8.1.1	$\alpha \neq 0$ . . . . .	88
8.1.2	$\beta \neq 0$ . . . . .	93
<b>9</b>	<b>Sincronización</b> . . . . .	<b>101</b>
9.1	Antecedentes recientes . . . . .	101
9.2	Un caso de baja dimensión . . . . .	103
9.3	Alta dimensión . . . . .	104
9.4	Modelo de prueba . . . . .	105
9.5	Problema general de alta a baja dimensión . . . . .	114
<b>10</b>	<b>Discusión y Conclusiones</b> . . . . .	<b>117</b>
<b>A</b>	<b>Cálculo del espectro de Lyapunov</b> . . . . .	<b>131</b>

# Capítulo 1

## INTRODUCCION

El principal objetivo de este trabajo es el desarrollo de formalismos aplicables al estudio de redes biológicas que permitan un mejor entendimiento de su comportamiento dinámico. Como ejemplos de estas redes podemos mencionar entre otras: redes neuronales, poblacionales, ecológicas, metabólicas e inmunológicas.

Los sistemas biológicos presentan una gran variedad de escenarios dinámicos. Variaciones de los parámetros del sistema pueden conducir a comportamientos que van desde estados estacionarios, evoluciones cíclicas, hasta dinámicas caóticas. El poder descifrar los orígenes de los comportamientos dinámicos de estos sistemas es uno de los retos científicos de nuestro tiempo. Como ejemplo de estos escenarios podemos mencionar, [1], [2], [3], los fenómenos de sincronización parcial observados en las respuestas oscilatorias de la corteza visual de los gatos, la coherencia oscilatoria de la actividad neuronal que da lugar a picos pronunciados en la transmisión nerviosa durante ataques epilépticos, la variación temporal en el acoplamiento efectivo de neuronas y su efecto en los cambios del grado de sincronización de oscilaciones. En el ámbito de la ecología se nos presenta la evolución de un ecosistema el cual, visto como un sistema dinámico disipativo extendido en el espacio, plantea el estudio de la dinámica de poblaciones de diversas especies. Problemas interesantes son la relación entre la complejidad del sistema y su estabilidad. Recientemente se ha planteado que los comportamientos asintóticos corresponden a procesos dinámicos cam-



biantes en lugar de comportamientos estacionarios de punto fijo [4], [5]. El proceso de evolución molecular se ha visualizado como un proceso dinámico [6], el cual puede describirse mediante ecuaciones cinéticas de replicación de nucleótidos. La dinámica resultante de este tipo de ecuaciones es consistente con los resultados de estudios experimentales de replicación *in vitro* de algunos virus. Otro ejemplo lo constituye la coevolución de especies en la que diferentes subsistemas tienen sus propios criterios de evolución, los cuales usualmente son conflictivos ya que la evolución o adaptación de un organismo modifica los criterios de los demás [7]. Podemos también considerar el proceso de diferenciación celular, en este problema las interacciones celulares juegan un papel muy importante, independientemente del control debido a la activación o desactivación de algunos genes. El sistema inmune también presenta una riqueza dinámica considerable, durante la respuesta inmune los anticuerpos interactúan a través de una complicada red y se ha observado que algunos de ellos exhiben oscilaciones que pueden ser caóticas en el tiempo [8].

Los sistemas descritos anteriormente presentan la característica común de ser sistemas extendidos, es decir constan de un elevado número de componentes interactuantes definidos en un espacio determinado, al estudiar su evolución espacio-temporal nos enfrentamos a un problema de muchos grados de libertad fuertemente acoplados que en general presentan una dinámica no-lineal. Tradicionalmente los sistemas que presentan estas características han sido estudiados en términos de ecuaciones diferenciales parciales no lineales. Recientemente, en parte debido a la disponibilidad de computadoras cada vez más poderosas, se utilizan modelos discretos como *autómatas celulares*, o parcialmente discretos como es el caso de las *redes de mapeos acoplados*. El estudio de este tipo de modelos es intrínsecamente interesante debido a la gran variedad de comportamientos que pueden exhibir y que van desde lo muy simple hasta lo muy complicado. Introducidos en la literatura hace escasamente diez años como una posible herramienta en el estudio de la dinámica de sistemas extendidos [9], [10], [11], las redes de mapeos acoplados son actualmente utilizadas como modelos para muchos problemas en muy diversas áreas como son física, química, biología, etc. [12], [13], [14].

Para cumplir con el objetivo planteado en este trabajo nos vamos a centrar en un problema particular. El problema que estudiaremos es el de la evolución de secuencias genéticas, en particular la del *Virus de Inmunodeficiencia Humana* (VIH), causante del *Síndrome de Inmunodeficiencia Adquirida* (SIDA). Sin embargo en los capítulos finales se presentarán resultados que son de relevancia en el contexto general del tipo de los problemas anteriormente descritos.

El estudio del VIH comenzó alrededor de 1983, año en que se tuvieron las primeras indicaciones de que es este virus el causante del SIDA [15]. Hoy en día la estructura molecular del virus es bien conocida sin embargo, los mecanismos por medio de los cuales la infección por VIH evoluciona hacia la etapa del síndrome son aún materia de estudio. El entendimiento de dichos mecanismos ofrece la posibilidad de disminuir los efectos y eventualmente encontrar un tratamiento terapéutico contra esta enfermedad.

Para el estudio de la evolución de la secuencia genética del VIH se han propuesto varios formalismos entre los cuales podemos mencionar los trabajos de Schuster [6] y Eigen [16] que hacen un tratamiento en términos de ecuaciones diferenciales parciales no lineales acopladas, Kauffman [7] quien aborda el problema desde el punto de vista de redes booleanas, etc. En nuestro caso el estudio del virus del SIDA surge en el marco del trabajo de Cocho y Martínez-Mekler [17] quienes proponen un formalismo de redes de mapeos acoplados.

Desde el punto de vista biológico, estamos interesados principalmente en relacionar el concepto de adecuación evolutiva con constricciones fisicoquímicas de los ácidos nucleicos y con los procesos moleculares que se presentan a nivel del ribosoma. En nuestro planteamiento el problema se desarrolla en un espacio discreto que corresponde a las posibles composiciones químicas de las secuencias genéticas y está sujeto a una dinámica no lineal (proceso de reproducción celular) en la cual hay un acoplamiento espacial que corresponde a las mutaciones del material genético.

La visión sobre el SIDA ha cambiado notable y muy rápidamente en los últimos meses. El descubrimiento de que uno de los tratamientos más prometedores contra el SIDA (a base del complejo AZT o sidobudina) es aparentemente inefectivo para prevenir la llegada de la etapa del síndrome, así como la presencia del virus en una gran variedad de células del organismo han revelado que el problema es más complejo de lo que se pensaba [15], [18]. El mecanismo preciso mediante el cual el virus VIH causa SIDA es aún desconocido y actualmente es materia de debate. Se han propuesto una serie de mecanismos complementarios [15], [19] con la esperanza de que eventualmente pueda llegarse a un entendimiento integral del problema. En este trabajo nos proponemos participar en este esfuerzo colectivo.

El trabajo que se desarrolla en esta tesis resalta la importancia del proceso de síntesis de proteínas a nivel del ribosoma. Cabe mencionar que la construcción del modelo que se propone toma en cuenta las propiedades fisicoquímicas de los ácidos nucleicos, con lo cual se establece una correlación con criterios de adecuación, así mismo se ponen de manifiesto algunas características del código genético. El trabajo está organizado de la siguiente manera:

En el capítulo 2 se da una introducción al formalismo de redes de mapeos acoplados como un modelo para el estudio de sistemas extendidos. A fin de dar una visión general del tipo de modelos que se utilizan en el estudio de dinámica de poblaciones, en este mismo capítulo se discuten algunos modelos sencillos. Como se verá, el formalismo de redes de mapeos acoplados resulta ser una herramienta adecuada para el estudio de este tipo de problemas.

El estudio de la evolución de la secuencia genética del virus del SIDA requiere de una comprensión más o menos detallada de una gran variedad de fenómenos que tienen lugar a nivel molecular. El capítulo 3 tiene como objetivo principal establecer un lenguaje adecuado para introducir posteriormente el modelo matemático. En él se describe la estructura de los ácidos nucleicos y se discuten algunos aspectos termodinámicos y estructurales de dichos ácidos, los cuales servirán como base para la construcción del modelo.

El capítulo 4 contiene una breve descripción de los principales componentes del sistema inmune y sus funciones. Se incluye una discusión de los aspectos más importantes del proceso mediante el cual un virus penetra en la célula hospedera y consigue que su material genético sea sintetizado por ésta. Finalmente se describen por una parte la estructura del virus del SIDA y por la otra las características más importantes de la enfermedad desde la infección hasta el desarrollo del síndrome.

En el capítulo 5 se describe detalladamente la construcción del modelo, el cual está inspirado en los modelos que se utilizan en el estudio de dinámica de poblaciones e incorpora los aspectos discutidos en los capítulos 3 y 4 sobre los mecanismos a nivel molecular, las propiedades termodinámicas de los ácidos nucleicos y las principales características del VIH.

En el capítulo 6 se hace un análisis de las simulaciones numéricas y su interpretación en términos del problema biológico, con especial atención a las predicciones del modelo sobre la existencia de una anticorrelación entre la composición de las secuencias genéticas y la variabilidad que presentan. En este análisis se consideran secuencias que codifican un solo gen así como secuencias que codifican varios genes o en las cuales se pueden identificar dominios con diferentes características evolutivas.

En el capítulo 7, teniendo en mente la variedad dinámica de las redes biológicas que mencionamos anteriormente, se consideran parámetros fuera del rango de relevancia para el problema del SIDA. La consideración de regímenes caóticos cobra especial relevancia en el contexto de discusiones recientes sobre la relación entre dinámica caótica y diversidad biológica [4], [5]. Esto nos lleva al estudio de bifurcaciones, que es el contenido de este capítulo.

Ante la complicación de la dinámica resultante de las ecuaciones planteadas, es conveniente buscar métodos aproximados que nos permitan entender algunos aspectos de dicha dinámica. En el capítulo 8 se introducen versiones de tipo campo medio. La dinámica de estos

modelos presenta características semejantes a la dinámica exhibida por algunos sistemas unidimensionales. Una posible explicación de lo anterior puede provenir de efectos de sincronización y formación de cúmulos. En vista de ello en el capítulo 9 se estudia esta posibilidad y se plantean algunas ideas sobre reducción dimensional.

Finalmente el capítulo 10 contiene una discusión de los principales resultados y las conclusiones del trabajo, tanto desde el punto de vista del problema biológico como desde el punto de vista de la teoría de los sistemas dinámicos.

## Capítulo 2

# BASES TEORICAS

### 2.1 SISTEMAS EXTENDIDOS

Cuando los grados de libertad de un sistema se encuentran distribuidos espacialmente, se dice que se trata de un *sistema extendido* ó *espacialmente extendido*. En contraposición con estos podemos mencionar otros sistemas, que pueden describirse por ejemplo en términos de ecuaciones diferenciales ordinarias, los cuales no contienen la noción de espacio. Cabe aclarar que el espacio al cual nos referimos no necesariamente es el espacio físico sino que puede tratarse de un espacio de composiciones químicas, un espacio de características fenotípicas, etc. Ejemplos de formalismos matemáticos para el estudio dinámico de sistemas extendidos son ecuaciones diferenciales parciales, autómatas celulares, redes de mapeos acoplados, etc., en los que el espacio juega un papel muy importante.

La naturaleza de dinámicas caóticas ha atraído la atención durante mucho tiempo y en años recientes el estudio de caos espacio-temporal. Por caos espacio-temporal entendemos una dinámica caótica en un sistema extendido. Como ejemplos de problemas en donde se observa la aparición de caos espacio-temporal podemos mencionar convección en cristales líquidos, convección de Bénard, turbulencia en fluidos, algunos sistemas biológicos en donde el acoplamiento de osciladores independientes puede dar lugar a la aparición de estructuras auto-organizadas,

Tabla 2.1: Clasificación de formalismos para el estudio de sistemas extendidos

Sistema	Espacio	Tiempo	Estados
Ecuaciones diferenciales parciales	continuo	continuo	continuos
Automatas celulares	discreto	discreto	discretos
Redes de mapeos acoplados	discreto	discreto	continuos
Ecs. diferenciales ordinarias acopladas	discreto	continuo	continuos

ritmos, etc.

Los modelos de sistemas extendidos pueden clasificarse de acuerdo a si las variables involucradas *tiempo*, *espacio* y *variables de estado* son discretas o continuas según se resume en la Tabla 2.1. En el caso particular de las redes de mapeos acoplados, las variables de tiempo y espacio son discretas mientras que las variables de estado son continuas, por lo que en estructura estas redes son más complicadas que los autómatas celulares, cuyos estados son discretos y resultan ser más simples que las ecuaciones diferenciales parciales en donde tanto el tiempo como el espacio son continuos. El formalismo utilizado en cada caso dependerá del problema particular de que se trate. En este trabajo estamos interesados en un formalismo particular para el estudio de sistemas extendidos, nos referimos a las redes de mapeos acoplados.

### 2.1.1 REDES DE MAPEOS ACOPLADOS

Las redes de mapeos acoplados fueron introducidos en la literatura, como modelos para el estudio de sistemas extendidos por Kaneko [9] y Waller y Kapral [10]. Aunque su fenomenología no está totalmente entendida, los modelos de redes de mapeos acoplados son ampliamente usados no solo en el estudio del surgimiento de patrones y caos espacio-temporal, sino también en problemas de física, química, biología, etc. [12], [13], [14].

Como se mencionó anteriormente, las redes de mapeos acoplados son sistemas dinámicos en donde las variables de tiempo y espacio son discretas mientras que las variables de estado son continuas. Se definen sobre una red regular en donde el estado de una variable  $x$  definida en el sitio  $i$  de la red evoluciona de acuerdo a un mapeo que actúa sobre una vecindad del sitio. Más concretamente la evolución de las redes de mapeos acoplados se da en dos pasos. En el primero la variable en cada sitio de la red evoluciona de acuerdo a un mapeo local, y en el segundo paso el resultado de dicho mapeo se acopla en una cierta vecindad del sitio dando como resultado un nuevo conjunto de variables definidas sobre la red. El ejemplo más sencillo se tiene en el caso unidimensional. Denotando por  $x_i(t)$  el valor de la variable de estado al tiempo  $t$ , la evolución no lineal está dada por el mapeo  $f(x)$ , el cual define una variable intermedia  $y_i(t)$  dada por:

$$y_i(t) = f(x_i(t)) \quad (2.1)$$

El valor  $x_i(t+1)$  está determinado por la función de acoplamiento  $G\{y_j(t)\}$  donde  $j$  pertenece a alguna vecindad del sitio  $i$ . La actualización de los sitios se hace de manera simultánea y en general se toman condiciones a la frontera periódicas y condiciones iniciales al azar, sin embargo esto depende del problema particular que se estudie. El ejemplo más estudiado en la literatura es aquel en donde la función  $G\{y_j(t)\}$  corresponde a un acoplamiento difusivo. Se llama difusivo a este acoplamiento debido a su semejanza con el operador de difusión discretizado. Este acoplamiento está dado por:

$$G\{y_j(t)\} = \frac{\epsilon}{2}[y_{i+1}(t) - 2y_i(t) + y_{i-1}(t)] \quad (2.2)$$

donde  $\epsilon$  es el parámetro de acoplamiento. En tal caso  $x_i(t+1)$  estará dada por:

$$x_i(t+1) = (1 - \epsilon)f(x_i(t)) + \frac{\epsilon}{2}[f(x_{i+1}(t)) - f(x_{i-1}(t))] \quad (2.3)$$

El mapeo definido por la ecuación (2.3), es un ejemplo en donde el acoplamiento es de tipo local debido a que el valor de la variable



en un sitio depende solamente de sus vecinos más cercanos. El caso opuesto se tiene cuando todos los sitios de la red determinan el valor de la variable en cada uno de los sitios, se dice entonces que se tiene un acoplamiento global. Un ejemplo de este tipo de acoplamiento está dado por:

$$x_i(t+1) = (1 - \epsilon)f(x_i(t)) + \frac{\epsilon}{N} \sum_{j=1}^N f(x_j(t)) \quad (2.4)$$

En la literatura se utiliza el término mapeos globalmente acoplados para referirse a este tipo de sistemas [20] y se introducen como una extensión del tipo teoría de campo medio de los mapeos localmente acoplados.

Las redes de mapeos acoplados se han utilizado en el estudio de sistemas muy diversos. Por ejemplo como modelo de la dinámica del ordenamiento de fases en sistemas con fases termodinámicamente inestables [21], en el estudio de la dinámica de reacciones químicas [22], convección en fluidos, crecimiento de cristales [23], etc.

En este trabajo se estudia un modelo de redes de mapeos acoplados como una herramienta en el estudio de la evolución de secuencias genéticas, con miras a aplicarse en general al estudio de redes biológicas tales como redes neuronales, poblacionales, ecológicas, metabólicas, inmunológicas, etc.

## 2.2 DINAMICA DE POBLACIONES

El presente trabajo se desarrolla en el marco general del estudio de dinámica de poblaciones, en donde el problema principal es describir los mecanismos básicos que regulan el crecimiento de las poblaciones naturales. Empezaremos pues por describir brevemente el tipo de problemas que se estudian en dinámica de poblaciones.

El sistema más simple que se puede estudiar en dinámica de poblaciones corresponde a una población en la cual las generaciones no se

traslapan. Muchas poblaciones naturales son de este tipo, por ejemplo algunas poblaciones de peces, artrópodos, insectos de zonas templadas, etc. [24], [25]. En este caso la información con que usualmente se cuenta es la población promedio o la población total en una generación. Desde el punto de vista teórico, se busca establecer una relación entre la magnitud de la población en una generación, con la magnitud de la población de la generación anterior. Si denotamos por  $x_n$  la magnitud de la población en la  $n$ -ésima generación, esta relación puede expresarse en forma general como:

$$x_{n+1} = F_{\mu}(x_n) \quad (2.5)$$

donde la función  $F_{\mu}(x_n)$  es no lineal y depende del parámetro  $\mu$ . La ecuación (2.5) resulta ser entonces una ecuación en diferencias no lineal. Entre las más usadas se encuentra el mapeo logístico:

$$x_{n+1} = \mu x_n(1 - x_n) \quad (2.6)$$

Como se discutirá en la sección 6.3.1, cuando se varía el parámetro  $\mu$ , se obtiene una gran variedad de comportamientos que van de lo regular o periódico a lo irregular o caótico.

En sistemas de dos o más poblaciones, en cada una de las cuales las generaciones no se traslapan, el análisis se hace en términos de ecuaciones en diferencias acopladas. En el caso de dos poblaciones se han estudiado poblaciones del tipo presa-depredador [26], dos poblaciones que compiten [24], etc. Para una población presa-depredador se utilizan ecuaciones del tipo Lotka-Volterra [27]:

$$\begin{aligned} x_{n+1} &= \mu x_n(1 - x_n - y_n) \\ y_{n+1} &= \beta x_n y_n \end{aligned} \quad (2.7)$$

en donde  $x_n$  y  $y_n$  denotan a las poblaciones de presa y depredador en la  $n$ -ésima generación respectivamente. Este modelo exhibe patrones de comportamiento similares a los observados en el caso de una sola especie. Las ecuaciones (2.7) corresponden a dos mapeos acoplados como los descritos en la sección anterior. Un esquema similar puede aplicarse al problema general de varias especies. Un ejemplo concreto se verá en la sección 5.2.

Otro problema de gran interés es el estudio de una población en donde los individuos presentan características fenotípicas diversas. El *fenotipo* de un individuo determinado es el conjunto de sus características físicas las cuales están determinadas por su *genotipo*, es decir el conjunto de sus cromosomas. Debido a esta diversidad de características fenotípicas, los individuos exhiben diferentes niveles de adecuación a su medio ambiente. Veamos como ejemplo el caso de los *pinzones de Darwin*. Estos pájaros que habitan en las islas Galápagos se alimentan de semillas. El tamaño de las semillas que consumen está en función del tamaño de sus picos, por lo que es de esperarse que la abundancia de semillas de un cierto tamaño propicie la abundancia de individuos cuyo tamaño de pico sea el más adecuado para consumir ese tipo de semillas. La longitud y anchura del pico del pájaro son características fenotípicas que se pueden incluir dentro del modelaje. En este caso se tendrá un sistema en un espacio de dos dimensiones, al cual llamaremos espacio fenotípico. En este espacio cada punto  $(k, \ell)$  corresponde a un largo y un ancho del pico del pájaro. La población de pinzones estará descrita por una variable  $x_{k,\ell}(t)$ , la cual denota el número de individuos al tiempo  $t$ , cuyas características fenotípicas están etiquetadas por  $(k, \ell)$ . En este caso el problema principal es entender el mecanismo de selección natural por medio del cual la población cambia su composición predominando los individuos con un nivel de adecuación más alto.

Hasta aquí no hemos tomado en cuenta el hecho de que los individuos a los cuales nos hemos referido, se encuentran en un espacio físico real. La noción de espacio físico se introduce en estos modelos como una red discreta de puntos. En un espacio bidimensional  $\mathbf{x}(t; i, j)$  denota el estado del punto  $(i, j)$  al tiempo  $t$ , donde  $\mathbf{x} = (x_1, \dots, x_e)$  es el conjunto de especies que interactúan en dicho espacio. Debido a la competencia entre especies la población de cada una de ellas, es decir  $x_k(t; i, j)$  estará limitada por una restricción ecológica que denotamos por  $F_{\mu,k}(\{x_\ell(t; i, j)\})$ . En este caso  $\mathbf{x}(i, j)$  está dado por una ecuación como la siguiente:

$$x_k(t+1; i, j) = F_{\mu,k}(\{x_\ell(t; i, j)\}) + D_k \left( F_{\mu,k}(\{x_\ell(t; i, j-1)\}) + \right.$$

$$F_{\mu,k}(\{x_t(t; i, j + 1)\}) + F_{\mu,k}(\{x_k(t; i - 1, j)\}) + F_{\mu,k}(\{x_k(t; i + 1, j)\}) - 4F_{\mu,k}(\{x_k(t; i, j)\}) \quad (2.8)$$

donde  $\mu$  es un parámetro y  $D_k$  es un coeficiente de difusión. En [27] Solé et. al. estudian un modelo de este tipo en donde, a fin de garantizar la estabilidad del sistema y que las poblaciones sean positivas, introducen una función tipo Heaviside, definida por:

$$H(w) = \begin{cases} w & \text{si } w > 0 \\ 0 & \text{si } w \leq 0 \end{cases} \quad (2.9)$$

de tal forma que la población de la  $k$ -ésima especie al tiempo  $t$  en el punto  $(i, j)$  del espacio estaría dada por  $H(x_k(t; i, j))$ .

En este trabajo estamos interesados en modelar la evolución de un conjunto de secuencias genéticas visto como una población en donde existe competencia ecológica y mutaciones. Las diferentes especies corresponden a composiciones químicas de las secuencias genéticas y el espacio corresponde a las células en las cuales se introducen estas secuencias y se espera poder aplicar este formalismo a otros problemas.

Antes de introducirnos de lleno en el planteamiento de un modelo es necesario discutir algunos aspectos importantes de la fenomenología involucrada en el problema que nos interesa. En los dos capítulos siguientes se introducen algunas definiciones y la terminología que se utiliza en la descripción de la fenomenología de la evolución de secuencias genéticas [28], [29] y del VIH [15], [18], [19].

$$(0.1) \quad \left( \begin{array}{c} ((1, 1, 1) \cdot (1, 1, 1))_1 \\ ((1, 1, 1) \cdot (1, 1, 1))_2 \\ ((1, 1, 1) \cdot (1, 1, 1))_3 \end{array} \right) = \left( \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right)$$

donde  $(1, 1, 1)$  es el vector de componentes  $1$ ,  $1$  y  $1$ , y  $(1, 1, 1)_i$  es el producto escalar de  $(1, 1, 1)$  con el vector de componentes  $1$ ,  $0$  y  $0$ ,  $0$ ,  $1$  y  $0$ ,  $0$ ,  $0$  y  $1$ , respectivamente.

$$(0.2) \quad \left( \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right) = \left( \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right)$$

lo que a su vez es la misma ecuación de un subconjunto de un sistema de ecuaciones lineales.

En un subconjunto de un sistema de ecuaciones lineales, una ecuación es redundante si puede eliminarse sin cambiar el conjunto de soluciones. En un sistema de ecuaciones lineales, una ecuación es redundante si puede eliminarse sin cambiar el conjunto de soluciones. En un sistema de ecuaciones lineales, una ecuación es redundante si puede eliminarse sin cambiar el conjunto de soluciones.

donde  $(1, 1, 1)$  es el vector de componentes  $1$ ,  $1$  y  $1$ , y  $(1, 1, 1)_i$  es el producto escalar de  $(1, 1, 1)$  con el vector de componentes  $1$ ,  $0$  y  $0$ ,  $0$ ,  $1$  y  $0$ ,  $0$ ,  $0$  y  $1$ , respectivamente.

## Capítulo 3

# FENOMENOLOGIA

### 3.1 ADN y ARN, LAS MOLECULAS DE LA VIDA

En casi todas las células vivas la información genética está almacenada en las cadenas de ácido desoxiribonucleico (ADN) y es esta información la que dirige la síntesis de proteínas, las cuales se fabrican a partir de un conjunto de veinte aminoácidos, (ver Tabla 3.1). El encargado de realizar la transferencia de dicha información es el ácido ribonucleico (ARN). Las moléculas de ADN son polímeros lineales formados por cuatro monómeros diferentes (Figura 3.1). Cada monómero tiene dos partes principales, una que es común a todos ellos y que consiste en un azúcar de cinco carbonos llamada desoxiribosa y un grupo fosfato. La otra parte, que es la que los hace diferentes, es una de las cuatro bases nitrogenadas Adenina (A), Guanina (G), Timina (T) y Citosina (C). La A y la G son purinas (R) y son más grandes que las pirimidinas (Y) T y C (Figura 3.2).

Dentro de las células el ADN se encuentra en forma de una doble hélice constituida por dos polímeros lineales antiparalelos con bases complementarias (Figura 3.1). Cada base está unida por dos o tres puentes de hidrógeno a una base en la cadena opuesta. Dados el tamaño y estructura de las bases este arreglo requiere que una A en una de las cadenas esté unida a una T en la cadena opuesta y que una G esté unida

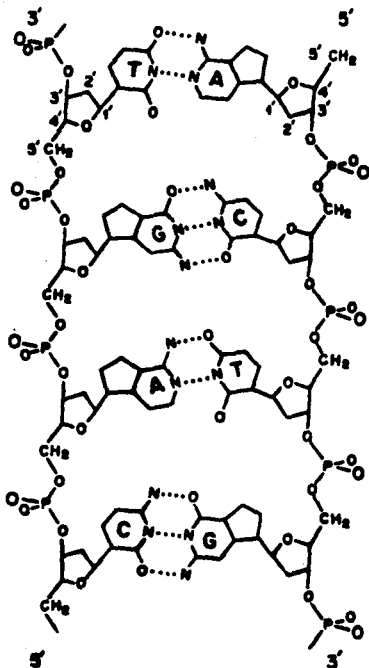


Figura 3.1: Diagrama esquemático de la doble hélice de ADN en donde puede observarse el esqueleto azúcar-fosfato y las cuatro bases nitrogenadas.

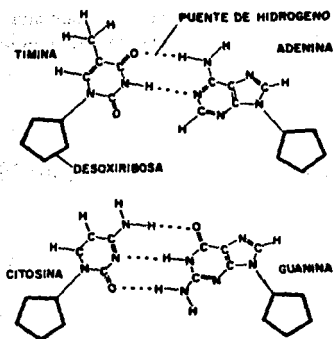


Figura 3.2: Las cuatro bases nitrogenadas Adenina, Guanina, Citosina y Timina. En la Figura se muestran los puentes de hidrógeno que las unen en la doble hélice de ADN.

a una C, así la secuencia de nucleótidos en una de las cadenas determina la secuencia en la otra. El par G-C está unido por tres puentes de hidrógeno, mientras que el par A-T está unido sólo por dos (Figura 3.2). En este sentido diremos que G y C son bases fuertes (F) y A y T son bases débiles (D).

Las moléculas de ARN son también polímeros lineales que presentan dos diferencias esenciales con respecto al ADN. La primera consiste en que el azúcar del ARN es ribosa en lugar de desoxirribosa y la segunda consiste en que la base pirimidica T se sustituye por el Uracilo (U) que es también una pirimidina. En general, el ARN se presenta como un polímero de cadena simple pero también puede presentar apareamiento de bases, en ese caso el Uracilo se une a la Adenina por medio de dos puentes de hidrógeno, por lo que resulta ser una base débil.

Entre las células podemos distinguir dos tipos, las eucariotes que son aquellas que poseen un núcleo bien definido que separa el material



genético del resto de la célula y las procariotes que carecen de dicho núcleo. La síntesis de proteínas sigue un esquema muy parecido en ambos tipos de células aunque con algunas diferencias que se discuten a continuación.

## 3.2 SINTESIS DE PROTEINAS Y EL CODIGO GENETICO

La síntesis de proteínas se inicia cuando la enzima ARN-polimerasa se adhiere a la doble hélice de ADN en algún sitio específico y comienza a transcribir una de las cadenas de ADN en una cadena de ARN. En células eucariotes existen tres tipos diferentes de esta enzima, las ARN-polimerasas I, II y III, mientras que en células procariotes hay un solo tipo. En ambos casos, la transcripción de ADN en ARN da lugar a tres tipos diferentes de ARN, el ARN mensajero (mARN), el ARN de transferencia (tARN) y el ARN ribosomal (rARN). Cada uno de los diferentes tipos de RNA cumple funciones específicas en el proceso de síntesis de proteínas: el mARN lleva el mensaje de las proteínas que han de sintetizarse, el tARN participa en la síntesis de proteínas, ya que lleva consigo los aminoácidos con los que se construyen las proteínas, y el rARN forma parte de los ribosomas que son las unidades en donde se lleva a cabo la síntesis. Cada conjunto de tres bases en la secuencia de mARN recibe el nombre de codón y cada codón codifica uno de los veinte aminoácidos que constituyen a las proteínas. A la identificación entre tripletes de bases (codones) y aminoácidos se le denomina código genético, el cual se muestra en la Tabla 3.1.

El ARN debe ser procesado antes de que pueda intervenir en la síntesis de proteínas. La transcripciones primarias de rARN y tARN son procesadas en ambos tipos de células mientras que el mARN es procesado únicamente en las células eucariotes. La razón de ello es que en células eucariotes los genes están formados por uno o varios segmentos llamados exones, que son los que en realidad codifican a las proteínas, los cuales están, por lo general, separados por secuencias que

Tabla 3.1: EL CODIGO GENETICO. Identificación entre tripletes de bases (codones) y aminoácidos.

