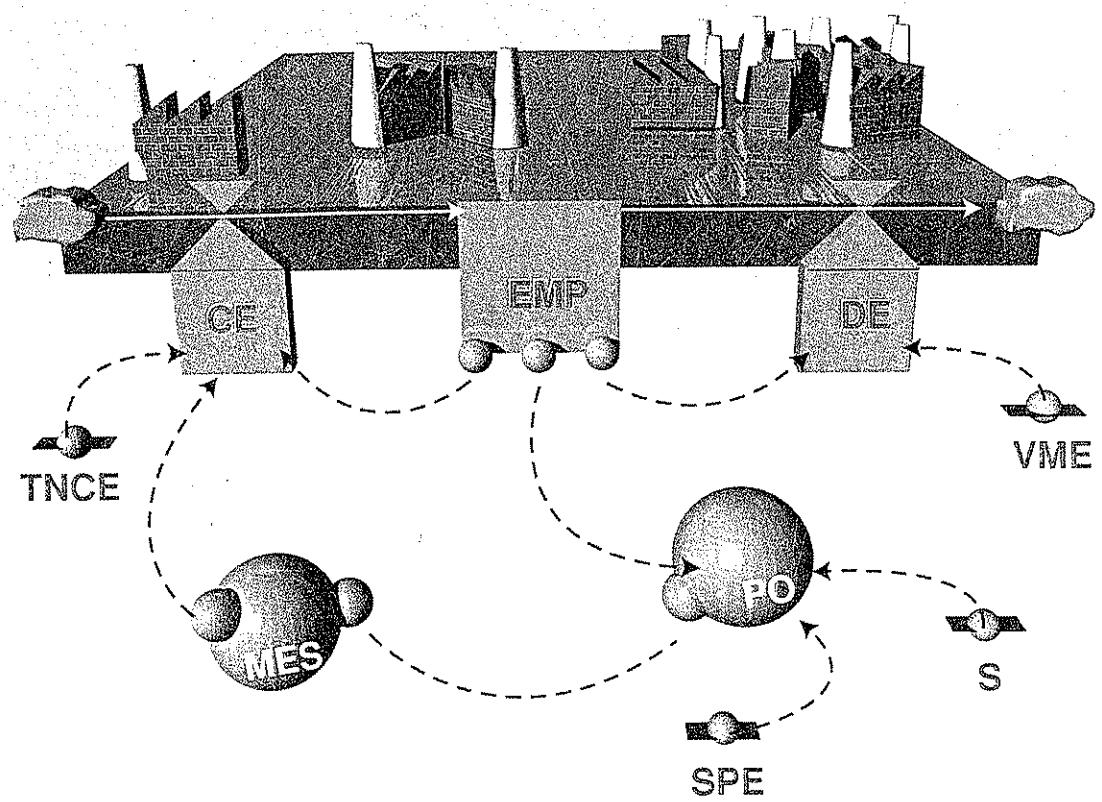


Javier Aracil y Francisco Gordillo Dinámica de sistemas



Alianza
Universidad Textos

Javier Aracil
y Francisco Gordillo

Dinámica de sistemas

Alianza
Editorial

Primera edición: 1997
Tercera reimpresión: 2007

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión y/o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicasen públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© Javier Aracil y Francisco Gordillo
© Alianza Editorial, S. A., Madrid, 1997, 2002, 2005, 2007
Juan Ignacio Luca de Tena, 15; teléf. 91 393 88 88; 28027 Madrid
www.alianzaeditorial.es
ISBN: 978-84-206-8168-9
Depósito legal: M. 46.541-2007
Fotocomposición: EPCA, S. A.
Parque Industrial «Las Monjas»
Torrejón de Ardoz - 28850 Madrid
Impreso en Anzós, S. L. - Fuenlabrada (Madrid)
Printed in Spain

SI QUIERE RECIBIR INFORMACIÓN PERIÓDICA SOBRE LAS NOVEDADES DE
ALIANZA EDITORIAL, ENVÍE UN CORREO ELECTRÓNICO A LA DIRECCIÓN:
alianzaeditorial@anaya.es

ÍNDICE

CAPÍTULO 1. MODELOS DE SIMULACIÓN INFORMÁTICA DEL COMPORTAMIENTO ...	11
1.1. Introducción	11
1.1.1. Concepto de sistema	11
1.1.2. La estructura de realimentación.....	14
1.1.3. Dinámica y comportamiento.....	15
1.2. Concepto de modelo.....	17
1.3. Simulación del comportamiento	19
1.4. Elementos básicos del método sistémico	19
1.5. Campos de aplicación de la dinámica de sistemas	21
CAPÍTULO 2. ELEMENTOS BÁSICOS DE UN LENGUAJE SISTÉMICO.....	25
2.1. Atributos y sistemas	25
2.2. Estructuras de realimentación simple	29
2.2.1. Bucle de realimentación negativa	29
2.2.2. Formulación matemática de un bucle elemental de realimentación negativa	31
2.2.3. Bucles de realimentación positiva	33
2.2.4. Formulación matemática de un bucle elemental de realimentación positiva	34
2.2.5. Retrasos	38
2.3. Estructuras complejas de realimentación	39
2.3.1. Arquetipo del crecimiento sigmoidal.....	40
2.3.2. Formulación matemática de un crecimiento sigmoidal	42
2.3.3. Arquetipo de la adicción	45
2.3.4. Arquetipo del crecimiento con inversión insuficiente	48

CAPÍTULO 3. MODELADO MEDIANTE DINÁMICA DE SISTEMAS.....	55
3.1. Introducción	55
3.2. De la estructura del comportamiento.....	56
3.3. Diagramas de Forrester	58
3.3.1. Variables de estado	61
3.3.2. Variables de flujo	62
3.3.3. Variables auxiliares	64
3.3.4. Otros símbolos empleados en los diagramas	65
3.4. Retrasos	66
3.4.1. Retrasos en las transmisiones de material.....	69
3.4.2. Retrasos en la transmisión de información	74
3.5. Programación del modelo en el computador	75
3.5.1. Los datos del modelo	77
3.6. Dinámica de sistemas y sistemas dinámicos	78
CAPÍTULO 4. EJEMPLOS ELEMENTALES DE MODELOS	83
4.1. Introducción	83
4.2. Bucle de realimentación negativa	83
4.3. Bucle de realimentación positiva	88
4.4. Crecimiento sigmoidal.....	91
4.5. Oscilaciones	93
4.6. Adicción	99
4.7. Inversión insuficiente	103
CAPÍTULO 5. CONSTRUCCIÓN DE MODELOS NO ELEMENTALES EN DINÁMICA DE SISTEMAS.....	107
5.1. Introducción	107
5.2. Las fases en la construcción de un modelo	108
5.2.1. Fase de conceptualización	109
5.2.2. Fase de formulación	109
5.2.3. Fase de evaluación	109
5.3. Estudio de una aplicación de la dinámica de sistemas	110
5.3.1. Fase de conceptualización	110
5.3.2. Fase de formulación	111
5.3.3. Fase de evaluación	114
5.4. Descripción de un modelo de demanda eléctrica: MEREA3	116
CAPÍTULO 6. MÉTODOS AUXILIARES DE LA DINÁMICA DE SISTEMAS	121
6.1. Introducción	121
6.2. Linealización de sistemas no lineales: método de las perturbaciones....	121
6.2.1. Linealización de funciones sin memoria.....	122

6.2.2. Ejemplo: linealización del modelo de desarrollo de un área industrial	123
6.3. Análisis de sensibilidad	126
6.3.1. Análisis de sensibilidad por linealización	127
6.3.2. Ejemplo: análisis de sensibilidad del modelo de crecimiento de un área industrial	129
6.4. Estimación de parámetros en modelos dinámicos	132
6.4.1. Ejemplo: estimación de los parámetros del modelo de crecimiento de un área industrial	135
CAPÍTULO 7. BIFURCACIONES Y ANÁLISIS CUALITATIVO EN DINÁMICA DE SISTEMAS	139
7.1. Análisis cualitativo de sistemas dinámicos	139
7.1.1. Ciclos límite	142
7.1.2. Comportamientos extraños de un sistema determinista.....	143
7.2. Bifurcaciones en sistemas dinámicos	146
7.2.1. Análisis cualitativo del modelo de un área industrial	149
CAPÍTULO 8. SISTEMAS COMPLEJOS	153
8.1. Estructuras jerarquizadas	153
8.1.1. Sistemas dinámicos con dos escalas de tiempo	154
8.2. La estructura de reacción-difusión y la morfogénesis	159
8.2.1. Mecanismo morfogenético de Turing	161
8.2.2. Aplicación al modelo de un área industrial.....	166
8.2.3. Orden mediante fluctuaciones	167
CAPÍTULO 9. EXPLOTACIÓN DE LOS MODELOS	171
9.1. Introducción	171
9.2. El modelo en un proceso de actuación racional.....	171
9.3. Uso de modelos para búsqueda de actuaciones racionales	174
9.3.1. Optimización estática.....	174
9.3.2. Control óptimo	175
9.3.3. Control óptimo en bucle cerrado	181
9.3.4. Ejemplos	182
9.3.5. Otras consideraciones	188
9.4. Principio de racionalidad acotada	189
9.5. El principio de racionalidad acotada y la dinámica de sistemas	191
APÉNDICE A. ENTORNOS INFORMÁTICOS EN DINÁMICA DE SISTEMAS	195
BIBLIOGRAFÍA	197

1.1. INTRODUCCIÓN

Este libro está dedicado al estudio de la dinámica de sistemas. Con esa locución se alude a un método para el estudio del comportamiento de sistemas mediante la construcción de un modelo de simulación informática que ponga de manifiesto las relaciones entre la estructura del sistema y su comportamiento. En esta introducción vamos a analizar los diferentes términos que aparecen en la anterior caracterización. Así, en primer lugar, explicaremos el concepto de sistema en el sentido usado aquí. Despues trataremos de precisar qué entendemos por comportamiento de un sistema, y de esbozar la relación existente entre ese comportamiento y la estructura del sistema, lo que nos llevará a considerar la dinámica de un sistema. Nos detendremos en una estructura particularmente interesante, como es la estructura de realimentación, que nos permitirá tener un primer ejemplo de una propiedad propiamente sistémica, en la que el modo de comportamiento depende de la estructura. Con ello ya estaremos en disposición de analizar qué entenderemos por un modelo y en particular de un modelo de simulación informática. Con todo tendremos un primer esbozo de la dinámica de sistemas que consideraremos como un método concreto para el estudio de sistemas dentro del campo disciplinario de más amplias pretensiones que es la sistémica. Cerraremos el capítulo con una rápida revisión de los campos de aplicación de este método de estudio de sistemas.

1.1.1. Concepto de sistema

El término *sistema* posee varias acepciones en su uso ordinario. Por ejemplo, decimos que tenemos un sistema o método para resolver una cierta situación problemática. No es ése el sentido que nos interesa aquí. Para nosotros un sistema es un objeto formado por un conjunto de partes entre las que se establece alguna forma de relación que las articula en la unidad que es precisamente el sistema. Un sistema se nos manifiesta como un aspecto de la realidad dotado de cierta complejidad precisamente por estar formado por partes en interacción. Esta interacción

12 Dinámica de sistemas

coordina a las partes dotando al conjunto de una entidad propia. Las partes y la interacción entre ellas son los elementos básicos en esta concepción de sistema. Un sistema se percibe como algo que posee una entidad que lo distingue de su entorno, aunque mantiene interacción con él. Esta identidad permanece a lo largo del tiempo y bajo entornos cambiantes.

Esta acepción del término sistema se aplica en múltiples campos. Por ejemplo, hablamos del sistema planetario formado por los planetas ligados por las fuerzas gravitatorias que se ejercen entre ellos y el Sol. Hablamos también de un sistema ecológico formado por diferentes poblaciones entre las que se establecen relaciones de cooperación o de predación, formándose en este último caso cadenas alimentarias. Asimismo aludimos a un sistema económico formado por los diferentes agentes entre los que se producen relaciones de intercambio de bienes y prestación de servicios. Análogamente decimos de una empresa que es un sistema en el que se coordinan sus distintos departamentos para, por una parte, producir el producto o servicio que justifica su existencia y, por otra, asegurar la adecuada retribución del trabajo y del capital en ella involucrados. La lista de ejemplos es interminable, ya que de casi todo lo que nos rodea se puede decir que es un sistema. Incluso en física fundamental el propio concepto de átomo con el que soñaba Demócrito ha sido desbancado por un enjambre de partículas elementales coordinadas.

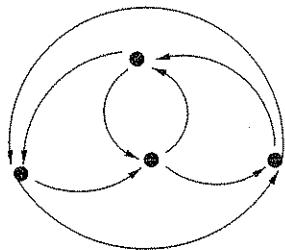


FIGURA 1.1. Grafo que representa un sistema.

Por tanto, para nosotros, un sistema es un cierto aspecto de la realidad al que podemos adscribir una descripción en la que básicamente se enuncian una serie de partes componentes y una forma de interacción entre ellas que suministre un vínculo que las organice en la unidad que es el sistema. La descripción más elemental que podemos hacer de un sistema es sencillamente enunciar el **conjunto C de sus partes y la relación R que establece la vinculación que se produce entre ellas**. Es decir, un sistema es el objeto que admite al menos una descripción mínima según la cual está formada por un conjunto C y una relación R entre los elementos de C . En esta anterior definición los elementos básicos son el par (C, R) . A esta descripción podemos asociar la imagen gráfica de un grafo como el que se muestra en la figura 1.1, cuyos nodos denotan esas partes y cuyas aristas representan las influencias que se producen entre ellas. Este grafo aporta una descripción de naturaleza estructural del sistema y por ello se dice que representa su estructura. Un ejemplo más concreto de grafo de un sistema aparece en la figura 1.2 que muestra el grafo de un sistema demográfico.

Las partes de un sistema se caracterizan mediante sus atributos, de modo que las influencias entre las partes se convierten en relaciones entre los atributos característicos. En el capítulo 2 estudiaremos con detalle estructuras elementales de sistemas.



FIGURA 1.2. Grafo de un sistema demográfico.

En todos los ejemplos de sistemas que hemos considerado anteriormente se trataba de cosas naturales o sociales, pero siempre materiales o concretas. A estos sistemas aludimos corrientemente diciendo que forman parte del mundo externo o extralingüístico que nos afecta con sus solicitudes. Por otra parte, existen los sistemas formales, formados por objetos abstractos (símbolos y relaciones) que aportan los útiles con los que representar a esos sistemas concretos que forman nuestro entorno. El objetivo del constructor de un modelo de dinámica de sistemas será precisamente el producir esos objetos formales con los que representar sistemas del mundo real, de forma análoga a como un mapa representa un ámbito geográfico.

En ese sentido conviene precisar los dos usos habituales del término *sistema*. Por una parte, se emplea para designar un aspecto del mundo, como en los ejemplos más arriba mencionados de sistemas ecológicos o sociales. Por otra, para referirse al objeto abstracto mediante el cual se describen esos sistemas reales. Esta distinción no es exclusivamente terminológica, sino que subyace en ella la distinción entre el objeto S del mundo real que se trata de estudiar, y su representación mediante un objeto matemático adecuado (C, R), al que nos referiremos como un modelo M . Sin embargo, no es extraño que se confundan las dos acepciones, lo que conduce a identificar la representación con lo representado. Los que defienden que esa identificación no es grave dirán que precisamente se construye el modelo para poder hacerla, para tener una representación que sea réplica exacta de lo representado. Sin embargo, esta afirmación hay que tomarla con las oportunas reservas, ya que nunca podemos agotar un cierto aspecto de la realidad con su representación.

Conviene observar que en todos los ejemplos mencionados más arriba el término sistema aparece adjetivado: sistema planetario, sistema ecológico, sistema económico... Esta necesidad de adjetivación sugiere que algunas de sus propiedades están subsumidas bajo el adjetivo, mientras que otras lo están bajo el propio término sistema. El adjetivo describe lo que es particular de la clase de sistemas que se están considerando, mientras que *sistema* alude a aquellas propiedades que son independientes de la naturaleza concreta del sistema en cuestión. En todo sistema S cabe considerar, por una parte, propiedades que dependen de la naturaleza de las partes C que lo forman; es decir, propiedades de las partes que se manifiestan en el sistema. Por otra, S posee propiedades ligadas a la forma de organizarse las partes en el sistema, propiedades que cabe asociar, en la descripción del sistema con la que lo conceptualizamos, a la relación R .

Estas propiedades son las propiamente sistémicas y de ellas emana la noción de sistemicidad. Vamos a ver en el concepto de realimentación una propiedad de esta naturaleza.

1.1.2. La estructura de realimentación

Una estructura básica en el estudio del comportamiento de un sistema es la estructura de realimentación, por lo que conviene dedicarle algún espacio en esta introducción. Para ilustrarla supongamos un proceso muy sencillo de la vida ordinaria como es el llenar un vaso (figura 1.3). En este caso, el grifo, el vaso y la persona que lo llena son las partes del sistema. El proceso de llenar un vaso de agua presupone que el que lo realiza abre el grifo, observa cómo aumenta el nivel del agua en el vaso y lo cierra cuando estima que el nivel alcanzado es el deseado. Podemos describir lo que sucede diciendo que el que llena el vaso compara el nivel alcanzado con el nivel deseado y actúa sobre el grifo en función de esta discrepancia, de modo que según disminuya la diferencia de niveles irá cerrando el grifo hasta hacerlo definitivamente cuando la discrepancia se anule.



FIGURA 1.3. *Proceso de llenar un vaso.*

En la figura 1.4 se ha superpuesto al dibujo de la figura 1.3 un diagrama en el que de forma esquemática se indican las influencias implícitas en la anterior descripción del proceso. Este diagrama básico se muestra de forma aislada en la figura 1.4b. En el diagrama se han indicado mediante flechas las influencias que se producen entre los distintos elementos que intervienen en la descripción del proceso. Así, se indica mediante una flecha que el flujo de agua influye sobre el nivel alcanzado en el vaso. Además, mediante un signo + se indica que la influencia es positiva, lo que quiere decir que a mayor flujo de agua, mayor nivel alcanzado. Análogamente al crecer el nivel disminuye la discrepancia, lo que se indica mediante una flecha a la que se asocia un signo -. Por último, la discrepancia se determina a partir del nivel deseado y del nivel alcanzado.

Este ejemplo constituye una muestra de cómo se puede analizar un sistema, descomponerlo en sus elementos esenciales y relacionar estos elementos mediante un bosquejo de cómo se producen las influencias entre ellos. Se tiene así la descripción más elemental que podemos tener de ese sistema, que se limita a establecer sus elementos componentes básicos y las influencias entre ellos. En este ejemplo concreto se ilustra también lo que se conoce como un

que se aplica en la programación de los sistemas de control y en la simulación informática. El desarrollo de la teoría de sistemas dinámicos es una de las principales contribuciones de la ciencia moderna.

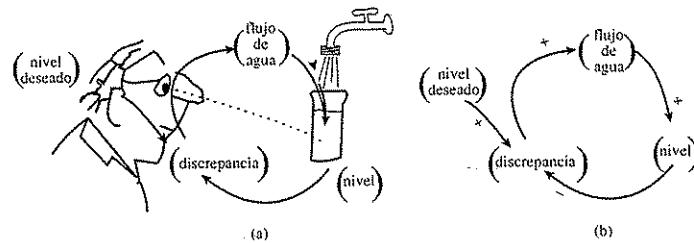


FIGURA 1.4. Diagrama básico del proceso de llenar un vaso de agua: (a) con un grafo orientado; (b) con un grafo signado.

bucle de realimentación, ya que en él se produce una transmisión de información circular de forma continua.

Esta estructura causal circular aparece en múltiples situaciones y está en el origen de comportamientos complejos, como veremos a lo largo de este libro. Existe una metáfora que de forma gráfica ilustra ese hecho. Se supone que alguien está inmerso entre dos tipos de problemas que se ilustran mediante losas como hace la figura 1.5a. Al tratar de liberarse de una de ellas, empujándola para apartarla, se produce el fenómeno que se ilustra en la figura 1.5b. La causalidad circular implícita en el proceso de realimentación hace que aparentes soluciones a unos problemas producen deterioros importantes en otros. Existen muchos ejemplos de estructuras de realimentación en los sistemas en los que estamos inmersos.

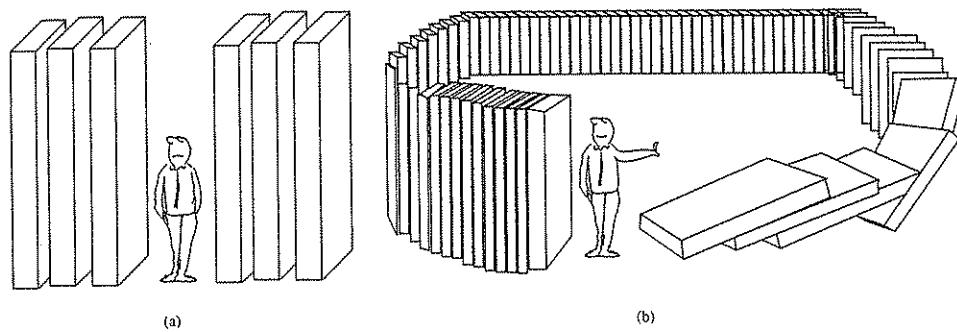


FIGURA 1.5. Ilustración de los problemas asociados a las estructuras de realimentación.

1.1.3. Dinámica y comportamiento

En la locución *dinámica de sistemas*, junto con el término sistema aparece el de dinámica. A un nivel superficial, el término dinámica lo empleamos por oposición al de estática, y queremos expresar con él el carácter cambiante de aquello que adjetivamos con ese término. A algo que

16 Dinámica de sistemas

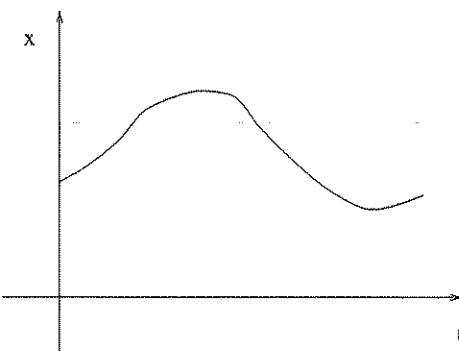


FIGURA 1.6. Trayectoria que describe el comportamiento de una magnitud x .

cambia con el tiempo asociamos una imagen como el de la figura 1.6 en la que se muestra la trayectoria de una magnitud.

El concepto de trayectoria es algo con lo que estamos familiarizados. En principio, evoca la imagen de un proyectil que se desplaza por el espacio. Pero también podemos emplearlo para referirnos a la evolución de un indicador económico (la figura que muestran los periódicos de la evolución de la Bolsa, por ejemplo) o, en general, de cualquier magnitud que cambie con el tiempo. La representación gráfica de las trayectorias muestra el comportamiento del sistema al que se asocian esas magnitudes, ya que es una imagen geométrica de ese comportamiento. Se emplea aquí comportamiento en un sentido preciso: entenderemos por comportamiento la evolución a lo largo del tiempo de las magnitudes que se consideran relevantes para caracterizar los objetos considerados.

Cuando hablamos de estudiar su comportamiento, estamos asumiendo que los sistemas cambian con el tiempo (es decir, que los atributos asociados a ellos sufren variaciones; por eso registramos sus trayectorias) y que nos interesa dar razón de esos cambios (buscamos una descripción racional del comportamiento). Resulta especialmente importante el caso en que esos cambios se generen endógenamente; es decir, resulten básicamente de las tensiones que se producen en el seno del sistema, mediante las relaciones entre las partes que lo forman. Ello no excluye la posibilidad de considerar la acción de eventuales factores externos, pero para nosotros lo más significativo en la descripción del comportamiento serán las propias tensiones internas. Estas tensiones vienen determinadas por las interacciones entre las partes del sistema, que suministran, a su vez, el vínculo que articula a esas partes en la entidad que es el sistema. En este sentido hay que resaltar que el término dinámico tiene una connotación no sólo de cambio, sino de la fuerza (del dinamismo) que lo engendra.

El conjunto de las trayectorias de las magnitudes asociadas a un sistema dan cuenta de su comportamiento durante un periodo de tiempo determinado. Muestran una imagen gráfica de qué ha hecho el sistema durante ese periodo. En esas trayectorias se ponen de manifiesto los cambios que se han producido en los atributos asociados a S . Si queremos profundizar en el conocimiento del sistema interesa conocer a qué pautas responden esos cambios (en el caso de que no sean enteramente arbitrarios o caprichosos). En ello reside la pretensión de un análisis racional del comportamiento de un sistema.

En algunos casos es posible *dar razón* de esas pautas a partir de la estructura del sistema; es decir, es posible asociar modos de comportamiento a determinadas formas de organizar las influencias entre las partes de un sistema. En el capítulo 2 se tratarán ejemplos de estos casos.

Para el estudio del comportamiento de los sistemas disponemos de un objeto matemático especialmente adecuado: el **sistema dinámico**. Formalmente un sistema dinámico es el objeto matemático formado por un espacio de estados X y una regla que prescribe como varían estos estados a lo largo del tiempo. Esta regla se puede expresar de diferentes formas, pero a nosotros nos interesará fundamentalmente la siguiente:

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

en donde la función f expresa precisamente la regla que rige el cambio dx/dt ¹ que se produce en el estado $x \in X$.

El concepto de *sistema dinámico* se origina en la mecánica clásica para describir cómo se produce la variación de la posición y la velocidad de las partículas materiales en función de las fuerzas que se producen entre ellas. El uso de este concepto se puede generalizar al caso en que en lugar de posiciones y velocidades de un sistema mecánico se consideren los distintos atributos asociados a las partes de un sistema; y en vez de las fuerzas que interactúan entre partículas se consideren las relaciones de influencia entre ellas que suministran la trabazón que articula las partes del sistema. De este modo el concepto de *sistema dinámico* aporta el lenguaje para la descripción del comportamiento de un sistema. Veremos cómo la dinámica de sistemas suministra un método para transcribir la descripción elemental de un sistema en un sistema dinámico. De este modo para el estudio de un sistema concreto construiremos un objeto formal, que será un sistema dinámico en el sentido matemático del término, y con el que tenemos un modelo de ese aspecto de la realidad que tratamos de estudiar.

1.2. CONCEPTO DE MODELO

Otro concepto que conviene revisar es el de modelo. Como sucede con sistema, el término modelo se emplea con múltiples sentidos. De estos usos ordinarios hay dos claramente contrapuestos que conviene recordar para precisar el que a nosotros nos interesa. Por una parte, del que posa para un pintor o para un fotógrafo decimos que es su modelo. En este caso se entiende por modelo el referente de lo representado. Por el contrario, también hablamos de modelo al referirnos a una maqueta que pretende reproducir un determinado aspecto de la realidad. En este segundo caso hablamos de modelo como representación. El modelo es un objeto que representa a otro. Éste es el sentido del término modelo que nos interesa aquí. Y así diremos que para un observador \mathcal{O} un objeto \mathcal{M} es un modelo de un objeto \mathcal{S} (un sistema), si \mathcal{O} se puede servir de \mathcal{M} para responder a cuestiones que le importan con relación a \mathcal{S} ². Es decir, un

¹ En este libro se utilizan indistintamente las notaciones dx/dt y \dot{x} para representar a la derivada de la variable x respecto del tiempo t .

² M. Minsky, 1968, "Matter, mind and models", en *Semantic Information Processing*, M. Minsky (ed.), pp. 425-432. The MIT Press.

modelo M es un *instrumento* que ayuda a \mathcal{O} a responder preguntas acerca de un aspecto de la realidad al que convenimos en considerar un sistema concreto S . Conviene resaltar el carácter de instrumento del modelo. Es un medio para algo (habitualmente una ayuda a la toma de decisiones, en un sentido amplio) y no un fin en sí. Sirve, aquí y ahora, para ayudar a resolver un problema concreto, que ha motivado su construcción. Normalmente no tiene un carácter definitivo.

En esta definición de modelo, la presencia del observador \mathcal{O} es básica, ya que toda descripción (y, en consecuencia, todo modelo M) lleva asociada un observador \mathcal{O} . El modelo lo es *para él*, y para aquellos con los que comparte un lenguaje con el que describir una determinada forma de ver un cierto aspecto de su entorno. El modelo no es —al menos no tiene por qué serlo— una copia (de aquel aspecto) que tenga un valor descriptivo independiente de cómo se realiza, sino que está asociada a una interpretación de la realidad construida con los útiles aportados por un lenguaje de modelado.

Con un modelo se pretende describir un cierto fenómeno o proceso S . Por tanto, recogerá sólo aquellos aspectos que —en opinión de su constructor, y según su buen saber y entender— resulten relevantes con relación a S . Presupone, por tanto, la adopción de un criterio de relevancia con respecto a aquello que se va a incluir en el modelo. No existen descripciones neutrales.

El proceso mediante el cual \mathcal{O} construye M recibe la denominación de *proceso de modelado*. Con su concurso se procede a la construcción de un objeto artificial: el modelo M . En todo proceso de modelado se pueden distinguir, al menos, tres aspectos:

- Una problemática concreta con relación a S . Como ya hemos apuntado, un modelo nunca puede pretender agotar la realidad de S , sino que sólo atiende a determinados aspectos suscitados por un problema concreto. Este problema es el que ha determinado la decisión de construir el modelo (normalmente se trata de decidir sobre unas acciones a realizar con respecto a S para alcanzar ciertas metas concretas).
- La experiencia previa relativa a otros S , análogos a S . Esta experiencia puede ser propia o ajena; puede, a su vez, constituir un cuerpo de doctrina organizado o ser simplemente un conjunto de opiniones. Pero, en todo caso, se trata de la información de que se dispone con relación a S .
- Un medio de expresión que permita a \mathcal{O} realizar M , al que podemos denominar *lenguaje de modelado*, que suministra los módulos básicos (conceptos y símbolos) a partir de los cuales se construye el modelo. Estos útiles suministran un marco desde el cual *ver* y, a partir de esa *visión*, representar a S . Combinando esos módulos se tiene un repertorio de posibilidades de representación entre las que hay que buscar aquella que *mejor se ajuste* al S concreto objeto de modelado.

Un modelo se dice, en la terminología habitual del modelista, que se *construye*, en el sentido de que se edifica; es decir, que se ensamblan módulos básicos para dar lugar al objeto artificial que es el modelo. Este aspecto de construcción, a lo largo del proceso de modelado, no debe ser subvalorado y, en consecuencia, cabe considerar a ese proceso como artesanal, en el sentido de arte de organizar los elementos básicos que suministra la técnica de modelado empleada de forma adecuada para conseguir el objetivo propuesto: una imagen aceptable, para un determinado propósito, de un cierto aspecto del mundo real.

1.3. SIMULACIÓN DEL COMPORTAMIENTO

El computador u ordenador es un instrumento de considerable universalidad que puede convertirse, adecuadamente programado, en una aparente réplica —en lo que al comportamiento se refiere— de objetos o procesos del mundo de los que se disponga de una descripción convenientemente formalizada. Se dice entonces que *simula* al objeto con cuya descripción se ha programado. También se dice que se tiene un *simulador* del objeto o proceso considerado.

En este libro nos vamos a ocupar de la simulación del comportamiento de sistemas. Ya hemos visto que el comportamiento de un sistema se representaba mediante las trayectorias de las diferentes magnitudes que se le asocian. Nuestro objetivo será construir un modelo matemático, mediante un sistema dinámico, que, programado en un computador, permita generar las trayectorias que ilustren su comportamiento.

Los modelos \mathcal{M} que aquí consideramos (modelos de simulación informática del comportamiento) no son una representación pasiva de S (al modo de un plano o mapa), sino que son de tal naturaleza que su programación en un computador permite que éste genere un comportamiento que es una réplica, empleando cierta convención, del de S . Mediante la conjunción de técnicas informáticas y matemáticas, es posible pasar de la descripción de un cierto sistema, en un lenguaje adecuado, a la generación del comportamiento correspondiente.

Con lo visto hasta aquí ya estamos en condiciones de vislumbrar cómo se pueden ligar estructura y comportamiento. Por una parte, la estructura de un sistema viene dada por la trama de relaciones de influencia, de modo que se puede escribir,

$$\text{estructura} \iff \text{trama de relaciones}$$

Por otra, su comportamiento se manifiesta mediante las trayectorias, por lo que se puede decir que:

$$\text{comportamiento} \iff \text{trayectorias}$$

Por tanto, si a partir de la trama de relaciones se pudiesen obtener las trayectorias, se tendrían relacionadas estructura y comportamiento. Veremos cómo mediante la dinámica de sistemas somos capaces de llevar a cabo esta relación.

1.4. ELEMENTOS BÁSICOS DEL MÉTODO SISTÉMICO

En las secciones anteriores se han presentado, de forma somera, los elementos básicos sobre los que trata la dinámica de sistemas, a la que cabe considerar como un método concreto para el estudio de los sistemas que forman nuestro entorno. Este método se puede considerar como una manifestación particular de lo que se ha venido a denominar *metodología de sistemas*, o conjunto de métodos mediante los cuales abordar los problemas en los que la presencia de sistemas es dominante.

La metodología sistémica forma parte de lo que podemos denominar de forma genérica *movimiento sistémico*, que incluye todas las aportaciones, de naturaleza muy variada —desde

filosóficas a metodológicas— relacionadas con el estudio de los objetos dotados de una cierta complejidad a los que hemos venido en denominar sistemas. Como ya hemos visto, para los sistemistas, un sistema es un objeto complejo —natural o artificial— susceptible de ser analizado —dividido— en partes, pero cuya entidad resulta precisamente de cómo esas partes se integran en la unidad sustantiva que es el propio sistema.

Entre el sistema —el todo— y sus partes, se establece una relación dialéctica que fue precursoramente apuntada por Pascal cuando dijo que era imposible entender el todo —el sistema, diríamos aquí— sin conocer las partes, así como el entender las partes sin conocer el todo. El movimiento sistémico trata precisamente de desarrollar útiles conceptuales y operativos específicos con los que llevar a cabo el programa así esbozado de estudio de sistemas complejos. Los resultados alcanzados por este movimiento pretenden articularse en torno a una teoría de sistemas o sistémica.

Esta teoría dista mucho de presentarse como algo unitario y claramente estructurado. Por el contrario, en la actualidad es más bien un punto de confluencia de estudiosos y especialistas de diferente procedencia, cada uno de los cuales comparece en el foro de la sistémica con su bagaje metodológico y los sesgos propios de la disciplina de la que procede. A todos les une una llamada común: sus objetos de estudio tienen la característica genérica de ser sistemas; de ser objetos complejos formados por múltiples partes en interacción, dotados de alguna forma de organización, y de tal naturaleza que existe la presunción de que su comportamiento (el del sistema, el del conjunto) se debe más a la forma de producirse las interacciones entre las partes que a las propiedades de las partes tomadas aisladamente.

La ciencia clásica ha sido fundamentalmente reduccionista, en el sentido de asumir como axioma metodológico básico el que para estudiar un objeto lo que había que hacer era reducirlo a sus partes y estudiarlas aisladamente. Una vez conocidas las propiedades de las partes, las propiedades del sistema se desprenden por sí solas. Se trataba, por tanto, de reducir el estudio de un sistema a su análisis, a su disección. Este principio analítico ha sido enormemente fecundo, en particular en física. Sin embargo, desde las disciplinas que se ocupan del estudio de objetos de una cierta complejidad pronto se empezó a presentir su insuficiencia. Tan importante como la disección y el análisis es la integración y la síntesis.

La teoría de sistemas pretende no subordinar en el sistema el todo a las partes, pero tampoco las partes al todo. El sistemista cuando analiza un sistema, lo disecciona en sus partes, pero sin perder de vista la unidad del sistema; las partes sólo tienen sentido en tanto que partes del sistema, y no como objetos aislados desvinculados de él. Por el contrario, cuando estudia el sistema como una unidad, no olvida las partes. Esta forma peculiar de integrar el análisis y la síntesis ha sido asumida lúcidamente por Simon al decir: “frente a la complejidad [de un sistema] un reduccionista convencido puede ser al mismo tiempo un holista pragmático”³. Análisis y síntesis se dan en una peculiar conjunción que constituye la esencia de la metodología sistémica. Sin embargo, estas ideas, por sugestivas que parezcan, pueden pecar de excesivamente abstractas y vagas. ¿Cómo articular el análisis y la síntesis en un método que permita estudiar sistemas concretos? Una respuesta a esta pregunta la aportan las técnicas de modelado y simulación informática y muy en especial la conocida como dinámica de sistemas, a la que se dedica este libro.

Veremos cómo la dinámica de sistemas es un método en el que se combinan el análisis y

³ H. Simon, 1981, *The sciences of the artificial*, MIT, Cambridge, Massachusetts.

la síntesis, suministrando un ejemplo concreto de una metodología sistémica. La dinámica de sistemas suministra un lenguaje que permite expresar las relaciones que se producen en el seno de un sistema, y explicar cómo se genera su comportamiento. De este modo nos va a permitir poner de manifiesto cómo están relacionadas la estructura de un sistema (la trama de relaciones que articulan sus partes, recuérdese la figura 1.1) y su comportamiento (los cambios que se producen en sus atributos, representados por sus trayectorias, según se vio en la figura 1.6). Su objetivo es conciliar estructura y comportamiento, de modo que aparezcan como las dos caras de una misma moneda.

La construcción de un modelo para simulación informática requiere el análisis del sistema que se va a modelar, para decidir, en primer lugar, qué partes son relevantes para tener una descripción significativa. Una vez decididas las partes relevantes hay que establecer los mecanismos que las interrelacionan; es decir, hay que postular la forma cómo se integran para dar lugar al sistema. Mediante el computador es posible generar el comportamiento de esas partes (resultado del análisis) articuladas mediante el mecanismo de integración (consecuencia de la síntesis). Ese comportamiento deberá ajustarse al observado en la realidad. De este modo se tienen metodológicamente articulados análisis, síntesis y contrastación empírica.

Conviene resaltar cómo los modernos estudios sobre sistemas se han visto enormemente enriquecidos con la aparición del computador. Esta máquina es, en sí misma, un sistema complejo que puede, mediante su adecuada programación, convertirse en una réplica o *modelo* de una amplia clase de sistemas concretos. De este modo se dispone de un instrumento mediante el que desarrollar algo parecido a un trabajo experimental con las modelaciones de los sistemas. La introducción del computador lleva consigo la necesaria formalización de los objetos que se están estudiando y eso, para la teoría de sistemas, ha contribuido a su desarrollo formal. Por todo ello, el computador se ha convertido en un instrumento esencial para esta disciplina.

1.5. CAMPOS DE APLICACIÓN DE LA DINÁMICA DE SISTEMAS

La dinámica de sistemas fue concebida para resolver un problema concreto: el que presentaba una empresa de productos electrónicos que teniendo pocos clientes, y todos ellos con unos pedidos muy estables y previsibles, sin embargo registraba considerables oscilaciones en la línea de producción. La situación era desconcertante. El entorno en el que actuaba la empresa era muy estable y sin embargo en su seno se producían oscilaciones. Del análisis del problema se llegó a la conclusión de que estas oscilaciones eran generadas endógenamente por la empresa. Más aún, eran debidas a la combinación de estructuras de realimentación y de retrasos en la transmisión de información a lo largo de estas estructuras. Para llevar a cabo este análisis Jay W. Forrester sentó las bases del método que hoy conocemos como dinámica de sistemas.

Originalmente el método se denominó dinámica industrial, por el tipo de aplicación que había suscitado su desarrollo. Los trabajos pioneros se desarrollaron a finales de los años 50, y durante el decenio posterior se produce su implantación en los medios profesionales. A mediados de los 60 se comienza a vislumbrar aplicaciones del método más allá del ámbito industrial. En particular se promueve la aplicación de esas técnicas al estudio de áreas urbanas,

dando lugar a lo que se denominó dinámica urbana⁴. Una aplicación análoga la constituye la dinámica regional. Con estas aplicaciones se desarrollan modelos que aportan una herramienta auxiliar para la planificación urbana y regional. Estos modelos representan las interacciones que se producen entre las principales magnitudes socioeconómicas del área correspondiente y permiten generar, a partir de ellas, las evoluciones de las magnitudes consideradas significativas: número de habitantes, indicadores económicos, etcétera, para, a partir de estas evoluciones, planificar las necesidades de infraestructura y de servicios.

Con todo ello se van consolidando las posibilidades del método para un análisis de las relaciones entre estructura y comportamiento en sistemas complejos. Por ello no es extraño que a finales de los 60 el Club de Roma considerase su idoneidad para llevar a cabo su estudio sobre los límites del crecimiento. Precisamente un modelo de dinámica de sistemas sirvió de base al primer informe al Club de Roma en el que se analizaba la previsible evolución de una serie de magnitudes agregadas a nivel mundial como son la población, los recursos y la contaminación. En el modelo se analizaba la interacción entre estas magnitudes y se ponía de manifiesto cómo en un sistema de esa complejidad, debido a las fuertes interacciones que se producen en su seno, la actuación sobre unos elementos, prescindiendo de los otros, conduce a resultados imprevistos, cuando no insatisfactorios. El informe correspondiente tuvo una gran incidencia en la opinión pública y ha sido objeto de múltiples debates, tanto favorables como en contra. Es indudable que todo ello contribuyó enormemente a la difusión del método empleado en su elaboración. Recientemente se ha publicado una reelaboración de sus conclusiones, a la vista de los veinte años transcurridos desde el primer informe, en la que prácticamente se mantienen las recomendaciones de aquel informe pionero⁵.

Con este estudio sobre la dinámica mundial se puso de manifiesto que el método empleado posee una cierta universalidad que va más allá de los campos concretos a los que se aplica. Por tanto, las denominaciones dinámica industrial o dinámica urbana resultaban insuficientes y se convino en empezar a denominar el método como dinámica de sistemas⁶.

Sus campos de aplicación han resultado ser muy variados durante sus más de treinta años de existencia. Además de los sistemas industriales y socioeconómicos de ámbito urbano o regional, se ha empleado en sistemas sociológicos en donde se han modelado desde aspectos teóricos como la dinámica social de Pareto y/o de Marx⁷ hasta cuestiones de implantación de la justicia⁸. Otra área en que ha servido para importantes aplicaciones es la de los sistemas ecológicos y medioambientales, en donde se ha aplicado tanto a problemas de dinámica de poblaciones⁹, como de difusión de la contaminación¹⁰. Otro campo interesante de aplicaciones es el que suministra los recursos energéticos, en donde se ha empleado para definir estrategias

⁴ J. W. Forrester, 1969. *Urban dynamics*, Productivity Press.

⁵ D. H. Meadows, D. L. Meadows, J. Randers, y W. W. Behrens III, 1972, *Los límites del crecimiento: informe al Club de Roma sobre el predicamento de la humanidad*, Fondo de Cultura Económica.

⁶ Esa denominación ha sido cuestionada, ya que aparecía ser excesivamente pretenciosa cuando, en realidad, se ocupa de una clase limitada de lenguajes de modelado. Recientemente se han sugerido denominaciones alternativas, como *simulación dinámica* o *simulación del comportamiento*. Estas denominaciones, sin embargo, han conseguido aún menos aceptación. Aquí hemos conservado la denominación original.

⁷ R. A. Haneman, 1988, *Computer-assisted theory building*, Sage.

⁸ C. Jacobsen y R. Bronson, 1985, *Simulating violators*, ORSA.

⁹ L. Gutiérrez y W. Fey, 1980, *Ecosystem succession*, MIT Press.

¹⁰ D. H. Meadows, D. L. Meadows, J. Randers y W. W. Behrens III, 1973, *Toward global equilibrium*, Wright-Allen Press.

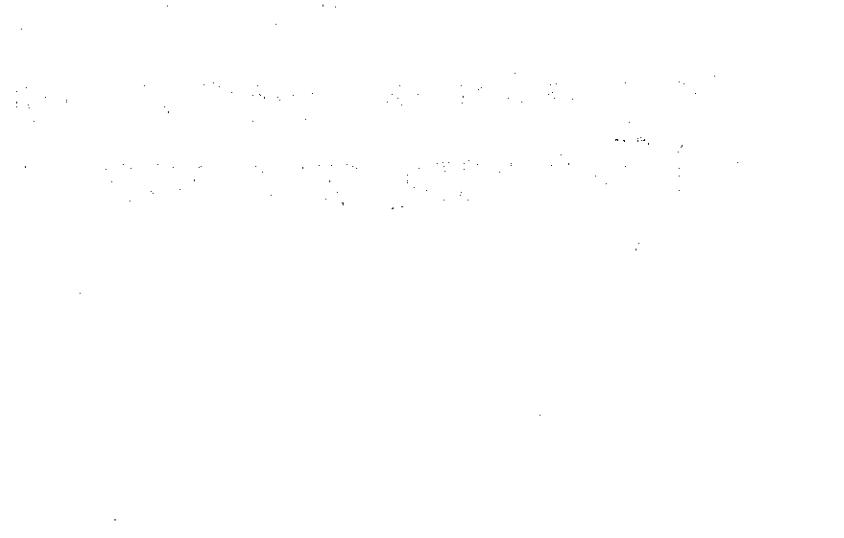
de utilización de esos recursos¹¹. Por terminar con esta sucinta enumeración de campos de aplicación, conviene también aludir que se ha empleado para problemas de defensa nacional¹².

La difusión de esta técnica de modelado y simulación informática ha sido muy amplia y en nuestros días se puede decir que constituye una de las herramientas sistémicas más sólidamente desarrolladas y que mayor grado de aceptación e implantación ha alcanzado.

Los capítulos de este libro constituyen una introducción a este método. En el capítulo 2 se exponen los elementos básicos del lenguaje sistémico, haciendo especial énfasis en cómo la estructura de un sistema determina su comportamiento. El método de la dinámica de sistemas se expone en el capítulo 3 y se ilustra con ejemplos en el capítulo 4. En el capítulo 5 se exponen diversas técnicas complementarias a la dinámica de sistemas que ayudan al análisis de los modelos. El capítulo 6 se dedica íntegramente a la descripción de una de estas técnicas: el análisis cualitativo de sistemas. La problemática asociada a los sistemas complejos se trata en el capítulo 7. Por último, el capítulo 8 se dedica a las técnicas de explotación de los modelos.

¹¹ N. Choucri, 1982, *International energy futures*, MIT Press. R. F. Naill, 1977, *Managing the energy transition*, Ballinger.

¹² E. Wolstenholme, 1990, *Systems enquiry*, Wiley, capítulos 8 y 10.



2.1. ATRIBUTOS Y SISTEMAS

En el capítulo anterior hemos visto cómo a un sistema se puede asociar una descripción mínima, que se reduce a un par (C, R) , y cómo este par se puede representar mediante un grafo como el de la figura 1.1, donde los nodos o vértices del grafo representan los elementos del sistema, y cuyos arcos representan las relaciones de influencia entre ellos. El grafo incorpora una *información cualitativa* muy elemental con respecto al sistema; sin embargo, suministra una visión global de la organización del sistema de indudable interés. Además, de esa información, como veremos, se pueden obtener indicaciones genéricas respecto a los modos de comportamiento. Vamos a dedicar este capítulo a analizar las posibilidades de describir un sistema mediante el lenguaje sistémico que aporta la descripción mínima (C, R) . Para ello conviene que nos detengamos en analizar los supuestos básicos de este lenguaje.

Al tratar de describir un sistema S , un observador \mathcal{O} le asocia un conjunto de *atributos* o características $\{X_i\}$. De una manera general se puede decir que un atributo X_i representa una cualidad perceptible de S , que da lugar a una unidad conceptual de representación. Los atributos son los perceptos mediante los cuales *vemos*, entramos en relación con un determinado objeto S . Por tanto, al mundo de S nos asomamos por medio de los atributos $\{X_i\}$, que se asocian a las distintas partes de S . Así, en el ejemplo del proceso de llenar un vaso de agua, el nivel que alcanza el agua en el vaso es un atributo del sistema.

En la definición de sistema está implícita la asunción de que entre las partes se produce alguna forma de relación. Las partes de un sistema no presentan un comportamiento autónomo e independiente unas de otras, sino que los vínculos que las articulan en la unidad del sistema se manifiestan en que los comportamientos de cada una de ellas presentan formas de dependencia o correlación con los de las otras. Por ello las trayectorias de los distintos atributos X_i no serán independientes entre sí, sino que, de alguna manera, estarán implícitas relaciones de dependencia entre ellas.

El buen especialista en el sistema S posee formas de conocimiento (que a veces es posi-

26 Dinámica de sistemas

Este capítulo introduce el análisis dinámico de sistemas, que es la descripción de cómo los sistemas cambian con el tiempo. Se exploran las interacciones entre las partes del sistema y cómo se manifiestan los cambios.

ble que no estén completamente estructuradas ni sean explícitas) sobre cómo se manifiestan los vínculos entre las partes del sistema; respecto a cómo variaciones de determinado atributo X_k afectan a las manifestaciones de otro X_j ; es decir, sobre cómo se producen las interacciones en el seno del sistema. Estamos suponiendo, de momento, un conocimiento —relativamente pobre— que establece únicamente la existencia de alguna forma de influencia entre atributos. Se emplea a veces el término *cuantitativo* para referirse a la naturaleza de este conocimiento. Se puede representar formalmente mediante enunciados de la forma:

$$X_k \rightarrow X_j \quad (2.1)$$

que se leen ‘el atributo X_k influye sobre el atributo X_j ’ y que reciben la denominación de *relaciones de influencia* entre atributos. Debe notarse que el concepto de relación de influencia es un concepto primitivo con vistas al establecimiento de la estructura de un sistema.

El concepto de influencia tiene connotaciones con el de causalidad. De hecho, la relación de influencia, tal como se acaba de enunciar, se corresponde con enunciados de la forma ‘ X_k causa X_j ’ que son muy comunes en descripciones informales. La noción de causalidad, en este sentido laxo, desempeña un papel importante en la descripción de los fenómenos del mundo real. Existen, sin embargo, importantes diferencias entre este empleo laxo del concepto de causalidad y el sentido fuerte con el que se emplea en la ciencia, y que ha sido cuestionado por muchos filósofos. Este sentido fuerte se refiere a la *causa total*, mientras que la causalidad que se recoge en las anteriores relaciones de influencia es una relación parcial, realizada en conjunción con otras ‘causas’. Por todo ello aquí se va a evitar hablar de causalidad, en la relación entre atributos, y se ha preferido emplear el término más débil de influencia.

La relación de influencia aporta un elemento esencial para la vertebración del sistema. A las diferentes partes de un sistema se asocian los correspondientes atributos (no se admiten partes sin atributos, porque en tal caso no pueden entrar en la descripción). Supongamos que a la parte A se asocia el atributo X_1 y a la B el X_2 . Si $X_1 \rightarrow X_2$ entonces diremos que ‘ A actúa sobre B ’. De este modo resulta que la articulación entre las partes y la influencia entre atributos resultan ser sinónimos. Por ello, el concepto de influencia es básico en la formulación conceptual de un sistema.

Entre el simple conocimiento de la relación de influencia entre dos atributos, como establece la expresión (2.1), y la forma matemática precisa de esta influencia, existen formas de conocimiento intermedias. Supóngase que los valores que toman los atributos pueden ordenarse, de modo que tenga significado el decir que el valor tomado por un atributo se incrementa o se decrementa. En tal caso, una relación de influencia puede tener un signo positivo o negativo. Así, se dice que X_k influye positivamente sobre X_j si sucede que un incremento en X_k produce un incremento en X_j , y un decremento en X_k produce un decremento en X_j . Se escribe entonces:

$$X_k \xrightarrow{+} X_j \quad (2.2)$$

Análogamente se asigna un signo negativo a la relación si a un incremento de X_k se sigue un decremento en X_j , y a un decremento de X_k un incremento en X_j . En este caso se escribe:

$$X_k \xrightarrow{-} X_j \quad (2.3)$$

Obsérvese que para asociar un incremento o un decremento a los atributos no se requiere que sean magnitudes; basta con que se produzca una relación ordinal entre sus manifestaciones. Al asociar un signo a la relación de influencia se tiene una información más rica sobre cómo se produce la interacción entre las partes del sistema, pero que continúa conservando el carácter cualitativo.

La relación de influencia puede ser multívoca en sus antecedentes

$$X_1, X_2, \dots, X_n \longrightarrow X_k \quad (2.4)$$

en cuyo caso se lee ‘ X_1, X_2, \dots y X_n influyen sobre X_k ’.

El conocimiento implícito en una relación de influencia puede ser de naturaleza muy variada, desde leyes científicas (con un nivel de aceptación prácticamente universal) hasta opiniones de expertos (de menor valor epistemológico), pero que, en todo caso, constituye la información *disponible* con relación a la clase de sistemas a la que pertenece \mathcal{S} . En todo caso, la relación de influencia constituye el conocimiento de naturaleza estructural más simple que posee el especialista en \mathcal{S} .

Entre los elementos que constituyen el sistema se establece un bosquejo esquemático de aquellos que están relacionados entre sí, lo cual se hace por medio de un diagrama en el cual los nombres de los distintos elementos están unidos entre sí por flechas. El diagrama que así se obtiene recibe las denominaciones de *diagrama de influencias* o *causal*. Aquí se empleará la primera de ellas.

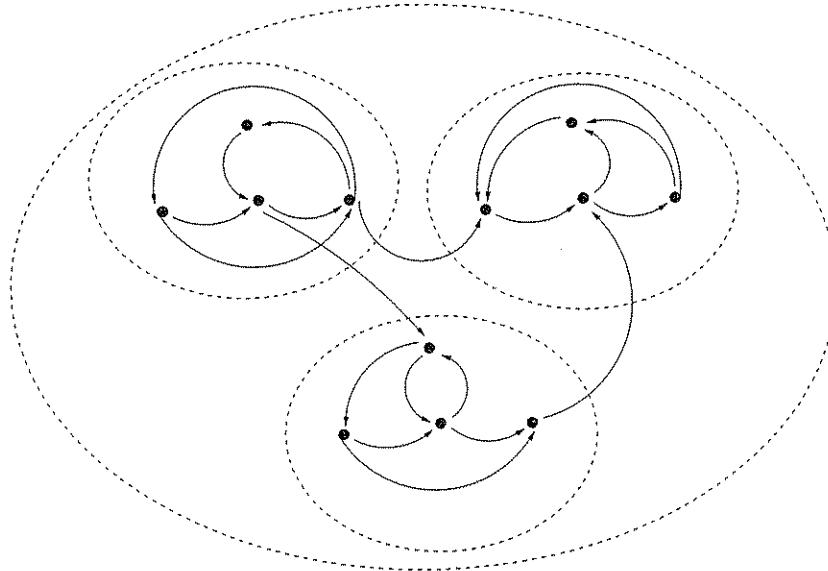


FIGURA 2.1. Estructura de un sistema complejo.

Mediante este diagrama, o lo que es lo mismo un conjunto de enunciados de la forma (2.3), se tiene una descripción del sistema \mathcal{S} , que puede servir de base para construir un modelo. En este conjunto de enunciados debe haber uno para cada elemento de \mathcal{S} , el cual aparecerá en su

consecuente. A este conjunto se puede asociar un grafo, como el de la figura 2.1, que suministra un esbozo de cómo se produce la vertebración del sistema, y permite tener una perspectiva global respecto a cómo se articulan las partes en la unidad del sistema: esto se conoce como *estructura* del sistema, como ya hemos visto en el capítulo anterior.

La estructura de un sistema recoge lo que podemos llamar su *forma sistémica*; es decir, la forma en que se organizan las partes para dar lugar al sistema, de acuerdo con la descripción adoptada (es la manera más simple y esquemática de formular la relación R). La forma sistémica es responsable de aquellas propiedades del sistema asociadas a sus manifestaciones como tal sistema y no relativas a las partes consideradas aisladamente.

Como ya hemos indicado en el capítulo anterior, mediante un modelo se pretende, entre otras cosas, explicar *endógenamente* la generación de un comportamiento. Vamos a ver cómo esta generación endógena está asociada a la existencia de *bucles de realimentación* en la estructura del modelo. Estos bucles son cadenas de influencias circulares cerradas, cuya presencia permite explicar determinadas formas del comportamiento, que son específicas de la propia estructura, e independientes de las solicitudes exteriores a las que se ve sometido el sistema.

Con ayuda de estos bucles podemos sentar las bases, aunque sea de forma muy laxa, de un lenguaje sistémico. Este lenguaje nos va a permitir no sólo realizar determinadas descripciones de los sistemas, sino que nos va a suministrar los elementos para realizar una primera organización de la percepción que tenemos de esos objetos. En cierto sentido, nos va a ayudar a *ver cosas que sin él no veríamos*. De este modo disponemos de un primer instrumento para iniciarnos en el estudio de las propiedades sistémicas.

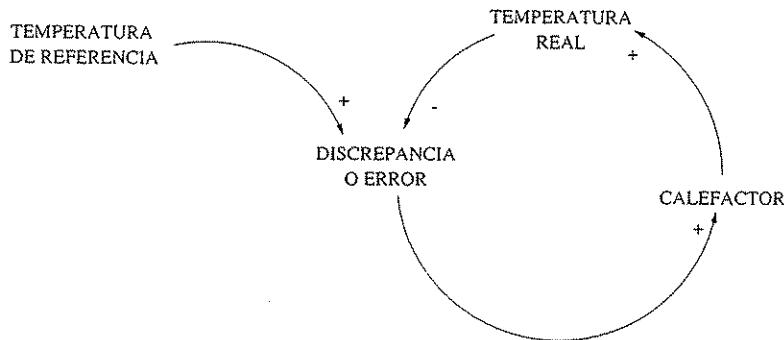


FIGURA 2.2. *Estructura de realimentación de un regulador de temperatura.*

Los elementos básicos de ese lenguaje son los bucles de realimentación positiva y negativa. Veremos cómo, a partir de ellos, podemos hacer algunas caracterizaciones elementales de los modos de comportamiento de los sistemas. Sin embargo, estas caracterizaciones pueden pecar de excesivamente esquemáticas, y resultar insuficientes. Para completarlas necesitamos un lenguaje más rico, del que nos ocuparemos en capítulos posteriores. Pero ese lenguaje es una reelaboración del que vamos a ver en este capítulo que, en cualquier caso, suministra las bases elementales de un lenguaje sistémico.

2.2. ESTRUCTURAS DE REALIMENTACIÓN SIMPLE

2.2.1. Bucle de realimentación negativa

En la figura 2.2 se tiene un ejemplo de un bucle de realimentación negativa al que ya se aludió en la sección 1.1.2 al analizar el proceso de llenado de un vaso. Ahora se trata de un sistema de regulación de temperatura.

El funcionamiento de un sistema de este tipo es extremadamente simple: si la temperatura se separa del valor deseado, aparece un error o discrepancia entre los valores de temperatura real y deseado, que activa un calefactor (o un refrigerador, en su caso), el cual, mediante la inyección de calorías (o frigorías) lleva la temperatura al valor deseado. Se trata, por tanto, de un comportamiento en el que el sistema corrige, de forma autónoma, las perturbaciones que tienden a separarlo del comportamiento deseado: el mantenimiento de la temperatura en un cierto valor. Este modo de comportamiento se conoce como comportamiento *autorregulador*. El sistema por sí solo se regula (él mismo *arregla* lo que la perturbación ha alterado). En la figura 2.3a se tiene la estructura general de un sistema de esta naturaleza.

Veamos con más detalle el comportamiento de un sistema con esta estructura. Si suponemos una variación en alguno de los elementos de la figura 2.3a, por ejemplo un incremento de *A*, este incremento determinará, de acuerdo con el signo de la influencia, un decremento de *C*, que a su vez determinará un decremento de *B*. Este último decremento de *B* producirá un decremento de *A*. Es decir, mediante la cadena causal circular, el incremento inicial de *A* se ha contrarrestado. Cualquier modificación (incremento o decremento) en cualquiera de los elementos vuelve a él, a lo largo de la cadena, con una acción de signo contrario. Se comprende así el carácter autorregulador del sistema que posee esta estructura. Por ello es la que incorporan a sus proyectos los ingenieros de control en el diseño de los sistemas automáticos. Las

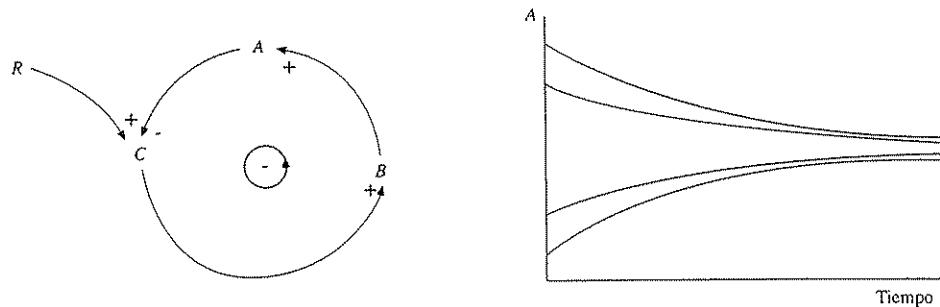


FIGURA 2.3. Grafo genérico de la estructura de realimentación negativa (a) y comportamiento correspondiente (b).

magnitudes tienden a mantenerse constantes, atenuándose los efectos de las perturbaciones (obsérvese que no se está afirmando que necesariamente los efectos se van a anular, sino que se atenúan). Suponiendo que en el instante inicial se produce la perturbación que separa a la magnitud regulada del valor deseado, las trayectorias correspondientes tienen la forma que se

muestra en la figura 2.3b. Los bucles de realimentación negativa se denominan también *bucles reguladores* o *estabilizadores*.

Se comprende también el porqué de la denominación de *realimentación negativa*: por una parte, se tiene una realimentación, ya que la información “circula” por el sistema (se realimenta, en el sentido de que se vuelve a alimentar) debido a la cadena circular de influencias; y la realimentación es negativa por la existencia de una influencia negativa (o de un número impar de ellas). Sin esta influencia negativa no se tendría el efecto corrector o compensador.