X	Y				Z
	T	C	A	G	
T	Fenilalanina	Serina	Tirosina	Cisteína	T
	Fenilalanina	Serina	Tirosina	Cisteína	C
	Leucina	Serina	Stop	Stop	A
	Leucina	Serina	Stop	Triptófano	G
C	Leucina	Prolina	Histidina	Arginina	T
	Leucina	Prolina	Histidina	Arginina	C
	Leucina	Prolina	Glutamina	Arginina	A
	Leucina	Prolina	Glutamina	Arginina	G
A	Isoleucina	Treonina	Asparragina	Serina	T
	Isoleucina	Treonina	Asparragina	Serina	C
	Isoleucina	Treonina	Lisina	Arginina	A
	Metionina	Treonina	Lisina	Serina	G
G	Valina	Alanina	Acido aspártico	Glicina	T
	Valina	Alanina	Acido aspártico	Glicina	C
	Valina	Alanina	Acido glutámico	Glicina	A
	Valina	Alanina	Acido glutámico	Glicina	G
	Alifático	Alcoholico	Polar	Polar	

no codifican y que reciben el nombre de intrones.

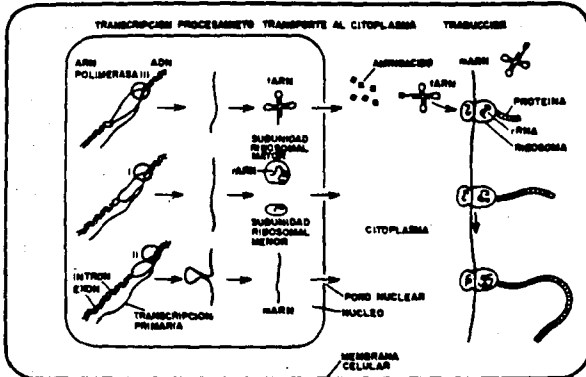
Durante el procesamiento del mRNA la transcripción que proviene de los intrones es eliminada a fin de dar lugar al mRNA maduro. En el procesamiento del rARN, este se corta en diferentes regiones para dar lugar a al menos tres tipos de rARN los cuales al unirse con algunas proteínas forman ribosomas. En células eucariotes los procesos hasta ahora descritos tienen lugar dentro del núcleo, posteriormente los tres tipos de ARN se desplazan hacia el citoplasma en donde tiene lugar la síntesis de proteínas (Figura 3.3). Durante el procesamiento de la transcripción primaria de tARN se cortan los extremos de la molécula y esta adquiere forma de horquilla. Una vez fuera del núcleo celular, cada tARN se une al aminoácido correspondiente de tal forma que la estructura final del tARN es la siguiente: en uno de los extremos de la horquilla se encuentra una secuencia de tres bases, que es complementaria a uno de los diferentes codones del mRNA (y por esta razón se le llama anticodón) y en el otro extremo se encuentra el aminoácido correspondiente al codón en cuestión según la Tabla 3.1. Algunas moléculas de tARN pueden reconocer más de un codón, lo cual podría deberse a la degeneración del código genético de la que hablamos anteriormente.

Si llamamos  $X, Y$  y  $Z$  a las posiciones de las bases en un codón, la posición menos importante para definir que aminoácido corresponde a un codón es la posición  $Z$ , ya que como puede apreciarse de la Tabla 3.1, en general es muy probable que si se altera la base en dicha posición el aminoácido que codifica sea el mismo. Por el contrario la posición determinante para la definición de un aminoácido es la posición  $Y$ . Al alterar la base en esta posición el aminoácido resultante no solamente será diferente sino que puede cambiar incluso su naturaleza polar. Por su parte la posición  $X$  queda en una situación intermedia. Este hecho se traduce en la existencia de diferentes probabilidades de que una mutación al azar sea transmitida a los descendientes. Si denotamos por  $p_X$  la probabilidad de que una mutación en la posición  $X$  se fije al genoma (con definiciones análogas para  $p_Y$  y  $p_Z$ ), podemos resumir este hecho por medio de la siguiente desigualdad:

$$p_Z > p_X > p_Y \quad (3.1)$$

### 3.2. SINTESIS DE PROTEINAS Y EL CODIGO GENETICO 21

EUCARIONTES



PROCARIONTES

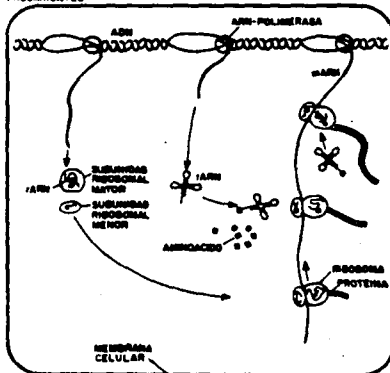


Figura 3.3: Síntesis de proteínas en células eucariotes y procariontes. Tomada de [30].

a estas probabilidades se les llama a veces probabilidades de mutación aceptadas.

Este hecho no significa de ninguna manera que la base en la posición *Z* sufra mutaciones con mayor frecuencia, sino que debido a que en general el aminoácido resultante realizará las mismas funciones que el original, esta mutación puede ser transmitida de generación en generación sin dañar a los descendientes. En el caso de mutaciones en las posiciones *X* o *Y*, estas son transmitidas a las generaciones posteriores con una probabilidad muy pequeña debido a la existencia de mecanismos de reparación que contribuyen a que aquellas proteínas que realizan funciones vitales dentro de un organismo se expresen con exactitud.

Es claro entonces que la dependencia de la probabilidad de mutación con la posición de la base en el codón es el resultado de la restricción selectiva sobre la funcionalidad de las proteínas resultantes. Por esta razón nos referiremos a este hecho como una restricción de selección natural externa.

### 3.2.1 DINAMICA RIBOSOMAL

Los ribosomas son complejos moleculares en los cuales el mRNA se traduce en proteínas. Cada ribosoma consta de dos subunidades, una más grande que la otra, constituidas por proteínas y rARN. En células procariotes (eucariotes) a la subunidad mayor se le denomina 50S (60S), mientras que la menor se identifica como la región 30S (40S). Cada una de estas subunidades del ribosoma está a su vez formada por diferentes ARNs y algunas proteínas. En particular la subunidad 30S está formada por una partícula conocida como 16S y 21 proteínas diferentes.

Existe evidencia de interacciones entre bases durante el proceso de traducción. En particular la región 16S de la subunidad 30S del ribosoma procariótico puede tener interacciones con cadenas de mRNA, previas a la traducción (Figura 3.4). Estas interacciones podrían tener el efecto de realzar la traducción de algunos mRNA particulares. Firpo y Dahlberg identificaron [31] algunos de los sitios de la región 16S del

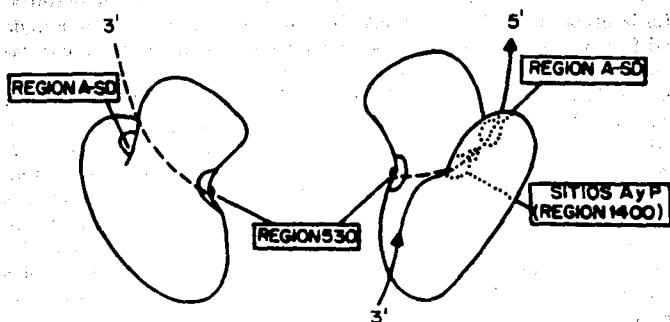


Figura 3.4: Modelo hipotético que muestra los puntos de contacto y el movimiento del mRNA con respecto al ribosoma previos a la etapa de iniciación. La línea punteada muestra la orientación propuesta de la porción 3' del mRNA durante la traducción. En la figura se indican además de la región Anti-Shine-Delgarno los sitios A y P.

rRAN que pueden ser los responsables de la traducción selectiva de algunos mRNA. En su trabajo presentan un modelo del ribosoma en el que muestran una trayectoria probable para el mRNA cuando este pasa por los supuestos sitios de contacto, los cuales se localizan en la superficie del ribosoma, durante el proceso de selección y durante la traducción misma. Cuando el mRNA hace su contacto inicial con el ribosoma las cadenas de mRNA buscan secuencias complementarias en el rARN 16S con las cuales establecer contactos transitorios. Se cree que estos enlaces transitorios tienen el efecto de incrementar la concentración de algunos tipos de mRNA en la región 30S, aumentando de esta manera su probabilidad de ser traducidos. Durante el proceso de traducción el mRNA vuelve a establecer contactos transitorios con el ribosoma.

En el proceso de traducción pueden distinguirse tres etapas princi-

pales: *iniciación, elongación y terminación* de la cadena polipeptídica. En la etapa de iniciación, el mRNA se posiciona sobre el ribosoma de tal forma que su mensaje pueda traducirse secuencialmente en una cadena de aminoácidos. Durante esta etapa, en células procariotes, son particularmente importantes las interacciones de la llamada secuencia de Shine-Delgarno y un codón de iniciación (que puede ser AUG o GUG) con el rARN así como la presencia de secuencias específicas adyacentes al codón de iniciación. La etapa de elongación consiste en un ciclo que se repite cada vez que un aminoácido es añadido a la cadena peptídica. Durante la última etapa el mRNA y la cadena peptídica recién sintetizada se separan del ribosoma. Al principio de cada iteración del ciclo que se repite durante el proceso de elongación, dos moléculas de tARN están unidas a dos sitios vecinos en el rARN 16S conocidos como los sitios *A* y *P*. Cuando se establece el enlace anticodón-mARN, el aminoácido respectivo se une por medio de enlaces peptídicos a la cadena peptídica que se está formando y la molécula de mRNA se desplaza avanzando un codón con respecto al ribosoma. Tan pronto como se expone el sitio de enlace para el siguiente tARN el enlace codón-anticodón se rompe, el tARN anterior es liberado y la síntesis de proteínas continúa.

Cuando solo se toma en cuenta el carácter débil (D) o fuerte (F) de las bases, los 64 diferentes codones (Tabla 3.1) se agrupan en ocho clases: DDD, DDF, DFD, FDD, DFF, FDF, FFD y FFF. En la Figura 3.5.a, se muestra la región final del rARN 16S del ribosoma de varios organismos [32]. Esta región se asocia con la iniciación de la traducción de proteínas y con el desplazamiento del mRNA con respecto al rARN. Las secuencias se muestran en orden inverso. Si las secuencias para *E. coli* y *X. laevis* se escriben en términos de D/F, se hace evidente la presencia de tripletes del tipo DDF y FDF (Figura 3.5.b).

En la Figura 3.6.a se muestran la secuencia de otra región del rARN 16S de cuatro organismos diferentes que resulta ser la misma para los cuatro. En la Figura 3.6.b la misma secuencia se escribe en la representación D/F, una vez más puede notarse la presencia de tripletes FDF. En ambas figuras, la aparición del par DF en las posiciones *Y* y *Z* del codón sugiere un cierto grado de regularidad que podría reflejarse

ORGANISMO	SECUENCIA
	1530
	• • • • •
E. coli	AOH UUC CUCCACUAG <u>GU</u> ---
B. brevis	AOH CUUCCUCCACUAG <u>GU</u> ---
Mitocondria humana	---GAAGCAG <u>GU</u> ---
H. volcanis	NDHU CCUCCACUAG <u>GU</u> ---
X. laevis	AOH UUA CUAG <u>GA</u> ---

(a)

E. coli	AOH <u>DDF</u> <u>EDF</u> <u>EDF</u> <u>DDF</u> <u>FD</u> ---
X. laevis	AOH D <u>DDF</u> <u>DDF</u> <u>FD</u> ---

(b)

Figura 3.5: a).-Región final del rARN 16S de diferentes organismos. Los puntos indican puntos susceptibles de ser atacados químicamente y las bases separadas por rectángulos son parte de regiones helicoidales. b).- Representación de las secuencias para E. coli y X. laevis listadas en (a) en la base degenerada D/F.



ORGANISMO	SECUENCIA
E. coli	--CAU <sup>530</sup> AAUGG <sup>520</sup> CGCCGACGACC <sup>GUG</sup>
Cloro plasmas	--CA <sup>GAAUGGCGCCGACGACCGUG</sup> (a)
H. volcani	--CAUAAUGGCGCCGCCGACCGUG
X. laevis	--CUAAUGGCGCCGACGACCGUG

---D<sup>DD</sup>DF<sup>FFFF</sup>DF<sup>DF</sup>DF<sup>FF</sup>F---

E. coli (b)

Figura 3.6: a).- Secuencia de bases en la región 515-535 del 16-ARN del ribosoma para diferentes organismos. b).- Representación de las secuencias listadas en (a) en el código D/F.

en las secuencias de mRNA.

### 3.3 FISICOQUIMICA DE LOS ACIDOS NUCLEICOS

Algunas mediciones termodinámicas en oligonucleótidos y polinucleótidos diseñados y sintetizados *ex profeso* han permitido hacer estimaciones de la entropía, la entalpía y la energía libre en pares de bases tanto para ADN [33] como para ARN [34], ambos en forma de doble hélice. En este tipo de mediciones la unidad básica la constituyen dos pares de bases que son vecinas a lo largo de la cadena, a cada pareja de bases se le denomina digrama. Cada digrama tiene una interacción horizontal debida a los puentes de hidrógeno, una energía vertical de apilamiento y una interacción diagonal (Fig. 3.7).

En cualquier doble hélice de ADN (o ARN), es posible identificar 16 diferentes pares de bases que sean vecinas cercanas a lo largo de

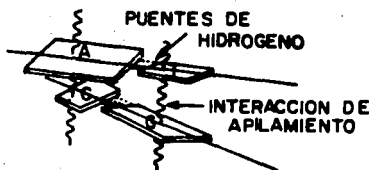


Figura 3.7: Par de digramas de una doble hélice de ADN en donde se muestran las interacciones de apilamiento y horizontal, esta última debida a los puentes de hidrógeno.

la hélice, sin embargo de estos 16 pares sólo 10 de ellos tienen diferentes energías de interacción debido a que la secuencia en una de las ramas determina la secuencia en la otra y a las simetrías que pueden establecerse (Fig. 3.7). Por lo tanto la energía de interacción de cualquier doble hélice (ya sea de ADN o ARN) depende de tan sólo 10 parámetros.

Breslawer et. al. [33] determinaron las propiedades termodinámicas de pares de bases en ADN mediante un análisis de espectroscopía ultravioleta y calorimetría diferencial. En el experimento realizado se utilizaron 19 oligómeros y 9 polímeros de ADN autocomplementarios (es decir cada una de las ramas de la doble hélice es idéntica a su complementaria), los cuales fueron cuidadosamente seleccionados para minimizar los inconvenientes que pudieran surgir en el experimento. Durante el experimento se determinaron las energías libres, entropías y entalpías asociadas a la transición de fase hélice-lazo de cada una de las 28 secuencias utilizadas y a partir de esas mediciones se determinaron los 10 valores de las energías de interacción. En la Tabla 3.2 se dan los valores de la entalpía  $\Delta H$  y la energía libre  $\Delta G$  que se predicen mediante el procedimiento descrito. Las diferencias entre dichos

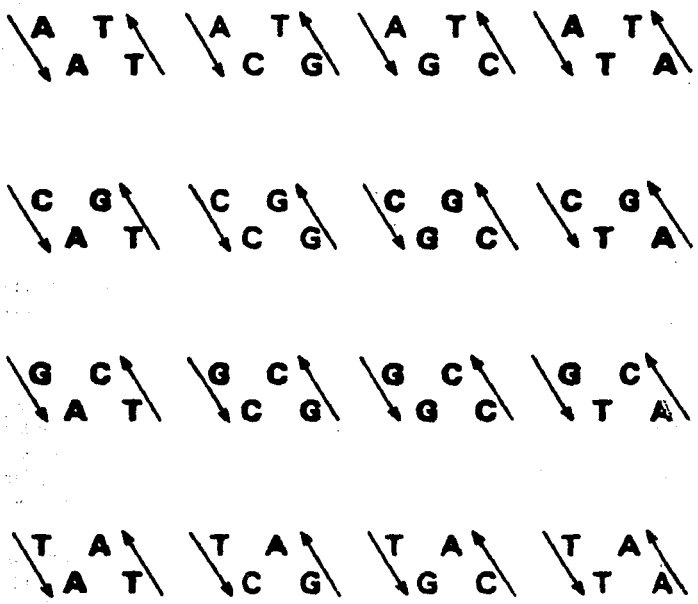


Figura 3.8: En esta figura se muestran los 16 diferentes pares de bases que se encuentran en cualquier doble hélice de ADN. Mediante un giro de 180 grados se puede ver que la energía de interacción del digrama TG es igual a la del digrama CA. Con negritas se indican los 10 digramas con energías de interacción diferentes. Las flechas indican la dirección en que debe leerse la secuencia.

Tabla 3.2: Entalpía  $\Delta H$  y Energía Libre  $\Delta G$  para digramas de ADN. Los valores de  $\Delta H$  y  $\Delta G$  estan tomados de la referencia [33] y están expresados en Kcal/mol.

Digramas	D/F	$-\Delta H$	$-\Delta G$
AA	DD	9.1	1.9
AT	DD	8.6	1.5
TA	DD	6.0	0.9
CG	FF	11.9	3.6
GC	FF	11.1	3.1
GG	FF	11.0	3.1
CA	FD	5.8	1.9
CT	FD	7.8	1.6
GA	FD	5.6	1.6
GT	FD	6.5	1.3

valores y algunos determinados experimentalmente es muy pequeña lo cual muestra que en una primera aproximación la entropía o la energía libre de una doble hélice de ADN es igual a la suma de las entropías o energías libres de los digramas que la forman.

Freier et. al. [34] determinaron las propiedades termodinámicas de ARN en forma de doble hélice por medio de un análisis de las curvas de fusión de las diferentes cadenas de secuencias. En el experimento se utilizaron 45 oligómeros de ARN los cuales fueron elegidos de tal forma que los errores durante el experimentos fueran mínimos. Los valores de la entalpía y la energía libre de los digramas de ARN se predicen en función de los cambios de la energía libre debidos a la formación de la doble hélice a partir de las dos cadenas de nucleótidos. En la Tabla 3.3, se dan los valores de la entalpía  $\Delta H$  y la energía libre  $\Delta G$  obtenidos por medio del experimento de Freier et. el. La comparación de dichos valores y algunos otros medidos experimentalmente muestra que las diferencias son del orden del 10%.

Tabla 3.3: Entalpía  $\Delta H$  y Energía Libre  $\Delta G$  para digramas de ARN. Los valores de  $\Delta H$  y  $\Delta G$  estan tomados de la referencia [34] y están expresados en Kcal/mol.

Digramas	D/F	$-\Delta H$	$-\Delta G$
AA	DD	6.6	0.9
AU	DD	5.7	0.9
UA	DD	8.1	1.1
CG	FF	8.0	2.0
GC	FF	14.2	3.4
GG	FF	12.2	2.9
CA	FD	10.5	1.8
CU	FD	7.6	1.7
GA	FD	13.3	2.3
GU	FD	10.2	2.1

Como puede verse de las Tablas 3.2 y 3.3, los valores de  $\Delta H$  y  $\Delta G$  dependen principalmente de la intensidad de los enlaces en los planos formados por las bases de cadenas complementarias (digramas), por lo que la energía libre de un polímero dependerá de la secuencia de bases más que de su composición química. En general se observa que:

- Los digramas de ADN (expresados en términos de D y F) que son de la forma DD o FF están más fuertemente unidos que los DF y los FD.
- En el caso del ARN la unión de los digramas FF es más fuerte que la de los DF y los FD, siendo los DD los más débilmente unidos.

En cadenas de ARN se observa una inhomogeneidad espacial en la densidad de energía [17], la cual podría estar relacionada con la eficiencia (en donde la eficiencia es una medida de la razón de síntesis de proteínas) y/o la exactitud (la exactitud está asociada con el grado de ausencia de errores) con que el mARN funciona en el proceso de traducción en el ribosoma. Thomas et. al. [35] y Dix et. al. [36], mostraron mediante experimentos realizados *in vitro*, que algunos codones

**GGA GAC CGG AAG CUU GGG CUG CAG GAG**  
**GAU UUA AUC AUG XYZ AAG AUC UCG**

**FFD FDF FFF DDF FDD FFF FDF FDF FDF**  
**FDD DDD DDF DDF XYZ DDF DDF DDF**

Figura 3.9: Secuencia del mRNA utilizado en el análisis de la referencia [35].

degenerados (cambio en la última base) pueden diferir en eficiencia. En particular, el codón UUC es más eficiente que el UUU, y el CUC más eficiente que el CUU. En términos del código D/F, este hecho se traduce en que codones del tipo NDF son más eficientes que los del tipo NDD, en donde N puede tomar cualquiera de los valores D o F.

En [35] también se menciona que la exactitud en el proceso de traducción de cadenas de mRNA inhomogéneas puede ser hasta doce veces mayor que la de cadenas homogéneas. Estos resultados sugieren que la eficiencia está relacionada principalmente con la interacción tARN-mRNA, mientras que la exactitud depende del "contexto" en el que se encuentre el codón y por lo tanto podría estar relacionada con la interacción rARN-mRNA. En las secuencias utilizadas en el experimento, se encuentra un dominio de codones del tipo NDF (Figura 3.7). Esto sugiere la existencia de una estrecha relación entre la presencia de codones NDF y la exactitud en los procesos de traducción.

Estas restricciones de tipo energético y estérico, pueden implicar

desviaciones locales del azar en la composición de los ácidos nucleicos. Nos referiremos a estas restricciones (locales) como restricciones de selección natural internas. Por otra parte del análisis de secuencias del banco de datos de secuencias genéticas GenBank, se sabe que en el ARN correspondiente a exones de células eucariotes la entalpía satisface la relación [37]:

$$\Delta H_{XY} < \Delta H_{YZ} < \Delta H_{ZX'} \quad (3.2)$$

en donde  $X$ ,  $Y$  y  $Z$  son las diferentes posiciones de las bases en un codón. Esto implica que la energía de amarre entre digramas tampoco se distribuye uniformemente, sino que depende de la posición del digrama dentro del codón. De aquí se infiere la existencia de regularidades en el arreglo de bases [37], [38]. Dado que estas regularidades pueden ser de origen termodinámico, el lenguaje degenerado D/F para las bases resulta ser una representación adecuada. Las desigualdades de la ecuación (3.2) sugieren que los enlaces DD deberían ser los más abundantes en los digramas  $XY$  mientras que los FF deberían encontrarse con mayor frecuencia en los digramas  $ZX'$ , sin embargo es imposible satisfacer ambas condiciones simultáneamente. De los ocho diferentes codones degenerados los que presentan menor conflicto son los de tipo NDF que corresponden a una menor energía, esto implica que el enlace que predomina en el par  $YZ$  es DF.

## Capítulo 4

# EL SISTEMA INMUNE Y EL VIRUS DEL SIDA

### 4.1 EL SISTEMA INMUNE

El sistema inmune de los animales vertebrados es una complicada red de células y moléculas capaces de distinguir entre *lo propio y lo extraño* proporcionando defensas al organismo contra bacterias, virus, hongos y otros agentes patógenos. La respuesta inmune puede ser de dos tipos: *humoral o celular*. La primera se desencadena contra los antígenos que permanecen en medios extracelulares, mientras que la segunda se dirige contra los antígenos presentes en la superficie de las células.

Las células más importantes en la respuesta inmune son los *linfocitos*, los cuales se encuentran en la sangre, el fluido linfático y en los órganos linfáticos. Los linfocitos se dividen en dos grupos principales: los de tipo **B** y los de tipo **T**. Al igual que todas las células que circulan por el torrente sanguíneo, los linfocitos se derivan de un tipo de célula que se origina en la médula. Los linfocitos **B** completan su ciclo de maduración en la médula mientras que los linfocitos **T** lo hacen en el timo. Aunque ambos tipos de células son muy parecidas en forma y tamaño participan de forma diferente en la respuesta inmune, las de tipo **B** toman parte la respuesta humoral y las de tipo **T** lo hacen principalmente en la respuesta celular.



## 34 CAPITULO 4. EL SISTEMA INMUNE Y EL VIRUS DEL SIDA

Los linfocitos **B** son los encargados de fabricar anticuerpos, o inmunoglobulinas, los cuales permanecen en su superficie y actúan como receptores. Cuando un antígeno se une a un anticuerpo localizado en la membrana de la célula ésta es estimulada, activándose una serie de procesos que dan lugar a dos tipos de células, las *plasmáticas* y las de *memoria*. Las células plasmáticas secretan anticuerpos los cuales se unen a los antígenos que circulan ya sea en la sangre o en el fluido linfático marcándolos para ser destruidos por otros elementos del sistema inmune entre los cuales se encuentran los macrófagos. Por su parte las células de memoria circulan en el organismo hasta por varios años activando rápidamente una respuesta inmune en caso de que el antígeno ataque nuevamente al organismo.

Entre las células **T** podemos distinguir dos grupos principales dependiendo del receptor que llevan en su superficie, las **T4** que funcionan como *inductoras* y *ayudadoras* y las **T8** con funciones *citotóxicas* y *supresoras*. La función de las células T inductoras consiste en desencadenar la maduración de las células T a partir de formas precursoras en sus diferentes formas funcionales, las citotóxicas se unen a los antígenos localizados en la superficie de las células que han sido infectadas por algún virus y destruyen a la célula, las supresoras se encargan de detener la respuesta inmune cierto tiempo después de que ha sido activada y las ayudadoras se unen a los antígenos localizados en la superficie de las células B. En general las de las células T no son capaces de reconocer antígenos que circulan libremente en la sangre o en el fluido linfático, pero pueden reconocer antígenos presentes en la superficie de otras células y su activación está condicionada a que el antígeno se presente junto con una proteína de la propia célula, la cual está determinada por el *Complejo Principal de Histocompatibilidad* (Major Histocompatibility Complex, MHC).

En la Figura 4.1 se muestra un esquema con los principales aspectos de la respuesta inmune ante una infección en la que los antígenos permanecen fuera de las células y en la Figura 4.2 se muestra la acción de una célula citotóxica contra una célula infectada por un virus.

#### 4.1. EL SISTEMA INMUNE

##### CELULA DE MEMORIA

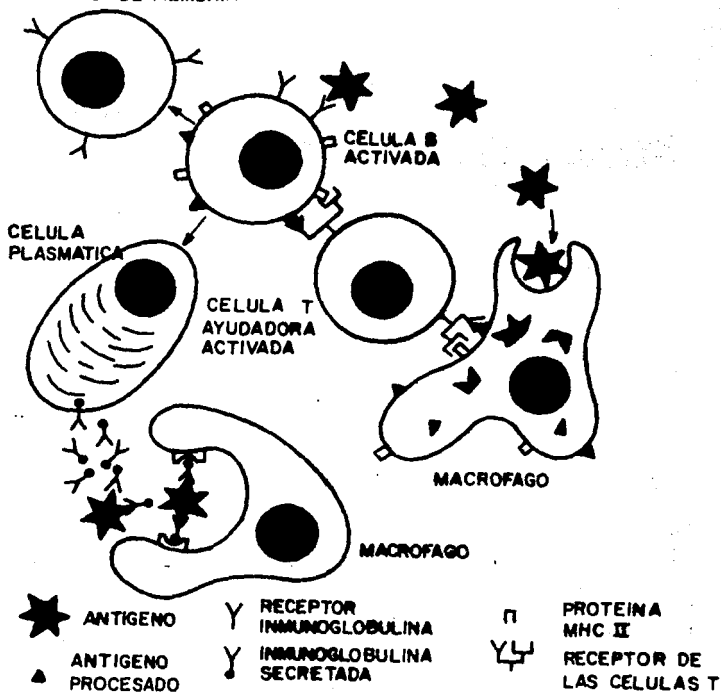


Figura 4.1: Diagrama esquemático que muestra las funciones que desempeñan algunos de los principales elementos del sistema inmune durante la respuesta ante una infección. Tomada de [40].

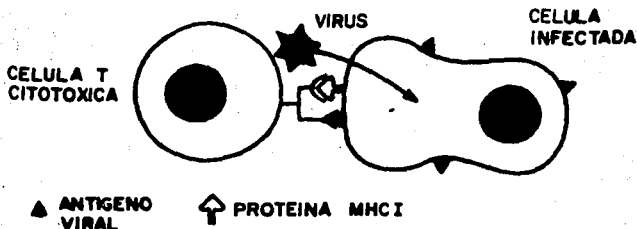


Figura 4.2: En esta Figura se muestra la manera en que las células T citotóxicas reconocen al antígeno localizado en la superficie de una célula infectada por un virus. Tomada de [40].

## 4.2 VIRUS Y RETROVIRUS

Los virus son entidades cuyo material genético está constituido por ADN y cuyos genes son capaces de integrarse al material genético de las células que infectan. Una vez que el ADN viral se integra al ADN celular, se replica conjuntamente con éste de modo que el ADN viral es heredado por las células hijas. Cuando el ADN viral se ha integrado a ADN celular se le llama provirus, debido a que bajo ciertas condiciones da lugar a la formación de nuevos virus.

Existe otra clase de virus, llamados retrovirus, cuyo genoma está constituido por ARN, el cual se transcribe en ADN por medio de una enzima llamada reversotranscriptasa (o retrotranscriptasa). Este fenómeno se conoce como transcripción inversa o retrotranscripción, debido a que se da en sentido inverso al usual, y es por esta razón que esta clase de virus se conocen como retrovirus. Una vez que el ADN proveniente del ARN viral se integra al ADN celular, también se le llama provirus. Es importante hacer notar que la retrotranscripción

del ARN viral en ADN, sólo tiene lugar cuando el retrovirus se encuentra en el interior de una célula.

El ciclo de vida de un retrovirus comienza cuando éste se adhiere y penetra en una célula inyectándole su material genético y sus proteínas. Típicamente el ARN de un retrovirus consta de tres regiones codificadoras principales conocidas como **gag**, **pol** y **env** (Figura 4.3), las cuales codifican las proteínas de la parte central del virus, la enzima reverso-transcriptasa y las proteínas que forman la envoltente respectivamente. Así mismo, el ARN viral contiene tres regiones no codificadoras dos en los extremos y otra llamada  $\psi$ . Los extremos terminales del ADN doble-helicoidal resultante de la retrotranscripción, llamados repeticiones terminales largas (*long terminal repeats LTR*), tienen influencia sobre la actividad de los genes virales y facilitan la inserción del ADN viral en el ADN celular (Figura 4.3). Después de que el ADN viral se ha integrado al material genético de la célula, éste dirige la síntesis de sus proteínas y su ARN.

Los procesos que tienen lugar durante el ciclo de vida de un retrovirus no se llevan a cabo con toda exactitud, sino que pueden ocurrir mutaciones. Las principales fuentes de mutaciones son la transcripción de ADN en ARN y la retrotranscripción de ARN en ADN. En el proceso de transcripción debido a la acción de una enzima reparadora, la probabilidad de mutación resulta ser del orden de  $10^{-9}$  por base y por ciclo de reproducción celular. Sin embargo el proceso de retrotranscripción no cuenta con una enzima reparadora, lo que ocasiona que la probabilidad de mutación sea del orden de  $10^{-3}$  por base y por ciclo.

## 4.3 EL RETROVIRUS DEL SIDA

### 4.3.1 LA VISION TRADICIONAL

En 1981 se identificaron los primeros casos del *síndrome de inmunodeficiencia adquirida* (SIDA) debido al cual el sistema inmune de las víctimas se ve deteriorado ocasionando que éstas mueran a causa de

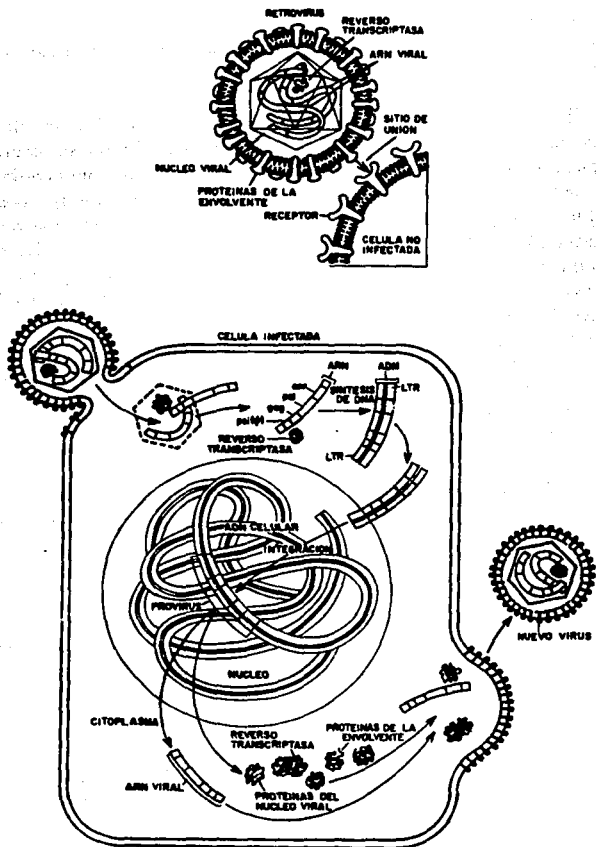


Figura 4.3: Ciclo de vida de un retrovirus. Tomado de [41].

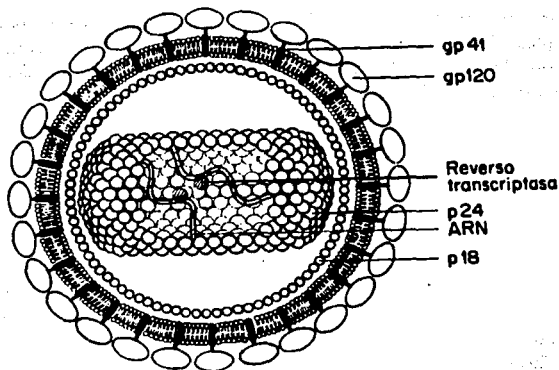


Figura 4.4: Diagrama esquemático del retrovirus VIH. Tomado de [42].

infecciones no relacionadas directamente con el virus.

Experimentos realizados en 1983 y 1984, mostraron [39] que el causante de esta enfermedad es el *Virus de la Inmunodeficiencia Humana VIH* (Human Immunodeficiency Virus, HIV). Este virus de ARN (o retrovirus) está relacionado con el virus linfotrópico humano I de las células T (human T-cell lymphotropic virus I, HTLV-I), causante de un tipo de leucemia poco común, y por lo general las células a las cuales ataca son los linfocitos **T4**, los cuales juegan un papel muy importante en la respuesta inmune de los organismos.

Como puede verse en la Figura 4.4, la membrana que envuelve el material genético del VIH consta de una bicapa de lípidos y dos glicoproteínas conocidas como **gp41** y **gp120**. La **gp41** atraviesa la membrana y está unida por medio de enlaces disulfuro a la **gp120** que está situada en la cara externa. La parte central consta de dos copias de ARN genómico, t-ARN's asociados y varias moléculas de la enzima reversotranscriptasa. Este conjunto se encuentra rodeado de un gran número de copias de las proteínas **p18** y **p24**. Mediante la interacción

#### 40 CAPITULO 4. EL SISTEMA INMUNE Y EL VIRUS DEL SIDA

de la proteína **gp120** de la membrana viral con un receptor de la membrana plasmática del linfocito **T4**, conocido como **CD4**, el VIH penetra al interior de la célula, las dos membranas se fusionan y el núcleo viral se libera dentro del linfocito.

En algunos casos el sistema inmune no consigue eliminar a los antígenos presentes en un determinado organismo ya sea porque los anticuerpos producidos no los atacan efectivamente o debido a una alta tasa de mutación, de tal forma que aunque los anticuerpos producidos sean los adecuados, se vuelven totalmente inefectivos para destruirlos. El virus del SIDA, sobrevive dentro de un organismo debido al deterioro causado en el sistema inmune. Una vez que el retrovirus se encuentra dentro de la célula que ha infectado, puede permanecer en su interior en vida latente hasta que ésta es estimulada inmunológicamente por alguna infección, en ese momento el retrovirus responde reproduciéndose en gran número. Cuando las nuevas partículas virales salen de la célula destruyen la membrana celular y ocasionan la muerte del linfocito. La disminución del número de linfocitos **T4** provocada de esta manera, deteriora el sistema inmune de la persona infectada dejándola vulnerable a toda clase de infecciones.

#### 4.3.2 LA VISION ACTUAL

Como se mencionó anteriormente en un principio se pensaba que el colapso total de la respuesta inmune de los enfermos de SIDA se debía a una sola causa: la disminución y alteración de las funciones de los linfocitos **T4**. Actualmente se cuenta con varias hipótesis que más que ser alternativas son complementarias [15], [19]. Entre estas hipótesis se encuentran:

- 1.- El VIH escapa al control del sistema inmune debido a una alta tasa de mutación.
- 2.- La infección por VIH induce el suicidio (apoptosis) de las células **T**.
- 3.- El virus del SIDA imita a algunas proteínas cruciales en el sistema inmune humano creando confusión en este y provocando que se ataque a sí mismo.
- 4.- La infección por VIH causa un desbalance en la células **T** ayudado-

ras provocando una alta producción de anticuerpos y una muy baja producción de células asesinas (variantes de las células T que atacan a los antígenos y los destruyen).

5.- El virus VIH causa una parálisis inmune impidiendo que se activen las células T (anergia).

6.- La infección por VIH induce una reacción autoinmunitaria.

Poco tiempo después del descubrimiento del VIH-1 se encontró en pacientes originarios del este de Africa un segundo tipo de virus, el VIH-2. Los análisis realizados mostraron que este nuevo virus difiere en más del 55% del VIH-1. La proliferación del VIH-2 no ha sido tan extendida como la del VIH-1, ya que principalmente se ha encontrado en pacientes de dicha región de Africa. Actualmente se tienen identificadas del orden de 25 sepas diferentes del VIH-1, en tanto que del VIH-2 solamente se han encontrado alrededor de 5. En este trabajo nos referiremos principalmente al VIH-1, ya que se cuenta con una mejor estadística.

Se sabe también que el virus no ataca únicamente a los linfocitos T4, ya que ha sido encontrado en diversos tipos de células (alrededor de 40) [15], aunque en algunas de ellas en muy pequeñas cantidades. Algunas de las células susceptibles de ser atacadas por el VIH son: *linfocitos T*, *macrófagos*, *células asesinas naturales*, *células del epitelio del timo*, *fibroblastos*, *células de Lagerhans (piel)*, *células de la mucosa intestinal*, etc, siendo el principal blanco de ataque los linfocitos T4.

En la Figura 4.5 se muestra un diagrama esquemático de los principales rasgos que se han identificado en la evolución de la infección por VIH. En la gráfica se muestran tres cantidades como función del tiempo: i).- la cantidad de virus en la sangre ii).- el número de linfocitos T4 (o ayudadores), que como ya se vio anteriormente juegan un papel vital en la respuesta inmune del organismo, y iii).- la producción de anticuerpos dirigidos contra proteínas externas. El esquema inicia en el momento que tiene lugar la infección. En la primera etapa, la cual tiene una duración de entre 4 y 8 semanas, se observa un incremento en el número de virus en sangre y disminuye el número de linfocitos T4. A continuación se puede identificar una etapa de latencia con una duración de hasta doce años durante la cual el organismo tolera la in-



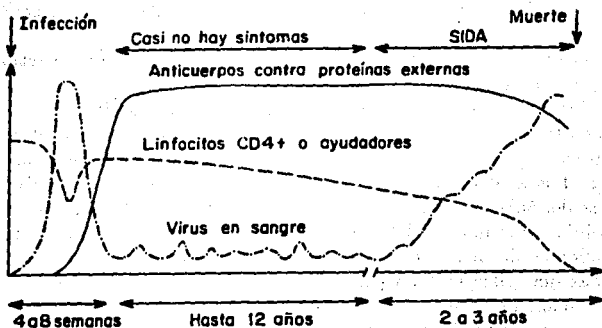


Figura 4.5: Diagrama esquemático de la evolución hacia el SIDA, a partir de la infección por VIH. Tomado de [19].

fección, es decir no reconoce al virus como un antígeno y por lo tanto no lo destruye. Durante este período, el virus prolifera infectando a una gran cantidad de células de diversos tipos y los síntomas de la enfermedad prácticamente no se manifiestan. Finalmente llega un momento en que el organismo reconoce al virus como un agente extraño y reacciona atacando activamente a aquellas células que han sido infectadas, las cuales en su mayoría serán linfocitos T4. Esta última etapa tiene una duración de entre dos y tres años y es la etapa del síndrome. El final de esta etapa está marcado por la muerte del individuo.

Se mencionó en la sección 4.2 que típicamente un retrovirus presenta una tasa de mutación del orden de  $10^{-3}$  por base y por ciclo. En el retrovirus VIH, es el gen *env* el que muta con mayor frecuencia. La secuencia de la proteína **gp120** de las partículas virales obtenidas de diferentes pacientes o incluso del mismo paciente a tiempos diferentes presenta un alto grado de variabilidad, por lo que la preparación de una vacuna eficaz contra el VIH es mucho más complicada que la de una vacuna para contrarrestar los efectos de virus más estables como el

de viruela o el de la poliomielitis.

Es claro de lo expuesto en esta sección que la evolución del VIH corresponde a la dinámica de una intrincada red de procesos interdependientes. En este trabajo aislaremos uno de estos procesos, al que llamaremos ciclo reproductivo, con la expectativa de que algunos rasgos esenciales estén condicionados por él. Consideraremos como ciclo reproductivo del retrovirus al proceso que da inicio con la incisión del ARN viral en la célula huésped y finaliza con la expulsión del virus por parte de dicha célula. Podemos identificar dos pasos en este ciclo: uno de multiplicación durante el cual no ocurren mutaciones y otro cuando la reversotranscriptasa convierte el ARN en ADN en el cual existe una alta probabilidad de ocurrencia de mutaciones. Estos dos procesos se dan en tiempos y lugares diferentes, lo cual como se verá más adelante permitirá la implementación de una red de mapeos acoplados para modelar su evolución.

#### 44 CAPITULO 4. EL SISTEMA INMUNE Y EL VIRUS DEL SIDA

... ..

HIV infecta a las células T CD4+ y a las células T CD8+ del sistema inmune. El sistema inmune responde a la infección de estas células T, pero cuando la concentración de estas células T disminuye por debajo de un nivel crítico, el sistema inmune no puede responder adecuadamente a la infección por el virus del SIDA. Este nivel crítico de concentración de células T CD4+ se conoce como el nivel crítico de células T CD4+. Cuando la concentración de células T CD4+ cae por debajo de este nivel crítico, el sistema inmune no puede responder adecuadamente a la infección por el virus del SIDA. Esto puede dar lugar a una serie de problemas de salud, como la neumonía, la tuberculosis y el cáncer. El nivel crítico de células T CD4+ puede variar entre individuos, pero generalmente se sitúa entre 350 y 500 células T CD4+ por milímetro cúbico de sangre. Cuando la concentración de células T CD4+ cae por debajo de este nivel crítico, el sistema inmune no puede responder adecuadamente a la infección por el virus del SIDA. Esto puede dar lugar a una serie de problemas de salud, como la neumonía, la tuberculosis y el cáncer.

## Capítulo 5

# EL MODELO

Este trabajo tiene como objetivo principal el estudio de la evolución de la secuencia genética del virus causante del **SIDA**, haciendo uso de un modelo de *redes de mapeos acoplados*. Desde el punto de vista biológico, estamos interesados principalmente en relacionar el concepto de adecuación evolutiva con constricciones fisicoquímicas de los ácidos nucleicos y con los procesos moleculares que se presentan a nivel del ribosoma. Hasta ahora nos hemos concentrado principalmente en describir características importantes tanto de los ácidos como de algunos procesos a nivel molecular. En este momento estamos en posibilidad de construir el modelo tomando como base dichas características.

Desde el punto de vista de sistemas complejos, el problema se desarrolla en un espacio discreto que corresponde a las posibles composiciones químicas de las secuencias genéticas y está sujeto a una dinámica no lineal (proceso de reproducción celular) en la cual hay un acoplamiento espacial que corresponde a las mutaciones del material genético. En la sección 6.1 vamos a describir con un poco más de detalle este espacio de composiciones químicas para en la sección 6.2 proceder a construir el modelo.

## 5.1 EL ESPACIO DE SECUENCIAS.

El estudio de la evolución de secuencias genéticas de  $m$  nucleótidos se hace por lo general en el espacio de secuencias el cual, debido a que existen cuatro diferentes nucleótidos, será  $4^m$  dimensional y la evolución en este espacio sigue un criterio de adecuación. Sin embargo en la mayoría de estos modelos no se toman en cuenta restricciones asociadas a características estructurales y termodinámicas de los ácidos nucleicos.

Como ha sido acertadamente señalado por algunos autores [43] el punto crucial en la formulación de estos modelos es la definición y la medición de la adecuación de un individuo sin recurrir al resultado de la selección, por lo que es deseable construir un criterio de adecuación basado en características físicas y químicas así como en aspectos de los mecanismos de los procesos biológicos involucrados. En este trabajo se da especial atención a los procesos que tienen lugar en el ribosoma. Aquellas secuencias que optimicen la síntesis de proteínas en el ribosoma favorecerán un alto grado de adecuación.

En 1931 Wright [44] introdujo el concepto de *espacio de posibles genotipos*. A cada genotipo se le puede asignar una adecuación y la distribución de estos valores de adecuación sobre el espacio de genotipos constituye una *superficie de adecuaciones*. El proceso de evolución o de adecuación se describe como un proceso de maximización o minimización en dicha superficie. Dependiendo de la distribución de valores, esta superficie puede ser más o menos rugosa y la introducción de constricciones puede alterar su rugosidad. En nuestro caso las constricciones internas y externas mencionadas anteriormente tienen el efecto de suavizar la superficie.

Si consideramos que las unidades fundamentales de las secuencias genéticas son los codones, cada secuencia de  $m$  nucleótidos estará caracterizada por los  $\mathcal{L} = \frac{m}{3}$  codones que la forman. El proceso evolutivo se visualiza entonces en términos de las mutaciones que ocurren en secuencias vecinas. Diremos que dos secuencias son vecinas cercanas si su composición difiere únicamente en un codón. En esta descripción no se toma en cuenta la posición de los diferentes codones dentro de la cadena

debido a que para bacterias y virus las correlaciones entre codones son pequeñas [17], es decir, en primera aproximación la frecuencia con que un tipo de codón aparece en la secuencia en una posición dada es una variable estadísticamente independiente.

Hemos visto ya que el lenguaje adecuado para tomar en cuenta las restricciones termodinámicas de los ácidos nucleicos, es aquel en el que solo se toma en cuenta el carácter débil (D) o fuerte (F) de las bases, por lo tanto la composición de una secuencia dada será un vector en un espacio de 8 dimensiones  $\mathbf{i} = (i_{DDD}, \dots, i_{FFF})$ , donde  $i_{\alpha\beta\gamma}$  ( $\alpha, \beta, \gamma = D, F$ ) es el número de codones de cada uno de los ocho tipos que tenga dicha secuencia.

Si tomamos en cuenta el hecho de que:

$$\sum_{\alpha, \beta, \gamma = D, F} i_{\alpha\beta\gamma} = \mathcal{L} \quad (5.1)$$

el espacio de configuraciones es una red inmersa en un hipertetraedro de siete dimensiones y está definido por  $i_{DDD} + \dots + i_{FFF} = \mathcal{L}$  donde cada  $i_{\alpha\beta\gamma}$  toma únicamente valores enteros. Es fácil ver que si solamente se tienen dos tipos de codones (por ejemplo DDD y DDF), el espacio de configuraciones está definido por  $i_{DDD} + i_{DDF} = \mathcal{L}$ , esto es, una cadena lineal (discretización de un tetraedro unidimensional). Si se permiten tres tipos de codones por ejemplo DDD, DDF y DFF), el espacio de configuraciones está determinado por  $i_{DDD} + i_{DDF} + i_{DFF} = \mathcal{L}$  y consiste en la discretización de la superficie de un triángulo (tetraedro bidimensional), etc.

Denotemos por  $N_i(t)$  el número de secuencias de  $\mathcal{L}$  codones que al tiempo  $t$  tienen composición  $\mathbf{i}$ . Al permitir la ocurrencia de mutaciones debemos considerar las diferentes probabilidades de mutación aceptadas  $p_X$ ,  $p_Y$  y  $p_Z$  introducidas en la sección 3.2, y que satisfacen las desigualdades:  $p_Z > p_X > p_Y$ . Si denotamos por  $n_{XYZ}$  el número de codones de tipo  $XYZ$  en una secuencia de codones, dado que la probabilidad de mutación aceptada en la segunda posición del codón es mucho menor que en la tercera posición, como una primera aproximación podemos suponer que la dinámica debida a las mutaciones no cambia

$n_{NDN}$  ni  $n_{NFN}$ . Podemos entonces fijarnos en cualquiera de los dos grupos de codones, ya sea  $NDN$  o  $NFN$ , y proponer un modelo. Para los fines de este trabajo vamos a considerar codones del grupo  $DDN$  y denotamos por  $L$  el número de codones de este tipo que contenga la secuencia genética en cuestión. En ese caso las secuencias estarán etiquetadas por un solo parámetro, a saber el número de codones ya sea de tipo  $DDF$  o de tipo  $DDD$ , por ejemplo  $i = i_{DDF} = i$ , ya que  $i_{DDD} + i_{DDF} = L$ .

Aquí no tiene sentido utilizar condiciones a la frontera periódicas, ya que una secuencia con  $L$  codones del tipo  $DDF$  no es vecina cercana de una que tenga cero codones del tipo  $DDF$  (o bien  $L$  codones del tipo  $DDD$ ) y viceversa. En este caso particular, el número de posibles composiciones  $Q$  (o equivalentemente de sitios en la red) es igual al  $L + 1$ , donde  $L$  es el número de codones en una secuencia genética. Es importante hacer notar que en general no se tiene una relación tan sencilla entre estos dos parámetros.

## 5.2 CONSTRUCCION DEL MODELO

En la sección 4.3 vimos que aunque el virus VIH puede infectar varios tipos de células, ataca principalmente a los linfocitos **T4**, por lo que en primer término vamos a plantear el modelo para el caso de una célula huésped y posteriormente se propondrá una generalización al caso de varias células huésped.

En este punto es necesario incorporar la información que se da en el primer capítulo referente a los procesos de multiplicación y mutación. Como se recordará, en células eucariotes la multiplicación y las mutaciones se dan simultáneamente, mientras que para los retrovirus, la multiplicación y las mutaciones tienen lugar en tiempos y lugares diferentes. A lo largo de varios ciclos de replicación estos dos fenómenos se dan en forma alternada, permitiendo la implementación natural de una dinámica de mapeo acoplado en dos pasos.

Denotemos por  $N_i(t)$  el número de secuencias genéticas que al tiempo  $t$  tienen composición  $i$ . Debido a la competencia entre las células infectadas la dinámica de  $N_i(t)$  está limitada por una restricción ecológica dada por:

$$f(\{N_i(t)\}) = A_i N_i(t) - c N_i^2(t) - d N_i \bar{N}(t) \quad (5.2)$$

donde

$$\bar{N}(t) = \frac{1}{L+1} \sum_{j=0}^L N_j(t) \quad (5.3)$$

y  $A_i$  es un factor de crecimiento.

Si únicamente se permiten mutaciones puntuales en la posición  $Z$  del codón, las cuales ocurren con probabilidad  $p_Z = p \neq 0$ , la ecuación de-evolución para  $N_i(t)$  se puede escribir como:

$$N_i(t+1) = f\{N_i(t)\} + p \left[ C_{i+1}^1 f\{N_{i+1}(t)\} + C_{L-(i-1)}^1 f\{N_{i-1}(t)\} - (C_i^1 + C_{L-i}^1) f\{N_i(t)\} \right] \quad (5.4)$$

donde:

$$C_r^s = \frac{r!}{s!(r-s)!} \quad (5.5)$$

es el número de formas en que los  $r$  codones del tipo DDF de una secuencia pueden sufrir  $s$  mutaciones,  $p$  es la razón de mutación por base y por ciclo,  $i_{DDF} + i_{DDD} = L$  e  $i_{DDF} = i$ . El superíndice en  $f\{N_i(t)\}$  indica que el proceso de reproducción puede repetirse  $n$  veces antes de que el virus salga de la célula, en el análisis de la dinámica del modelo vamos a limitarnos al caso  $n = 1$ .

Introduciendo la forma explícita de los coeficientes  $C_r^s$  la ecuación (4.6) se puede escribir como (con  $n = 1$ ):

$$N_i(t+1) = f\{N_i(t)\} + p \left[ (i+1) f\{N_{i+1}(t)\} + (L+1-i) f\{N_{i-1}(t)\} - L f\{N_i(t)\} \right] \quad (5.6)$$



Si  $c \neq 0$  y  $d \neq 0$  diremos que el sistema evoluciona sujeto a una *constricción ecológica genérica*. Si  $c \neq 0$  y  $d = 0$  nos referiremos a una *constricción ecológica local* y si  $c = 0$  y  $d \neq 0$  hablaremos de una *constricción ecológica global*.

Este modelo puede generalizarse al caso en que el virus del SIDA invade diversos tipos de células y esta invasión depende del tipo de célula de que se trate. En vista de esta situación el problema se puede tratar como un problema de dinámica de poblaciones en donde se tienen diversos tipos de células y diversos tipos de secuencias que buscan infectar a dichas células. Denotemos por  $N_{i,\gamma}(t)$  el número de células de tipo  $\gamma$  que al tiempo  $t$  están infectadas por el  $i$ -ésimo tipo de secuencia genética. Debido a la competencia entre los diversos tipos de células, es necesario introducir una restricción ecológica que controle el aumento de la población de células. Cada célula compite con células no solamente de su tipo sino de todos los tipos presentes por lo tanto el crecimiento de la población de células limitado por esta constricción ecológica está dado por:

$$f(\{N_{i,\gamma}(t)\}) = A_{i,\gamma}N_{i,\gamma}(t) - c_{\gamma}N_{i,\gamma}^2(t) - d_{\gamma}N_{i,\gamma}(t)\bar{N}_{\gamma}(t)g_{\gamma}N_{i,\gamma}(t)\bar{N}_i(t) \quad (5.7)$$

donde

$$\begin{aligned} \bar{N}_{\gamma}(t) &= \frac{1}{L+1} \sum_{i=0}^L N_{i,\gamma}(t) \\ \bar{N}_i(t) &= \frac{1}{\mathcal{M}} \sum_{\gamma=1}^{\mathcal{M}} N_{i,\gamma}(t) \end{aligned} \quad (5.8)$$

con  $\mathcal{M}$  el número total de tipos diferentes de células y  $A_{i,\gamma}$  un factor de crecimiento.

Además para cada tipo de célula se tendrá una cierta probabilidad  $u_{\gamma}$  de producir más virus y una probabilidad  $v_{\gamma}$  de ser infectada, por lo que la ecuación de evolución para  $N_{i,\gamma}(t)$  estará dada por:

$$N_{i,\gamma}(t+1) = f(\{N_{i,\gamma}(t)\}) + v_{\gamma} \sum_{\xi \neq \gamma} u_{\xi} f(\{N_{i,\xi}(t)\}) +$$

$$v_i \sum_{\xi=1}^M u_{\xi} p \left[ C_{i+1}^1 f(\{N_{i+1,\xi}(t)\}) + C_{L-(i-1)}^1 f(\{N_{i-1,\xi}(t)\}) - (C_i^1 + C_{L-i}^1) f(\{N_{i,\xi}(t)\}) \right] \quad (5.9)$$

Este sería un modelo más realista del problema pero debido a su complejidad sería difícil de analizar. Como primera aproximación, en este trabajo únicamente se analizará al modelo dado por las ecuaciones (5.6).

Reordenando términos en la ecuación (5.6), se obtiene:

$$N_{i+1}(t) = \left[ 1 + \frac{2\epsilon}{L+2} \right] f\{N_i(t)\} + \mathcal{A} + \mathcal{B} \quad (5.10)$$

donde:

$$\mathcal{A} = \frac{\epsilon}{2} [f\{N_{i+1}(t)\} - 2f\{N_i(t)\} + f\{N_{i-1}(t)\}] \quad (5.11)$$

$$\mathcal{B} = \frac{\epsilon(2i-L)}{2(L+2)} [f\{N_{i+1}(t)\} - f\{N_{i-1}(t)\}] \quad (5.12)$$

y  $\epsilon = p(L+2)$ .

Estas ecuaciones muestran que la ecuación de evolución obtenida es del tipo *reacción-difusión*. En la ecuación (5.10) el primer término es un término fuente no lineal,  $\mathcal{A}$  es un término difusivo discretizado y  $\mathcal{B}$  un término tipo gradiente discreto. Esta ecuación presenta la peculiaridad de que el coeficiente de difusión  $\frac{\epsilon}{2}$  está relacionado también con los términos fuente y gradiente.

Si se relajan algunas de las restricciones que conducen a la ecuación (5.10) se obtendrá una jerarquía de ecuaciones. Por ejemplo si se permite la ocurrencia simultánea de varias mutaciones, esta ecuación pierde su carácter local. Por otra parte si se considera que más de una de las  $p_s \neq 0$  tanto el espacio de composiciones como las ecuaciones de evolución resultantes tendrán una estructura más compleja.

Evidentemente sería deseable que cualquier modelo construido para estudiar un problema biológico fuera lo suficientemente completo para

describir el problema sin inconvenientes, sin embargo eso es muy difícil de lograr. En modelos del tipo del que estamos planteando en este trabajo generalmente tiene el problema de que en algún momento las poblaciones resultan ser negativas, sobre todo en rangos de parámetros en donde o bien la no linealidad o el acoplamiento son fuertes. Por esta razón y con el fin de garantizar la estabilidad del sistema, se introduce una función tipo Heaviside  $\mathbf{H}(w)$  (ec. (2.9)), de tal forma que el mapeo acoplado que vamos a analizar está dado por la ecuación:

$$N_{i+1}(t) = \mathbf{H} \left( \left[ 1 + \frac{2\epsilon}{L+2} \right] f\{N_i(t)\} + \mathcal{A} + \mathcal{B} \right) \quad (5.13)$$

con  $\mathcal{A}$  y  $\mathcal{B}$  dadas por las ecuaciones (5.11) y (5.12). Cabe señalar que se hizo un esfuerzo por evitar la introducción de esta función pero a la larga resultó ser más complicado y los tiempos de cómputo requeridos aumentaron drásticamente. Es importante mencionar en este punto que esta función se utilizará en todos los casos durante la realización de las simulaciones numéricas.

Existen evidencias [45] en estudios de redes de mapeos acoplados relacionados con dinámica de poblaciones de que la introducción de este tipo de restricciones pueden conducir a comportamientos disímiles. En nuestro caso, para el intervalo de parámetros en donde se tiene una dinámica de punto fijo, el hecho de introducir esta función es totalmente irrelevante. Sin embargo, es importante hacer notar que, el haberla introducido generará gran parte de la dinámica del mapeo acoplado.

Nótese que la ecuación (5.10) define un mapeo globalmente acoplado del tipo de los introducidos en la sección 2.1, sin embargo en este caso los coeficientes tienen una dependencia "espacial" explícita y se contempla una contribución local asociada a un concepto de "distancia" entre las secuencias genéticas.

Consideremos ahora el caso en que se permite que varias mutaciones ocurran simultáneamente. Al igual que en el caso anterior nos limitaremos a mutaciones puntuales que ocurran en la base en la posición  $Z$  del codón. Dado que la probabilidad de que ocurra una mutación es  $p$ , la probabilidad de ocurrencia de  $k$  mutaciones simultáneas será  $p^k$ .

Si se permite que ocurran hasta tres mutaciones simultáneas,  $N_i(t+1)$  estará dada por:

$$\begin{aligned}
 N_i(t+1) = & f\{N_i(t)\} + p \left[ C_{i+1}^1 f\{N_{i+1}(t)\} + \right. \\
 & C_{L-(i-1)}^1 f\{N_{i-1}(t)\} - (C_i^1 + C_{L-i}^1) f\{N_i(t)\} \left. \right] + \\
 & p^2 \left[ C_{i+2}^2 f\{N_{i+2}(t)\} + C_{L-(i-2)}^2 f\{N_{i-2}(t)\} - \right. \\
 & (C_i^2 + C_{L-i}^2) f\{N_i(t)\} \left. \right] + p^3 \left[ C_{i+3}^3 f\{N_{i+3}(t)\} + \right. \\
 & C_{L-(i-3)}^3 f\{N_{i-3}(t)\} + (C_{i+1}^2 C_{L-(i+1)}^1 f\{N_{i+1}(t)\} + \\
 & C_{i-1}^1 C_{L-(i-1)}^2 f\{N_{i-1}(t)\}) - C_L^3 f\{N_i(t)\} \left. \right] \quad (5.14)
 \end{aligned}$$

en donde el término independiente de  $p$  y el lineal en  $p$  son los incluidos en la ecuación (5.4). Los términos cuadráticos y cúbicos en  $p$  son la contribución debida a la ocurrencia de dos y tres mutaciones respectivamente y los coeficientes  $C_i^k$  están dados por la ecuación (5.5).

En general se puede ver que si se permiten hasta  $m$  mutaciones simultáneas,  $N_i(t+1)$  estará dada por:

$$N_i(t+1) = \left( f\{N_i(t)\} + \sum_{k=1}^m p^k \left[ \sum_{\{r\}} \left( C_{i+r}^{\frac{k+r}{2}} C_{L-i-r}^{\frac{k-r}{2}} f\{N_{i+r}(t)\} + \right. \right. \right. \\
 \left. \left. \left. C_{i-r}^{\frac{k-r}{2}} C_{L-i+r}^{\frac{k+r}{2}} f\{N_{i-r}(t)\} \right) - \mathcal{R} f\{N_i(t)\} \right] \right) \quad (5.15)$$

donde  $\{r\} = \{2, 4, \dots, k\}$  y  $\mathcal{R} = \sum_{l=1}^{\frac{k}{2}-1} (C_i^{k-l} C_{L-l}^l + C_l^l C_{L-l}^{k-l})$  si  $k$  es par y  $\{r\} = \{1, 3, \dots, k\}$  y  $\mathcal{R} = C_L^k$  si  $k$  es impar.

De esta última ecuación es claro que el alcance de la interacción aumenta a medida que aumenta el valor de  $m$ , hasta que la interacción se convierte en global cuando  $m = L$ . Esta forma de pasar de un sistema con interacciones locales o de corto alcance a uno con interacciones

globales, nos hace pensar en un modelo tipo *campo medio* como los utilizados en física estadística para estudiar sistemas con interacciones muy complicadas como es el caso de fluidos, ferromagnetos, etc. En un modelo de campo medio, la suposición esencial es que el alcance de las interacciones se supone muy largo de tal forma cada uno de los componentes del sistema interactúa con un campo medio producido por el resto de los elementos.