Es claro que el modo de comportamiento autorregulador depende de la estructura del sistema; es decir, de la forma de organizarse los distintos elementos que lo forman en una cadena de influencias circular. Cualquiera que sea la naturaleza de los componentes, siempre que se tenga una estructura de realimentación negativa, se tendrá un comportamiento autorregulador. De hecho, así sucede en múltiples ámbitos de la realidad, y se tienen procesos autorregulados tanto en sistemas artificiales (los ya mencionados sistemas de regulación o los servomecanismos) como en sistemas naturales (como, por ejemplo, los procesos homeostáticos en los seres vivos por los cuales mantienen sus constantes en los valores que permiten su supervivencia, con independencia de los cambios que se produzcan en el entorno). El hecho de que el comportamiento de un sistema con realimentación negativa dependa de la estructura y no de la naturaleza de los elementos componentes pone de manifiesto que se trata de un ejemplo paradigmático de una propiedad sistémica, como ya se indicó en el capítulo anterior.

Este tipo de bucles subyace en los comportamientos orientados a un objetivo como son los ejemplos antes mencionados. En muchas ocasiones los objetivos se encuentran implícitos, por lo que no es fácil detectar el bucle de realimentación negativa correspondiente.

Los bucles de realimentación negativa, aunque simples desde un punto de vista conceptual, pueden generar sorpresas y comportamientos problemáticos si no se detectan. La resistencia al cambio, tratando de mantener un objetivo implícito, es la principal manifestación del comportamiento de un sistema con un bucle de esta naturaleza. Hasta que no se haga explícita esa estructura los intentos de cambiar el comportamiento del sistema están condenados al fracaso. En general, siempre que se detecte una resistencia al cambio se debe suponer que existe “escondido” un proceso (o varios) de realimentación negativa subyacente.

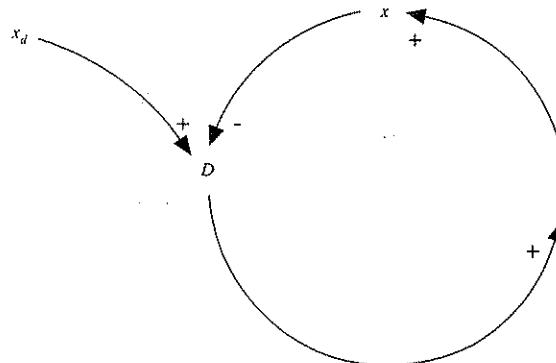


FIGURA 2.4. Elementos básicos de un bucle de realimentación negativa elemental.

2.2.2. Formulación matemática de un bucle elemental de realimentación negativa

Vamos a ver ahora una formulación matemática elemental de un sistema dotado de un bucle de realimentación negativa. Con ello dispondremos de una primera muestra de cómo asociar un sistema dinámico (en el sentido preciso que tiene en matemáticas esta locución) a un sistema dotado de realimentación negativa y descrito mediante un grafo. En la figura 2.4 se tiene un bucle de realimentación negativa elemental. Los elementos básicos de este bucle son:

- el estado del sistema x ,
- la acción o flujo F ,
- la discrepancia D ,
- el objetivo, meta o estado deseado x_d .

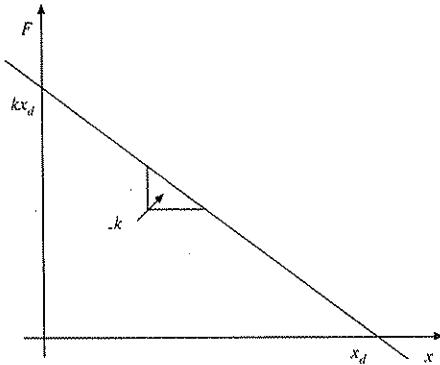


FIGURA 2.5. Relación entre el flujo F y el estado x .

Con el fin de introducir una hipótesis dinámica en el sistema, se supone que el estado representa la acumulación de acciones pasadas.

Se asume, además, que la relación entre el estado x y la acción F viene dada por una ecuación de la forma

$$\frac{dx}{dt} = F \quad (2.5)$$

lo que es consistente con la hipótesis de que el estado representa la acumulación de acciones pasadas:

$$x = \int_0^t F dt$$

Los otros elementos del bucle vienen dados por las ecuaciones:

$$\begin{aligned} F &= kD \\ D &= x_d - x \end{aligned}$$

De estas dos expresiones se desprende que:

$$F = k(x_d - x) \quad (2.6)$$

32 Dinámica de sistemas

La relación entre el flujo o acción F y el estado x se representa gráficamente como se hace en la figura 2.5. Integrando la expresión (2.5), teniendo en cuenta (2.6), se tiene

$$x(t) = x_d + [x(0) - x_d]e^{-kt} \quad (2.7)$$

que se representa gráficamente en la figura 2.6. El único equilibrio del sistema es la meta x_d .

Conviene observar que este comportamiento cualitativo, en el cual el estado x tiende al objetivo x_d , lo presentan todos los sistemas en los que la relación entre el flujo F y el estado x es una función monótona decreciente, de forma arbitraria, tal como se indica en la figura 2.7, y no una función lineal como la de la figura 2.5. Se dice entonces que todos los sistemas con un bucle de realimentación negativa elemental presentan un comportamiento que tiende a un objetivo, como el que muestra la figura 2.6.

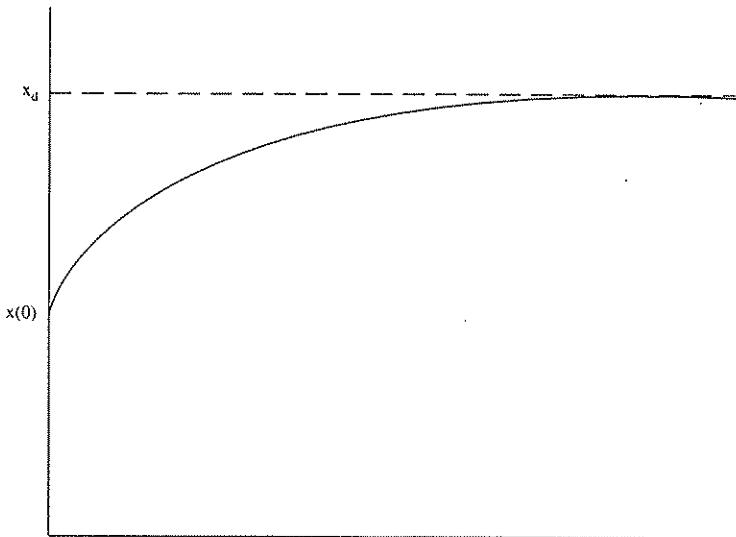


FIGURA 2.6. *Trayectoria de un sistema elemental de realimentación negativa.*

Se entiende que el comportamiento al que se está aludiendo es el que *cualitativamente* aparece en la figura 2.6, y no el que *cuantitativamente* muestra esta figura, que es el dado por la expresión (2.7). En el estudio del comportamiento de sistemas en muchas ocasiones interesa más el aspecto cualitativo de ese comportamiento que el estrictamente cuantitativo.

El comportamiento asintótico de la figura 2.6 lo presentan exclusivamente los sistemas en los que en el bucle aparece una única variable de estado; es decir, para los que la ecuación diferencial es de primer orden. Para sistemas de orden superior, como veremos más adelante, el comportamiento puede ser oscilatorio aunque sigue manteniéndose la pauta de tender asintóticamente a un objetivo.

En el capítulo siguiente generalizaremos el método empleado en esta subsección a un sistema cuya estructura sea más compleja que la aquí considerada. Veremos entonces cómo los elementos básicos que nos han servido para construir el sistema dinámico asociado a un grafo de influencias se pueden generalizar a sistemas más complejos.

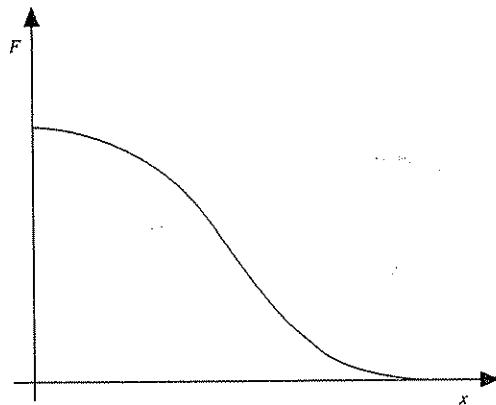


FIGURA 2.7. Función monótona decreciente que liga el flujo F y el estado x .

2.2.3. Bucles de realimentación positiva

Además del bucle de realimentación negativa, existe el de realimentación positiva. En la figura 2.8a se tiene un bucle de esta naturaleza. En él la perturbación de cualquier elemento tiende a reforzarse a lo largo de la cadena, por lo que, por ejemplo, un incremento de A determina a su vez su propio reforzamiento. De este modo se tiene un comportamiento autorreforzado o catalítico. Por ello estos bucles se conocen también como *bucles reforzadores*. En la figura 2.8b se tiene el comportamiento de un sistema cuya estructura es la de la figura 2.8a. El comportamiento que resulta de un bucle de esta naturaleza consiste en acelerar o bien el crecimiento, o el declive. El ejemplo más simple de un sistema que posea esta estructura es el de una población que crece sin ninguna limitación. Cuanto mayor sea el número de individuos, mayor será su descendencia, que incrementará a su vez el número de individuos, realimentándose el bucle sin cesar (figura 2.8b). Se tiene, entonces, un comportamiento explosivo. De nuevo tenemos un comportamiento que se explica mediante la estructura.

Otro ejemplo de bucle de realimentación positiva se tiene en la figura 2.9.

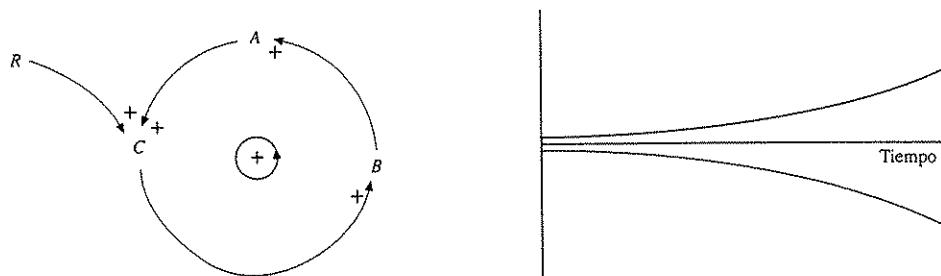


FIGURA 2.8. Grafo genérico de la estructura de realimentación positiva de un sistema (a) y comportamiento correspondiente (b).

34 Dinámica de sistemas

De acuerdo con este diagrama si un producto tiene la calidad adecuada, un incremento en las ventas implica un incremento de clientes satisfechos, los cuales contribuyen a la difusión oral del producto, lo que a su vez determina un incremento de las ventas. El bucle de realimentación positiva se asocia a procesos que son comúnmente conocidos como de “bola de nieve” o de “círculo vicioso o virtuoso”.

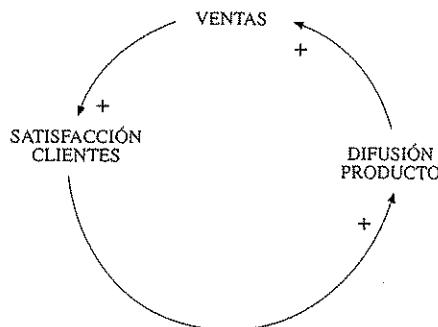


FIGURA 2.9. *Estructura de realimentación positiva de un sistema de incremento de las ventas de un producto.*

En general, los procesos de crecimiento se pueden explicar mediante bucles de realimentación positiva. Por ello, en los procesos de nuestro entorno en los que se manifiesta crecimiento (por ejemplo, en sistemas económicos y sociales), la presencia de estos bucles es dominante. Sin embargo, es fácil comprender que un bucle de realimentación positiva tiene un carácter fuertemente inestabilizador, al contrario de lo que sucede con los de realimentación negativa, que son, como hemos visto, estabilizadores.

En realidad, los procesos de crecimiento (o de declive) acelerado no se producen en la naturaleza, o en los sistemas sociales, hasta sus últimas consecuencias, porque todo proceso de crecimiento tarde o temprano encuentra unos límites. Para representar la aparición de estos límites se suele introducir un bucle de realimentación negativa adicional, como veremos en la sección 2.3.1.

Conviene observar que en los bucles de realimentación positiva y negativa se tiene exclusivamente información de la que más arriba hemos dicho que era de tipo cualitativo; es decir, el simple establecimiento de relaciones de influencia, con un signo determinado. Acabamos de ver cómo a partir de esa información se puede inferir el modo de comportamiento del sistema, al menos a grandes rasgos; por tanto, en estos casos es posible establecer el modo de comportamiento de un sistema a partir exclusivamente de información cualitativa.

2.2.4. Formulación matemática de un bucle elemental de realimentación positiva

Análogamente, como en el caso del bucle de realimentación negativa, es posible tener una formulación matemática del bucle de realimentación positiva en su caso más elemental. Los elementos básicos de esa formulación son:

- el estado x , y
- la acción (o flujo) F

y se organizan como se indica en la figura 2.10.

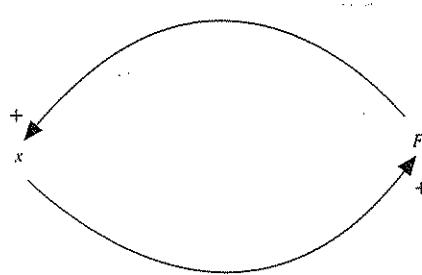


FIGURA 2.10. Elementos básicos de un bucle de realimentación positiva elemental.

Procedemos de forma análoga a como se hizo en el caso de un bucle con realimentación negativa, y adoptaremos la hipótesis de una relación entre el estado y la acción de la forma

$$\frac{dx}{dt} = F \quad (2.8)$$

es decir,

$$x = \int_0^t F dt$$

Si se asume que la acción es proporcional al estado

$$F = kx \quad (2.9)$$

se tiene que la ecuación que gobierna la evolución del estado del sistema viene dada por

$$\frac{dx}{dt} = kx \quad (2.10)$$

cuya integración conduce a

$$x(t) = x(0)e^{kt} \quad (2.11)$$

Conviene observar que el único equilibrio del sistema es el origen, que en este caso es inestable.

36 *Dinámica de sistemas*

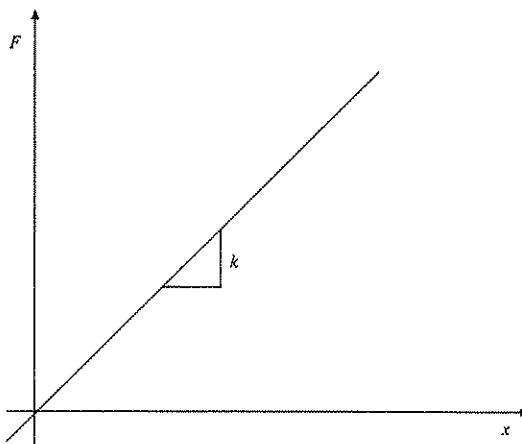


FIGURA 2.11. Relación que liga el flujo F con el estado x .

De acuerdo con (2.9) la relación que liga el flujo F con el estado x se puede representar gráficamente como se hace en la figura 2.11.

El comportamiento del sistema viene dado por la figura 2.12; es decir, que crece exponencialmente con el tiempo. Se pueden hacer aquí unos comentarios análogos a los que se hicieron en el caso del bucle de realimentación negativa: el comportamiento cualitativo que se muestra en la figura 2.12 lo presentan también aquellos sistemas no lineales en los que la relación entre el flujo y el estado es sencillamente una función monótona creciente, del tipo de la que se muestra en la figura 2.13.

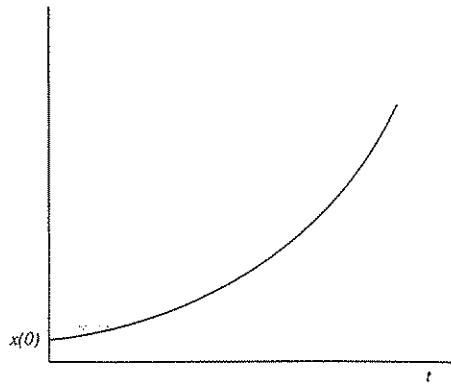


FIGURA 2.12. Crecimiento exponencial de un bucle de realimentación positiva.

Detengámonos a analizar el modo de comportamiento que estamos considerando. En la expresión (2.9) la constante k representa la *tasa de crecimiento* de x . Significa el tanto por uno de crecimiento de x , por unidad de x . Se expresa también en tanto por ciento. Así, una

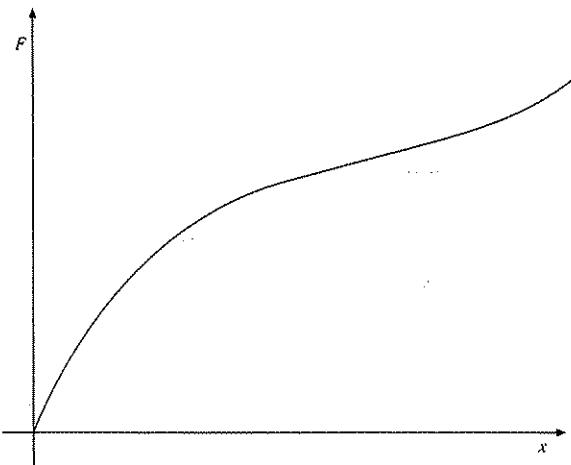


FIGURA 2.13. Relación entre el flujo y el estado función monótona creciente.

tasa de crecimiento del 5% representa que la magnitud x crece $0,05x$ en la unidad de tiempo. Por tanto, en las unidades en que se mide k debe indicarse explícitamente el tiempo al que se refiere el crecimiento, aunque normalmente está implícito y es en la unidad en que se mide t . Así, si t se mide en años, y se habla de una tasa de crecimiento del 5%, se entiende que x crece un 5% en un año. La expresión (2.10) se puede leer diciendo que la variación dx de x en la unidad de tiempo dt es k veces el valor alcanzado por x .

El comportamiento representado en la figura 2.12 recibe la denominación de *crecimiento exponencial*. Una característica interesante de este crecimiento es el denominado *tiempo de duplicación* t_d , que es el tiempo que tarda en duplicarse $x(t)$; es decir,

$$x(t_d) = 2x(0) \quad (2.12)$$

Conviene observar que el tiempo de duplicación es independiente del tiempo t inicial que se considere. En efecto,

$$x(t + t_d) = 2x(t)$$

ya que como $x(t) = x(0)e^{kt}$ y $x(t + t_d) = x(0)e^{kt}e^{kt_d}$, el tiempo de duplicación resulta independiente del tiempo inicial. De ahí su interés como parámetro para caracterizar el comportamiento exponencial. En la figura 2.14 se ilustra este proceso.

De las expresiones (2.11) y (2.12) es inmediato que

$$t_d = \frac{\ln 2}{k} \quad (2.13)$$

es decir

$$t_d \approx \frac{0.7}{k} = \frac{70}{k'} \quad (2.14)$$

38 Dinámica de sistemas

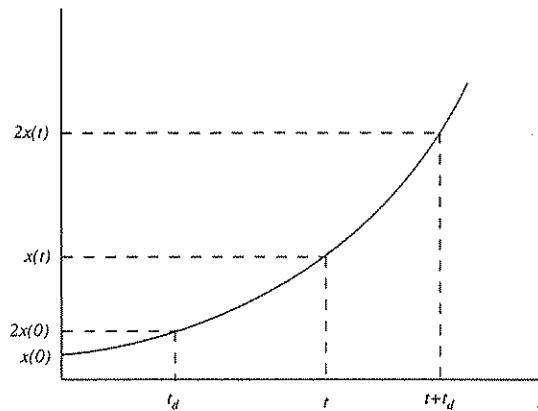


FIGURA 2.14. Tiempo de duplicación de un proceso con un crecimiento exponencial.

en donde k está expresada en tanto por uno y k' en tanto por ciento. Esta expresión tiene un indudable interés práctico, ya que nos dice que el tiempo de duplicación de un proceso con crecimiento exponencial es igual a 70 dividido por la tasa crecimiento (en tanto por ciento).

2.2.5. Retrasos

En todos los sistemas se producen retrasos entre las acciones y sus consecuencias. De hecho en todos los procesos de realimentación se produce alguna forma de retraso.

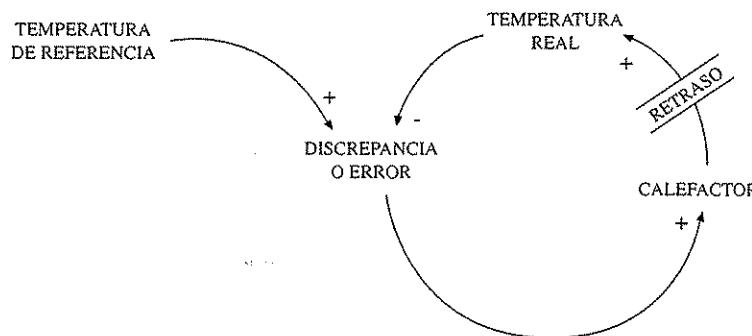


FIGURA 2.15. Bucle estabilizador con un retraso: regulación de la temperatura con retraso en la información.

Los retrasos, especialmente cuando son importantes, pueden producir inestabilidad en sistemas con realimentación negativa. Por ejemplo, si en el proceso de regulación de temperatura

que se ha visto anteriormente se introduce un retraso, como se hace en la figura 2.15, es claro que se producirán oscilaciones en el comportamiento como las que muestra la figura 2.16. Este fenómeno es bien conocido por la persona que regula la temperatura del agua preparándola para una ducha, cuando el calefactor está alejado. Obsérvese que cuanto más agresivo es el comportamiento del que actúa sobre la llave o grifo, mayores serán las oscilaciones que se producirán antes de alcanzar el equilibrio y, por tanto, mayor será el tiempo que se tardará en alcanzar la temperatura deseada. Ésta es una de las lecciones que se pueden extraer de los bucles de realimentación negativa con retrasos: una acción excesivamente agresiva a menudo produce un resultado que es justamente el opuesto del que se pretendía. Se producen oscilaciones e inestabilidad, en lugar de alcanzar de forma razonablemente rápida el objetivo que se pretende.

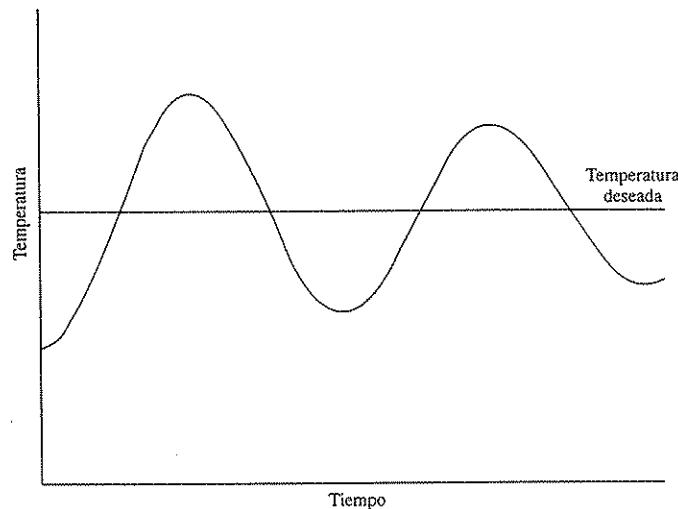


FIGURA 2.16. Evolución de la temperatura con retraso en la información.

En los problemas en los que se trata de estudiar tendencias a largo plazo, los retrasos y los bucles de realimentación resultan de gran importancia. A corto plazo, a veces incluso se pueden ignorar, pues tal vez no tengan consecuencias significativas.

2.3. ESTRUCTURAS COMPLEJAS DE REALIMENTACIÓN

Hasta aquí hemos considerado bucles de realimentación positiva o negativa aisladamente. En la práctica, sin embargo, es frecuente que la estructura de un sistema presente múltiples bucles de realimentación entrelazados, de modo que los bucles aparecen en combinación. Se tienen entonces comportamientos que son mezcla de los asociados a cada uno de estos bucles.

De estas combinaciones hay algunas particularmente interesantes, a las que se alude como *arquetipos sistémicos*.

Los arquetipos sistémicos son determinadas pautas estructurales que se presentan en múltiples situaciones. Se denominan también *estructuras genéricas*. Suministran "plantillas" que nos ayudan a establecer una primera estructura de determinadas situaciones. Sin embargo, al tratarse de estructuras complejas de realimentación, la relación entre estructura y comportamiento no es tan clara como en los casos elementales expuestos hasta ahora: aunque el estudio de los arquetipos clarifica ciertas situaciones, no permite definir totalmente el comportamiento de estos sistemas.

Los arquetipos sistémicos nos ayudan a reorganizar nuestra percepción, de modo que seamos capaces de ver las estructuras que se manifiestan en distintas situaciones problemáticas. En la actualidad se ha publicado una docena de estos arquetipos¹. En lo que sigue vamos a considerar dos de ellos, denominados la *dinámica del crecimiento sigmoidal* y la *dinámica de la adicción*, que, a su vez, pueden combinarse dando lugar a un tercer arquetipo: la *dinámica del crecimiento con inversión insuficiente*.

2.3.1. Arquetipo del crecimiento sigmoidal

Al considerar el proceso de crecimiento asociado a un bucle de realimentación positiva se indicó que en realidad todo proceso de crecimiento, más pronto o más tarde, se encuentra con unos límites. Ello es debido a que la espiral de crecimiento produce, aunque sea de forma no deseada, efectos secundarios que eventualmente conducen al agotamiento del proceso de crecimiento.

Consideremos como ejemplo el crecimiento de una población en un hábitat que es capaz de sustentarla, pero en el que los recursos son limitados. En la figura 2.17 se tiene la estructura elemental de este sistema. Se observan en ella dos bucles. Por una parte, un bucle de realimentación positiva que liga la población con su crecimiento vegetativo, y que en el diagrama de la figura 2.17 se muestra en la parte izquierda. Este bucle es el responsable del proceso de crecimiento.

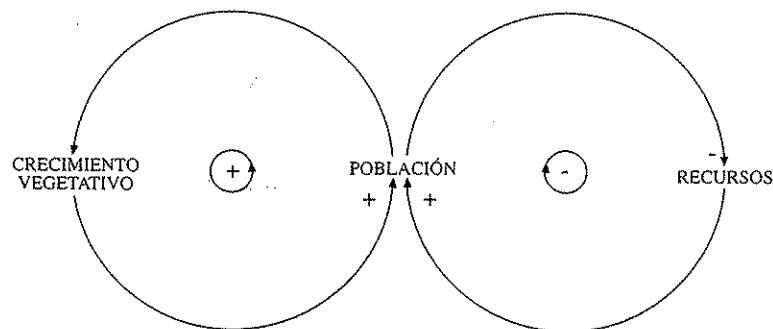


FIGURA 2.17. *Diagrama de un sistema de crecimiento de una población.*

¹ P. Senge, 1990, *The Fifth Discipline*, DoubleDay.

Por otra parte, las disponibilidades de recursos (la capacidad de sustentación del hábitat) limitan el crecimiento de la población mediante el bucle de realimentación negativa que se muestra en la figura 2.17. De acuerdo con este bucle, el crecimiento de la población hace disminuir los recursos *per capita* disponibles. La influencia de estos recursos sobre el crecimiento de la población es positiva, por lo que su disminución determinará que esa población disminuya su ritmo de crecimiento, tal como se indica en la figura 2.17, cerrándose así la segunda cadena circular de influencias. Este bucle de realimentación negativa actúa de acuerdo con una condición limitadora.

En la figura 2.18 aparece el diagrama genérico de un sistema que presenta crecimiento sigmoidal. El proceso de crecimiento sigmoidal se presenta al interactuar los dos bucles. En la fase inicial del proceso, cuando se desencadena el crecimiento de la población, la limitación de los recursos no es perceptible, por lo que el bucle de realimentación positiva es el dominante, y se produce un crecimiento casi exponencial de la población. Al crecer ésta, la limitación de los recursos empieza a manifestarse, de modo que la dominancia de los bucles va pasando del positivo al negativo. En la medida en que se agoten las posibilidades de crecimiento, el bucle de realimentación negativa resulta dominante hasta que, al final del proceso, se corta la posibilidad de crecimiento. Por tanto, el comportamiento de un sistema con la estructura de la figura 2.18 presenta la forma que se ilustra en la figura 2.19, y es una combinación de los comportamientos representados en las figuras 2.3 y 2.8. Resulta interesante observar que el crecimiento que resulta de la interacción de los bucles positivo y negativo es el crecimiento sigmoidal, por lo que se tiene una interpretación estructural de ese tipo de crecimiento.

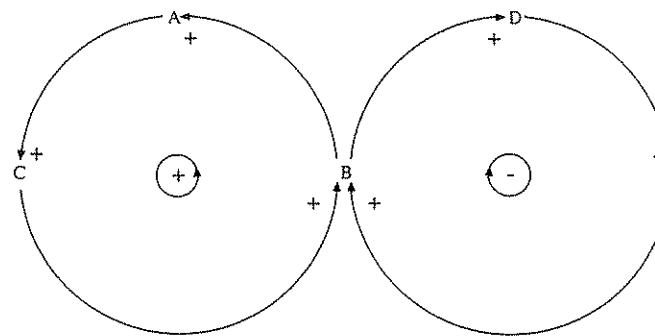
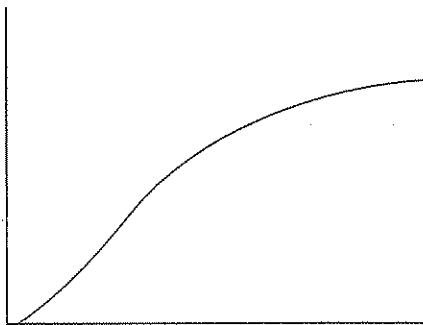


FIGURA 2.18. Diagrama genérico de un sistema que presenta crecimiento sigmoidal.

Existen múltiples ejemplos de procesos con crecimiento sigmoidal, como son el proceso de difusión de una innovación tecnológica, el de introducción de un nuevo producto en el mercado o el de difusión de un rumor en un determinado medio social. En todos ellos se tiene una fase inicial en la que se produce un crecimiento —prácticamente exponencial— de la innovación tecnológica, del nuevo producto o del rumor. Sin embargo, a medida que se alcanzan los límites de crecimiento, como consecuencia del carácter finito del medio en el que se desarrollan (el nuevo producto satura el mercado) se produce una limitación que aborta el crecimiento. La

FIGURA 2.19. *Crecimiento sigmoidal.*

pauta de comportamiento a lo largo del tiempo presenta en todos estos casos la misma forma que la de la figura 2.19, y a todos ellos podemos aplicar la plantilla estructural de la figura 2.18.

Esta interpretación estructural del crecimiento sigmoidal nos ayuda a entender cómo actuar para tratar de atenuar los efectos negativos que pueda tener la limitación al crecimiento. Intuitivamente existe una tendencia a tratar de continuar un proceso de crecimiento, cuando éste parece abortarse, tratando de fomentar el propio crecimiento; es decir, actuando sobre el bucle de realimentación positiva. Sin embargo, ese tipo de actuación se manifiesta ineficiente. El modo de actuación adecuado consiste en actuar sobre el bucle de realimentación negativa, y en particular se debe identificar y cambiar el factor limitador, que es el verdadero responsable de los límites al crecimiento, si es que ello es posible. En otro caso hay que resignarse a aceptar estos límites.

2.3.2. Formulación matemática de un crecimiento sigmoidal

En la figura 2.20 se muestra la estructura básica de un sistema en el que coexisten un bucle de realimentación positiva, a la izquierda del diagrama, con un bucle de realimentación negativa, a la derecha. Tanto un bucle como otro están formados por los elementos básicos discutidos en las secciones anteriores.

A este diagrama puede asociarse la ecuación

$$\frac{dx}{dt} = F_1 + F_2 \quad (2.15)$$

en donde los flujos F_1 y F_2 , en el caso más simple de una influencia lineal, vienen dados por

$$\begin{aligned} F_1 &= k_1 x \\ F_2 &= k_2 D = k_2 (x_d - x) \end{aligned}$$

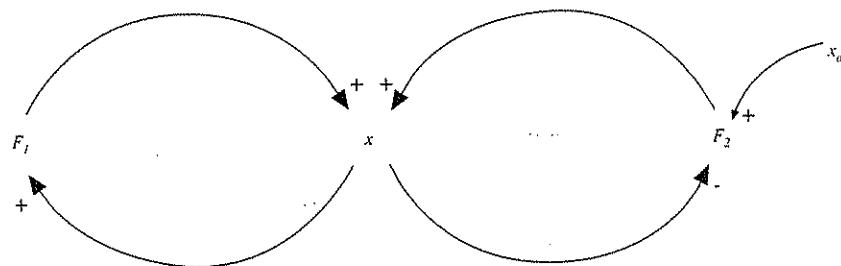


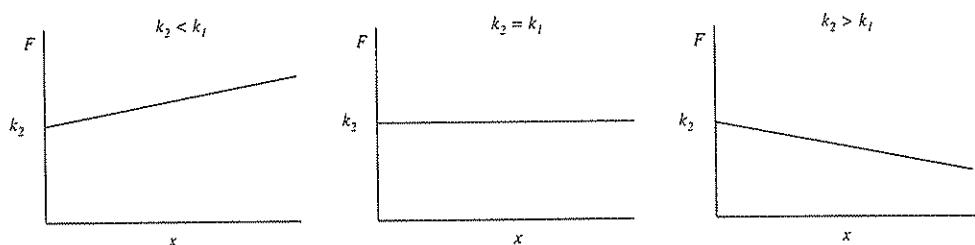
FIGURA 2.20. Estructura básica de un sistema en el que coexisten dos bucles de realimentación.

siendo F_1 el flujo asociado al bucle de realimentación positiva y F_2 el correspondiente a la negativa. La ecuación que rige la evolución del estado viene dada por

$$\frac{dx}{dt} = (k_1 - k_2)x + k_2x_d. \quad (2.16)$$

Según el valor relativo de k_1 y k_2 pueden darse los tres tipos de situaciones que se indican en la figura 2.21. En el primer caso, es decir, cuando $k_1 > k_2$, el bucle de realimentación positiva es el dominante, mientras que en el tercer caso, o sea, cuando $k_1 < k_2$, entonces el dominante es el negativo. El caso intermedio $k_1 = k_2$ es un caso de transición.

Los comportamientos asociados a los tres casos representados en la figura 2.21 se muestran en la figura 2.22. Por tanto, en ninguno de ellos se produce el crecimiento sigmoidal.

FIGURA 2.21. Forma gráfica de la relación entre F y x , según los valores relativos de k_1 y k_2 .

Un caso en el que el crecimiento sigmoidal se produce es aquel en el cual las relaciones entre los flujos F_1 y F_2 , aunque monótonas crecientes, se cortan entre sí tal como se muestra en la figura 2.23. En tal caso el comportamiento del sistema requiere una consideración especial. El flujo conjunto $F = F_1 + F_2$ será el que se muestra en la figura 2.24, en la que se observa que F es creciente con x para valores pequeños de esta variable, hasta alcanzar un valor máximo a partir del cual empieza a decrecer hasta anularse y alcanzar valores negativos.

Lo anterior se interpreta diciendo que para valores pequeños de x , antes de alcanzar el máximo de F de la figura 2.24, el bucle de realimentación positiva domina sobre el negativo;

44 Dinámica de sistemas

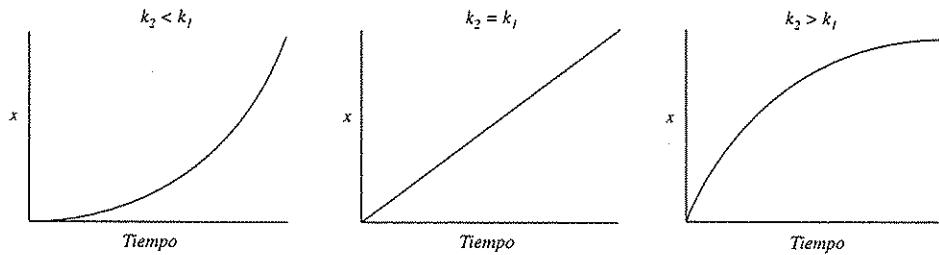


FIGURA 2.22. Comportamientos de x según los valores relativos de k_1 y k_2 .

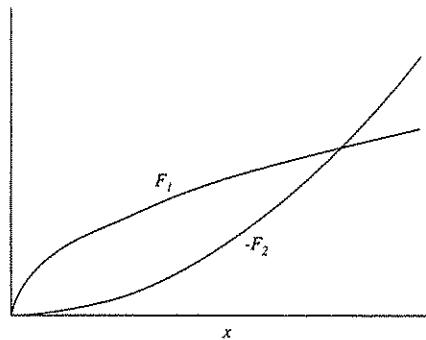


FIGURA 2.23. Formas no lineales de las relaciones entre F y x .

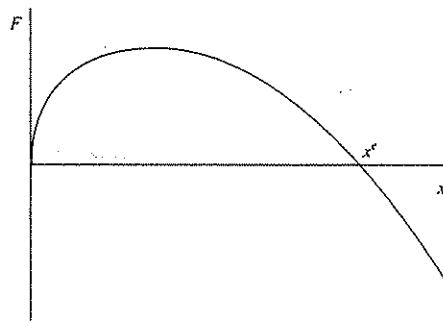


FIGURA 2.24. Relación conjunta entre F y x .

mientras que para valores de x superiores a ese máximo sucede lo contrario, el bucle de realimentación negativa domina sobre el positivo. El sistema tiende al punto x^e tal que $F(x^e) = 0$. En consecuencia el comportamiento resultante, tal como se indica en la figura 2.25, será el resultado de la combinación de estos dos comportamientos. En la fase inicial del proceso el bucle dominante será el positivo, mientras que en la final será el negativo. El comportamiento inicial será un crecimiento exponencial —cuando el bucle dominante sea el de realimentación positiva— para continuar con un comportamiento del tipo asintótico, cuando lo sea el negativo. La combinación de los dos comportamientos da lugar al sigmoidal.

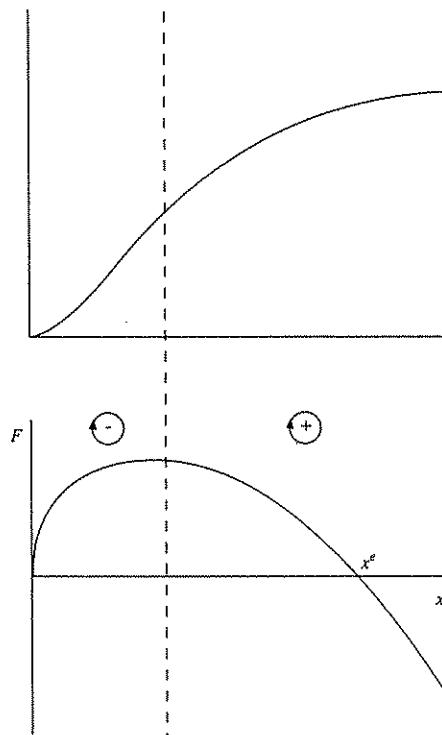


FIGURA 2.25. *Comportamiento sigmoidal con especificación de la dominancia relativa de los bucles.*

2.3.3. Arquetipo de la adicción

Un tipo de situaciones con el que frecuentemente nos encontramos es aquel en el que se manifiesta un síntoma problemático que requiere atención. Este síntoma está producido por un problema subyacente que resulta difícil de tratar, bien sea porque no se conoce bien, o porque resulte costoso el afrontarlo. En ese tipo de situaciones es frecuente que se traten los síntomas sin resolver el problema fundamental. Con esas actuaciones es posible que se obtengan éxitos

a corto plazo sobre el síntoma; pero se mantiene inalterado el problema fundamental. Es más, posiblemente se vaya atrofiando la propia capacidad de afrontarlo.

Un ejemplo de ese tipo de situaciones lo presenta el problema del estrés que produce el exceso de trabajo. Si nuestro trabajo se incrementa más allá de nuestra capacidad de acometerlo razonablemente y como consecuencia de ello caemos en una situación de estrés, la única solución es autolimitarnos con respecto a la aceptación de más trabajo. Tal vez no estemos dispuestos a ello, ya que podría afectar a nuestra promoción profesional. Por ello puede que tratemos de aliviar el estrés mediante actuaciones duras, como el alcohol o los fármacos, o mediante formas más suaves de compensación, como el ejercicio físico. Pero, en todo caso, estas actuaciones no resuelven el problema del exceso de trabajo; sólo lo enmascaran temporalmente aliviando el estrés. El problema sigue latente y periódicamente reaparece, con lo que se incrementa la dependencia y se llega, por último, a la adicción. A este proceso, que se conoce como *arquetipo de la adicción*, se asocia un diagrama de influencias como el de la figura 2.26.

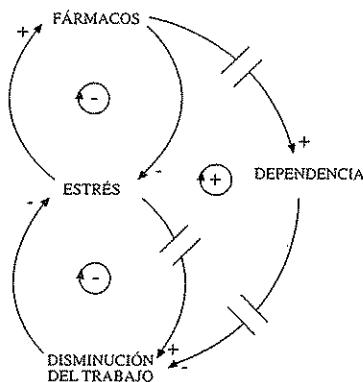


FIGURA 2.26. *Diagrama de influencias del problema del estrés.*

Esta figura es un caso particular de la figura 2.27 que muestra la estructura general subyacente en un proceso de este tipo. De acuerdo con esta estructura se tienen dos procesos de realimentación negativa. Los dos tratan de corregir el mismo síntoma problemático. El superior representa la intervención sintomática, mediante la que se persigue una solución rápida y superficial. Resuelve el problema con rapidez, pero sólo temporalmente. El bucle inferior tiene un retraso. Representa una actuación de carácter más básico sobre el problema, cuyos efectos tardan más tiempo en convertirse en evidentes. Con frecuencia, aunque no siempre, se presenta un proceso adicional de realimentación positiva generado por los efectos colaterales de la solución sintomática. Cuando esto sucede, esos efectos secundarios hacen aún más difícil llegar a la solución fundamental.

Este tipo de estructura permite explicar una amplia variedad de comportamientos en los que soluciones bien intencionadas a corto plazo empeoran las cosas a largo plazo. Al solucionar exclusivamente los síntomas se reduce la percepción de la necesidad de buscar soluciones de carácter fundamental. Con el tiempo, se confía más y más en la solución sintomática, que acaba por parecer la única solución posible. Con ello, sin embargo, se hacen crónicas situaciones de dependencia o adicción.

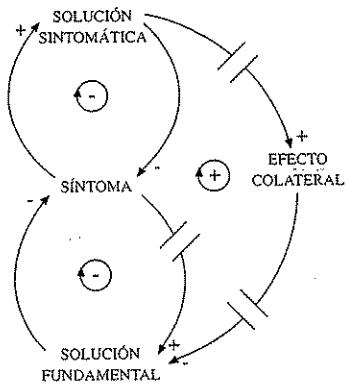


FIGURA 2.27. Diagrama de influencias del arquetipo de la adicción.

Situaciones como las que se acaban de describir surgen en problemas de otra naturaleza, como el representado en la figura 2.28, en la que se interpreta la dependencia energética como una forma de adicción al petróleo.

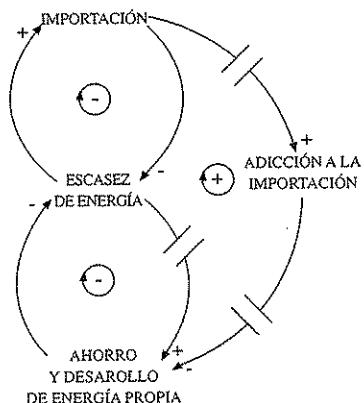


FIGURA 2.28. Diagrama de influencias del arquetipo de la dependencia energética.

También se interpretan con esa estructura situaciones en las que se desplaza la responsabilidad de resolver un problema a la intervención exterior a la propia organización (por ejemplo, el problema de la ayuda a países subdesarrollados —y de los subsidios, en general— si no conduce a un desarrollo de los propios recursos, acaba por crear situaciones de dependencia a largo plazo).

Un caso especial del arquetipo de la adicción se presenta cuando se produce un proceso de *erosión de metas*. Cuando existe una discrepancia entre nuestros objetivos, y el estado que hemos alcanzado con respecto a ellos, podemos tratar de anularla mediante dos tipos de actua-

ciones: o bien se mejora nuestra situación (con relación a esos objetivos) o bien se modifican, a la baja, los propios objetivos. En la figura 2.29 se tiene la estructura correspondiente.

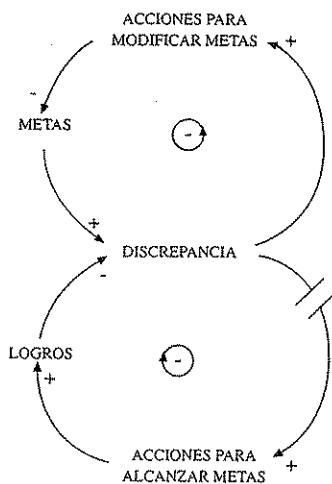


FIGURA 2.29. *Diagrama de influencias del arquetipo de la erosión de metas.*

Para tratar con eficacia los procesos del arquetipo de la adicción se requiere una combinación de actuaciones que fortalezcan la respuesta fundamental y debiliten la respuesta sintomática. Para fortalecer la respuesta fundamental se precisa una orientación a largo plazo en las actuaciones. En cuanto a la necesidad de soluciones sintomáticas, éstas son necesarias en muchas situaciones, pero siempre que se reconozcan como tales y se combinen con estrategias que permitan rehabilitar la capacidad para resolver el problema fundamental.

2.3.4. Arquetipo del crecimiento con inversión insuficiente

En esta sección se va a analizar un arquetipo que resulta de la combinación de los anteriores. Con este arquetipo se contempla un tipo de situación que se da con frecuencia y que se ilustrará con lo sucedido a una empresa que, en determinado momento, desarrolla un nuevo tipo de producto, de gran éxito inicial². Hay una gran demanda para su producción y no existen aparentemente problemas de financiación.

Durante sus tres primeros años las ventas se duplican anualmente. Sin embargo, la cartera de pedidos pendientes empieza a crecer de forma peligrosa a mediados del segundo año. La demanda crece tan rápidamente que el suministro de los pedidos comienza a retrasarse ligeramente. Inicialmente, se sirven las máquinas en aproximadamente ocho semanas, pero este periodo empieza a ampliarse considerablemente.

Se requiere aumentar la capacidad de producción. Sin embargo, la oportuna toma de decisiones se retrasa de modo que, cuando se toman medidas, el mercado ya ha acusado los efectos de los retrasos excesivos en el suministro de las máquinas, y las ventas han iniciado una cierta recesión (figura 2.30).

² J. W. Forrester, 1975, *Collected Papers*, Productivity Press, capítulo 7.

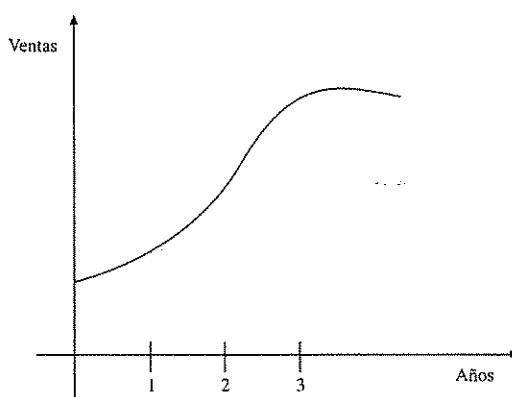


FIGURA 2.30. Evolución de las ventas en los tres primeros años.

Se realizan entonces intentos para que las ventas aumenten, mediante incentivos y campañas publicitarias, y así sucede en principio. Pero de nuevo la cartera de pedidos pendientes vuelve a incrementarse, los retrasos en el servicio crecen de nuevo, y se plantea nuevamente la necesidad de ampliación de la capacidad de producción.

En la figura 2.31 se muestra el comportamiento de las ventas de esa empresa, en el que se pone de manifiesto la intermitencia de períodos de crecimiento y crisis. Finalmente se produce la bancarrota y la empresa no puede adaptarse a un mercado potencial que, en condiciones adecuadas, le hubiese permitido un crecimiento sostenido y un éxito a largo plazo.

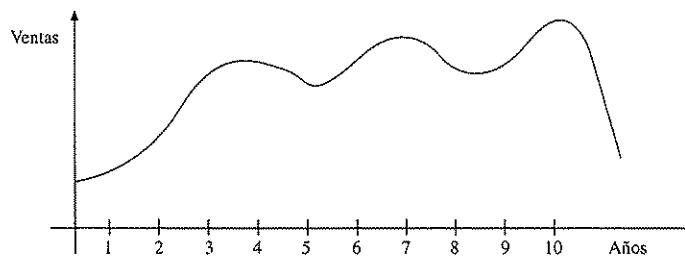


FIGURA 2.31. Evolución de las ventas durante los diez años de existencia de la empresa.

Veamos cómo se puede analizar lo sucedido a esta empresa con ayuda de los arquetipos sistemáticos. En primer lugar tenemos un proceso de crecimiento. Sabemos que un proceso de esta naturaleza se asocia a un bucle de realimentación positiva. En el caso que nos ocupa, la empresa cuidó especialmente su departamento de ventas, reinvertiendo parte de sus beneficios en incrementar la capacidad de ese departamento. En la figura 2.32 se muestra el diagrama de realimentación positiva correspondiente a ese proceso.

Hemos visto también que todo proceso de crecimiento lleva asociado el proceso que lo limita. En general, en el caso de la introducción de un nuevo producto en el mercado —como

50 Dinámica de sistemas

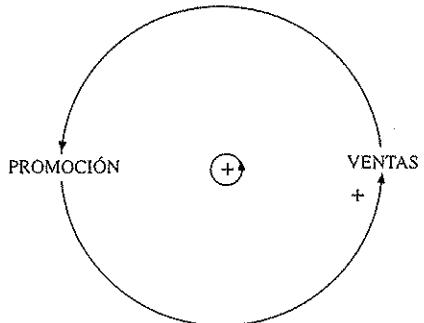


FIGURA 2.32. *Diagrama de realimentación positiva asociado al proceso de crecimiento de las ventas.*

es el caso que nos ocupa— la limitación al crecimiento suele venir dada porque el mercado alcanza sus propios límites. Sin embargo, en el caso de la empresa que estamos considerando, este crecimiento se vio cercenado —como acabamos de ver— por el incremento de los retrasos en el suministro de las máquinas. Este hecho puede ser interpretado mediante el diagrama de la figura 2.33.

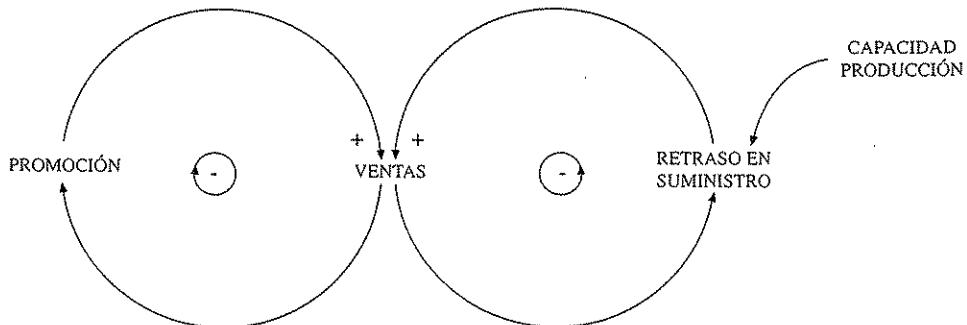


FIGURA 2.33. *Estructura de la limitación al crecimiento por el incremento de los retrasos en el suministro de las máquinas.*

Tenemos, por tanto, la combinación de un bucle de retroalimentación positiva con otro de retroalimentación negativa, en la estructura que hemos denominado de crecimiento sigmoidal. Al incrementarse la cartera de pedidos pendientes, con relación a la capacidad de producción, los tiempos de suministro se incrementan. Se produce entonces una reputación de mal servicio por parte de la empresa, por lo que las nuevas ventas se dificultan. Hemos visto, al analizar la estructura del crecimiento sigmoidal, que la peor manera de actuar era forzar el bucle de retroalimentación positiva. Pero, de hecho, eso es lo que hace la empresa al incentivar las ventas. Lo que debía hacerse, con suficiente prontitud, es disminuir el retraso en el tiempo de suministro a los clientes mediante el incremento de la capacidad de producción. La actuación sobre estas estructuras debe hacerse sobre el bucle de retroalimentación negativa.

¿Por qué no se detecta a tiempo el problema del retraso en los suministros? Posiblemente

porque los órganos de gestión de la empresa están más interesados en los problemas financieros (manejan fundamentalmente datos de esta naturaleza) que en un detalle aparentemente sin importancia como son los retrasos en el suministro de las máquinas a los clientes. Cuando se detecta el problema ya es difícil actuar. Son víctimas de la dificultad, presente en múltiples procesos del aprendizaje, de detectar la causa y el efecto cuando están considerablemente separados en el tiempo.

Además, una vez detectado el problema de que los retrasos en el suministro son debidos a la insuficiente capacidad de producción, esta última se aborda de forma extremadamente cautelar. Este hecho lo podemos interpretar mediante el arquetipo de la adicción. En efecto, si consideramos el retraso en el suministro como el síntoma problemático, y la ampliación de la capacidad de producción como la solución fundamental, podemos representar el proceso correspondiente mediante el diagrama de la figura 2.34. El bucle superior representa actua-

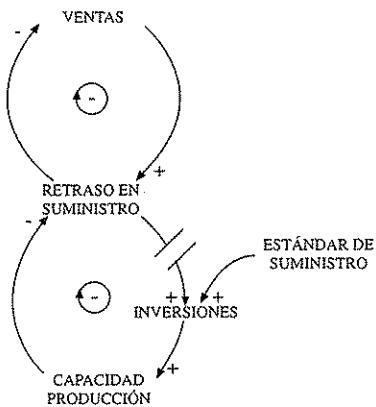


FIGURA 2.34. *Aplicación del arquetipo de la adicción al problema sintomático del retraso en el suministro de equipos.*

ciones directas sobre las ventas, mientras que el bucle inferior representa actuaciones sobre la capacidad productiva. De nuevo aquí la empresa adoptó el mal camino. En lugar de actuar sobre el bucle inferior (que es el modo de actuación adecuado a un arquetipo de la adicción, como hemos visto anteriormente) actuó fundamentalmente sobre el superior. Sobre el inferior actúa tímidamente y con retraso. Con los anteriores esquemas se ha construido un modelo de simulación que muestra que el comportamiento insatisfactorio de la empresa (representado en la figura 2.31) se hubiese podido cambiar, con una política adecuada, por el de la figura 2.35.

En la figura 2.36 se muestra el diagrama de influencias conjunto del proceso. Las dos estructuras que hemos visto en las figuras 2.33 y 2.34 se sobreponen en una sola al compartir el bucle de realimentación negativa de la izquierda de la figura 2.33, y superior de la figura 2.34. Este bucle estabilizador es el responsable de los límites al crecimiento, en el proceso de crecimiento sigmoidal y, al mismo tiempo, lo es de la imagen distorsionadora según la cual aumentando las ventas (aun a costa de erosionar los criterios de calidad, como son los tiempos de suministro), y no incrementando la capacidad de producción, se consigue un crecimiento sostenido de la empresa. Todo sucede como si la empresa hubiese acabado siendo víctima de

52 Dinámica de sistemas

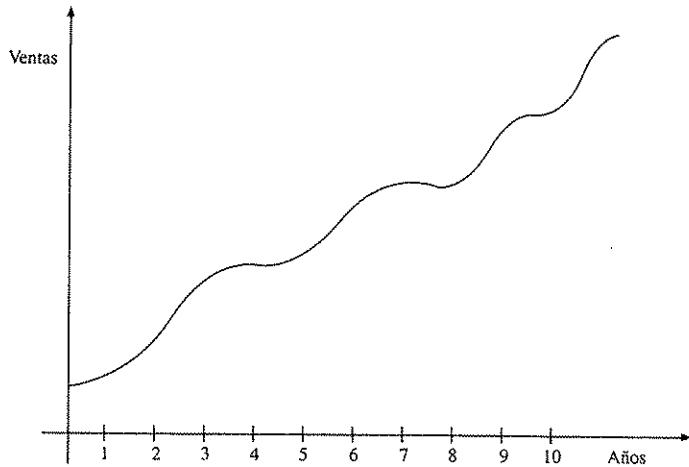


FIGURA 2.35. Comportamiento satisfactorio de la empresa mediante un modelo de simulación.

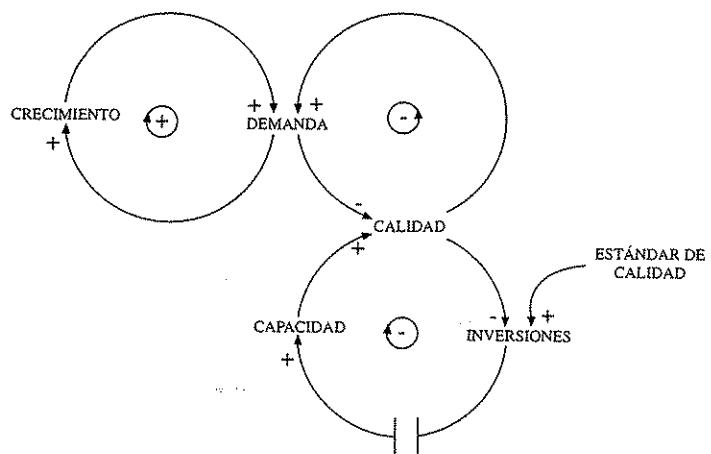


FIGURA 2.36. Diagrama de influencias conjunto del proceso de crecimiento con inversión insuficiente.

un proceso de erosión de metas, limitando su propio crecimiento, al adoptar unas pautas de inversión inadecuadas por insuficientes.

Para terminar con este capítulo conviene recordar que lo que aquí se ha pretendido es mostrar algunas estructuras elementales y los comportamientos a ellas asociados. Hemos visto que con estructuras de un solo bucle el comportamiento del sistema está bien determinado. Con dos bucles las cosas ya no están tan claras. Y en el último ejemplo, ya con tres bucles, la estructura tal como la hemos considerado hasta aquí es insuficiente para dar cuenta, de forma unívoca, del comportamiento, aunque nos ayuda a comprenderlo.

En la práctica, sin embargo, lo frecuente es encontrar sistemas cuya estructura sea mucho más compleja que las anteriores, y esté formada por múltiples bucles de realimentación entrelazados entre sí. El lenguaje elemental de descripción de sistemas mediante relaciones de influencia, al que hemos dedicado el capítulo, es insuficiente en estos casos. Se requieren entonces para determinar el comportamiento útiles más elaborados que la simple inspección de los grafos que representan la estructura. Disponemos, como ya hemos visto, al tratar las estructuras de realimentación positiva y negativa y la de crecimiento sigmoidal y —como generalizaremos a sistemas más complejos en los capítulos siguientes— de esquemas conceptuales más potentes que permiten una descripción más rica de la estructura de un sistema, con vistas a determinar su comportamiento. En concreto, disponemos del concepto de sistema dinámico, y de un método —la dinámica de sistemas— que nos va a permitir transcribir la relación de influencias en un sistema dinámico. Todo ello lo vamos a ver en el próximo capítulo.

3.1. INTRODUCCIÓN

En el capítulo anterior hemos visto distintos ejemplos de situaciones susceptibles de ser modeladas mediante un lenguaje sistemático elemental, formado esencialmente por diagramas de influencias, y en el que se ponía especial énfasis en los bucles de realimentación. Hemos visto también cómo la generación del comportamiento, a partir de los bucles que formaban la estructura, sólo era posible en casos especialmente sencillos, y aun en ellos sólo admitía una caracterización esencialmente cualitativa. Para generar las trayectorias se requiere que la relación entre las partes de un sistema posea un contenido matemático más rico que el simple establecimiento de las relaciones de influencia. El concepto de sistema dinámico aporta un lenguaje más elaborado que permite generar el comportamiento, como hemos visto ya en algunos casos particularmente sencillos dotados de uno o dos bucles de realimentación.

Cabe, por tanto, preguntarse a qué reelaboración hay que someter en general al diagrama de influencias para poder obtener un sistema dinámico; es decir, qué mayor riqueza se requiere en la descripción de un sistema para obtener una estructura que permita ser interpretada mediante un sistema dinámico.

Por otra parte, la información contenida en un diagrama de influencias se reduce a una serie de enunciados que establecen vínculos de dependencia entre los elementos básicos de un sistema. Estos enunciados se pueden formular mediante el lenguaje ordinario, de modo que a la descripción mínima de un sistema, mediante el enunciado de sus elementos y las relaciones de influencia entre ellos, se puede asociar una descripción del sistema en lenguaje ordinario.

Ello invita a pensar que sea posible concebir una transición de una descripción verbal de un sistema a un diagrama de influencias, y de éste a un sistema dinámico, si somos capaces de establecer las condiciones oportunas. Este libro está dedicado a un método, denominado *dinámica de sistemas*, que permite realizar esta transición, en la que partiendo de una descripción en lenguaje ordinario del sistema que se trata de modelar se llega a un sistema dinámico. Este último, además de los aspectos estructurales que comparte con la descripción verbal,

incorpora aspectos cuantitativos que no están explícitamente en esa descripción. De hecho, la dinámica de sistemas se encuentra en una encrucijada entre lo cuantitativo y lo cualitativo: emplea instrumentos cuantitativos, propios de las ciencias físicas, como son los sistemas dinámicos; pero involucra una forma de conocimiento en los modelos que se construyen con su concurso, en la que los rasgos cualitativos pueden ser dominantes.