En nuestro problema las interacciones están relacionadas con las mutaciones y por lo tanto tenemos dos formas naturales de construir un modelo tipo campo medio dependiendo de si las mutaciones ocurren simultáneamente con la reproducción o en forma alternada. Estos dos modelos estarán dados por:

$$N_i(t+1) = f\{N_i(t)\} + \frac{\alpha}{(L+1)} \sum_{i=0}^L N_i(t) \quad (5.16)$$

$$N_i(t+1) = f\{N_i(t)\} + \frac{\beta}{(L+1)} \sum_{i=0}^L f\{N_i(t)\} \quad (5.17)$$

respectivamente, donde  $\alpha$  y  $\beta$  son los parámetros que miden la intensidad del correspondiente campo medio. Estos modelos corresponden al mapeo globalmente acoplado definido por la ecuación (2.4).

En el siguiente capítulo se analiza la dinámica del modelo descrito por la ecuación (5.13) así como de su extensión de largo alcance dada por la ecuación (5.15). En el capítulo 7 se analizará la dinámica de los modelos descritos por las ecuaciones (5.16) y (5.17), su relación con los modelos anteriores y sus implicaciones en el contexto biológico que nos interesa.

## Capítulo 6

# DINAMICA DEL MODELO

## 6.1 SECUENCIAS QUE CODIFICAN UN SOLO GEN

### 6.1.1 ACOPLAMIENTO DE CORTO ALCANCE

Para determinar completamente el modelo sólo nos resta especificar la forma funcional del factor de crecimiento  $A_i$ , la cual está determinada por los procesos que tienen lugar en el ribosoma. G. Cocho y G. Martínez-Mekler realizaron un estudio fenomenológico [17] en el que se propone que en el caso de retrovirus (como el del SIDA)  $A_i$  debe ser una función monótona no decreciente de  $i$ . Como se mencionó anteriormente, estudios experimentales de propiedades termodinámicas y el análisis de secuencias genéticas del Gene-Bank, sugieren que los codones del tipo NDF son más eficientes y más exactos y que este comportamiento puede asociarse con las interacciones tARN-mARN y rARN-mARN [17], [31], [32], [35], [36]. Cabe hacer notar que los resultados experimentales corresponden a un régimen en el que las mutaciones son despreciables, es decir  $p \approx 0$ . Al considerar únicamente codones de los tipos DDF y DDD, en ausencia de mutaciones, aquellas secuencias con un mayor contenido de codones DDF deberán tener una razón de crecimiento mayor. Dado que  $i$  es una medida de la concentración de codones DDF y  $A_i$  es la razón de crecimiento de las cadenas

cuya composición es  $i$ , una elección razonable es que  $A_i$  sea una función monótona creciente de  $i$ . En este sentido la forma funcional de  $A_i$  refleja características biológicas y no se introduce como una suposición, sin embargo la forma particular de esta función se elige arbitrariamente. En nuestros cálculos:

$$A_i = a + b \left( \frac{i}{L} \right) \quad (6.1)$$

en donde  $a$  y  $b$  son constantes positivas y la eficiencia y la exactitud mencionadas anteriormente crecen al aumentar el parámetro  $b$ .

Consideremos ahora el comportamiento de la ecuación (5.13), con la restricción ecológica dada por  $f\{N_i(t)\}$  (ecuación (5.2)) y  $A_i$  por (6.1), suponiendo que hay mutaciones pero solamente se permite una mutación por base y por ciclo. El análisis de la dinámica se llevó a cabo por medio de simulaciones numéricas por computadora. A continuación se encuentran algunas de las gráficas obtenidas.

En la Figura 6.1, se muestra la distribución  $N_i(t)$  sujeta a una restricción ecológica global ( $c = 0, d \neq 0$ ) para varios valores del parámetro  $\epsilon$ .  $N_i(t)$  se grafica como función de  $i$  durante 300 iteraciones consecutivas después de un transitorio de 2000 iteraciones y en todos los casos la condición inicial se toma al azar. El parámetro  $\epsilon$  toma valores del orden de 1 y  $L = 500$ . Cabe señalar que los valores de estos parámetros están relacionados con la proteína externa **gp120** del virus VIH-1 del SIDA, la cual tiene una longitud aproximada de 1500 bases (500 codones) y presenta una razón de mutaciones  $p$  del orden de  $10^{-3}$ . Esta región codifica la proteína env del virus y es la región más activa en su interacción con la célula huésped. Dado que hemos supuesto una sola mutación por base y por ciclo,  $\epsilon = p(L + 2)$  será del orden de 1. Los parámetros  $a$  y  $b$  cumplen con la condición necesaria para tener una dinámica de punto fijo ( $a + b$ ) < 3. Esta condición se obtiene del análisis de la evolución temporal del promedio de todos los mapeos  $\bar{N}(t)$ . Si  $c = 0$   $\bar{N}(t)$  se puede expresar como:

$$\bar{N}(t+1) = \frac{1}{L+1} \sum_{i=0}^L N_i(t+1) = (a - d\bar{N}(t))\bar{N}(t) + \frac{1}{L+1} \sum_{i=0}^L b \left( \frac{i}{L} \right) N_i(t) \quad (6.2)$$

y como  $0 \leq i \leq L$ , entonces:

$$(a - d\bar{N}(t))\bar{N}(t) \leq \bar{N}(t+1) \leq (a+b) - d\bar{N}(t)\bar{N}(t) \quad (6.3)$$

Haciendo  $\bar{M}(t) = d\bar{N}(t)$ , se tiene que:

$$a\bar{M}(t) - \bar{M}^2(t) \leq \bar{M}(t+1) \leq (a+b)\bar{M}(t) - \bar{M}^2(t) \quad (6.4)$$

es decir  $\bar{M}(t)$  está acotada entre dos mapeos logísticos. El mapeo de la derecha presenta una dinámica de punto fijo siempre y cuando  $(a+b) < 3$ , por lo que una condición necesaria para que  $\bar{M}(t)$  (y  $\bar{N}(t)$ ) tenga una dinámica de punto fijo es que  $(a+b) < 3$  (ver sección 6.3.1).

En la Figura 6.2 se muestran algunas gráficas de  $N_i(t)$  para varias condiciones iniciales, como función de  $i$  para los mismos valores de  $a, c$  y  $d$  usados en las gráficas de la Figura 6.1 con  $b = 0$ ,  $L = 10$  y un valor especial de  $\epsilon = \epsilon^*$ , que en este caso es  $\epsilon^* = 1.2$ . El efecto mostrado en esta gráfica se observa para cualquier  $L$  si se elige el valor adecuado  $\epsilon^*(L)$ . Por mayor claridad se eligió  $L = 10$ .

En la Figura 6.3 se muestran las gráficas correspondientes al caso en que todas las especies tienen el mismo factor de crecimiento (es decir  $b = 0$ ), para dos valores diferentes de  $\epsilon$  en el intervalo  $[0, 1]$ . Los valores de los demás parámetros son  $a = 1.5$ ,  $c = 0$ ,  $d = 1$  y  $L = 500$ .

Las simulaciones numéricas muestran las siguientes características:

a).- Para  $\epsilon = 0$  la única especie que sobrevive es aquella con  $i = L$  y a medida que se incrementa el valor de  $\epsilon$  se forma una distribución de secuencias (Figura 6.1). Si  $\epsilon$  está en el intervalo  $0 < \epsilon < \epsilon^* = \frac{L+2}{L}$  la distribución muestra una deslocalización que puede medirse ya sea en términos de su desviación estandar o de su soporte y que crece al aumentar  $\epsilon$ . También se observa que el máximo de la distribución se recorre hacia valores menores de  $i$  a medida que aumenta el valor de  $\epsilon$ . En la sección 7.3 nos referiremos nuevamente al valor de  $\epsilon^*$ .

b).- Cuando  $\epsilon = \epsilon^* = \frac{L+2}{L}$  se pierde la dinámica de punto fijo y surge un comportamiento alternante de período dos que muestra multiestabilidad, es decir, el estado asintótico depende de la condición inicial



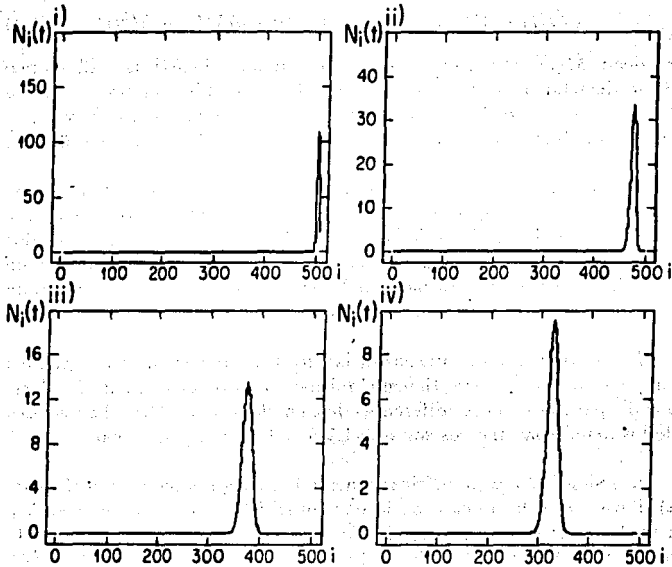


Figura 6.1: Efecto de la variación del parámetro  $\epsilon$  sobre la distribución de secuencias genéticas. Se grafica el número  $N_i(t)$  de secuencias de  $L$  codones, que al tiempo  $t$  tienen composición  $i$ , como función de  $i$  en 300 iteraciones consecutivas después de un transitorio de 2000 pasos temporales. Los parámetros  $a$ ,  $b$ ,  $c$ ,  $d$  y  $L$  se mantienen constantes ( $a = 0.5$ ,  $b = 1.5$ ,  $c = 0$ ,  $d = 1$  y  $L = 500$ ). El valor de  $\epsilon$  es: i).- 0.01, ii).- 0.10, iii).- 0.50 y iv).- 1.00.

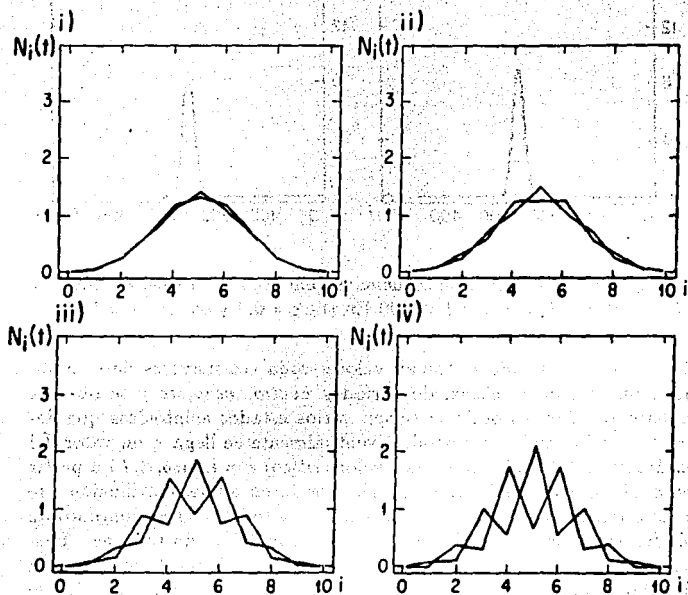


Figura 6.2: Dinámica alternante de la distribución de secuencias en donde puede apreciarse la multiestabilidad encontrada en el punto de bifurcación  $\epsilon^* = \frac{(L+2)}{L}$ . Las gráficas corresponden a los mismos valores de los parámetros ( $a = 1.5$ ,  $b = 0$ ,  $c = 0$ ,  $d = 1$ ,  $\epsilon = 1.2$  y  $L = 10$ ) y diferentes condiciones iniciales. En las gráficas se muestran 300 iteraciones consecutivas después de un transitorio de 1000.

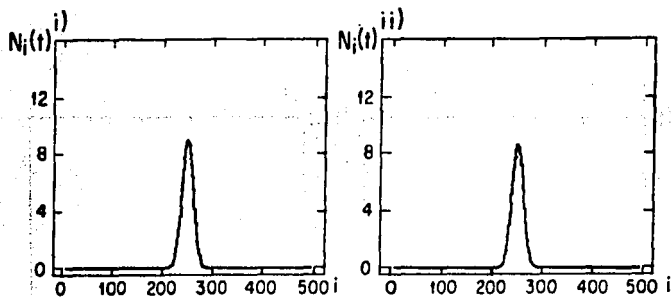


Figura 6.3: Distribución de secuencias genéticas en el caso en que  $a = 1.5$ ,  $b = 0$ ,  $c = 0$ ,  $d = 1$  y  $L = 500$  En i).-  $\epsilon = 0.5$  y en ii).-  $\epsilon = 1.0$ .

(Figura 6.2). Cuando se toman valores cada vez mayores de  $\epsilon$  se obtiene una dinámica cíclica de períodos *cuatro*, *ocho*, etc y se observa multiestabilidad, es decir se tienen varios estados asintóticos que dependen de la condición inicial. Eventualmente se llega a un valor del parámetro  $\epsilon$  (al cual llamaremos valor crítico)  $\epsilon = \epsilon_c(a, b, d, L)$  a partir del cual, algunas condiciones iniciales conducen a una distribución que se colapsa, es decir  $N_i(t) = 0$  para toda  $i$  y todo  $t$ . Otro conjunto de condiciones iniciales conducen a estados asintóticos no triviales. Posteriormente se encuentra un valor de  $\epsilon$  a partir del cual la distribución se colapsa para todas las condiciones iniciales. En la sección 7.3 se analiza con mayor detalle la cascada de bifurcaciones, la cual da lugar a la aparición de caos y eventualmente al colapso.

c).- Cuando todas las especies tienen el mismo factor de crecimiento (es decir  $b = 0$ ), se obtiene una distribución simétrica con un máximo en  $i = \frac{L}{2}$  si  $L$  es par o en  $i = \frac{L-1}{2}$  si  $L$  es impar. La distribución es independiente del valor de  $\epsilon$  si  $0 < \epsilon < \frac{L+2}{L}$  (Figura 6.3). Este comportamiento puede corroborarse analíticamente (ver sección 7.3).

La característica a).- indica que en ausencia de mutaciones, la especie que tiene el factor de crecimiento mayor es la que domina. Este resultado puede obtenerse analíticamente y es consecuencia directa del criterio con el cual se eligió  $A_i$ . La forma de la distribución cuando  $b = 0$  indica que en presencia de mutaciones (es decir  $\epsilon \neq 0$ ), aún cuando todas las especies tengan el mismo factor de crecimiento, no todas se reproducen de igual forma.

Estos resultados pueden reexaminarse en términos de la teoría de la evolución de Darwin basada en la sobrevivencia de la especie que mejor se adapta. En algunos trabajos [6], [43] se argumenta que en el problema de crecimiento de una población en un medio ambiente restringido, la continua generación de mutantes produce un conjunto de variantes de la especie con mayor adecuación el cual constituye la fuente de los procesos de optimización y adaptación. De esta forma en lugar de tener una sola especie con mejor adecuación se genera una familia de especies muy parecidas entre sí, a esta familia se le denomina *cuasiespecie*. Algunos autores han señalado que la gran variedad de secuencias genéticas que se ha encontrado del virus VIH confirma que éste evoluciona como una cuasiespecie [46].

Por otra parte como menciona Schuster [43] la sugerencia de que mayores tasas de error producen más mutaciones que permitirán un proceso de optimización más eficiente es válida solo hasta cierto límite. Cuando las tasas de mutación son muy altas, las mutaciones dominan, la herencia deja de funcionar de tal forma que el proceso de selección no encuentra nada sobre que actuar. La razón de mutaciones a la cual ocurre la transición entre optimización eficiente y selección aleatoria, se denomina umbral de error. Como ha sido señalado en trabajos anteriores [47], los virus de ARN evolucionan bajo condiciones cercanas a este umbral. Nuestro modelo apoya este punto de vista, el umbral de error corresponde a  $\epsilon_c$  y depende de los parámetros que están directamente relacionados con mecanismos específicos de reproducción a nivel molecular.

Las gráficas de la Figura 6.4 se muestran el comportamiento de  $N_i(t)$

al variar el parámetro  $b$ . Los parámetros corresponden a una condición ecológica global y a un régimen de punto fijo. En estas gráficas se observa que a medida que  $b$  crece, la distribución se ensancha (aumenta su dispersión) y tanto el valor medio como el más probable decrecen. Como veremos más adelante las secuencias genéticas del virus del SIDA exhiben este comportamiento.

Anteriormente se mencionó que la forma particular del factor de crecimiento se eligió arbitrariamente, en [48] y [49] se ha utilizado una función con dependencia cuadrática en  $(i/L)$  y se encuentra que los resultados hasta ahora discutidos no se alteran significativamente.

### 6.1.2 ACOPLAMIENTO DE LARGO ALCANCE

Hasta ahora hemos considerado el caso en que sólo se permite una mutación por base y por ciclo. Como se discutió en la sección 5.1, cuando se permite un número arbitrario de mutaciones la dinámica está gobernada por la ecuación (5.15). En la Figura 6.5 se muestran algunas de las gráficas de  $N_i(t)$  para diferentes valores de  $m$  (número de mutaciones), las cuales se obtienen al iterar las ecuaciones (5.15). En general se observa que la forma de la distribución va cambiando y se estabiliza para  $m \approx 10$ . Este hecho puede entenderse si recordamos que la probabilidad de que ocurran  $m$  mutaciones es  $p^m$  y  $p$  es del orden de  $10^{-3}$ . Comparando dichas gráficas con las de las Figuras 6.1, y 6.4, se observa que la distribución de secuencias presenta un comportamiento similar sin embargo, como se verá más adelante, fuera del rango en donde se observa una dinámica de punto fijo se encuentran algunas diferencias.

En la Figura 6.6, se presentan las gráficas de  $\bar{N}(t)$  con  $L = 50$ , como función de  $\epsilon$  para  $m = 2, 4, 6$  y  $10$ . En estas gráficas puede observarse que  $\bar{N}(t)$  tiene un comportamiento más regular a medida que aumenta  $m$ . Si estas gráficas se construyen tomando varias condiciones iniciales para cada valor de  $\epsilon$  se observa la existencia de multiestabilidad en la zona en que  $1.2 < \epsilon < 2$ , por lo que la regularización del compor-

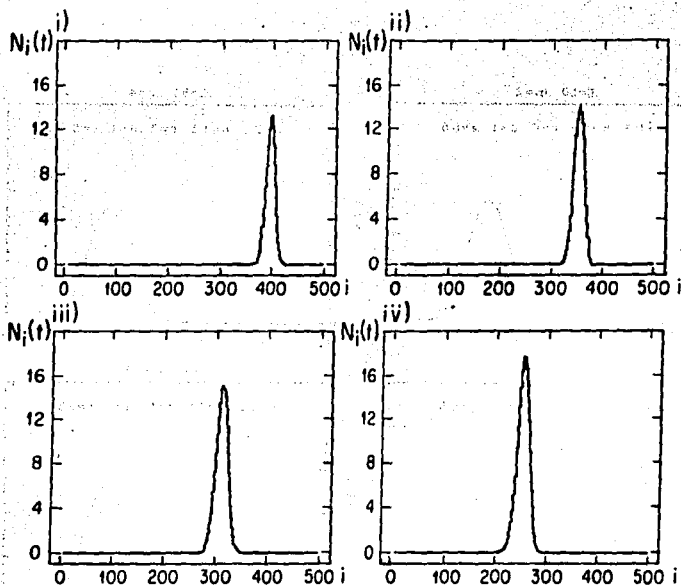


Figura 6.4: Efecto de la variación del parámetro  $b$  sobre la distribución de secuencias genéticas. Se grafican 300 iteraciones consecutivas después de un transitorio de 3000. Los parámetros  $c$ ,  $d$ ,  $\epsilon$  y  $L$  se mantienen constantes ( $c = 0$ ,  $d = 1$ ,  $\epsilon = 0.5$  y  $L = 500$ ). En i).-  $a = 0.005$ ,  $b = 1.995$ , ii).-  $a = 1.000$ ,  $b = 1.000$ , iii).-  $a = 1.500$ ,  $b = 0.500$ , iv).-  $a = 1.995$ ,  $b = 0.005$ .  $a$  y  $b$  satisfacen la condición de punto fijo  $a + b < 3$ .

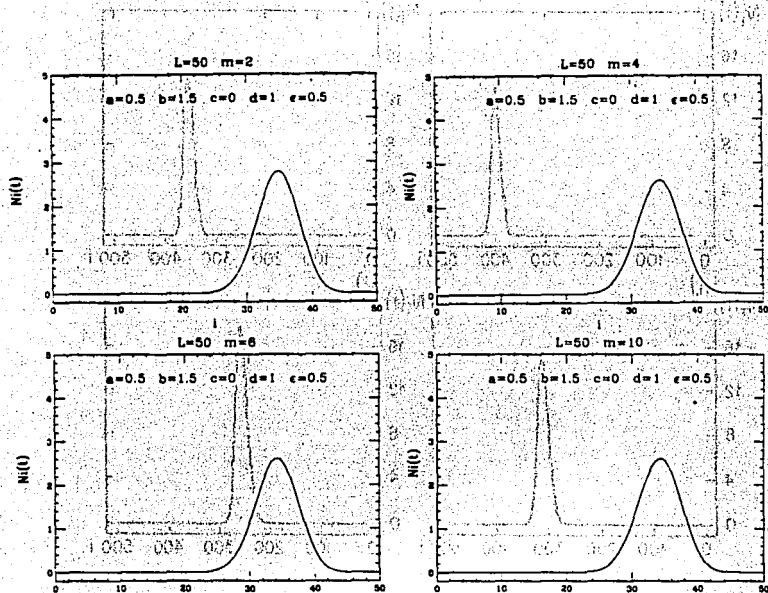


Figura 6.5: Distribución de secuencias genéticas para i).-  $m = 2$ , ii).-  $m = 4$ , iii).-  $m = 6$  y iv).-  $m = 10$  mutaciones, para una cadena de  $L = 50$  codones. Los valores de los parámetros son:  $a = 0.5$ ,  $b = 1.5$ ,  $c = 0$ ,  $d = 1$  y  $\epsilon = 0.5$ . Se grafican 100 iteraciones consecutivas después de un transitorio de 1000.

tamiento a medida que aumenta  $m$  es menos clara. En el caso en que  $m = 1$  esta regularización en el comportamiento se observó cuando aumenta el tamaño  $L$  de la red. En la Figura 6.7 se muestra el caso en que  $L$  aumenta de 50 a 500.

## 6.2 EXTENSION A K GENES.

Hasta aquí hemos considerado únicamente la evolución de secuencias que codifican un solo gen. Si la secuencia codifica dos genes uno con longitud  $L_1$  y otro con longitud  $L_2$ , la ecuación (5.4) se convierte en:

$$N_{i,j}(t+1) = f_{i,j} + p \left[ (i+1)f_{i+1,j} + (L_1+1-i)f_{i-1,j} + (j+1)f_{i,j+1} + (L_2+1-j)f_{i,j-1} - Lf_{i,j} \right] \quad (6.5)$$

donde:

$$f_{\gamma,\delta} = A_{\gamma,\delta} N_{\gamma,\delta}(t) - d N_{\gamma,\delta}(t) \bar{N}(t) \quad (6.6)$$

con:

$$\bar{N}(t) = \sum_{\xi,\eta=0}^{L_1,L_2} \frac{N_{\xi,\eta}(t)}{(L_1+1)(L_2+1)} \quad (6.7)$$

$L = L_1 + L_2$  y  $A_{\gamma,\delta}$  es el factor de crecimiento de  $N_{\gamma,\delta}$ . En estas expresiones siempre que se encuentre un doble subíndice se entenderá que el primero toma valores en el intervalo  $[0, L_1]$  y el segundo en el intervalo  $[0, L_2]$ .

Es posible obtener una solución factorizada de punto fijo (es decir  $N_{\gamma,\delta}(t+1) = N_{\gamma,\delta}(t) = N_{\gamma}(t)N_{\delta}(t)$ ) de la ecuación (6.5), suponiendo que  $A_{\gamma,\delta} = A_{\gamma}^{(1)} + A_{\delta}^{(2)}$  donde el superíndice se refiere a los genes 1 y 2. Esta última igualdad establece que la eficiencia del conjunto de dos genes es igual a la suma de las eficiencias de los dos genes tomados por separado. Si los coeficientes  $A_{\gamma}^{(1)}$  y  $A_{\delta}^{(2)}$  se substituyen por constantes efectivas determinadas por métodos variacionales o perturbativos, la



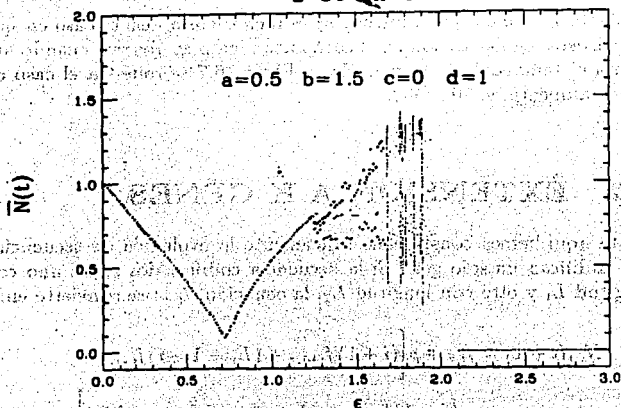
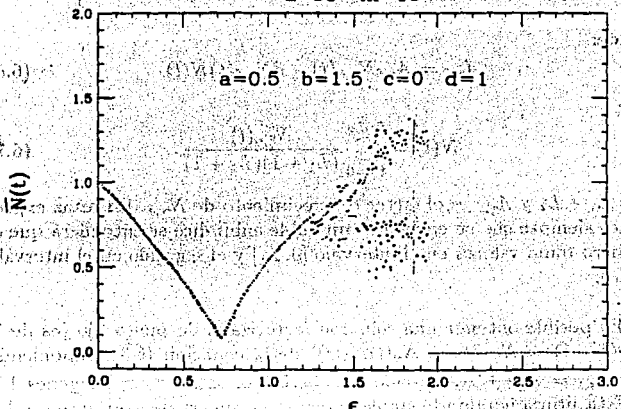
L=50  $\leftarrow$  m=6L=50  $\leftarrow$  m=10

Figura 6.6: (Cont.) Gráficas de  $\bar{N}(t)$  como función del parámetro  $\epsilon$  para i).-  $m = 6$  y iv).-  $m = 10$  mutaciones. Se grafican 100 iteraciones después de un transitorio de 3000. Los valores de los parámetros se muestran en las gráficas.

CAPITULO 6. DINAMICA DEL MODELO  
 $L=50$   $m=2$

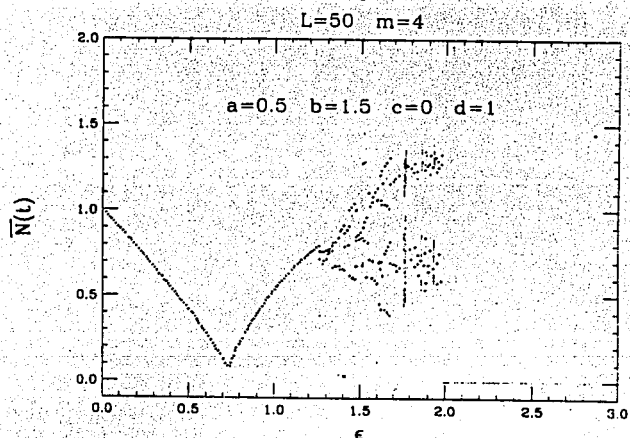
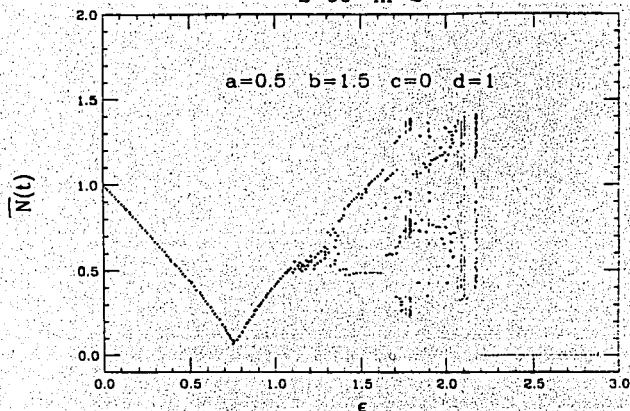
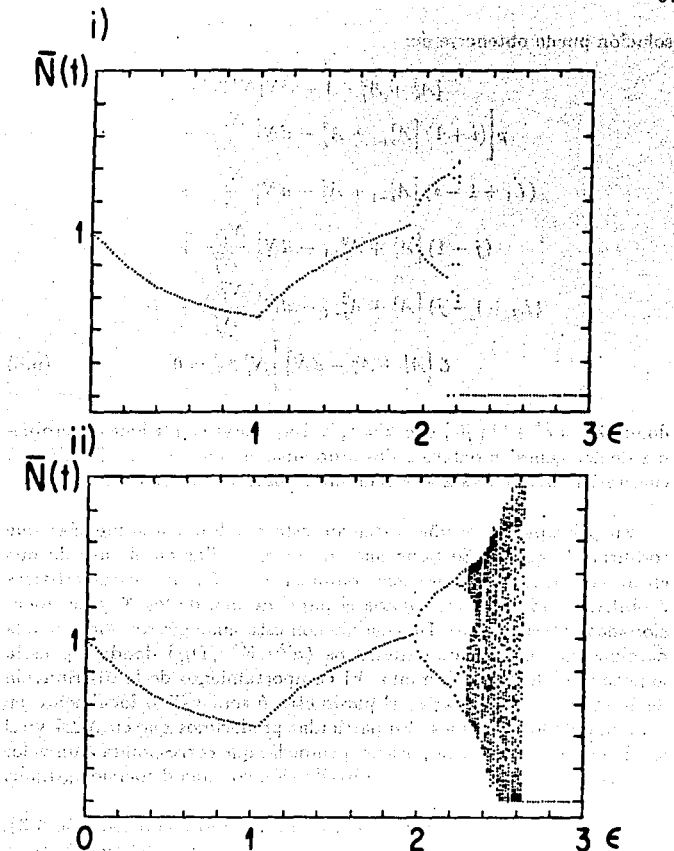


Figura 6.6: Gráficas de  $\bar{N}(t)$  como función del parámetro  $\epsilon$  para i).-  $m = 2$  y iv).-  $m = 4$  mutaciones. Se grafican 100 iteraciones después de un transitorio de 3000. Los valores de los parámetros se muestran en las gráficas.



solución puede obtenerse de:

$$\begin{aligned}
 & [A_i^1 + A_j^2 - 1 - d\bar{N}] N_i^1 N_j^2 + \\
 & p \left[ (i+1) [A_{i+1}^1 + A_j^2 - d\bar{N}] \frac{N_{i+1}^1}{N_i^1} + \right. \\
 & (L_1 + 1 - i) [A_{i-1}^1 + A_j^2 - d\bar{N}] \frac{N_{i-1}^1}{N_i^1} + \\
 & \left. (j-1) [A_i^1 + A_{j+1}^2 - d\bar{N}] \frac{N_{j+1}^2}{N_j^2} + \right. \\
 & \left. (L_2 + 1 - j) [A_i^1 + A_{j-1}^2 - d\bar{N}] \frac{N_{j-1}^2}{N_j^2} - \right. \\
 & \left. L [A_i^1 + A_j^2 - d\bar{N}] \right] N_i^1 N_j^2 = 0 \quad (6.8)
 \end{aligned}$$

donde  $A_k^k = a^k + b^k \left( \frac{\gamma}{L_k} \right)$  con  $k = 1, 2$ . Bajo estas condiciones el problema de dos genes se reduce a dos problemas de un gen y la descripción cualitativa dada hasta ahora será válida para cada uno de ellos.