El proceso de modelado mediante dinámica de sistemas tiene un carácter eminentemente constructivo en el que después de analizar el sistema, y partiendo del conocimiento disponible respecto a cómo se articulan sus partes —bien sea en forma de leyes ampliamente aceptadas o bien sencillamente a partir de opiniones de expertos expresadas en lenguaje ordinario—, mediante un proceso de sucesivas reelaboraciones, se llega a un modelo en forma de sistema dinámico que permite generar su comportamiento y del que, además, se extraen eventualmente conclusiones cualitativas.

La dinámica de sistemas, introducida, como ya se mencionó en el capítulo 1, por Jay W. Forrester, permite analizar la estructura del sistema y, a partir de este análisis, construir un sistema dinámico mediante el cual se pueda generar su comportamiento. Se parte del modelo mental que se tiene de una cierta situación problemática y se llega a construir un sistema dinámico que contribuya a explicarla.

El método de Forrester entraña directamente con la teoría de sistemas, en la medida que suministra un método para estudiar sistemas dotados de una cierta complejidad estructural, donde el bucle de realimentación es el bloque básico. Utiliza tanto instrumentos de matemática aplicada (los grafos y, sobre todo, los sistemas dinámicos) como útiles informáticos, entre los que se encuentran los entornos DYNAMO, STELLA, ITHINK, VENSIM y POWERSim.

Los modelos construidos mediante la dinámica de sistemas son sistemas dinámicos, por lo que todo el amplio y rico bagaje de conocimientos matemáticos que se tienen en la teoría de sistemas dinámicos pueden explotarse en el ámbito de la dinámica de sistemas. Esto es especialmente interesante si se considera que los modelos de dinámica de sistemas son normalmente sistemas fuertemente no lineales, por lo que pueden presentar formas de comportamiento muy complejas, para cuyo análisis los recientes resultados de la teoría matemática de sistemas dinámicos no lineales resultan de un gran interés, como veremos en un capítulo posterior.

3.2. DE LA ESTRUCTURA DEL COMPORTAMIENTO

En el capítulo anterior se ha presentado un lenguaje sistémico que aporta los elementos básicos para una descripción esquemática de un sistema. Hemos visto también en la subsección 2.2.2 cómo era posible asociar un sistema dinámico elemental a un sistema que poseyera la estructura de realimentación negativa, y generar, a partir de ese sistema dinámico, el comportamiento autorregulador de la figura 2.6. De este modo, a un sistema descrito mediante un grafo se asociaba otro descrito mediante un sistema dinámico. Lo mismo se hizo con los sistemas con realimentación negativa, en la subsección 2.2.4, y con los de crecimiento sigmoidal, en la 2.3.2.

Vamos a tratar de generalizar estos resultados, proponiendo un método que permita asociar un sistema dinámico a un sistema descrito mediante un grafo. Trataremos con ello de explicar la generación endógena del comportamiento de un sistema. Partiremos de la observación de que entre los distintos elementos que aparecen en los nodos de un diagrama de influencias, algunos representan variaciones con respecto al tiempo de otras magnitudes consideradas en

ese mismo diagrama. Por ejemplo, en el diagrama de la figura 2.10 la variable de flujo representa la variación con respecto al tiempo del valor del estado x . Recordando esa figura se tiene:

FLUJO → ESTADO

Esta influencia es un caso particular de otra más general que podemos expresar de la forma:

$$\frac{dX}{dt} \longrightarrow X \quad (3.1)$$

En la que dX/dt denota la variación con respecto al tiempo de la magnitud X . Esta expresión representa una relación trivial: la variación con respecto al tiempo de X influye en el crecimiento de la propia variable X . Sin embargo, lo que interesa por el momento resaltar es que la existencia —en el diagrama de influencias— de variables que representan la variación con respecto al tiempo de otras, comporta que estas últimas cambien a lo largo del tiempo. En este sencillo hecho se basa el que podemos afirmar que en la estructura está implícito el comportamiento del sistema.

Conviene también observar que siempre que existe una variable del tipo dX/dt , que representa la variación de una magnitud X con respecto al tiempo, se tendrá una relación de influencia como la de la expresión (3.1). La variable X resulta de la acumulación del cambio implícito en la variable dX/dt . Por tanto, siempre que aparezca una variable como la dX/dt aparecerá otra X , y entre ambas se establecerá una relación como la (3.1). La variable X se denomina *variable de estado* y la dX/dt *variable de flujo*. A las variables de estado se las conoce también, en dinámica de sistemas, como *variables de nivel* por razones que se verán en la sección siguiente.

Las anteriores consideraciones llevaron a Forrester a postular una clasificación de las distintas variables que aparecen en un diagrama de influencias en tres grupos: variables de estado, variables de flujo y variables auxiliares. Las variables de estado son normalmente las variables más importantes y representan las magnitudes cuya evolución es especialmente significativa. Asociadas a cada variable de estado se encuentran una o varias variables de flujo, que determinan su variación a lo largo del tiempo. Por último, las variables auxiliares constituyen las restantes variables que aparecen en el diagrama, y representan pasos intermedios para la determinación de las variables de flujo a partir de las variables de estado. Si se recuerdan las formulaciones matemáticas de los bucles de realimentación positiva y negativa, y del crecimiento sigmoidal, se verá cómo en todos estos casos los modelos matemáticos correspondientes admitían la anterior clasificación entre sus variables.

La distinción entre variables de estado y variables auxiliares, a partir del diagrama de influencias, no siempre está clara y a veces es difícil decidir si una variable debe ser un estado o una variable auxiliar. Ya se ha indicado que un estado representa un punto de acumulación. Una regla aceptable para decidir el carácter de una variable se basa en considerar cómo se comporta esta variable ante un cambio en el sistema. Las variables de estado varían lentamente acumulando los flujos. Las variables auxiliares varían instantáneamente en respuesta a los valores que toman las variables de estado a lo largo del sistema. También puede ayudar en esta distinción tener en consideración que, si se detiene el proceso que se pretende modelar, los flujos se anulan mientras que las variables de estado conservan su valor. Puede suceder que

una variable, representada por una variable auxiliar cuando se emplea un horizonte temporal muy grande, deba ser representada como un estado cuando el horizonte temporal sea menor.

3.3. DIAGRAMAS DE FORRESTER

Una vez clasificados los elementos que aparecen en el diagrama de influencias en variables de estado, flujo y auxiliares estamos en disposición de obtener, a partir del diagrama de influencias, lo que se conoce como el diagrama de Forrester, que es uno de los instrumentos básicos de la dinámica de sistemas.

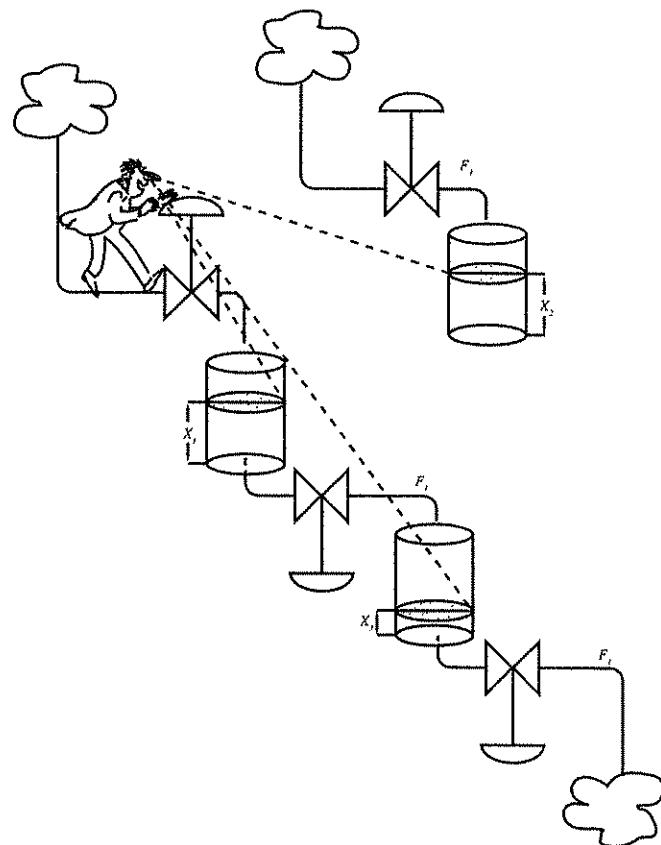


FIGURA 3.1. Símil hidrodinámico de un sistema de ecuaciones diferenciales de primer orden.

Para ayudar a comprender el significado de las tres clases de variables es conveniente recurrir a un símil hidrodinámico como el mostrado en la figura 3.1. En esta figura se representan tres depósitos en los que se acumulan tres niveles X_1 , X_2 y X_3 . Las variaciones de los niveles vienen determinadas por las actuaciones sobre unas ciertas válvulas que regulan los caudales que alimentan a cada uno de los depósitos. La decisión sobre la apertura de estas válvulas se toma teniendo como única información los valores alcanzados por los niveles, en cada uno

de los depósitos, en el instante de tiempo considerado. En la figura 3.1 esto se representa con ayuda de un observador que teniendo como única información el conocimiento de los niveles en el resto de los depósitos determina la apertura de la válvula correspondiente. Aunque en la figura sólo aparece el observador en una de las válvulas, debe considerarse que hay uno en cada una de ellas.

De acuerdo con lo anterior está claro que el valor tomado por la variable de flujo en cada instante depende exclusivamente de los valores alcanzados por los niveles en dicho instante; de forma análoga, los valores alcanzados por los niveles dependen de los valores tomados por las variables de flujo que alimentan a dichos niveles.

Con el símil hidrodinámico se obtiene una forma intuitiva —apropiada para una mentalidad que busca imágenes físicas— de representar un sistema de ecuaciones diferenciales de primer orden. En efecto, se ve de inmediato que haciendo ciertas simplificaciones, en la figura 3.1 no se hace sino representar —de forma analógica— un sistema de ecuaciones diferenciales tal como

$$\frac{d(X_1)}{dt} = -F_1 - F_2 \quad (3.2)$$

$$\frac{d(X_2)}{dt} = -F_4 \quad (3.3)$$

$$\frac{d(X_3)}{dt} = -F_2 - F_3 \quad (3.4)$$

siendo

$$F_i = f_i(X_1, X_2, X_3) \quad i = 1, 2, 3 \quad (3.5)$$

Estas funciones f_i pueden ser lineales o no lineales. La determinación del valor tomado por una variable de flujo, por ejemplo F_1 , a partir de los estados X_1 , X_2 y X_3 puede que sea conveniente hacerla en distintas etapas, requiriéndose para ello el establecimiento de unas variables auxiliares; por ejemplo, la función f_1 puede descomponerse en tres etapas, empleando dos variables auxiliares A_1 , y A_2 , teniendo,

$$A_1 = \varphi_1(X_1, X_2) \quad (3.6)$$

$$A_2 = \varphi_2(A_1, X_3) \quad (3.7)$$

$$F_1 = \varphi_3(A_2). \quad (3.8)$$

En efecto

$$F_1 = \varphi_3[\varphi_2(A_1, X_3)] \quad (3.9)$$

$$= \varphi_3[\varphi_2[\varphi_1(X_1, X_2), X_3]] \quad (3.10)$$

$$= f_1(X_1, X_2, X_3) \quad (3.11)$$

Es decir, las variables auxiliares representan etapas intermedias en la determinación de los flujos a partir de los estados y, en último extremo, pueden ser eliminadas.

El símil hidrodinámico de la figura 3.1 se puede completar con la inclusión de variables exógenas. Éstas suministran información adicional y exterior, que debe considerarse para decidir el valor que toman las variables de flujo F_i . Es decir, llamando E a una variable exógena, las expresiones (3.5) se convertirían en

$$F_i = f_i(X_1, X_2, X_3, E)$$

En el símil hidrodinámico se pone claramente de manifiesto que se pueden concebir dos tipos esenciales de variables, los estados y los flujos, y una clase secundaria, las variables auxiliares. Empleando esta analogía, en dinámica de sistemas, las variables que aparecen en un modelo se clasifican en variables de estado, variables de flujo y variables auxiliares. De esta manera se consigue dar una forma intuitiva al proceso de construir un modelo que, en último extremo, no va a ser sino un sistema de ecuaciones diferenciales de primer orden.

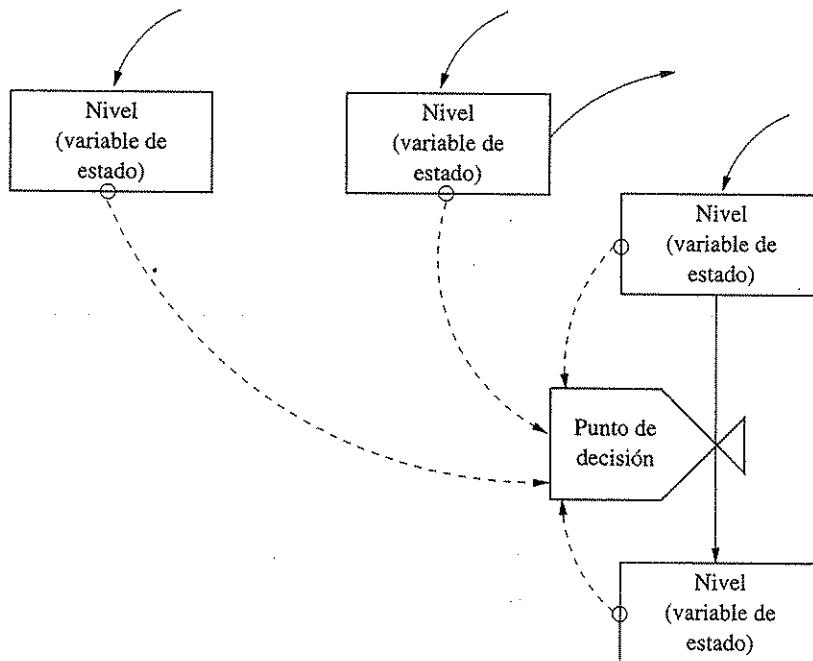


FIGURA 3.2. *Conexión entre las variables de nivel (de estado) y los puntos de decisión (variables de flujo).*

En la figura 3.2 se tiene un diagrama que muestra, de forma gráfica, las ideas que se acaban de exponer. Se emplean en el diagrama los símbolos que se muestran en la figura 3.3 y que se explicarán posteriormente. Es necesario indicar que estos símbolos son los que inicialmente propuso Forrester, pero que, en la actualidad, no son empleados de forma universal, y se

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO. FACULTAD DE INGENIERÍA. ESTACIÓN BIOMÉDICA. ESTACIÓN DE INGENIERÍA QUÍMICA. ESTACIÓN DE INGENIERÍA MECÁNICA. ESTACIÓN DE INGENIERÍA ELECTRÓNICA. ESTACIÓN DE INGENIERÍA MATERIALES. ESTACIÓN DE INGENIERÍA INDUSTRIAL. ESTACIÓN DE INGENIERÍA AERONÁUTICA. ESTACIÓN DE INGENIERÍA MECATRÓNICA. ESTACIÓN DE INGENIERÍA EN SISTEMAS Y COMPUTACIÓN. ESTACIÓN DE INGENIERÍA EN PETROQUÍMICA. ESTACIÓN DE INGENIERÍA EN PETRÓLEO. ESTACIÓN DE INGENIERÍA EN PETROQUÍMICA. ESTACIÓN DE INGENIERÍA EN PETRÓLEO.

dibujan estos diagramas con cierta libertad a la hora de escoger los símbolos. En cualquier caso, el significado de la figura 3.2 es claro a primera vista. Se observa en la misma cómo las variaciones de un estado son el resultado de una decisión tomada a partir de información que proviene del resto de los estados. En lo que sigue se estudiará, de forma sistemática y detenida, este proceso. Para ello se discuten en primer lugar, con detalle las variables de estado de flujo y auxiliares, así como las interconexiones que se establecen entre ellas.

3.3.1. Variables de estado

Las variables de estado o niveles constituyen aquel conjunto de variables cuya evolución es significativa para el estudio del sistema. Los estados representan magnitudes que acumulan los resultados de acciones tomadas en el pasado. Esta función de acumulación puede asimilarse a la del nivel alcanzado por un líquido en un depósito; de ahí proviene la denominación de nivel, siguiendo el símil hidrodinámico.

La elección de los elementos que se representan por variables de estado, en un modelo determinado, depende del problema específico que se esté considerando. En la elección de estas variables desempeña un papel primordial la experiencia del diseñador del modelo. Una característica común a todos los estados es que cambian lentamente en respuesta a las variaciones de otras variables.

En los diagramas de Forrester los niveles se representan por medio de rectángulos (figura 3.3).

	<i>Nube:</i>	representa una fuente o un pozo; puede interpretarse como un nivel que no tiene interés y es prácticamente inagotable.
	<i>Estado:</i>	representa una acumulación de un flujo.
	<i>Flujo:</i>	variación de un nivel; representa un cambio en el estado del sistema.
	<i>Canal de material:</i>	canal de transmisión de una magnitud física que se conserva.
	<i>Canal de información:</i>	canal de transmisión de una cierta información, que no es necesario que se conserve.
	<i>Variable auxiliar:</i>	una cantidad con un cierto significado físico en el mundo real y con un tiempo de respuesta instantáneo.
	<i>Constante:</i>	un elemento del modelo que no cambia de valor.
	<i>Retraso:</i>	un elemento que simula retrasos en la transmisión de información o de material.
	<i>Variable exógena:</i>	variable cuya evolución es independiente de las del resto del sistema. Representa una acción del medio sobre el sistema.

FIGURA 3.3. Símbolos utilizados originalmente en los diagramas de Forrester.

62 Dinámica de sistemas

A cada estado X se le puede asociar un flujo de entrada F_e y uno de salida F_s , de modo que la ecuación que representa su evolución es la siguiente,

$$X(t) = X(0) + \int_0^T (F_e - F_s) dt$$

o lo que es lo mismo

$$\frac{dX}{dt} = F_e - F_s$$

3.3.2. Variables de flujo

Las variables de flujo determinan las variaciones en los estados del sistema. Las variables de flujo caracterizan las acciones que se toman en el sistema, las cuales quedan acumuladas en los correspondientes estados; es decir, determinan cómo se convierte la información disponible en una acción o actuación.

Originalmente se representaban por medio de los símbolos que se indican en la figura 3.3, aunque hoy en día se utilizan iconos distintos. Estos símbolos están inspirados en el símil hidrodinámico, según el cual las variables de flujo se pueden asociar a válvulas que regulen los caudales que alimentan determinados depósitos, cuyos niveles materializan el estado del sistema.

A las variables de flujo se asocian ecuaciones que definen el comportamiento del sistema. El bloque representativo de un flujo admite, como señal de entrada, la información proveniente de los estados, o de las variables auxiliares del sistema y suministra como salida el flujo que alimenta a un estado. Por ejemplo, en la figura 3.4 se representa, utilizando dos símbolos alternativos, el bloque que representa el flujo F_a , al que se puede asociar una ecuación de la forma,

$$F_a = \frac{A \times B}{D}$$

siendo A , B y D variables de estado o auxiliares. Las ecuaciones asociadas a una variable de flujo reciben la denominación de *ecuaciones de flujo* o *funciones de decisión*.

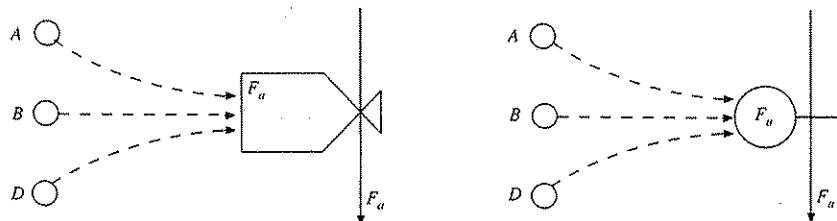


FIGURA 3.4. Representaciones alternativas de un flujo en un diagrama de Forrester.

La ecuación de flujo representa la función desarrollada por el observador del símil hidrodinámico de la figura 3.1. Es decir, con ayuda de la ecuación de flujo el observador calcula en

cada instante la abertura de la válvula, o sea el flujo; de ahí la denominación de función de decisión.

A todo estado se asocia una variable de flujo, o incluso varias, lo que gráficamente, y empleando los símbolos de la figura 3.3, se puede representar como se hace en la figura 3.5.

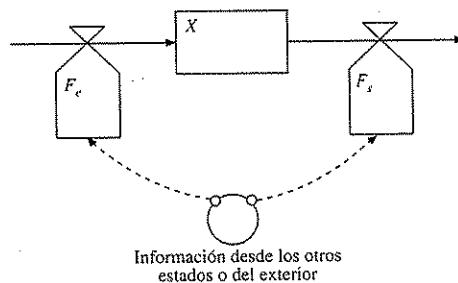


FIGURA 3.5. Conexión de un estado X a los flujos de entrada F_e y de salida F_s .

Una forma que toma muy frecuentemente la ecuación de un flujo es la que se representa en la figura 3.6 y que toma la forma

$$F(t) = T_n M(t) X(t)$$

donde T_n es una tasa normal y M es lo que se denomina un multiplicador de flujo normal. Si $M(t) = 1$ se tiene una situación neutral en la que $F(t) = T_n X(t)$, es decir, el flujo es una fracción constante y normal del nivel (por ejemplo, el número de nacimientos anuales de un modelo de población es una tasa normal de natalidad multiplicada por el nivel de la población).

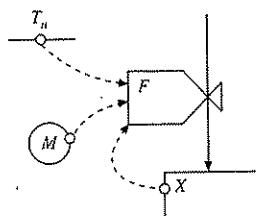


FIGURA 3.6. Representación en el diagrama de Forrester de un flujo F cuyo valor viene dado por una tasa normal T_n afectada por un multiplicador M .

Normalmente una variable de flujo dependerá no de una, sino de varias variables de acuerdo con una expresión de la forma

$$F = f(V_1, V_2, \dots, V_k)$$

es decir, el flujo F es una función de varias variables. Es frecuente que esta dependencia pueda descomponerse multiplicativamente de modo que se tenga

$$F = f_1(V_1) \times f_2(V_2) \times \dots \times f_k(V_k)$$

Así, en las funciones f_i se recoge el efecto de cada una de las variables que influyen sobre F . Una variable de flujo vendrá siempre medida por la unidad del estado al que alimenta, partida por el tiempo.

Las variables de flujo tienen como entradas exclusivamente a estados y a variables auxiliares. Es decir, dos variables de flujo no pueden conectarse entre sí. Siguiendo el símil hidrodinámico es fácil concebir cómo la decisión respecto a la abertura de la válvula, que alimenta a un cierto nivel, se toma exclusivamente en función de los valores de los otros estados; y cómo una variable de estado no puede influir directamente a otra variable de estado, sino a través del flujo que proporcione la primera.

La evolución del sistema en el tiempo comporta variaciones en los distintos estados. Estas variaciones se deben no sólo a la acción de factores externos (variables exógenas), sino, y sobre todo, a decisiones en un sentido amplio, tomadas en el interior del sistema, que se interpretan con ayuda de las funciones de decisión asociadas a las variables de flujo. En este sentido es cómo debe entenderse el que el sistema genere su propio comportamiento.

3.3.3. Variables auxiliares

Las variables auxiliares representan pasos o etapas en los que se descompone el cálculo de una variable de flujo a partir de los valores tomados por los estados. Se representan por medio de círculos como los que aparecen en la figura 3.3. Por ejemplo, en la figura 3.7 se tiene la representación, por medio de diagramas, del empleo de variables auxiliares que se indicó en las expresiones (3.6), (3.7) y (3.8).

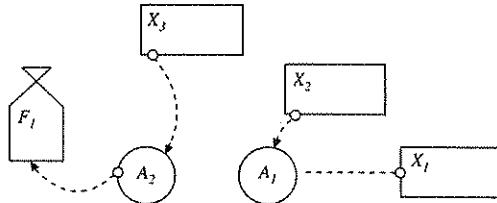


FIGURA 3.7. *Variables A_1 y A_2 como pasos intermedios en la determinación de F_1 en función de X_1 , X_2 y X_3*

Las variables auxiliares unen los canales de información entre variables de estado y de flujo; en realidad, son parte de las variables de flujo. Sin embargo, se distinguen de ellas en la medida en que tengan un significado real por sí mismas, o sencillamente porque hacen más fácil la comprensión de las ecuaciones de flujo.

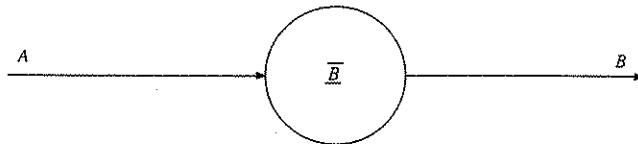


FIGURA 3.8. *Forma simbólica de representar que la variable B es una función no lineal o tabla de A .*

Las variables auxiliares se pueden emplear para representar las no-linealidades que aparecen en el sistema.

Si las variables A y B están ligadas por una expresión de la forma $B = f(A)$, en donde $f(A)$ es una función no lineal, entonces se utiliza un símbolo como el empleado para las variables auxiliares, tal como se indica en la figura 3.8. Estas funciones no lineales es habitual que se den mediante tablas de puntos, de modo que la función se completa por interpolación entre ellos. Por ello es frecuente referirse a estas funciones denominándolas *tablas*.

3.3.4. Otros símbolos empleados en los diagramas

Un estado se puede alimentar o bien desde otro estado, a través de la correspondiente variable de flujo, o bien desde una fuente exterior al sistema. En este último caso si, además, la fuente puede considerarse infinita —es decir, no agotable— se representa en los gráficos por medio de una “nube”. En la figura 3.3 se tiene este símbolo.

Por otra parte, un estado, al disminuir, puede evacuar sobre otro estado, a través de la correspondiente variable de flujo, o sobre un pozo exterior al sistema. En este último caso, y si se supone que la capacidad del pozo es infinita, se representa por medio de una “nube”. En la figura 3.9 se representa un estado, junto con las correspondientes variables de flujo, y las fuentes y pozos a él asociados.

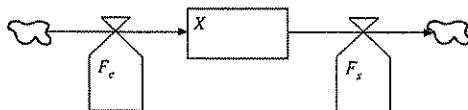


FIGURA 3.9. Diagrama de Forrester de un estado X , con los flujos de entrada F_e y de salida F_s , y las “nubes” que representan los pozos y sumideros infinitos.

Las variables de estado y de flujo están ligadas entre sí por medio de canales. En las primeras formulaciones de la dinámica de sistemas se distinguían entre dos clases de canales:

- canales materiales, los cuales se representan por un trazo continuo; y,
- canales de información, los cuales se representan por medio de un trazo discontinuo.

En la actualidad se prescinde de esa distinción y los modernos entornos informáticos ni siquiera permiten realizarla.

Con los símbolos de la figura 3.3 se puede construir un diagrama que represente el símil hidrodinámico de la figura 3.1, el cual a su vez no es sino una interpretación analógica del sistema de ecuaciones diferenciales de primer orden de las expresiones (3.2), (3.3) y (3.4). Ello es lo que se hace en la figura 3.10. En esta figura se han considerado exclusivamente las variables auxiliares definidas en la figura 3.7, aunque se hubiesen podido concebir otras más.

Un diagrama construido con ayuda de los símbolos de la figura 3.3, tal como el de la figura 3.10, recibe la denominación de *diagrama de Forrester*¹ o *diagrama de flujos-estados*. En

¹ Se emplea en este libro la denominación de diagrama de Forrester para referirse a lo que se conoce también como diagrama Dynamo. Se considera más adecuada la denominación adoptada aquí, ya que la segunda hace referencia a un

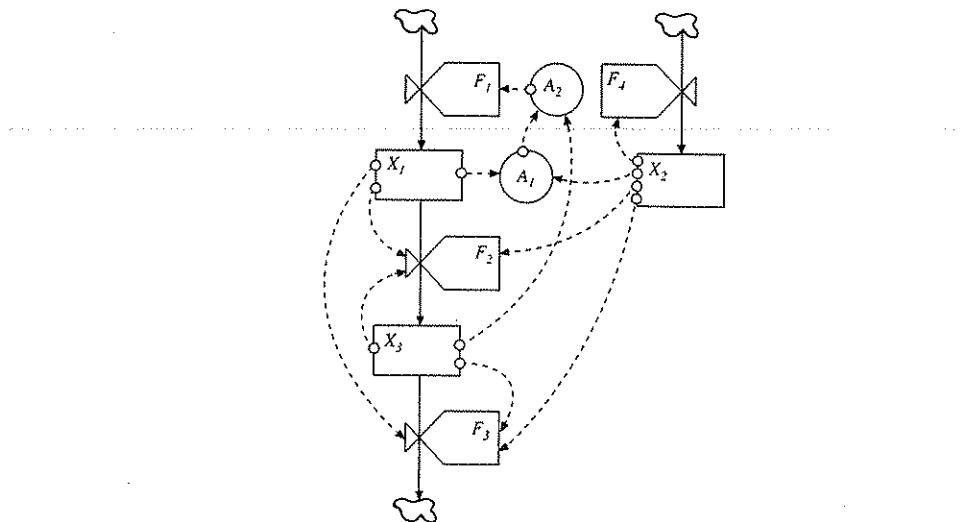


FIGURA 3.10. *Diagrama de Forrester del símil hidrodinámico de la figura 3.1.*

estos diagramas se ligan entre sí variables de estado y de flujo, a través de las correspondientes variables auxiliares.

Cualquier trayecto a través del diagrama de un sistema debe encontrar alternativamente estados y flujos y nunca dos variables del mismo tipo en sucesión (excepto variables auxiliares).

Resumiendo todo lo anterior cabe hacer dos consideraciones:

- En primer lugar debe considerarse que los procesos fundamentales que tienen lugar en un sistema pueden ser caracterizados por flujos y por estados (acumulaciones). Por ejemplo, los nacimientos se acumulan en la población, los flujos de producción se acumulan en "stocks", el personal contratado se acumula en la plantilla, etc. En este contexto es claro que integración es sinónimo de acumulación.
- En segundo lugar, aunque el flujo y la integración son inherentes a los sistemas, solamente se puede observar la integración. Los flujos son instantáneos y sólo pueden ser medidos como promedios sobre un determinado periodo. Por consiguiente, las integraciones cobran un interés singular, puesto que son las variables que pueden ser medidas y que suministran las bases prácticas para la actuación sobre el sistema.

3.4. RETRASOS

Una característica importante que debe considerarse en el estudio de los sistemas son los retrasos que se producen en la transmisión de la información o de los bienes materiales en su

lenguaje concreto de programación denominado DYNAMO, del que se hablará más adelante, restándole generalidad a estos diagramas, que como se pone de manifiesto en lo que sigue, dan lugar a modelos susceptibles de ser programados en cualquier lenguaje de alto nivel. Por otra parte, está justificada la terminología aquí adoptada ya que en estos diagramas se encuentra la aportación más original de Forrester al modelado de sistemas.

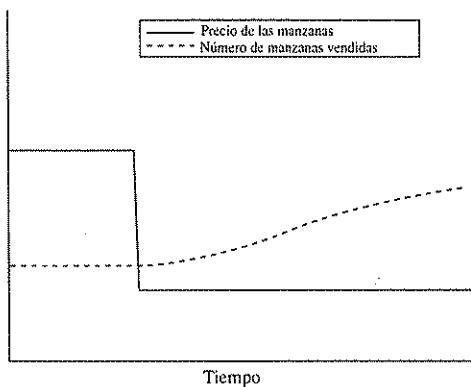
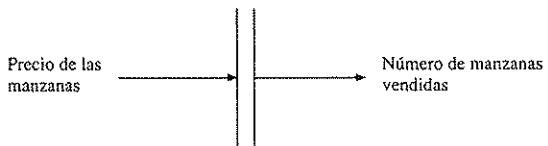


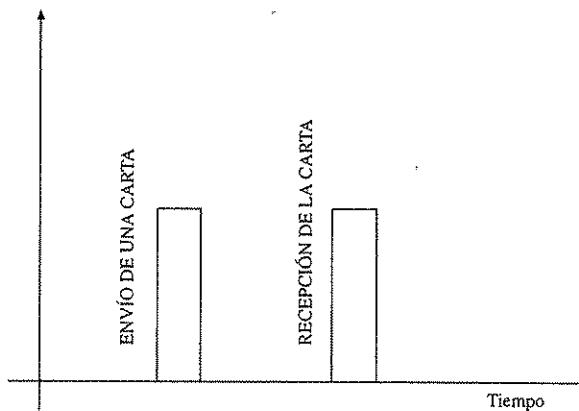
FIGURA 3.11. Retraso que se produce en la percepción de la variación del precio de un producto en el mercado.

seno. Al construir el diagrama de influencias de un sistema hay que tener en cuenta que la relación de influencia que liga a dos variables puede implicar una transmisión para la que se requiera el transcurso de un cierto tiempo. Se está entonces en presencia de un retraso. Se han propuesto también las denominaciones de retardo y de demora, aunque aquí se ha adoptado la de retraso.

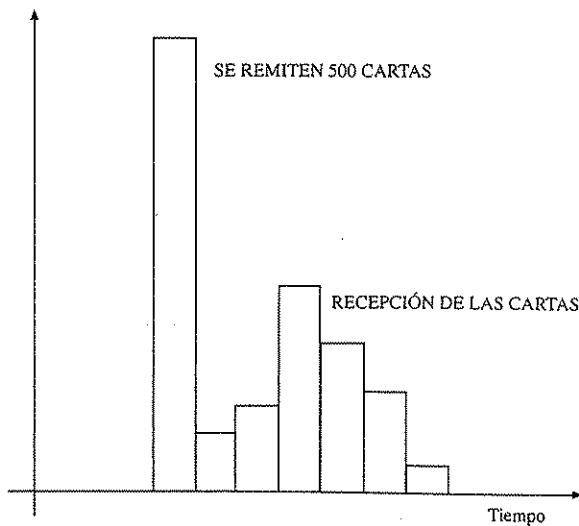
De hecho, los retrasos se producen en cualquier descripción del mundo real. Por ejemplo, la gente basa normalmente sus decisiones en la percepción que tiene del mundo, y no en su estado actual. Se necesita un cierto tiempo para formarse una idea sobre la situación real de un determinado problema antes de tomar una decisión con respecto al mismo. Por otra parte, una vez tomada una decisión, debe transcurrir algún tiempo hasta que se observen sus efectos.



Supóngase que en un instante determinado desciende de forma significativa el precio de un producto del mercado, las manzanas por ejemplo. El número de manzanas vendidas aumentará, como resultado de esta disminución de precio, sin embargo el número de manzanas vendidas no responderá de manera instantánea a la variación del precio, sino que más bien sucederá algo parecido a lo que se indica en la figura 3.11. Se justifica la forma de la curva considerando que la percepción por parte de los compradores de la disminución del precio de las manzanas requiere un cierto tiempo, que hace que se produzca un retraso entre la disminución del precio de las manzanas y el aumento de sus ventas. Uno de los símbolos que representan un retraso se muestra en la figura 3.12.

FIGURA 3.13. *Retraso de un acontecimiento aislado.*

Para los microprocesos los retrasos se convierten en pausas. Es decir, para un comprador aislado en el mercado, el retraso consiste en el tiempo que tarda en llegarle la noticia de la reducción del precio de las manzanas. Sin embargo, para variables agregadas —que son las que aquí interesan— los retrasos producen ajustes graduales entre las variables relacionadas. Por ejemplo, la figura 3.13 muestra el retraso que se presenta en la recepción de una carta aislada, mientras que en la figura 3.14 se tiene el retraso que se produce en la distribución de un número elevado de ellas. Se muestra así el efecto de agregación de distintos retrasos, que da lugar a un retraso global.

FIGURA 3.14. *Retraso de un acontecimiento agregado.*

¹ The author would like to thank the editor and anonymous referees for their useful comments and suggestions.

Un proceso como el de distribución de las cartas puede ser modelado con ayuda de variables de flujo y de estado, tal como se indica en la figura 3.15. Se tiene así un ejemplo de la función que se denomina *retraso* en dinámica de sistemas. Los ejemplos anteriores muestran que un retraso implica siempre una acumulación del material o de la información que se retrasa; por ejemplo, las cartas se acumulan en alguna parte entre el instante en que son depositadas en el buzón de correos y el momento en que llegan a manos del destinatario. Los retrasos, por consiguiente, implican la aparición de variables de estado adicionales en la construcción de un modelo. Segundo se trate de flujos de transmisión de bienes materiales o de información, los retrasos pueden ser también de materiales o de información.

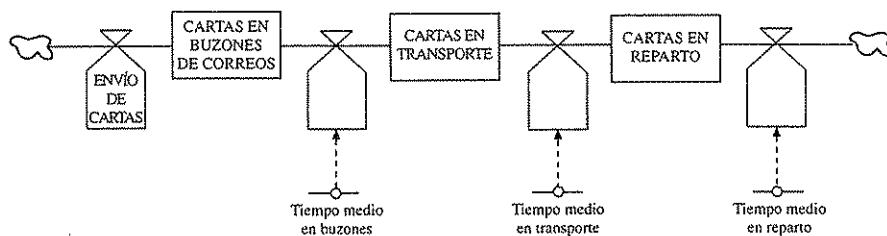


FIGURA 3.15. Ejemplo de mecanismo para la generación de retrasos.

3.4.1. Retrasos en las transmisiones de material

Los retrasos en la transmisión de materiales, o simplemente retrasos de materiales, se producen cuando existen elementos en el sistema que almacenan el material que fluye por el mismo produciendo con ello el correspondiente retraso. Un ejemplo lo constituye la distribución de cartas de correos que se ha considerado más arriba.

La forma más sencilla de obtener una aproximación a un retraso en dinámica de sistemas es mediante lo que se conoce como un retraso de primer orden. Este retraso se consigue mediante una ecuación diferencial de la forma

$$\dot{x} = a(u - x) \quad (3.12)$$

en donde $a = 1/T_a$, siendo T_a la constante de tiempo del sistema. El comportamiento de un sistema que responda a esta ecuación puede interpretarse de modo que la señal de salida x trata de ajustarse a la de entrada u , con una constante de tiempo T_a .

A continuación veremos cómo se interpreta en dinámica de sistemas el efecto del retraso del sistema de distribución de cartas. Aquí aparece una única variable de estado x responsable de la acumulación que produce el retraso. Lo que se retrasa es un flujo (el flujo de cartas) porque se acumula en la variable de estado correspondiente, variable que absorbe la diferencia entre el flujo de entrada y el de salida. El flujo de salida F_s depende del nivel alcanzado en la variable de estado x y del tiempo de retraso medio T_a (que puede ser constante o variable), de acuerdo con la expresión

$$F_s = \frac{x}{T_s}$$

70 Dinámica de sistemas

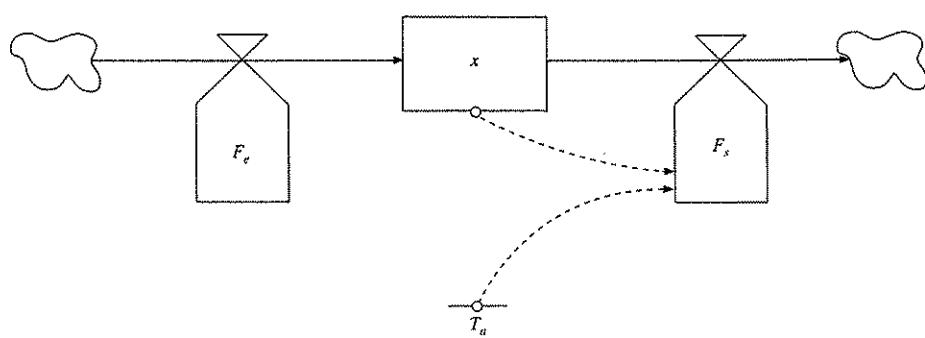


FIGURA 3.16. Diagrama de Forrester de un retraso de primer orden.

Un retraso de primer orden consiste, por tanto, en un flujo de entrada F_e que se acumula en una variable de estado x y que fluye “retrasado” a la salida mediante un flujo F_s . Su diagrama de influencias básico se muestra en la figura 3.17. El correspondiente diagrama de Forrester es el de la figura 3.16. De acuerdo con este diagrama las ecuaciones correspondientes son:

$$\frac{dx}{dt} = F_e - F_s = F_e - \frac{x}{T_a} \quad (3.13)$$

es decir

$$\frac{dx}{dt} + \frac{x}{T_a} = F_e$$

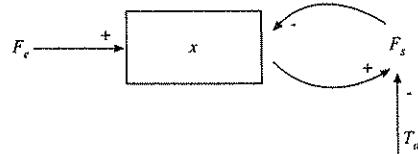


FIGURA 3.17. Diagrama de influencias básico de un retraso de primer orden.

Esta ecuación define la variación de x , variable que representa la acumulación de bienes, dinero o gente en tránsito, y que sufre el retraso. Si u es la señal que se pretende retrasar se toma $F_e = u/T_a$, con lo que la ecuación anterior se convierte en

$$\frac{dx}{dt} + \frac{x}{T_a} = \frac{u}{T_a}$$

por tanto

$$u = x + T_a \frac{dx}{dt}$$

y se tiene la expresión (3.12).

La respuesta de este sistema a un escalón unitario de entrada es la que se muestra en la figura 3.18.

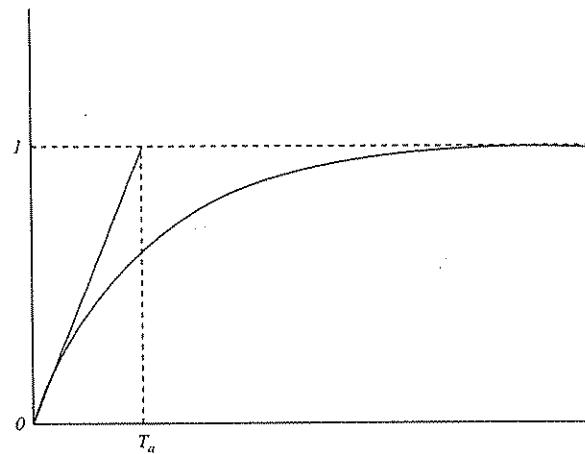


FIGURA 3.18. Respuesta de un retraso de primer orden a un escalón unitario de entrada.

Para obtener una mejor aproximación a un retraso se emplean varios sistemas de primer orden en serie. En la figura 3.19 se tiene un retraso de tercer orden, es decir, formado por tres variables de estado. De hecho es parecido al que se encontró más arriba al tratar el problema de la distribución de las cartas. Debe observarse que el tiempo de ajuste $1/a$ de cada uno de los flujos es un tercio del correspondiente al conjunto del retraso T_a . Las ecuaciones del retraso son en este caso:

$$\begin{aligned}\frac{dz_2}{dt} &= a(u - z_2) \\ \frac{dz_1}{dt} &= a(z_2 - z_1) \\ \frac{dx}{dt} &= a(z_1 - x)\end{aligned}\tag{3.14}$$

donde $1/a = T_a/3$.

De manera análoga se tiene un retraso de orden n . Cada uno de los flujos tendrá un tiempo de ajuste $T_a/n = 1/a$, en donde T_a es el tiempo de ajuste del conjunto.

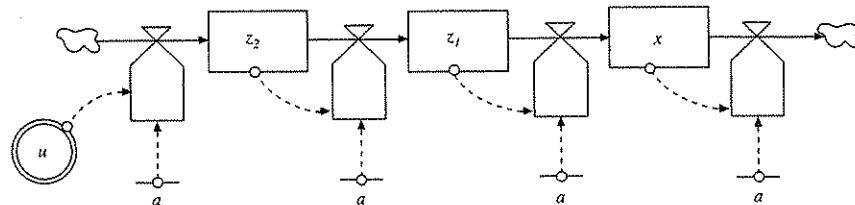


FIGURA 3.19. Retraso de tercer orden en la transmisión de material.

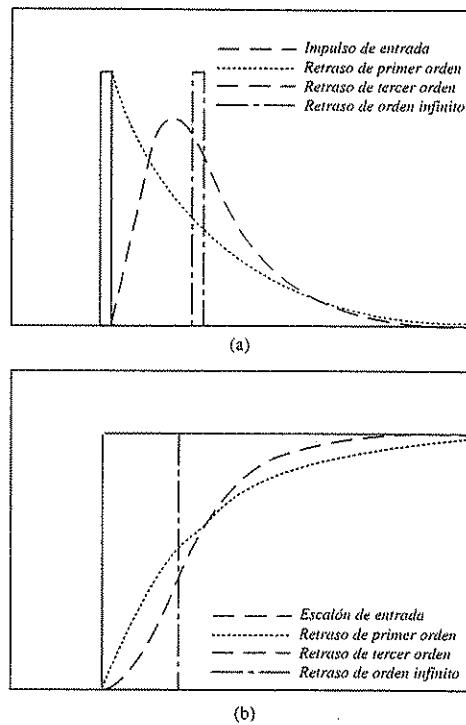


FIGURA 3.20. Evolución en el tiempo de retrasos de primer orden y tercer orden: (a) respuesta a un impulso y (b) respuesta a un escalón.

En las figuras 3.20a y 3.20b se representa la respuesta de un retraso de material para distintos valores de n a una entrada en escalón, es decir, a un cambio súbito en el flujo de entrada, y a un impulso. Se observa cómo al aumentar el orden n del retraso se alcanza una mayor fidelidad en la reproducción de los impulsos.

Con relación a la inclusión de retrasos en las relaciones de influencia deben tenerse en cuenta los siguientes efectos:

1. para $t \rightarrow \infty$, $x = u$, por lo que los equilibrios del sistema no vienen alterados. Es decir, a largo plazo la entrada se hace igual a la salida con independencia de los retrasos que se produzcan;
2. por el contrario, durante el transitorio, el efecto de un retraso puede ser inestabilizador.

Para ilustrar este último punto veamos qué sucede en un sistema de primer orden si se introduce un retraso en su cadena de realimentación. Sea el sistema de primer orden con realimentación negativa de la figura 3.21, cuya ecuación es

$$\frac{dy}{dt} = k(r - y)$$

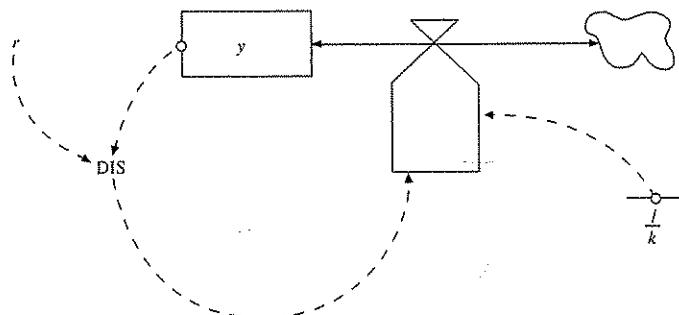


FIGURA 3.21. Sistema de primer orden con realimentación negativa.

donde y es la variable de salida del sistema, r la referencia (valor deseado de la salida) y k una constante. Obsérvese que si se hace $k = a$ esta ecuación es idéntica a la ecuación (3.12). Por tanto, el comportamiento del sistema de primer orden ante una entrada en escalón es el que se vio en la figura 3.18 y que se vuelve a reproducir en la figura 3.22a.

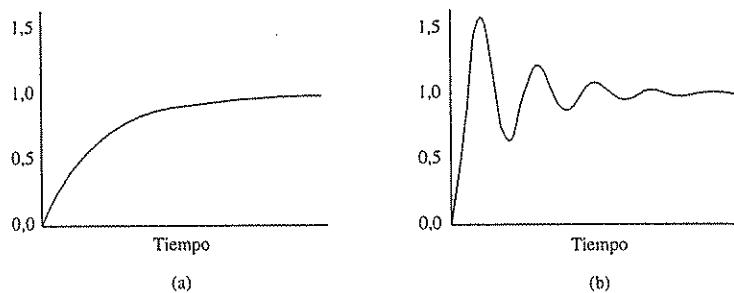


FIGURA 3.22. Comportamiento de un sistema de primer orden: a) sin retraso; b) con retraso.

Si se introduce un retraso en la cadena de realimentación se tendrá el sistema que se muestra en la figura 3.23, y cuyas ecuaciones son:

$$\frac{dy}{dt} = kx \quad (3.15)$$

$$\frac{dx}{dt} = a(u - x) \quad (3.16)$$

$$u = r - y \quad (3.17)$$

siendo $a = 1/\tau$. Para valores suficientemente grandes del retraso $1/a$, este sistema presenta oscilaciones como se refleja en la figura 3.22b y como se puede comprobar utilizando técnicas convencionales de análisis matemático.

En consecuencia, la introducción del retraso puede modificar el comportamiento de un sistema de primer orden y convertirlo en oscilatorio.

74 Dinámica de sistemas

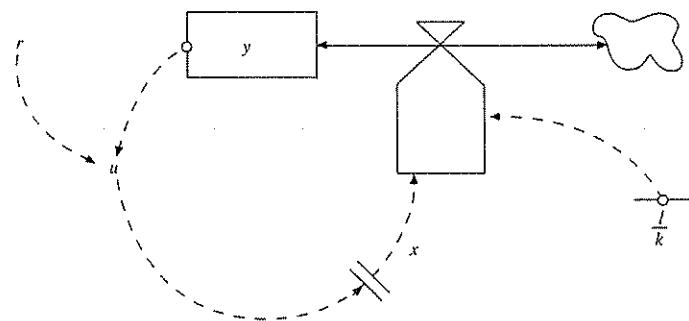


FIGURA 3.23. Sistema con retraso en la realimentación.

3.4.2. Retrasos en la transmisión de información

Los retrasos en la transmisión de información resultan de la necesidad de procesar información sobre el sistema con el fin de averiguar las tendencias subyacentes, antes de proceder a una toma de decisión. Este proceso introduce un retraso en la toma de decisiones, ya que éstas no se toman hasta que las pautas correspondientes han sido reveladas.

De esta naturaleza es el retraso considerado más arriba en el problema del precio de las manzanas. De hecho, los retrasos en la transmisión de información actúan como filtros alisadores (*smoothing*) que son capaces de alisar los picos que presenta la evolución de una variable tomando un valor promedio de la misma, tal como se representa en la figura 3.24. En el proceso de promediar se ponderan los datos disponibles de manera que los más recientes influyan en el promedio de forma más significativa que los más antiguos.

La ecuación de un promedio exponencial asociado a un retraso de información es la misma que la de un retraso de material.

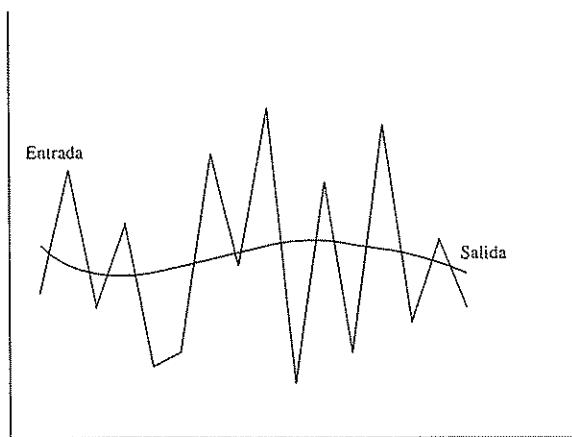


FIGURA 3.24. Efecto de un filtro alisador sobre la evolución de una variable.

3.5. PROGRAMACIÓN DEL MODELO EN EL COMPUTADOR

Hasta ahora se han descrito las dos primeras etapas para el modelado mediante la dinámica de sistemas: la construcción del diagrama de influencias y del diagrama de Forrester. El siguiente paso es desarrollar el modelo en el computador, es decir, realizar un programa que permita simularlo. Para ello, previamente es necesario definir, de forma unívoca, el sistema dinámico que se pretende simular. Esta definición no se consigue con el diagrama de Forrester, puesto que en él solamente están definidas las ecuaciones de las variables de estado. Para el resto de las variables no se definen las ecuaciones que rigen su comportamiento (aunque se especifica el conjunto de variables que interviene en estas ecuaciones). Por tanto, para continuar con el proceso de modelado es necesario especificar completamente estas relaciones, es decir, definir totalmente el sistema dinámico que se pretende programar. Así, si en el diagrama de Forrester aparece que un cierto flujo F depende de un estado X , es preciso definir esta relación de forma que sea posible evaluar F conociendo X (aunque no es necesario expresar esta relación de forma analítica: bastaría, por ejemplo, con expresarla en forma de tabla de valores). Cuando no se disponga de una información tan detallada se pueden realizar ciertas hipótesis adicionales para construir el modelo. Finalmente se tendrá definido unívocamente un sistema dinámico (no necesariamente de forma analítica).

Una vez definidas todas las relaciones entre las variables del modelo, éste ya se puede programar en un computador. Para ello se puede utilizar cualquier lenguaje de alto nivel (como FORTRAN, BASIC o C) o alguno de los entornos informáticos de simulación que utilizan lenguajes específicos para la dinámica de sistemas (como DYNAMO, STELLA, POWERSIM o VENSIM).

Utilizando un lenguaje de alto nivel el programador se libera de la tarea de aprender un nuevo entorno de programación y un nuevo lenguaje, e incluso se ahorrará el coste de adquisición del nuevo paquete informático. Sin embargo, esta opción lleva consigo bastantes inconvenientes que pueden ser definitivos a la hora de elegir, sobre todo cuando se desean realizar modelos de cierta envergadura. Entre las desventajas de utilización de un lenguaje de alto nivel se pueden citar las siguientes:

- Es necesario programar métodos numéricos de integración de ecuaciones diferenciales.
- El orden con el que se escriben las ecuaciones no es indiferente, por lo que se debe invertir un cierto tiempo y esfuerzo en ordenarlas.
- No se dispone de herramientas adicionales para la gestión de los resultados de las diversas simulaciones que se realicen, de los conjuntos de parámetros que las definen, etc.

Por el contrario, los entornos informáticos específicos de dinámica de sistemas liberan al programador de implementar los algoritmos de integración (incluso, frecuentemente, permiten cambiar de método con tan sólo pulsar una tecla). También realizan la ordenación automática de las ecuaciones, de manera que el programador se puede despreocupar de este problema. Por último, suelen incluir herramientas que facilitan la realización de simulaciones y la presentación de resultados.

La programación de modelos en estos entornos se puede realizar de dos maneras distintas:

- Escribiendo las ecuaciones en un lenguaje determinado, de manera similar a como se hace en un lenguaje de alto nivel, con las diferencias reseñadas anteriormente.

76 Dinámica de sistemas

- Utilizando un entorno gráfico con el que se dibuja el diagrama de influencias o el de Forrester y se escriben después las ecuaciones que definen cada variable.

Algunos entornos de programación permiten elegir entre ambas modalidades, mientras que otros están limitados a una de ellas.

No es el objetivo de este libro explicar cómo se utiliza alguno de estos entornos informáticos. Tampoco se va a suponer que el lector conoce ninguno de ellos. Sin embargo, si es necesario expresar las ecuaciones de alguna forma determinada en los ejemplos que se van a desarrollar. Esto no es un inconveniente grave, ya que las relaciones entre las variables se pueden expresar con notación matemática convencional de manera que sean fácilmente comprensibles. Además, la traducción de estas expresiones a un lenguaje de simulación suele ser inmediata (en la mayoría de los casos bastará sustituir el signo del producto \times por un asterisco $*$, y el de la operación de la potenciación x^y por algún símbolo propio del lenguaje que se esté utilizando, por ejemplo, x^y ; también se deberá sustituir la coma decimal por un punto y eliminar los acentos). Una ventaja adicional de esta adopción es la independencia respecto a cualquier lenguaje de simulación, con lo que se evita que la notación pueda envejecer al evolucionar los lenguajes.

Una expresión que necesita una atención especial es la que define las variables de estado o de nivel. Ya se ha comentado que esta ecuación es de la forma:

$$\frac{dX}{dt} = F(t) \quad (3.18)$$

donde $F(t)$ es el flujo que alimenta el estado $X(t)$. Ésta será la notación que se utilizará en los ejemplos. Para definir totalmente la evolución de X es necesario especificar una condición inicial:

$$X(t_0) = X_0$$

A continuación se va a exponer cómo se expresan estas ecuaciones en algunos de los lenguajes de simulación más utilizados:

- En DYNAMO:

$$L X . K = X . J + DT * F . JK$$

$$N X = X_0$$

El lector con conocimiento de métodos numéricos de integración de ecuaciones diferenciales reconocerá, en la primera de estas ecuaciones, el método de Euler si tiene en cuenta que, en DYNAMO, la variable DT es el paso de integración y las letras J y K detrás de un punto hacen referencia a instantes de tiempo. En cualquier caso, la ecuación no indica la forma en la que se integrará el sistema: es solamente la expresión en DYNAMO de la ecuación (3.18).

- La programación en ITHINK y STELLA se hace en entorno gráfico. Los estados se implementan dibujando el diagrama de la figura 3.25a y especificando que la condición inicial para X es X_0 .

Las ecuaciones resultantes, generadas automáticamente, aparecen en la figura 3.25b.

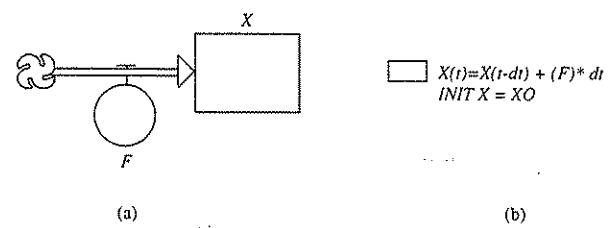


FIGURA 3.25. Definición de un estado en ITHINK/STELLA: (a) diagrama, (b) ecuaciones resultantes.

- En POWERSIM, al igual que en el caso anterior, la programación se hace en entorno gráfico. El diagrama que define un estado aparece en la figura 3.26, así como las ecuaciones generadas automáticamente.

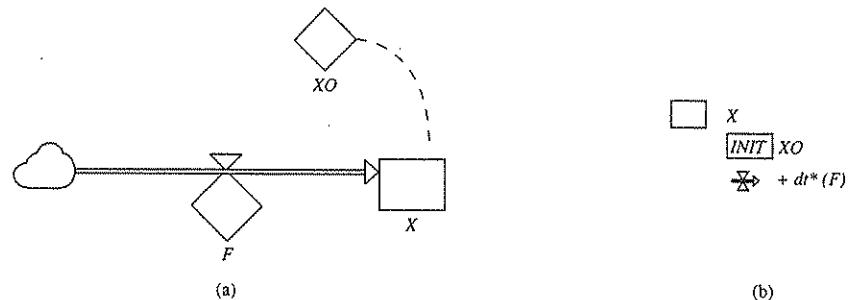


FIGURA 3.26. Definición de un estado en POWERSIM: (a) diagrama, (b) ecuaciones resultantes

- En el entorno VENSIM la programación se puede realizar de forma gráfica o escribiendo las ecuaciones. Por cualquiera de ambos métodos la ecuación resultante para el caso de un estado es:

$$X = \text{INTEG}(F, X_0)$$

Además de estas expresiones que definen las ecuaciones de las variables de estado, y de las expresiones anteriormente mencionadas para traducir otro tipo de ecuaciones, los lenguajes de simulación ofrecen otras funciones. Algunas de ellas permiten utilizar elementos más complejos de la dinámica de sistemas —por ejemplo, funciones de retraso—, mientras que otras son similares a ciertas funciones de los lenguajes de propósito general —como la asignación alternativa IF que se comentará en el próximo capítulo. Con todo ello, estos lenguajes son una herramienta poderosa que permite programar cómodamente los modelos.

3.5.1. Los datos del modelo

Ya se ha indicado que, para programar informáticamente el modelo, es necesario especificar completamente la relación entre las distintas variables; es decir, al diagrama de Forrester que

define la estructura del modelo es necesario añadirle más información. Entre esta información se encuentran lo que podemos denominar datos del modelo y que se pueden clasificar en tres grupos:

- Constantes o datos estáticos: son parámetros que se corresponden con variables que no varían en el tiempo y, por tanto, se definen con un solo valor. Normalmente estos datos se incluyen dentro del modelo con sentencias del tipo $A = 1$ (o con sintaxis similares, dependiendo del lenguaje escogido). Los entornos de simulación suelen incluir herramientas para cambiar estos valores en distintas simulaciones sin tener que cambiar las ecuaciones correspondientes, para facilitar el estudio de la sensibilidad a estos parámetros.
- Valores numéricos que definen las funciones no lineales: como ya se ha comentado anteriormente, en dinámica de sistemas es habitual definir las relaciones no lineales entre distintas variables por medio de tablas de valores que quedan caracterizadas por una serie de datos numéricos. Así si la variable B depende de A por medio de una función no lineal f ($B = f(A)$), la función f se puede aproximar por medio de un conjunto de N pares de valores: $\{(a_1, b_1), (a_2, b_2), \dots, (a_N, b_N)\}$.
- Datos dinámicos $[x(t_1), x(t_2), \dots]$: corresponden a variables x que cambian de valor a lo largo del tiempo pero que no son calculadas por el modelo, sino que son datos de éste. Por tanto, para definirlas sería necesario especificar la trayectoria de las mismas. Esto se hace normalmente indicando los valores de estas variables en ciertos instantes de tiempo t_i , de manera que entre estos instantes la trayectoria será calculada por interpolación. Normalmente, esto se puede realizar de dos formas distintas:
 - Definiendo una tabla en la que la variable independiente es el tiempo. Esto es, especificar una función f de tal manera que la variable dinámica D se calcularía haciendo $D = f(t)$, donde t es el tiempo.
 - Especificando estos datos en un fichero anexo al modelo. De esta manera se pueden realizar varias simulaciones del modelo con distintos datos cambiando el fichero de datos.

Estos datos se denominan frecuentemente *variables exógenas* ya que son magnitudes dinámicas que no son calculadas por el modelo.

La bondad del modelo dependerá sobre todo de la exactitud de los datos incluidos en los dos primeros puntos de esta clasificación, ya que son datos intrínsecos al modelo. Por el contrario, los datos dinámicos son externos al modelo y, aunque si partimos de variables exógenas incorrectas obtendremos resultados erróneos, esto se deberá a la imperfección de los datos con los que alimentamos el modelo y no a las deficiencias del mismo.

3.6. DINÁMICA DE SISTEMAS Y SISTEMAS DINÁMICOS

En esta sección se va a estudiar la forma del sistema dinámico asociado a un diagrama de Forrester. Veremos cómo a este diagrama corresponde una clase bastante amplia de sistemas dinámicos, con la ventaja de aportar una representación gráfica de la estructura del sistema

dinámico. Recordemos que un sistema dinámico en variables continuas se escribe mediante la expresión:

$$\frac{dx}{dt} = f(x) \quad (3.19)$$

donde x es un vector de dimensión n . La ecuación (3.19) es una ecuación vectorial que representa n ecuaciones escalares de la forma:

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \dots, x_n) \quad i = 1, \dots, n$$

En esta expresión f_i es una función no lineal de las variables x_1, x_2, \dots, x_n . En dinámica de sistemas se admiten únicamente no-linealidades en funciones de una única variable (correspondientes a la influencia de una única variable sobre otra). Estas funciones no lineales de una única variable se combinan mediante las operaciones aritméticas elementales para formar la función de varias variables. Recuérdense, por ejemplo, las consideraciones realizadas en la sección 3.3.2 en las que se combinaban de forma multiplicativa las funciones no lineales de una variable para formar la ecuación de las variables de flujo.

Con el fin de conseguir representar las funciones f_i a partir de funciones de una sola variable se introducen las variables auxiliares w_1, \dots, w_m , y las funciones auxiliares g_1, \dots, g_m . Cada función auxiliar g_i puede tener una de las dos formas siguientes:

- o bien es una función no lineal de una única variable x_i o w_i ;
- o bien representa una combinación aritmética de variables x_i o w_i , de modo que entre ellas se tienen únicamente las operaciones suma, resta, multiplicación o división.

Entonces f_i resulta de la aplicación de las cuatro operaciones aritméticas elementales a las variables auxiliares w_1, \dots, w_m y a las funciones auxiliares g_1, \dots, g_m .

Un ejemplo aclarará este proceso de descomposición. Sea el sistema dinámico

$$\frac{dx}{dt} = (\beta_1 x - \alpha_1)^{\alpha_1} (\alpha_2 - \beta_2 x)^{\alpha_2} x \quad (3.20)$$

en donde x es la variable cuya evolución se estudia y $\alpha_1, \alpha_2, \beta_1$ y β_2 son parámetros cuyo valor permanece constante. Se pueden definir las variables auxiliares w_1, w_2, w_3 y w_4 , de acuerdo con las expresiones

$$\begin{aligned} w_1 &= \beta_1 x - \alpha_1 = g_1(x) \\ w_2 &= \alpha_2 - \beta_2 x = g_2(x) \\ w_3 &= w_1^{\alpha_1} = g_3(w_1) \\ w_4 &= w_2^{\alpha_2} = g_4(w_2) \end{aligned}$$

Con ayuda de estas variables auxiliares, el sistema (3.20) se puede escribir

$$\frac{dx}{dt} = w_3 w_4 x$$

Debe observarse que las funciones auxiliares g_i están definidas recurrentemente, de modo que resolviéndolas para las variables w_i (sencillamente por sustitución) se obtiene de nuevo la función f .

El proceso de descomposición puede ilustrarse también considerando el grafo que representa una relación de influencia. Por ejemplo, considérese el grafo de la figura 3.27a en el que se muestra la relación de influencia de x_1, x_2 y x_3 sobre dx_1/dt . A esta relación de influencia corresponderá una ecuación diferencial de la forma:

$$\frac{dx_1}{dt} = f(x_1, x_2, x_3)$$

Supongamos ahora que descomponemos esta influencia, mediante la introducción de dos variables auxiliares w_1 y w_2 , tal como se indica en la figura 3.27b. Con ello se quiere decir que la secuencia de cálculos que se establece es la siguiente:

$$\begin{aligned} w_1 &= \varphi_1(x_1) \\ w_2 &= \varphi_2(w_1, x_2) \\ \frac{dx_1}{dt} &= h(x_3, w_2) \end{aligned}$$

de modo que se cumplan las condiciones más arriba expuestas; es decir, las dependencias funcionales lo son de una sola variable excepto cuando se refieren a operaciones aritméticas. En particular, la función φ_2 será la suma, la resta, la multiplicación o la división de las variables w_1 y x_2 (resp. la función h de las variables x_3 y w_2).

Cuando la función f_i representa una relación de influencia multivariable ($dx_i/dt = f_i(x_1, x_2, \dots, x_n)$) es habitual que, con la descomposición anterior, se llegue a una estructura en la que los efectos de las distintas variables x_1, x_2, \dots, x_n se formulen de forma separada y relacionados entre sí por medio de operaciones de suma o multiplicación (aunque esta formulación sólo sea una aproximación que permite simplificar las expresiones). Así se tienen dos tipos habituales de separaciones: la multiplicativa y la aditiva. De acuerdo con la primera se escribe:

$$f_i(x_1, x_2, \dots, x_n) = f_{i1}(x_1) \times f_{i2}(x_2) \dots \times f_{in}(x_n)$$

En donde las f_{ij} son funciones de una única variable. La separación aditiva se escribe

$$f_i(x_1, x_2, \dots, x_n) = f_{i1}(x_1) + f_{i2}(x_2) \dots + f_{in}(x_n)$$

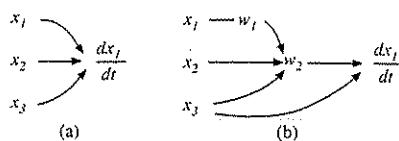


FIGURA 3.27. Grafo de relaciones de influencia.

En la construcción de modelos se emplean habitualmente estos dos tipos de separaciones. La separación aditiva se encuentra en la base de las linealizaciones que conducen al empleo de los sistemas dinámicos lineales. En los modelos de dinámica de sistemas, en los que las características no lineales se consideran muy relevantes, se emplea habitualmente la separación multiplicativa, como ya se indicó en la sección 3.3.2. Ello se debe a que se considera un supuesto comportamiento normal, y se trata de ver cómo se desvía el comportamiento cuando se perturba en torno a ese comportamiento normal, conservando los efectos no lineales.

La introducción de las variables auxiliares y la descomposición de la función f conduce a un aumento considerable del número de variables, y puede dar la impresión de que se incrementa la complejidad del sistema en términos del número de interconexiones. Sin embargo, esto no es así, ya que cada relación individual, que resulta después de la separación, es mucho más simple. Además, se puede asociar un diagrama que permite visualizar la estructura del sistema.

Este hecho posee un considerable interés metodológico. En efecto, el diagrama de Forrester constituye una reelaboración del diagrama de influencias, el cual, a su vez, es una formalización de la descripción verbal del modelo mental de los especialistas en S . Se establece así un continuo que va desde la descripción en lenguaje ordinario del modelo mental al diagrama de influencias, de éste al de Forrester, y de este último a un sistema dinámico (el cual incorpora además valores cuantitativos de los parámetros del sistema). El diagrama de Forrester, por tanto, es un paso intermedio en el proceso que va desde la descripción verbal al sistema dinámico, con la ventaja de ser accesible a todos los especialistas que participan en la construcción del modelo M . Con él se tiene un diagrama, relativamente sencillo, que puede ser entendido incluso por quien carezca de familiaridad con la notación matemática, y al que se asocia un objeto matemático bastante elaborado, como un sistema dinámico.

Para acabar este capítulo conviene realizar ciertas consideraciones sobre el proceso de modelado. Pudiera parecer que este proceso consiste en la realización secuencial de tres pasos: primero se parte del modelo verbal; de él se obtendría el diagrama de influencias, y, a su vez, de este último el sistema dinámico, pasando por el diagrama de Forrester. Sin embargo, esto no es así de simple, y el proceso de modelado es más complejo y comporta sucesivas reelaboraciones entre estas diferentes descripciones del sistema. En realidad, en el proceso de modelado se puede considerar que interactúan diferentes lenguajes (el ordinario, en la descripción verbal; el de los grafos, en el diagrama de influencias; el de los sistemas dinámicos, en las ecuaciones del modelo; así como los datos y registros cuantitativos), y que el modelista emplea cada uno de ellos en cada momento según el problema concreto que le ocupa, tratando, en último extremo, de alcanzar un razonable grado de coherencia *interna* entre las diferentes descripciones. El modelista es, en cierta forma, un traductor entre lenguajes, y sabe que en toda traducción se gana y se pierde. Su arte consiste en preservar el mayor contenido descriptivo entre las diferentes traducciones, tratando de enriquecer el producto final.

4.1. INTRODUCCIÓN

Se va a dedicar este capítulo a desarrollar algunos ejemplos elementales de aplicación de la dinámica de sistemas con el fin de ilustrar los conceptos vertidos anteriormente.

4.2. BUCLE DE REALIMENTACIÓN NEGATIVA

Como ejemplo de un bucle de realimentación negativa considérese un sistema de control de un inventario, en el que un almacénista trata de mantener su inventario en un determinado nivel. Cuando el stock de productos almacenados desciende del nivel deseado, entonces realiza pedidos al distribuidor con el fin de mantener dicho nivel. Si la cantidad de producto almacenado es superior al nivel deseado, entonces devuelve este exceso de producto al distribuidor. Se supone, además, que las ventas están determinadas con independencia del vendedor (son una variable exógena).

El diagrama de influencias de este sistema aparece en la figura 4.1, mientras que en la figura 4.2 se representa su diagrama de Forrester. Debe observarse que, de acuerdo con las observaciones hechas en el párrafo anterior, el flujo de pedidos puede ser positivo o negativo.

El flujo de pedidos depende de la diferencia (discrepancia) entre el inventario deseado y el inventario real. La cuestión que inmediatamente se presenta es la de establecer la naturaleza de esta dependencia. En la figura 4.3 se muestran distintas posibles respuestas a esta cuestión. Cuando el inventario real es igual al inventario deseado, entonces el flujo de pedidos es cero; sin embargo, cuando esto no sucede así existen distintas maneras de actuar que se reflejan en la figura 4.3.

Cada una de las curvas *A*, *B* y *C* de la figura 4.3 representa distintos modos de tomar la decisión respecto a qué flujo de pedidos realizar cada semana, en función del nivel del inventario en la misma. Las gráficas *A* y *B* son lineales, es decir, el flujo de pedidos es proporcional a

84 Dinámica de sistemas

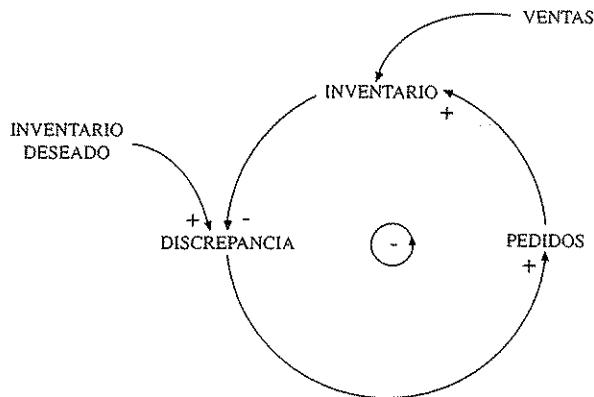


FIGURA 4.1. Diagrama de influencias del sistema de control de un inventario.

la discrepancia entre los inventarios real y deseado. La recta *B* representa un flujo de pedidos más rápido que el de la *A*. La curva *C* es no lineal.

Obsérvese que, teniendo en cuenta el comportamiento creciente de las curvas de la figura 4.3, el bucle que aparece en la figura 4.1 (el que liga las variables INVENTARIO, DISCREPANCIA y PEDIDOS) es un ejemplo de bucle de realimentación negativa, según se explicó en la sección 2.2.1. Por tanto, este modelo estará dotado de un mecanismo autorregulador (que es precisamente el que se pretende modelar: el almacenista desea que el inventario permanezca en un valor deseado).

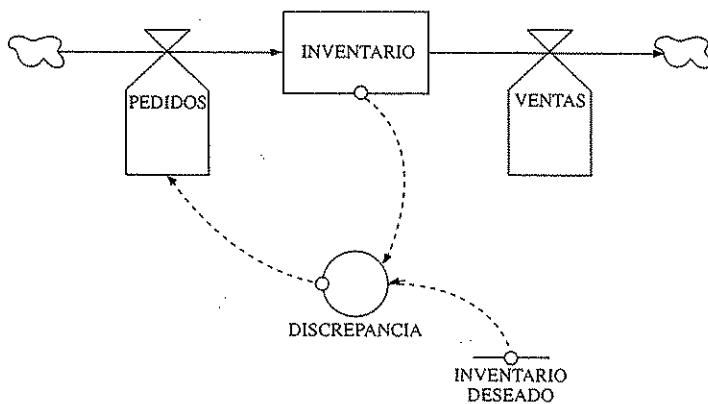


FIGURA 4.2. Diagrama de Forrester del sistema de control de un inventario.

Para determinar completamente las ecuaciones del modelo es necesario dar valores numéricos a las distintas constantes y condiciones iniciales del mismo. Supóngase que el sistema se encuentra inicialmente en equilibrio y con el inventario igual al deseado, el cual es de 100 unidades. Supóngase que las ventas, de repente, varían de 0 a 20 unidades/semana en la semana cero de la simulación. Se supone que las unidades pedidas por semana son proporcionales a la discrepancia (es decir, la política de pedidos es lineal) con un factor igual a 0,5.

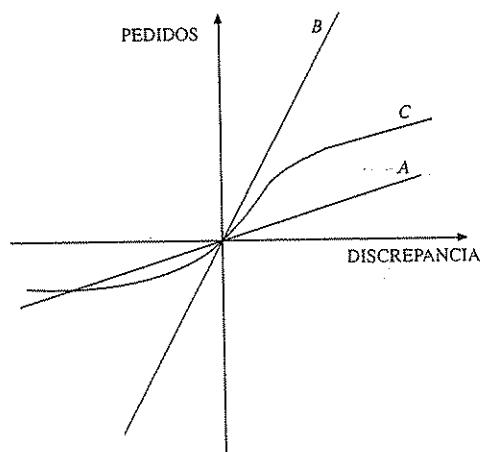


FIGURA 4.3. Distintas políticas de pedidos en función de la discrepancia del inventario respecto al deseado.

A continuación se van a presentar las ecuaciones del modelo con la notación escogida, en la que se van a utilizar las variables que se describen en la tabla 4.1.

Nombre	Significado
INV	Inventario
PED	Pedidos
VEN	Ventas
DIS	Discrepancia
INVD	Inventario deseado
FPED	Factor para el cálculo de los pedidos

TABLA 4.1. Significado de las variables del modelo de control de un inventario.

Las ecuaciones resultantes son:

$$\begin{aligned}
 \frac{d(\text{INV})}{dt} &= \text{PED}-\text{VEN} \\
 \text{INV}(t_0) &= 100 \\
 \text{DIS} &= \text{INVD}-\text{INV} \\
 \text{PED} &= \text{FPED} \times \text{DIS} \\
 \text{FPED} &= 0,5 \\
 \text{VEN} &= 20 \\
 \text{INVD} &= 100
 \end{aligned}$$

La primera de estas ecuaciones define la evolución de la variable de estado INV utilizando la notación de la expresión (3.18) —cuando se programe esta ecuación utilizando un lenguaje determinado será necesario traducirla siguiendo las pautas que se indicaron en la sección 3.5. En este caso el flujo neto que alimenta esta variable es la diferencia entre los pedidos PED y las ventas VEN. La siguiente ecuación establece el valor inicial de la variable INV. La siguiente ecuación determina la discrepancia en función del inventario deseado INVD y del inventario real INV. La ecuación correspondiente a los pedidos PED define la política con la que se realizan los mismos. El resto de las ecuaciones define las constantes del modelo.

La evolución de las principales magnitudes que intervienen en el sistema se representa en la figura 4.4.

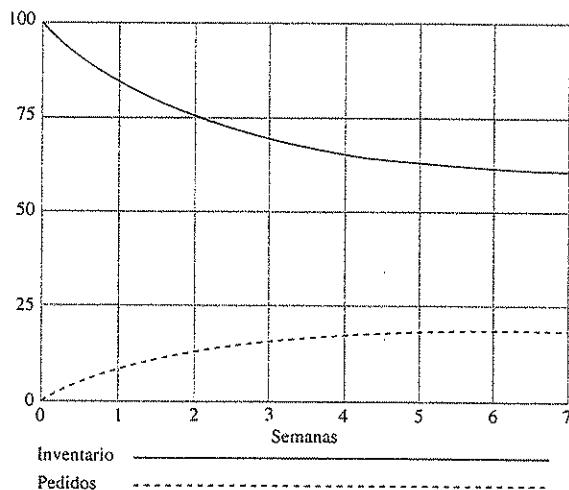


FIGURA 4.4. *Evolución del inventario y del flujo de pedidos.*

El resultado de la simulación que se muestra en la figura 4.4 puede parecer aparentemente inesperado¹. En efecto, se observa que el nivel del inventario se estabiliza en 60 unidades, no alcanzando nunca el nivel deseado de 100. Esto muestra que una política de pedidos basada exclusivamente en la discrepancia entre los niveles de inventario deseado y real no es adecuada para cumplir el objetivo de mantener el inventario a un nivel deseado.

Es fácil ver por qué se produce este desajuste. Observando la figura 4.4 se ve cómo entre las semanas 1 a 6 el número de unidades pedidas es inferior a 20. Sin embargo, el número de unidades vendidas en estas semanas es 20. Prácticamente hasta la semana 7 el flujo de pedido no se estabiliza en 20 unidades, que son precisamente las unidades vendidas en cada semana, por lo que se habrá producido una merma en el inventario inicial debida precisamente a los desajustes entre ventas y pedidos entre las semanas 1 a 6. Además, la política de pedidos

¹ Este resultado no extrañará al especialista en regulación automática, que conoce perfectamente que con una acción proporcional no se corrigen los errores en régimen permanente tratando con sistemas de tipo cero.

establecida no consigue recuperar esta disminución, pues en ningún caso los pedidos superan al valor de las ventas. Se concluye, como ya se decía más arriba, que esta política de pedidos no es satisfactoria.

En cualquier caso, este resultado no contradice el hecho de que los bucles de realimentación negativa son autorregulados. Esta autorregulación ha de entenderse como la propiedad del sistema (que emana de su estructura) de evolucionar hacia valores constantes. De hecho, la evolución de las variables de la figura 4.4 se corresponde con la descrita en la sección 2.2.1. Sin embargo, el valor que se alcanza no tiene que ser igual al de las variables deseadas o de referencia que suelen aparecer en este tipo de bucles.

Discrepancia	- 50	- 40	- 30	- 20	- 10	0	10	20	30	40	50	60
Pedidos	- 12	- 11	- 9	- 7	- 4	0	8	14	18	21	23	25

TABLA 4.2. *Tabla de valores que define un política de pedidos no lineal.*

La dinámica de sistemas y los lenguajes de simulación basados en ella no solamente contemplan funciones lineales como la correspondiente al flujo de pedidos anterior, sino que permiten incorporar otro tipo de funciones como la curva C de la figura 4.3. Imagínese, por ejemplo, que del análisis de un caso real de control de inventario se desprende que la política de pedidos es una función no lineal y que se puede definir mediante la tabla de valores que aparece en la tabla 4.2 que definiría una curva del tipo de las que aparecen en la figura 4.3.

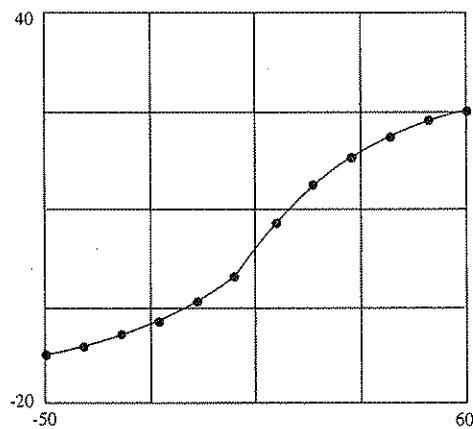


FIGURA 4.5. *Política no lineal de pedidos .*

Esta función se representa en la figura 4.5 y una posible explicación de su forma es la siguiente (no se debe olvidar que se trata de un ejemplo y que, en cualquier caso, esta información la proporcionaría un experto en el sistema que se pretende modelar): por un lado, la curva no es simétrica y presenta valores mayores para discrepancias positivas que para negativas, es decir, el almacenista tiene una mayor facilidad para pedir nuevas unidades

(debidas a un inventario deficiente) que para devolverlas (al tener un inventario excesivo); por otro lado, la pendiente de la curva va disminuyendo en valor absoluto a medida que se aleja del origen, lo que puede ser el reflejo del temor a pedir o devolver un gran número de unidades.

En los lenguajes de simulación existen sentencias específicas para definir funciones del tipo de la que se está tratando. Frecuentemente aportan herramientas gráficas para que esta definición sea más cómoda.

El resultado de la simulación correspondiente a este nuevo caso aparece en la figura 4.6.

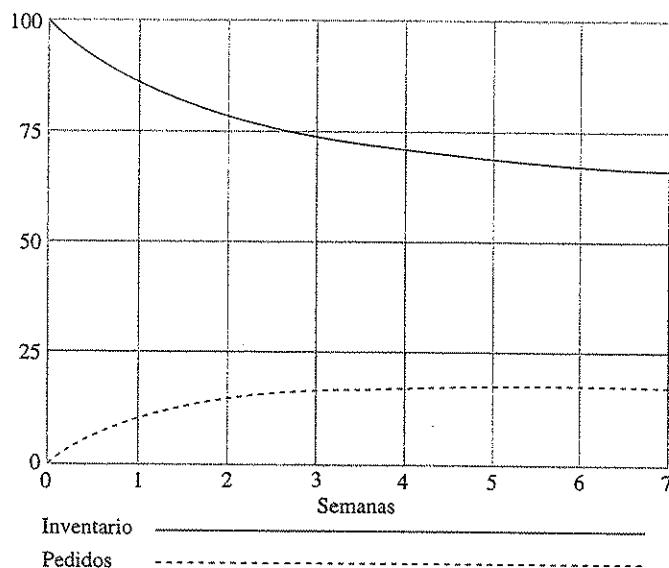


FIGURA 4.6. *Evolución del inventario y del flujo de pedidos para el caso de política no lineal de pedidos.*

Como colofón de este ejemplo, merece la pena detenerse en un aspecto que quizás haya extrañado al lector. Se trata del hecho de que se está tratando un sistema discreto —es decir, los pedidos se hacen en distintos instantes de tiempo, separados entre sí y, probablemente, no podrán ser de cualquier cantidad, sino cantidades enteras— modelándolo como un sistema continuo. Sin entrar de lleno en este problema se puede decir que esta aproximación será válida o no dependiendo de las magnitudes de las variables con las que se está tratando. Así, por ejemplo, si el tiempo entre pedidos es de un día y se pretenden realizar simulaciones de varios meses, se puede suponer, como una aproximación, que el tiempo es una variable continua.

4.3. BUCLE DE REALIMENTACIÓN POSITIVA

Para ilustrar un ejemplo de bucle de realimentación positiva se puede acudir al ya mencionado de crecimiento de una población que se encuentra en un medio sin ninguna limitación. En este caso, y asumiendo ciertas hipótesis simplificadoras, se puede suponer que la variación

de la población es proporcional al número de individuos, ya que tanto el número de nacimientos como el de defunciones serán proporcionales a la población. Si se supone que la tasa de natalidad supera a la de mortalidad se tendrá una tasa de crecimiento positiva y, de acuerdo con las ideas expresadas anteriormente, se tendrá un crecimiento exponencial.

Los diagramas de influencias y de Forrester de este sistema se representan en las figuras 4.7 y 4.8, respectivamente.

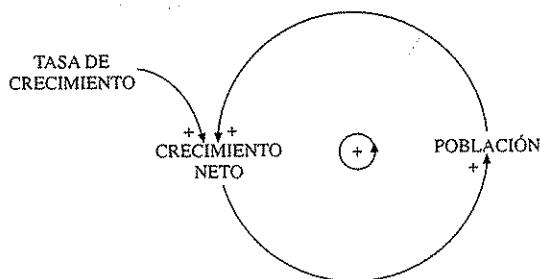


FIGURA 4.7. Diagrama de influencias de un modelo demográfico elemental.

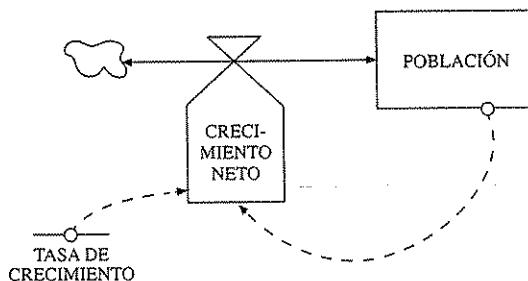


FIGURA 4.8. Diagrama de Forrester de un modelo demográfico elemental.

Supóngase que se pretende modelar la evolución de la población mundial desde el año 1850. Se conoce que este año la población mundial era aproximadamente de 1.250 millones de personas y se supone que la tasa neta de crecimiento es igual a 0,009 (que corresponde a un tiempo de duplicación de 77 años). A continuación se exponen las ecuaciones que definen este modelo. El significado de las variables se muestra en la tabla 4.3.

Nombre	Significado
CREC	Crecimiento neto
TCRE	Tasa de crecimiento
POB	Población

TABLA 4.3. Significado de las variables del modelo demográfico elemental.

$$\begin{aligned}
 \frac{d(\text{POB})}{dt} &= \text{CREC} \\
 \text{POB}(t_0) &= 1250 \times 10^6 \\
 \text{CREC} &= \text{TCRE} \times \text{POB} \\
 \text{TCRE} &= 0,009
 \end{aligned}$$

Las dos primeras ecuaciones definen la variable de estado POB. La primera determina su dinámica diciendo que el flujo que alimenta esta variable es CREC, mientras que la segunda establece su condición inicial. Estas ecuaciones se deben traducir al lenguaje adoptado siguiendo las pautas indicadas en la sección 3.5. Las dos últimas ecuaciones definen el flujo CREC utilizando la tasa de crecimiento TCRE, que es constante.

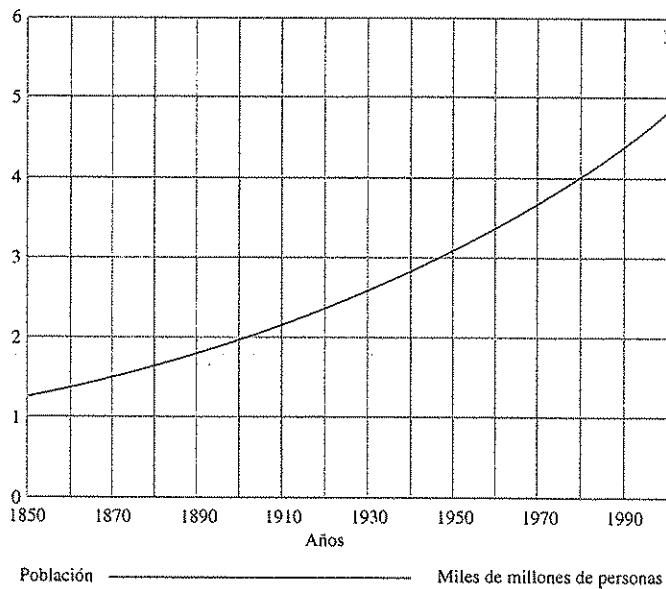


FIGURA 4.9. *Evolución de la población.*

La evolución de la población según este modelo aparece en la figura 4.9, en la que se observa el comportamiento explosivo de esta variable. Por supuesto, esta evolución no se corresponde con la que realmente ha tenido la población mundial. Esto se debe a que en el modelo se han realizado una serie de hipótesis que lo hacen extremadamente simple, pero lo alejan de la realidad. En primer lugar se ha supuesto que la tasa de crecimiento TCRE es constante. En segundo lugar debe recordarse el comentario realizado en la sección 2.2.3, respecto a la inexistencia de los comportamientos explosivos en la naturaleza. El que aquí aparece se debe a que no se ha incluido ningún proceso limitador como podría ser el agotamiento de los recursos naturales (merma de la capacidad de sustentación).

4.4. CRECIMIENTO SIGMOIDAL

La propagación de enfermedades infecciosas puede presentar un crecimiento sigmoidal. En este apartado se va a modelar la difusión de una epidemia. Supónganse las siguientes hipótesis:

1. La población es constante, es decir, el saldo de nacimientos, defunciones y fenómenos migratorios es nulo.
2. La enfermedad es lo suficientemente suave como para que los enfermos no dejen de hacer una vida normal, y éstos no se curan completamente durante el periodo de la epidemia; con ello se evita la reinfección.
3. La población enferma y la sana se encuentran homogéneamente mezcladas.

Este tipo de circunstancias se dan, por ejemplo, en las epidemias de catarros, gripes, etcétera.

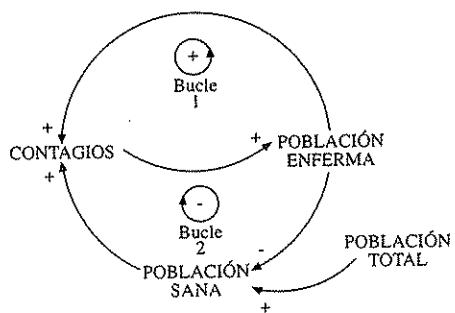


FIGURA 4.10. Diagrama causal del proceso de difusión de una epidemia.

En la figura 4.10 se muestra un diagrama causal del proceso epidémico. El flujo (tasa) de contagio depende a la vez de la población enferma y de la población sana, que es la diferencia entre la población total y la población enferma. El bucle 1 es un bucle de realimentación positiva, que corresponde al crecimiento de la población enferma en el seno de la población total. Si la población total fuese infinita, el crecimiento de la epidemia sería exponencial. El bucle 2 es un bucle de realimentación negativa, determinado por el hecho de que la población sea finita. Al aumentar la población enferma, la población sana, que es la única susceptible de enfermar, disminuye; lo cual conduce a una disminución del flujo de contagio. El flujo de contagio será nulo cuando toda la población esté enferma. El diagrama de Forrester de la figura 4.11 constituye un modelo simple del proceso epidémico. Las ecuaciones correspondientes son las siguientes:

$$\begin{aligned} \frac{d(PE)}{dt} &= FC \\ FC &= CPI \times PCD \times PS \times PE \\ PS &= P - PE \end{aligned}$$

92 *Dinámica de sistemas*

En donde PE denota la población enferma, FC el flujo de contagio, PCD el porcentaje de contactos diarios, CPI los contactos que producen infección por día, PS la población sana y P la población total.

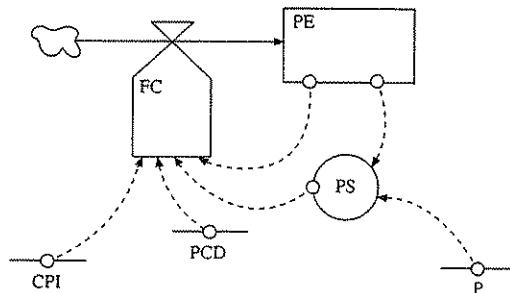


FIGURA 4.11. *Diagrama de Forrester del modelo de un proceso de difusión de una epidemia.*

La ecuación de flujo se interpreta de la manera siguiente. El producto de la población enferma PE por la población sana PS suministra el número total de posibles contactos que pueden darse. De este número total, por consideraciones estadísticas, se concluye que sólo se da en la realidad una cierta fracción que viene dada por la constante PCD, porcentaje de contactos diarios. El producto de los tres términos da el total de contactos que ocurren cada día. Una fracción de estos contactos dará lugar a la infección; esta fracción está definida por la constante CPI (contactos que producen infección por día).

La relación que liga al flujo de contagio FC con la población infectada PE, resulta ser

$$FC = PCD \times CPI \times PE \times [P - PE]$$

Esta expresión puede representarse gráficamente como aparece en la figura 4.12. De la observación de esta relación entre el flujo y el estado se concluye que la difusión de una epidemia se hace de acuerdo con una curva sigmoidal. En efecto, si se simula el modelo se obtiene el resultado de la figura 4.13 en la que se observa cómo la población infectada aumenta con un crecimiento sigmoidal. También se ha representado en esta figura la evolución del flujo de contagio.

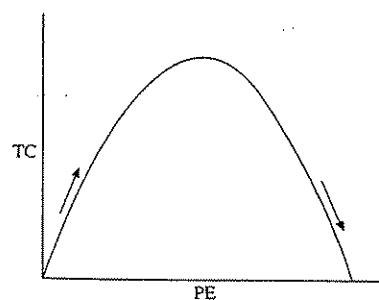


FIGURA 4.12. *Relación entre el flujo y el nivel en el modelo de difusión de una epidemia.*

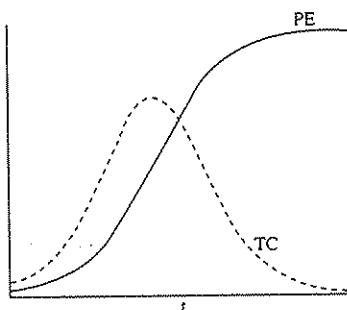


FIGURA 4.13. Evolución de la población enferma PE y de la tasa de contagio TC.

En general, los procesos epidémicos y otros similares, como la difusión de rumores, que se inician exponencialmente y a la vez “consumen” una cierta población, tienen un crecimiento sigmoidal tal como el considerado en este ejemplo.

4.5. OSCILACIONES

Como se comentó en el capítulo 1, la motivación de Forrester para estudiar los problemas de la dinámica industrial fue la observación de las oscilaciones que mostraban los pedidos de una empresa. Sin embargo, en los ejemplos presentados hasta ahora no aparece ningún tipo de oscilación. La explicación es que todos los modelos expuestos son de primer orden (solamente tienen una variable de nivel o estado) y en este tipo de sistemas, como se razonará a continuación, no pueden aparecer oscilaciones. Éstas sí se presentan en sistemas de orden igual o superior a 2, aunque éstos también pueden seguir un comportamiento similar al de los sistemas anteriores, sin ningún tipo de oscilación.

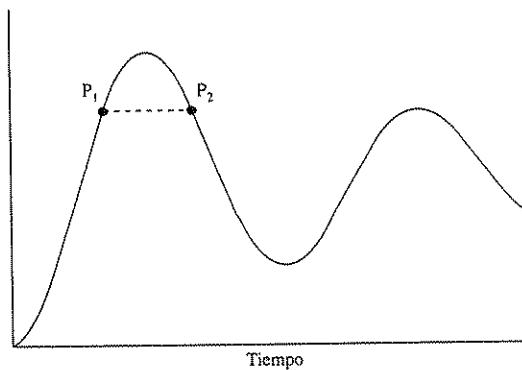


FIGURA 4.14. Comportamiento oscilatorio.

La explicación intuitiva por la que, en un sistema de primer orden, no pueden aparecer oscilaciones es la siguiente. Supóngase que en un sistema de primer orden la variable de nivel

presenta un comportamiento oscilatorio, tal como aparece en la figura 4.14. Si esto es así, se tendrían pares de puntos como los P_1 y P_2 de la figura 4.14 en los que el valor de la variable de estado es idéntico. Sin embargo, en P_1 el flujo neto es positivo, ya que la curva es creciente —la variable de estado crece con el tiempo— mientras que en P_2 es negativo. Como los flujos solamente pueden depender de las variables de estado y , en este caso, solamente existe una de ellas, el flujo vendrá únicamente determinado por el valor de la variable de estado. Dado que P_1 y P_2 corresponden al mismo valor de esa variable el flujo neto tiene que tomar el mismo valor en dichos puntos. Por lo tanto, es imposible que se dé la situación de la figura 4.14 .

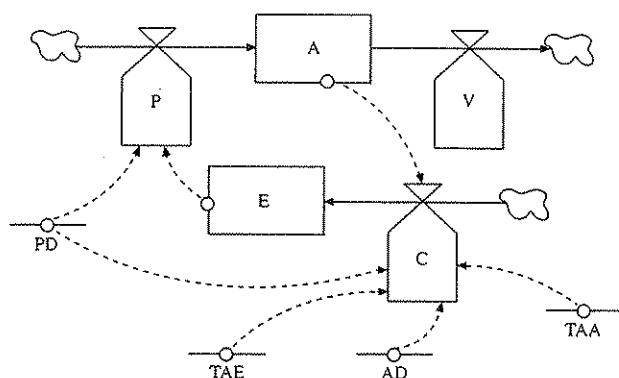


FIGURA 4.15. Diagrama de Forrester de un sistema de segundo orden con comportamiento oscilatorio.

Para ilustrar un ejemplo de un sistema de orden superior a uno con oscilaciones, considérese un modelo elemental de las interacciones que se producen entre el empleo y el almacén en una empresa. En la figura 4.15 se representa el diagrama de Forrester de este modelo. El significado de los nombres de las variables aparece en la tabla 4.4.

Nombre	Significado
A	Almacén
P	Producción
V	Ventas
E	Empleo
C	Contratación
PD	Productividad
AD	Almacén deseado
TAA	Tiempo de ajuste del almacén
TAE	Tiempo de ajuste del empleo

TABLA 4.4. Significado de las variables del ejemplo de sistema con comportamiento oscilatorio.

Unidad didáctica 4. Modelos dinámicos de sistemas económicos. Ejemplos elementales de modelos

La variable de estado almacén A representa el número de bienes almacenados por la empresa. Este estado se alimenta por medio de la variable de flujo producción P y se vacía por la variable de flujo ventas V. La producción P se determina a partir del empleo E y de la productividad PD; la variable E representa el número de empleados de la empresa y constituye la segunda variable de estado del modelo. Ésta se nutre del flujo contratación C.

La determinación de la variable de flujo C, es decir, el establecimiento de la política de contratación por parte de la empresa, constituye uno de los puntos cruciales del modelo. Se pueden considerar distintas alternativas para el establecimiento de la política de contratación. Se va a contemplar una especialmente simple que se basa en suponer que la empresa pretende ajustar el nivel de bienes en el almacén a un cierto valor que se representa por la variable almacén deseado AD. Se define la discrepancia en el almacén como la diferencia entre las variables que representan el almacén deseado y el almacén realmente existente, es decir, (AD - A). Si se divide la discrepancia del almacén (AD - A) por el tiempo de ajuste del almacén TAA se obtiene la producción deseada para corregir la discrepancia del almacén.

La corrección de esta discrepancia determina una contratación (o descontratación) que viene dada por el cociente (AD - A)/TAA dividido por la productividad por empleado PD. Si se divide el empleo necesario para corregir la discrepancia del almacén, tal como se acaba de calcular, por el tiempo de ajuste del empleo TAE se obtiene el flujo de contratación.

$$C(t) = \frac{AD - A(t)}{TAE \times TAA \times PD} \quad (4.1)$$

El conjunto de las ecuaciones del modelo resultan ser las siguientes:

$$\frac{d(A)}{dt} = P - V \quad (4.2)$$

$$P = E \times PD \quad (4.3)$$

$$\frac{d(E)}{dt} = C \quad (4.4)$$

$$C = (AD - A)/(TAE \times TAA \times PD) \quad (4.5)$$

La característica esencial del comportamiento de este modelo consiste en una serie de oscilaciones mantenidas con independencia de los valores tomados por las distintas constantes que aparecen en el modelo, tal y como se muestra en la figura 4.16.

Este hecho se puede demostrar de forma muy sencilla recurriendo a consideraciones de tipo analítico. Aquí se va a prescindir de estas consideraciones y en su lugar se van a hacer otras de tipo intuitivo y no formal, que si bien son más laboriosas y largas, tienen la ventaja de contribuir a una mejor comprensión intuitiva del comportamiento real del proceso.

Supóngase que el modelo de interacciones entre el almacén y el empleo que se muestra en la figura 4.17 se encuentra inicialmente en equilibrio; es decir, la producción es igual a las ventas y el almacén de bienes es justamente el deseado. Supóngase ahora que en el instante t_1 , las ventas aumentan instantáneamente a un nuevo valor que se supone que permanece constante.

Al ocurrir este incremento de ventas, se producirá un desequilibrio entre las ventas y la producción, inicialmente iguales entre sí. Como consecuencia de este desequilibrio se producirá

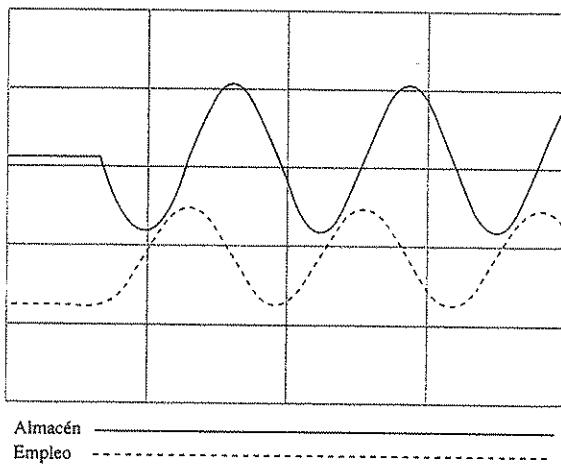


FIGURA 4.16. *Evolución del almacén y el empleo.*

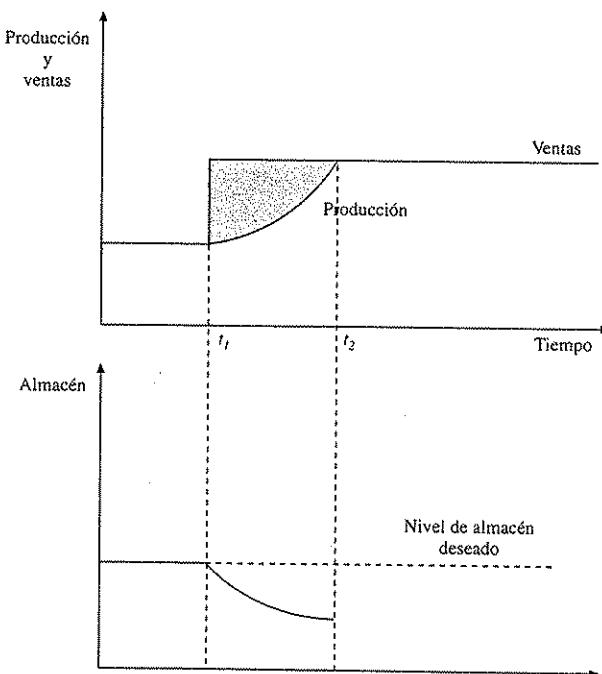


FIGURA 4.17. *Desajuste en la producción y las ventas*

una disminución de los bienes almacenados, ya que las ventas exceden a la producción. La disminución de los bienes almacenados, es decir, el valor de la variable A, por debajo del valor deseado, es decir, el valor de la variable AD, determinará de acuerdo con la expresión 4.1 una contratación de nuevos empleados con los que aumentar la producción para llevar de nuevo el nivel de bienes almacenados al valor deseado. Por tanto, el empleo y la producción —que es proporcional al empleo— crecen gradualmente; en la figura 4.17 se representa el crecimiento de la producción.

A pesar del crecimiento de la producción que se acaba de indicar, el nivel de bienes almacenados continuará decreciendo mientras que la producción se mantenga por debajo de las ventas. Sin embargo, esta disminución se va atenuando tal como indica la figura 4.17. Es decir, según la producción se aproxime al nuevo valor de las ventas, la disminución del almacén se hace más lenta. Por el contrario, la contratación, que es proporcional a la discrepancia entre el almacén real y el deseado, va aumentando con el tiempo. Al alcanzar el instante t_2 , representado en la figura 4.17, la producción se hace igual a las ventas, por lo que el número de bienes en el almacén deja de disminuir. Es decir, en la representación gráfica de la evolución en el tiempo del almacén de la figura 4.17 se alcanza un mínimo; sin embargo, el almacén se encuentra en un valor sensiblemente inferior al deseado; existe una discrepancia entre el almacén real y el deseado que determina una política de contratación para aumentar la producción y llevar el almacén al nivel deseado. Resumiendo el párrafo anterior cabe decir

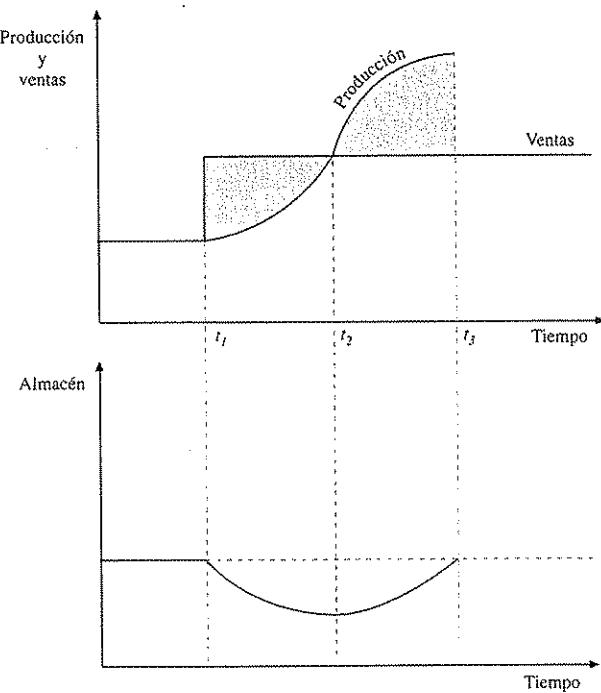


FIGURA 4.18. El ajuste no se produce y se inicia una oscilación.

que, como consecuencia de un aumento de las ventas, se ha producido una merma importante de los bienes almacenados durante el tiempo necesario para reajustar la producción al nuevo valor de las ventas. La superficie sombreada de la figura 4.17 representa la disminución del almacén entre los instantes t_1 y t_2 . En el instante t_2 se mantiene la política de contratación de empleados para aumentar la producción de manera que se recupere la pérdida de nivel en el almacén. Este aumento de la contratación determina que la producción sea superior a las ventas, y como consecuencia, el nivel del almacén empieza a recuperarse a partir de t_2 , tal como se indica en la figura 4.18. El nivel del almacén aumenta suavemente cuando la producción es ligeramente superior a las ventas; pero cuando la producción se hace sensiblemente superior a las mismas, el nivel del almacén crece rápidamente; por el contrario, según el nivel del almacén se vaya acercando al valor deseado, el empleo y consecuentemente la producción, crecen más lentamente; en la figura 4.18 se representa este proceso. Al alcanzarse el instante t_3 , el nivel del almacén alcanza el valor deseado; en dicho instante, al alcanzar la discrepancia entre los niveles real y deseado del almacén el valor nulo, se deja de contratar nuevos empleados y finaliza el proceso de crecimiento de la producción tal como se indica en la figura 4.18.

En el instante t_3 el nivel de bienes almacenados es el deseado; sin embargo, la producción excede sensiblemente a las ventas. La producción ha debido crecer por encima de las ventas para que el almacén recuperase el nivel perdido entre t_1 y t_2 . Este aumento de la producción sobre las ventas debe ser tal que el área sombreada entre t_2 y t_3 de la figura 4.18 (que representa el incremento del almacén en este intervalo) sea igual al área sombreada entre t_1 y t_2 .

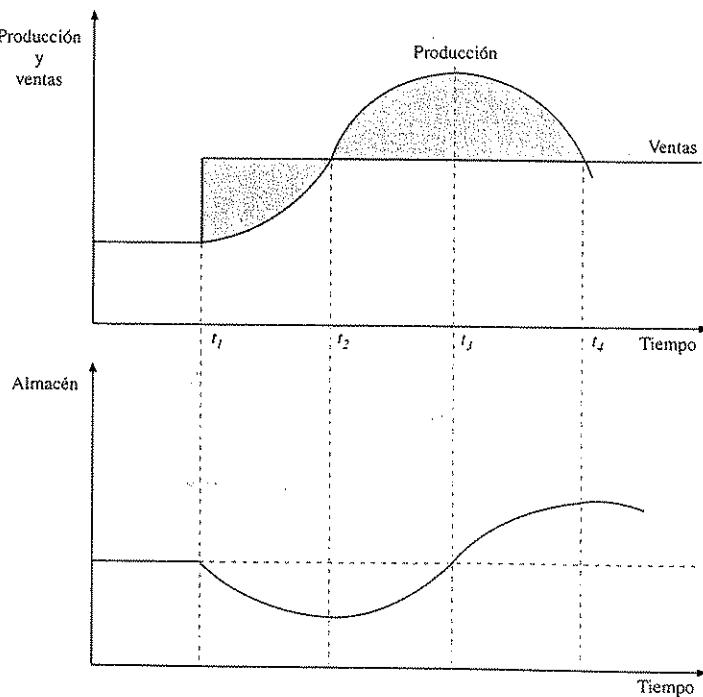


FIGURA 4.19. *La oscilación tiende a mantenerse.*

Algunas curvas de oscilación representan comportamientos que se extinguirán, otros permanecen estables y otros crecen.

A partir del instante t_3 , puesto que la producción excede a las ventas, el nivel del almacén tenderá a crecer por encima del valor deseado. Al crearse un exceso de bienes en el almacén se tenderá a disminuir la producción, lo que en este modelo simplificado equivale al despido de empleados. La producción y el empleo disminuirán hasta que el nivel de bienes almacenados sea de nuevo el deseado.

En la figura 4.19 se representa la evolución de la producción hasta el instante t_4 , en que nuevamente la producción y las ventas se equilibra. Puede demostrarse que la curva de producción de la figura 4.19 entre t_2 y t_3 es simétrica a la de t_3 a t_4 . De este modo se completa la ilustración de cómo se genera un comportamiento oscilatorio.

Las oscilaciones que han aparecido en este sistema se llaman *mantenidas* porque la magnitud de la amplitud de las mismas se mantiene a lo largo del tiempo. Sin embargo, éste no es el único caso que se puede dar: en otros sistemas la amplitud puede amortiguarse o crecer tal y como aparece en la figura 4.20.

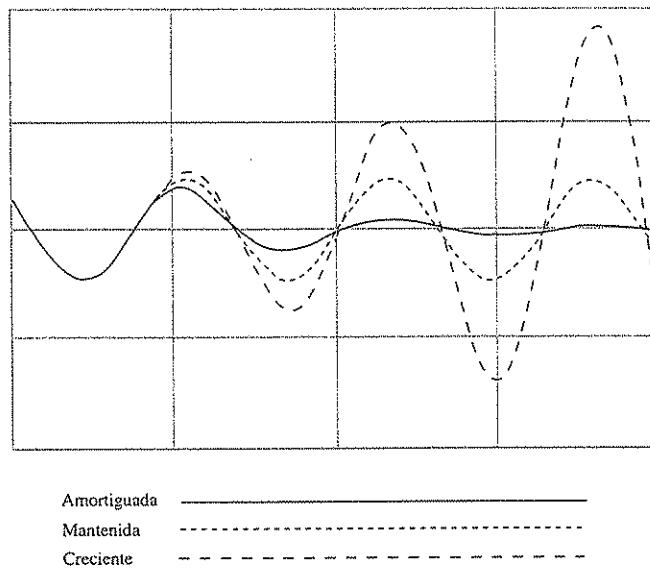


FIGURA 4.20. Tipos de oscilaciones.

4.6. ADICCIÓN

A continuación retomaremos el ejemplo expuesto en la sección 2.3.3 referente al arquetipo de la adicción. El ejemplo considera las situaciones en las que el exceso de trabajo conduce al estrés y éste puede llevar a la utilización de fármacos. Se trata ahora de plantear las ecuaciones que modelan esta situación y comprobar, mediante simulación, que aparecen los comportamientos comentados anteriormente.

Para ello supondremos que se puede medir y expresar mediante números tanto la cantidad de trabajo que se realiza TRAB, como la cantidad de fármacos que se ingiere FARM así como el estrés ESTRÉS.

En primer lugar se tratará el sistema representado mediante el diagrama de influencias de la figura 4.21. En él aparece que el trabajo produce estrés y que éste lleva al uso de fármacos que hacen disminuir el estrés. Sin embargo, los fármacos producen un aumento de la percepción de la capacidad de aceptar más trabajo (capacidad subjetiva de aumento del trabajo CSAT) que lleva a un aumento del trabajo (que se va a denominar VT, variación del trabajo).

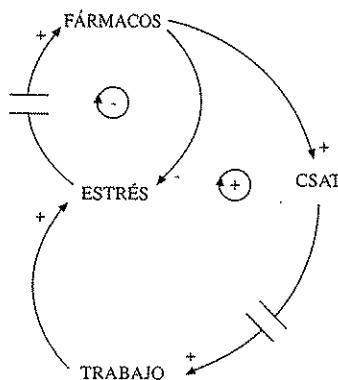


FIGURA 4.21. *Diagrama de influencias de un ejemplo de adicción.*

Supóngase que la relación de influencias del trabajo y los fármacos sobre el estrés se puede expresar mediante la ecuación:

$$\text{ESTRÉS} = \text{MAX}(\text{TRAB} - 5 - 2, 5 \times \text{FARM}, 0) \quad (4.6)$$

en donde MAX representa la función máximo incorporada en la mayoría de los lenguajes de simulación:

$$\text{MAX}(a, b) = \begin{cases} a & \text{si } a \geq b \\ b & \text{si } a \leq b \end{cases}$$

La ecuación (4.6) obliga a que ESTRÉS sea siempre mayor o igual que cero y que éste aumente al crecer TRAB y decrezca al aumentar FARM. Es fácil comprobar que en esta ecuación, el valor 5 es el máximo trabajo que se puede asumir sin que aparezca estrés (en ausencia de fármacos). Se han supuesto relaciones lineales aunque podría utilizarse otro tipo de funciones.

El resto de las ecuaciones que se pueden adoptar es fácil de entender:

$$\begin{aligned} \frac{d(\text{TRAB})}{dt} &= \text{VT} \\ \text{VT} &= \text{DELAY}(\text{CSAT}, 3) \\ \text{CSAT} &= \text{FARM} \times 2 \\ \text{FARM} &= \text{KFARM} \times \text{DELAY}(\text{ESTRÉS}, 1) \\ \text{KFARM} &= 0.4 \end{aligned}$$

En estas ecuaciones se ha utilizado la función `DELAY` que representa un retraso de primer orden y que está incorporada en la mayoría de los lenguajes específicos de dinámica de sistemas (aunque en éstos se denomina a veces con otros nombres y se suele diferenciar entre retrasos de material y de información). El primer argumento de la función `DELAY` aquí utilizada corresponde a la entrada del retraso, mientras que el segundo indica el tiempo de retraso. La estructura de estas ecuaciones y el valor de los parámetros que en ellas aparecen se obtendrían estudiando un caso real y planteando nuevas hipótesis, con el concurso de especialistas.

Dando un valor inicial igual a 10 para la variable de nivel `TRAB`, el modelo queda totalmente definido. En la figura 4.22 aparece la evolución de las variables `ESTRÉS`, `TRAB` y `FARM`. Se puede observar el aumento continuado del uso de los fármacos y que, tras una leve mejoría del estrés, éste no hace sino aumentar.

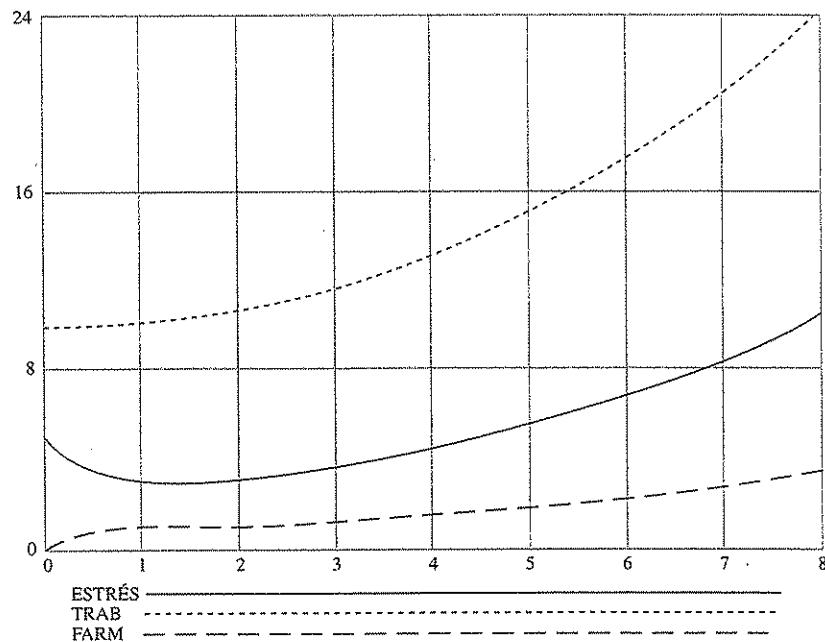
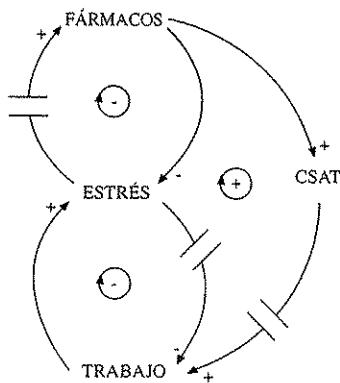


FIGURA 4.22. Evolución de las variables del ejemplo de la adicción.

Como se comentó en la sección 2.3.3, el problema es que se pretenden resolver los síntomas (el estrés) de un problema de fondo (el exceso de trabajo), sin intentar soslayar este último. Supóngase ahora que, en el instante 8 de la simulación anterior, se decide atacar la raíz del problema. En este caso el diagrama de influencias se modifica para llegar al de la figura 4.23 en el que se ha introducido una nueva influencia: la que va desde `ESTRÉS` a `VT`: un aumento del estrés se intentará corregir con una variación negativa del trabajo asumido.

Como otra medida de corrección del efecto indeseado se pretende disminuir la ingesta de fármacos disminuyendo la constante `KFARM` de 0,4 a 0,2 en $t = 8$.

Para plantear las nuevas ecuaciones del modelo, dado que éste varía con el tiempo, es conveniente emplear dos nuevas funciones que están incorporadas en la mayoría de los lenguajes

FIGURA 4.23. *Diagrama de influencias modificado del ejemplo de la adicción.*

de simulación: la función IF y la función STEP. Se va a utilizar la función IF con tres parámetros de manera que, por ejemplo, la expresión $a = \text{IF}(b > c, d, e)$ va a significar:

$$a = \begin{cases} d & \text{si } b > c \\ e & \text{si } b \leq c \end{cases}$$

La función STEP tiene dos parámetros con el significado siguiente $\text{STEP}(a, b)$ es equivalente a $\text{IF}(t \geq b, a, 0)$, siendo t el tiempo.

La sintaxis de estas dos funciones (e incluso el nombre) en los diversos lenguajes de simulación puede ser distinta a la expuesta.

Con estas dos funciones el modelo resulta:

$$\begin{aligned} \text{ESTRÉS} &= \text{MAX}(\text{TRAB} - 5 - 2,5 \times \text{FARM} > 0, 0) \\ \frac{d(\text{TRAB})}{dt} &= \text{VT} \\ \text{VT} &= \text{DELAY}(\text{CSAT}, 3) - \text{KEST} \times \text{ESTRÉS} \\ \text{CSAT} &= \text{FARM} \times 2 \\ \text{FARM} &= \text{KFARM} \times \text{DELAY}(\text{ESTRÉS}, 1) \\ \text{KFARM} &= \text{IF}(t < 8, 0, 4, 0, 2) \\ \text{KEST} &= \text{STEP}(2, 8) \end{aligned}$$

La evolución de las variables aparece en la figura 4.24. Se observa que en el instante en el que se decide cambiar la política se produce un aumento del estrés, consecuencia de la disminución de la constante KFARM, y por tanto de los fármacos ingeridos. Sin embargo, a la larga, el estrés y la ingesta de fármacos disminuyen (y tienden a anularse), y el trabajo asumido tiende al valor máximo admisible.

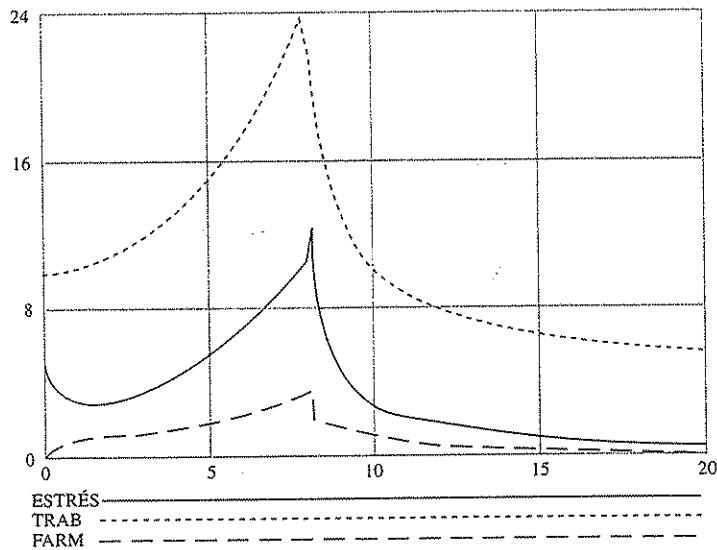


FIGURA 4.24. Evolución de las variables en el ejemplo de la adicción modificada.

4.7. INVERSIÓN INSUFICIENTE

En esta sección realizaremos un modelo del caso expuesto en la sección 2.3.4, donde se trataba el problema de la bajada de la demanda de un producto al aumentar el plazo de entrega del mismo.

En la figura 4.25 se muestra el diagrama de Forrester del sistema. En él se puede observar que las órdenes pendientes ORD corresponde a un nivel que se llena con las nuevas órdenes NUEORD y se vacía con las entregas ENT. La capacidad de producción CAP se alimenta de las nuevas inversiones INV. La demanda DEM se llena gracias al marketing MARK y se vacía con la disipación de la demanda DISIP debida al posible descontento con el plazo de entrega. La disipación de la demanda depende del descontento DESC producido por tiempos de ciclo elevados.