En principio es posible extender este resultado a secuencias que codifican  $K$  genes. Se tiene una situación similar en el caso de que en un gen puedan identificarse dominios con diferentes características evolutivas, cada dominio jugaría el papel de uno de los  $K$  genes mencionados anteriormente. De acuerdo con esta analogía se asocia a cada dominio un conjunto de parámetros  $\{a^{(K)}, b^{(K)}, D_K\}$  donde  $D_K$  es la longitud del  $K$ -ésimo dominio. El comportamiento de la distribución de la cuasiespecie al variar el parámetro  $b$  será válido localmente en cada uno de los dominios. En particular predcimos que en ARN viral los dominios con una composición promedio que corresponda a un valor pequeño de  $i_k$ , presentarán una distribución con una dispersión grande.

Los resultados experimentales descritos anteriormente (sección 3.3), sugieren una relación entre la ocurrencia de codones del tipo NDF, y la exactitud del proceso de traducción de proteínas. Dada la alta tasa de mutabilidad en los retorvirus, es razonable esperar una disminución en el peso relativo de codones NDF. Para explorar la generalidad de

estas observaciones experimentales en [50] y [51] se analiza el comportamiento estadístico de secuencias genéticas tomadas de células eucariotes y del virus del SIDA. La cantidad que se estudia es la frecuencia local de aparición  $P_\ell(NDF, k)$  de codones NDF en secuencias de mRNA.  $P_\ell(NDF, k)$  está dada por:

$$P_\ell(NDF, k) = \frac{n_{NDF}(k)}{n_{NDN}} \quad (6.9)$$

donde  $k$  denota la posición del codón a la largo de la cadena y  $n_{NDF}(k)$  es el número de codones NDF en una ventana de longitud  $\ell$  centrada en la posición  $k$ . Algunos resultados se muestran en la Figura 6.8. La línea continua representa el valor promedio  $\bar{P}_\ell(NDF, k)$  de  $P_\ell(NDF, k)$  para la secuencia de mRNA que codifica las proteínas gp120 y gp40 de la envoltura del virus. El promedio se tomó sobre las 25 secuencias del VIH-1 encontradas en [52] y  $\ell = 20$ . Con línea punteada se indica el promedio  $\bar{V}_\ell(k)$  de la variabilidad  $V_\ell(k)$  definida en cada posición  $k$  como el número de casos en que la composición del codón difiere de la composición encontrada con mayor frecuencia.  $V_\ell(k)$  se calcula dentro de una ventana de longitud  $\ell$  centrada en  $k$  y se toma el promedio sobre las 25 secuencias. De esta gráfica se observa que para valores de  $i$  menores que 500, que corresponden a la región de la gp120, los dominios con valores grandes de  $\bar{P}_\ell(NDF, k)$  presentan, en general, valores pequeños de  $\bar{V}_\ell(k)$  y viceversa. Esta relación es menos pronunciada en la región correspondiente a la gp41, lo cual puede deberse a la presencia de una codificación múltiple. Esto quiere decir que en algunas regiones del virus, el mensaje se puede leer no solamente en un marco de referencia sino en dos y en partes hasta en tres marcos de referencia diferentes. En dichas zonas el mRNA codifica no solo una sino dos y hasta tres proteínas diferentes y esto hace que la base en la tercera posición que sería casi irrelevante para un marco dado, es importante para los otros dos y está restringida para que se tengan los aminoácidos correspondientes a los otros marcos.

Este análisis estadístico refuerza las determinaciones experimentales mencionadas en las secciones anteriores, algunas de las cuales se realizaron *in vitro*. El dominio de codones NDF parece estar fuertemente

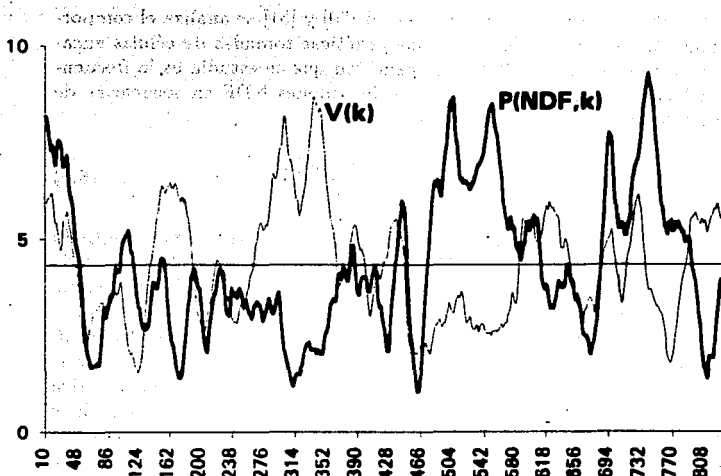


Figura 6.8: Gráfica de  $\bar{P}_\ell(NDF, i)$  (línea continua) y  $\bar{V}_\ell(i)$  (línea punteada) determinadas a partir de 25 secuencias del virus VIH-1 como función de la posición  $k$  del codón.

correlacionado con la baja variabilidad en las secuencias genéticas.

La tendencia observada en la Figura 6.8 corrobora la predicción de nuestro modelo de que regiones con una variabilidad  $\bar{V}_\ell(i)$  grande coinciden con concentraciones bajas de codones NDF, es decir valores pequeños de  $\bar{P}_\ell(NDF, i)$ . La comparación de los valores globales de  $P_\ell(NDF)$  para células eucariotes y el VIH revela una marcada diferencia entre estos valores. Se ha encontrado [50] que para células eucariotes  $\bar{P}_\ell(NDF) \approx 0.42$  mientras que para VIH  $\bar{P}_\ell(NDF) \approx 0.24$ . Esta diferencia en los valores de  $\bar{P}_\ell(NDF)$  sugiere que mediante una modulación química del ribosoma se podría bloquear la traducción de VIH sin afectar la síntesis de proteínas eucariotes. Experimentos recientes

realizados *in vitro* [53] muestran que mediante el uso de algunas cadenas de RNA llamadas *ribozimas* diseñadas *ex profeso* es posible alterar el proceso de traducción del ARN correspondiente a virus específicos. Aún cuando no se tiene la seguridad de que el proceso funcione igualmente *in vivo*, la exploración de esta indicación es de vital importancia ya que representa una posibilidad de encontrar un tratamiento terapéutico de la enfermedad.

## Capítulo 7

# BIFURCACIONES

Siempre que la solución de una ecuación o sistema de ecuaciones cambie cualitativamente cuando un parámetro tome un valor muy particular -que llamaremos *valor crítico*- diremos que el sistema exhibe una *bifurcación*. Un punto en el espacio de parámetros en donde ocurre esto es llamado un punto de bifurcación. De un punto de bifurcación emergen varias soluciones las cuales pueden ser estables o inestables. Existen diversos mecanismos por medio de los cuales el punto fijo o la órbita periódica de un sistema dinámico pueden bifurcarse. Entre los más comunes están: 1.-la *bifurcación subarmónica*, o de *duplicación de período*, 2.-la *bifurcación tangente*, 3.-la *bifurcación de Hopf* y 4.-la *tipo tridente*. A continuación se da una breve descripción de estas bifurcaciones tanto en mapeos unidimensionales como en mapeos multidimensionales. Para una revisión detallada ver [54], [55] y [56].

### 7.1 Mapeos unidimensionales

Consideremos la familia de mapeos unidimensionales definida por:

$$f : \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R} \quad (7.1)$$

donde  $(\mu, x) \mapsto f_\mu(x)$  y  $\mu$  es un parámetro. Sin pérdida de generalidad podemos considerar que el punto fijo de  $f$  está localizado en  $x = 0$  y que el valor de  $\mu$  para el cual el punto fijo se vuelve inestable, dando



Tabla 7.1: Bifurcaciones en mapeos unidimensionales.

Tipo	$\chi(0,0)$	Punto fijo	Ejemplos
Tangente	1	$f(0,0) = 0$	$\mu + x + x^2$
Subarmónica	-1	$f(0,0) = 0$	$\mu - x - x^2$
Tridente	1	$f(\mu, -x) = -f(\mu, x)$	$(\mu + 1)x - x^3$

lugar a una *bifurcación*, es  $\mu = 0$ . La estabilidad de los puntos fijos de este tipo de mapeos está gobernada por el valor de su derivada  $\chi(x, \mu) = \left(\frac{df}{dx}\right)_\mu$  en ese punto. Supongamos que  $f$  tiene un punto fijo en  $(x = 0, \mu = 0)$  y que queremos determinar la naturaleza de las órbitas en puntos cercanos a ese punto fijo. Entonces resulta natural considerar el mapeo lineal asociado:

$$y \mapsto Ay \quad y \in \mathbf{R}. \quad (7.2)$$

donde  $A = \chi(0,0)$ . Si  $|A| < 1$  entonces  $x = 0$  es estable, y si  $|A| > 1$ ,  $x = 0$  es inestable.

Los tipos más comunes de bifurcaciones pueden resumirse en la Tabla 7.1. Además de las condiciones que se indican en esta Tabla deben satisfacerse algunas otras para cada caso particular<sup>1</sup>. Los correspondientes diagramas de bifurcación se muestran en la Figura 7.1. Es importante hacer notar que aún cuando los diagramas de bifurcación para las bifurcaciones subarmónica y tridente son a primera vista iguales (Figura 7.1.a y 7.1.b), existe una diferencia muy importante entre ellos y es que mientras que el primero se pasa de un punto fijo a una órbita de período dos, el segundo se pasa a dos órbitas de período uno.

<sup>1</sup>Tangente:  $\frac{\partial^2 f}{\partial x^2}(0,0) > 0$ ,  $\frac{\partial^2 f}{\partial \mu}(0,0) > 0$ .

Feigenbaum:  $\frac{\partial^2 f}{\partial x \partial \mu}(0) > 0$ ,  $\frac{\partial^2 f}{\partial x^2}(0,0) < 0$ .

Tridente:  $\frac{\partial^2 \chi}{\partial \mu}(0) > 0$ ,  $\frac{\partial^3 f}{\partial x^3}(0,0) < 0$

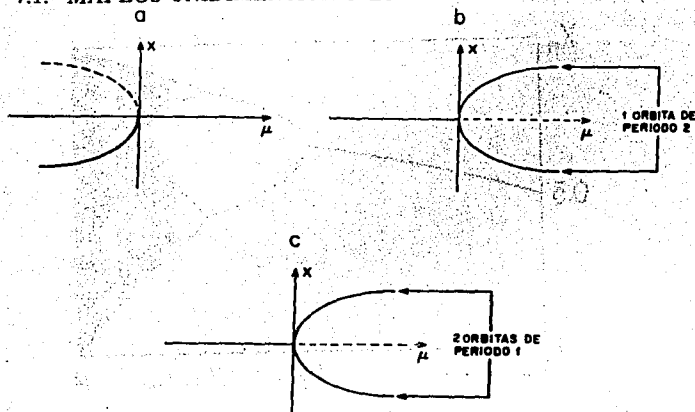


Figura 7.1: Diagramas de bifurcación para las bifurcaciones a).- Tangente, b).-Subarmónica y c).-Tridente. Las líneas punteadas indican líneas de puntos inestables.

Por ahora estamos interesados en analizar con un poco más de detalle la bifurcación subarmónica o de doblamiento de período. Un ejemplo representativo de este tipo de bifurcación se da en el mapeo logístico definido por:

$$x_{i+1} = \mu x_i(1 - x_i) \quad (7.3)$$

en donde  $x \in [0, 1]$  y  $\mu \in [0, 4]$ . Mediante un par de cambios de variable se puede ver que el mapeo  $f(x) = \mu - x - x^2$  de la Tabla 7.1, es equivalente al de la ecuación (7.3).

En la Figura 7.2, se muestra el diagrama de bifurcaciones de mapeo logístico (ec. 7.3), en el cual se grafican los puntos visitados durante la evolución de  $x_i$  para  $i > M$ , donde  $M$  es un entero suficientemente grande como para garantizar que estamos observando el comportamiento asintótico, como función de  $\mu$ . En esta gráfica se observa que en el intervalo de valores entre 0 y  $\mu_1$ , el sistema visita un solo valor. Al pasar por  $\mu_1$  observamos una bifurcación y en todo el inter-

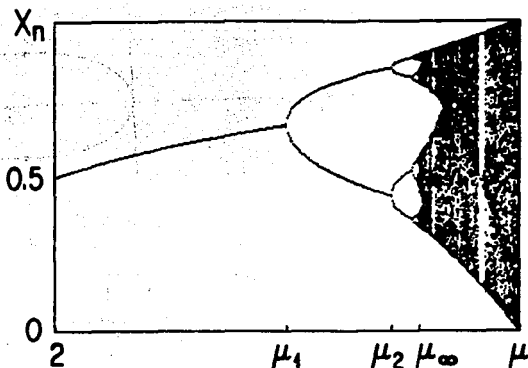


Figura 7.2: Diagrama de bifurcaciones del mapeo logístico, en donde se grafican 100 iteraciones después de un transitorio de 1000. En la figura se indican los valores de  $\mu_1$ ,  $\mu_2$  y  $\mu_\infty$  (ver Tabla No. 7.2)

valo entre  $\mu_1$  y  $\mu_2$  el sistema visita dos valores. Nuevamente al pasar por  $\mu_2$  se observa otra bifurcación. A estas bifurcaciones se les conoce como subarmónicas debido a que el número de puntos del atractor es cada vez mayor por un factor de dos. Los valores  $\mu_n$  para los cuales se presentan estas bifurcaciones se pueden obtener numéricamente. Dichos valores forman una serie creciente que converge rápidamente hacia un punto de acumulación  $\mu_\infty$  (ver Tabla 7.2). Este valor corresponde a un órbita de período infinito que es el resultado de la cascada de doblamientos de período. La convergencia hacia el punto de acumulación obedece una ley rigurosa: la diferencia entre los valores de  $\mu$  asociados con dos bifurcaciones consecutivas se reduce cada vez por un factor casi constante:

$$\delta_n = \frac{\mu_n - \mu_{n-1}}{\mu_{n+1} - \mu_n} \quad (7.4)$$

Se encuentra que el valor  $\delta$  definido por:

$$\delta = \lim_{n \rightarrow \infty} \delta_n \quad (7.5)$$

Tabla 7.2: Valores de  $\mu_n$  para los cuales el mapeo logístico sufre una bifurcación de una órbita de período  $2^{n-1}$  a una de período  $2^n$ . Se incluyen también los valores de  $\delta_n$ .

Periodicidad del atractor	$\mu$	$\delta$
$2^0 = 1$	$\mu_1 = 3.0$	...
$2^1 = 2$	$\mu_2 = 3.449489 \dots$	$\delta_2 = 4.7514 \dots$
$2^2 = 4$	$\mu_3 = 3.544090 \dots$	$\delta_3 = 4.6562 \dots$
$2^3 = 8$	$\mu_4 = 3.564407 \dots$	$\delta_4 = 4.6682 \dots$
$2^4 = 16$	$\mu_5 = 3.568759 \dots$	$\delta_5 = 4.6687 \dots$
$2^5 = 32$	$\mu_6 = 3.569692 \dots$	...
...	...	...
$2^\infty = \infty$	$\mu_\infty = 3.5699456 \dots$	...

es una constante universal independiente de la función  $f$ , siempre que ésta sea una función no lineal con un máximo de segundo orden [56]. El valor de esta constante universal es:

$$\delta = 4.669\,201\dots \quad (7.6)$$

Realizando simulaciones numéricas se encuentra que el valor  $\mu_\infty$  marca el comienzo de una dominio muy complicado y un análisis detallado muestra que en la región  $\mu_\infty < \mu < 4$  se encuentran alternadamente atractores periódicos y zonas caóticas. Para entender esta cascada de doblamientos de período Feigenbaum [57] llevó a cabo un análisis en el que la idea central es que la iteración  $f_\mu^{2^n}$  se comporta igual que la iteración  $f_\mu^{2^{n-1}}$  salvo por un reescalamiento. Este resulta ser un punto de vista bastante natural para un físico familiarizado con ideas del grupo de renormalización.

Una bifurcación tipo tangente se observa en el mismo mapeo logístico cuando aparece por primera vez el punto fijo que no es trivial ( $\mu = 1$ ). Un ejemplo muy claro de una bifurcación tipo tridente, se tiene en el caso de la función  $f_\mu^2(x)$ , cuando  $\mu = \mu_1$ , donde  $f_\mu$  es el mapeo logístico dado por la ecuación (7.3). Sabemos que en  $\mu_1$ ,  $f$  presenta una bifurcación de doblamiento de período, es decir el punto fijo

se convierte en una órbita de período dos. En este punto, la segunda iteración de  $f_\mu(x)$ , es decir  $f_\mu^2(x)$ , continua teniendo período uno, con la diferencia de que ahora existen dos valores asintóticos de período uno (Figura 7.1). El valor asintótico al cual converge una trayectoria dada depende de la condición inicial.

A continuación se da una breve descripción de los mecanismos que dan lugar a bifurcaciones en mapeos multidimensionales.

## 7.2 Mapeos multidimensionales

Consideremos la familia de un parámetro de mapeos (si el problema incluye más de un parámetro, se mantienen todos fijos menos uno) definida por:

$$g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n \quad (7.7)$$

donde  $(y, \mu) \mapsto g_\mu(y)$  y  $g$  es una función con derivada de orden  $r$  continua ( $C^r$ ),  $r > 1$ . Supongamos que  $g$  tiene un punto fijo en  $(y, \mu) = (\bar{y}, \mu_0)$ , es decir  $g_{\mu_0}(\bar{y}) = \bar{y}$ . Cuando el punto fijo es hiperbólico (es decir cuando ninguno de los valores propios de la matriz Jacobiana  $J = D_y g_{\mu_0}(\bar{y})$  tiene módulo igual a uno), el mapeo lineal asociado:

$$\xi \mapsto J\xi \quad (7.8)$$

donde  $\xi \in \mathbb{R}^n$ , caracteriza la estabilidad de los puntos fijos: si el mapeo lineal asociado es estable (inestable), el punto fijo  $\bar{y}$  será estable (inestable). El problema surge cuando se quiere caracterizar la estabilidad de un punto fijo que es *no hiperbólico*, ya que en ese caso la aproximación lineal (7.8) no puede usarse para determinar su estabilidad. Los siguientes son los ejemplos más sencillos en los que el punto fijo es no hiperbólico:

- 1.-  $J$  tiene un solo valor propio igual a 1 y los restantes valores propios tienen módulo diferente de 1.
- 2.-  $J$  tiene un solo valor propio igual a -1 y los restantes valores propios tienen módulo diferente de 1.
- 3.-  $J$  tiene dos valores propios complejos conjugados con módulo igual

a 1 y los restantes valores propios tienen módulo diferente de 1.

El primer caso corresponde a una bifurcación *nodo-punto silla*, el segundo corresponde a una bifurcación de *doblamiento de período o subarmónica* y el tercer caso a una bifurcación de *Hopf*.

Si se desea obtener información sobre la existencia de divergencias exponenciales, se puede calcular el espectro de exponentes de Lyapunov. Estos exponentes juegan un papel muy importante en el estudio de sistemas dinámicos, ya que proporcionan una caracterización cualitativa y cuantitativa del comportamiento dinámico del sistema. Los valores de los exponentes de Lyapunov son una medida de la razón exponencial promedio de la divergencia o convergencia de trayectorias vecinas en el espacio fase. Los signos de los exponentes caracterizan cualitativamente la dinámica del sistema. Por ejemplo un mapeo unidimensional está caracterizado por un solo exponente de Lyapunov, el cual es positivo para un comportamiento caótico, cero para una órbita marginalmente estable y negativo para una órbita periódica. En sistemas de dimensionalidad más alta, la presencia de un exponente positivo indica que las trayectorias divergen exponencialmente en alguna dirección, esto significa que el comportamiento a tiempos largos de una trayectoria que se especifica con cierta incertidumbre no puede predecirse. Se tiene como condición necesaria para que un sistema sea caótico el que tenga al menos un exponente de Lyapunov positivo. La magnitud de dicho exponente refleja la escala de tiempo a la cual el sistema se vuelve impredecible.

El espectro de exponentes de Lyapunov está definido por:

$$\lambda_j = \ln |\Gamma_j| \quad (7.9)$$

donde:

$$\Gamma_j = \lim_{n \rightarrow \infty} (j - \text{ésimo valor propio de } \prod_{k=1}^n (\mathbf{J}_k \mathbf{J}_k^*)^{\frac{1}{2n}}), \quad (7.10)$$

$\mathbf{J}_k$  es la matriz Jacobiana asociada al sistema, evaluada en el  $k$ -ésimo punto de la trayectoria y  $\mathbf{J}_k^*$  su adjunta. Usualmente estos exponentes se dan en orden decreciente, es decir:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M. \quad (7.11)$$

**ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA**

Las propiedades de estos exponentes son fáciles de entender en el caso especial de una órbita periódica ya que en ese caso es posible definir valores y vectores propios asociados a la matriz Jacobiana del sistema. Para una órbita de período  $p$  los exponentes están dados por:

$$\lambda_i = \frac{1}{p} \sum_{k=1}^p \ln \Gamma_i(\mathbf{x}_k) \quad (7.12)$$

donde  $\Gamma_i(\mathbf{x}_k)$  es el  $i$ -ésimo valor propio de  $\mathbf{J}_k \mathbf{J}_k^*$ .

Para una órbita no periódica, no es posible definir estos valores ni los correspondientes vectores propios, sin embargo Oseledec [58] demostró la existencia de un conjunto de vectores base que permite el cálculo de los exponentes para órbitas no periódicas.

Benetin et. al. [59] desarrollaron un método para calcular este espectro de exponentes, el cual está basado en el llamado teorema de Oseledec [58]. Los exponentes a los cuales nos hemos referido hasta ahora son también llamados exponentes de orden 1; en su teorema, Oseledec generalizó este concepto y definió los exponentes de Lyapunov de orden  $p$  que están relacionados con el incremento exponencial de volúmenes en el espacio fase. En el algoritmo desarrollado por Benetin et. al., la determinación del espectro de exponentes de Lyapunov de orden 1 se basa en el cálculo de exponentes de orden superior.

Como se dijo anteriormente, este conjunto de exponentes caracteriza el comportamiento dinámico de un sistema, sin embargo cuando se desea determinar características más específicas de la dinámica como puede ser la naturaleza de las bifurcaciones, es necesario evaluar los valores propios de la matriz  $\mathbf{J}$ .

## 7.3 CASCADA DE BIFURCACIONES

En la sección 6.1 se mencionó que cuando aumenta el valor del parámetro  $\epsilon$ , se observa una serie de bifurcaciones en la distribución de secuencias. En esta sección vamos a analizar la cascada de bifurcaciones

CENTRO DE INVESTIGACIONES Y ESTADÍSTICAS

observada. Para visualizar mejor estas bifurcaciones, nos concentramos en el análisis del término de biomasa  $\bar{N}(t)$ , definido por la ecuación (5.3), como función de  $\epsilon$ . En la Figura 7.3 se muestra el diagrama de bifurcaciones de  $\bar{N}(t)$  para una red de 10 mapeos acoplados. Los valores de los parámetros son  $a = 1.5, b = 0, c = 0, d = 1$  y  $\epsilon \in [0, 3]$ . Claramente esta gráfica presenta un gran parecido con el diagrama de bifurcaciones del mapeo logístico (Figura 7.2). Es posible obtener numéricamente los valores  $\epsilon_1, \epsilon_2$ , etc. para los cuales se observan las primeras bifurcaciones. Es conveniente aclarar que en la gráfica de la Figura 7.3, el punto de inflexión en  $\epsilon_1$  corresponde a la primera bifurcación en la distribución  $\{N_i(t)\}$  sin embargo a partir de la segunda bifurcación, es decir cuando se pasa de una órbita de período  $2^1$  a una de período  $2^2$ , con cada bifurcación de la distribución se observa una bifurcación en el término de biomasa. Por esta razón en este análisis nos limitamos a los valores de  $\epsilon_i$  para  $i > 1$ . Como se mencionó anteriormente, gran parte de la dinámica del mapeo acoplado se genera debido al hecho de haber introducido la función tipo Heaviside. El punto de inflexión que aparece en la gráfica  $\bar{N}(t)$  vs  $\epsilon$  se debe precisamente a este hecho. Mientras nos encontramos en una dinámica de punto fijo, este hecho es irrelevante, sin embargo para valores de  $\epsilon$  mayores que  $\epsilon^*$ , las dos curvas que se muestran en la Figura (6.2), comienzan a separarse. De no haber introducido la función tipo Heaviside, estas seguirían separándose indefinidamente yendo una de ellas hacia infinito y la otra hacia menos infinito, sin embargo debido a la acción de dicha función, cuando una de ellas llega a cero la separación se detiene, provocando que el valor de  $\bar{N}(t)$  aumente debido a que no tiene la contribución de la parte negativa correspondiente. A partir de  $\epsilon^*$  es fundamental el introducir la función tipo Heaviside ya que de otra manera al iterar el mapeo acoplado eventualmente se tienen divergencias hacia  $\pm\infty$ .

En la Tabla 7.3 se muestran los valores de  $\epsilon_n$  (para  $n = 2, 3, 4, 5, 6, 7$  y 8) que se obtienen de la gráfica de la Figura 7.3. A partir de estos valores se puede verificar si se satisface una ley semejante a la que obedece el parámetro  $\mu$  en el mapeo logístico. Como puede verse de la



Tabla 7.3: Valores del  $\epsilon$  en los cuales  $\bar{N}(t)$  presenta bifurcaciones. En la tercera columna se dan los valores de  $\Delta_n^\epsilon$ .

Periodicidad del atractor	$\epsilon$	$\Delta^\epsilon$
$2^1 = 2$	$\epsilon_2 = 2.398774\dots$	$\dots\dots$
$2^2 = 4$	$\epsilon_3 = 2.759591\dots$	$\Delta_2^\epsilon = 4.7587561\dots$
$2^3 = 8$	$\epsilon_4 = 2.835224\dots$	$\Delta_3^\epsilon = 4.6429098\dots$
$2^4 = 16$	$\epsilon_5 = 2.851514\dots$	$\Delta_4^\epsilon = 4.6676217\dots$
$2^5 = 32$	$\epsilon_6 = 2.855004\dots$	$\Delta_5^\epsilon = 4.6676474\dots$
$2^6 = 64$	$\epsilon_7 = 2.85575173\dots$	$\Delta_6^\epsilon = 4.6626496\dots$
$2^7 = 128$	$\epsilon_8 = 2.855912095\dots$	$\dots\dots$
$\dots\dots$	$\dots\dots$	$\dots\dots$

Tabla 7.3 el valor del parámetro  $\Delta_n^\epsilon$  definido por:

$$\Delta_n^\epsilon = \frac{\epsilon_n - \epsilon_{n-1}}{\epsilon_{n+1} - \epsilon_n} \quad (7.13)$$

es siempre muy cercano al valor de  $\delta$ .

Al igual que en el caso del mapeo logístico la cascada de doblamientos de período da lugar a la aparición de una órbita de período infinito. Como se mencionó anteriormente de las simulaciones numéricas se observa que para algunas condiciones iniciales la distribución se colapsa, esta característica parece estar relacionada con el hecho de haber introducido la función tipo Heaviside (ec. (5.13), sección 5.2). Cuando  $\epsilon$  toma valores cada vez mayores, el cero se visita artificialmente con una frecuencia que aumenta con  $\epsilon$ , por lo que el peso del punto fijo metaestable crece hasta volverse determinante. En este sentido diremos que es la introducción de un “atractor anómalo” que lo provoca el colapso de la distribución. En la Figura 7.4 se muestra un histograma del número de condiciones iniciales (de un total de 100) para las cuales la distribución se colapsa. Como se puede ver dicho número aumenta al aumentar el parámetro  $\epsilon$  hasta que eventualmente para todas las configuraciones iniciales la distribución se colapsa, los valores de los

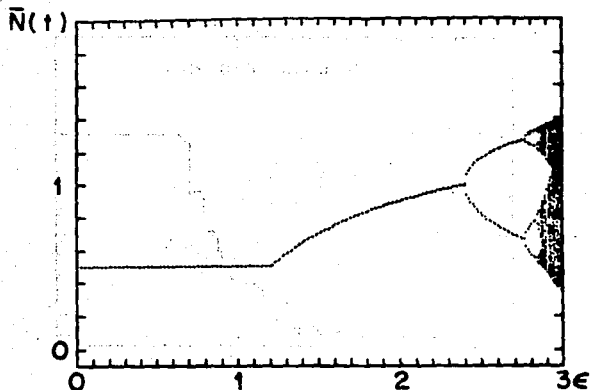


Figura 7.3: Diagrama de bifurcaciones de  $\bar{N}(t)$ . Los valores de los parámetros son  $a = 1.5$ ,  $b = 0$ ,  $c = 0$ ,  $d = 1$ ,  $\epsilon \in [0, 3]$  y  $L = 10$ .

parámetros se muestran en la figura.