La mayor parte de las ecuaciones de este modelo se formulan inmediatamente a partir del diagrama de Forrester:

$$\begin{aligned}
 \frac{d(CAP)}{dt} &= INV \\
 \frac{d(ORD)}{dt} &= NUEORD - ENT \\
 \frac{d(DEM)}{dt} &= MARK - DISIP \\
 NUEORD &= DEM \\
 ENT &= \text{MIN}(CAP, ORD).
 \end{aligned}$$

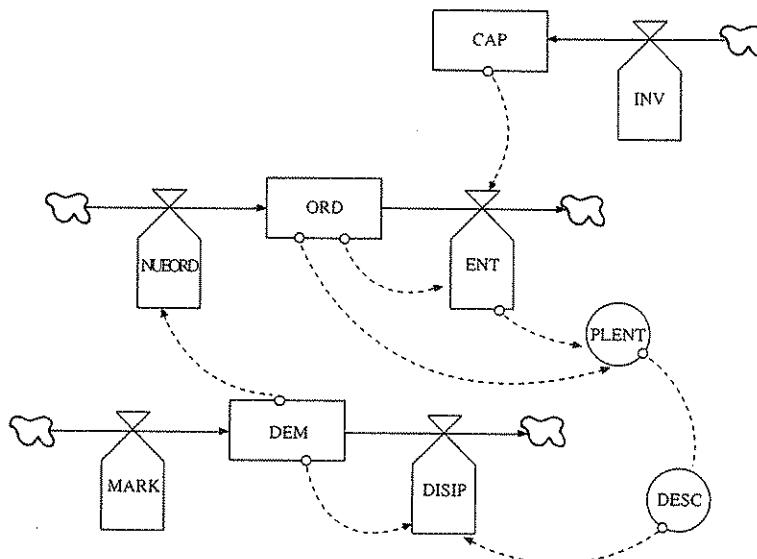


FIGURA 4.25. Diagrama de Forrester del ejemplo de inversión insuficiente.

Para estimar el descontento es preciso evaluar el plazo de entrega. Si en una unidad de tiempo se suministran ENT unidades de las órdenes (ORD) existentes, el plazo medio de entrega (PLENT) vendrá dado por la ecuación:

$$\text{PLENT} = \frac{\text{ORD}}{\text{ENT}}$$

Supóngase que se conoce una medida del descontento producido por distintos plazos de entrega, como la que aparece en la figura 4.26. Como se ha comentado anteriormente, es posible incluir este tipo de relaciones en un modelo programado con los lenguajes de simulación más usuales.

El descontento influirá en la disipación de la demanda. Supóngase una relación lineal entre estas variables:

$$\text{DISIP} = (0,01 + \text{DESC}) \times \text{DEM}$$

Con esta ecuación, además, se supone que existirá una pequeña disipación de la demanda aunque el descontento sea nulo.

Para acabar de definir el modelo es necesario especificar el comportamiento de los gestores de la empresa, es decir, definir la política de inversiones y la de marketing. Supóngase que se opta por una política de marketing que da lugar al bucle de realimentación positiva comentado en la sección 2.3.4, es decir, una relación entre la demanda y el marketing monótona creciente. Por ejemplo, se utilizará la de la figura 4.27. Esta curva se puede representar analíticamente mediante la ecuación:

$$\text{MARK} = 0,1 + \text{MAX}[(\text{DEM} - 2)/20, 0]$$

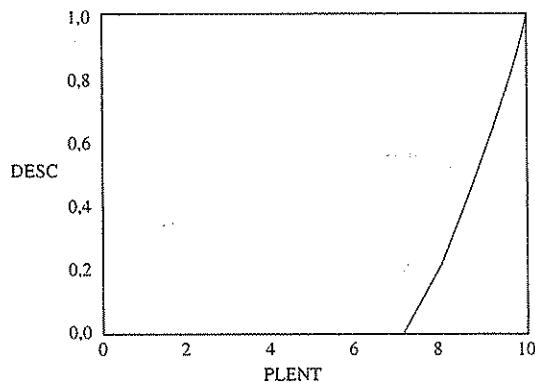


FIGURA 4.26. Relación entre el descontento (DESC) y el plazo de entrega (PLENT).

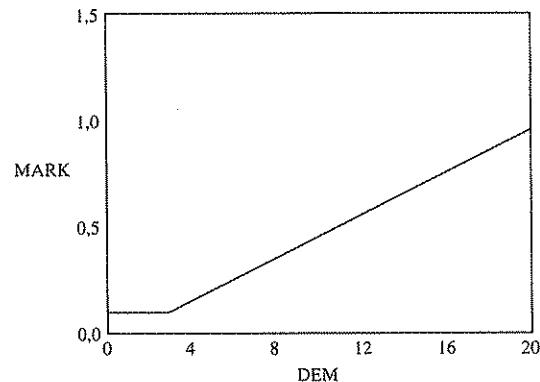


FIGURA 4.27. Política de marketing.

Supóngase ahora que se decide no invertir en el aumento de la capacidad de producción, es decir, $INV = 0$. Simulando el modelo se obtiene un comportamiento como el reflejado en la figura 4.28, que presenta las oscilaciones comentadas en la sección 2.3.4. Sin embargo, en este ejemplo sencillo no llega a producirse el colapso. Si, por el contrario, se opta por una política de inversiones que intente evitar los plazos de entrega elevados, se conseguirá un crecimiento sostenido de demanda, como el reflejado en la figura 4.29 obtenida con una política de inversiones dada por la ecuación:

$$INV = \text{MAX}(PLENT - 5, 0),$$

que supone que si el plazo de entrega es mayor que 5 se debe invertir proporcionalmente a esta cantidad.

106 *Dinámica de sistemas*

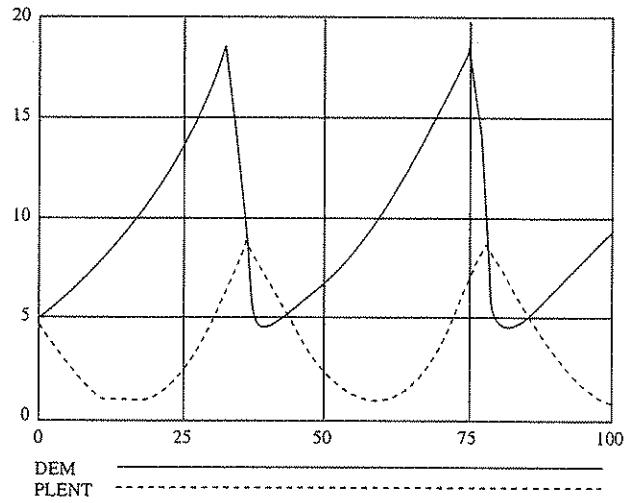


FIGURA 4.28. *Comportamiento en ausencia de inversiones.*

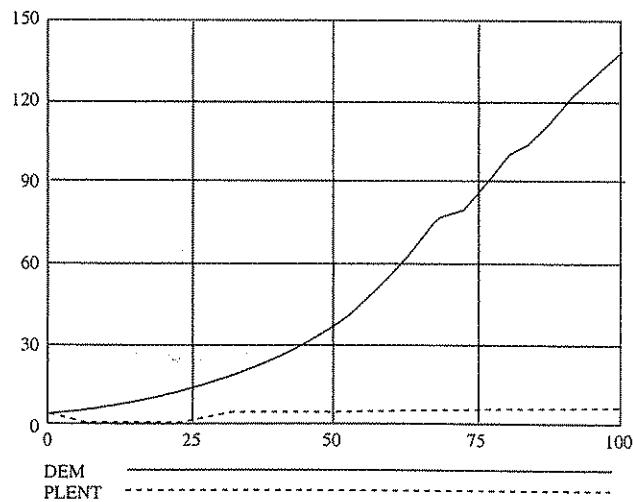


FIGURA 4.29. *Comportamiento con inversiones suficientes.*

5.1. INTRODUCCIÓN

Con el material contenido en los capítulos anteriores el lector se encuentra en disposición de comprender los elementos básicos que aparecen en un modelo de dinámica de sistemas, y aún de construir por sí mismo modelos elementales. Con estos conocimientos se puede abordar con provecho la lectura de obras dedicadas a describir modelos basados en dinámica de sistemas. Sin embargo, al interesado no sólo en la lectura de trabajos de otros, sino en la creación de modelos propios de carácter no elemental, se le plantean inmediatamente las cuestiones sobre cuál es la génesis de uno de estos modelos, cómo se llega a elegir las variables, a establecer las relaciones y a decidir todos los aspectos que conducen a un modelo acabado. Es decir, ¿cómo se consiguen captar los elementos esenciales para la reproducción de un comportamiento real complejo?

La respuesta a algunas de estas cuestiones ya ha sido insinuada, en forma dispersa, a lo largo de las páginas anteriores. Por otra parte, no debe perderse de vista que cualquier fase de construcción de un modelo de la naturaleza de los que aquí se estudian, en los que se trata de sintetizar en unas pocas variables el funcionamiento de un aspecto sumamente complejo de la realidad, está presidida, en gran parte, por la experiencia, la intuición y la inspiración. Sin embargo, es posible llegar a una cierta sistematización en las fases de la construcción de un modelo con ayuda de la dinámica de sistemas, y a ello se va a dedicar la sección 5.2. A continuación se ilustrarán estas fases por medio de un ejemplo (sección 5.3). El modelo que se va a utilizar es muy simple, al igual que los descritos hasta ahora, para no dificultar la exposición. Por supuesto, la aplicación de la dinámica de sistemas no se reduce a modelos elementales, sino que su interés radica en el desarrollo de modelos complejos. En la sección 5.4 se describe a grandes rasgos uno de estos modelos.

5.2. LAS FASES EN LA CONSTRUCCIÓN DE UN MODELO

De una forma general se puede decir que en el proceso de desarrollo de un modelo se incluyen tres fases principales:

- *conceptualización*, que consiste en la adopción de una perspectiva y en el esbozo de una comprensión de un cierto fenómeno del mundo real;
- *formulación del modelo*, que trata de la representación de los elementos intuitivos elaborados en la fase de conceptualización por medio de un lenguaje formal;
- *evaluación del modelo*, consistente en un análisis del mismo, así como su sometimiento a varios criterios de aceptabilidad.

En la parte izquierda de la figura 5.1 se muestran, de forma esquemática, las tres fases a las que se acaba de aludir. En la mitad derecha de esta figura se indica el carácter iterativo de la construcción de un modelo, en virtud del cual no se pasa de una forma progresiva y única por las tres fases indicadas, sino que se puede ir de una fase a otra, sin ningún orden especial, cuantas veces sea necesario. A continuación se van a describir con detalle cada una de las fases enunciadas.

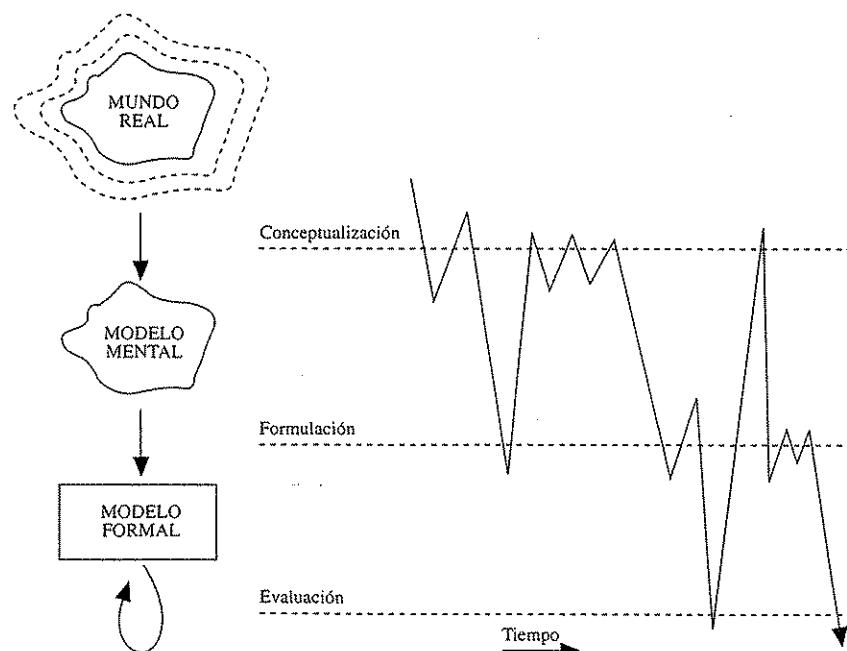


FIGURA 5.1. Carácter iterativo de las fases de construcción de un modelo.

5.2.1. Fase de conceptualización

La fase de conceptualización se inicia, normalmente, con una familiarización con el problema que se va a estudiar, que incluye el tratamiento de la literatura al respecto, opiniones de expertos, experiencias propias, etc.; en resumen, se trata de sumergirse en el problema en cuestión.

Tras esta familiarización, hay que definir con precisión los aspectos del problema que se quiere resolver y describirlos en forma precisa y clara. En toda la fase de conceptualización debe intentarse llegar al máximo de concisión, claridad y precisión.

A partir del conocimiento previo sobre estructuras simples de comportamiento de sistemas dinámicos se tratará de particularizar el comportamiento dinámico relevante del sistema bajo estudio, así como la estructura más simple que pueda generar este comportamiento.

De esa manera, y de una forma progresiva, se van identificando los distintos elementos que formarán el sistema, lo que conduce de modo natural al establecimiento de los límites del sistema y a una descripción primaria de los bucles de realimentación. Surge así el diagrama de influencias del sistema, con lo que se puede considerar finalizada la fase de conceptualización.

5.2.2. Fase de formulación

Después de construir el diagrama de influencias se procede a su formulación con ayuda de un lenguaje formal preciso. En dinámica de sistemas ello consiste en primer lugar en el establecimiento del diagrama de Forrester, a partir del cual se escriben las ecuaciones del modelo; éstas pueden expresarse en un lenguaje que permita su formulación informática.

En esta fase debe procederse a asignar valores a los parámetros que intervienen en el modelo. Se trata de un punto sumamente delicado e importante del que, en muchos casos, dependerá la utilidad que consiga. Sobre este punto se volverá más adelante al estudiar la estimación de parámetros.

La fase de formulación concluye cuando se dispone de un modelo del sistema bajo estudio en forma de ecuaciones programadas en un computador.

5.2.3. Fase de evaluación

Una vez construido el modelo se procede a ensayar, por medio de simulaciones, las hipótesis sobre las que se ha construido, así como la consistencia entre las mismas.

Un aspecto muy importante de esta fase es el análisis de sensibilidad del modelo, en virtud del cual se estudia la dependencia de las conclusiones que se extraen del modelo, con relación a posibles variaciones que sufran los valores de los parámetros que aparecen en él.

Cuando se consideran satisfactorios los análisis de consistencia de las hipótesis y los de sensibilidad, se procede a estudiar el comportamiento del modelo ante distintas políticas alternativas, con el fin de elaborar unas recomendaciones respecto a la actuación futura sobre la realidad.

5.3. ESTUDIO DE UNA APLICACIÓN DE LA DINÁMICA DE SISTEMAS

Conviene ilustrar las fases de construcción de un modelo con ayuda de la dinámica de sistemas por medio de una aplicación concreta. A ello se va a destinar la presente sección.

El proceso de construcción de un modelo reviste, en cada aplicación, características peculiares. Así, la utilización de las indicaciones desarrolladas en las secciones anteriores requiere una enorme capacidad de iniciativa y de imaginación, y en cada caso concreto se hará de forma distinta.

La aplicación que se presenta en esta sección pretende tener un carácter fundamentalmente ilustrativo y no puede, en ningún caso, mostrar la pauta de construcción de un modelo que sea literalmente transcribible a cualquier otra aplicación. Sin embargo, se considera de sumo interés el análisis detenido de este ejemplo del proceso de construcción, puesto que en él se muestran las líneas generales que lo presiden.

El problema que se trata de analizar consiste en el estudio del crecimiento económico de un área industrial. Supóngase un proceso de crecimiento al que se puede asociar una curva de crecimiento sigmoidal o logístico (sección 2.3.1). Ese tipo de proceso abunda en la realidad y se presenta cuando nos encontramos en las fases iniciales de un proceso que aparentemente tiende a crecer de forma ilimitada. Sin embargo, más pronto o más tarde aparecen factores limitativos que condicionan ese crecimiento (el infinito es una ficción matemática: en la práctica todo crecimiento acaba manifestando alguna forma de saturación). De hecho, la estructura básica de ese tipo de proceso ya se vio en la sección 2.3.1. Aquí, para concretar, lo referiremos al proceso de crecimiento de la actividad económica en un área industrial, mediante la creación de empresas; aunque igual pudiera haberse tratado del crecimiento de una población en un hábitat, de la introducción de un nuevo producto en un mercado, de la difusión de un rumor en un grupo social, de la propagación de una enfermedad contagiosa en una población previamente sana, etcétera. Este modelo está tomado del libro de Alfeld y Graham¹, donde se denomina BSNSS3.

5.3.1. Fase de conceptualización

Veamos una descripción elemental del sistema que se pretende modelar. Se supone un área de superficie limitada en la que se produce un proceso de establecimiento de empresas (por ejemplo, un polo de promoción industrial). Como consecuencia de ello el número de las empresas crece. Este crecimiento viene afectado por el grado de ocupación del área, de modo que cuando hay todavía pocas empresas es fácil el establecimiento de nuevas, mientras que cuando el área está prácticamente ocupada se producen fenómenos de saturación que tienden a limitar el crecimiento. Por otra parte, se supone que se produce un flujo, relativamente pequeño, de desaparición de empresas, por obsolescencia, quiebra o razones análogas. De acuerdo con esta descripción, la estructura básica del modelo está formada por los cinco enunciados siguientes:

- *R₁: La creación de empresas influye sobre el número de empresas.* Este enunciado es completamente natural: cuantas más empresas se crean, más empresas habrá.

¹ Alfeld y Graham, 1976, *Introduction to Urban Dynamics*, Wright-Allen Press.

- R_2 : El factor de ocupación del terreno influye sobre la creación de empresas. Este enunciado es importante y, como se ha indicado anteriormente, si el factor de ocupación es pequeño su influencia sobre la creación de empresas es creciente; mientras que si es grande es decreciente. Se trata, por tanto, de un enunciado que no tiene un signo perfectamente definido. Para valores pequeños del factor de ocupación la influencia implícita es positiva, mientras que para valores elevados es negativa.
- R_3 : El número de empresas influye sobre el factor de ocupación. Se trata de un enunciado que establece una relación positiva: a mayor número de empresas mayor factor de ocupación.
- R_4 : El número de empresas influye sobre su eventual desaparición. Se supone que desaparece una fracción, en principio muy pequeña, de las empresas establecidas.
- R_5 : La desaparición de empresas influye sobre su número. Se trata, como el enunciado R_1 , de una relación natural, pero de signo negativo.

Esta descripción verbal es la forma más simple que puede tomar el supuesto modelo mental que poseen los especialistas en el sistema que se trata de modelar. Si observamos los enunciados R_i veremos que en ellos se pueden distinguir cuatro variables: el número de empresas EMP, la creación de empresas CE, la desaparición de empresas DE y el factor de ocupación del terreno FO. Con ayuda de estas variables los anteriores enunciados se pueden reescribir en forma de relaciones de influencia, recordando la notación de (2.2) o de (2.3), como sigue:

$$\begin{aligned}
 R_1 & : \text{CE} \xrightarrow{+} \text{EMP} \\
 R_2 & : \text{FO} \xrightarrow{+/-} \text{CE} \\
 R_3 & : \text{EMP} \xrightarrow{+} \text{FO} \\
 R_4 & : \text{EMP} \xrightarrow{+} \text{DE} \\
 R_5 & : \text{DE} \xrightarrow{-} \text{EMP}
 \end{aligned}$$

A la relación R_2 se ha asociado $+/-$ para indicar que su signo no está únicamente definido.

Estas cinco relaciones de influencia constituyen una primera reelaboración de la *descripción verbal* del modelo, y suministran un primer bosquejo cualitativo de su estructura. Las relaciones R_i se pueden representar mediante un diagrama de influencias, tal como el que se indica en la figura 5.2. Este diagrama posee el considerable interés de su comunicabilidad a todos los implicados en el proceso de modelado. El paso siguiente en el proceso de modelado, como vamos a ver, consiste en elaborar estas relaciones hasta convertirlas en relaciones cuantitativas.

5.3.2. Fase de formulación

Para realizar el diagrama de Forrester es necesario clasificar adecuadamente las distintas variables en variables de estado, de flujo y auxiliares.

Las variables de estado constituyen, normalmente, aquel conjunto de variables cuya evolución es significativa para el estudio del sistema. Representan magnitudes que acumulan los resultados de acciones tomadas en el pasado. Un ligero análisis de la descripción verbal del modelo conduce a la conclusión de que únicamente la variable empresas EMP es una variable de estado (acumula las decisiones respecto a la creación y desaparición de empresas). Las

112 *Dinámica de sistemas*

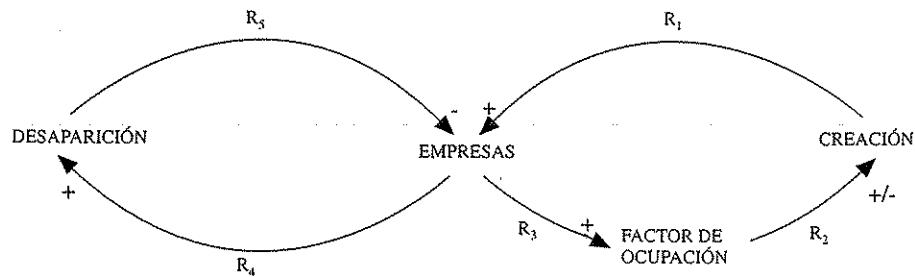


FIGURA 5.2. *Diagrama de influencias de un modelo elemental de evolución de las empresas en un área industrial.*

variables creación de empresas CE y desaparición de empresas DE son variables de flujo, que determinan el ritmo de variación de la variable de estado EMP. El resto de las variables son variables auxiliares, que representan pasos o etapas en que se descompone el cálculo de una variable de flujo a partir de los valores tomados por las variables de estado.

Con estas consideraciones se puede construir el diagrama de Forrester (figura 5.3) correspondiente al de influencias de la figura 5.2. En este diagrama se han empleado los iconos habituales para los distintos tipos de variables: rectángulos para las de estado, válvulas para las de flujo y círculos para las auxiliares.

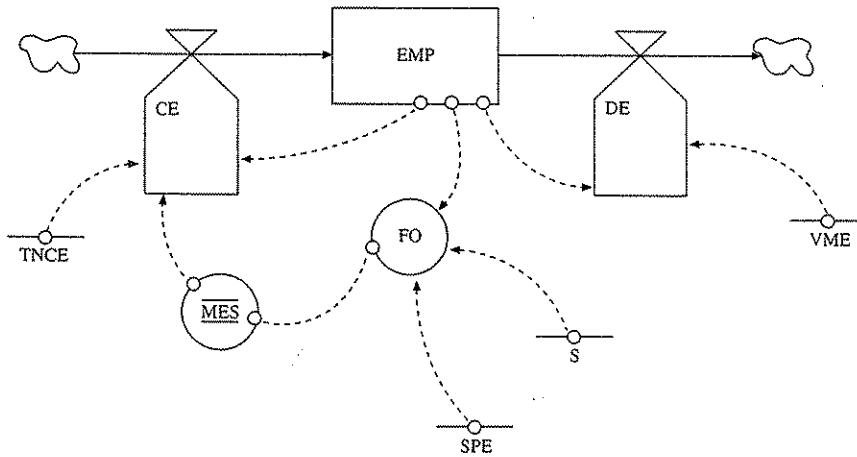


FIGURA 5.3. *Diagrama de Forrester del modelo de un área industrial.*

En el diagrama de la figura 5.3 se muestra cómo la variable de estado empresas EMP se nutre de la variable de flujo creación de empresas CE, y se vacía mediante la variable de flujo desaparición de empresas DE. Al mismo tiempo están los canales de información, que a lo largo del modelo representan las influencias entre las distintas variables. Se indican, así mismo, los parámetros del modelo: la tasa normal de creación de empresas TNCE, la superficie edifi-

cable S, la superficie media por empresa SPE y la vida media de las empresas VME. En ese diagrama es patente que las variables de estado y los flujos desempeñan un papel relevante, de tal forma que estos diagramas se han denominado también diagramas de *flujos y niveles* de la estructura de un sistema.

A partir del diagrama de Forrester es posible establecer las ecuaciones del modelo. Consideremos, en primer lugar, la evolución de la variable empresas EMP. Su crecimiento a lo largo del tiempo dependerá de la diferencia entre la creación de empresas CE y la desaparición de empresas DE, de acuerdo con la ecuación:

$$\frac{d(\text{EMP})}{dt} = \text{CE} - \text{DE} \quad (5.1)$$

Esta ecuación subsume las relaciones de influencia R_1 y R_5 que hemos visto anteriormente. La expresión (5.1) es un ejemplo de lo que se conoce como *ecuación de estado o de nivel* y expresa cómo evoluciona una variable de estado, en este caso la variable empresas EMP.

De acuerdo con las hipótesis utilizadas para construir el modelo —y que representarían la opinión de los especialistas con cuyo concurso se supone que se ha modelado el área industrial correspondiente— la desaparición de empresas DE se determina con ayuda de la expresión:

$$\text{DE} = \frac{\text{EMP}}{\text{VME}} \quad (5.2)$$

en donde VME representa la vida media de las empresas; es decir, en un determinado periodo de tiempo desaparecerán un número de empresas que vendrá determinado por su vida media. Esta expresión representa la relación R_4 .

La ecuación (5.2) es un ejemplo de lo que se denomina una *ecuación de flujo*, y representa cómo se determina una variable de flujo, en este caso DE. Esta ecuación se puede considerar como una razonable aproximación al comportamiento, en promedio, de una política de desaparición de empresas por obsolescencia. Se supone aquí que la vida media de las empresas es constante. Esta hipótesis podría relajarse, si se estimara conveniente, en una versión ulterior del modelo. Normalmente este efecto será muy pequeño, pero se ha considerado interesante el incluirlo para reforzar el interés didáctico del modelo.

Para la determinación del flujo creación de empresas CE se supone que el número de ellas que se establecen en un determinado periodo resulta ser una fracción de las ya establecidas. Es como si el propio éxito en la promoción del área atrajese nuevas empresas. Esa fracción viene afectada por un factor MES —multiplicador de empresas en función de la superficie edificada— que depende del factor de ocupación de la superficie del área industrial. La expresión que da el valor de CE es:

$$\text{CE} = \text{EMP} \times \text{TNCE} \times \text{MES} \quad (5.3)$$

en donde TNCE es la tasa normal de creación de empresas. Esta expresión suministra un nuevo ejemplo de una ecuación de flujo.

A su vez el multiplicador MES, según se ha indicado un poco más arriba, depende del factor de ocupación de la superficie FO, de acuerdo con la expresión:

$$\text{MES} = f(\text{FO}) \quad (5.4)$$

Esta ecuación es un ejemplo de una *ecuación auxiliar*, y establece cómo se determina la variable auxiliar MES. La relación funcional que liga al factor de ocupación FO y al multiplicador MES se representa gráficamente en la figura 5.4 y representa el efecto que ejerce el factor de ocupación del área industrial sobre la creación de empresas. De acuerdo con esta gráfica se observa que para valores pequeños del factor de ocupación FO, el multiplicador MES crece con FO; sin embargo, una vez alcanzado un determinado valor del factor de ocupación, el multiplicador MES decrece hasta anularse, cuando toda el área está edificada, es decir, cuando $\text{FO} = 1$. Las expresiones (5.3) y (5.4) representan la relación R_2 de la estructura básica del modelo. La función f introduce una no-linealidad en el modelo, y constituye un ejemplo de lo que en dinámica de sistemas se denomina una *función tabla* (denominación que alude al hecho de que la no-linealidad se introduce por puntos, mediante una tabla de valores).

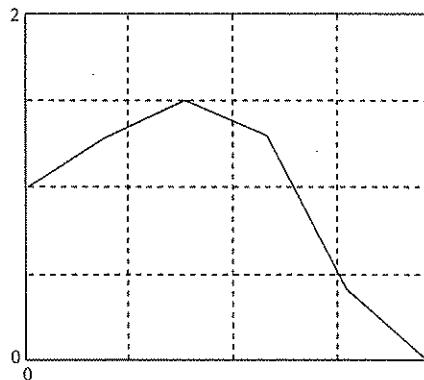


FIGURA 5.4. *Tabla del multiplicador MES.*

La determinación del factor de ocupación FO se hace mediante la ecuación:

$$\text{FO} = \frac{\text{EMP} \times \text{SPE}}{\text{S}} \quad (5.5)$$

donde SPE representa la superficie por empresa y S la superficie del área industrial. Esta expresión constituye otra ecuación auxiliar y representa la relación R_3 .

5.3.3. Fase de evaluación

Con la construcción del modelo se culmina una descripción racional del proceso que se está estudiando. A partir de este modelo se genera su comportamiento, reflejado en la figura 5.5.

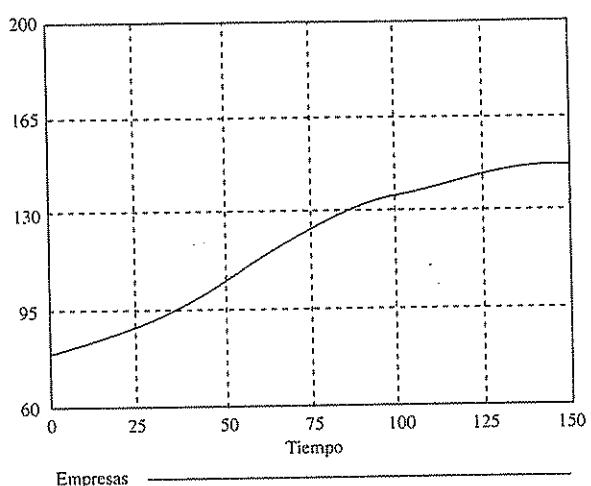


FIGURA 5.5. Comportamiento del modelo de un área industrial.

Si el modelo es correcto, el comportamiento representado en esta figura deberá reproducir, con un grado de precisión que se considere aceptable, el de la descripción empírica del sistema modelado. Normalmente, en dinámica de sistemas, se dispone de un *modo de referencia*, que constituye la descripción empírica a la que se tratará de ajustar al comportamiento del modelo. En este caso, el modo de referencia estará formado por la información de la que se disponga respecto a la evolución del número de empresas en algún periodo de tiempo, o en un área de características análogas a las estudiadas (podría darse el caso de que se careciera de ese modo de referencia, por tratarse de un tipo de situación en la que confluyesen una serie de factores que previamente no se hubiesen dado; en tal caso, la evaluación del modelo se haría por parte de los expertos, en tanto se considerase que incorporara todas las hipótesis que razonablemente se pudiesen hacer con relación a esa situación). Una técnica complementaria de evaluación del modelo es la del *análisis de sensibilidad* que se describirá en el capítulo 6.

El comportamiento que presenta el modelo (figura 5.5) tiene la forma, indicada más arriba, de crecimiento logístico o sigmoidal. En él se observan dos fases perfectamente diferenciadas. En la primera, el comportamiento es de crecimiento aproximadamente exponencial. Ello es debido a que para valores suficientemente pequeños de la variable empresas la tabla que representa el multiplicador MES (recuérdese la figura 5.4) presenta un crecimiento, por lo que si se analiza el correspondiente bucle de realimentación, éste resulta ser positivo, y por tanto el comportamiento es de crecimiento autorreforzado o exponencial. Por el contrario, en la segunda fase de crecimiento, la tabla de la figura 5.4 muestra una pendiente negativa, a la que corresponde una relación negativa, de modo que el bucle de realimentación pasa de positivo a negativo, dando lugar a un comportamiento que tiende a estabilizarse. De este modo, como ya se vio en la subsección 2.3.1, es posible dar una interpretación estructural al crecimiento logístico.

El modelo que se acaba de ver constituye una versión enormemente simplificada de lo que sucede en un área industrial. Se ha considerado únicamente la evolución del número de

empresas. Se ha prescindido de factores tales como la evolución de la población o de la disponibilidad de viviendas (en la propia área industrial o en torno a ella, del que potencialmente se nutra). Además, la evolución del número de empresas se ha considerado de una forma agregada, sin tener en cuenta los distintos tipos de empresas que puedan establecerse. Tampoco se han tenido en cuenta factores de otra naturaleza, como pueden ser los medioambientales, que pueden afectar al desarrollo del área industrial en cuestión. Lo que se pretendía, como se ha indicado al principio de la sección, era sencillamente ilustrar, mediante un caso sencillo, la posibilidad de construir un modelo (un sistema dinámico) a partir de determinadas hipótesis relativas a cómo se produce el crecimiento económico de un área industrial. La metodología propuesta permite tener en cuenta otros factores, como los que se acaban de mencionar, y construir así modelos de considerable complejidad y detalle (Alfeld y Graham, 1976).

Por último conviene señalar que, por razones pedagógicas, aquí se ha expuesto la construcción del modelo de una forma secuencial. En realidad, para llegar a las ecuaciones del modelo normalmente habrá que acudir al proceso iterativo que se mencionó en la sección 5.2 y que se ilustra en la figura 5.1.

5.4. DESCRIPCIÓN DE UN MODELO DE DEMANDA ELÉCTRICA: MEREA3

Esta sección se va a dedicar a describir, a grandes rasgos, un modelo desarrollado mediante la técnica de dinámica de sistemas: el modelo MEREA3. El objetivo del modelo es reflejar el comportamiento de la demanda de energía eléctrica en España. El interés del modelo es el siguiente: el proceso de construcción de centrales eléctricas introduce grandes retrasos, lo que obliga a predecir la demanda para decidir las inversiones en capacidad de generación de energía. Además, la propia decisión de inversión aumenta los costes financieros de las compañías eléctricas, lo que afectará a los precios y, por tanto a la demanda, con lo que la predicción que llevó a la decisión de inversión deberá ser revisada. Se observa que existe una serie de influencias endógenas en el sistema que hacen que el análisis del problema sea complejo. Un modelo realizado con la técnica de la dinámica de sistemas puede ayudar en este análisis.

En realidad, el modelo MEREA3 no sólo predice la evolución de la demanda de electricidad, sino también la demanda de energía total y su distribución entre electricidad, gas natural y otras energías. Los datos del modelo (variables exógenas) son la actividad económica del país (medida, por ejemplo, mediante el PIB), y los precios de las distintas energías excepto de la electricidad. Este precio lo calcula el propio modelo simulando el comportamiento de las compañías eléctricas, calculando todos los costes asociados a la generación y distribución de energía eléctrica y simulando el marco legal estable que fija la tarifa eléctrica en función de los costes de las compañías.

La figura 5.6 presenta la estructura simplificada del modelo. De acuerdo con el diagrama de esta figura, el Marco Legal Estable determina los precios que influyen sobre la demanda y ésta, a su vez, influye sobre el Marco Legal Estable. Por otra parte esa demanda de electricidad, junto con su precio, son considerados por el Sector Eléctrico en su proceso de suministro. Este sector genera unos costes que afectan al Marco Legal Estable, con lo que se completa el diagrama de la figura 5.6.

A continuación se van a tratar con mayor detalle los bucles del modelo MEREA3. En primer lugar, en la figura 5.7, se muestra el bucle de demanda. En él se pone de manifiesto cómo sobre

Algunas variables tienen un efecto directo sobre otras, pero otras tienen un efecto indirecto que se produce a través de la acción de otras variables. Los bucles de retroalimentación positiva y negativa son los mecanismos fundamentales que permiten que las variables se modifiquen y evolucionen.

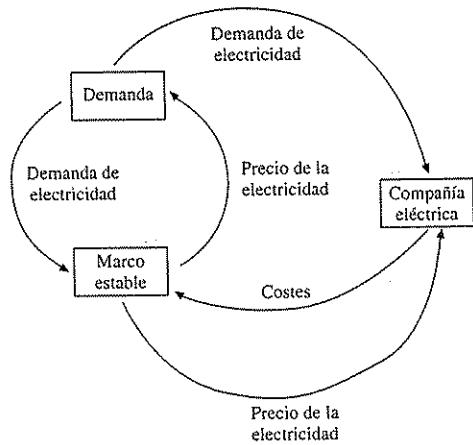


FIGURA 5.6. Estructura simplificada del modelo MEREA3.

la demanda de electricidad influyen tanto la actividad económica como la modernidad, ambos factores exógenos, y el precio como factor endógeno.

Como ya se ha indicado más arriba, el precio viene determinado por el Marco Legal Estable, que lo regula a partir de la demanda de electricidad y de los costes de generación, transporte y distribución. Se trata de un bucle de realimentación positiva en el que se tienen dos retrasos considerables, que se indican en el diagrama mediante dos trazos situados sobre la flecha que representa la influencia correspondiente.

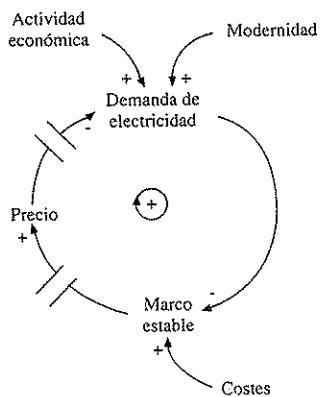


FIGURA 5.7. Estructura básica del bucle de demanda.

En la figura 5.8 se observa el bucle de construcción o planificación. A partir de una cierta demanda de electricidad, en un instante de tiempo determinado, se hace una previsión de la capacidad de generación que se requerirá en el futuro. La comparación entre esta capacidad de generación en el futuro y la que actualmente se dispone (corregida con las construcciones ya en

curso y los retiros del parque) desencadena la necesidad de una planificación para satisfacer ese aumento previsible en la demanda futura. Estas interacciones son las que se han representado en la figura 5.8.

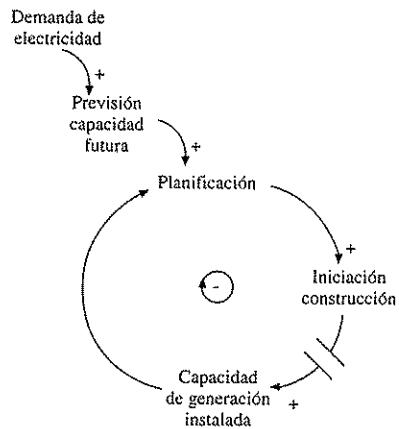


FIGURA 5.8. *Estructura básica del bucle de planificación.*

Se trata de un diagrama de realimentación negativa de carácter regulador, que trata de ajustarse a la previsión de capacidad necesitada en el futuro.

La conexión del bucle de demanda con el de planificación da lugar al diagrama global del modelo que aparece en la figura 5.9. Conviene observar que la interconexión de los dos subsistemas introduce un nuevo bucle adicional, que es de realimentación negativa, y, por tanto, de carácter autorregulador.

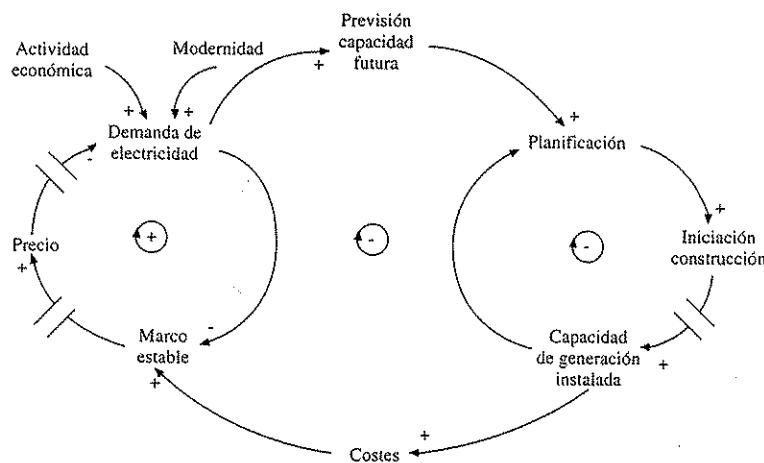


FIGURA 5.9. *Acoplamiento de los bucles de demanda y planificación, para formar la estructura básica del modelo MEREA3.*

Los diagramas anteriores completan un primer acercamiento a la fase de conceptualización del modelo MEREA3. Como se puede entender fácilmente, el modelo detallado es muy complejo y escapa del alcance de este libro, por lo que nos vamos a detener en este punto sin avanzar más en la descripción del modelo. Con esta somera descripción se ha pretendido ilustrar, mediante un ejemplo de carácter profesional, que la utilidad de la dinámica de sistemas supera ampliamente el carácter elemental y didáctico del resto de los ejemplos que se han utilizado hasta ahora.

6.1. INTRODUCCIÓN

En los capítulos anteriores se han expuesto las técnicas de la dinámica de sistemas para el desarrollo de modelos. Estas técnicas pueden complementarse con otras que, sin pertenecer específicamente a la dinámica de sistemas, ayudan a analizar y mejorar los modelos. De algunas de estas técnicas se ocupa este capítulo.

6.2. LINEALIZACIÓN DE SISTEMAS NO LINEALES: MÉTODO DE LAS PERTURBACIONES

Aunque los modelos que se construyen con ayuda de la dinámica de sistemas son no lineales, es decir, poseen no-linealidades en su estructura, sucede que los problemas no lineales presentan una serie de limitaciones, como son:

- Las soluciones no pueden generalizarse. Las conclusiones que se extraen sólo son válidas para las condiciones iniciales y los valores de los parámetros con que han sido determinadas.
- En general no existen soluciones analíticas, sino que deben obtenerse de forma numérica o gráfica; la complejidad que presentan requiere normalmente el empleo del computador.
- Las teorías de la estabilidad y de la optimización son aplicables sólo en forma muy limitada.

Estas razones, entre otras, invitan a estudiar formas linealizadas de los modelos. Estas linearizaciones se hacen en torno a un punto de operación nominal, de forma análoga a como en geometría se sustituye una superficie por su plano tangente en un punto. Obviamente este tipo

de aproximaciones sólo son válidas para pequeñas variaciones en torno al punto de operación nominal.

Por ello, la linealización de un modelo no lineal debe tomarse con enormes precauciones. En resumen, cabe decir que el interés de la linealización reside en la posibilidad de estudiar mediante el modelo linealizado determinadas propiedades del modelo original, pero en todo caso las conclusiones que se extraigan del estudio de un modelo linealizado deben confirmarse sobre el modelo no lineal. La ventaja de la linealización estriba en que permite obtener, de una forma sencilla, una orientación respecto a cómo actuar, pero nunca una solución definitiva.

Para realizar la linealización de un sistema no lineal se emplea el llamado método de las perturbaciones. Consiste éste, esencialmente, en considerar pequeñas variaciones en torno al punto de operación nominal de manera que estas perturbaciones puedan considerarse lo suficientemente pequeñas para admitir un tratamiento de infinitésimos. Para aplicar el método de perturbaciones, en las ecuaciones se realizan sustituciones de la forma

$$x = x_0 + \delta x \quad (6.1)$$

en donde x es la variable perturbada, x_0 el valor de referencia de x o punto de operación nominal y δx la perturbación o variación de x .

Si a todas las variables del modelo del sistema escrito en forma analítica se aplica una sustitución de la forma anterior y se desprecian los productos y potencias de las perturbaciones, se obtiene un modelo lineal con respecto a las variaciones. Estas ecuaciones, denominadas de perturbación, al ser lineales son resolubles analíticamente, aparte de presentar otros tipos de ventajas. No debe olvidarse, sin embargo, que las ecuaciones de perturbación sólo son válidas para "pequeñas perturbaciones". La forma linealizada de un modelo dinámico puede ser aceptablemente válida en un cierto periodo para el que se diseña el modelo.

Un tipo de estudios para los que los modelos linealizados son sumamente interesantes son los análisis de sensibilidad con respecto a los parámetros. A esta cuestión se dedicará la sección 6.3 del presente capítulo. El empleo de modelos lineales en ese contexto, habida cuenta de la propiedad de superposición, garantiza que este tipo de análisis pueda realizarse separadamente para cada uno de los parámetros y posteriormente determinar, por superposición, los efectos conjuntos.

A continuación se explica más detalladamente cómo se aplican los métodos de perturbaciones.

6.2.1. Linealización de funciones sin memoria

Sea una función no lineal del tipo

$$z = f(x, y)$$

Se desea linealizar esta función en torno a un punto nominal (x_0, y_0) , $z_0 = f(x_0, y_0)$. Para ello se puede desarrollar la función f en serie de Taylor en torno al punto (x_0, y_0) y despreciar los términos de orden superior a uno:

$$z = f(x, y) \approx f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x} \Big|_{(x_0, y_0)} + (y - y_0) \frac{\partial f}{\partial y} \Big|_{(x_0, y_0)} \quad (6.2)$$

Teniendo en cuenta la definición de perturbación (6.1) y que $z_0 = f(x_0, y_0)$, la ecuación (6.2) puede expresarse del siguiente modo:

$$\delta z \approx \delta x \frac{\partial f}{\partial x} \Big|_{(x_0, y_0)} + \delta y \frac{\partial f}{\partial y} \Big|_{(x_0, y_0)}$$

Se obtiene así una ecuación que es lineal. Obsérvese que la ecuación obtenida depende de los valores nominales (x_0, y_0) .

Las funciones no lineales que suelen aparecer con más frecuencia son funciones multiplicativas o de tipo cociente. En el primer caso:

$$z = f(x, y, t) = x(t) \cdot y(t)$$

el anterior procedimiento lleva a

$$\delta z = y_0 \delta x + x_0 \delta y$$

Para funciones del tipo cociente:

$$z = f(x, y, t) = \frac{x(t)}{y(t)}$$

se llega a:

$$\delta z = \frac{x_0}{z_0} \left(\frac{\delta x}{z_0} - \frac{\delta y}{x_0} \right)$$

6.2.2. Ejemplo: linealización del modelo de desarrollo de un área industrial

Como ejemplo de la técnica de linealización expuesta anteriormente se va desarrollar su aplicación en el modelo de crecimiento de un área industrial mostrado en la sección 5.3. Este modelo se puede formular en una sola ecuación:

$$\frac{d(\text{EMP})}{dt} = \text{TNCE} \times \text{EMP} \times f \left(\frac{\text{EMP} \times \text{SPE}}{S} \right) - \frac{\text{EMP}}{\text{VME}}$$

donde f es la función no lineal que modela la influencia del factor de ocupación ($\text{FO} = \text{EMP} \times \text{SPE}/S$) en la construcción de empresas. Para simplificar la notación, y adoptar una formulación más matemática se va adoptar el siguiente cambio de variables:

x	EMP
m	$1/\text{VME}$
n	TNCE
k	SPE/S

Con estos cambios la ecuación anterior resulta:

$$\frac{dx}{dt} = nx f(kx) - mx \quad (6.3)$$

Donde $FO = kx$. Obsérvese que la única variable que depende del tiempo es x , permaneciendo n , k y m constantes. Para concretar se tomarán los siguientes valores: $n = 0,5$, $k = 375 \times 10^{-3}$ y $m = 0,066$. La función f viene dada por la tabla de valores que aparece en la tabla 6.1.

FO	0	0,2	0,4	0,6	0,8	1
$f(FO)$	1	1,3	1,5	1,3	0,4	0

TABLA 6.1. *Definición de la función f.*

Se va a linealizar este modelo respecto a uno de sus puntos de equilibrio aunque, como se ha indicado anteriormente, no es necesario que el punto en torno al cual se linealiza sea de equilibrio.

Los equilibrios se obtienen igualando a cero el segundo término de la expresión (6.3) (ya que en estado de equilibrio las derivadas respecto al tiempo son nulas):

$$nx f(kx) - mx = 0$$

ecuación que tiene tres soluciones:

$$\begin{aligned} x_1 &= 0 \\ f(kx) &= \frac{m}{n} = 1,33 \Rightarrow \begin{cases} kx_2 = 0,233 \Rightarrow x_2 = 62 \\ kx_3 = 0,566 \Rightarrow x_3 = 151 \end{cases} \end{aligned}$$

Se va a linealizar en torno al equilibrio $x_0 = x_3 = 151$. Eligiendo, como se ha explicado anteriormente, la variable de perturbaciones δx de manera que $x = x_0 + \delta x$ la ecuación (6.3) queda:

$$\frac{d(\delta x)}{dt} = \left[n \left(f(kx_0) + x_0 \frac{df}{dx} \Big|_{x_0} \right) - m \right] \delta x \quad (6.4)$$

Teniendo en cuenta que

$$\frac{df}{dx} \Big|_{x_0} = \frac{df}{d(FO)} \Big|_{x_0} \frac{d(FO)}{dx} \Big|_{x_0} = (-4,5)k$$

y sustituyendo en la ecuación (6.4) resulta:

$$\frac{d(\delta x)}{dt} = -0,1637 \delta x \quad (6.5)$$

que es la expresión linealizada del modelo. De esta ecuación se puede obtener la evolución de δx y haciendo $x = x_0 + \delta x$ se obtiene la de las empresas. La expresión 6.5 es lineal y

está ampliamente estudiada, por lo que es inmediato obtener importantes conclusiones sobre su solución.

Con el fin de comprobar el grado de bondad alcanzado con la linealización se van a mostrar los resultados de dos experimentos distintos. En primer lugar se simulan tanto el modelo original como el linealizado partiendo de la condición inicial $x(t_0) = 100$. La evolución de las empresas para ambos casos se muestra en la figura 6.1a. En ella se puede observar que existe un gran parecido entre ambas curvas (aunque existen diferencias cuantitativas). Además, no debe olvidarse que en los modelos manejados en dinámica de sistemas lo que interesa fundamentalmente es la forma cualitativa de la respuesta. Por lo tanto, el modelo linealizado es una buena aproximación (en este caso) del modelo original y además tiene la ventaja de carecer de los inconvenientes de los sistemas no lineales señalados anteriormente.

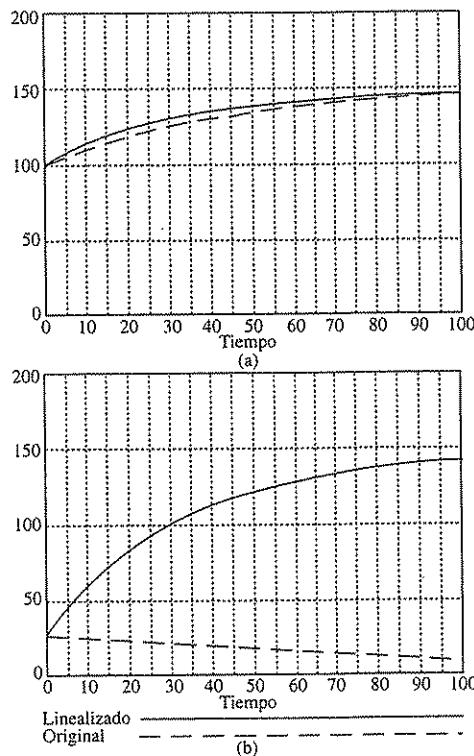


FIGURA 6.1. Comparación de la evolución del número de empresas con el modelo original y linealizado para (a) $x(t_0) = 100$ y (b) $x(t_0) = 30$.

El segundo experimento corresponde a la condición inicial $x(t_0) = 30$. Los resultados aparecen en la figura 6.1b. En este caso existe una gran discrepancia entre ambas curvas. Esto es debido a que la linealización es una aproximación válida en las proximidades del punto respecto al que se linealiza y, en este caso, se parte de un punto muy lejano (el punto de linealización corresponde a un valor de 151). La diferencia entre ambos comportamientos se

puede aclarar si se tiene en cuenta que un sistema lineal solamente puede tener un punto de equilibrio, mientras que un modelo no lineal puede tener varios. Es lo que sucede en las gráficas de la figura 6.1: mientras que en la 6.1a el sistema tiende al valor de equilibrio 151, en la 6.1b el equilibrio al que tiende es 0. Esta distinción no la capta el modelo lineal, que siempre tiende al primero de estos atractores. Por tanto, el modelo linealizado no puede tener toda la riqueza de comportamiento de los sistemas no lineales que tengan más de un equilibrio. En cualquier caso, en la sección 6.3 y en el capítulo 7 se estudiará más detenidamente qué es lo que sucede en este modelo.

6.3. ANÁLISIS DE SENSIBILIDAD

El análisis de sensibilidad de un sistema dinámico consiste en el estudio de los efectos de las variaciones de los valores que toman sus parámetros y sus condiciones iniciales, sobre el comportamiento del mismo. Al estudiar la sensibilidad de un modelo a los cambios que se pueden producir en sus parámetros, o en las condiciones iniciales, no sólo se obtiene una mayor comprensión del comportamiento del modelo, sino que, además, este análisis contribuye al establecimiento tanto de la validez como de la utilidad de ese modelo. La contribución a la validez del modelo es obvia porque una cualidad intrínseca del mismo es precisamente su insensibilidad a las variaciones de los parámetros. Esta falta de sensibilidad permite que las conclusiones que se extraigan del modelo no estén excesivamente ligadas a unos determinados valores de los parámetros, que en algunos casos pueden ser discutibles. Una de las metas perseguidas por la dinámica de sistemas es que los resultados deben depender exclusivamente de la estructura del modelo.

Por lo que respecta a la utilidad del modelo, en el sentido de ayuda a la toma de decisiones respecto a la implementación de determinadas políticas, el análisis de sensibilidad tiene una singular importancia porque permite conocer cuáles son los parámetros que tienen mayor influencia para actuar sobre una determinada variable.

En resumen, el análisis de sensibilidad permite saber en qué medida una determinada conclusión que se extrae del análisis de un modelo depende de una cierta hipótesis sobre el valor de algún parámetro. Permite saber qué precisión se exigirá a cada uno de los parámetros que intervienen en el modelo para que las conclusiones que se extraigan de éste sean aceptables. Por lo que respecta a la fase de explotación del modelo para el establecimiento de una política de actuación sobre el sistema real, el análisis de sensibilidad permite decidir cuáles son los parámetros que tienen mayor influencia en el comportamiento del sistema, y, por tanto, sobre los que debe actuarse en la realidad para obtener un determinado efecto en el comportamiento del mismo.

Existen dos maneras de enfocar el problema de la determinación de la sensibilidad de un modelo. En primer lugar, se puede proceder a variar los parámetros cuya sensibilidad se trata de analizar y a realizar simulaciones para cada una de estas variaciones. Así, por ejemplo, si se trata de estudiar la sensibilidad de un modelo a las variaciones de un parámetro p , se realizan simulaciones del modelo para variaciones de este valor de p de, por ejemplo, un 5, un 10%, ..., de su valor nominal. Observando los gráficos del comportamiento del sistema ante estas variaciones del parámetro, se obtiene una imagen de la sensibilidad del modelo frente a este parámetro. Es obvio que este procedimiento de análisis de sensibilidad puede

resultar tedioso si no se dispone de herramientas informáticas adecuadas para realizarlo de forma automática. Sin embargo, los entornos habituales suelen incorporar estas herramientas, por lo que en la práctica es un procedimiento muy empleado.

6.3.1. Análisis de sensibilidad por linealización

El segundo enfoque del análisis de sensibilidad de un modelo consiste en la aplicación de técnicas analíticas. Estas técnicas conducen a realizar el análisis de sensibilidad de forma más simple que por simulaciones sucesivas. Con este enfoque es posible llegar a un algoritmo para el análisis de sensibilidad con el que se obtienen trayectorias temporales de los cambios que se producen en las variables de estado debido a variaciones, pequeñas y constantes, en los parámetros y en las condiciones iniciales. Supóngase un sistema dinámico descrito por¹

$$\dot{x}(t) = f[x(t), p] \quad (6.6)$$

$$x(t_0) = x^0 \quad (6.7)$$

en donde el vector $p \in \mathbb{R}^{n_p}$ representa el conjunto de parámetros cuya sensibilidad se quiere analizar. El modelo anterior va acompañado por un conjunto de condiciones iniciales x^0 susceptibles de sufrir pequeñas variaciones. Se define el vector s como sigue

$$s = \begin{pmatrix} x^0 \\ p \end{pmatrix}$$

La variación de la derivada de la trayectoria del sistema $x(t)$ frente a una pequeña variación de los elementos del vector s vendrá dada por

$$\frac{\partial \dot{x}}{\partial s} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial f}{\partial s} \quad (6.8)$$

Si se deriva cada elemento de x^0 respecto a cada elemento de s resulta:

$$\frac{\partial x_i^0}{\partial s_j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (6.9)$$

Supóngase una trayectoria nominal x^* , correspondiente a un conjunto determinado de parámetros y condiciones elementales s^* . Se puede definir

$$\sigma(x^*, t) = \left. \frac{\partial x(t)}{\partial s} \right|_{x^*(t)}$$

que representa la variación de la trayectoria $x(t)$, con respecto a la trayectoria nominal x^* , por efecto de la variación del vector s . Los elementos de la matriz $\sigma(x^*, t)$ reciben la denominación de *coeficientes de sensibilidad*; también se denominan *funciones de sensibilidad*, o *coeficientes de influencia*. El elemento $\sigma_{ij}(t) = \partial x_i / \partial s_j$ representa la variación del nivel x_i ,

¹ Se usará aquí la notación \dot{x} para representar la derivada de x respecto al tiempo $\dot{x} = dx/dt$

en el instante t , por efecto de la variación de s_j , es decir, de un parámetro o de una condición inicial. La expresión (6.8) se puede escribir

$$\frac{d\sigma}{dt}(x^*, t) = F(x^*, t)\sigma(x^*, t) + G(x^*, t) \quad (6.10)$$

siendo

$$F(x^*, t) = \left. \frac{\partial f}{\partial x} \right|_{x^*(t)} \quad G(x^*, t) = \left. \frac{\partial f}{\partial s} \right|_{x^*(t)} \quad (6.11)$$

La matriz $F(x^*, t)$ es el jacobiano de la función f .

La expresión (6.10) representa cómo se propaga el efecto sobre la trayectoria x de la variación de un elemento del vector s y recibe la denominación de *ecuación de sensibilidad*. Las condiciones iniciales de la ecuación diferencial (6.10) vendrán dadas por la ecuación (6.9):

$$\sigma_{ij}(x^*, t_0) = \sigma_{ij}^0 = \left. \frac{\partial x_i^0}{\partial s_j} \right|_{s^0} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (6.12)$$

Los elementos de la matriz de coeficientes de sensibilidad pueden interpretarse de la forma siguiente. Supóngase una pequeña variación Δs del vector s . En vista de la definición de σ se puede escribir, de forma aproximada,

$$x(s + \Delta s, t) \approx x(s, t) + \sigma(x^*, t)\Delta s \quad (6.13)$$

Esta aproximación corresponde a una linealización del modelo respecto a la trayectoria de referencia, utilizando el mismo método descrito en la sección 6.2. Gracias a esta linealización, los coeficientes σ_{ij} son una medida del efecto que tendría la variación del elemento j del vector s en la trayectoria del elemento i del vector de estados. Sin embargo, no debe olvidarse que la linealización es válida en el entorno del punto respecto al cual se hace (en este caso respecto a la trayectoria de referencia), dejando de ser una aproximación cuando nos alejamos de éste.

Supóngase un componente s_i del vector s . La expresión (6.13) se puede escribir

$$\Delta x = \sigma_i \Delta s_i \quad (6.14)$$

donde σ_i es la columna i de la matriz σ . La anterior expresión se puede interpretar diciendo que una norma apropiada para el vector σ_i es una medida del cambio introducido en el vector x debido a pequeñas variaciones en el parámetro s_i .

La expresión de los coeficientes de sensibilidad de la forma $\partial x_i(t)/\partial s_j$ suministra una matriz de valores que no pueden compararse entre sí. Por este motivo interesa introducir una expresión de los coeficientes de sensibilidad de la forma

$$\sigma'_{ij}(t, x^*) = \left. \frac{\partial [\ln x_i(t)]}{\partial (\ln s_j)} \right|_{x^*}$$

que da valores comparables ya que

$$\sigma'_{ij}(t, x^*) = \left. \frac{\partial [\ln x_i(t)]}{\partial (\ln s_j)} \right|_{x^*} = \left. \frac{\partial x_i(t)/x_i(t)}{\partial s_j/s_j} \right|_{x^*}$$

con lo que para pequeñas variaciones porcentuales de los parámetros y condiciones iniciales se obtiene, directamente, la variación porcentual de las variables de estado.

El cálculo de estos nuevos coeficientes de sensibilidad es inmediato a partir de los considerados más arriba

$$\sigma'_{ij}(t, x^*) = \frac{\delta[\ln x_i(t)]}{\delta(\ln s_j)} \Big|_{x^*} = \frac{s_j}{x_i(t)} \frac{\delta x_i(t)}{\delta s_j} \Big|_{x^*} = \frac{s_j}{x_i(t)} \sigma_{ij}(t, x^*) \quad (6.15)$$

Esta expresión de la sensibilidad tiene el inconveniente de la posible anulación de $x_i(t)$, con lo que el coeficiente correspondiente tendería a infinito. Sin embargo, el hecho de que los elementos sean comparables, junto con lo intuitivo de las respuestas porcentuales a variaciones porcentuales, compensa ampliamente esta posible deficiencia.

El análisis formal de sensibilidad que se acaba de desarrollar es un análisis de primer orden, de acuerdo con la interpretación dada a la expresión (6.14). Es posible concebir análisis de sensibilidad de orden superior, pero entonces la complejidad formal que se alcanza dificulta la obtención de resultados sencillos.