Ahora lo que resta es saber de que tipo son las bifurcaciones observadas. Las características mencionadas anteriormente parecen sugerir que las bifurcaciones son del tipo de doblamiento de período. Para la primera bifurcación es posible calcular los valores propios de matriz  $\mathbf{J}$  asociada a la dinámica. Los elementos de la matriz  $\mathbf{J}$  están dados por:

$$\begin{aligned}
 (\mathbf{J})_{ij} &= \frac{\partial N_i(t+1)}{\partial N_j(t)} = (a - d\bar{N}(t)) \left[ \left(1 - \frac{\epsilon L}{L+2}\right) \delta_{ij} + \right. \\
 &\quad \left. \epsilon \left(\frac{1-i}{L+2}\right) \delta_{i+1,j} + \epsilon \left(\frac{L+1-i}{L+2}\right) \delta_{i-1,j} \right] - \\
 &\quad \frac{d}{L+1} \left[ \left(1 - \frac{\epsilon L}{L+2}\right) N_i(t) + \epsilon \left(\frac{1-i}{L+2}\right) N_{i+1}(t) + \right. \\
 &\quad \left. \epsilon \left(\frac{L+1-i}{L+2}\right) N_{i-1}(t) \right] \quad (7.14)
 \end{aligned}$$

Dado que las bifurcaciones aparecen independientemente del valor

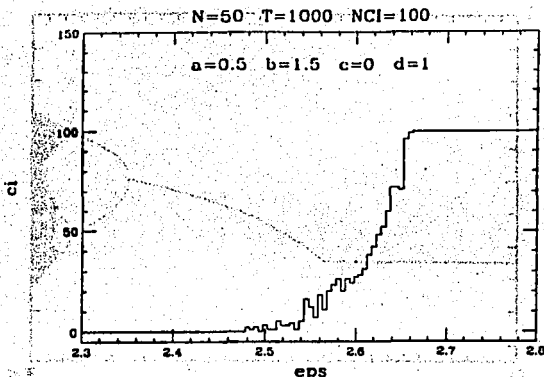


Figura 7.4: Histograma de condiciones iniciales para las cuales la distribución se colapsa.

de  $b$ , por simplicidad, vamos a hacer el análisis en el caso en que  $b = 0$ . Se puede ver que bajo condiciones de punto fijo, si  $b = 0$ , la distribución asintótica  $N_i(t)$  (vista como función de  $i$ ) es simétrica (Figura 6.3). Supongamos, sin pérdida de generalidad, que  $L$  es un número par, entonces de (5.10) se tiene que los valores de las diferentes  $N_i(t)$  están dados por:

$$N_i = \frac{L(L-1)(L-2)\cdots(L-i+1)}{i!} N_0 \quad i = 1, 2, \dots, \frac{L}{2} \quad (7.15)$$

con:

$$N_i = N_{L-i} \quad i = 0, 1, \dots, \frac{L}{2} \quad (7.16)$$

Si  $L$  es impar entonces en las ecuaciones (5.24) y (5.25)  $i = 1, 2, \dots, \frac{L-1}{2}$  e  $i = 0, 1, \dots, \frac{L-1}{2}$ , respectivamente. Cuando  $b = 0$  se puede demostrar que  $\bar{N}(t)$  está dada por:

$$\bar{N}(t+1) = (a - d\bar{N}(t))\bar{N}(t) \quad (7.17)$$

y este mapeo tiene un punto fijo dado por:

$$\bar{N} = \frac{(a-1)}{d} \quad (7.18)$$

Haciendo uso de esta información en (7.14), se puede demostrar que los valores propios de la matriz  $\mathbf{J}$  en el punto fijo están dados por:

$$\begin{aligned} \lambda_0 &= 2 - a \\ \lambda_1 &= 1 - \frac{2L}{L+2}\epsilon \\ \lambda_2 &= 1 - \frac{2(L-1)}{L+2}\epsilon \\ \lambda_3 &= 1 - \frac{2(L-2)}{L+2}\epsilon \\ &\vdots \\ \lambda_L &= 1 - \frac{2}{L+2}\epsilon \end{aligned} \quad (7.19)$$

donde  $\mathbf{J}$  es una matriz de  $(L+1) \times (L+1)$ . Así vemos que cuando mantenemos fijo el parámetro  $a$  y variamos  $\epsilon$ , los valores propios que cambian son únicamente  $\lambda_r$  con  $r > 0$ . Además  $\lambda_i < \lambda_j$  si  $i < j$ , por lo tanto, al aumentar  $\epsilon$  a partir de cero, el primer valor propio que cruzará el círculo unitario será  $\lambda_1$ . La condición para que esto ocurra es que  $\lambda_1 = -1$ , es decir:

$$\epsilon = \frac{L+2}{L} \quad (7.20)$$

que es el valor  $\epsilon^*$  mencionado en la sección 6.1.1.

Hemos encontrado así que cuando el punto fijo pierde su estabilidad, la matriz  $\mathbf{J}$  tiene un valor propio igual a  $-1$  y por lo tanto podemos concluir que la bifurcación de un punto fijo a una órbita de período dos, es del tipo de doblamiento de período. Respecto a la segunda bifurcación, solo se cuenta con argumentos cualitativos para identificar su tipo. Observando el diagrama de bifurcaciones para una sola de las  $N_i(t)$  (Figura 7.5), se encuentra que cuando  $\epsilon \approx 2.4$ , de la rama superior surgen dos ramas, sin embargo, la órbita de período dos continua siendo una órbita de período dos, pero con la diferencia de que ahora el valor

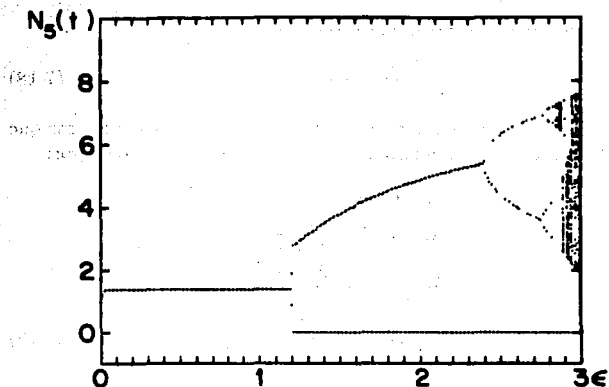


Figura 7.5: Diagrama de bifurcaciones de  $N_5(t)$  para  $a = 1.5$ ,  $b = 0$ ,  $c = 0$ ,  $d = 1$ ,  $\epsilon \in [0, 3]$  y  $L = 10$ .

asintótico en la rama superior depende de la condición inicial, tal como sucede en la bifurcación tipo tridente (sección 7.1). Finalmente para las demás bifurcaciones se observa que cada una de las ramas sufre una bifurcación tipo Feigenbaum o doblamiento de período. El análisis de las series de valores  $\{\epsilon_n\}$  y  $\{\Delta_n^{\epsilon}\}$  apoya el hecho de que las bifurcaciones tanto de  $\bar{N}(t)$  como las de cada  $N_i(t)$  son del tipo de doblamiento de período. Cuando  $b \neq 0$ , la distribución es asimétrica, como puede verse de la Figura 6.1, sin embargo la fenomenología observada es muy parecida.

## Capítulo 8

# MODELOS TIPO CAMPO MEDIO

### 8.1 DESCRIPCION DE LA DINAMICA

En este capítulo vamos a concentrarnos en el análisis de la dinámica de los modelos tipo campo medio introducidos en el capítulo 5, los cuales se proponen como una aproximación del modelo definido por la ecuación (5.15). En el análisis de estos modelos y el que se hará en el capítulo 9, se consideran parámetros fuera del rango de relevancia para el problema del SIDA teniendo en mente su potencial para el estudio de otras redes biológicas, en particular las discusiones recientes sobre la relación entre dinámicas caóticas y diversidad biológica [4], [5].

Consideremos en primer término la dinámica que surge de (5.16). Posteriormente nos referiremos a la ecuación (5.17). Cuando  $b = 0$  y  $c = 0$  es posible hacer un estudio analítico de estos modelos por lo tanto, se considerará primero ese caso y posteriormente se analizarán casos en los que  $b \neq 0$  y  $c \neq 0$ .

8.1.1  $\alpha \neq 0$ 

Por ahora nos concentraremos en el análisis del mapeo globalmente acoplado dado por la ecuación (5.16), que reescribimos a continuación:

$$N_i(t+1) = f\{N_i(t)\} + \frac{\alpha}{(L+1)} \sum_{i=0}^L N_i(t)$$

en donde  $f\{N_i(t)\}$  está dada por la ecuación (5.2) y  $A_i$  por (6.1). Es importante recordar que cuando se realizan las simulaciones numéricas, se introduce la función tipo Heaviside que previene la aparición de cantidades negativas.

Cuando  $b = 0$  y  $c = 0$ , el mapeo (5.16) se reduce a:

$$N_i(t+1) = (a - d\bar{N}(t))N_i(t) + \alpha\bar{N}(t) \quad (8.1)$$

donde como antes:

$$\bar{N}(t) = \frac{1}{L+1} \sum_{i=0}^L N_i(t)$$

Los atractores más sencillos de este sistema son atractores coherentes, es decir,  $N_i(t) = N_j(t) = \bar{N}(t)$  para toda  $i$  y toda  $j$ . Se puede ver que la dinámica estará gobernada entonces por un mapeo logístico unidimensional dado por:

$$\bar{N}(t+1) = [a + \alpha] \bar{N}(t) - d[\bar{N}(t)]^2 \quad (8.2)$$

y es posible obtener analíticamente una relación entre los parámetros, la cual caracteriza la estabilidad del atractor coherente. El mapeo (8.2) tiene punto fijo dado por:

$$\bar{N} = \frac{a + \alpha - 1}{d} \quad (8.3)$$

el cual se vuelve inestable cuando  $|2 - a - \alpha| \geq 1$ , es decir cuando  $3 - a \leq \alpha$ . El exponente de Lyapunov del mapeo (8.2), está dado por:

$$\lambda = \ln |2 - a - \alpha| \quad (8.4)$$

de los puntos de la órbita. En un ejemplo numérico, los puntos de una órbita de un sistema de ecuaciones diferenciales autónomas con una dimensión se muestran en la Figura 8.1. Los puntos de la órbita se muestran en la Figura 8.1. Los puntos de la órbita se muestran en la Figura 8.1. Los puntos de la órbita se muestran en la Figura 8.1.

Los puntos de la órbita se muestran en la Figura 8.1. Los puntos de la órbita se muestran en la Figura 8.1. Los puntos de la órbita se muestran en la Figura 8.1. Los puntos de la órbita se muestran en la Figura 8.1.

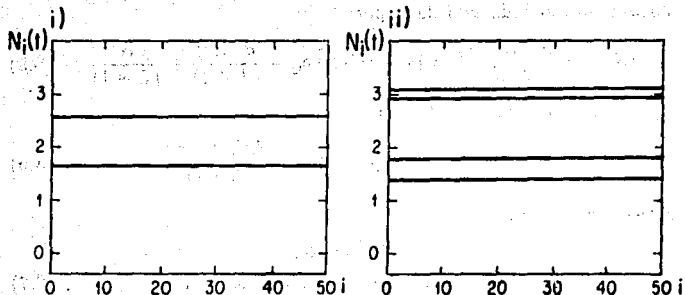


Figura 8.1: Algunos atractores coherentes del mapeo (8.1). Los valores de los parámetros son:  $d = 1$ ,  $\alpha = 1$  y  $L = 50$ . En i).-  $a = 2.2$  ii).-  $a = 2.5$ . Se grafican 50 iteraciones después de un transitorio de 1000.



En la Figura 8.1, muestran algunos de los atractores coherentes que se obtienen con esta dinámica. Los valores de los parámetros son:  $b = 0$ ,  $c = 0$ ,  $d = 1$  y  $L = 50$ . En i)  $a = 2.2$  y en ii)  $a = 2.35$ . Numéricamente se verifica que bajo un régimen de punto fijo y a partir de condiciones iniciales tomadas al azar, el punto fijo del mapeo acoplado (8.1) coincide con el punto fijo del mapeo unidimensional (8.2).

En general, como ya se mencionó en el capítulo 7, la estabilidad de los atractores del sistema está determinada por los valores propios de la matriz Jacobiana  $\mathbf{J}_n$  de las diferentes iteraciones, o alternativamente por los exponentes de Lyapunov. En el caso del mapeo (8.1), la matriz Jacobiana asociada está dada por:

$$(\mathbf{J}_t)_{ij} = \frac{\partial N_i(t+1)}{\partial N_j(t)} = [a - d\bar{N}(t)] \delta_{ij} - \frac{dN_i(t)}{(L+1)} + \frac{\alpha}{(L+1)} \quad (8.5)$$

o bien si el atractor es coherente:

$$(\mathbf{J}_t)_{ij} = [a - d\bar{N}(t)] \delta_{ij} - \frac{(d\bar{N}(t) - \alpha)}{(L+1)} \quad (8.6)$$

y sus valores propios son:

$$\begin{aligned} \Lambda_0^t &= a - 2d\bar{N}(t) + \alpha \\ \Lambda_i^t &= a - d\bar{N}(t) \quad i = 1, 2, \dots, L \end{aligned} \quad (8.7)$$

Ahora bien, cuando nos encontramos en un punto fijo,  $\bar{N}(t) = \bar{N}$  y entonces:

$$(\mathbf{J})_{ij} = [a - d\bar{N}] \delta_{ij} - \frac{(d\bar{N} - \alpha)}{(L+1)} \quad (8.8)$$

y sus valores propios serán:

$$\begin{aligned} \Lambda_0 &= a - 2d\bar{N} + \alpha \\ \Lambda_i &= a - d\bar{N} \quad i = 1, 2, \dots, L \end{aligned} \quad (8.9)$$

Dado que el punto fijo de  $\bar{N}(t)$  está dado por (8.3), los valores propios son:

$$\begin{aligned} \Lambda_0 &= 2 - a - \alpha \\ \Lambda_i &= 1 - \alpha \quad i = 1, 2, \dots, L \end{aligned} \quad (8.10)$$

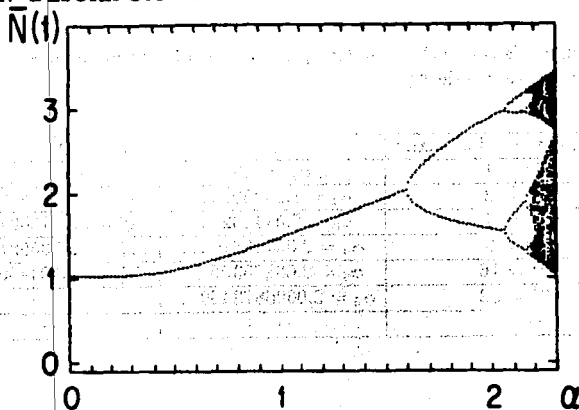


Figura 8.2: Diagrama de bifurcaciones de  $\bar{N}(t)$  para el mapeo (8.1). Los valores de los parámetros son:  $a = 1.5$ ,  $d = 1$ ,  $L = 10$  y  $\alpha \in [0, 2.3]$ . Se grafican 100 iteraciones consecutivas después de un transitorio de 1000.

La estabilidad de este punto fijo se perderá cuando alguno de los valores propios tenga módulo igual a 1. Por ejemplo si los valores de los parámetros son tales que  $\Lambda_0 = -1$ , entonces se observará una bifurcación de doblamiento de período.

Al variar el parámetro  $\alpha$  que controla la intensidad del campo promedio, se observa una serie de bifurcaciones que, al igual que en el caso del mapeo definido por la ecuación (5.13), parecen ser del tipo de doblamiento de período. En la Figura 8.2 se muestra la gráfica de  $\bar{N}(t)$  vs  $\alpha$  en donde se pueden apreciar claramente estas bifurcaciones. Cuando se trata de atractores coherentes, las bifurcaciones de  $\bar{N}(t)$  coinciden con las bifurcaciones de la distribución  $\{N_i(t)\}$ .

Sea  $\alpha_n$  el valor de  $\alpha$  para el cual se observa la bifurcación de una órbita de período  $2^{n-1}$  a una órbita de período  $2^n$ , y definamos  $\Delta_n^*$

Tabla 8.1: Valores del  $\alpha$  en los que  $\bar{N}(t)$  exhibe bifurcaciones. Se incluyen los vaores de  $\Delta_n^\alpha$

Periodicidad del atractor	$\alpha$	$\Delta^\alpha$
$2^0 = 1$	$\alpha_1 = 1.5$	.....
$2^1 = 2$	$\alpha_2 = 1.949592 \dots$	$\Delta_2^\alpha = 4.7570839 \dots$
$2^2 = 4$	$\alpha_3 = 2.044102 \dots$	$\Delta_3^\alpha = 4.6773004 \dots$
$2^3 = 8$	$\alpha_4 = 2.06440816 \dots$	$\Delta_4^\alpha = 4.6735621 \dots$
$2^4 = 16$	$\alpha_5 = 2.06875306 \dots$	$\Delta_5^\alpha = 4.6504986 \dots$
$2^5 = 32$	$\alpha_6 = 2.0696873469 \dots$	.....
.....	.....	.....

como:

$$\Delta_n^\alpha = \frac{\alpha_n - \alpha_{n-1}}{\alpha_{n+1} - \alpha_n} \quad (8.11)$$

En la Tabla 8.1 se dan los valores de las correspondientes  $\alpha_n$  y de los valores de  $\Delta_n^\alpha$  calculados por medio de (8.11). De esta Tabla se observa que los valores  $\alpha_n$  obedecen una ley similar a la que se mencionó en la sección 7.1 para el parámetro  $\mu$  en el caso del mapeo logístico.

Cuando  $b \neq 0$  y  $c \neq 0$ , la dinámica se vuelve un poco más complicada así como el estudio analítico del problema. Sin embargo es posible entender algunos aspectos en términos del análisis para  $b = 0$  y  $c = 0$ . Como puede verse de la Figura 8.3, cuando  $b \neq 0$  y  $c = 0$ , las distribuciones que en el caso  $b = 0$  y  $c = 0$  eran homogéneas (con respecto a  $i$ ) pierden esa homogeneidad y reflejan el hecho de que  $A(i)$  es una función creciente de  $i$ . Así mismo los valores de  $\alpha$  en los cuales se observan las bifurcaciones son menores en este caso que cuando  $b = 0$ . Los valores de los parámetros son:  $b = 0.3$ ,  $c = 0$ ,  $d = 1$  y  $L = 50$ . En i)  $a = 2.2$  y en ii)  $a = 2.35$ .

De la ecuaciones (5.2) y (5.16) es claro que si  $b = 0$  y  $c \neq 0$ , entonces el mapeo globalmente acoplado de la ecuación (5.16) se reduce a un mapeo del tipo de los estudiados por Kaneko [20], en donde se ob-

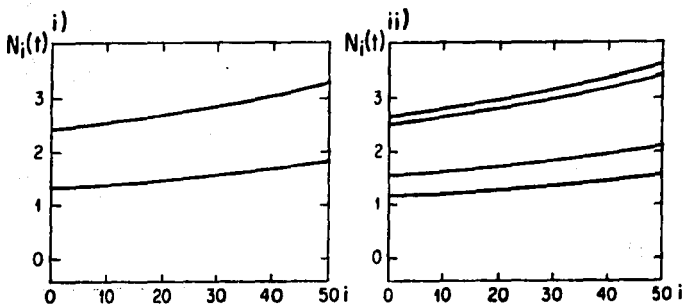


Figura 8.3: Distribuciones  $N_i(t)$  correspondientes al mapeo (5.16) para  $b = 0.3$ ,  $c = 0$ ,  $d = 1$  y  $L = 50$ . En i).-  $a = 2.2$ , ii).-  $a = 2.35$ . Se grafican 50 iteraciones después de un transitorio de 1000.

serva el surgimiento de dominios que evolucionan conjuntamente. En la Figura 8.4.i se muestra uno de los atractores para  $b = 0$ ,  $c \neq 0$   $d = 0$ , en donde se pueden ver claramente los dominios antes mencionados. Nótese que este comportamiento también se observa si  $d \neq 0$  (Figura 8.4.ii).

### 8.1.2 $\beta \neq 0$

Consideremos ahora el mapeo globalmente acoplado dado por la ecuación (5.17), es decir:

$$N_i(t+1) = f\{N_i(t)\} + \frac{\beta}{(L+1)} \sum_{i=0}^L f\{N_i(t)\}$$

donde como antes  $f\{N_i(t)\}$  está dada por la ecuación (5.2). Al igual que en el caso anterior el estudio analítico de este mapeo se puede realizar con mayor facilidad si  $b = 0$  y  $c = 0$ , en cuyo caso el mapeo

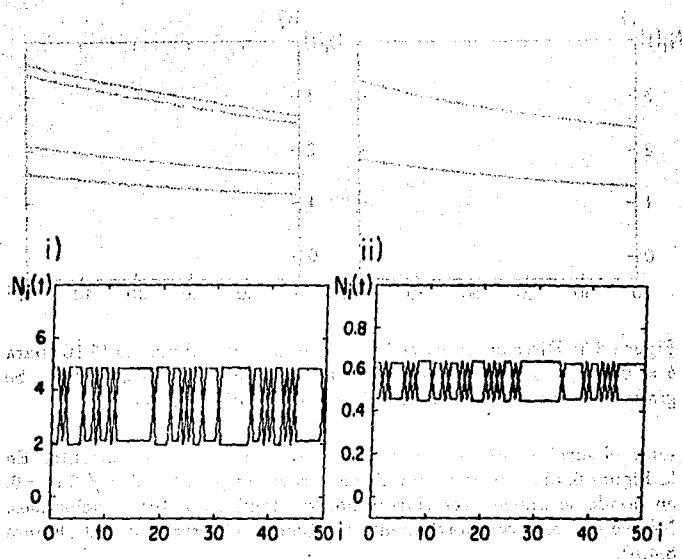


Figura 8.4: Gráficas de las distribuciones  $N_i(t)$  correspondientes al mapeo (5.16) para i).-  $a = 1.4$ ,  $b = 0$ ,  $c = 0.35$ ,  $d = 0$  y ii).-  $a = 1.6$ ,  $b = 0$ ,  $c = 1.9$ ,  $d = 1$ . En ambos casos  $\alpha = 1.0$  y  $L = 50$ .

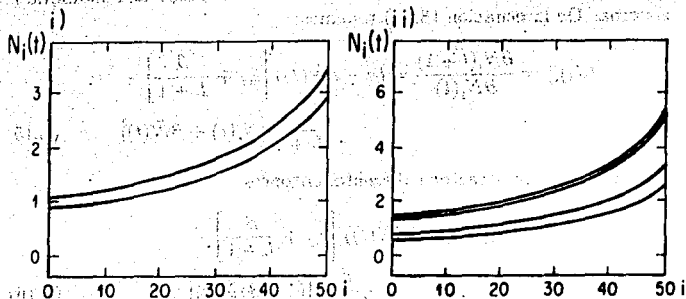


Figura 8.5: Algunos atractores coherentes del mapeo (8.12). Los valores de los parámetros son:  $d = 1$ ,  $\beta = 0.2$  y  $L = 50$ . En i).-  $a = 2.51$  y en ii).-  $a = 2.88$ . Se grafican 50 iteraciones después de un transitorio de 1000.

(5.17) se reduce a:

$$N_i(t+1) = (a - d\bar{N}(t))N_i(t) + \beta(a - d\bar{N}(t))\bar{N} \quad (8.12)$$

donde:

$$\bar{N}(t) = \frac{1}{L+1} \sum_{i=0}^L N_i(t).$$

Las simulaciones numéricas realizadas muestran que el mapeo (5.17) también puede exhibir atractores coherentes como los del mapeo (5.16). Algunos ejemplos se muestran en la Figura 8.5. Si el atractor del sistema es un atractor coherente, el mapeo (5.17) también se reduce a un mapeo logístico unidimensional, el cual está dado por:

$$\bar{N}(t+1) = (1 + \beta)(a - d\bar{N}(t))\bar{N}(t) \quad (8.13)$$

Este mapeo tiene un punto fijo dado por:

$$\bar{N} = \frac{(a - \frac{1}{\beta+1})}{d} \quad (8.14)$$

y la estabilidad del punto fijo se pierde cuando  $|2 - a(\beta + 1)| > 1$ , es decir cuando  $\beta > (\frac{2}{a} - 1)$ . Nuevamente para caracterizar la estabilidad de los atractores necesitamos calcular la matriz Jacobiana asociada al sistema. De la ecuación (8.12) tenemos:

$$(\mathbf{J}_t)_{ij} = \frac{\partial N_i(t+1)}{\partial N_j(t)} = (a - d\bar{N}(t)) \left[ \delta_{ij} + \frac{\beta}{L+1} \right] - \frac{d}{L+1} [N_i(t) + \beta \bar{N}(t)] \quad (8.15)$$

Si se trata de un atractor coherente, entonces:

$$(\mathbf{J}_t)_{ij} = (a - d\bar{N}(t)) \left[ \delta_{ij} + \frac{\beta}{L+1} \right] - \frac{d}{L+1} [(1 + \beta)\bar{N}(t)] \quad (8.16)$$

y sus valores propios están dados por:

$$\begin{aligned} \Lambda_0^t &= a(1 + \beta) - 2d(1 + \beta)\bar{N}(t) \\ \Lambda_i^t &= a - d\bar{N}(t) \quad i = 1, 2, \dots, L \end{aligned} \quad (8.17)$$

Si nos encontramos en un punto fijo entonces  $\bar{N}(t) = \bar{N}$  y:

$$(\mathbf{J})_{ij} = (a - d\bar{N}) \left[ \delta_{ij} + \frac{\beta}{L+1} \right] - \frac{d}{L+1} [(1 + \beta)\bar{N}] \quad (8.18)$$

y los valores propios serán:

$$\begin{aligned} \Lambda_0 &= a(1 + \beta) - 2d(1 + \beta)\bar{N} \\ \Lambda_i &= a - d\bar{N} \quad i = 1, 2, \dots, L \end{aligned} \quad (8.19)$$

Sustituyendo el valor de punto fijo de  $\bar{N}$  (ec. 8.14), se encuentra que:

$$\begin{aligned} \Lambda_0 &= 2 - a(1 + \beta) \\ \Lambda_i &= \frac{1}{\beta + 1} \quad i = 1, 2, \dots, L \end{aligned} \quad (8.20)$$

Tabla 8.2: Valores del parámetro  $\beta$  en los que  $\bar{N}(t)$  sufre una bifurcación. En la última columna se dan los valores de  $\Delta_n^\beta$ .

Períodicidad del atractor	$\beta$	$\Delta^\beta$
$2^0 = 1$	$\beta_1 = 1.0$	.....
$2^1 = 2$	$\beta_2 = 1.299796 \dots$	$\Delta_2^\beta = 4.7602534 \dots$
$2^2 = 4$	$\beta_3 = 1.362775 \dots$	$\Delta_3^\beta = 4.6649383 \dots$
$2^3 = 8$	$\beta_4 = 1.37627551 \dots$	$\Delta_4^\beta = 4.6553448 \dots$
$2^4 = 16$	$\beta_5 = 1.37917551 \dots$	$\Delta_5^\beta = 4.6836138 \dots$
$2^5 = 32$	$\beta_6 = 1.37979469 \dots$	.....
.....	.....	.....

y la estabilidad del punto fijo se pierde cuando alguno de los valores propios tiene módulo igual a 1. Dada la forma de los valores propios, esto sucederá cuando  $\Lambda_0 = -1$ , es decir cuando  $\beta > \frac{3}{\alpha} - 1$ .

También en este caso se observa una serie de bifurcaciones al variar el parámetro  $\beta$  que controla la intensidad del campo medio. En la Figura 8.6 se muestra el diagrama de bifurcaciones de  $\bar{N}(t)$ . Por tratarse de atractores coherentes, las bifurcaciones de  $\bar{N}(t)$  coinciden con las bifurcaciones de la distribución  $\{N_i(t)\}$ . En la Tabla 8.2 se dan los valores de las correspondientes  $\beta_n$  para las cuales se observan las bifurcaciones de una órbita de período  $2^{n-1}$  a una órbita de período  $2^n$ . Si definimos  $\Delta_n^\beta$  como:

$$\Delta_n^\beta = \frac{\beta_n - \beta_{n-1}}{\beta_{n+1} - \beta_n} \quad (8.21)$$

de esta Tabla se observa que los valores  $\beta_n$  también obedecen una ley similar a la que se mencionó en la sección 7.1 para el parámetro  $\mu$  en el caso del mapeo logístico, ya que el valor del parámetro  $\Delta_n^\beta$  calculado según (8.21) es cercano al valor del parámetro  $\delta$  definido para el mapeo logístico.

Cuando  $b \neq 0$ , el comportamiento observado es muy parecido al que se encuentra en el caso del mapeo (8.1), es decir las distribuciones que



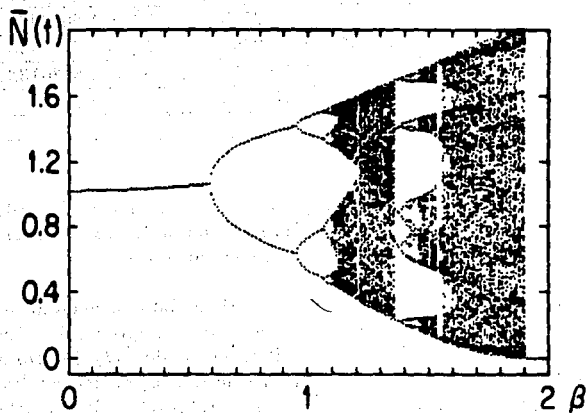


Figura 8.6: Diagrama de bifurcaciones de  $\bar{N}(t)$  para el mapeo (8.12). Los valores de los parámetros son:  $a = 1.5$ ,  $d = 1$ ,  $L = 10$  y  $\beta \in [0, 2]$ . Se grafican 100 iteraciones consecutivas después de un transitorio de 1000.

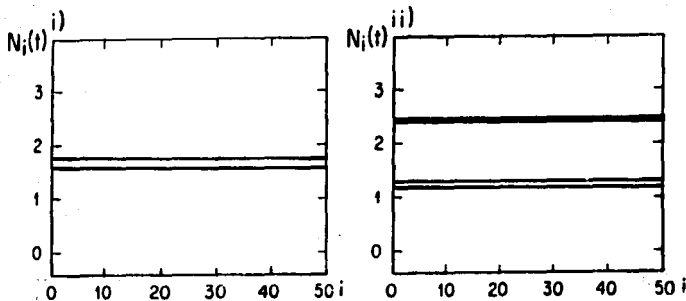


Figura 8.7: Gráficas de las distribuciones  $N_i(t)$  para  $b = 0.2$ ,  $c = 0$ ,  $d = 1$ ,  $L = 50$  y  $\beta = 0.2$  correspondientes al mapeo (5.17). En i).-  $a = 2.4$ , ii).-  $a = 2.8$ . Se grafican 50 iteraciones después de un transitorio de 1000.

eran homogéneas cuando  $b = 0$  se vuelven inhomogéneas y reflejan el hecho de que  $A(i)$  es una función monótona creciente de  $i$ . Algunos de los atractores observados se muestran en la Figura 8.7. Si  $b = 0$ ,  $c \neq 0$  y  $d = 0$ , el mapeo (5.16) también se convierte en un mapeo globalmente acoplado del tipo de los analizados en [20]. La Figura 8.8.i muestra uno de los atractores obtenidos en donde se aprecia claramente la formación de dominios que evolucionan en fase. Este comportamiento también se encuentra si  $d \neq 0$  (Figura 8.8.ii).

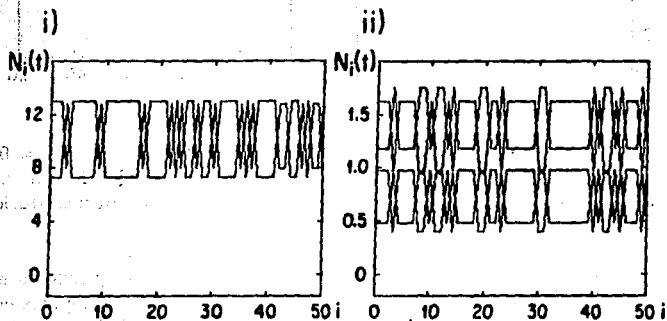


Figura 8.8: Gráficas de las distribuciones  $N_i(t)$  para el mapeo (5.17).  
i).- corresponde a  $a = 3.1$ ,  $b = 0$ ,  $c = 0.2$ ,  $d = 0$  y ii).- corresponde a  $a = 3.3$ ,  $b = 0$ ,  $c = 1$ ,  $d = 1$ . En ambos casos  $\beta = 0.1$  y  $L = 50$ .