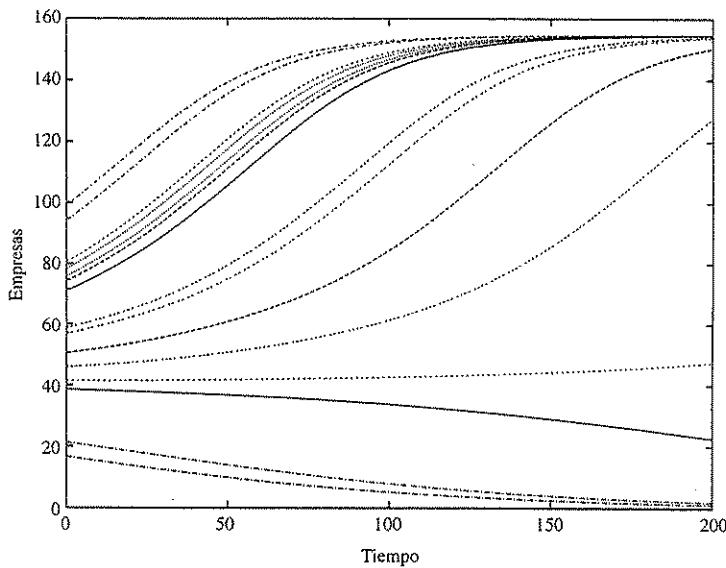
6.3.2. Ejemplo: análisis de sensibilidad del modelo de crecimiento de un área industrial

Como ejemplo de aplicación del análisis de sensibilidad se va a considerar el modelo de crecimiento de un área industrial expuesto en la sección 5.3.

Mediante simulaciones intensivas

En primer lugar se va a realizar un estudio basado en simulaciones intensivas utilizando un entorno de simulación que permite analizar la sensibilidad de forma automática. Para ello basta definir el número de simulaciones que se quieran realizar y los parámetros que se desean variar, especificando los límites entre los que van a variar éstos y con qué función de distribución. Por ejemplo, si se realizan 20 simulaciones variando únicamente el número inicial de empresas entre 10 y 100, con una distribución uniforme, se ha obtenido el resultado que aparece en la figura 6.2. En esta figura se pueden observar dos tipos de comportamiento completamente diferenciados: en primer lugar, para valores altos del número inicial de empresas, el sistema evoluciona hacia un valor de equilibrio relativamente alto igual para todos estos casos. Sin embargo, si la población inicial es pequeña, el sistema evoluciona hacia el colapso, es decir, desaparecen todas las empresas. La frontera entre estos dos comportamientos se encuentra en un valor de la población inicial cercano a 40. Estos dos tipos de comportamiento ya aparecieron en la subsección 6.2.2. A la misma conclusión se puede llegar con un estudio más riguroso que se hará en el capítulo 7.

Por tanto, con este sencillo análisis se puede extraer una idea bastante clara sobre la sensibilidad de la evolución del número de empresas respecto al parámetro *número inicial de empresas* NIE: es trascendental si NIE es superior o inferior a 40, ya que el equilibrio al que tiende el sistema es radicalmente distinto en ambos casos. Sin embargo, el valor de la diferencia entre NIE y 40 no afecta al equilibrio final, aunque sí influye en la velocidad con la que el sistema se acerca a este equilibrio; cuanto más cerca esté NIE de 40, más lentamente evolucionará el número de empresas.

FIGURA 6.2. *Resultado del análisis de sensibilidad por simulaciones intensivas.*

Mediante linealización

Como complemento del análisis de sensibilidad anterior se puede seguir la técnica expuesta en la sección 6.3.1. Para ello se parte de la ecuación analítica del modelo ya expuesta anteriormente:

$$\frac{dx}{dt} = nx f(kx) - mx$$

Para simplificar, se va a considerar una expresión analítica para la función f , aunque también sería posible utilizar la expresión en forma de tabla de valores que aparece en la tabla 6.1. Se va a tomar:

$$f(\xi) = -3,75\xi^2 + 2,75\xi + 1$$

Por tanto:

$$\begin{aligned} \frac{dx}{dt} &= nx(-3,75k^2x^2 + 2,75kx + 1) - mx = \\ &= -3,75nk^2x^3 + 2,75nkx^2 + (n-m)x \end{aligned} \tag{6.16}$$

En este ejemplo se va a calcular la sensibilidad respecto a la condición inicial $x_0 = x(t_0)$ y el parámetro $n = \text{TNCE}$. Por tanto, el vector s será:

$$s = \begin{pmatrix} x_0 \\ n \end{pmatrix}$$

Algunas de las ecuaciones que aparecen en este apartado se refieren a sistemas de control. Los sistemas de control son sistemas dinámicos que responden a estímulos o señales de entrada para producir una respuesta deseada.

El cálculo de los coeficientes de sensibilidad se realiza utilizando las ecuaciones (6.10) y (6.11) (téngase en cuenta que en estas expresiones la función f es igual al segundo miembro de la definición del sistema 6.17):

$$F(x^*, t) = \frac{\partial f}{\partial x(t)} \Big|_{x^*} = -3,75nk^23x^{*2} + 2,75nk2x^* + n - m \quad (6.17)$$

donde x^* corresponde a la evolución de las empresas en una trayectoria nominal. Por otro lado, si se descompone $G(x^*, t)$ en $G = (g_1, g_2)$, se tiene:

$$g_1(x^*, t) = \frac{\partial f}{\partial s_1} \Big|_{x^*} = \frac{\partial f}{\partial x_0} \Big|_{x^*} = 0 \quad (6.18)$$

$$g_2(x^*, t) = \frac{\partial f}{\partial s_2} \Big|_{x^*} = \frac{\partial f}{\partial n} \Big|_{x^*} = -3,75k^2x^{*3} + 2,75kx^{*2} + x^* \quad (6.19)$$

Como se puede observar F y G dependen de x^* . Para obtener los coeficientes de sensibilidad se puede ampliar el modelo incluyendo las ecuaciones (6.10), (6.17), (6.18) y (6.19), de manera que se utiliza $x^*(t)$ a medida que se va calculando. Además se debe incluir la condición inicial de la ecuación (6.10) que viene dada por (6.12), por lo que resulta:

$$\sigma(x^*, t_0) = (1 \ 0)$$

Con estas ecuaciones se puede obtener fácilmente la evolución de cada uno de los componentes del vector de sensibilidad. Incluso, mediante la expresión (6.15) se puede hallar el valor del vector σ' . En la figura 6.3 se presenta la evolución de la primera componente de este vector, es decir, la sensibilidad respecto al número inicial de empresas, cuando la trayectoria nominal corresponde a un número inicial de empresas igual a 70. Obsérvese que esta curva parte, como es obvio, del valor 1, que significa que un cambio en el número inicial

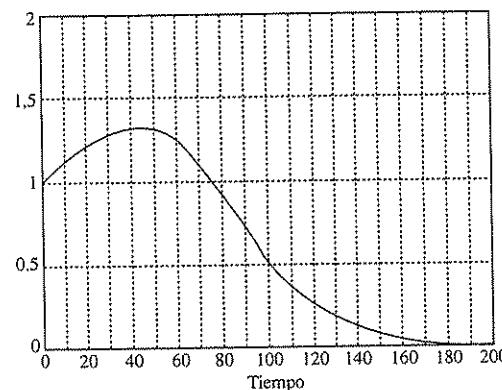


FIGURA 6.3. Sensibilidad respecto al número inicial de empresas. Trayectoria nominal con $x(t_0) = 70$.

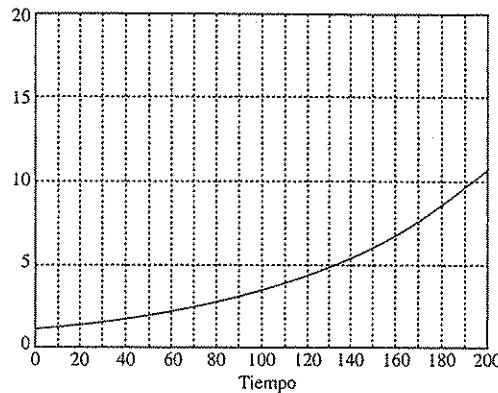


FIGURA 6.4. *Sensibilidad respecto al número inicial de empresas. Trayectoria nominal con $x(t_0) = 40$.*

de empresas afectará en la misma cuantía al número de empresas. Se observa también que este efecto crece al aumentar el tiempo, hasta llegar a un cierto valor en el que el efecto empieza a decrecer y, finalmente, se anula. Este fenómeno concuerda con el observado mediante simulaciones intensivas: obsérvese en la figura 6.2 que las trayectorias que empiezan en los alrededores de 70 son ligeramente divergentes en los años iniciales para converger más tarde todas en el mismo valor.

Otro fenómeno que se observaba en el estudio mediante simulaciones intensivas era el cambio cualitativo de la respuesta cuando el número inicial de empresas está por encima o por debajo de 40. Veamos cómo este fenómeno aparece reflejado en el estudio de sensibilidad por linearización. Si se calcula la evolución de la sensibilidad σ' con respecto al número inicial de empresas cuando la trayectoria nominal corresponde a un número inicial igual a 40, se obtiene la gráfica de la figura 6.4. Se observa que en este caso la evolución del coeficiente de sensibilidad es totalmente distinta que la anterior, de manera que aumenta constantemente, reflejando que el estado final que se alcanza, en caso de variar el número inicial de empresas, será diferente.

Por supuesto, mediante este procedimiento también se puede calcular la evolución del coeficiente de sensibilidad respecto al parámetro TNCE, como aparece reflejado en la figura 6.5 para los casos de número inicial de empresas igual a 70 y a 40. En esta figura se observa también el diferente comportamiento en ambos casos.

6.4. ESTIMACIÓN DE PARÁMETROS EN MODELOS DINÁMICOS

Como se comentó anteriormente, en la fase de formulación del modelo es necesario asignar valores a las constantes del mismo. Ésta es una etapa complicada en la que el modelista frecuentemente tropieza con la ausencia de la información adecuada. Aquí se va a tratar el problema de cómo dar valor a estos datos a los que denominaremos *parámetros*, que corresponderán a constantes o a los valores que definen las tablas (recuérdese la sección 3.5.1).

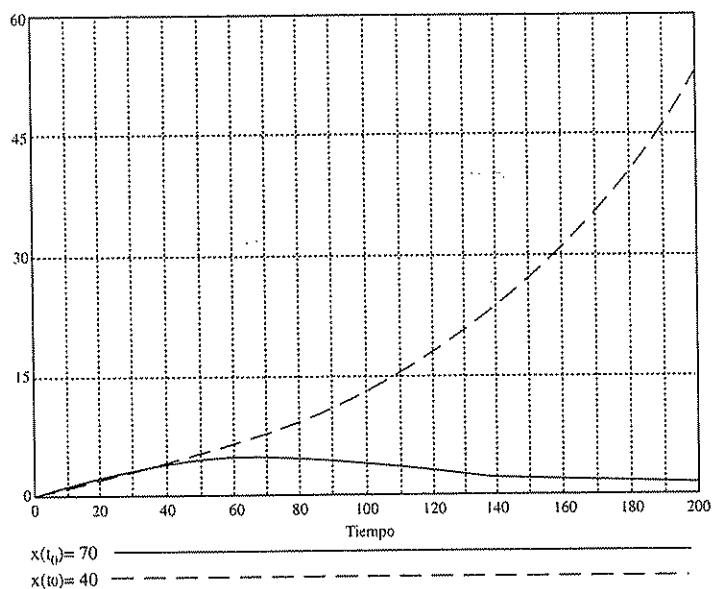


FIGURA 6.5. Sensibilidad respecto a TNCE.

En este punto cabe distinguir dos tipos de situaciones. La primera corresponde al caso en que un parámetro del modelo tiene un significado real perfectamente establecido —por ejemplo, un parámetro que representa una tasa de nacimientos— y se dispone de suficiente información como para asignar a este parámetro un valor con la precisión requerida. Esta precisión puede estar condicionada, por ejemplo, por el análisis de sensibilidad del modelo. En este caso, la asignación de un valor al parámetro en cuestión no ofrece ningún problema; se asigna el valor que se desprende de los datos disponibles. Este tipo de situaciones es el que resalta Forrester al enunciar la ventaja de la dinámica de sistemas, permitiendo asignar un significado real a los parámetros que intervienen en los modelos.

Sin embargo, se presenta un segundo tipo de casos en el que, aunque el parámetro tenga un significado real, no es fácil asignarle un valor numérico. Esto puede ser debido a distintas causas como, por ejemplo, la propia naturaleza del parámetro, la falta de suficientes datos respecto al mismo o a que el análisis de sensibilidad prescriba, para el parámetro en cuestión, una precisión imposible de alcanzar con la información disponible. Es decir, se pueden presentar casos en los que la asignación de un valor numérico a un parámetro no es posible hacerla de una forma aislada con ayuda del conocimiento previo.

En estos casos se puede suplir la carencia de información sobre los parámetros con la información cuantitativa disponible del comportamiento del sistema que se pretende modelar, a la que denominaremos *datos históricos*. Para ello, se puede resolver el problema de la asignación de valores a los parámetros de manera que el comportamiento del modelo se acerque lo más posible a los datos históricos. Para medir la cercanía entre las trayectorias de las variables históricas y las resultantes del modelo se suele emplear un criterio cuadrático como medida de distancia:

$$J_i = \int_{t_0}^{t_1} [x_i^h(t) - x_i^m(t)]^2 dt$$

donde $x_i^h(t)$ es el dato histórico correspondiente a la variable x_i , $x_i^m(t)$ la variable correspondiente del modelo y J_i la distancia entre ellas: es la integral (suma) del cuadrado de los errores cometidos en la simulación de la variable i . Los tiempos t_0 y t_1 son los instantes inicial y final de la simulación, respectivamente.

La distancia total entre el comportamiento del modelo y los datos históricos se obtiene por combinación lineal de las distancias J_i :

$$\mathcal{J} = \sum_{i=1}^N a_i J_i \quad (6.20)$$

donde N es el número total de variables históricas disponibles, \mathcal{J} la medida de la distancia y los coeficientes a_i (mayores que cero) los escogeremos de manera que el resultado del problema de optimización tenga sentido. Para ello:

- Se debe procurar que los distintos sumandos de \mathcal{J} tengan el mismo orden de magnitud con lo que se soluciona el problema de la elección desproporcionada de las unidades. Por ejemplo, supóngase que la variable x_1^h mide, en pesetas, los ingresos de una empresa constructora de automóviles y tiene un valor característico de miles de millones, y la variable x_2^h corresponde al número de automóviles vendidos, con un valor característico del orden de los miles. Si se elige $a_2 = 1$, entonces a_1 se debe escoger del orden de 10^{-6} (así el resultado sería equivalente a escoger el millón de pesetas como unidad para los ingresos) para que $a_1 J_1$ y $a_2 J_2$ sean números del mismo orden.
- Se debe dar valores mayores a los coeficientes a_i asociados a las variables x_i^h que se conozcan con mayor precisión, con lo que nos fijamos más en las discrepancias asociadas a estas variables.
- También se debe dar más peso (esto es, coeficientes a_i mayores) a las discrepancias asociadas a las variables que estimemos que miden mejor el comportamiento del modelo. Así si en el ejemplo anterior el objetivo del modelo es simular el comportamiento de los beneficios de la empresa, se le dará mayor peso al error cometido en la simulación de esta variable.

El problema de estimación de parámetros queda reducido a un problema de optimización (minimización) de una expresión cuadrática como la (6.20), estando los parámetros ligados por las ecuaciones que rigen el comportamiento dinámico del sistema. Este problema tiene solución analítica cuando estas ecuaciones son lineales respecto a los parámetros (la solución se conoce como principio de los mínimos cuadrados). Sin embargo, en la mayoría de los casos de interés práctico esto no sucede, por lo que es necesario acudir a técnicas numéricas que resuelvan el problema con ayuda del computador.

Algunos de los entornos de simulación de dinámica de sistemas incorporan herramientas que permiten plantear cómodamente estos problemas y los resuelven numéricamente.

En cualquier caso, no se debe abusar de estos métodos “automáticos” olvidando el procedimiento ideal que sería el poder conocer el valor de los parámetros directamente de datos

Autor: D. J. Gómez. Título: Dinámica de Sistemas. Edición: 1. Año: 2002. ISBN: 84-7829-560-0. ISSN: 1578-518X. ISSN-L: 1578-518X. © 2002, Ediciones Paraninfo, S.A. Madrid. Reservados todos los derechos. No está permitida la impresión, la fotocopia ni la transmisión en cualquier forma o por cualquier medio, ni la conversión a otro lenguaje electrónico, de todo o parte de esta obra sin la autorización escrita de los titulares del copyright.

reales. Aunque por este método no se pueda conocer cuantitativamente el valor de los mismos, sí se tendrá una idea cualitativa. Esta información no debe olvidarse al utilizar un método numérico de optimización y debe utilizarse para:

- rechazar soluciones dadas por el método numérico que contradigan la información disponible;
- guiar al método numérico en la búsqueda introduciendo restricciones de desigualdad, si el entorno informático permite plantearlas, o modificando los coeficientes a_i del índice \mathcal{J} .

6.4.1. Ejemplo: estimación de los parámetros del modelo de crecimiento de un área industrial

Para ilustrar cómo se pueden estimar los parámetros de un sistema se va a realizar la aplicación al modelo de crecimiento de un área industrial ya utilizado previamente. Los parámetros de este modelo son:

- La superficie total disponible (S)
- El número inicial de empresas (NIE)
- La superficie ocupada por cada empresa (SPE)
- La tasa normal de construcción de empresas (TNCE)
- La vida media de las empresas (VME)

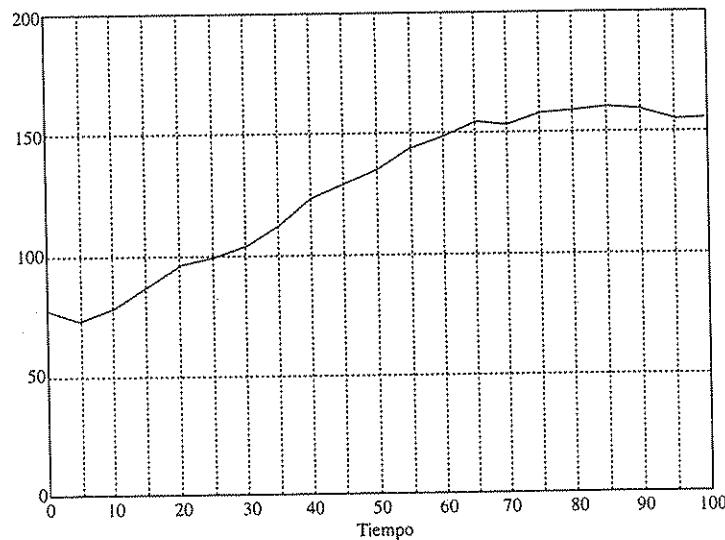


FIGURA 6.6. Datos históricos sobre la evolución del número de empresas.

Supóngase que los dos primeros parámetros son conocidos y que de los tres últimos solamente se conoce un orden de magnitud. Así se sabe que SPE está entre 5 y 15, VME entre 20 y 35 y que TNCE es menor que 0,1. Además, se dispone de una serie de datos históricos

sobre la evolución de las empresas que aparecen en la figura 6.6. Se trata de obtener el mejor valor de estos parámetros de manera que la trayectoria generada por el modelo se acerque lo más posible a la de los datos históricos. Para ello se va a utilizar un entorno de simulación que permite resolver el problema de optimización presentado en la sección previa.

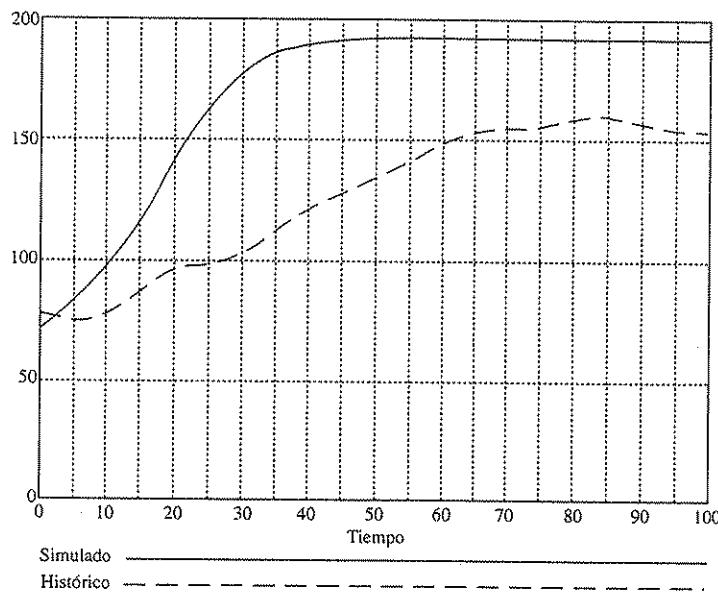


FIGURA 6.7. *Evoluciones del número de empresas correspondientes a los datos históricos y al modelo original.*

En primer lugar, como ilustración de la utilidad del método, se presenta en la figura 6.7 la evolución del número de empresas generada por el modelo cuando se toman unos valores arbitrarios de los parámetros ($SPE = 7,5$, $VME = 27,5$ y $TNCE = 0,05$), comparada con los datos históricos. Se puede observar una gran discrepancia entre ambas trayectorias, lo que hace presumir que estos valores de los parámetros no son adecuados. Se escoge como índice de bondad de una trayectoria el índice siguiente:

$$\mathcal{J} = \int_0^{100} [\text{EMP}^h(t) - \text{EMP}^m(t)]^2 dt$$

donde $\text{EMP}^h(t)$ representa la evolución de los datos históricos y $\text{EMP}^m(t)$ los generados por el modelo. Como se observa, en este ejemplo el número de variables N que se desean ajustar es igual a 1, por lo que resulta indiferente el valor de a_1 . Para los parámetros anteriores el valor del índice asociado es 47287.

Planteando el problema de optimización se realizan automáticamente 131 simulaciones (en este caso) y se obtiene el valor óptimo encontrado, correspondiente a los valores $VME = 20,18$, $SPE = 8,13$ y $TNCE = 0,44$. El valor del índice asociado es 170. La evolución del modelo para este caso aparece en la figura 6.8 junto a las trayectorias de la figura 6.7. Se puede observar

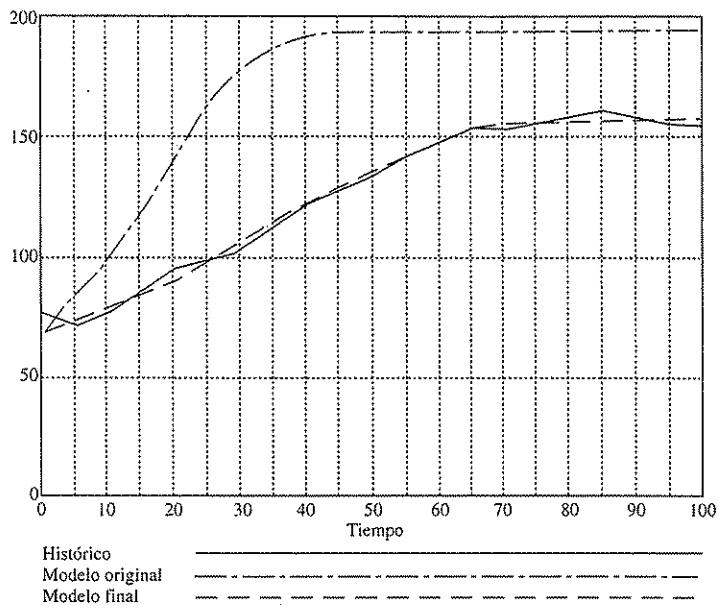


FIGURA 6.8. Evoluciones del número de empresas correspondientes a los datos históricos, al modelo original y al modelo con los parámetros ajustados.

cómo se han encontrado unos valores para los parámetros que hacen que el comportamiento del modelo se acerque mucho a la realidad, al menos en el periodo de tiempo para el que se tienen datos disponibles. Además debe notarse que, una vez planteado el problema de optimización, el resultado se obtiene de forma automática.

7.1. ANÁLISIS CUALITATIVO DE SISTEMAS DINÁMICOS

En los capítulos anteriores hemos visto cómo construir un modelo de simulación del comportamiento de un sistema empleando los útiles que suministra la dinámica de sistemas. Estos útiles son tanto conceptuales como operativos y permiten tener informáticamente el modelo de un proceso problemático que queremos estudiar. De este modo podemos obtener las trayectorias que muestran el comportamiento del sistema. Una vez fijados los valores de los parámetros y establecidas unas condiciones iniciales el computador nos devuelve las trayectorias correspondientes a esos valores.

Pudiera parecer con ello que la relación entre modelo y comportamiento es tal que a un modelo dado corresponde un único comportamiento. Sin embargo, esto no es exactamente así. A un mismo modelo pueden corresponder varios comportamientos, como vamos a ver en este capítulo.

Recordemos que un sistema dinámico viene dado por:

$$\frac{dx}{dt} = f(x) \quad (7.1)$$

donde $x \in \mathbb{R}^n$. Esta ecuación define un campo vectorial de manera que, a cada punto $x \in \mathbb{R}^n$ le asocia un vector $f(x)$, que representa la velocidad del cambio de x . La integración de esta ecuación conduce al conjunto de trayectorias que representan el comportamiento del sistema. Las trayectorias dependen de los valores de partida: las llamadas *condiciones iniciales* para realizar la integración. Si las cambiamos obtendremos otras trayectorias. Estas diferentes trayectorias, en principio, diferirán entre sí en lo que se conoce como *comportamiento transitorio*, pero normalmente todas tienden a un mismo comportamiento a largo plazo. Confluyen en un comportamiento asintótico hacia un mismo estado constante final, o hacia un comportamiento oscilatorio periódico. El comportamiento a largo plazo se conoce también como *comportamiento en régimen permanente*. La pregunta que inmediatamente surge es: el hecho de haber

identificado un comportamiento en régimen permanente, ¿excluye la posibilidad de que haya otros? Y si esto no es así, es decir, si un mismo sistema puede presentar distintos modos de comportamiento en régimen permanente, ¿es posible tener una perspectiva global respecto a todos estos modos de comportamiento?

Existe una rama de la teoría matemática de sistemas dinámicos que se ocupa de estas cuestiones: la teoría cualitativa de sistemas dinámicos. Sus orígenes se remontan a los trabajos de Henri Poincaré, y constituye un intento de síntesis entre métodos geométricos y topológicos con métodos de análisis clásico, para el estudio de sistemas dinámicos. Esta teoría pretende alcanzar una visión global de los modos de comportamiento del sistema, y no una trayectoria particular. Con el fin de aclarar el modo de proceder en el ámbito de esta teoría conviene que volvamos al sistema dinámico (7.1) y consideremos el caso especialmente sencillo en el que x es un escalar. En tal caso f será una función que podemos representar gráficamente. Supongamos que esta representación es la de la figura 7.1. Esta figura se puede interpretar

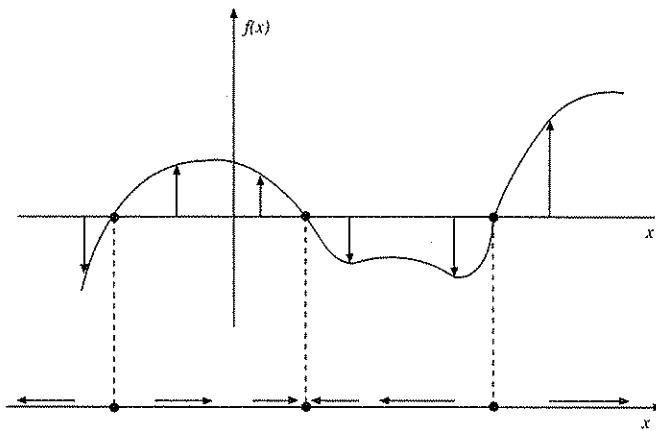


FIGURA 7.1. *Representación gráfica de la función f que define un sistema dinámico.*

como formada por los vectores que constituyen el campo vectorial asociado al sistema (7.1) en el que los distintos vectores f se han girado 90 grados en sentido contrario a las agujas de un reloj. Es inmediato que los equilibrios del sistema (7.1) serán los valores de x para los que se satisfaga la ecuación $f(x) = 0$ (ya que, para estos valores, la derivada de x respecto del tiempo es cero). Estos puntos son precisamente aquellos en que en la figura 7.1 la curva f corta al eje horizontal.

Resulta interesante observar que las tangentes a la curva f en la figura 7.1 se corresponden con las linealizaciones a las que se aludía en la sección 6.2 del capítulo anterior. En los equilibrios, las pendientes de estas tangentes son los autovalores del sistema dinámico y determinan la estabilidad de aquéllos.

Para concretar más estas ideas sea el sistema dinámico

$$\frac{dx}{dt} = -x \quad (7.2)$$

al que podemos asociar una representación geométrica como la que se muestra en la figura 7.2, en la que se ha representando el espacio de estados X (la recta real \mathbb{R}) y el vector f en cada uno de los puntos (en realidad en la figura se han dibujado únicamente unos pocos vectores, pero el lector imaginará que se asocia un vector a cada uno de los puntos de \mathbb{R}). Ésta es la representación del campo vectorial definido por la ecuación (7.2).

En la figura 7.2 es especialmente notable el origen, en el cual se hace $\frac{dx}{dt} = 0$. Este punto constituye el punto de equilibrio del sistema, ya que en él se anula la variación de x . Además, de la figura se desprende que el equilibrio es estable.



FIGURA 7.2. Campo vectorial asociado al sistema $dx/dt = -x$

Observando la ecuación que define los puntos de equilibrio, es decir, $f(x) = 0$, inmediatamente surge la sospecha de que si esta función es no lineal podrá tener varias soluciones en x , por lo que podrán coexistir varios puntos de equilibrio. En realidad, no todas las soluciones aportan comportamientos estables a largo plazo, ya que algunas de ellas pueden ser inestables. Pero una vez hecho el análisis de estabilidad de estas soluciones encontraremos un subconjunto de ellas que son estables y que representan los posibles comportamientos del sistema a largo plazo. Estas soluciones de equilibrio estable se conocen también con la denominación de *atractores puntuales* (el punto $x = 0$, en el sistema dinámico (7.2), es un ejemplo de un atractor puntual). Es claro que si el sistema dinámico presenta varios atractores puntuales, ello significa que tiene otros tantos comportamientos estacionarios a largo plazo. Veamos, por ejemplo, el sistema dinámico

$$\frac{dx}{dt} = -(x^3 - x) \quad (7.3)$$

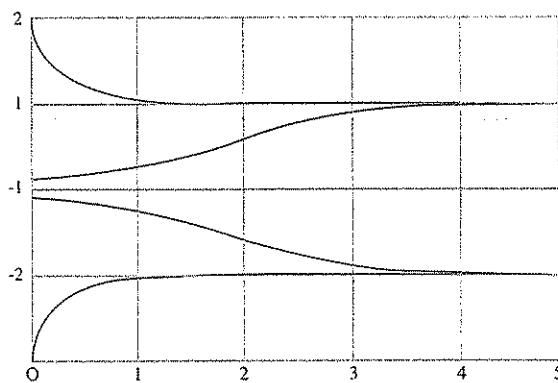
La figura 7.3 muestra el campo vectorial asociado a este sistema dinámico, que tiene dos equilibrios estables en -1 y $+1$, y uno inestable en el origen. La mitad izquierda del espacio de estados (la semirecta negativa) fluye hacia -1 , y la derecha hacia $+1$. Los equilibrios -1 y $+1$ son los atractores del sistema.



FIGURA 7.3. Campo vectorial asociado al sistema $dx/dt = -(x^3 - x)$

A cada atractor se asocia, en el espacio de estados, una *cuenca de atracción*, que está formada por todos aquellos puntos del espacio de estados tales que la trayectoria iniciada en ellos converja al atractor que define la cuenca. En el sistema dinámico (7.3) la cuenca de atracción del atractor -1 es el semieje negativo; y la del atractor $+1$, el semieje positivo. Los dos tipos de comportamiento asociados a cada uno de los atractores se reflejan en la figura 7.4. El espacio de estados se divide en las cuencas de atracción de todos los atractores, separadas entre sí por *separatrices*. En el sistema (7.3) la separatrix es un punto, el origen.

La representación gráfica del espacio de estados con los diferentes atractores, y sus correspondientes cuencas de atracción recibe la denominación de *retrato de estados* o de fases del

FIGURA 7.4. Comportamientos del sistema $dx/dt = -(x^3 - x)$.

sistema dinámico. Constituye una partición del espacio de estados en las cuencas de atracción de los distintos atractores que posee un sistema. Suministra una visión geométrica y global de los modos de comportamiento del sistema, y es el instrumento básico para su análisis cualitativo. Con su ayuda disponemos de un mapa que reúne todos los modos de comportamiento del sistema, organizados mediante sus cuencas de atracción. No nos suministra un comportamiento particular del sistema, sino una visión sintética de todos. Se comprende su carácter de instrumento básico para el análisis cualitativo.

En las figuras 7.2 y 7.3 se muestran los retratos de estados de los sistemas (7.2) y (7.3), respectivamente.

En los ejemplos anteriores se ha supuesto que el espacio de estados era de dimensión 1 (la recta \mathbb{R}). En realidad, el espacio de estados puede ser de cualquier dimensión. Si es de dimensión 2, entonces será el plano \mathbb{R}^2 (en general, una superficie). En cada uno de los puntos de \mathbb{R}^2 se obtiene un vector correspondiente a f . Los puntos en los que $f(x) = 0$ serán los equilibrios del sistema. Análogamente sucede cuando la dimensión del espacio de estados sea superior a 2.

7.1.1. Ciclos límite

En espacios de estados de dimensión mayor que uno, además de los atractores puntuales existen otros atractores más complejos, como los denominados *ciclos límites*, a los que se asocian formas de comportamiento oscilatorio periódico. Estos atractores dan lugar a curvas cerradas en el espacio de estados (mientras que los atractores puntuales eran sencillamente puntos).

En la figura 7.5 se muestra el retrato de estados del sistema dinámico

$$\frac{dx}{dt} = x - y - x(x^2 + y^2)$$

$$\frac{dy}{dt} = x + y - y(x^2 + y^2)$$

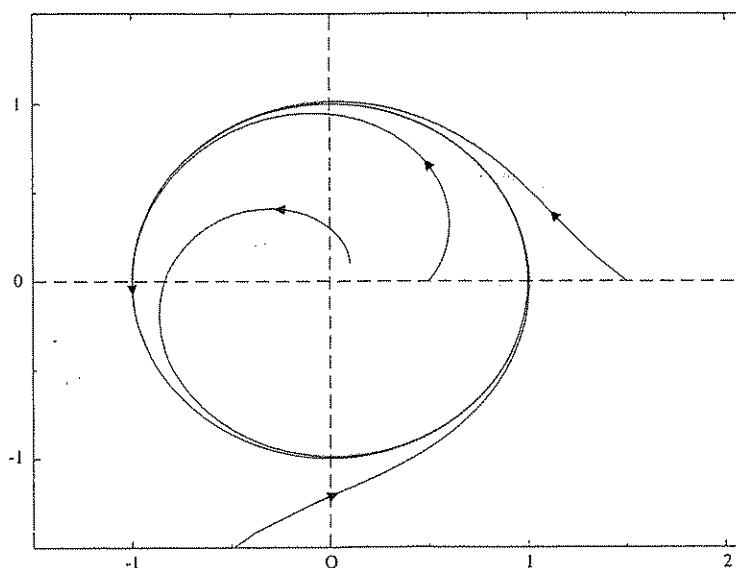


FIGURA 7.5. Retrato de estados de un sistema de dimensión 2, mostrando un ciclo límite.

cuyo atractor es de este tipo; es decir, las trayectorias, con el tiempo, no tienden al equilibrio del sistema (el origen) que es inestable, sino que convergen hacia una circunferencia de radio unidad, en el que permanece la trayectoria una vez que lo alcanza. Esta circunferencia es precisamente el ciclo límite. Las trayectorias de este sistema dinámico no alcanzan nunca el reposo, sino que están oscilando, de forma periódica, indefinidamente. Esta oscilación periódica es la pauta de comportamiento que adquiere el sistema a largo plazo (figura 7.6).

7.1.2. Comportamientos extraños de un sistema determinista

Hasta aquí hemos considerado exclusivamente sistemas dinámicos con atractores puntuales o con ciclos límite. Si con estos atractores se agotase el tema, las cosas serían relativamente sencillas y los modos de comportamiento, aunque variados, perfectamente claros y definibles. Lamentablemente, las cosas no son así. Fue Poincaré quien al estudiar el llamado problema de los tres cuerpos —Sol, Tierra y Luna, por ejemplo— se encontró con resultados sorprendentes. Las soluciones no eran “integrables”. No podía resolver analíticamente el problema. Las trayectorias no se ajustaban a las pautas conocidas que acabamos de ver: los atractores puntuales o los ciclos límite. Las trayectorias, en lugar de converger hacia atractores de estos dos tipos, presentaban una trayectoria retorcidamente compleja, a la que denominó un *enredo homoclinico*, cuya geometría (y la divergencia creciente con el tiempo que mostraba) resultaba difícil de entender.

El trabajo de Poincaré constituye una de las primeras publicaciones en las que pone de manifiesto la existencia de lo que luego se han denominado *atractores extraños*. Estos atractores son regiones del espacio que no son ni puntos, ni curvas, ni superficies, ni ninguna figura geométrica conocida (figura 7.7).

144 Dinámica de sistemas

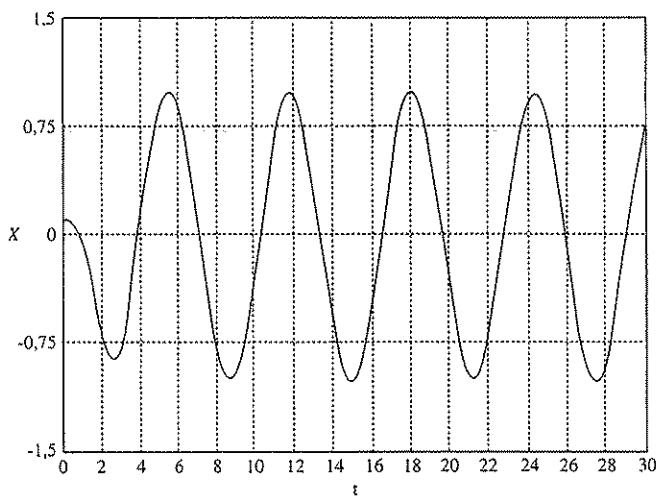


FIGURA 7.6. Comportamiento oscilatorio de un sistema con un atractor tipo ciclo límite.

Para su estudio ha habido que acudir a una nueva clase de objetos geométricos, los *fractales*, de dimensión no entera¹. Se trata de regiones del espacio acotadas en cuyo interior la trayectoria se comporta erráticamente. No es posible discernir ningún tipo de pauta aparente en el comportamiento de un sistema que presenta un atractor de este tipo, más allá del confinamiento de la trayectoria a una cierta región del espacio de estados (el atractor extraño). Aunque de hecho, sí existen algunas pautas que revelan algunas formas de orden en el seno de

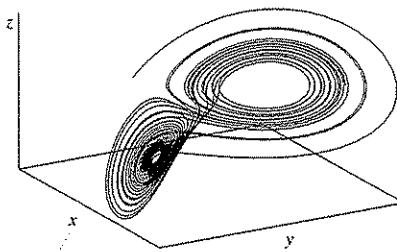


FIGURA 7.7. Ejemplo de atractor extraño (atractor de Lorenz).

objetos tan complejos, sin embargo, estas pautas no son aparentes. Si se determina el espectro de frecuencia de una trayectoria de este tipo, aparecen componentes planas análogas a las que se obtienen si se analiza una señal completamente aleatoria. Lo importante y lo sorprendente es que una señal de este tipo esté generada por un mecanismo al que conceptuamos como determinista.

Pero con ello no se agotan las sorpresas. Un sistema dinámico con este tipo de atractor es de tal modo que a una discrepancia —todo lo pequeña que se quiera— en las condiciones

¹ M. Mandelbrot, 1985, *The fractal geometry of nature*, Freeman.

iniciales responde con trayectorias que son no sólo erráticas, sino crecientemente divergentes. Este fenómeno se conoce como *hipersensibilidad a las condiciones iniciales*. Sus implicaciones para el problema de la causalidad y el determinismo son sustanciales.

El carácter extraño de estos atractores es bien patente: no son ni curvas ni superficies ordinarias, sino objetos de dimensión no entera a los que se conoce como *fractales*. Aunque tengan una dimensión finita, su análisis en el dominio de la frecuencia revela que poseen un espectro de frecuencias casi plano. Y por último, lo que es más notable, presentan lo que hemos denominado hipersensibilidad a las condiciones iniciales. Estas tres propiedades caracterizan a un atractor extraño.

Consideremos uno de los ejemplos recientes que llevaron a restituir actualidad e interés al estudio de los atractores extraños. A principios de los años 60, un meteorólogo llamado Eduard N. Lorenz se propuso estudiar un modelo simplificado de los fenómenos de circulación atmosférica con el fin de mejorar las previsiones meteorológicas. Llegó a establecer un sistema dinámico de dimensión tres que parecía ser el modelo más simple posible que retuviese las características del fenómeno estudiado. Pues bien, este modelo presenta formas de comportamiento caótico del tipo de las encontradas por Poincaré para el problema de los tres cuerpos. Al fenómeno de hipersensibilidad a las condiciones iniciales lo denominó Lorenz de forma muy expresiva: *efecto mariposa*. El aleteo de una mariposa, en cualquier lugar del planeta, representa una variación en las condiciones iniciales que alterará, tarde o temprano, la evolución de las magnitudes asociadas al modelo en cualquier otro lugar de la Tierra. La previsión a largo plazo en meteorología queda con ello profundamente cuestionada.

Este ejemplo nos sitúa de lleno ante la brecha que la aparición de atractores extraños provoca entre causalidad y determinismo. Tradicionalmente estos dos conceptos han estado profundamente vinculados. La posibilidad de explicar un determinado fenómeno, asociándole una estructura causal, parecía que haría posible predeterminar su futuro comportamiento y, de hecho, es lo que sucede si el sistema dinámico que modela al proceso en cuestión presenta atractores puntuales o ciclos límite. En tal caso, el estado inicial —que está situado en una cuenca de atracción— determina completamente el atractor al que evolucionará el sistema. Una pequeña modificación de este estado inicial —tan pequeña que no lo “saque” de la cuenca— no tendrá ningún efecto importante sobre el comportamiento que presentará el sistema a largo plazo. En este sentido decimos que el comportamiento es determinista: está determinado por sus condiciones iniciales. A eso lo llamamos determinismo. La evolución futura es previsible —está determinada— de manera unívoca.

No sucede lo mismo si el sistema presenta un atractor extraño. Aparte del carácter caótico —carente de pautas— de la trayectoria asociada a cada condición inicial, está el hecho de la hipersensibilidad a las condiciones iniciales. Es claro que si éstas se reproducen con precisión infinita, la trayectoria del sistema —aún caótica— será la misma. Pero esta precisión infinita en la medida de las condiciones iniciales es inalcanzable (no sólo porque es imposible *medir* con precisión infinita, sino porque también lo es *manejar* un número con infinitas cifras). Por tanto, con un modelo de un sistema que presente este tipo de atractores no podemos hacer ninguna predicción de su futura evolución, más allá del confinamiento de la trayectoria a la región que define el atractor extraño (aunque sí podamos hacer predicciones a corto plazo; de hecho se puede establecer una relación entre la precisión de las condiciones iniciales y el plazo en el que se pueden hacer razonables predicciones).

El descubrimiento del caos determinista constituye un reto considerable a nuestra visión

convencional del mundo. En los últimos trescientos años se había forjado el presupuesto de que el comportamiento de todo sistema podría predecirse si se conocían las leyes del comportamiento y las condiciones iniciales. Hoy, sin embargo, nos vemos obligados a reemplazar esta concepción del mundo por otra en la que la aleatoriedad y el carácter imprevisible pueden generarse internamente mediante un proceso determinista y no lineal.

El interés de estas cuestiones para la dinámica de sistemas es patente, ya que todo modelo de dinámica de sistemas es un sistema dinámico que eventualmente puede tener un atractor extraño del tipo de los que estamos considerando en esta sección. Sin embargo, en este libro, dado su carácter introductorio, no volveremos a hablar de estos atractores, aunque el lector debe quedar alertado sobre la posibilidad de que su modelo los tenga.

7.2. BIFURCACIONES EN SISTEMAS DINÁMICOS

En todos los modelos que se han visto hasta ahora aparecen magnitudes que toman valores variables (y que hemos clasificado como variables de estado, de flujo o auxiliares) y otras que toman valores constantes a lo largo del proceso de simulación y a las que de una manera genérica nos referimos como parámetros del modelo. Si llamamos a estos parámetros p , entonces un sistema dinámico (7.1) se puede escribir:

$$\frac{dx}{dt} = f(x, p) \quad (7.4)$$

Esta expresión se conoce como *familia de sistemas dinámicos* y representa el modelo genérico de todos los sistemas que comparten una misma estructura (los mismos diagramas de influencias y de Forrester). Para cada valor concreto de los parámetros p se tendrá un determinado modelo. Hemos visto que el retrato de estados suministra una imagen global de los modos de comportamiento de un sistema. La cuestión que surge de inmediato es la siguiente: ¿cómo afectan al retrato de estados las posibles variaciones del parámetro p ?

Para analizar este hecho consideremos, en primer lugar, la familia de sistemas dinámicos:

$$\frac{dx}{dt} = -(x^3 + x - p) \quad (7.5)$$

La figura 7.8 muestra los retratos de estados para distintos valores del parámetro p . En esta figura se tiene una curva que representa el equilibrio del sistema para cada valor de p . En la vertical de cada valor de p se tiene el retrato de estados del sistema dinámico correspondiente a ese valor de p (es decir, cada línea vertical se obtiene girando noventa grados una figura similar a la 7.2, que será el retrato de estado del sistema con el valor correspondiente de p). Se observa que al variar p varía el punto de equilibrio del sistema, pero la forma del retrato de estados no se altera (solamente se desplaza). La curva de la figura 7.8 no es sino la representación de $x^3 + x - p = 0$, y se denomina *curva de equilibrios* de la familia de sistemas dinámicos (7.5). Conviene observar que esta figura se obtiene disponiendo figuras como la 7.2, una al lado de otra, verticalmente, cada una en el lugar que le corresponde según el valor del parámetro p .

Consideremos ahora una nueva familia de sistemas dinámicos:

$$\frac{dx}{dt} = -(x^3 - x - p)$$

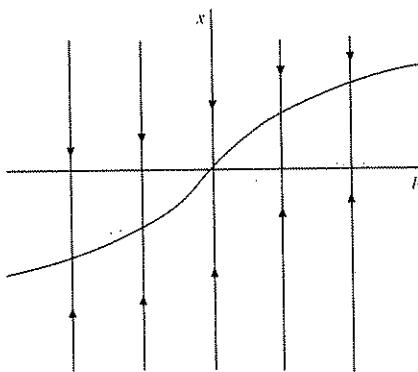


FIGURA 7.8. Diagrama de bifurcaciones de la familia de sistemas dinámicos $dx/dt = -(x^3 + x - p)$.

En la figura 7.9 se ha representado su diagrama de bifurcaciones, y el retrato de estados para los distintos valores del parámetro p . En esta figura se ha adoptado la convención de representar con trazo continuo los equilibrios estables, y con discontinuo los inestables. La característica notable de esta familia es que para diferentes valores del parámetro p el comportamiento del sistema dinámico puede diferir de una manera sustancial. Se observa que si p es menor que p_1 , o mayor que p_2 , el sistema dinámico posee un único atractor (es de la misma forma del que se tenía en la figura 7.2). Sin embargo, si p está comprendido entre p_1 y p_2 , entonces el sistema dinámico posee dos atractores y una separatrix (es del mismo tipo del representado en la figura 7.3).

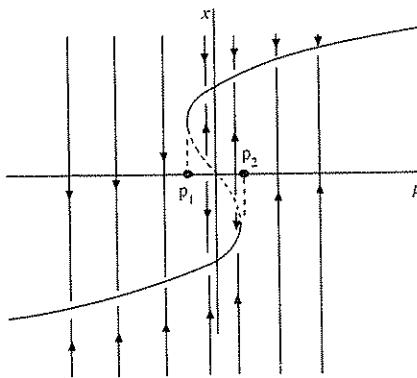


FIGURA 7.9. Diagrama de bifurcaciones de la familia de sistemas dinámicos $dx/dt = -(x^3 - x - p)$.

Para los valores p_1 y p_2 del parámetro p se produce una modificación cualitativa del retrato de estados, ya que el número de atractores pasa de uno a dos, y aparece, en el segundo caso, un punto separatrix. Los valores p_1 y p_2 definen los llamados *puntos de bifurcación*

del sistema dinámico, y constituyen un ejemplo de las bifurcaciones más elementales que se pueden presentar en un sistema dinámico: las *catástrofes de pliegue*, empleando la terminología de la teoría de catástrofes, o los *puntos límites*, en la jerga de la teoría de bifurcaciones. Los puntos de bifurcación en este caso concreto tienen la notable propiedad de que si se varía el parámetro p —de modo que cruce uno de los valores p_1 ó p_2 — se produce la aparición o desaparición de atractores del sistema; es decir, se ramifican los comportamientos. Por ejemplo, si atravesamos p_1 en sentido creciente del parámetro p aparece (de la nada) un par atractor-repulsor. Lo opuesto ocurre en p_2 . Resulta claro que la estructura topológica del retrato de estados (y por tanto los modos de comportamiento que presenta el sistema) se modifica al atravesar un punto de bifurcación.

Es importante observar que el diagrama de bifurcaciones suministra una perspectiva global con relación a todos los modos de comportamiento que pueda presentar una familia de sistemas dinámicos (7.4). Al mismo tiempo suministra información con relación a los valores de los parámetros p , para los que el comportamiento del sistema será robusto ante variaciones de los valores de estos parámetros. En efecto, para valores de p alejados de los puntos de bifurcación p_1 ó p_2 el comportamiento cualitativo del sistema será insensible a variaciones de p . Por el contrario, para valores de p cercanos a p_1 o a p_2 , las variaciones de p pueden hacer que se atraviese el punto de bifurcación y que se produzca la alteración del retrato de estados del sistema, con la consiguiente alteración de sus modos de comportamiento. Por tanto, el diagrama de bifurcaciones suministra información respecto a los valores del parámetro p para los que el sistema muestra una especial sensibilidad con relación a los modos de comportamiento.

El caso que hemos considerado es el más sencillo posible, ya que en él sólo están involucrados atractores puntuales. Otra bifurcación de gran interés es la llamada *bifurcación de Hopf*, en la que un atractor puntual se transforma en un ciclo límite (figura 7.10).

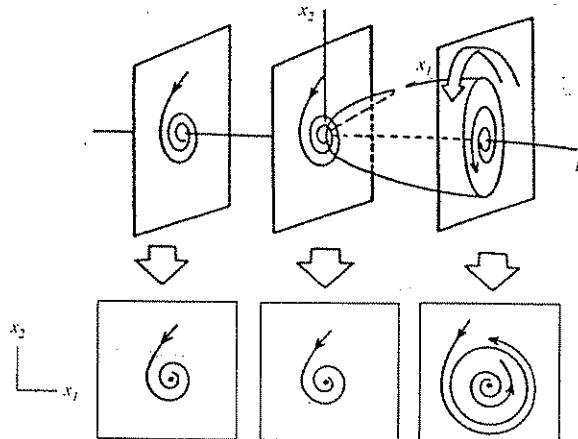


FIGURA 7.10. *Bifurcación de Hopf*.

En esta bifurcación se produce un cambio del tipo de atractor que pasa de ser un punto de equilibrio a un ciclo límite; es decir, de un atractor estable a uno dinámico. En ello resi-

de precisamente su interés. Se trata de la bifurcación más sencilla en la que se produce un fenómeno de este tipo.

Además de los dos tipos de bifurcaciones que acabamos de ver, existe todo un catálogo de bifurcaciones correspondientes a las posibles transiciones entre los distintos tipos de atractores². El estudio de las bifurcaciones es un campo de investigación al que se ha dedicado un gran esfuerzo en estos últimos años. En la actualidad se empieza a disponer de todo un cuerpo de resultados a los que se puede recurrir en aplicaciones prácticas del tipo de las que aquí se están comentando.

7.2.1. Análisis cualitativo del modelo de un área industrial

El modelo de evolución de la actividad económica en un área industrial, que hemos visto en la sección 5.3, nos va a permitir ilustrar con un ejemplo concreto las posibilidades y el interés del análisis cualitativo. Este modelo se puede escribir, mediante notación matemática convencional, recordando (6.3), de la forma siguiente:

$$\frac{dx}{dt} = nx f(kx) - mx = x[nf(kx) - m] \quad (7.6)$$

Estamos, por tanto, ante un sistema dinámico con una variable de estado x (las empresas del polígono industrial) y los parámetros n , k y m .

Para realizar el análisis cualitativo de este modelo se dibuja, en primer lugar, en el espacio $X \times P$ el diagrama de bifurcaciones del sistema. Este lugar vendrá dado por las soluciones en x de la ecuación,

$$x[nf(kx) - m] = 0$$

Las soluciones resultan ser:

$$x = 0 \quad (7.7)$$

$$f(kx) = \frac{m}{n} = p \quad (7.8)$$

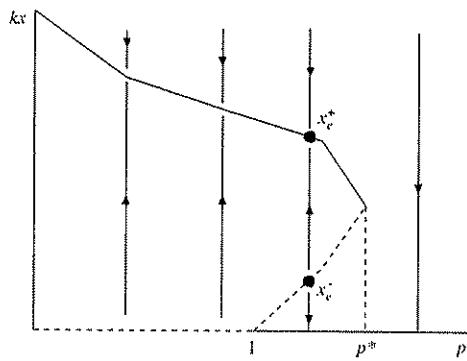
Esta última expresión equivale a

$$kx = f^{-1}(p) \quad (7.9)$$

por tanto, el diagrama de bifurcaciones estará formado en el plano (x, p) por el eje $x = 0$ y por los valores de x que satisfacen la ecuación (7.9). La inversa de la función f se obtiene gráficamente reflejando la curva correspondiente en torno a la bisectriz del primer cuadrante en el plano (x, p) . Para algunos valores de p la expresión (7.9) da, además de $x = 0$, un único valor de x : x_e^+ , mientras que para otros dos, x_e^+ y x_e^- . De este modo se llega al diagrama de bifurcaciones de la figura 7.11 (en realidad esta figura no representa exactamente el diagrama de bifurcaciones, ya que se ha tomado kx como variable de ordenadas en lugar de x ; dado que $k > 0$, el diagrama de bifurcaciones real será cualitativamente el mismo)³.

² J. Guckenheimer y P. Holmes, 1983, *Nonlinear oscillations, dynamical systems and bifurcations of vector fields*, Springer-Verlag; R. H. Abraham y C. D. Shaw, 1987, "Dynamics: A Visual Introduction", en F. E. Yates (ed.), *Self-Organizing Systems*, Plenum Press.

³ Puede sorprender el carácter abrupto de la curva de la figura 7.11, que viene determinado por la definición de la función f por medio de la tabla 6.1 y por el método de interpolación que se adopte; aquí se ha optado por una interpolación lineal que es la que realizan la mayoría de los lenguajes de simulación.

FIGURA 7.11. *Diagrama de bifurcaciones del modelo de un área industrial.*

La observación de esta figura conduce a dividir los valores de p en tres regiones con diferente comportamiento cualitativo: para $p \in [0, 1)$ existe, además de $x = 0$, un equilibrio adicional x_e^+ ; para $p \in (1, p^*)$, siendo p^* el valor máximo de f , aparece otro equilibrio x_e^- ; y para $p > p^*$ no hay ningún equilibrio adicional debido a (7.9). Es interesante observar que para $p = p^*$ se produce la coalescencia de los equilibrios x_e^+ y x_e^- que desaparecen.

De acuerdo con la figura 7.11 se observan dos puntos de bifurcación para p . Para $p = 1$ se tiene una bifurcación (que se conoce como bifurcación transcritica) y se produce el cambio de las estabilidades de las ramas que se cortan. Por otra parte, para $p = p^*$ se llega a una bifurcación del tipo pliegue o punto límite que se ha visto anteriormente. La figura 7.11 recuerda a la 7.9. Consideraciones análogas a las que se hicieron allí se aplican aquí. Como allí, se pueden dar tres tipos de situaciones que se reflejan en la figura 7.12. En la parte inferior de esta figura aparecen las trayectorias seguidas por la variable número de empresas en cada uno de los tipos de comportamiento; en la parte superior se dibuja esta trayectoria en un plano perpendicular a un nuevo eje asociado al parámetro p . Los tres tipos de comportamiento se dan según:

1. $p < 1$, en cuyo caso existe un único equilibrio estable (figura 7.12a);
2. $p^* < p$, en cuyo caso no existe ningún equilibrio del tipo (7.8), siendo el único equilibrio $x = 0$ (figura 7.12b);
3. $1 < p < p^*$, en cuyo caso existen dos equilibrios estables (figura 7.12c).

En el caso 1 se produce el crecimiento económico del área industrial, ya que el equilibrio al que tiende la trayectoria es del tipo (7.8), mientras que en el 3 se produce su declive, ya que el único equilibrio posible es $x = 0$ (actividad económica inexistente). Pero además, los dos modos de comportamiento pueden presentarse en un mismo modelo; es decir, para un mismo valor del parámetro p . Precisamente esto es lo que sucede cuando este parámetro está comprendido entre 1 y p^* , en cuyo caso el crecimiento económico o el declive del área urbana dependerá de las condiciones iniciales a partir de las cuales se desencadene el proceso. Estas condiciones iniciales vendrán dadas por la política de “cebado” con la que se pretenda promocionar el desarrollo del área urbana en cuestión. Si esta política de promoción no alcanza

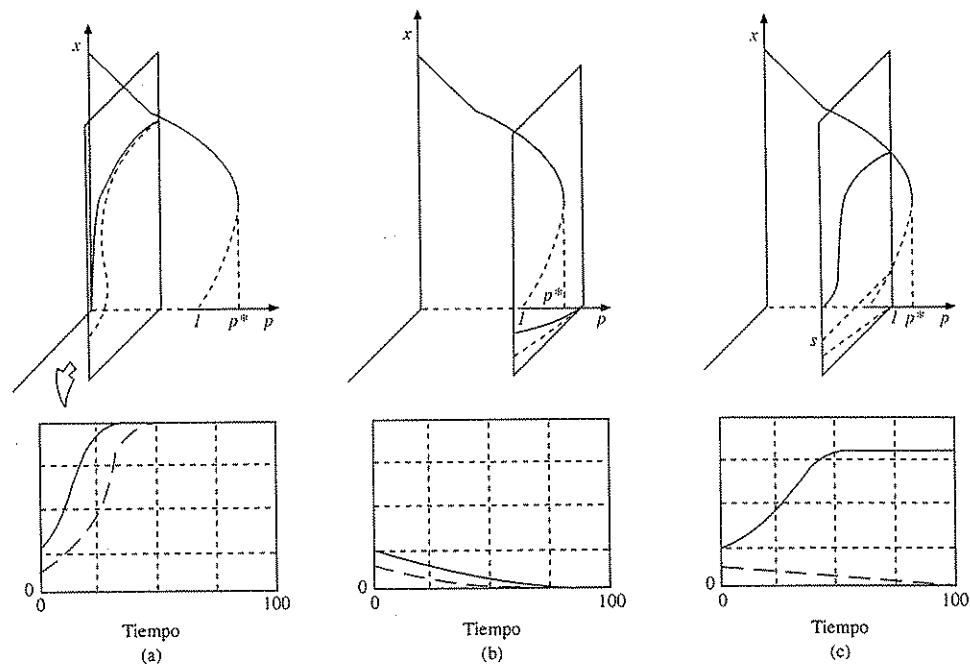


FIGURA 7.12. Modos de comportamiento del modelo de un área industrial.

un determinado umbral, el desarrollo de la actividad económica del área urbana está condenado al fracaso. Éste era el caso estudiado en el análisis de sensibilidad de la sección 6.3.2.

Resumiendo, podemos decir que el sistema presenta dos modos de comportamiento cualitativamente diferenciados:

- o bien se produce el crecimiento de la actividad económica del área urbana;
- o bien tiene lugar su declive hasta su extinción.

Los dos comportamientos pueden comprobarse por simulación —recuérdese la figura 6.2— si se asignan a los parámetros y a las condiciones iniciales valores apropiados. Estos valores se reflejan en el diagrama de bifurcaciones. En este sentido se dice que el análisis cualitativo suministra una guía para la simulación de un modelo. Conviene observar que los dos modos de comportamiento los genera la misma estructura (el mismo modelo mental). Se ilustra con ello el hecho que se anunciaba al principio de este capítulo: un mismo modelo puede generar varios comportamientos.

Conviene también observar que la existencia de estos dos modos de comportamiento depende de información cualitativa contenida en el modelo. Es decir, en el anterior análisis lo único realmente relevante es, aparte de la forma matemática de la ecuación (7.6), el hecho de que la función f tuviese la forma cualitativa que posee: creciente al principio para decrecer luego (recuérdese la figura 5.4). Esta información es la contenida en el conjunto de enunciados R_i de la sección 5.3.1. Es decir, de los enunciados que representan una caracterización muy

152 *Dinámica de sistemas*

Este capítulo es una adaptación del trabajo de A. Gómez-Pérez, "Dinámica de sistemas y desarrollo urbano: un análisis cualitativo", en el que se analiza la dinámica de sistemas en el desarrollo urbano. El autor es profesor de la Facultad de Ciencias Políticas y Sociales de la Universidad de Valencia.

general del proceso de desarrollo de la actividad económica en un área urbana se desprenden los dos modos de comportamiento anteriores: que el área urbana se puede desarrollar o no. Que se produzca uno u otro comportamiento dependerá de los valores concretos que tomen determinados parámetros.

Se comprende que la información que suministra el análisis cualitativo de un modelo es de un valor inapreciable para adquirir una perspectiva global con relación a los modos de comportamiento compatibles con el modelo mental a partir del cual se ha construido el modelo. Sin embargo, antes de concluir el capítulo hay que añadir que la posibilidad de llevar a cabo este tipo de análisis se complica en modelos de gran dimensión.