## Capítulo 9

# SINCRONIZACION

En este capítulo vamos a centrar nuestra atención en el análisis del modelo como sistema dinámico. Esto ya se hizo en alguna medida en capítulos anteriores, en donde se dieron evidencias de un comportamiento coherente de un sistema de más baja dimensionalidad. Esto puede apreciarse claramente en los patrones de bifurcaciones presentados en las secciones 7.3 y 8.1. Recordemos que cuando se hizo el análisis de estos diagramas de bifurcaciones, se calculó el análogo de la  $\delta$  de Feigenbaum, encontrándose que los valores de  $\Delta^c$ ,  $\Delta^\alpha$  y  $\Delta^\beta$ , son muy cercanos al valor de la  $\delta$  de Feigenbaum. Se hizo también un análisis de los tipos de bifurcaciones encontradas. Ante estas evidencias podemos preguntarnos cual es el mecanismo operante en la dinámica que es capaz de producir todo esto, ¿no será cuestión de un problema de sincronización?

## 9.1 ANTECEDENTES RECIENTES

El problema de sincronización es un problema de interés en diversas áreas tales como circuitos eléctricos, comunicaciones, robótica, biología (cronobiología), etc. En mecánica clásica, se estudia el fenómeno de oscilaciones periódicas sincronizadas en donde es posible distinguir dos tipos de sincronización: la sincronización forzada por acciones externas y la sincronización mutua entre sistemas. También, aunque en

mucha menor medida, se ha estudiado el fenómeno de sincronización de osciladores no lineales más complicados. A grandes rasgos podemos decir que las componentes esenciales para que se de el fenómeno de sincronización en un sistema de osciladores son dos:

Primera, es necesario tener un conjunto osciladores más o menos independientes que puedan ser acoplados. Y segunda, debe haber disipación de energía.

El fenómeno de sincronización puede observarse en la naturaleza en una gran variedad de sistemas, quizás el más conocido se refiere al movimiento de la tierra y la luna. En este ejemplo, la frecuencia de rotación de la luna alrededor de su eje está sincronizada con su período de revolución alrededor de la tierra de tal forma que, desde la tierra, observamos siempre la misma cara de la luna. En este caso, los osciladores consisten en: uno, un cuerpo rígido (la luna) que rota a una frecuencia fija alrededor de un eje de inercia y el otro un cuerpo que se mueve en una órbita de Kepler (la tierra). El mecanismo de acoplamiento es menos obvio y tiene como origen la distribución asimétrica de la masa de la luna, la cual está acoplada a la asimetría producida en la tierra por la propagación de las mareas. El amortiguamiento está dado en este caso por la fuerza de fricción que ejerce la tierra sobre las mareas.

La sincronización de dinámicas caóticas, que es el fenómeno que nos interesa por el momento, es algo mucho más complejo y está asociado con la complejidad de los atractores de los sistemas dinámicos. A este fenómeno se le conoce como sincronización caótica. La trayectorias de dos sistemas caóticos independientes que evolucionan a partir de condiciones iniciales muy cercanas rápidamente se encontrarán totalmente descorrelacionadas. Sin embargo se ha encontrado [60], que algunos sistemas caóticos pueden sincronizarse si se les acopla con una señal común. Por sincronización debe entenderse que las trayectorias de dichos sistemas coincidan y sean recorridas con un retraso fijo. Esta capacidad de algunos sistemas caóticos para sincronizarse no es obvia y puede caracterizarse por medio de los exponentes de Lyapunov de los subsistemas. Pecora y Carroll [60] demostraron que subsistemas de una dinámica caótica se sincronizan únicamente si sus exponentes de Lyapunov son todos negativos, sin embargo ésta es una condición

necesaria pero no suficiente para la sincronización, ya que no hace referencia alguna al conjunto de condiciones iniciales que permitirán la sincronización.

En años recientes el problema sincronización para dinámicas no-lineales bajo regimenes caóticos ha sido estudiado con bastante interés [61]. Ejemplos de ello los encontramos en : i).- El estudio de estructuras coherentes en sistemas de alta dimensionalidad. Por ejemplo el análisis que hace Kaneko [20] sobre redes de mapeos acoplados estudiando principalmente las condiciones de estabilidad de dichas estructuras coherentes más que su origen mismo. ii).- Problemas de comunicaciones [62], [63] y de control de ritmos cardiacos [64] en donde se estudian sistemas de baja dimensionalidad y se habla de sincronización distinguiendo el sistema forzante del sistema de respuesta. iii).- Regularización, basada en variaciones de los parámetros incluidos en las ecuaciones [65]. iv).- El efecto del ruido en la regularización del caos en mapeos unidimensionales [66]. En la siguiente sección vamos a discutir con un poco más de detalle un caso particular de esto, que ha sido estudiado por Anishchenko et. al. [67].

## 9.2 UN CASO DE BAJA DIMENSION

Consideremos por ejemplo un caso particular del mapeo (8.1) en dos dimensiones ( $L = 1$ ,  $i = 0, 1$ ). Sabemos por la ecuación (8.7) que los valores propios de la matriz Jacobiana son:

$$\begin{aligned}\Lambda_0 &= a - 2d\bar{N} + \alpha \\ \Lambda_1 &= a - d\bar{N}\end{aligned}\quad (9.1)$$

y los correspondientes vectores propios estan dados por:

$$\begin{aligned}\mathbf{v}_1 &= (1, 1) \\ \mathbf{v}_2 &= (-1, 1)\end{aligned}\quad (9.2)$$

Si el valor absoluto de alguno de los valores propios es mayor que uno, una perturbación en la dirección del vector propio correspondiente tenderá a amplificarse. En este caso encontramos que uno de los vectores

propios de la matriz Jacobiana tiene todas sus componentes iguales entre sí, lo cual como ha sido indicado por Kaneko [20] es una condición suficiente para que la amplificación de una perturbación en la dirección de este vector no destruya la coherencia. Por lo tanto la condición suficiente para que una perturbación cualquiera no destruya la coherencia es que el otro valor propio tenga módulo menor o igual que uno, es decir:

$$|\Lambda_1| = |a - d\bar{N}| \leq 1 \quad (9.3)$$

Posteriormente vamos a discutir la posibilidad de aprovechar este resultado para analizar un sistema de alta dimensión por medio de la implementación de un esquema tipo grupo de renormalización. Un resultado similar se obtiene cuando se considera el mapeo (8.12).

### 9.3 ALTA DIMENSION

Los sistemas que se han estudiado en este trabajo son sistemas de alta dimensionalidad, sin embargo, vamos a utilizar el método propuesto por Anishchenko et. al. [67] para baja dimensionalidad para analizar nuestro problema. En su trabajo Anishchenko et. al. analizan un sistema de dos osciladores acoplados en donde cada oscilador está descrito por las ecuaciones:

$$\begin{aligned} \dot{x} &= mx + y - xz \\ \dot{y} &= -x \\ \dot{z} &= -gz + gf(x) \end{aligned} \quad (9.4)$$

con  $m$  el parámetro de excitación del oscilador y  $f(x)$  una función dada por:

$$f(x) = \begin{cases} x^2 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (9.5)$$

El sistema de dos osciladores acoplados estará descrito entonces por el sistema de ecuaciones:

$$\dot{\mathbf{x}}_1 = (m_1 - z_1)\mathbf{x}_1 + \mathbf{y}_1 + G_1(x_2 - x_1 + y_1 - y_2/p),$$

$$\begin{aligned}
 \dot{y}_1 &= -x_1, \\
 \dot{z}_1 &= g_1(f(x_1) - z_1), \\
 \dot{x}_2/p &= (m_2 - z_2)x_2 + y_2 + G_2(Bx_1 - x_2 + y_2 - Bpy_1), \\
 \dot{y}_2/p &= -x_2, \\
 \dot{z}_2/p &= g_2(f(x_2) - z_2),
 \end{aligned} \tag{9.6}$$

donde  $p$  es la razón de las frecuencias de los osciladores,  $B$  es un coeficiente de transmisión y  $G_1$  y  $G_2$  son los parámetros de acoplamiento. En particular el caso que nos interesa es el de un acoplamiento unidireccional, es decir  $G_1 = 0$  y  $G_2 = G$ , en cuyo caso estamos tratando con un problema de sincronización forzada. En la Figura 9.1.a se muestra el espectro de potencia que corresponde al sistema forzante, (en este caso  $x_1$ ), la Figura 9.1.b corresponde al espectro de potencia de las oscilaciones autónomas de  $x_2$ , mientras que en la Figura 9.1.c se muestra el espectro de potencia de las oscilaciones forzadas de  $x_2$  para  $m_1 \approx 0.7$ ,  $m_2 \approx 1.2$ ,  $g_1 = g_2 = 0.3$  y  $p = 1.08$ , los valores del parámetro de acoplamiento  $G$  aumentan de izquierda a derecha.

## 9.4 MODELO DE PRUEBA

Para analizar nuestro problema vamos a construir un modelo de prueba que nos permita hacer un análisis semejante al realizado por Anishchenko et. al. Hemos visto en el capítulo precedente dos modelos de tipo campo medio que muestran una fenomenología similar a la de sistemas unidimensionales. En esta sección vamos a referirnos únicamente al modelo representado por la ecuación (5.17). A partir de este modelo vamos a construir otro modelo que nos permita hacer una análisis similar al discutido en la sección anterior. Para ello vamos a desacoplar parcialmente uno de los mapeos, de tal forma que nuestro nuevo sistema está descrito por las ecuaciones:

$$\begin{aligned}
 N_i(t+1) &= \mathbf{H} \left[ f\{N_i(t)\} + \frac{\beta}{(L+1)} \sum_{i=0}^L f\{N_i(t)\} \right] \quad i = 0, \dots, L-1 \\
 N_L(t+1) &= f\{N_i(t)\}
 \end{aligned} \tag{9.7}$$



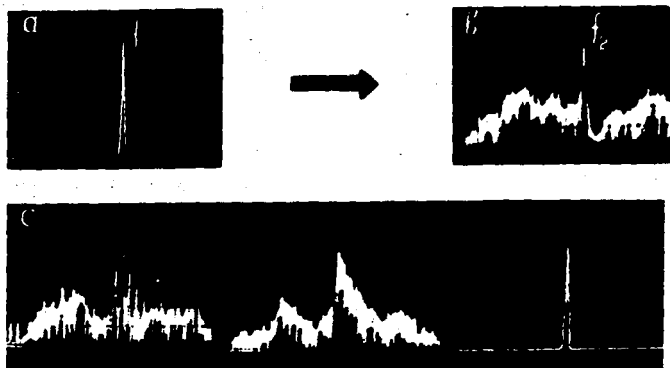


Figura 9.1: Espectros de potencia para a).- las oscilaciones del sistema forzante,  $x_1(t)$ , b).- las oscilaciones autónomas de  $x_2(t)$  y c).- las oscilaciones forzadas de  $x_2(t)$ .

donde como antes:

$$f\{N_i(t)\} = A_i N_i(t) - c N_i^2(t) - d N_i \bar{N}(t), \quad (9.8)$$

con la diferencia de que:

$$A_i = a + b \left( \frac{i}{L} \right)^2 \quad (9.9)$$

Consideremos a la segunda de las ecuaciones (9.7) como el sistema forzante y a la primera como el subsistema. A diferencia del ejemplo descrito por las ecuaciones (9.6), en este caso el sistema forzante continúa acoplado a todos los grados de libertad del sistema original a través del término con el coeficiente  $d$ . Como siguiente paso, con valores fijos de los parámetros  $a$ ,  $c$ ,  $d$  y  $\beta = 0$ , variando el parámetro  $b$  localizamos una zona en donde el mapeo  $N_L(t)$  tenga un atractor periódico. Con los valores de los parámetros  $a$ ,  $b$ ,  $c$  y  $d$  para los cuales el sistema tiene este atractor periódico, empezamos a aumentar el valor del parámetro de acoplamiento  $\beta$  (que juega el papel de  $G$  del caso anterior), y se encuentra que para algunos valores de este parámetro, la variable  $\bar{N}(t)$  (que es una propiedad global del sistema), tiene algunas ventanas del mismo período que el del mapeo parcialmente desacoplado. Es decir este oscilador está forzando a todo el sistema a sincronizarse con su frecuencia. Nótese que el comportamiento observado para  $\bar{N}(t)$  es una condición necesaria pero no suficiente para que cada uno de los mapeos esté sincronizado con el sistema forzante, sin embargo en los casos estudiados se encontró que en efecto cada uno de los mapeos está sincronizado.

El análisis se limitó a dos casos: 1).- Cuando los parámetros toman los valores:  $a = 2.0$ ,  $b = 1.972$ ,  $c = 0.2$ ,  $d = 1$ ,  $L = 10$  y  $\beta = 0$ ; el mapeo parcialmente desacoplado  $N_L(t)$  cae en un atractor de período 6 (Figura 9.2.a). En la gráfica de la Figura 9.2.b se muestra el diagrama de bifurcaciones de  $\bar{N}(t)$  como función del parámetro de acoplamiento  $\beta$ . En esta gráfica se puede observar claramente una ventana de período 6 alrededor de  $\beta = 0.2$ . En la Figura 9.3, se muestran los espectros de potencia para a).-  $\bar{N}(t)$  en las proximidades de la ventana y b).-  $\bar{N}(t)$  en la ventana. En la figura 9.3,  $a = 2.0$ ,  $b = 1.972$ ,  $c = 0.2$  y  $d = 1$ .

En a).- $\beta = 0.199$  y en b).- $\beta = 0.2$ . En todos los casos el espectro de potencia se calculó a partir de una serie temporal de 1024 datos. Cabe mencionar que todos los espectros de potencia (salvo los de la Figura 9.1) se ha suprimido el pico correspondiente a un período dos.

2).- El segundo caso corresponde a valores de los parámetros  $a = 1.5$ ,  $b = 2.36$ ,  $c = 1$  y  $d = 1$ , en cuyo caso se observan atractores de período 10. La fenomenología es semejante a la descrita en el punto 1).-

Comparando los espectros de potencia de la Figura 9.3 con los de la Figura 9.1 se encuentra un comportamiento muy parecido, es decir en el espectro de potencia empieza a ser evidente el advenimiento de un comportamiento periódico antes de que éste pueda detectarse de la simple observación de los valores de la serie temporal. En el caso del mapeo logístico unidimensional (ecuación 7.3) se encuentra un comportamiento similar (Figura 9.4). Sin embargo cuando nos fijamos en los espectros de potencia del oscilador parcialmente desacoplado, no se encuentra este comportamiento (Figura 9.5). Esto puede deberse a que, como se mencionó anteriormente, este oscilador no está totalmente desacoplado.

El análisis anterior se ve reforzado cuando se calcula el espectro de exponentes de Lyapunov. Como se indicó previamente, una condición necesaria para que los subsistemas se sincronicen es que sus exponentes de Liapunov sean todos negativos. Utilizando el algoritmo propuesto por Benetin et. al. [59] y basandonos en una implementación de Livi et. al. [68], se calculó el espectro de exponentes. En la Figura 9.6 se muestran las gráficas de los espectros de exponentes para todo el sistema así como para el subsistema de respuesta, en donde puede verse que efectivamente todos los exponentes tanto del sistema total como del subsistema de respuesta son negativos. En el apéndice A se encuentra el programa utilizado para calcular estos espectros de Lyapunov, con el fin de que sirva como guía al lector interesado. Pequeñas modificaciones en dicho programa permiten calcular los exponentes para los mapeos (5.13), (5.16) y (5.17).

Ahora bien, en el caso general el sistema es algo más complicado

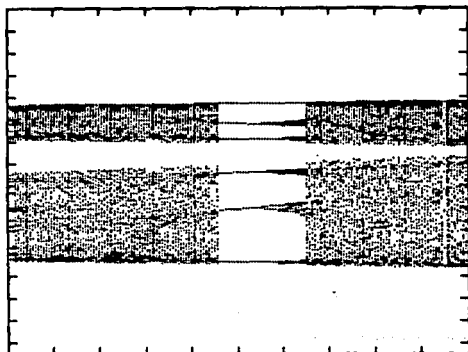
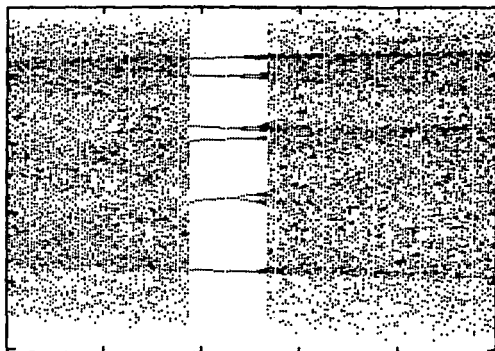


Figura 9.2: Diagramas de bifurcaciones para a).-  $N_L(t)$  y b).-  $N(t)$  en donde se muestran las ventanas de período 6. Los valores de los parámetros se dan en el texto.

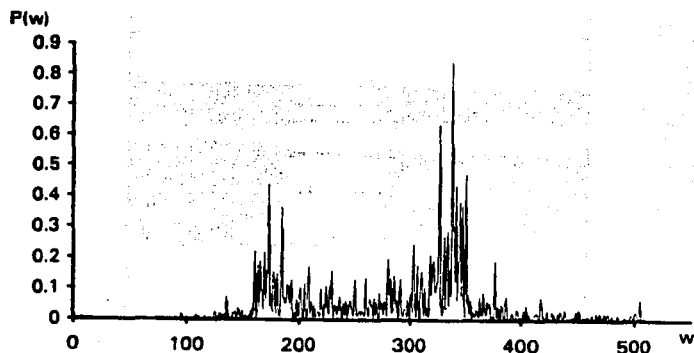
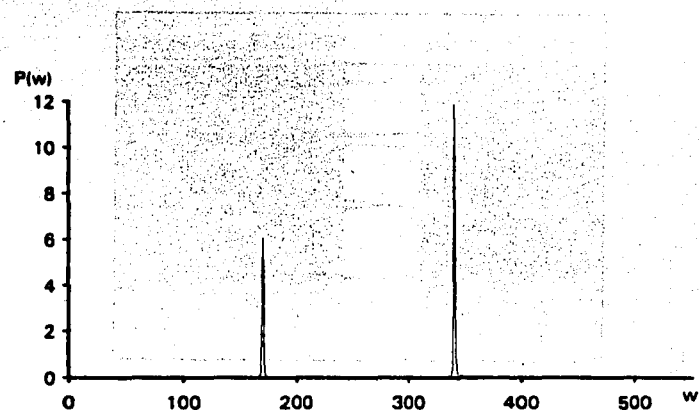


Figura 9.3: Espectros de potencia de  $\bar{N}(t)$  a).-en las proximidades de una ventana de período 6 y b).-en la ventana de período 6.

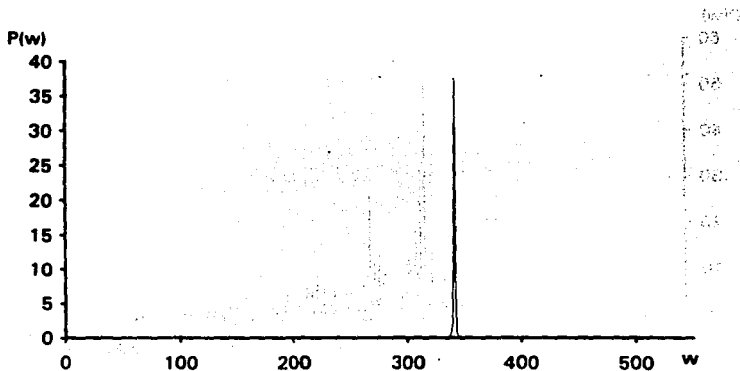
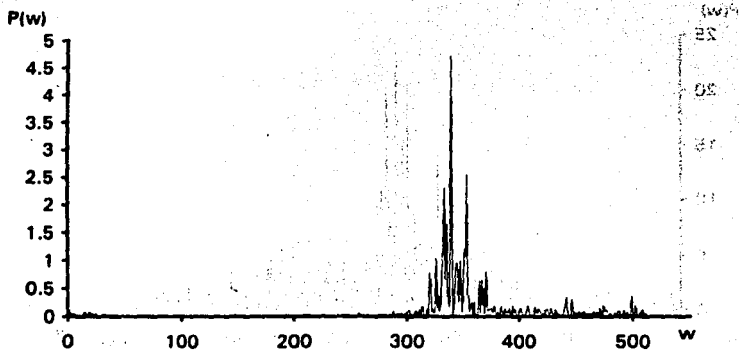


Figura 9.4: Espectros de potencia del mapeo logístico a).-cerca de la ventana de período 3 ( $\mu = 3.828$ ) y b).- en la ventana de período 3 ( $\mu = 3.829$ ).

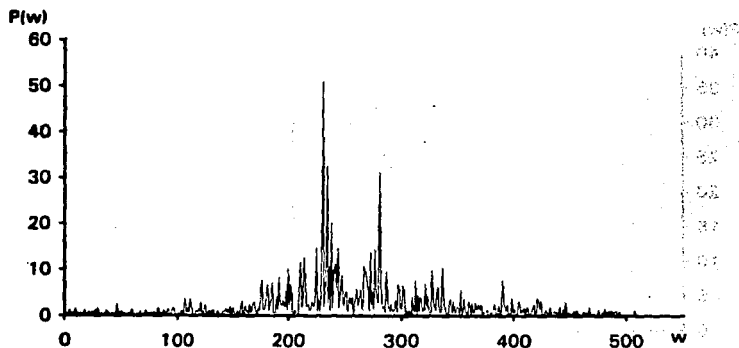
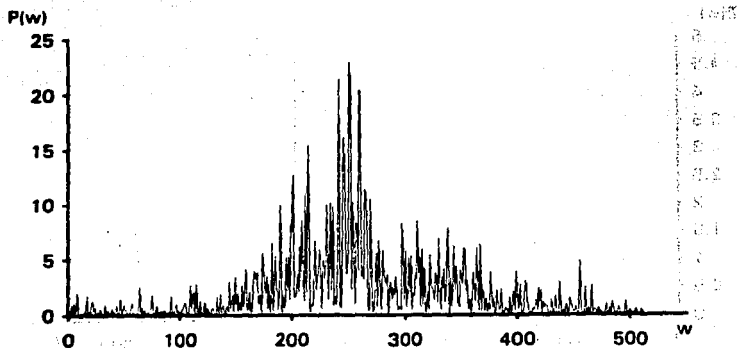


Figura 9.5: Espectros de potencia del oscilador parcialmente desacoplado cerca de la ventana de período 6;  $a = 2.0$ ,  $c = 0.2$ ,  $d = 1$ ,  $\beta = 0$  y  $L = 10$ . En i).-  $b = 1.965$  y en ii).-  $b = 1.968$ .

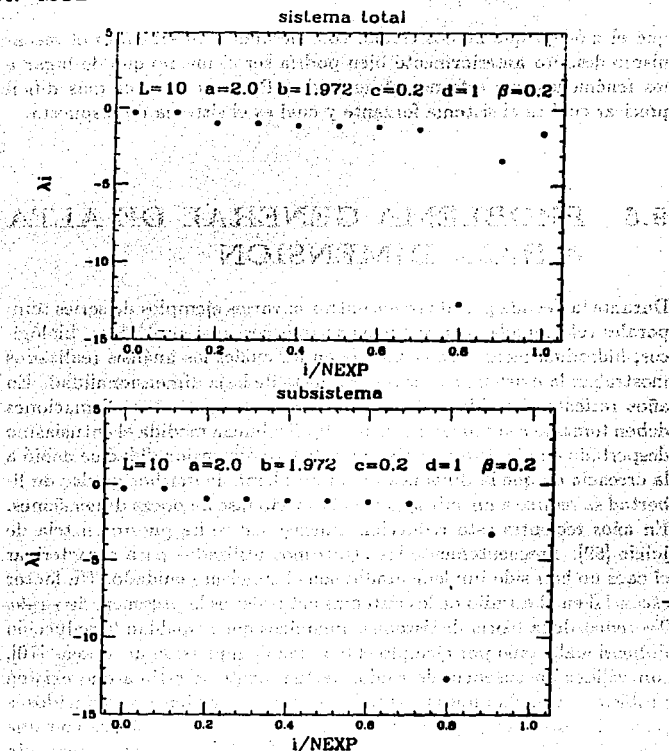


Figura 9.6: Gráficas de los espectros de Lyapunov para a).-el sistema completo dado por las ecuaciones (9.7) y b).-el subsistema de respuesta dado por la primera de las ecuaciones (9.7). Los valores de los parámetros se muestran en las gráficas.



que el modelo que hemos estado considerando, sin embargo el mecanismo descrito anteriormente bien podría ser el mismo que de lugar a los fenómenos que estamos observando. Para ese caso es más difícil precisar cual es el sistema forzante y cual es el sistema de respuesta.

## 9.5 PROBLEMA GENERAL DE ALTA A BAJA DIMENSION

Durante la década pasada se encontraron varios ejemplos de series temporales relacionadas con sistemas astrofísicos, meteorológicos, biológicos, hidrodinámicos y en economía en los cuales los análisis realizados mostraban la existencia caos determinista de baja dimensionalidad. En años recientes ha sido cada vez más evidente que estas afirmaciones deben tomarse con mucho cuidado [69]. En buena medida el entusiasmo despertado en el estudio de sistemas de bajadimensionalidad se debió a la creencia de que la dinámica de un problema de muchos grados de libertad se reduce a un subespacio del espacio fase de pocas dimensiones. En años recientes esta reducción dimensional se ha puesto en tela de juicio [69]. Frecuentemente los algoritmos utilizados para caracterizar el caos no han sido implementados con el suficiente cuidado. Un factor esencial en el estudio de los sistemas naturales es la presencia de *ruido*. Teoremas de la teoría de sistemas dinámicos que respaldan la reducción dimensional, como por ejemplo el teorema de inmersión de Takens [70], son válidos en ausencia de ruido. Actualmente se estima que existen problemas con dinámicas extendidas que no pueden ser reducidos a pocas dimensiones y que requieren de un estudio en espacios con una dimensionalidad menor que el número de grados de libertad pero aún considerable. Bajo esta premisa cabe la pregunta de si la sincronización constituye un mecanismo de importancia en una reducción dimensional.

En la sección 9.1.2 vimos que en el caso de un sistema de dos mapeos acoplados, existen condiciones para las cuales el problema se reduce a un problema de un sistema unidimensional. En la sección 7.1 se mencionó que el esquema del Grupo de Renormalización se introdujo en el

## 9.5. PROBLEMA GENERAL DE ALTA A BAJA DIMENSION 115

análisis de sistemas no lineales con el trabajo de Feigenbaum. Dado que este esquema resulta útil en dicho análisis, es razonable pensar que también podría aplicarse al estudio de redes de mapeos acoplados, las cuales están constituidas por mapeos unidimensionales del tipo de los analizados por Feigenbaum, ya existen trabajos en esta dirección [71]. Los principales elementos del esquema del *Grupo de Renormalización* son:

- a.- Un proceso de reducción dimensional en el sistema. Por ejemplo se pasa de un sistema de  $N$  dimensiones a uno de  $N/2$  dimensiones con una prescripción semejante a la dada en la sección 9.1.2.
- b.- Se busca un nuevo acoplamiento efectivo entre los nuevos componentes.
- c.- Se itera el proceso.

La existencia de una dimensión mínima podría hacerse aparente con la obtención de un punto fijo de la iteración de este proceso. Sin embargo este análisis está más allá del objetivo de este trabajo y es un campo de trabajo abierto.



... de la evolución de las secuencias genéticas, en particular del virus del SIDA. A lo largo del desarrollo del trabajo vimos que este es un problema que requiere de un análisis interdisciplinario y quizás debido a ello existe una gran variedad de tratamientos alternativos y complementarios. El tratamiento utilizado en este trabajo es uno de ellos y está basado en un análisis fenomenológico. La información obtenida a partir de dicho análisis fue de vital importancia para tener criterios adecuados para introducir las aproximaciones y simplificaciones que permitieran obtener algunos resultados analíticos y numéricos.

## Capítulo 10

# DISCUSION Y CONCLUSIONES

En este trabajo se realizó el análisis de un modelo para el estudio de la evolución de secuencias genéticas, en particular del virus del SIDA. A lo largo del desarrollo del trabajo vimos que este es un problema que requiere de un análisis interdisciplinario y quizás debido a ello existe una gran variedad de tratamientos alternativos y complementarios. El tratamiento utilizado en este trabajo es uno de ellos y está basado en un análisis fenomenológico. La información obtenida a partir de dicho análisis fue de vital importancia para tener criterios adecuados para introducir las aproximaciones y simplificaciones que permitieran obtener algunos resultados analíticos y numéricos.

Un aspecto muy importante del presente trabajo se refiere al tipo de modelaje utilizado. Tradicionalmente la evolución dinámica de sistemas extendidos ha estado asociada con el estudio de ecuaciones diferenciales parciales no lineales, sin embargo recientemente se han propuesto herramientas alternativas tales como *autómatas celulares* y *redes de mapeos acoplados*. Es importante recalcar que estos tratamientos alternativos no deben considerarse como discretizaciones de una ecuación diferencial parcial, sino como formalismos que pueden, intrínsecamente, ser más adecuados para el modelaje de algunos fenómenos. El problema tratado en esta tesis es un ejemplo de ello.

La simulaciones numéricas analizadas en el capítulo 6 muestran que, en el caso del VIH, la población de secuencias genéticas presenta una distribución que corresponde a una cuasiespecie. Algunos autores señalan que el hecho de que exista una gran variedad de secuencias del VIH indica que el virus efectivamente evoluciona como una cuasiespecie [46]. Del análisis de las simulaciones obtienen indicaciones de que existe una relación entre el comportamiento de máquinas moleculares, en particular del ribosoma, y propiedades de regularidad en la composición de las secuencias genéticas. En la sección 6.2 se mencionó el hecho de que existe una anticorrelación entre el contenido de codones NDF y la variabilidad de las secuencias. Las gráficas presentadas en dicha sección muestran que aquellas secuencias, o dominios de las secuencias, en donde la composición promedio corresponde a un valor pequeño del peso local  $i$  de codones NDF con respecto a los del tipo NDN, la dispersión es grande. Dada la alta tasa de mutación en el virus del SIDA, se espera una baja frecuencia de aparición de codones de tipo NDF en el VIH-1. Análisis estadísticos realizados con las 25 secuencias de mRNA del VIH-1 tomadas de [52] indican que la anticorrelación predicha se observa en las secuencias genéticas de los bancos de datos [50], [51]. Lo anterior nos hace pensar que en la formulación del modelo se han tomado en cuenta aspectos esenciales del mecanismo de la evolución de secuencias genéticas, a pesar de las simplificaciones introducidas.

Conviene recalcar que la anticorrelación antes mencionada se observa gracias a que las cadenas de ácidos nucleicos se expresaron en el lenguaje degenerado D/F. La adopción de este lenguaje, que es reflejo de las propiedades fisicoquímicas de los ácidos nucleicos, permitió así mismo descubrir algunas características de la dinámica ribosomal y de alguna forma resalta la importancia de las interacciones mRNA-rRNA a nivel ribosomal. Los resultados sugieren que el mRNA se ancla al ribosoma debido a estas interacciones, permitiendo una síntesis de proteínas más precisa. Bajo esta óptica el ribosoma adquiere un valor selectivo. El trabajo de Cocho et. al [50] apoya esta conjetura al demostrar que la anticorrelación arriba mencionada tiene un máximo antes y después del sitio de traducción. Las observaciones experimentales descritas en la sección 3.2 pusieron de manifiesto que algunas características de la composición del RNA ribosomal, tanto en los sitios activos de síntesis de

proteínas como en una vecindad de ellos, regulan el buen desempeño de la síntesis. Las determinaciones de Cocho et. al. [50] muestran que propiedades locales de composición química en regiones fuera de la vecindad de los sitios activos son determinantes en la variabilidad del virus. Estos resultados llevan al planteamiento de un estudio detallado a nivel de máquinas moleculares brownianas de los procesos llevados a cabo por el ribosoma. Para el caso de células eucariotes, el conocimiento de la estructura y funcionamiento del ribosoma es aún muy precario. De confirmarse, en general, las observaciones obtenidas hasta el momento, se tendría una indicación sobre la complejidad estructural del rRNA en el ribosoma.

En búsqueda de una pista terapéutica vimos que la cantidad  $\bar{P}_L(NDF)$  es una medida de la distribución local del triplete  $NDF$  y cuantifica las regularidades estructurales de las secuencias genéticas. Se encuentra que  $\bar{P}_L(NDF)$  calculada para toda la secuencia genética de longitud  $L$ , toma valores diferentes para eucariotes y el VIH-1, sugiriendo una diferencia en las interacciones del ARN que tienen lugar en el ribosoma. Esta diferencia en los valores promedio de  $P_L(NDF)$  sugieren que un ataque químico al ribosoma que altere las interacciones mRNA-rARN podría bloquear la traducción del VIH sin afectar considerablemente la síntesis de proteínas eucariotes. Este ataque químico debería tender a inhibir las funciones del ribosoma para secuencias de mRNA con una concentración baja de codones  $NDF$ . Este efecto podría ser complementario a resultados recientes [53] en los que se muestra que mediante el empleo de *ribozimas* (cadenas de RNA específicas), es posible bloquear la traducción del RNA correspondiente a algunas proteínas del virus. Aún cuando dichos experimentos se realizaron *in vitro* y no se tiene la seguridad de que el proceso pueda tener éxito *in vivo*, la exploración de esta indicación representa la posibilidad de encontrar un tratamiento terapéutico contra el SIDA.

Dada la hipermutabilidad del VIH, es poco probable que la inhibición de un proceso particular de su ciclo vital logre interrumpir la reproducción del virus. En nuestra opinión un tratamiento que involucre varios procesos del ciclo reproductivo tiene mayores posibilidades de éxito. El ataque al ribosoma en los términos que mencionamos es

una alternativa que, hasta donde llega nuestro conocimiento de la literatura, no ha sido planteada. La idea de obstruir el funcionamiento del ribosoma con fines terapéuticos no es nueva. Muchos antibióticos operan de esta forma, [72], [73], sin embargo, en nuestra propuesta se trata de un ataque al ribosoma de la célula hospedera y no al ribosoma de una bacteria o virus. Es por ello que el ataque químico deberá ser muy selectivo y preciso. Para ello se requerirá de un estudio más general y detallado.

Es interesante puntualizar que la anticorrelación a la que hemos hecho referencia, solo mide los efectos de cambio en el nucleótido que ocupa la tercera posición en un codón. Esto sugiere que, apesar de que la información retenida en esa posición es en general irrelevante en la determinación del aminoácido seleccionado, si es de consecuencia en el "mecanismo de lectura" del código genético. Una investigación en la generalidad del fenómeno podría ayudar al entendimiento de la forma en que opera el código genético. La relevancia de nuestras observaciones para el entendimiento del uso de codones, así como su relación con la disponibilidad de tARN, requiere de un estudio discriminatorio detallado.

Desde el inicio en esta tesis nos planteamos como objetivo el estudio de la evolución temporal de sistemas extendidos, con énfasis en redes biológicas y en particular en la evolución de la secuencia genética del virus del SIDA. En la introducción se mencionó que dinámicas caóticas o irregulares pueden ser ilustrativas para problemas biológicos. Bajo esta perspectiva en los capítulos 8 y 9, nos concentramos en el estudio de la dinámica de las ecuaciones que se plantearon en el capítulo 5 para la evolución del VIH, en un régimen de los parámetros para los cuales las condiciones de estacionariedad dejan de cumplirse. En el contexto del estudio generalizado de sistemas dinámicos nuestras ecuaciones son espacialmente inhomogéneas y presentan acoplamiento tanto local como global. El entendimiento presente de la dinámica bajo estas condiciones es muy reducido. En los capítulos 8 y 9 se propusieron dos esquemas con el propósito de contribuir a este entendimiento.

En el capítulo 8 se plantea la relación entre dinámicas en una di-

mención y las de sistemas extendidos. Este punto ha sido tratado de diversas formas en la literatura. Por ejemplo Willeboordse [74] introduce retardo en ecuaciones unidimensionales para recuperar la fenomenología observada en sistemas de muchos grados de libertad. Nuestro tratamiento sigue las líneas tradicionalmente trazadas en física estadística de introducir un campo promedio para reducir, o más bien aproximar, el problema de muchos cuerpos al de un cuerpo inmerso en un campo medio producido por los demás.

En el capítulo 9, abordamos el problema más general de la reducción dimensional en términos de mecanismos de sincronización. Para casos más sencillos se cuenta con esquemas para la formación de cúmulos por medio de sincronización [20] e inclusive de renormalización [71]. En nuestra búsqueda de un esquema apropiado para nuestras ecuaciones intentamos dar algunos indicios de como detectar mecanismos de sincronización. Casos con menos grados de libertad ya han sido estudiados por diversos métodos [60]-[62], [65], [67]. Nuestras observaciones llevan al planteamiento de un esquema de renormalización que resulta ser consistente con trabajos en progreso para dinámicas en el continuo, esto es, para ecuaciones diferenciales parciales, en las cuales se ha demostrado, para algunos sistemas disipativos, la existencia de una *variedad inercial* que recoge la dinámica asintótica [75]. Cabe enfatizar que en este capítulo solamente se plantean las semillas de una futura investigación y no contiene resultados concluyentes.



El presente trabajo de tesis tiene como objetivo principal analizar el comportamiento de los indicadores de desarrollo humano en el Perú, considerando el impacto de las políticas económicas y sociales implementadas durante el periodo de 1980 a 1995. Para ello se realizó un estudio de caso en el departamento de Arequipa, donde se aplicaron cuestionarios a una muestra representativa de la población. Los resultados obtenidos muestran que, a pesar de los esfuerzos realizados por el gobierno peruano, persisten importantes brechas de desarrollo humano, especialmente en el área de educación y salud. Se concluye que es necesario implementar políticas más integrales y sostenibles que permitan mejorar el nivel de vida de la población y reducir las desigualdades sociales.

El presente trabajo de tesis tiene como objetivo principal analizar el comportamiento de los indicadores de desarrollo humano en el Perú, considerando el impacto de las políticas económicas y sociales implementadas durante el periodo de 1980 a 1995. Para ello se realizó un estudio de caso en el departamento de Arequipa, donde se aplicaron cuestionarios a una muestra representativa de la población. Los resultados obtenidos muestran que, a pesar de los esfuerzos realizados por el gobierno peruano, persisten importantes brechas de desarrollo humano, especialmente en el área de educación y salud. Se concluye que es necesario implementar políticas más integrales y sostenibles que permitan mejorar el nivel de vida de la población y reducir las desigualdades sociales.

## Bibliografía

- [1] Kaneko, K.: "Relevance of Dynamic Clustering to Biological Networks." *Physica D* **75** (1994) 55.
- [2] "Constructive Complexity and Artificial Reality", eds. Kaneko, K., Tsuda, I. e Ikegami, T. *Physica D* **75**, Nos. 1-3 (1994).
- [3] Hopfield, J. J.: "Neurons, Dynamics and Computation", *Physics Today*, February (1994) 40.
- [4] Solé, R. V., Bascompte, J. y Valls, J.: "Nonequilibrium Dynamics in Lattice Ecosystems: Chaotic Stability and Dissipative Structures", *Chaos* **2(3)** (1992) 387.
- [5] Kaneko, K.: "Evolution of Host-Parasitoid Network Through Homeochaotics Dynamics", *Chaos* **2(3)** (1992) 397.
- [6] Schuster, P.: "The Physical Basis of Molecular Evolution", *Chemica Scripta* **26B** (1986) 27.
- [7] Kauffman, S. y Johnsen, S.: "Coevolution to the Edge of Chaos: Coupled Fitness Landscapes, Poised States and Coevolutionary Avalanches", *J. Theor. Biol.* **149** (1991) 467.
- [8] Kaneko, K.: "Overview of Coupled Map Lattices", *Chaos* **2(3)** (1992) 279.
- [9] Kaneko K.: "Period Doubling of Kink-Antikink Patterns, Quasiperiodicity in Antiferro-Like Structures and Spatial Intermittency in Coupled Logistic Lattice", *Prog. Theor. Phys.* **72** (1984) 480.

- [10] Waller, I. y Kapral, R.: "Spatial and Temporal Structure in Systems of Coupled Nonlinear Oscillators", *Phys. Rev. A* **30** (1984) 2047.
- [11] Crutchfield, J. P.: "Space Dynamics in Video Feedback", *Physica D* **10** (1984) 229.
- [12] *Nonlinear Science: The Next Decade, Proceedings of the Tenth Annual International Conference of the Center of Nonlinear Studies Los Alamos*, eds. Campbell, D., Ecke, R., y Hyman, J.M. *Physica D* **41**, Nos. 1-3 (1991) Volumen especial.
- [13] *Chaos Focus Issue on Coupled Map Lattices*, ed. Kaneko, K. *Chaos* **Vol.2, No.3** (1992) Número especial.
- [14] *Future Directions of Nonlinear Dynamics in Physical and Biological Systems*, eds. Christiansen, P.L., Eilbeck, J. C., y Parmentier, R. D. *Physica D* **68**, No.1 (1993) Número especial.
- [15] Levy, J. A.: "Pathogenesis of Human Immunodeficiency Virus Infection", *Microbiol. Rev.* **57** (1993) 183.
- [16] Eigen, M.: "The Physics of Molecular Evolution", *Chemica Scripta* **26 B** (1986) 10.
- [17] Cocho, G. y Martínez-Mekler G.: "On a Coupled Map Lattice Formulation for the Evolution of Genetic Sequences", *Physica D* **51** (1991) 119.
- [18] Greene, W. C.: "AIDS and the Immune System", *Sci. Am.* **269(3)** (1993) 67.
- [19] Weiss, R. A.: "How Does HIV Cause AIDS?", *Science* **260**, 28 May 1993, 1273.
- [20] Kaneko, K.: "Clustering, Coding, Switching, Hierarchical Ordering and Control in a Network of Chaotic Elements", *Phys. D* **41** (1990) 137.

- [21] a).-Oono, Y. y Puri, S.: "Study of Phase-Separation Dynamics by Use of Cell Dynamical Systems. I. Modelling", *Phys. Rev. A* **38** (1988) 434.  
b).-Puri, S. y Oono, Y.: "Study of Phase-Separation Dynamics by Use of Cell Dynamical Systems. II. Two Dimensional Demonstrations" *Phys. Rev. A* **38** (1988) 1542.
- [22] a).-Kapral, R.: "Discrete Models for Chemically Reacting Systems" *J. Math. Chem.* **6** (1991) 113.  
b).-Kapral, R.: "Chemical Waves and Coupled Map Lattices", *preprint*.
- [23] Kaneko K.: "Simulating Physics with Coupled Map Lattices", en: *Formation, Dynamics and Statistics of Patterns* eds. Kawasaki, K., Oniki, A., y Suzuki, M., World Scientific (1990) 1.
- [24] May, R. M.: "Biological Populations with Nonoverlapping Generations: Stable points, Stable Cycles, and Chaos", *Science* **186** (1974) 645.
- [25] May, R. M.: "Nonlinear Problemas in Ecology and Resource Management", en: *Comportement Chaotique des Systems Déterministes, XXXVI Session of the Les Houches Summer School 1981*, eds. Iooss, G., Helleman, R. H. G. y Stora, R. North-Holland Publishing Comapany 1983, Amsterdam.
- [26] Beggington, J. R., Free, C. A. y Lawton, J. H.: "Dynamic Complexity in Predator-Prey Models Framed in Diffrence Equations", *Nature* **255** (1975) 58.
- [27] Solé, R. y Valls, J.: "Order and Chaos in a 2D Lotka-Volterra Coupled Map Lattice" *Phys. Lett. A* **153** (1991) 330.
- [28] Styer, L.: *Bioquímica*, Reverté S.A., 1991, España.
- [29] Bohinski R. C.: *Bioquímica*, Addison-Wesley, 1991, U.S.A.
- [30] Darnell, J. E. Jr.: "RNA", *Sci.Am.* 1985, **253** (4), 54.

- [31] Firpo, M. A. y Dahlberg, A. E.: "Post-Transcriptional Control of Gene Expression", eds. McCarthy J. E. G. y Tuite, M. F. *NATO ASI Series H* 49. Springer-Verlag, 1991, Berlin, 185.
- [32] Woose, C. R., Gutell, R. G., Gupta, R. y Noller H.: "Detailed Analysis of the Higher-Order Structure of the 16S Like Ribosomal Ribonucleic Acids", *Microbiol. Rev.* 47 (1983) 621.
- [33] Breslauer, K. J., Frank, B., Blocker, H. y Markey, L. A.: "Predicting DNA Duplex Stability from the Base Sequence", *Proc. Natl. Acad. Sci. USA* 83 (1986) 3746.
- [34] Freier, S. M., Kierczek, R., Jaeger, J. A., Sugimoto, N., Carruthers, M. H., Nelson, T. y Turner, D. H.: "Improved Free-Energy Parameters for Predictions of RNA Duplex Stability", *Proc. Natl. Acad. Sci. USA* 83 (1986) 9373.
- [35] Thomas, L. K., Dix, D. B., y Thompson R. C.: "Codon Choice and Gene Expression: Synonymous Codons differ in their Ability to Direct Aminoacylated-Transfer RNA Binding to Ribosomes *in vitro*." *Proc. Natl. Acad. Sci. USA* 85 (1988) 4242.
- [36] Dix, D. B. y Thompson R. C.: "Codon Choice and Gene Expression: Synonymous Codons Differ in Translational Accuracy" *Proc. Natl. Acad. Sci. USA* 86 (1989) 6888.
- [37] Cocho, G y Rius, J. L.: "Discrete Aspects of Morphogenesis and Gene Dynamics" *Theoretical Biology*, eds. Goodwin, B. y Saunders, P. Edimburgh University Press 1989, 177.
- [38] Cocho, G., Rius, J. L., Medrano, L. y Miramontes P.: "Structural Constraints, DNA Periodicities and Gene Dynamics" en: *Quasicrystals and Incommensurate Structures in Condensed Matter*, eds. José-Yacamán, M., Romeu, D., Castaño, V. y Gómez, A. World Scientific 1990, 465.
- [39] Laurence, J.: "The Immune System in AIDS", *Sci. Amer.* 1985, 253(6), 70.

- [40] Tonegawa, S.: "The Molecules of the Immune System", *Sci. Amer.*, 104.
- [41] Verma, I. M.: "Gene Therapy", *Sci. Amer.* 1990, **263(5)**, 34.
- [42] Gallo, R. C.: "The AIDS Virus", *Sci. Amer.* **256(1)** (1987) 38.
- [43] Schuster, P.: "New Scenarios for Evolution" *Bulletin of the Santa Fe Institute* **6** (1991) 4.
- [44] Kauffman, S.: "Adaptation on Rugged Fitness Landscapes" *Lectures in the Sciences of Complexity, SFI Studies in the Sciences of Complexity, Lect. Vol. I*, ed. Stein, D. Addison-Wesley 1989, 527.
- [45] Martínez-Mekler, G. y Jacobs, L.: Comunicación privada.
- [46] Eigen, M.: "Viral Quasispecies" *Sci. Am.* **269(1)** (1993) 32.
- [47] Eigen, M. y Schuster, P.: "The Hipercycle, a Principle of Natural Self Organization" *Naturwissenschaften* **64** (1977) 541.
- [48] Martínez-Mekler, G., Cocho G., Gelover, A. y Bulajich, R.: "Modelling Genetic Evolution with Coupled Map Lattices", *Rev. Mex. Fís.* **38** Suplemento 1 (1992) 127.
- [49] Cocho, G., Gelover-Santiago, A. y Martínez-Mekler, G.: "An Interplay Between Local and Global Dynamics in Biological Networks: The Case of Genetic Sequences" en: *Cellular Automata and Cooperative Systems*, eds. Boccara, N., Goles, E., Servet-Martínez, S., et. al. Kluwer Academic Publishers, Dordrecht (1993) 85.
- [50] Cocho, G., Martínez-Mekler, G. y Rodin, A.: "Relevance of the Ribosome Structure and Dynamics as a Clue for an Alternative Therapeutical Approach to AIDS", *preprint IFUNAM DSC-93/1*.
- [51] Cocho, G., Gelover-Santiago, A., Martínez-Mekler, G. y Rodin, A.: "Modelling the AIDS Virus Genetic Sequence with Coupled Map Lattices", *preprint IFUNAM DSC-93/2*, en prensa.

- [52] Myers, G. L., Korber, B., Bersovfsky, J. A., et. al.: "Human Retroviruses and AIDS" 1992. Los Alamos National Laboratory. [HIV database].
- [53] Sullenger, B. A. y Cech, T. R.: "Tethering Ribozomes to a Retroviral Pacing Signal for Destruction of Viral ARN", *Science* **262** (1993) 1566.
- [54] Ruelle, D.: *Elements of Differentiable Dynamics and Bifurcation Theory*. Academic Press, Inc., 1989, Boston, MA.
- [55] Wiggins, S.: *Introduction to Applied Nonlinear Dynamical Systems and Chaos. Texts in Applied Mathematics 2*. Springer-Verlag, 1990, New York.
- [56] Lichtenberg, A. J. y Lieberman, M. A.: *Regular and Chaotic Dynamics*. Springer-Verlag, 1983, New York.
- [57] Feigenbaum, M. J.: "The Universal Metric Properties of Nonlinear Transformations", *J. Stat. Phys.* **21** (1979) 669.
- [58] Oseledec V. I.: "A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems" *Trans. Moscow Math. Soc* **19** (1968) 197.
- [59] a).- Benetin, C., Glagani, L., Giorgillo, A. y Strelcyn, J-M.: "Lyapunov Characteristic Exponents for Smooth Dynamical Systems; A Method for Computing All of Them. Part 1: Theory", *Mechanica*, **March 1980** 9.  
b).- Benetin, C., Glagani, L., Giorgillo, A. y Strelcyn, J-M.: "Lyapunov Characteristic Exponents for Smooth Dynamical Systems; A Method for Computing All of Them. Part 2: Numerical Application", *Mechanica*, **March 1980** 21.
- [60] Pecora, L. M., Carroll, T. L.: "Sincronization in Chaotic Systems" *Phys. Rev. Lett* **64(8)** (1990) 821.
- [61] Ditto, W. L. y Pecora L. M.: "Mastering Chaos" *Sci. Am.* **269(2)** (1993) 62.

- [62] Ditto, W. L., Rauseo, S.N. y Spano, M. L. "Experimental Control of Chaos" *Phys. Rev. Lett* **65** (26) (1990) 3211.
- [63] *Proceedings of The Experimental Chaos Conference*. Eds. Vohra, S., Spano, M., Shlesinger, M., Pecora, L. y Ditto W. World Scientific, 1992.
- [64] Garfinkel, A., Spano, M. L., Ditto, W. L. y Weiss J. N.: "Controlling Cardiac Chaos" *Science* **257** (1992) 1230.
- [65] Shinbrot, T., Grebogi, C., Ott, E. y Yorke, A.: "Using Small Perturbations to Control Chaos". *Nature* **263**, 2 June 1993, 411.
- [66] Matsumoto, K. y Tsuda, I.: "Noise-Induced Order". *J. Stat. Phys.* **31** (1) (1983) 87.
- [67] Anishchenko, V. S., Vadivasova, T.E., Postnov, D. E., y Safonova, M. A.: "Synchronization of Chaos". *Int. J. Bifurcation Chaos* **2** (3) (1992) 633.
- [68] Livi, R., Martínez-Mekler, G. y Ruffo, S.: "Periodic Orbits and Long Transients in Coupled Map Lattices" *Physica D* **45** (1990) 452.
- [69] Grassberger, P., Schreiber, T. y Schaffrath, C.: "Non-Linear Time Sequence Analysis". *Int. J. Bifurcation Chaos* **1** (3) (1991) 521.
- [70] Takens, F.: "Detecting Strange Attractors in Turbulence", en: *Lecture Notes in Mathematics Vol. 898*. Springer-Verlag, 1981, New York.
- [71] Kuznetsov, S. P.: "Renormalization Group, Universality and Scaling in Dynamics of Coupled Map Lattices", en: *Theory and Applications of Coupled Map Lattices*, eds. Kaneko, K. John Wiley & Sons Ltd (1993) 51.
- [72] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. y Watson, J. D.: *Molecular Biology. The Cell*. Garland Publishing, Inc., 1989, New York.



- [73] Lewin, B.: *Genes V*, Oxford University Press, 1994, Oxford.
- [74] Willeboordse, F. H.: *Selection of Windows, Attractors and Self-similar Patterns in a Coupled Map Lattice, Chaos, Solitons & Fractals 2(6)* (1992) 609.
- [75] Temam, R.: "Inertial Manifolds", *The Mathematical Intelligencer* Vol. 12, No. 4 (1990) 68.

## Apéndice A

# CALCULO DEL ESPECTRO DE EXPONENTES DE LYAPUNOV.

El cálculo del espectro de exponentes de Lyapunov de un sistema dinámico no-lineal resulta ser de gran ayuda en el entendimiento de la dinámica de dicho sistema. Sin embargo, este cálculo dista mucho de ser trivial debido a que precisamente por ser un sistema no-lineal, su dinámica es lo suficientemente complicada como para dar lugar a divergencias de algunas cantidades durante el cálculo mismo. El programa que se lista a continuación calcula el espectro de exponentes de Lyapunov para un sistema dado tomando en cuenta estas las posibles divergencias. El algoritmo fue propuesto por Benetin et. al [59] y adaptado para un sistema semejante por Livi et. al. [68]. La implementación que aquí se presenta esta diseñada para calcular los exponente de Lyapunov para los mapeos globalmente acoplados dados por las ecuaciones (5.16) y (5.17). El programa esta escrito en lenguaje PASCAL y está implementado en una estación de trabajo IBM-Risc System/6000. El principal objetivo de incluir este listado es que sirva como una guía al lector interesado en el cálculo de estos exponentes.

Al correr el programa se deben proporcionar los siguientes datos:

- i).- los valores de los parámetros involucrados en los mapeo:  $a, b, c, d,$

## 132 APENDICE A. CALCULO DEL ESPECTRO DE LYAPUNOV

- $\alpha$  y  $\beta$ . En el programa *cl* corresponde a *c* y *d1* a *d*.
- ii).- el número de mapeos *L* y el número de exponentes *NEXP* que se desea calcular. En el programa *NE* corresponde a *L*.
- iii).- *IRA* que inicializa el generador de números al azar.
- iv).- El número de iteraciones a lo largo de las cuales se realiza el cálculo está dado por:

$$NS * \sum_{i=0}^{NFI-1} 2^i$$

Los mapeos se inicializan al azar así como los vectores del espacio tangente.

Program LYAPUNOV(Input,Output);

```

Var      i,j,NE,NEXP,NFI,MFI,NS,k,n1,ls,IT,
          T11,iflag,IRA,ICONT,t0,t1:Integer;
          a,b,b1,c,c1,d1,e,e2,alfa,beta,suma,s,
          z:Real;
          N:Array[0..1,0..10] of Real;
          d,SL,EXPON,f:Array[0..10] of Real;
          del,v:Array[0..10,0..10] of Real;
          fi:Text;
Label    2,3;
Const    archivo='NAME=ab.dat';

```

Function RAN(y:Real):Real;

Begin

  y:=(y\*b1+1)-c\*trunc((y\*b1+1)/c);

  RAN:=y/c;

  s:=y;

End;

Procedure GRAM\_SCHMIDT;

Var i,j,im1,m:Integer;  
 s1:Real;  
 s:Array[0..10] of Real;

Label 1;

Begin

For i:=0 to NEXP do

d[i]:=0;

(\*i=0\*)

For j:=0 to NE do

d[0]:=d[0]+v[0,j]\*v[0,j];

d[0]:=sqrt(d[0]);

If d[0]<1e-30 then

Begin

iflag:=0;

Goto 1;

End;

If d[0]=0 then

For j:=0 to NE do

v[0,j]:=0

Else

For j:=0 to NE do

v[0,j]:=v[0,j]/d[0];

For i:=1 to NEXP do

Begin

im1:=i-1;

For m:=0 to im1 do

Begin

s[m]:=0;

For j:=0 to NE do

s[m]:=s[m]+v[m,j]\*v[i,j];

End;

For j:=0 to NE do

Begin

s1:=0;

For m:=0 to im1 do

s1:=s1+v[m,j]\*s[m];

```

    v[i,j]:=v[i,j]-s1;
  End;
  For j:=0 to NE do
    d[i]:=d[i]+v[i,j]*v[i,j];
  d[i]:=sqrt(d[i]);
  If d[i]=0 then
    For j:=0 to NE do
      v[i,j]:=0
    Else
      For j:=0 to NE do
        v[i,j]:=v[i,j]/d[i];
      End;
    End;
  I:
  End;

```

Procedure ELABORACION;

```

Var   ij,k:Integer;
      fl,temp,sumad,sumalfa,sumbeta,sumk:Real;
      k1,JB,vv:Array[0..100,0..100] of Real;

```

Begin

```

  sumad:=suma*d1;
  sumalfa:=suma*alfa;
  sumbeta:=0;
  suma:=0;
  For j:=0 to NE do
    Begin
      fl:=(f[j]-c1*N[t0,j]-sumad)*N[t0,j];
      If fl<0 then
        For i:=0 to NE do
          k1[j,i]:=0
        Else
          Begin
            sumbeta:=sumbeta+fl;
            For i:=0 to NE do
              k1[j,i]:=(f[j]-2*c1*N[t0,j]-sumad)*del[j,i]-
                d1*N[t0,j]/(NE+1);
            End;
          End;
        End;
      End;
  End;

```

```

      End;
    End;
    sumbeta:=beta*sumbeta/(NE+1);
    For j:=0 to NE-1 do
      Begin
        temp:=(f[j]-c1*N[t0,j]-sumad)*N[t0,j] + sumalfa +sumbeta;
        If temp<0 then N[t1,j]:=0 Else N[t1,j]:=temp;
        suma:=suma+N[t1,j];
        If temp<0 then
          For i:=0 to NEXP do
            JB[j,i]:=0
          Else
            For i:=0 to NEXP do
              Begin
                sumk:=0;
                For k:=0 to NE do
                  sumk:=sumk+k1[k,i];
                  JB[j,i]:=k1[j,i]+(alfa/(NE+1))
                    +(beta/(NE+1))*sumk;
                End;
              End;
            j:=NE;
            temp:=(f[j]-c1*N[t0,j]-sumad)*N[t0,j];
            If temp<0 then N[t1,j]:=0 Else N[t1,j]:=temp;
            suma:=suma+N[t1,j];
            If temp<0 then
              For i:=0 to NEXP do
                JB[j,i]:=0
              Else
                For i:=0 to NEXP do
                  Begin
                    sumk:=0;
                    For k:=0 to NE do
                      sumk:=sumk+k1[k,i];
                      JB[i,j]:=k1[j,i]*(alfa/(NE+1))+(beta/(NE+1))*sumk;
                    End;
                  End;
                For i:=0 to NEXP do

```

```

For j:=0 to NE do
  vv[i,j]:=0;
For i:=0 to NEXP do
  For j:=0 to NE do
    For k:=0 to NE do
      vv[i,j]:=vv[i,j]+JB[j,k]*v[i,k];
For i:=0 to NEXP do
  For j:=0 to NE do
    v[i,j]:=vv[i,j];
t1:=t0;
t0:=(t1+1) mod 2;
suma:=suma/(NE+1);
End;

```

Begin

```

Rewrite(fi,archivo);
Write('NFI='); Readln(NFI);
Write('NS='); Readln(NS);
Write('IRA='); Readln(IRA);
Write('NE='); Readln(NE);
Write('NEXP='); Readln(NEXP);
Write('a='); Readln(a);
Write('b='); Readln(b);
Write('c1='); Readln(c1);
Write('d1='); Readln(d1);
Write('alfa='); Readln(alfa);
Write('beta='); Readln(beta);
Writeln(fi,'NFI=',NFI:3,' NS=',NS:3,' IRA=',IRA:3);
Writeln(fi,'NE=',NE:3,' NEXP=',NEXP:3);
Writeln(fi,'a=',a:1:5,' b=',b:1:5,' c1=',c1:1:5,' d1=',d1:1:5);
Writeln(fi,'alfa=',alfa:1:5,' beta=',beta:1:5);
b1:=0.1048581E+07;
c:=0.34359738368E+11;
s:=0.11995147264E+11;
t0:=0;
t1:=1;

```

```

For j:=0 to NE do
  f[j]:=a+b*sqr(j/NE);
For i:=0 to NEXP do
  For j:=0 to NE do
    If i=j then del[i,j]:=1 Else del[i,j]:=0;
  iflag:=1;
  For i:=1 to IRA do
    z:=RAN(s);
  suma:=0;
  For j:=0 to NE do
    Begin
      N[t0,j]:=RAN(s);
      suma:=suma+N[t0,j];
    End;
  suma:=suma/(NE+1);
  For i:=0 to NEXP do
    For j:=0 to NE do
      v[i,j]:=1-2*RAN(s);
  GRAM_SCHMIDT;
  For i:=0 to NEXP do
    SL[i]:=0;
  ICONT:=0;
  For k:=1 to NFI do
    Begin
      MFI:=round(exp((k-1)*ln(2)));
      For n1:=1 to MFI do
        Begin
          For ls:=1 to NS do
            ELABORACION;
          GRAM_SCHMIDT;
          If iflag=0 then
            Begin
              IT:=5*(MFI-1+n1);
              Writeln(fi,'d[0]<1E-30 en:',IT:3,' iterac');
              Goto 2;
            End;
          For i:=0 to NEXP do

```



## 138 APENDICE A. CALCULO DEL ESPECTRO DE LYAPUNOV

```

If d[i]=0 then SL[i]:=-1e30
Else SL[i]:=SL[i]+ln(d[i]);
For i:=0 to NEXP do
  if SL[i]>0 then goto 3;
ICONT:=ICONT+1;
If ICONT > 1 then goto 3
Else
  Begin
    IT:=5*(MFI-1+n1);
    Writeln(fi,'Todos neg. en:',IT:3,' iterac.');
```

3:

```

T11:=(2*MFI-1)*NS;
For i:=0 to NEXP do
  EXPON[i]:=SL[i]/T11;
Writeln(fi,T11:3);
For i:=0 to NEXP do
  Begin
    Write(fi,EXPON[i]:1:10,' ');
    If i mod 5 = 0 then Writeln(fi);
  End;
Writeln(fi);
End;
```

2:

```

End;
End.
```