8.1. ESTRUCTURAS JERARQUIZADAS

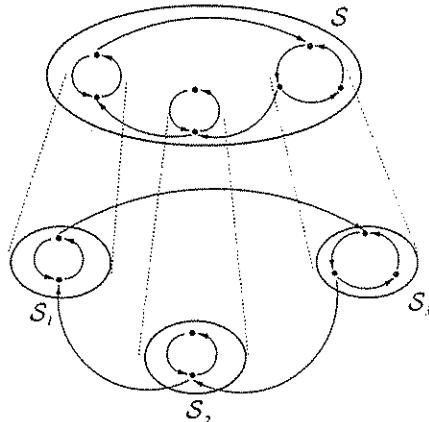
Los métodos de análisis de modelos que se han visto en el capítulo anterior funcionan razonablemente bien cuando la dimensión de los mismos no es excesivamente grande. Sin embargo, cuando su estructura alcanza un cierto nivel de complejidad no es fácil que, por esta misma complejidad, el modelo nos ayude a entender fácilmente qué sucede en el sistema. En tal caso puede servir de ayuda el organizar la estructura global del modelo en subestructuras jerarquizadas entre sí, de modo que la complejidad del modelo (y por ende, la del propio sistema) se subdivida, por decirlo así, en subcomplejidades debidamente coordinadas.

Se recurre entonces a los resultados de la *teoría de los sistemas jerarquizados* mediante la cual un sistema de una cierta complejidad se descompone en subestructuras que, a partir de su adecuada coordinación, restituyen la complejidad original del sistema, pero de modo que estas subestructuras están dotadas de suficiente sentido en sí mismas como para contribuir a la comprensión del comportamiento global. En lo que sigue hablaremos indistintamente de las complejidad del sistema y de su modelo (repárese en que al hablar de un sistema siempre estamos hablando de *una* descripción de ese sistema).

En la figura 8.1 se ilustra gráficamente la descomposición de un sistema complejo S en tres subsistemas S_1 , S_2 y S_3 coordinados entre sí. Cada uno de los subsistemas posee la suficiente entidad como para poder ser objeto de estudio por sí mismo (ello, claro está, sin perder de vista lo que se decía más arriba con relación a las partes de un sistema, cuya consideración no tiene sentido más que en la medida en que forman parte del propio sistema). Un subsistema es, por tanto, un sistema al que, bajo la óptica del sistema global, consideramos como una parte, pero que en sí mismo es, a su vez, un sistema (está dotado de una estructura, una composición y un entorno).

El subsistema es el componente esencial de un sistema jerarquizado. Por ello Arthur Koestler propuso la denominación de *holon* para referirse a los subsistemas¹. Con esta denominación aludía a su doble carácter de sistema —todo, *holos*— y de partícula elemental —por ello la terminación *on*.

¹ A. Koestler y J.R. Smyties, 1969, *Beyond reductionism*, Hutchinson, Londres.

FIGURA 8.1. *Estructura jerarquizada.*

La descomposición de un sistema complejo en subsistemas puede hacerse de acuerdo con diferentes criterios de organización. Se puede, por ejemplo, dividir un sistema complejo en subsistemas por criterios de proximidad. Otro procedimiento, considerado uno de los más interesantes y sugestivos, consiste en contemplar las diferentes *escalas de tiempo* en las que se desenvuelven los diferentes procesos que tienen lugar en el sistema complejo. En el caso más simple se puede considerar que al menos existen dos escalas de tiempo en un proceso complejo: una *rápida* y otra *lenta*². Por una parte, estarían los profundos cambios estructurales que se desarrollan a largo plazo. La evolución de estos procesos se produce en lo que llamamos la *escala de tiempos lenta*. Mientras que, por otra, tienen lugar los fenómenos más o menos coyunturales que corresponden al desenvolvimiento del sistema a corto plazo, y que se desarrollan en lo que llamamos *escala rápida*. Se trata, en principio, de una descomposición de carácter intuitivo. Afortunadamente, dentro de la teoría de sistemas dinámicos existe un amplio capítulo dedicado a los sistemas dinámicos singularmente perturbados que aporta los útiles matemáticos necesarios para el estudio de los sistemas con dos escalas de tiempo.

8.1.1. Sistemas dinámicos con dos escalas de tiempo

Un sistema dinámico singularmente perturbado es aquel que se puede escribir como:

$$\frac{dx}{dt} = f(x, y) \quad (8.1)$$

$$\varepsilon \frac{dy}{dt} = g(x, y) \quad (8.2)$$

siendo ε un parámetro cuyo valor es pequeño y el vector (x, y) el estado. La separación, por reagrupamiento, de los distintos componentes del vector de estado en x e y se hace por algún criterio adecuado al caso.

² H. Simon, 1981, *The Sciences of the Artificial*, M.I.T., Cambridge, Mass.; cap. 7.

Para simplificar el estudio de un sistema dinámico tal como el (8.1-8.2) se emplean dos sistemas dinámicos auxiliares:

1. Un sistema dinámico, llamado *sistema degenerado* de (8.1-8.2), que se obtiene haciendo $\varepsilon = 0$ en (8.1-8.2), con lo que se obtiene

$$\frac{dx}{dt} = f(x, y) \quad (8.3)$$

$$0 = g(x, y) \quad (8.4)$$

Al conjunto $S = \{(x, y) : g(x, y) = 0\}$ se le llama *espacio de configuración*. El comportamiento del sistema (8.3-8.4) puede considerarse regido por (8.3) en el espacio de configuración definido por (8.4). Geométricamente se puede interpretar como se hace en la figura 8.2. Obsérvese que de (8.4) eventualmente se puede eliminar y , obteniéndose $y = \phi(x)$, que llevado a (8.3), resulta

$$\frac{dx}{dt} = f(x, y) = f[x, \phi(x)] = \Phi(x)$$

El sistema dinámico degenerado es un sistema dinámico cuya dimensión es igual a la del subespacio $\{x\}$.

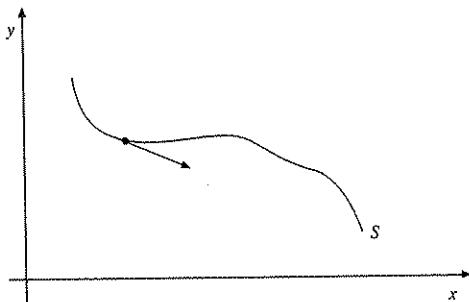


FIGURA 8.2. *Espacio de configuración de un sistema con dos escalas de tiempo.*

Puesto que el comportamiento tiene lugar a lo largo del espacio de configuración, la cuestión que se suscita inmediatamente es la de qué sucede si las condiciones iniciales $[x(0), y(0)]$ no están en dicho espacio. En tal caso hay que recurrir al segundo de los sistemas dinámicos asociados a (8.1-8.2) que anunciábamos más arriba.

2. Supongamos que en (8.1-8.2) se hace el cambio de variable tiempo $\tau = t/\varepsilon$, lo que implica que τ es una medida del tiempo a pequeña escala (por ejemplo, si t mide el tiempo en segundos y $\varepsilon = 10^{-3}$, entonces τ tendrá unidades de milésimas de segundo). Con ello se obtiene

$$\frac{dx}{d\tau} = \varepsilon f(x, y) \quad (8.5)$$

$$\frac{dy}{d\tau} = g(x, y) \quad (8.6)$$

Si en estas ecuaciones se hace $\varepsilon = 0$ se convierten en

$$\begin{aligned}\frac{dx}{d\tau} &= 0 \\ \frac{dy}{d\tau} &= g(x, y)\end{aligned}\tag{8.7}$$

que es el nuevo sistema dinámico aproximado asociado a (8.1-8.2). En este sistema está claro que

$$x = x_0 = x(0) = \text{constante}$$

$$\frac{dy}{d\tau} = g(x_0, y)$$

Debe notarse que el comportamiento de este segundo sistema es $1/\varepsilon$ veces más rápido que el correspondiente al sistema (8.3-8.4). Es decir, el comportamiento de este segundo sistema es un comportamiento *rápido* del sistema (8.1-8.2), mientras que al sistema (8.3-8.4) cabe considerarlo como un comportamiento *lento*.

Por tanto, el sistema original (8.1-8.2) se ha dissociado en dos comportamientos: uno lento y otro rápido. La yuxtaposición de estos dos comportamientos para realizar una aproximación del comportamiento original de (8.1-8.2) se puede interpretar gráficamente tal como se hace en la figura 8.3. Es decir, supuestas unas condiciones iniciales $x(0), y(0)$ que no se encuentren en el espacio de configuración, la trayectoria desde este punto se inicia con un movimiento rápido, en el que $x = x_0 = x(0) = \text{constante}$. Este movimiento rápido es tal que sólo varía y , de acuerdo con la ecuación (8.7). Los equilibrios de este movimiento rápido se hallan situados en el espacio de configuración, por lo que el movimiento rápido tiende precisamente a dicho espacio. Una vez que se alcanza, cesa el movimiento rápido y la trayectoria discurre sobre el espacio de configuración con movimiento lento.

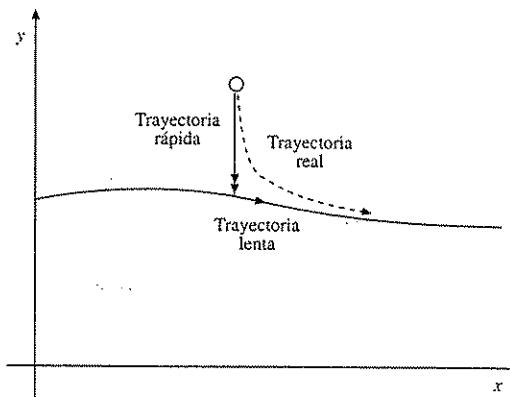


FIGURA 8.3. *Comportamiento de un sistema con dos escalas de tiempo.*

Las ventajas del planteamiento anterior son dobles. Por una parte, el estudio de un sistema de cierta dimensión se descompone en el estudio de dos subsistemas de menor dimensión (que

a su vez se pueden separar cada uno de ellos en otros dos, y así sucesivamente). Por otra parte, las trayectorias asociadas a cada uno de estos subsistemas tienen una interpretación geométrica, de acuerdo con lo que se ha indicado en la figura 8.3. En esta interpretación tiene un papel relevante la *geometría* del espacio de configuración S . En la figura 8.3 se ha supuesto que este espacio de configuración es monótono. Sin embargo, si este espacio presenta *pliegues*, tal como se indica en la figura 8.4, entonces se dan curiosos fenómenos de *cambio abrupto* que tienen una especial significación. Este modelado de los cambios abruptos establece una conexión de las ideas que aquí se están exponiendo con la *teoría de catástrofes*³.

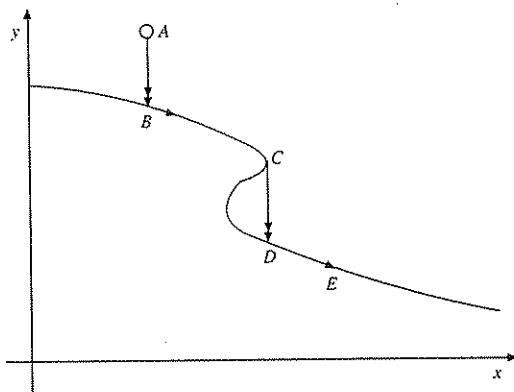


FIGURA 8.4. *Cambio abrupto en un sistema con dos escalas de tiempo.*

En efecto, suponiendo que la trayectoria se inicia en A , aparece, de acuerdo con lo visto anteriormente, una trayectoria rápida hasta alcanzar el espacio de configuración en B . Seguidamente la trayectoria discurre por el espacio de configuración, con un movimiento lento, desde B hasta C . Pero al llegar a este último punto el espacio de configuración se convierte en inestable para el movimiento lento, por lo que la trayectoria se ve impelida —digámoslo así— a abandonarlo, quedando sometida a un movimiento rápido, análogo al que se vio sometida en A , hasta llegar a una nueva porción estable del espacio de configuración, lo que sucede al alcanzar el punto D , a partir del cual la trayectoria vuelve a discurrir lentamente hacia el punto E .

Observamos que en el punto C se ha producido un “salto” de la trayectoria hasta el punto D . Este “salto” representa un cambio aparentemente abrupto en la evolución del sistema. Este cambio tiene particular interés cuando se trata de modelar cambios cualitativos sustanciales en el comportamiento de un sistema.

Veamos un ejemplo elemental en el que se pone de manifiesto el proceso anterior. Este ejemplo constituye además una muestra de la aplicación de lo que se acaba de ver a la dinámica de sistemas. Recordemos el modelo de un área industrial, que viene dado por la expresión (6.3)

$$\frac{dx}{dt} = x[nf(kx) - m] = nx[f(kx) - p] \quad (8.8)$$

³ R. Thom, 1977, *Stabilité structurelle et morphogenèse*, InterEditions; E. C. Zeeman, 1977, *Catastrophe Theory*, Addison Wesley.

donde $p = m/n$. Supóngase ahora que, por las circunstancias que sean, el parámetro m varíe lentamente, de modo que el parámetro p crezca con el tiempo muy lentamente y de manera uniforme. Formalmente esto último se puede escribir:

$$\frac{dp}{d\tau} = \varepsilon > 0 \quad (8.9)$$

Esta hipótesis de crecimiento lento del parámetro p tiene un interés fundamentalmente ilustrativo de las ideas que estamos exponiendo, que pueden interpretarse simplemente considerando que la tasa de desaparición de empresas crece a lo largo del tiempo. Por tanto, se está contemplando una situación de deterioro, muy lento pero progresivo, del área industrial. Las ecuaciones del nuevo modelo estarán formadas por la (8.8) y la (8.9), que, escritas conjuntamente, conducen a

$$\begin{aligned} \frac{dx}{d\tau} &= nx(f(kx) - p) \\ \frac{dp}{d\tau} &= \varepsilon \end{aligned}$$

Estas ecuaciones poseen precisamente la misma forma que las (8.5) y (8.6) anteriores⁴. Se tiene, por tanto, un sistema con dos escalas de tiempo al que se puede aplicar lo dicho hasta aquí.

El espacio de configuración para este sistema vendrá dado por la curva representada en la figura 8.5⁵. El comportamiento del sistema, en este caso, se muestra en la figura 8.6. Supóngase que la evolución del área industrial se inicia en un punto tal como el A . Se produce un proceso de crecimiento sigmoidal, relativamente rápido, tendiendo a alcanzar el punto B . Este comportamiento es, esencialmente, el considerado en los capítulos anteriores.

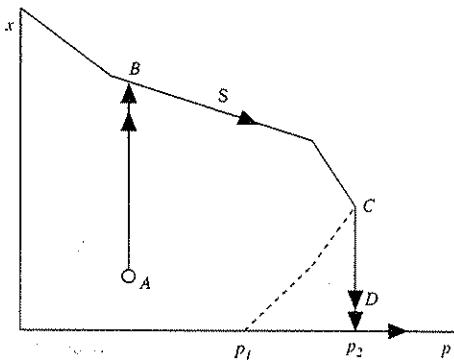


FIGURA 8.5. *Espacio de configuración del modelo de un área industrial con dos escalas de tiempo.*

⁴ Sin embargo, la notación adoptada puede llevar a confusión: obsérvese que en este caso x es la variable rápida, mientras que en el sistema (8.5-8.6) x era la variable lenta.

⁵ Se pueden hacer aquí los mismos comentarios que se hicieron en la nota 3 del capítulo 7.

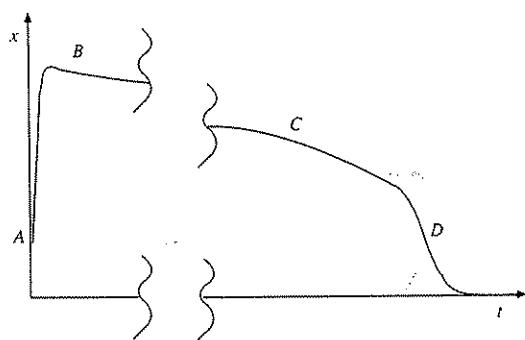


FIGURA 8.6. Comportamiento del modelo de un área industrial con dos escalas de tiempo.

Al mismo tiempo, lentamente, se va produciendo un proceso de crecimiento del parámetro p , por lo que el equilibrio al que tiende el número de empresas se va desplazando desde B a C (el salto que aparece en la figura se debe a que el tiempo que tarda el sistema en evolucionar de B a C es muy grande comparado con los demás tiempos característicos, por lo que se ha optado por no representar gran parte de esta evolución). Este desplazamiento se traduce en una lenta disminución de la tendencia al crecimiento del número de empresas. Pero obsérvese que, pese a este lento decrecimiento tendencial, la actividad industrial se mantiene. Hay una tendencia al declive en el área industrial, pero su percepción es muy lenta.

Sin embargo, al alcanzar el punto C se produce un cambio súbito en las expectativas de evolución del sistema, de modo que en un periodo relativamente corto se produce el salto de C a D , y la consiguiente desaparición del área industrial. Un proceso de decrecimiento lento del área industrial se convierte de forma súbita en un colapso total. Además, el proceso adquiere un carácter adicional de irreversibilidad.

La única forma de evitar el colapso del área industrial consiste en modificar la variación de p de modo que el equilibrio se mantenga en el tramo superior (región del espacio de configuración entre B y C); es decir, que p no crezca más allá de p_2 .

Éste es un ejemplo muy simple pero que pone de manifiesto la enorme riqueza de posibilidades que presenta el análisis de un sistema complejo que presente dos escalas de tiempo. En él resultan engarzados los dos modos de comportamiento (crecimiento y declive) a los que se ha aludido en el apartado 7.2.1. Conviene resaltar que el cambio súbito producido en el modelo está intimamente asociado a la presencia en él de mecanismos de interacción no lineales. En un sistema lineal no puede producirse el tipo de situación que se está analizando.

8.2. LA ESTRUCTURA DE REACCIÓN-DIFUSIÓN Y LA MORFOGÉNESIS

El corto catálogo de estructuras generales que hasta la fecha han sido estudiadas se ha enriquecido con una de especial interés: la estructura de *reacción-difusión*. Su descubrimiento corresponde a Alan M. Turing que la propuso como mecanismo básico para dar cuenta de los

fenómenos morfogenéticos⁶. El mecanismo correspondiente es de una sorprendente simplicidad y se conoce como estructura de reacción-difusión.

Las ideas originales de Turing fueron posteriormente reelaboradas por Ilya Prigogine y su escuela de Bruselas⁷, quienes les dieron un considerable contenido termodinámico. Es sabido que la segunda ley de la termodinámica establece que la evolución del Universo tiende hacia la desorganización progresiva. Por el contrario, en nuestro entorno se manifiestan fenómenos de signo contrario, de aparición de formas espontáneas de autoorganización, de las cuales las más patentes son los seres vivos y las formas de organización social. Esta tendencia parece contradecir los supuestos de la segunda ley de la termodinámica. Corresponde a Prigogine el haber reconciliado ambos conceptos, partiendo de la estructura de reacción-difusión de Turing.

Según el segundo principio de la termodinámica, todo sistema físico aislado tiende a evolucionar de modo que en él se produzca un incremento del desorden. Situaciones previamente ordenadas, dejando fluir libremente al tiempo, se convierten en situaciones progresivamente desordenadas. Este proceso de evolución desde el orden al desorden lo podemos representar de forma muy esquemática como se hace en la figura 8.7. En esta figura se muestra un sistema formado por dos pequeñas celdas comunicadas entre sí, en las que originariamente se suponen dispuestas moléculas del mismo tipo en cada una de ellas (en una moléculas "aspas" y en la otra "cruces").

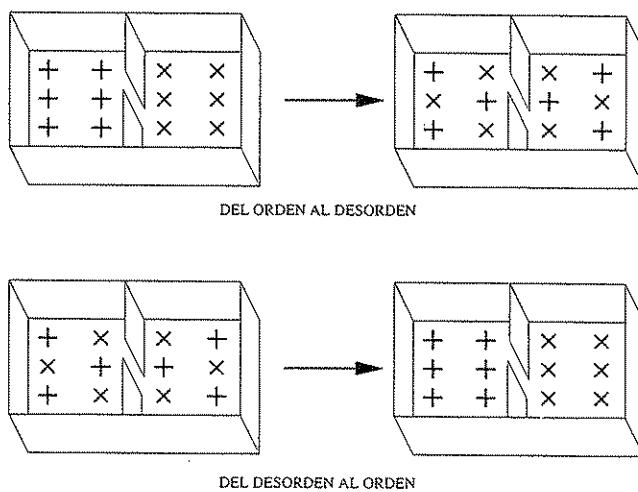


FIGURA 8.7. Evolución del orden al desorden y del desorden al orden.

El proceso que se produce al cabo de un tiempo es que la ordenación inicial de las moléculas desaparece y tenemos, en cada celda, tantas de un tipo como del otro. Es decir, una situación inicialmente ordenada (cada tipo de moléculas en una celda) se convierte en una situación completamente desordenada (hay las mismas moléculas de los dos tipos en ambas celdas). Este proceso nos suministra una ilustración elemental de lo que se esconde detrás del

⁶ A. M. Turing, 1952, *The chemical basis of morphogenesis*, Phil. Trans. R. Soc. B 237, 37-72.

⁷ I. Prigogine, 1980, *Physique, temps et devenir*, Mason; I. Prigogine, 1983, *La nueva alianza*, Alianza Editorial; G. Nicolis e I. Prigogine, 1977, *Self-Organization in Nonequilibrium Systems*, Wiley, Nueva York.

segundo principio de la termodinámica, cuando se dice que un sistema evoluciona libremente del orden al desorden.

A Prigogine, sin embargo, le intrigó extremadamente un proceso del mismo tipo que el anterior, pero de sentido contrario, que se manifiesta con profusión en la naturaleza. Aunque el segundo principio nos diga que la evolución de la realidad tiende al desorden progresivo, sin embargo, observamos en nuestro entorno fenómenos de signo contrario, mediante los cuales se está produciendo continuamente una emergencia de orden. La vida, las formas de organización social, son manifestaciones de un principio ordenador que parece contravenir lo establecido en el segundo principio de la termodinámica. ¿Cómo es posible que, a lo largo de la evolución de los seres vivos, se haya producido la emergencia de formas cada vez más complejas en las que subyace un mayor orden?

8.2.1. Mecanismo morfogenético de Turing

Para responder a esta cuestión cabe recurrir al mecanismo de reacción-difusión debido a Turing. A este último le había interesado un problema análogo al que interesaría a Prigogine. Se trata del problema de la morfogénesis. En el proceso de división celular, ¿cómo es posible que células inicialmente homogéneas den lugar a órganos diferenciados? ¿Cómo se genera la aparición de diversidad en algo que es inicialmente homogéneo? Para explicarlo Turing ideó un mecanismo extremadamente simple. Este mecanismo es suficientemente sencillo como para que nos detengamos a considerarlo con algún detalle.

Situándonos en el caso más simple, considérese un sistema formado por dos células. En cada una se considera que su estado está representado por el valor que alcanza una variable x —denominada por Turing *morfogen*— que representa, por ejemplo, la concentración de una determinada hormona en dicha célula (responsable de la función que se asignará posteriormente al órgano que resulte de la evolución posterior de esta célula). El comportamiento dinámico del morfogén x viene dado por un sistema dinámico lineal de primer orden, como es:

$$\frac{dx}{dt} = -ax + u \quad (8.10)$$

En donde a es un parámetro constitutivo del sistema y u es la señal de mando o control mediante la cual se actúa sobre las células (por ejemplo, la inyección de sustancias exteriores). Esta ecuación diferencial lineal representa el sistema dinámico más sencillo que se puede concebir. El comportamiento de x a lo largo del tiempo cuando u es constante viene representado en la figura 8.8, en la que se pone de manifiesto que la concentración del morfogén x evoluciona hacia una posición de equilibrio $x^e = u/a$.

El sistema dinámico (8.10), representa el comportamiento de una de las células. Hemos supuesto, sin embargo, que tenemos dos de ellas. A cada una podemos asociar una ecuación como la (8.10) con lo que, si además hacemos la convención de denominar x_1 a la concentración del morfogén en la célula 1, y x_2 a la correspondiente a la célula 2, las ecuaciones serán:

$$\begin{aligned} \frac{dx_1}{dt} &= -ax_1 + u \\ \frac{dx_2}{dt} &= -ax_2 + u \end{aligned}$$

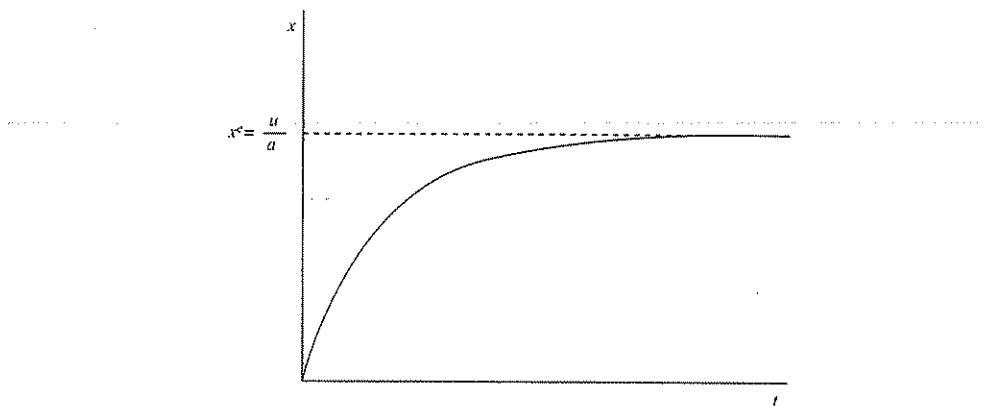


FIGURA 8.8. Comportamiento del morfogén en una célula aislada.

En estas ecuaciones se manifiesta un comportamiento autónomo para cada una de las células. Supongamos ahora que esto no es así, sino que ambas células están vinculadas entre sí por un fenómeno de difusión, mediante el cual se produce el transporte de morfogén x de una a otra, dependiendo de la concentración que tenga cada una de ellas. En el caso más sencillo, supongamos que el fenómeno de difusión viene regulado por una ley lineal, de modo que se puede representar por un término proporcional, mediante un factor de difusión D , a la diferencia de concentraciones entre ambas células. De este modo las ecuaciones del proceso se convierten en:

$$\frac{dx_1}{dt} = -ax_1 + D(x_1 - x_2) + u$$

$$\frac{dx_2}{dt} = -ax_2 + D(x_2 - x_1) + u$$

Empleando notación matricial, el anterior sistema de ecuaciones se puede escribir:

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -(a-D) & -D \\ -D & -(a-D) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \quad (8.11)$$

Los equilibrios de este sistema dinámico, que se obtienen haciendo en (8.11) $dx_1/dt = dx_2/dt = 0$, resultan ser los mismos que se tienen cuando las dos células están desvinculadas entre sí; es decir, $x_1^e = x_2^e = u/a$. Por tanto, la introducción del término de difusión no afecta a los equilibrios del sistema. Sí lo hace, sin embargo, a su estabilidad. Para estudiarla se recurre a la ecuación característica del sistema (8.11), que resulta ser

$$[s + (a - D)]^2 - D^2 = 0$$

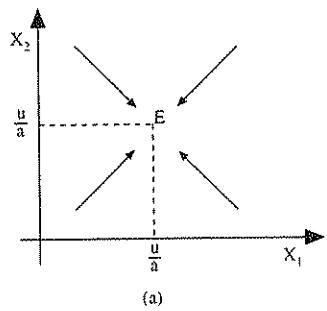
Es decir,

$$s^2 + 2(a - D)s + a^2 - 2aD = 0 \quad (8.12)$$

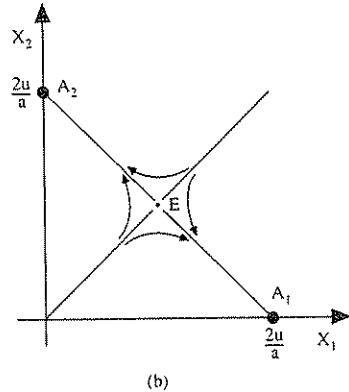
De esta ecuación se desprende inmediatamente que si $D < a/2$, entonces el sistema es estable, ya que las dos soluciones de esta ecuación de segundo grado resultan tener la parte real negativa.

La estabilidad del sistema en este caso indica que el equilibrio $x_1^e = x_2^e = u/a$ es un atractor estable. Esta solución está caracterizada por el hecho de que los valores alcanzados por el morfogén en ambas células son iguales. Se observa, por tanto, una situación de equilibrio caracterizada por la homogeneidad: las dos células alcanzan el mismo estado de equilibrio. No se produce ninguna diferenciación entre ellas.

Por el contrario, si el valor tomado por el coeficiente de difusión D es tal que $D > D_c = a/2$, entonces de las dos soluciones de la ecuación (8.12) una será de signo positivo y otra negativo. El equilibrio es, por tanto, inestable. Quiere ello decir que la situación de equilibrio en la cual el nivel alcanzado por el morfogén es el mismo en ambas células, ya no es posible que se produzca de una manera estable. Aunque se lleve al sistema a esa posición de equilibrio, cualquier pequeña perturbación lo apartará de él.



(a)



(b)

FIGURA 8.9. Comportamiento del sistema de dos células en el espacio de estados.

Los dos tipos de situaciones que estamos considerando se pueden ilustrar gráficamente mediante el comportamiento del sistema en el plano de fase, tal como se indica en la figura 8.9. En la figura 8.9a se muestra el caso en el que el equilibrio homogéneo E es estable, de modo que todas las trayectorias del sistema, cualesquiera que sean las condiciones iniciales, convergen a

él. Por el contrario, en la figura 8.9b se representa el caso en el cual el equilibrio E no es estable. En tal caso el sistema evolucionará, dependiendo de las condiciones iniciales, hacia una de las dos soluciones posibles que en esa figura se han indicado como A_1 o A_2 .

En cada uno de esos casos lo que se tiene es que todo el morfogén se ha desplazado a una de las células, de modo que en la situación de equilibrio final una de las células posee todo el morfogén, mientras que en la otra está completamente desprovista de él. Qué célula sea la que al final posea todo el morfogén dependerá de las condiciones iniciales (resultado de la fluctuación que separa al sistema del equilibrio inestable). Considerando la figura 8.9b, si estas condiciones iniciales se encuentran en la parte superior de la bisectriz que divide al primer cuadrante, entonces el equilibrio alcanzado será el A_2 , lo que quiere decir que todo el morfogén se encontrará en la célula 2; es decir, el equilibrio final será $x_2^e = 2u/a$, $x_1^e = 0$. La situación simétrica se producirá si las condiciones iniciales se encuentran por debajo de la mencionada bisectriz.

La figura 8.10 ayuda a comprender lo que sucede en la figura 8.9b. Se tiene una superficie convexa cuya cumbre representa el punto E, que es inestable, y que puede verter en dos cuencas de atracción (la A_1 y la A_2) según que una ligera modificación de las condiciones iniciales en E incline la trayectoria hacia una u otra cuenca. Es fácil concebir un simulacro mecánico de esta situación; una bola situada en lo alto de una de las acanaladuras de una superficie ondulada como la de la figura 8.10.

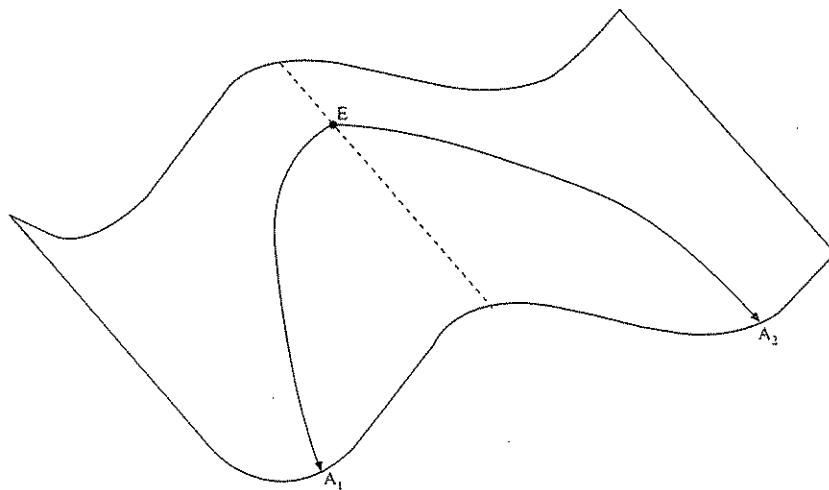


FIGURA 8.10. *Las trayectorias generadas en E pueden alcanzar el atractor A_1 o el A_2 .*

Así pues, una situación inicial de homogeneidad, mediante la adición de un mecanismo de difusión, puede conducir a una situación final de heterogeneidad (las dos células al final del proceso alcanzan estados distintos; se ha generado, por tanto, una heterogeneidad en cuanto al estado de las mismas). Obsérvese la profunda semejanza entre el proceso que estamos describiendo y el de transición del desorden al orden, que anteriormente hemos comentado

(recuérdese figura 8.7). En efecto, la situación en la cual ambas células poseen el mismo nivel de concentración del morfogén —es decir, la situación de homogeneidad— podemos asimilarla a la situación en la cual la distribución de las “cruces” y de las “aspas” es la misma en los dos módulos que considerábamos al comentar el fenómeno de paso del desorden al orden, de acuerdo con Prigogine. Así mismo, la situación en la cual todo el morfogén se ha concentrado en una de las células es asimilable a aquella en la cual todas las “cruces” o las “aspas” estaban en uno de los dos modos de la figura 8.7. Por tanto, a lo que en el modelo de Turing se denomina *homogeneidad*, en el de Prigogine lo llamamos *desorden*; mientras que para lo que en este último se emplea el término *orden*, en aquél utilizamos *heterogeneidad*. Pero, si bien se considera, se verá que en ambos casos estamos hablando de lo mismo. Tenemos, por consiguiente, en el proceso que acabamos de describir un mecanismo que permite explicar cómo se puede generar orden a partir del desorden, o heterogeneidad a partir de la homogeneidad. A este mecanismo se le ha denominado *metáfora morfogenética*⁸.

Los equilibrios A_1 y A_2 son en realidad seudoequilibrios, ya que vienen determinados por la imposibilidad de existencia de soluciones negativas físicamente realizables para las variables x_1 y x_2 en el proceso que estamos considerando. Son seudoequilibrios de frontera. En un sistema lineal, como es el caso de la expresión (8.11), sólo puede existir un equilibrio (salvo el caso degenerado en el que exista una infinitud de ellos). Por tanto, el ejemplo que acabamos de comentar es un tanto artificioso, ya que al modelo esencialmente lineal que describe el proceso, y sobre el que se basa la sencillez del modelo, hay que añadir un mecanismo de saturación mediante el cual se pongan de manifiesto múltiples equilibrios por los que se hace patente la existencia de la heterogeneidad. En general, con sistemas no lineales, en los que el formalismo matemático es más complejo que el anterior, la aparición de los equilibrios es, sin embargo, más natural.

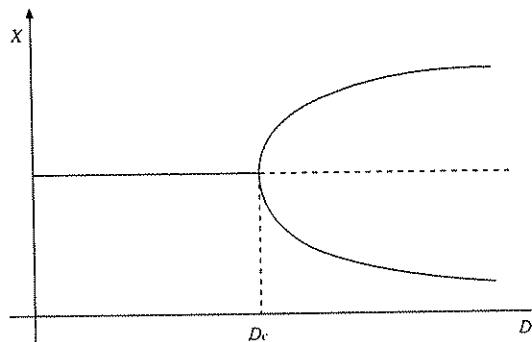


FIGURA 8.11. Punto de bifurcación.

En el ejemplo anterior, el valor crítico del parámetro de difusión $D_c = a/2$ se conoce como *valor de bifurcación* del parámetro, y tiene la notable propiedad de representar el punto del espacio de los parámetros en el cual se produce en la transición de una situación (la de homogeneidad estable) a otra (la de homogeneidad inestable). Para ilustrar el significado del punto de bifurcación considérese la figura 8.11. En esta figura se supone que en ordenadas se

⁸ R. Rosen, 1985, *Anticipatory systems*, Pergamon Press.

representa el parámetro D , mientras que en abscisas se tiene el equilibrio alcanzado x_i^e . Para valores del parámetro inferiores al valor crítico D_c aparece una única situación de equilibrio (la situación de homogeneidad). Sin embargo, para valores del parámetro superiores al valor crítico, el equilibrio anterior se convierte en inestable (por eso se representa gráficamente mediante un trazo discontinuo) y aparecen, para cada valor del parámetro, dos situaciones de equilibrio posibles. El diagrama de la figura 8.11 recibe la denominación de *bifurcación en tridente*, y ha sido objeto de una peculiar interpretación por Prigogine, como veremos en la sección 8.2.3.

El ejemplo que acabamos de ver constituye una muestra muy sencilla de un proceso de enorme generalidad mediante el cual se puede dar cuenta de la aparición de la diversidad en el seno de un medio inicialmente homogéneo. Si en vez de considerar exclusivamente dos células se consideran una gran cantidad de ellas, es decir, si en lugar de limitarnos a un proceso bimodular, consideramos una distribución espacial de módulos conectados entre sí por mecanismos de difusión, entonces se obtiene una generalización del fenómeno que acabamos de ver. En este caso se produce exactamente el mismo tipo de fenómenos que hemos observado, sólo que ahora con la complejidad inherente a la mayor dimensionalidad del proceso.

La estructura de reacción-difusión invita a adoptar una distribución espacial en el análisis de un sistema, mediante una malla de módulos ligados entre sí por fenómenos de difusión o transporte. En cada uno de los módulos puede tener lugar un proceso de los tradicionalmente modelados mediante la dinámica de sistemas. El conjunto constituye un sistema complejo en el que pueden darse fenómenos de pérdida de homogeneidad.

8.2.2. Aplicación al modelo de un área industrial

En la subsección 7.2.1 hemos visto el análisis cualitativo del modelo de un área industrial. Supongamos ahora que el área se descompone en una serie de módulos, cada uno de los cuales se articula con un modelo como el anterior, al que se añade un término responsable del transporte (difusión) entre módulos. En el caso más sencillo podemos considerar dos de tales módulos unidos por un fenómeno de difusión lineal de la forma $D(x_1 - x_2)$, siendo D el coeficiente de difusión⁹. Se obtiene entonces un modelo de la forma:

$$\begin{aligned}\frac{dx_1}{dt} &= x_1[nf(x_1) - m] + D(x_1 - x_2) \\ \frac{dx_2}{dt} &= x_2[nf(x_2) - m] + D(x_2 - x_1)\end{aligned}$$

Mediante métodos numéricos apropiados se determina el diagrama de bifurcaciones de este modelo. Este diagrama muestra la bifurcación tridente (figura 8.12), que es un tipo de bifurcación que no hemos considerado anteriormente y cuyo análisis es enormemente interesante.

Si se denota por D_c el valor de D para el que, en la figura 8.12, se produce la bifurcación, es claro que para valores de $D < D_c$ el sistema evoluciona de modo que el equilibrio final en ambos módulos es el mismo. El sistema tiende en ese caso a un estado estacionario uniforme:

⁹ J. Aracil, E. Ponce y E. Freire. 1984. *Order through fluctuations and systems dynamics models*, Int. J. System Science 15:575–599.

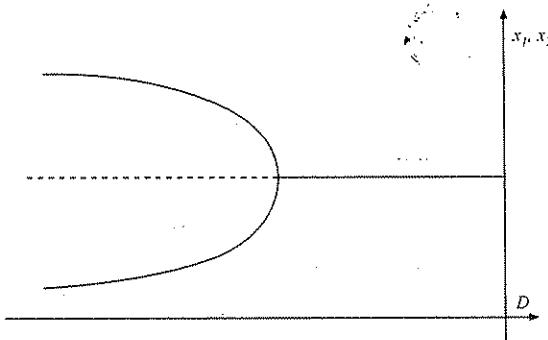


FIGURA 8.12. Diagrama de bifurcación del modelo de un área industrial con dos módulos.

los valores alcanzados por las variables x_1 y x_2 , que miden la actividad económica en cada uno de los módulos, coinciden. Sin embargo, si $D > D_c$ entonces el sistema evoluciona hacia un estado de equilibrio en el que el número de empresas en cada módulo es diferente: el estado estacionario ya no es uniforme. Se obtiene una estructuración espacial que ha sido bautizada por Prigogine como *estructura disipativa* u *orden mediante fluctuaciones*, sobre la que volveremos en la sección siguiente. En tal caso el comportamiento que resulta, muestra —para determinados valores del coeficiente de difusión D — la pérdida de homogeneidad del área industrial, lo que produce la diferente especialización de las subáreas. Se han propuesto modelos de esta naturaleza para representar las diferenciaciones tipo Norte-Sur en el desarrollo económico¹⁰. También pueden emplearse modelos de esta naturaleza cuando se estudia el proceso que se produce al competir diferentes tecnologías por su implantación en el mercado. Sigue entonces que pequeños acontecimientos de naturaleza aleatoria, que tengan lugar en las fases iniciales del desarrollo de estas tecnologías, pueden afectar decisivamente sobre cuál de ellas será la que definitivamente triunfe¹¹.

Conviene observar que el hecho de haber desagregado el área, aunque sea simplemente en dos subáreas, permite que se presenten situaciones que para el modelo agregado pasaban inadvertidas: aquellas en la que el comportamiento de las subáreas es distinto, aunque la suma de los dos sea el mismo. Estos comportamientos distintos serían ignorados por el que considera conjuntamente el área. En este sentido se puede afirmar que el modelo que hemos considerado ahora es más rico que el modelo sin estructuración espacial.

8.2.3. Orden mediante fluctuaciones

Las formas no homogéneas que surgen como consecuencia del mecanismo que acabamos de considerar fueron denominadas por Prigogine *estructuras disipativas*. Con el término *estructuras* se alude a que se trata de formas de organización no homogénea, y con el de *disipativas* se pone de manifiesto que estas estructuras sólo pueden presentarse en sistemas abiertos en

¹⁰ P. Allen, J. L. Deneubourg, M. Sanglier, F. Boon y A. de Palma, 1979, *Dynamic Urban Growth Models*, Report DOT-PB80-12245 0.

¹¹ W. Brian Arthur, 1990, *La retroacción positiva en la economía*, Investigación y Ciencia, abril, número 163, 86-93.

situaciones alejadas del equilibrio termodinámico. Precisamente este concepto constituye la aportación fundamental de Prigogine sobre las ideas previamente desarrolladas por Turing. Las estructuras disipativas sólo pueden darse en sistemas abiertos, y, por tanto, los enclaves localmente antientrópicos en los que tiene lugar la aparición del orden en el Universo deben ser sistemas abiertos. De este modo, las ideas de Prigogine vienen a dar forma a intuiciones de Bertalanffy, que en el momento en que este autor las formuló carecían del suficiente sustrato termodinámico. Con las aportaciones de Prigogine, sin embargo, adquieren un contenido científicamente aceptable.

La génesis de una estructura disipativa se produce cuando pierde el equilibrio una situación previamente homogénea. Esta pérdida del equilibrio hace que cualquier perturbación o fluctuación (recuérdense la figura 8.9 y los comentarios que se hicieron de ella) pueda "cebarse" en el seno del sistema, adquiriendo un desarrollo propio. De este modo, la situación heterogénea final comienza con una fluctuación del equilibrio homogéneo inestable previo. Por tanto, se puede considerar que la situación de equilibrio final es el resultado del crecimiento de una fluctuación. El orden final se alcanza a partir de un proceso que se desencadena mediante una fluctuación. Por ello Prigogine llamó a este proceso *orden mediante fluctuaciones*.

Debe notarse que se trata de un proceso esencialmente determinista una vez que se han dado las condiciones iniciales. Lo que sucede es que estas condiciones son el resultado de una fluctuación que tiene un carácter esencialmente estocástico. Por ello es imprevisible cuál de los dos estados posibles, el A_1 o el A_2 de la figura 8.9b, se alcanzará al final del proceso. Así pues, en el fenómeno considerado se asocian un proceso estocástico (una fluctuación) que da lugar a una desviación del sistema del estado E y un proceso determinista que, a partir del estado resultante de la fluctuación, da lugar a una trayectoria que conduce al sistema a uno de los dos estados de equilibrio A_1 o A_2 .

Es decir, cuando se rompe el equilibrio, debido a que el proceso de difusión hace inestable al sistema, se pone de manifiesto una forma elemental de autoorganización consistente en que los dos módulos evolucionan hacia un equilibrio caracterizado porque los valores alcanzados por el morfogén en cada uno de ellos, es diferente. Ese fenómeno de autoorganización es más sensible cuanto más lejos se está del equilibrio.

En el mecanismo mediante el cual se producen las estructuras disipativas se conjugan *necesidad* y *azar*. La necesidad fuerza al comportamiento del sistema a desenvolverse de una forma preestablecida (*Necesidad la Forzada*, la llama García Bacca¹²). El azar introduce el elemento de contingencia, indefinición, libertad en el comportamiento... mediante el cual se manifiesta la aparición de la variedad. La necesidad viene dada por las ecuaciones dinámicas del comportamiento del sistema que rigen sus trayectorias, mientras que el azar se manifiesta en la fluctuación a partir de la cual se desencadena el proceso que conduce al estado heterogéneo (ordenado) que es una estructura disipativa.

Precisamente la introducción de este elemento de azar en la génesis de un sistema complejo impide que el concepto de estado, en su sentido clásico de resumen o sumario del pasado de un sistema, tenga sentido en el contexto de los sistemas complejos. En efecto, en un sistema clásico (mecánico y determinista) sabemos que su estado resume la información necesaria con relación al pasado de este sistema para prever su futura evolución. De hecho, a ello es a lo que se refería Laplace cuando decía que conociendo la posición y velocidad de todas las partículas

¹² J. D. García Bacca, 1985. *Necesidad y azar*, Anthropos.

del Universo (es decir, su estado) sería capaz de predecir su futura evolución. Sin embargo, hoy sabemos que, en un sistema complejo, esta pretensión debe ser matizada.

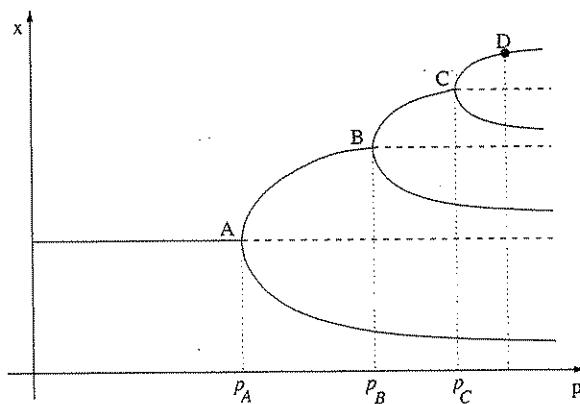


FIGURA 8.13. Comportamiento de un sistema con múltiples puntos de bifurcación.

En efecto, consideremos la figura 8.13 que muestra un diagrama de bifurcaciones en el que aparecen varias de las consideradas en la figura 8.11. Supóngase que el sistema se encuentra en el equilibrio correspondiente al punto D de dicha figura. La génesis de dicho estado —a partir de una situación originaria en la que el valor del parámetro p fuese menor que p_A — implica que en la evolución del sistema se hubiesen producido las bifurcaciones A, B y C. Pero no sólo eso, sino que, además, en cada una de esas bifurcaciones la fluctuación asociada hubiese conducido al sistema a la rama superior de las dos estables posibles que se generan en cada una de ellas. Por tanto, no es posible concebir la situación representada por D con un estado en un sentido clásico, sino que hay que narrar la historia previa mediante la cual se ha llegado a él (es decir, la indicación expresa de cuál es el camino que ha tomado el sistema en cada una de las bifurcaciones). De este modo, el concepto clásico de estado, propio del mecanicismo y del determinismo, queda fuera de lugar y necesitamos, para describir la situación en la que se encuentra el sistema en un instante determinado, recurrir a su *historia*. Solamente mediante la narración de lo acaecido al sistema es posible caracterizar completamente su situación actual. No tiene sentido plantearse un concepto abstracto como el de estado como resumen o sumario abstracto del pasado del sistema, sino que hay que registrar el conjunto de su evolución (o al menos los momentos significativos más importantes)¹³.

En el mecanismo de génesis de las estructuras disipativas podemos encontrar implícitamente desarrollada una *teoría de la autoorganización*. En efecto, mediante la estructura de reacción-difusión podemos dar cuenta de los fenómenos endógenos de organización que tienen lugar en un sistema complejo. Esta teoría puede aplicarse tanto a sistemas físicos o químicos, como a sistemas biológicos o sociales¹⁴. En realidad Prigogine la estudió originariamente en sistemas químicos pero el reconocimiento de su ubicuidad constituye una aportación a la teoría

¹³ Se puede intuir una cierta conexión entre estas ideas y la *razón histórica* de Ortega.

¹⁴ H. Haken, 1985, *Fórmulas del éxito en la naturaleza*, Salvat.

de sistemas de la misma naturaleza de la que hace unos cincuenta años realizó Norbert Wiener cuando, después de estudiar la estructura de realimentación en el posicionamiento de un cañón antiaéreo, descubrió asimismo su ubicuidad, al encontrarla en descripciones de lo que sucedía en otros dominios de la realidad. Precisamente estas dos estructuras, la de realimentación negativa —que suministra un mecanismo que permite dar razón del comportamiento propositivo (y con ello de comportamientos aparentemente teleológicos)— y la de reacción-difusión (que, como acabamos de ver, hace algo análogo con relación a los procesos morfogenéticos) son las dos estructuras básicas que ha estudiado la teoría de sistemas.

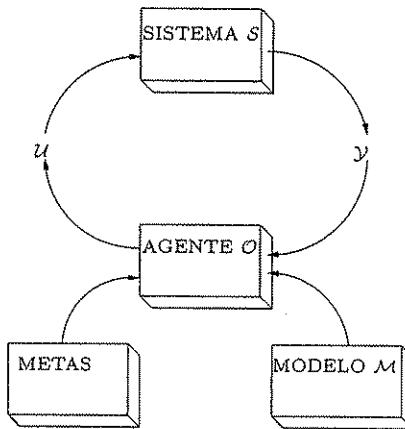
9.1. INTRODUCCIÓN

En capítulos anteriores nos hemos ocupado de los modelos como artefactos que construíamos para representar determinados aspectos problemáticos de los sistemas que pueblan nuestro entorno. Este capítulo se va a dedicar a la explotación de modelos, esto es, a cómo obtener provecho del modelo de una manera racional, una vez que éste ha sido construido y verificado.

Se acaba de emplear el término *racionalidad* para caracterizar el tipo de actuación que aquí pretendemos analizar. Este término, sin embargo, posee múltiples acepciones. La que ahora nos interesa es aquella *racionalidad* que se predica de actuaciones y comportamientos. Para nosotros, y de forma genérica, una actuación *racional* es aquella en la que el agente actúa con plena conciencia de cuáles son los fines que trata de alcanzar, de modo que sea capaz de adecuar los medios de que dispone al logro de sus fines y evaluar en qué medida se han alcanzado éstos para, en su caso, modificar sus pautas de comportamiento. En ese sentido decimos que la gente que actúa de ese modo puede *dar razones* de lo que hace; en cuyo caso calificamos de *racional* su comportamiento. Para caracterizar un proceso de actuación *racional* vamos a auxiliarnos del diagrama de la figura 9.1.

9.2. EL MODELO EN UN PROCESO DE ACTUACIÓN RACIONAL

La figura 9.1 muestra un diagrama que representa una forma esquemática de un proceso de *actuación racional*. Se supone que el que hasta ahora era un observador pasivo \mathcal{O} , que construía un modelo \mathcal{M} de \mathcal{S} , es ahora un *agente*, al que seguiremos denominando \mathcal{O} . Este agente \mathcal{O} pretende actuar sobre \mathcal{S} , de modo que el comportamiento de \mathcal{S} se ajuste a determinadas metas. Para actuar sobre \mathcal{S} , el agente \mathcal{O} dispone de ciertos instrumentos de actuación, que se concretan en el conjunto de variables U . En el caso de una máquina, la actuación sobre \mathcal{S} se realiza

FIGURA 9.1. *Esquema de un proceso de actuación racional.*

mediante mandos de actuación (botones, manivelas, teclados, palancas,...). En el caso de sistemas de otra naturaleza, se hace mediante la asignación de valores a lo largo del tiempo a parámetros que representan políticas de actuación sobre el sistema (por ejemplo, decisiones sobre el precio que un fabricante asigna a un determinado producto, o sobre el nivel de contratación de empleados; decisiones del gobierno sobre el tipo de redescuento bancario, u otro parámetro de política económica ...). En todo caso \mathcal{O} decide los valores que tomarán las U , a las que se denomina (recuérdese la figura 9.1) *señales de entrada*, o sencillamente, *entradas* al sistema S . Por la función que desempeñan se denominan también *señales de control* o de *mando*.

El agente \mathcal{O} obtiene información de S registrando las señales Y , denominadas *señales de salida* o *salidas* del sistema. Los registros de las señales Y le permiten conocer el comportamiento de S . Por tanto, la interacción de \mathcal{O} con S se realiza mediante las señales U e Y , que forman el conjunto de atributos asociados al sistema (recuérdese lo dicho en el capítulo 2). El esquema de la figura 9.1 presenta la tantas veces mencionada estructura de realimentación. El agente \mathcal{O} va tomando decisiones respecto a los valores que asigna a las señales U teniendo en cuenta las lecturas de los valores Y que, en cierta medida, son resultado de las acciones U que ha tomado previamente. Es decir, se realimenta (en el sentido concreto de “se vuelve a” alimentar) información sobre los resultados de acciones pasadas para decidir las acciones posteriores que se van tomando.

En el diagrama de la figura 9.1, además de \mathcal{O} y S , aparecen el modelo M de S y las metas que \mathcal{O} persigue de S . El modelo M es un elemento esencial en la actuación de \mathcal{O} . El agente pretende un cierto comportamiento de S . Este comportamiento deseado lo podemos concretar en que la señal de salida $y(t)$ tome un valor determinado $y^*(t)$ (para $0 < t < T$; de aquí en adelante, si no se dice lo contrario, se entenderá que toda señal está referida a este intervalo genérico de tiempo). Esta trayectoria deseada para $y(t)$ se obtendrá cuando se aplique una señal de control $u^*(t)$ adecuada. El agente \mathcal{O} , para saber qué $u(t)$ debe aplicar, debe conocer —al menos de forma virtual— todos los pares posibles $\{u(t), y(t)\}$ que se pueden

asociar al sistema S ; es decir, el agente debe ser capaz de prever cómo se comportará el sistema S ante cualquier actuación U que él pueda realizar (gobernar es prever) para, a partir de esa información, decidir cuál de las $u(t)$ aplicará para alcanzar la meta que persiga. Esa información que liga a las diferentes actuaciones $u(t)$ con los resultados $y(t)$ se la aporta el modelo M . El modelo M , en la medida que es una representación adecuada de S , constituye un sumario abstracto de los pares $\{u(t), y(t)\}$; o, lo que es lo mismo, una expresión formal que permite generarlos. En consecuencia, O necesita el modelo M para poder decidir su actuación sobre S .

El papel del modelo M en la conceptualización de una actuación racional es, por tanto, básico. Este hecho no ha sido, sin embargo, suficientemente resaltado en la literatura. La actuación se ha vinculado tradicionalmente a los objetivos que se trata de alcanzar, y no a la representación que se disponga del objeto sobre el que se actúa. Estamos viendo cómo en una actuación racional el agente trata de adecuar sus medios a los fines que persigue. El modelo M le permite precisamente realizar esta adecuación. Le indica cuál, entre sus posibilidades de actuación representadas por las diferentes políticas $u(t)$ que puede decidir adoptar, es la adecuada para los fines que persigue. De este modo se pone de manifiesto la importancia del modelo en toda actuación racional.

Conviene recordar que todas las decisiones habituales, tanto individuales como sociales, se toman en relación a algún modelo, al que genéricamente se alude como *modelo mental*, y que está formado por el conocimiento, más o menos fiable y estructurado, que el agente posee sobre el aspecto de la realidad sobre el que actúa. Este modelo mental se puede convertir en un útil de una potencia considerablemente mayor, por lo que respecta a la ayuda en la toma de decisiones, si somos capaces de formalizarlo. Es lo que podemos hacer, como hemos visto en los capítulos anteriores, con ayuda de métodos de modelado como la dinámica de sistemas.

Un modelo formalizado posee ciertas características que lo hacen particularmente interesante en el proceso de toma de decisiones:

- contiene un considerable grado de precisión;
- es explícito y transparente y, en consecuencia, puede ser objeto de un análisis crítico;
- procede desde los supuestos (hipótesis o asunciones) a las conclusiones de una manera lógica y reconstruible;
- puede ser fácilmente alterado para representar hipótesis diferentes o políticas alternativas, y
- es comprehensivo; es decir, interrelaciona muchos factores simultáneamente.

El otro elemento que aparece en la figura 9.1 junto con el modelo M está constituido por las metas que se persiguen con la actuación. Las metas o fines son lo que pretendemos con nuestras actuaciones, los objetivos que conscientemente deseamos. En la figura 9.1 se asume un proceso de actuación teleológica, en el sentido de que está orientada a la consecución de unas metas que posee el que desencadena la acción. Estas metas, en algunos casos, podrán describirse con precisión, e incluso podrán ser objeto de una cuantificación mediante una función objetivo que mida de forma cuantitativa su grado de consecución. En otros casos, esto no será posible, pero el número de casos en que sí lo es resulta ser suficientemente importante como para que nos ocupemos detenidamente de él. Es lo que vamos a hacer en las secciones siguientes.

9.3. USO DE MODELOS PARA BÚSQUEDA DE ACTUACIONES RACIONALES

En la sección anterior se ha analizado el empleo de modelos como instrumento en la búsqueda de actuaciones racionales. Sin embargo, no se ha descrito la manera de llevar a cabo esta búsqueda. Es éste un problema abierto que no tiene una solución definitiva. Existen diversos procedimientos que pueden ayudar en esta tarea, pero para los que no está garantizada su aplicabilidad en todos los casos. En esta sección se estudiarán estos procedimientos.

En el planteamiento original de Forrester, para la búsqueda de actuaciones racionales utilizando modelos de dinámica de sistemas no se deben utilizar procedimientos automáticos, sino que el proceso debe ser realizado por la persona que dispone del modelo. Así, mediante un procedimiento de prueba y error se llevaría a cabo la búsqueda: ejecutando el modelo, observando los resultados obtenidos y decidiendo los cambios a realizar para una nueva simulación. Repitiendo este proceso una y otra vez se puede observar el efecto de cada una de las actuaciones ensayadas y decidir cuál es la mejor.

Evidentemente, este proceso requiere un conocimiento profundo del sistema que se pretende controlar, y quizás grandes dosis de intuición para poder orientar la búsqueda adecuadamente, sobre todo si el modelo es relativamente complejo. Pero, además, el número de pruebas necesarias puede hacer que la tarea sea muy tediosa o incluso inabordable en modelos grandes, en los que el tiempo necesario para realizar una simulación puede ser relativamente alto y la complejidad del sistema obligará a un número elevado de pruebas.

Sin embargo, y dado que los modelos de dinámica de sistemas se implantan en un computador, podría plantearse la posibilidad de programar un algoritmo informático que, de forma automática, realice esta búsqueda. Las siguientes subsecciones se van a dedicar al estudio de diversos algoritmos que pueden ser de utilidad para conseguir este objetivo.

9.3.1. Optimización estática

En primer lugar, se podría pensar en utilizar diversos procedimientos de optimización de parámetros relativamente sencillos, como la búsqueda de máximos y mínimos de funciones, o la programación lineal. Con estos métodos, se pueden obtener los valores de ciertas variables que optimizan una función.

De esta forma sería posible resolver dos tipos distintos de problemas:

1. En primer lugar, se pueden considerar aquellos problemas en los que la bondad de una actuación se define en función de la evolución del sistema a lo largo de un intervalo de tiempo (y por tanto, no se puede tener una medida de esta bondad hasta finalizar una simulación del sistema en ese intervalo de tiempo). El objetivo sería determinar los parámetros que maximizan esta bondad. Por ejemplo, se podría calcular el precio de venta de un producto, de manera que se maximicen los beneficios obtenidos en el intervalo de tiempo. Obsérvese que los valores obtenidos de las variables de control o actuación (el precio de venta en este ejemplo) han de mantenerse constantes durante toda la simulación.

Algunos de los entornos informáticos de simulación tienen herramientas que permiten resolver este tipo de problemas. De hecho, este problema es prácticamente el mismo que el de ajuste de parámetros que se expuso en la sección 6.4. La única diferencia estriba en

la función objetivo: en el caso de ajuste de parámetros se desea minimizar la diferencia entre dos señales (la de referencia y la simulada); en el caso que nos ocupa ahora se tiene más libertad en cuanto a la forma de la función que se desea minimizar (o maximizar).

2. En otros casos se puede plantear un problema de optimización en cada instante de tiempo, con una función de bondad que dependa de los datos conocidos hasta ese instante. De esta forma se obtendría una evolución de las variables de control que cambiaría con el tiempo. Sin embargo, este procedimiento no tiene en cuenta el carácter dinámico del problema, ya que la optimización planteada depende del estado del sistema en el instante actual, y no tendrá en cuenta los estados futuros a los que se llegará con la señal de control aplicada. Así se puede obtener el precio de venta de un producto que maximice el beneficio en el instante considerado, pero no se podrán tener en cuenta los efectos de la decisión tomada en las ganancias futuras porque para ello habría que continuar con la simulación del modelo, cosa que no se puede hacer hasta que no se resuelva el problema de optimización planteado.

En cualquier caso este tipo de optimización no suele estar incorporado de una forma general en los paquetes informáticos de simulación. Sin embargo, si se puede programar en un lenguaje de propósito general e incluirlo en lenguaje de simulación en aquellos entornos que permitan la inclusión de funciones escritas en otros lenguajes.

Como es fácil deducir, ninguno de estos enfoques permite considerar el problema en su globalidad. El primero de ellos tiene en cuenta el carácter dinámico del sistema, pero obliga a señales de control constantes; por el contrario, el segundo tipo de problemas permite la variación de las señales de control, pero no tiene en cuenta el carácter dinámico del sistema.

Sería conveniente disponer de un método que calcule las señales de control como variables en el tiempo y contemple el efecto dinámico de estas decisiones, es decir, que tenga en cuenta tanto el efecto inmediato de las mismas como el efecto a más largo plazo. Éste es precisamente el caso que se trata en el problema del *control óptimo*.

9.3.2. Control óptimo

Planteamiento del problema

El cálculo variacional y su aplicación al control de sistemas permite llegar a un algoritmo automático que, al menos en principio, resuelve el problema de la búsqueda de decisiones racionales.

Supongamos que disponemos de un modelo \mathcal{M} del comportamiento de \mathcal{S} mediante un sistema dinámico:

$$\frac{dx}{dt} = f(x, u) \quad (9.1)$$

Este formalismo hace explícitas las señales u mediante las que se actúa sobre el sistema \mathcal{S} y es el habitualmente empleado en la formulación del problema del control óptimo.

Por otra parte, supongamos también que existe una función de la forma

$$J = \int_0^T L(x, u) dt \quad (9.2)$$

de modo que se pretende que el comportamiento de \mathcal{S} optimice (minimice o maximice) el valor tomado por J . Este valor es un índice del comportamiento del sistema, ya que depende del conjunto de valores tomados por las trayectorias en el intervalo $(0, T)$ de interés. La elección de este índice es determinante para que la solución que se obtenga sea adecuada: este índice es la única medida que se va a utilizar para discriminar las acciones racionales de las que no lo son o, mejor, define una relación de orden entre las distintas evoluciones del sistema, de manera que valores menores del índice corresponden a trayectorias mejores (en el caso de minimización).

Mediante el control óptimo se trata de determinar la señal de actuación $u^*(t)$, y la correspondiente trayectoria para el estado $x^*(t)$, de modo que la expresión (9.2) alcance su valor óptimo.

El formalismo (9.1-9.2) se presenta en múltiples aplicaciones técnicas en las que se trata de optimizar un determinado índice asociado al funcionamiento de un sistema; por ejemplo, el consumo de energía en un periodo de tiempo, el tiempo en el que se trata de alcanzar una meta, etc. En ciencias sociales también se encuentran problemas que se pueden abordar con este formalismo. Así, por ejemplo, se ha aplicado a problemas de optimización financiera, producción e inventarios óptimos, planificación de la producción, políticas óptimas de promoción de productos, explotación óptima de recursos naturales, problemas de transporte, etc.¹.

Una vez formulado el problema mediante las expresiones (9.1) y (9.2) sabemos, en general, cómo determinar la señal de control $u^*(t)$ que optimice J . Para ello formamos la denominada *función de Hamilton* del proceso, o *Hamiltoniana*, que viene dada por:

$$H(x, u, \lambda) = L(x, u) + \lambda f(x, u) \quad (9.3)$$

en donde las variables λ representan las variables adjuntas o de coestado del sistema, de acuerdo con la teoría del control óptimo. A partir de (9.3) podemos proceder a una resolución progresiva del problema. Para ello, en primer lugar, se determina el mínimo de H con relación a u . Ello, en el caso más simple, conduce a resolver en u la ecuación:

$$\frac{\partial H}{\partial u} = 0 \quad (9.4)$$

lo que permite obtener $u^*(x, \lambda)$. Este valor de u se lleva a la función de Hamilton $H(x, u, \lambda)$, y con la función $H(x, \lambda)$ que resulta de esta sustitución se forma el sistema de ecuaciones diferenciales de Hamilton:

$$\frac{d\lambda}{dt} = -\frac{\partial H}{\partial x} \quad (9.5)$$

$$\frac{dx}{dt} = \frac{\partial H}{\partial \lambda} \quad (9.6)$$

cuya integración permite obtener las trayectorias $x^*(t)$ y $\lambda^*(t)$. Estas trayectorias se llevan a $u^*(x, \lambda)$, con lo que se tiene la señal de mando óptima $u^*(t)$.

Para la integración de las ecuaciones (9.5) y (9.6) es necesario disponer de condiciones de contorno. En el caso de la variable de estado x , éstas son claramente las referidas al instante

¹ M. I. Kamien y N. Schwartz Lou, 1995, *Dynamic optimization: the calculus of variations and optimal control in economics and management*, North-Holland

inicial $x(0) = x_0$. Las condiciones de contorno para las variables de coestado λ se obtienen de la llamada *condición de transversalidad*. Esta condición escapa del carácter generalista con el que se está exponiendo aquí este tema. Baste decir que con la condición de transversalidad se obtienen condiciones de contorno en el instante final $\lambda(T) = \lambda_T$.

La resolución de las anteriores ecuaciones no es trivial. La resolución analítica está restringida a sistemas de pequeña dimensión, en la mayoría de los casos lineales. Para sistemas de gran dimensión, con no-linealidades, es preciso acudir a la ayuda de la informática mediante la utilización de algoritmos numéricos o incluso heurísticos para resolver el problema. A continuación se exponen las ideas fundamentales sobre estos algoritmos.

Algoritmos numéricos

Las ecuaciones (9.4), (9.5) y (9.6) forman un sistema de ecuaciones diferenciales con condiciones de contorno. Aunque aparece menos frecuentemente que los problemas con condiciones iniciales, su resolución numérica está suficientemente estudiada. Este problema es bastante más difícil de resolver que los problemas con condiciones iniciales. Las dificultades, como se puede intuir, aparecen porque si se desea aplicar un método de integración usual (como el método de Euler o de Runge-Kutta) es necesario partir de unas condiciones iniciales, las cuales no se conocen para todas las variables en un problema de este tipo (en concreto, no se conocen los valores de λ en $t = 0$, o, si se desease integrar hacia atrás, no se dispondría de los valores de x para $t = T$).

A continuación se va a esbozar un método numérico que se puede utilizar para integrar numéricamente el anterior sistema de ecuaciones diferenciales, basado en el método de la relajación:

1. Se discretiza el tiempo en n_t instantes en los que se van a calcular los valores de $u(t)$. En el interior de los intervalos así formados los valores que toma $u(t)$ se calculan por interpolación. Los intervalos de discretización deben ser mayores o iguales que los de integración en las integraciones de los pasos posteriores (normalmente el paso de discretización es un múltiplo del paso de integración).
2. Se postula una solución inicial para las señales de control $u(t)$ en los instantes de discretización. Dado que se han tomado n_t instantes y existen n_u variables de control, es necesario postular $n_t \times n_u$ valores.
3. Puesto que la ecuación (9.6) no depende de λ es posible realizar la integración de ésta, con lo que se puede calcular la evolución de las trayectorias $x(t)$. Las derivadas parciales $\partial H / \partial \lambda$ se pueden calcular por métodos numéricos. Los valores de x calculados se deben almacenar para utilizarlos en lo que sigue. Obsérvese que es preciso almacenar $n \times n_i$ valores, siendo n_i el número de pasos de integración (normalmente mucho mayor que n_t). Este almacenamiento impone una fuerte exigencia de memoria que se justifica por la mayor velocidad de este método frente a otros que no necesitaran tanta memoria.
4. Una vez terminada esta integración hacia adelante se puede integrar hacia atrás la ecuación de coestado (9.5), partiendo de las condiciones finales $\lambda(T)$. Para ello es necesario utilizar los valores de $x(t)$ calculados en el paso anterior.
5. Al llegar a cada instante t_k de discretización del tiempo donde se definen los valores de u se impone la condición (9.4). Para ello se determina el valor de $\|\partial H / \partial u\|_{t_k}$. Si

este valor es más pequeño que una tolerancia prefijada se continúa con la integración hacia atrás. En caso contrario se modifican los valores $u_i(t_k)$; $i = i, \dots, n_u$ buscando la minimización de $\|\partial H/\partial u\|_{t_k}$. Éste es un problema de minimización de una función escalar que depende de n_u parámetros (obsérvese que se ha huido de una minimización de funciones de $n_t \times n_u$ parámetros que son las incógnitas del problema). Para esta minimización se puede utilizar cualquier método numérico, siendo preferibles los que no necesitan calcular derivadas.

Una vez calculados los valores óptimos de $u(t_k)$ se continúa con la integración hacia atrás.

6. Al finalizar la integración hacia atrás se habrán calculado los nuevos valores de $u(t)$ que permiten que se cumpla (9.4). Como, por otro lado, también se cumplen (9.5) y (9.6), podría pensarse que el problema ya está resuelto. Sin embargo, esto no es cierto, ya que se han ido variando los valores de $u(t)$ en la integración hacia atrás, por lo que los valores de $x(t)$ almacenados dejan de ser válidos. A pesar de ello, los nuevos valores de $u(t)$ previsiblemente estarán más cerca de la solución buscada. En este paso se calcula una medida de la diferencia entre la nueva $u(t)$ y la anterior.
7. Si los nuevos valores de $u(t)$ no divergen de los que se tenían anteriormente más que una tolerancia fijada de antemano, se da por finalizado el proceso y $u(t)$ se considera solución. En caso contrario se vuelve a 3.

Con este algoritmo se obtienen buenos resultados con sistemas de mediana dimensión, pero no resuelve el problema cuando se tratan sistemas de muy gran dimensión. En concreto, el algoritmo puede que no converja al no cumplirse nunca la condición de finalización.

Algoritmos heurísticos

En los últimos tiempos, dado el aumento en potencia de cálculo de los computadores, se tiende a utilizarlos intensivamente de manera que problemas complejos puedan ser resueltos aproximadamente, sin necesidad de teorías específicas (como sería la utilización de las ecuaciones de Hamilton en el caso que nos ocupa). Para ellos existen diversos métodos o algoritmos que se pueden englobar bajo la calificación de *heurísticos*. Estos algoritmos resuelven problemas de optimización *paramétrica*, en el sentido en el que la solución de estos problemas es un conjunto de parámetros y no un conjunto de funciones (temporales, en el caso del control óptimo). Sin embargo, ya con los algoritmos numéricos se trata de resolver el problema del control óptimo discretizando las variables continuas, de manera que el problema se reduce a encontrar ciertos valores puntuales de la solución buscada. Con ello la esencia del problema a resolver cambia drásticamente: la resolución del problema del control óptimo se reduce a la del caso de optimización paramétrica. Sin embargo, el problema resultante de esta simplificación tiene dos características importantes, que hacen que siga siendo complejo: por un lado, y debido a que el sistema a controlar puede ser de gran dimensión, la evaluación de la función de coste, para un conjunto dado de los parámetros, conlleva un tiempo no despreciable de cálculo con el computador; por otro lado, el número de parámetros que define una posible solución suele ser bastante alto.

Una característica común de los métodos que se van a exponer es que imponen pocos requerimientos sobre el problema que se pretende resolver; así no es preciso hacer suposiciones

sobre linealidad, derivabilidad o continuidad de funciones o sus derivadas para poder aplicar estos algoritmos.

A continuación se van a describir someramente dos de estos métodos: el *método de Montecarlo* y los *algoritmos genéticos*. Además, existen otros métodos heurísticos como son el enfriamiento simulado (“simulated annealing”), la búsqueda tabú, etc., que aquí no se tratarán.

Método de Montecarlo Un primer algoritmo heurístico es conocido por el nombre de *método de Montecarlo*. Con este método se generan aleatoriamente un conjunto de posibles soluciones (o puntos del espacio de las señales u) y se prueba (se simula el sistema para cada u elegida) para obtener el valor del índice J asociado a cada una de ellas. La solución que, entre las probadas, tenga asociado un menor valor del índice se considera una aproximación del óptimo buscado.

Evidentemente, para que este método funcione correctamente es necesario que el número de puntos ensayados sea grande y, en general, de acuerdo con la complejidad del problema (a mayor complejidad debe corresponder un número mayor). En el caso del control óptimo, por las razones expuestas, el número de parámetros suele ser grande, lo que lleva a la necesidad de generar un número excesivamente numeroso de puntos. Este hecho, unido a que posiblemente el tiempo de una simulación sea apreciable, hace que este método sea inviable en muchas ocasiones, por requerir mucho tiempo para la resolución.

Algoritmos genéticos Un segundo grupo de algoritmos heurísticos son los conocidos con el nombre de *algoritmos genéticos*. Estos algoritmos son semejantes al método de Montecarlo en el sentido que se generan posibles soluciones (trayectorias de las señales de control) y se prueban. Sin embargo, el mecanismo de generación de éstas no es totalmente aleatorio (aunque sí tiene algo de aleatoriedad), de manera que existe, de alguna forma, una búsqueda dirigida por lo que se reduce el número de simulaciones necesarias.

Las leyes que dirigen la búsqueda están inspiradas en los mecanismos de la selección natural. De la misma manera que a lo largo de la evolución de las especies, éstas se van acercando a un ajuste óptimo al medio, así las trayectorias de las señales de control generadas por el algoritmo se irán acercando a la trayectoria óptima. A continuación se esboza el algoritmo genético básico.

Para ello las trayectorias (o, en general, las posibles soluciones del problema de optimización) se codifican en forma de “cadenas” (por ejemplo, como un conjunto de unos y ceros). A estas cadenas se les denomina *individuos*. El algoritmo maneja una *población* de individuos y la hace evolucionar de manera que los individuos de generaciones posteriores sean “mejores” de acuerdo con un cierto criterio. Cada individuo representa una posible solución del problema de optimización que se intenta resolver. Por tanto, el criterio que decide la bondad relativa de los individuos es la función objetivo del problema. La población inicial se escoge de forma aleatoria de manera que, con ello, se está empleando el método de Montecarlo. Sin embargo, el tamaño de la población —es decir, el número de individuos— es mucho menor que el que se emplearía para resolver el método de Montecarlo. Por ello lo más probable es que la inmensa mayoría de los individuos de la población inicial tenga asociado un coste muy elevado. Además, probablemente el mejor individuo no será excesivamente “bueno” (su coste asociado no se aproximará demasiado al óptimo). El coste asociado a cada individuo, es decir

el valor correspondiente del índice de funcionamiento, se denomina *ajuste* (“fitness”)² (nótese que, así definida, esta cantidad en realidad mide lo contrario de lo que se espera de su nombre: un individuo es mejor cuanto menor sea el ajuste).

A partir de esta población inicial y por medio de ciertas operaciones algunas de las cuales se describen más abajo, se van generando poblaciones de individuos *descendientes*. Las operaciones que generan las generaciones descendientes son tales que la probabilidad de que aparezcan en ellas individuos mejores que los ascendientes es relativamente alta. Esto se puede comprobar empíricamente mediante ejemplos. El hecho de que estas operaciones puedan ser variadas, y que queden definidas por una serie de parámetros que, a su vez, también pueden variar, es lo que hace que se hable de *algoritmos genéticos*, en lugar de un solo *algoritmo genético*.

Para el algoritmo genético básico se ha encontrado una justificación teórica del hecho de que en generaciones descendientes es más probable que existan individuos mejores; es el llamado *teorema de los esquemas*.

Por tanto, a medida que se van generando poblaciones sucesivas, la media del *ajuste* de la población mejora. También lo hace el *ajuste* del mejor individuo de cada generación. El algoritmo se detiene tras un número predeterminado de generaciones y se toma al mejor individuo de la última generación como una aproximación del óptimo buscado. Al igual que con el método de Montecarlo, no se garantiza que la solución encontrada sea verdaderamente la óptima, ni siquiera se puede saber si esto es así.

Las operaciones básicas para generar sucesivas generaciones son:

- *Réplica* (“reproduction”)³: Esta operación simula la reproducción asexual de algunas especies naturales. Mediante la réplica, se crea un individuo en la generación posterior como copia exacta de un individuo de la generación precedente. El individuo que se debe copiar se selecciona mediante algún procedimiento de manera que los individuos “mejores” sean escogidos más fácilmente.
 - *Cruce* (“crossover”): Mediante esta operación se simula la reproducción sexual, es decir, el nuevo individuo (o hijo) hereda características de dos individuos de una generación precedente (o progenitores) mientras que otras se pierden. Los progenitores se seleccionan mediante el mismo procedimiento que en el caso de la réplica. Aquí se describe el cruce más básico aunque existen operaciones más complicadas basadas en esta idea. Para realizar el cruce, en primer lugar y de forma aleatoria, se escoge una posición de la cadena que codifica a los individuos y se separan las cadenas de ambos progenitores por este punto. A continuación se generan dos hijos, el primero de ellos con la cabeza del primer progenitor y la cola del segundo; el otro hijo está formado con los restantes elementos: cabeza del segundo progenitor y cola del primero. En la figura 9.2 se representa gráficamente un ejemplo de este proceso.
- Esta operación es la base de los algoritmos genéticos. Como los progenitores se han escogido entre los mejores, existe cierta probabilidad de que, entre los hijos, aparezca

² El término inglés “fitness” se suele traducir como “ajuste” en ámbitos matemáticos y como “aptitud” en los biológicos. Se ha adoptado aquí la primera traducción dado que se está resolviendo un problema de optimización (a pesar de que el algoritmo empleado tenga una base biológica).

³ El término “reproduction” se puede traducir también como *reproducción*, pero, en este caso, en la acepción de *realizar una copia o réplica*. Dado que el término puede llevar a confusión con la idea de apareamiento o cruce, preferimos emplear el término *réplica*.

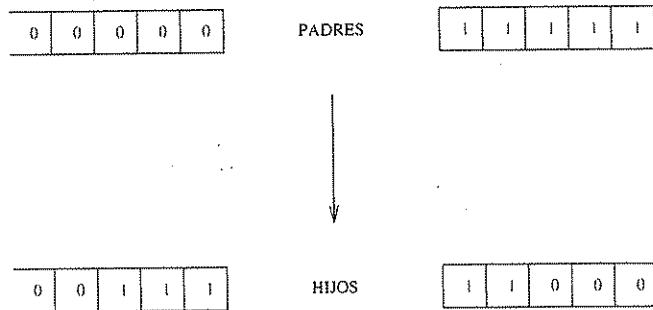


FIGURA 9.2. Ejemplo de cruce.

un individuo de características mejores que los padres. Por supuesto, esta mejora no está garantizada y muy frecuentemente aparecen individuos “nefastos”, pero, dado el criterio de selección que se utiliza, éstos difícilmente pueden sobrevivir más de una generación.

- **Mutación:** Mediante la mutación se pretende generar y mantener cierta diversidad en la población de individuos. Supóngase que, por ejemplo, todos los individuos de una generación tienen idéntico el “gen” (o valor de una posición de la cadena) correspondiente a una determinada posición. Esto puede ocurrir o bien porque en la generación inicial, generada aleatoriamente, se daba este hecho, o bien como resultado de cruces en generaciones precedentes. Ante esta situación, ni la réplica ni el cruce pueden hacer que aparezca un nuevo individuo en el que no se dé esta característica. Esta situación es indeseable, ya que no permite explorar ciertas regiones del espacio de búsqueda y nada impide el que el óptimo se encuentre precisamente en estas regiones.

Para evitar este fenómeno se introduce el operador mutación. En cada generación cada gen de cada individuo tiene una probabilidad muy pequeña de ser “mutado”. Cuando esto sucede el gen es cambiado. En el caso binario solamente se puede cambiar un “1” por un “0”, o viceversa. Con otros códigos se suele introducir un nuevo elemento aleatorio para generar el nuevo valor del gen.

Evidentemente, la probabilidad de la mutación debe ser muy pequeña, ya que es una operación totalmente aleatoria que no contiene ningún mecanismo de selección; por tanto los individuos “mutados” suelen presentar *ajustes* bajos, aunque nada impide que resulten individuos “buenos”.

9.3.3. Control óptimo en bucle cerrado

El problema de control que se ha recordado en la sección anterior es el que se conoce como el de la determinación de la señal de control óptima en bucle abierto. Es decir, se obtiene una señal de control $u^*(t)$ que está determinada para todo el periodo considerado $[0, T]$. Existe un planteamiento más interesante del problema del control en el que se pretende determinar la denominada ley de control óptima, que es una ley de la forma $u(x)$; es decir, es una ley que establece la actuación u en función del estado alcanzado por el sistema x . Se trata, por

tanto, de una ley de control en la que la actuación en todo instante depende del estado; es decir, realimentando el estado —mediante la ley de control— se determina la actuación óptima sobre el sistema. Surge así, de forma natural, la estructura de realimentación al poner en práctica la ley de control. El problema de determinar la ley de control óptima es considerablemente más complejo que el de la señal de mando óptima.

En realidad sólo se dispone de un método general de solución para el caso en el que el sistema dinámico (9.1) sea lineal en x y en u , y la función $L(x, u)$, de la expresión (9.2), sea cuadrática en x y u . Es decir, para el caso en que se tenga:

$$\frac{dx}{dt} = Ax + Bu$$

$$L(x, u) = x^T Qx + u^T Ru$$

Para este caso se demuestra que la ley de control óptima es de la forma

$$u^*(x) = kx$$

es decir, la señal de control óptima es una función lineal del estado x del sistema. Se obtiene así, como solución, una ley de control $u^*(x)$, que conduce a un control en bucle cerrado, y no una señal de la forma $u^*(t)$, que daría lugar a un control en bucle abierto. En la figura 9.3 se ilustra la diferencia entre el control en bucle abierto y en bucle cerrado.

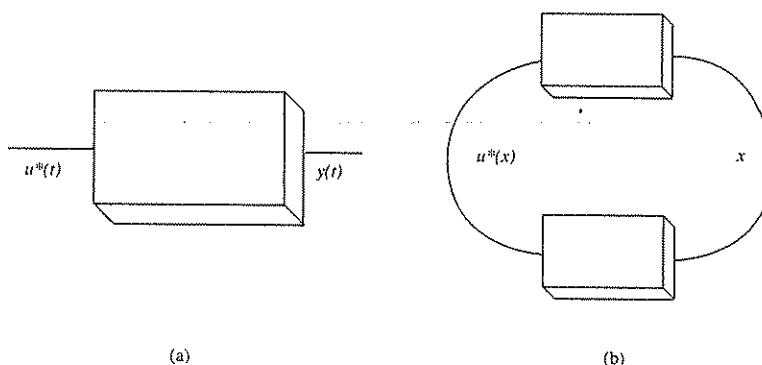


FIGURA 9.3. *Control óptimo en bucle abierto (a) y en bucle cerrado(b).*

9.3.4. Ejemplos

Para ilustrar el problema del control óptimo se van a exponer tres ejemplos sencillos aplicados al modelo de crecimiento de un área industrial. En ellos se va a utilizar el conjunto de valores de los parámetros del modelo utilizado anteriormente en varias ocasiones. En la figura 5.5 se tenía una muestra gráfica del comportamiento del modelo. Como se ha visto en la sección 6.2.2, el número de empresas a las que tiende el modelo depende del valor de los parámetros $n = \text{TNCE}$, $p = \text{SPE/S}$ y $m = 1/\text{VME}$ (además del número inicial, de empresas,

que determinará a cuál de los equilibrios se llegarán). Se supondrá que el gestor del área no puede actuar sobre estos parámetros.

Como señal de control se considerará que el gestor puede hacer variar la superficie del área, S , a lo largo del tiempo (es decir, no se abre toda la superficie desde el instante inicial, sino que se haría de forma progresiva). Por supuesto, se impondrá la restricción de que el área, que debe ser positiva, no debe ser mayor que el área total. Este tipo de restricciones puede ser tratado con los métodos anteriormente expuestos, incluso con los analíticos.

Ejemplo 1

Supóngase que se parte de una situación inicial con un número de empresas igual a 30. Recordando el análisis de sensibilidad por simulaciones intensivas que se reproduce de nuevo en la figura 9.4 se observa que en este caso el sistema tiende al declive si la superficie del área es

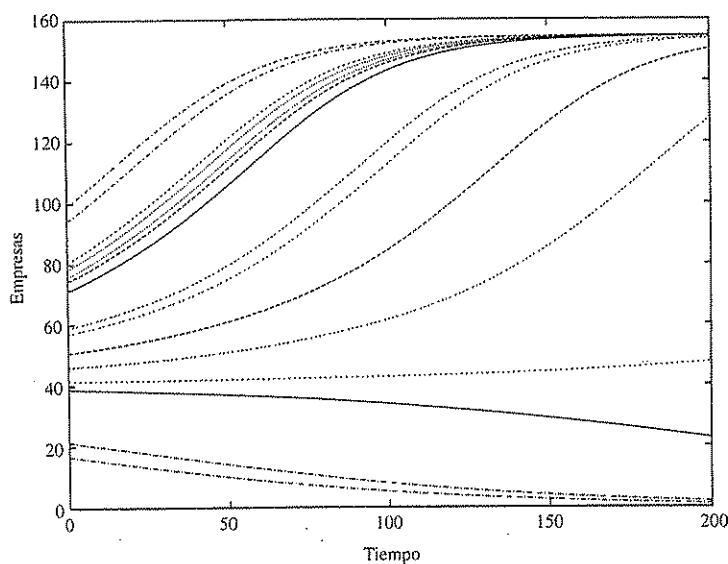


FIGURA 9.4. Modos de comportamiento del modelo de un área industrial

constante (igual a 2.000). Esto es debido a que se parte de un factor de ocupación muy pequeño que no permite una creación de empresas adecuada. Puesto que, como se está suponiendo ahora, la superficie es susceptible de variar, se puede hacer que el factor de ocupación inicial sea más alto (simplemente disminuyendo la superficie inicial). Incluso con unos sencillos cálculos se puede llegar al valor óptimo de la superficie inicial. Sin embargo, no es tan sencillo obtener el valor óptimo en los siguientes instantes de tiempo, para lo cual es necesario resolver el problema del control óptimo.

Para ello, por ejemplo, se puede plantear maximizar el número de empresas existentes en el instante $t = 150$, o lo que es lo mismo, maximizar la creación neta de empresas desde el instante inicial $t = 0$ hasta $t = 150$. Por tanto, el índice de optimización es

$$\mathcal{J}_1 = \int_0^{150} (\text{CE} - \text{DE})dt$$

Si se pretende plantear como un problema de minimización basta escoger

$$\mathcal{J} = - \int_0^{150} (\text{CE} - \text{DE})dt \quad (9.7)$$

El problema es encontrar la forma que debe tener esta evolución de la superficie de manera que se minimice el índice (9.7).

Las ecuaciones del sistema dinámico son las del modelo de un área industrial. Como este sistema es no lineal la resolución analítica de las ecuaciones (9.4), (9.5) y (9.6) no es en absoluto fácil. Sin embargo, sí se puede aplicar el método de optimización esbozado en la sección 9.3.2. Para ello, se debe discretizar el tiempo, por ejemplo en siete instantes $\{0, 25, 50, 75, 100, 125, 150\}$, y ejecutar el algoritmo. También se puede aplicar otro tipo de métodos como el de Montecarlo o los algoritmos genéticos. El resultado numérico aparece en la tabla 9.1 y se representa gráficamente en la figura 9.5.

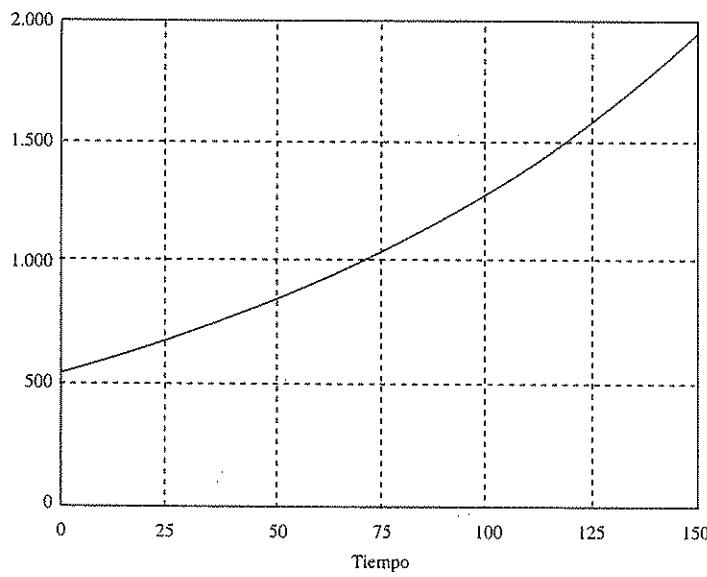


FIGURA 9.5. Evolución de la superficie del área industrial resultante del problema del control óptimo para el ejemplo I.

La evolución del número de empresas correspondiente a esta política de control aparece en la figura 9.6. Como se puede observar, se ha evitado el declive y se tiende hacia el punto de equilibrio superior. Si se representara la curva correspondiente al factor de ocupación se obtendría un valor constante igual a 0,4 que es el valor óptimo.

t	0	25	50	75	100	125	150
$S(t)$	562	692	852	1.049	1.292	1.591	1.958

TABLA 9.1. Solución del problema del control óptimo para el ejemplo 1.

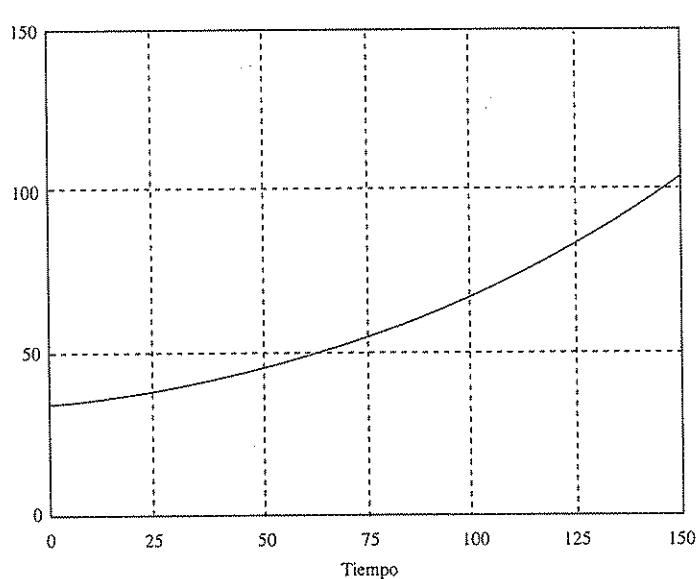


FIGURA 9.6. Evolución del número de empresas en el ejemplo 1.

Ejemplo 2

Supóngase ahora que el número inicial de empresas es igual a 70. Según la figura 6.2 se puede mantener la superficie el área industrial igual a su valor máximo desde el comienzo sin que se produzca el declive. Sin embargo, sí resultaría viable intentar que el área tuviera éxito lo antes posible. Por ejemplo, se puede pretender que en el instante $t = 50$ el número de empresas sea máximo. En este caso el índice que se desea minimizar será:

$$\mathcal{J} = - \int_0^{50} (\text{CE} - \text{DE}) dt$$

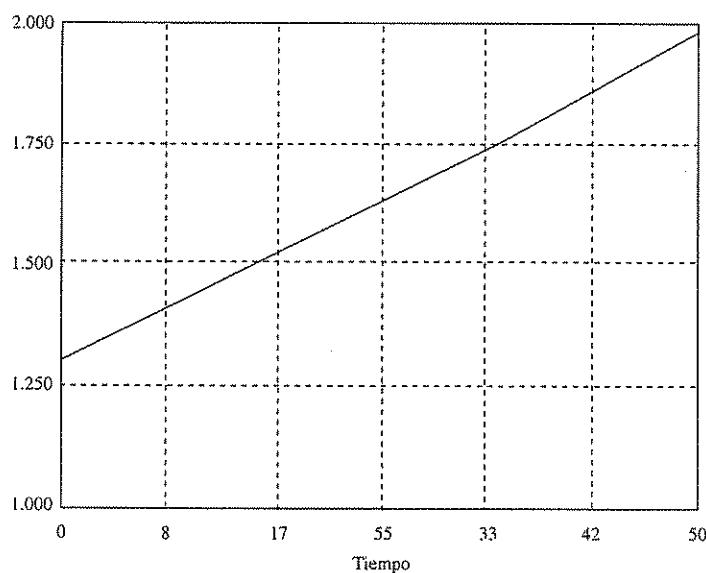
En este caso se va a discretizar el tiempo en cinco instantes: $\{0, 12, 5, 25, 37, 5, 50\}$. Resolviendo este problema de control óptimo se obtiene la señal de control que aparece en la tabla 9.2 y en la figura 9.7.

En la figura 9.8 se compara la evolución de las empresas en el caso base (superficie constante igual a 2.000) con la correspondiente a la solución del problema de control óptimo (dado que esta simulación se extiende a valores del tiempo más allá de $t = 50$, se ha completado la tabla 9.2 con el valor 2.000 para $t = 60$). Es evidente en esta figura la mejora introducida.

t	0	12,5	25	37,5	50
$S(t)$	1.312	1.456	1.616	1.793	1.990

TABLA 9.2. *Solución del problema del control óptimo para el ejemplo 2.*

Al igual que en el ejemplo 1, con la señal de control obtenida el factor de ocupación se mantiene en el valor óptimo durante el horizonte de control ($0 \leq t \leq 50$), tal y como se observa en la figura 9.9.

FIGURA 9.7. *Evolución de la superficie del área industrial resultante del problema del control óptimo para el ejemplo 2.*

Ejemplo 3

Los anteriores ejemplos corresponden a políticas de control en bucle abierto: en el instante inicial, el gestor calcula la evolución que debe seguir la superficie del área resolviendo el problema del control óptimo y ésta será la que aplicará aún en el caso en el que la evolución del sistema no sea exactamente igual a la predicha por el modelo.

Sin embargo, en este ejemplo tan simple se puede calcular, por procedimientos heurísticos, la política de control óptima en bucle cerrado (nótese que, a pesar de su sencillez, el modelo es no lineal, con lo que la obtención analítica de políticas de control en bucle cerrado no es nada sencilla). Como ya se ha mostrado en los ejemplos anteriores y, además, es fácil de intuir, la política óptima será aquella que mantenga al factor de ocupación en su valor óptimo (igual a

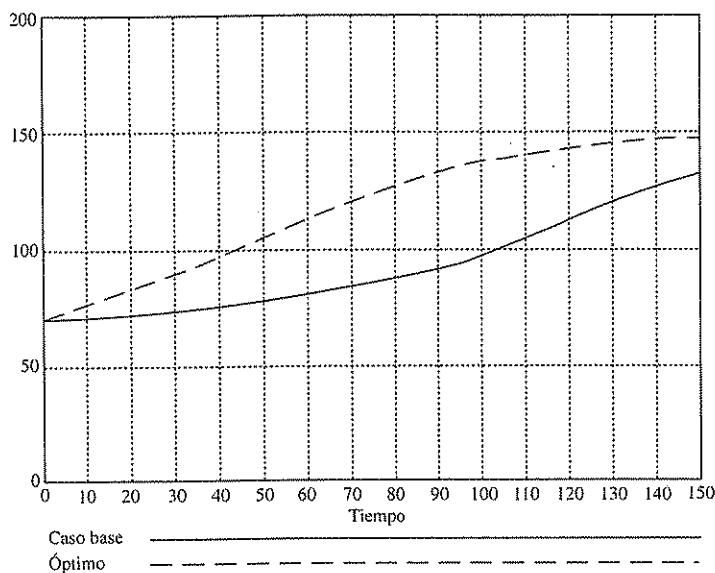


FIGURA 9.8. *Evolución del factor de ocupación en el caso base y con la señal de control óptima.*

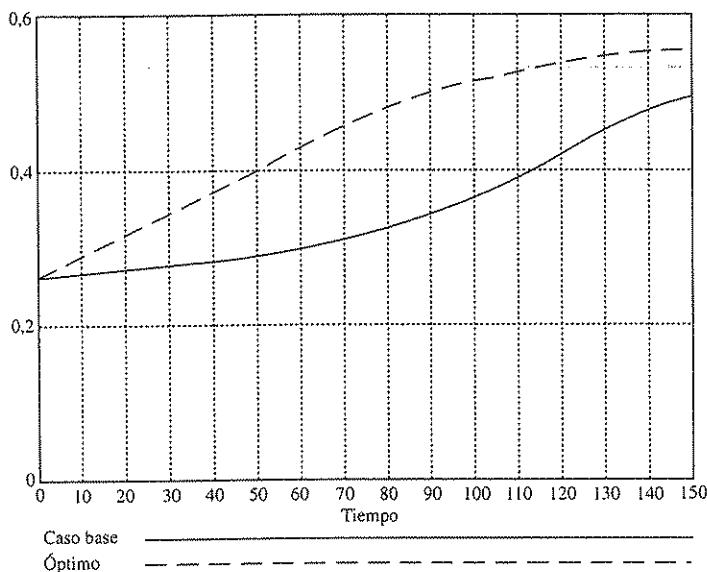


FIGURA 9.9. *Evolución del número de empresas en el caso base y con la señal de control óptima.*

0,4 en el ejemplo). De esta condición se puede extraer la política de control en bucle cerrado:

$$0,4 = FO = \frac{EMP \times SPE}{S} \implies S(EMP) = \frac{SPE}{0,4} \times EMP \quad (9.8)$$

Para mostrar la ventaja de este tipo de control vamos a volver al ejemplo 2. Supóngase que, en el instante $t = 20$, aparece un factor no contemplado en el modelo que hace que la tasa normal de creación de empresas TNCE disminuye a 0,047 (el valor anterior era 0,05). Evidentemente, la señal de control obtenida en el ejemplo 2 dejará de ser la óptima para este caso. Si se simula el modelo para esta nueva situación con la evolución de la superficie obtenida en el ejemplo 2 y la resultante de la ecuación (9.8) se obtienen las curvas que aparecen en la figura 9.10. Evidentemente, la política en bucle cerrado se adapta mejor a la nueva situación. Incluso se podría dar el caso (por ejemplo, si TNCE se hace igual a 0,045) en el que la política de control en bucle abierto llevara al sistema al declive y con la de bucle cerrado el área industrial tuviera éxito. La evolución de la superficie para los dos casos de la figura 9.10 se refleja en la figura 9.11.

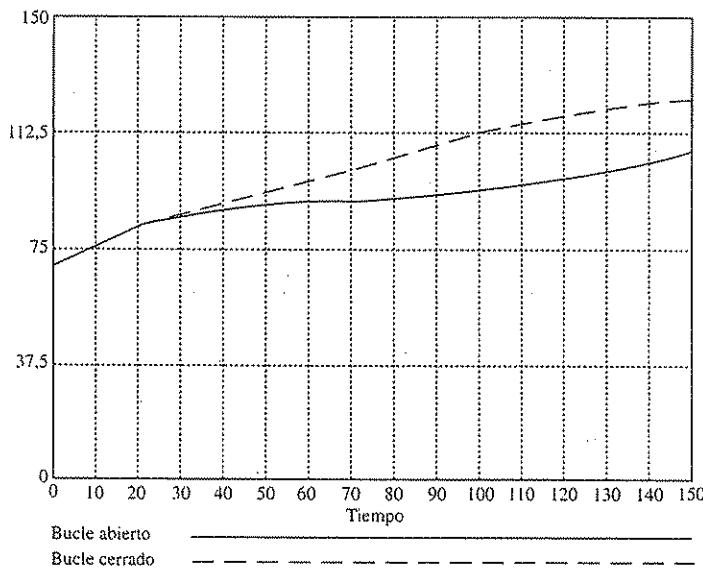


FIGURA 9.10. Evolución del número de empresas con control en bucle abierto y bucle cerrado.

9.3.5. Otras consideraciones

Resolviendo el problema del control óptimo se determina la actuación u^* que permite alcanzar el objetivo propuesto; es decir, optimizar el índice de funcionamiento J . Obsérvese que el resultado de la resolución de un problema de control óptimo es un enunciado del tipo 'hágase u^* ', sobre el mejor modo de alcanzar la meta propuesta. Es, por tanto, el caso paradigmático para la determinación de la acción racional, en un sentido limitado solamente

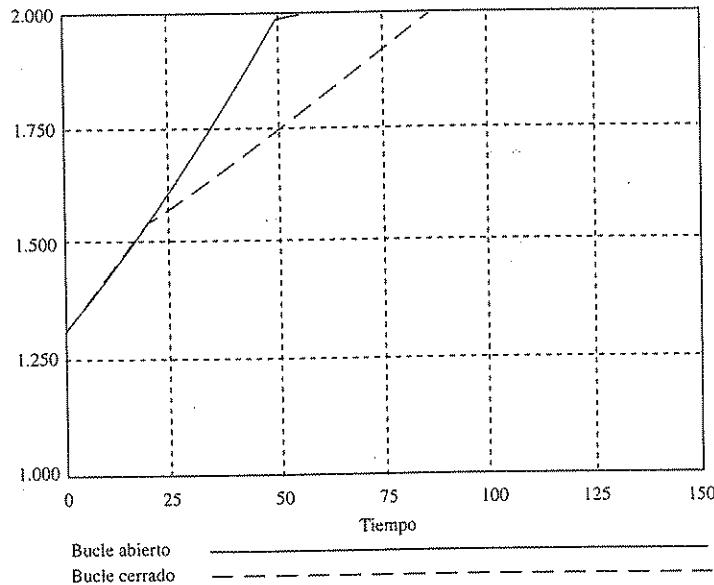


FIGURA 9.11. Evolución de la superficie con control en bucle abierto y bucle cerrado.

por la posibilidad de formalización (aunque no sea nada desdeñable esta limitación, como vamos a ver).

La aplicación práctica del formalismo del control óptimo a situaciones concretas adolece, sin embargo —además de las dificultades de resolución antes referidas—, del importante problema de que no siempre se dispone, por la propia naturaleza de los problemas, de formalizaciones como las requeridas por las expresiones (9.1) y (9.2) que posean suficiente precisión. Es decir, habitualmente se carece de un modelo \mathcal{M} y de una formulación de los objetivos \mathcal{J} que sean suficientemente ricos y precisos. Podría entonces proponerse el obtener a partir de un modelo (9.1) y de un índice de funcionamiento (9.2), ambos aproximados, una indicación, también aproximada, de la señal de actuación conveniente sobre el sistema \mathcal{S} . En algunos casos es lo que se hace, a falta de otra solución. Pero ello puede presentar importantes problemas, ya que el problema del control óptimo, tal como se ha descrito, es muy sensible a variaciones en el modelo o en el índice. Por ello parece más interesante afrontar el problema de la imprecisión en el modelo y en el índice de forma directa, para lo que, entre otras cosas, podemos recurrir al principio de racionalidad acotada de Simon⁴.

9.4. PRINCIPIO DE RACIONALIDAD ACOTADA

Ya hemos visto que la racionalidad se refiere a un modo de comportamiento que es apropiado para el logro de determinadas metas, dentro de los límites impuestos por ciertas condiciones

⁴ H. Simon, 1982, *Models of Bounded Rationality*, MIT Press.

y restricciones. Simon ha propuesto distinguir entre lo que llama *racionalidad substantiva* y *racionalidad procedimental*.

De un comportamiento se dice que es *substantivamente racional* si es el adecuado al logro de las metas que lo motivan. Esta forma de racionalidad presupone que el agente es capaz de conocer exhaustivamente las características objetivas de su entorno y de tenerlas presentes en el proceso de toma de decisión; es decir, se dispone de descripciones exhaustivas del sistema y de los objetivos a alcanzar. Se denomina también *racionalidad objetiva*. El proceso de control óptimo, que hemos visto en la sección anterior, presupone esta forma de racionalidad. En ese caso se puede decir con propiedad que la toma de decisiones es racional y consistente.

La racionalidad objetiva o substantiva puede caracterizarse por:

1. la existencia de un conjunto de alternativas perfectamente establecidas, entre las que debe escoger el agente decisor;
2. este agente posee conocimiento e información suficientes para predecir las consecuencias de cada una de las alternativas, y,
3. existe un criterio que permite decidir entre las distintas consecuencias que se puedan presentar.

Cuando se dan estas circunstancias es patente que la toma de decisiones es simple. La racionalidad objetiva es la racionalidad ideal, pero requiere un conocimiento exhaustivo, global y muy integrado. Por ello se da rara vez en la práctica, excepto en casos triviales.

Las posibilidades reales de decisión racional están acotadas por limitaciones en la información disponible, tanto en el modelo como en la formalización de los objetivos, y en la capacidad de procesarla. Estas limitaciones hacen que el comportamiento real se aleje mucho del considerado óptimo. Se recurre entonces a una *racionalidad procedimental* o *acotada*, que es la que de hecho se aplica cuando, en circunstancias normales, nos encontramos con una información y con una capacidad de procesarla limitadas. Esta forma de racionalidad trata de encontrar procedimientos eficientes para calcular soluciones efectivas a problemas concretos.

Los términos *optimización* y *satisfacción* han sido propuestos para matizar las diferencias entre los dos conceptos de racionalidad que estamos considerando. Para las situaciones en las que la complejidad y la falta de una información precisa hacen que un comportamiento de una racionalidad global no sea practicable, se propone un criterio de satisfacción, mediante el cual se trata de identificar soluciones accesibles y satisfactorias, mediante el empleo de procedimientos heurísticos, diferentes de la solución óptima que se determinaría algorítmicamente.

La distinción entre óptimo y satisfactorio puede ilustrarse con ayuda del juego del ajedrez. El ajedrez es un juego finito, puesto que lo son el número de situaciones que pueden darse y el de movimientos a partir de cada una de ellas, y, en consecuencia, es concebible la existencia de un algoritmo para llegar al jaque mate o a las tablas. Para ello habría que trazar el árbol de decisión para todas las jugadas posibles, que sería inmenso pero finito. Sin embargo, el número de ramas del árbol alcanzaría tal magnitud que el algoritmo no sería practicable. Un algoritmo, así concebido, corresponde a lo que hemos denominado racionalidad substantiva con respecto al juego del ajedrez. Presupone un conocimiento exhaustivo de todas las posibilidades, y una eventual optimización de este conocimiento para alcanzar el fin propuesto: el jaque mate.

Existen, sin embargo, programas informáticos que permiten ganar al ajedrez con un elevado nivel de fiabilidad, y en unos razonables períodos de tiempo. Estos programas no consideran el

árbol completo de posibilidades, sino exclusivamente unos cuantos niveles, e incluyen reglas heurísticas que priman aquellas jugadas que previamente han conducido a resultados satisfactorios. Con estos programas no existe garantía de que se conseguirá el jaque mate, sino sólo de que se jugará correctamente al ajedrez. Jugando correctamente se gana muchas veces, y en esa medida se dice que el programa es satisfactorio, aunque no óptimo. En el logro de resultados satisfactorios desempeñan un papel determinante las reglas heurísticas.

En la racionalidad objetiva subyace, más o menos explícitamente, la presunción de que existe un conocimiento —llamado *conocimiento objetivo*— con cuya posesión se está en disposición de decidir cuál será la mejor manera de actuar para alcanzar unas determinadas metas (que a su vez se supone que se pueden establecer de manera única). Simon cuestiona este punto de vista, no sólo porque duda de la posible existencia de ese conocimiento objetivo de la realidad (ese magno supermodelo que agotase su conocimiento), sino también porque considera que, aún en el caso de que existiera, las propias limitaciones humanas en su manejo lo harían poco útil. Debemos, por tanto, propone Simon, partir de una racionalidad acotada con todas sus consecuencias; establecer estrategias para decidir las reglas de acción buscando caminos alternativos en los que, desde un principio, se asuman esas limitaciones (como en el caso del ajedrez antes mencionado).

La necesidad de este enfoque se hace especialmente patente cuando se trata con lo que de una manera genérica se denominan sistemas complejos. Los sistemas que se modelan mediante la dinámica de sistemas pertenecen a esta clase, y, en consecuencia, a ellos (a sus modelos) se pueden aplicar las ideas desarrolladas en torno al concepto de racionalidad acotada que estamos comentando. A ello vamos a dedicar la siguiente sección.

9.5. EL PRINCIPIO DE RACIONALIDAD ACOTADA Y LA DINÁMICA DE SISTEMAS

Existe una conexión entre el principio de racionalidad acotada de Simon y la dinámica de sistemas⁵. Como hemos visto, el principio de racionalidad acotada establece que la capacidad de la mente humana para formular y resolver problemas complejos es pequeña comparada con la magnitud del problema cuya solución se requiere para alcanzar un comportamiento objetivamente racional, o aún para una razonable aproximación a esa racionalidad objetiva. La racionalidad acotada tiene importantes implicaciones en el modelado del comportamiento de sistemas en los que aparecen agentes humanos.

Puesto que los seres humanos tenemos una capacidad limitada de procesar información, hemos desarrollado una aptitud en virtud de la cual la percepción de información no es comprensiva, sino selectiva. Ante un determinado problema sólo consideramos como relevantes unos pocos factores. Aquellos que la experiencia que se tiene de situaciones análogas a la que ahora nos ocupa ha decantado como suficientemente sobresalientes. Al mismo tiempo, se rechaza todo aquello que es impreciso y remoto (aunque pudiera ser importante).

Cuando un cierto sistema alcanza un elevado nivel de complejidad hemos visto en el capítulo 8 que se puede recurrir a jerarquizar su estructura. Es lo que se hace cuando se

⁵ J. D. Morecroft, 1983, "System Dynamics: Portraying Bounded Rationality", *Omega*, Vol. 11, No. 2, 131-142; J. D. Morecroft. 1985, "Rationality in the Analysis of Behavioral Simulation Models", *Management Science*, volumen 31, No. 7, 900-916.

tiene una organización compleja (grandes empresas, instituciones u otras organizaciones), y se decide organizarla en unidades menores, cada una de las cuales pueda ser suficientemente abarcada por los gestores puestos al frente de ellas. De este modo se llega a una estructura jerarquizada en la que cada subunidad de la organización tiene asignadas metas concretas. Lo habitual es que la toma de decisiones en las subunidades se haga mediante reglas heurísticas.

Supongamos ahora que estamos representando esta organización compleja mediante un modelo de dinámica de sistemas. Las decisiones tomadas en cada una de las subestructuras se representan mediante las denominadas funciones de decisión, que no son sino ecuaciones de flujo, del tipo de las que hemos visto en el capítulo 3 (figura 9.12). Estas funciones de decisión están basadas normalmente en reglas heurísticas y conducen a decisiones que no son óptimas. Esto último podría invitar a descalificar a los modelos de dinámica de sistemas. Sin embargo, estos modelos pretenden ser descriptivos y no normativos. Pretenden simular el comportamiento de un sistema reproduciendo los procesos de toma de decisión en una forma lo más cercana posible a la real, y no incorporando supuestas decisiones óptimas.

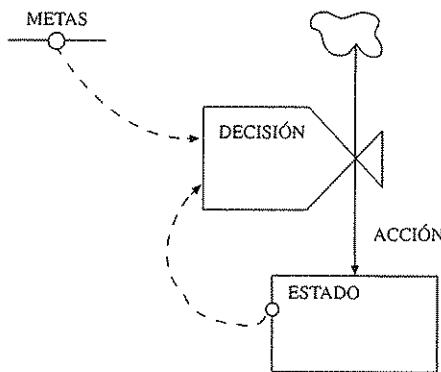


FIGURA 9.12. *Las ecuaciones de flujo como funciones de decisión.*

¿Cómo juzgar, en ausencia de un criterio óptimo de funcionamiento, la validez de una regla de decisión? Mediante lo que se conoce como *racionalidad local*. Se dice que una regla de decisión es localmente racional si da lugar a un comportamiento racional, enjuiciando éste en el supuesto de que el entorno sea tan simple como supone el que toma la decisión; es decir, si las premisas aceptadas por el agente fueran correctas.

Las funciones de decisión que representan los mecanismos de toma de decisión de diferentes agentes pueden ser localmente racionales, pero al considerar el sistema globalmente pueden producir un comportamiento irracional (en el sentido de no deseado; por ejemplo, un comportamiento oscilatorio en torno a la meta que se persigue). Es decir, se puede tener un comportamiento disfuncional como consecuencia de acciones localmente bien intencionadas.

Se ha realizado un interesante experimento con el fin de ilustrar cómo el comportamiento localmente racional de los agentes económicos puede conducir a resultados globalmente insatisfactorios. Este experimento está basado en un juego ideado hace unos treinta años en la Alfred Sloan School of Management del MIT, para ilustrar cómo la combinación de retardos en la transmisión de información y la estructura de realimentación negativa producen oscilaciones en el comportamiento de un sistema.

En ese juego se considera un proceso de distribución de cerveza, desde el fabricante al consumidor, mediante las cuatro etapas escalonadas que se indican en la figura 9.13. El fabricante atiende los pedidos del distribuidor, el distribuidor los del mayorista; este último los del minorista; y, finalmente, el minorista los del consumidor. La demanda de los consumidores se considera que inicialmente tiene un cierto valor, y que a las pocas jugadas sufre un incremento, que se mantiene constante el resto del juego.

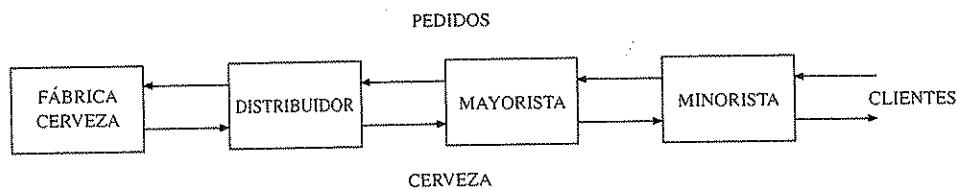


FIGURA 9.13. Los cuatro módulos del juego de la cerveza.

El juego consiste en que cuatro jugadores tratan de asumir los papeles de fabricante, distribuidor, mayorista y minorista de modo que cada uno de ellos satisface la demanda del nivel inferior y cursa pedidos al superior tratando, al mismo tiempo, de conseguir que su inventario se mantenga en un nivel razonable. Se penaliza tanto la acumulación excesiva como la escasez que determine que en determinadas circunstancias no se satisfaga la demanda. Para ello se establecen tanto costes asociados con el mantenimiento del inventario como costes correspondientes al no libramiento de pedidos (puesto que, en tal caso, se supone que se cursarían los pedidos a otros suministradores, y se perdería clientela). Por consiguiente, cada jugador debe mantener el nivel de inventario al menor nivel compatible con la restricción de no quedarse sin existencias. Si el inventario empieza a caer por debajo del nivel estimado como prudente deben cursarse los pedidos adecuados para restablecer el inventario. Si, por el contrario, empieza a producirse una acumulación excesiva, por disminución de la demanda, los pedidos deben reducirse.

El juego consiste en que, en cada ronda, cada jugador decide un pedido, al tiempo que se le satisface el que realizó un cierto tiempo atrás. Precisamente en la existencia de un retraso de tiempo entre la realización de un pedido y su cumplimiento por el nivel superior correspondiente está el origen de los desajustes que debe resolver el jugador agente decisor.

Cada jugador pretende tomar decisiones que sean racionales, en el sentido de que le permitan alcanzar su objetivo: mantener el inventario en el nivel deseado (con lo que disminuye los costes que penalizan un juego deficiente). Para ello dispone de una información consistente en el conocimiento del nivel de su inventario y de su cartera de pedidos, de los valores previos de esas variables, y otra información análoga. Al mismo tiempo, los participantes conocen el funcionamiento global de la red de distribución. A partir de ello toman decisiones que resultan ser sólo localmente racionales.

La característica más notable de este juego son las oscilaciones que se derivan del comportamiento de los jugadores, oscilaciones que crecen en amplitud del minorista al mayorista, del mayorista al distribuidor y del distribuidor al fabricante. Resulta notable el observar que, como consecuencia de un incremento en la demanda de los consumidores, la producción sufre una amplificación por un factor que normalmente es mayor que seis. Este fenómeno de ampli-

ficación está íntimamente asociado con los retrasos que se producen a lo largo de la cadena. Supóngase que un determinado jugador percibe un incremento considerable en su demanda. Antes de decidirse a incrementar significativamente sus pedidos, dudará tratando de determinar el carácter más o menos permanente de ese incremento (ya que si se trata de un incremento en la demanda de corta duración podrá satisfacerla con el inventario disponible; precisamente el objetivo de un inventario es absorber las componentes de alta frecuencia en las fluctuaciones de la demanda). Sin embargo, debido a ese titubeo, al que se suman los retrasos en los pedidos y en los libramientos, la demanda puede hacer que se desborde la capacidad del inventario para satisfacerla. Por ello, y para que el inventario vuelva a su nivel deseado, el jugador incrementará sensiblemente los pedidos a su suministrador, por encima de los pedidos que él recibe. De este modo se convierte en un factor de amplificación de la fluctuación en la demanda a la que se ha visto sometido.

Como consecuencia de todo ello, se tiene un comportamiento que dista considerablemente del óptimo. Los jugadores comprueban que son incapaces de alcanzar un modo de operación estable para el sistema y que se producen fluctuaciones de gran amplitud.

Los resultados del experimento muestran que los agentes reales, aunque pretenden comportarse racionalmente, lo hacen de modo que su comportamiento está muy alejado del óptimo. Los agentes son víctimas de una percepción errónea de la realimentación; es decir, no alcanzan a tener en cuenta los resultados de acciones de control que han sido iniciadas, pero que aún no han producido todo su efecto, debido a los retrasos a lo largo del bucle de realimentación. De este modo aparecen comportamientos oscilatorios, que en algunos casos pueden ser particularmente erráticos. Aunque los agentes pretenden comportarse racionalmente lo hacen de modo que su comportamiento resulta estar muy alejado no ya del óptimo, sino, en muchos casos, del aceptable. La interacción entre las acciones individuales con la estructura de la red de distribución simulada produce una dinámica conjunta que difiere significativa y sistemáticamente del comportamiento óptimo.

Con ello se tiene una ilustración del hecho siguiente. En determinadas situaciones, las causas de un comportamiento no deseado (por ejemplo, la presencia de oscilaciones) suelen imputarse a factores externos, cuando en realidad está generado por las propias acciones tomadas en el seno del sistema, que, aunque se pretenden racionales, en realidad son la causa de ese comportamiento no deseado. Todo sucede como si los agentes que toman las decisiones poseyeran un modelo mental del sistema sobre el que actúan en el que se prescinde del efecto combinado de las realimentaciones y los retrasos a lo largo de la estructura. Los agentes se comportan como si asumieran que el comportamiento se generase a partir de actuaciones exógenas, prescindiendo de la realimentación presente en el sistema. Este modelo mental recoge la tendencia a achacar a factores externos la causa del comportamiento, sin percibir que es en su propia estructura donde residen las causas de ese comportamiento. Es el propio modelo mental de los agentes, limitado e incorrecto, el responsable de la ineficiencia de sus acciones.

Se va a dedicar este apéndice a describir las características principales de los entornos informáticos específicos de la dinámica de sistemas. No se pretende comparar entre sí estos entornos, entre otros motivos porque se trata de productos que evolucionan continuamente, de manera que constantemente salen al mercado nuevas versiones con nuevas características que harían que muy pronto la comparación quedara obsoleta.

Existen entornos para diversos tipos de computadores (normalmente PC o Macintosh). Entre los más extendidos se encuentran DYNAMO, STELLA, ITHINK, POWERSIM y VENSIM.

Ya se ha comentado que existen en estos entornos dos formas distintas de introducir un modelo: en algunos el modo de trabajar es similar al de los lenguajes de programación de propósito general, es decir, con un editor de texto se escriben las instrucciones que definen el modelo; en otros, sin embargo, la programación es de forma gráfica: se empieza dibujando en el computador el diagrama de influencias o el diagrama de Forrester del modelo y después se completa la información introduciendo las ecuaciones o constantes necesarias para cada uno de los elementos. Algunos entornos permiten trabajar con ambas modalidades.

Los lenguajes de programación que utilizan estos entornos, además de permitir escribir las ecuaciones de las variables de nivel, de flujo y auxiliares, llevan incorporadas diversas funciones para hacer más cómoda la escritura de un modelo. Así, normalmente tienen funciones para modelar retrasos, calcular el mínimo o máximo de dos cantidades, realizar operaciones alternativas (como la que se denominó IF en la sección 4.6), etc. En el caso de que el programador necesite utilizar una función compleja que no viene incorporada en el lenguaje es posible, a veces, escribir esta función en un lenguaje de propósito general (como el C). Sin embargo, éste es un procedimiento complejo que solamente se recomienda a programadores avanzados.

Otra característica de estos lenguajes es la posibilidad de utilizar la notación de vectores y matrices. Con ello se consigue evitar escribir ecuaciones muy parecidas abreviándolas en una sola. Por ejemplo, al escribir un modelo demográfico, con la población desagregada por grupos de edades (niños, jóvenes, adultos y ancianos), si alguna ecuación es idéntica para los cuatro grupos se puede resumir en una sola ecuación indicando que se debe aplicar a cada

uno de los grupos. De esta manera se pueden escribir de forma cómoda modelos en los que el número de grupos es muy grande. Además, suelen existir funciones para el tratamiento de estas matrices como, por ejemplo, para calcular la suma de todos los elementos.

Respecto al tamaño máximo de los modelos que se pueden construir, los entornos han evolucionado mucho desde sus orígenes. En un principio, el tamaño máximo venía fijado por las limitaciones del entorno. Hoy en día, con el uso de sistemas operativos con memoria virtual, este límite no viene impuesto por la herramienta informática, sino por la capacidad del ordenador que se esté utilizando. Esto permite, en caso de llegar al límite, seguir trabajando con tan sólo ampliar o cambiar el computador, en lugar de tener que cambiar de lenguaje (con la consiguiente pérdida de tiempo en traducir el modelo).

Otras características de los entornos informáticos para la dinámica de sistemas son las relacionadas con la manipulación de los resultados de las simulaciones. Suelen incorporar funciones de presentación gráfica de estos resultados, incluyendo la posibilidad de comparar los de varias simulaciones entre sí. También permiten presentar los datos de forma tabular. Otra alternativa es la posibilidad de exportar los datos para su tratamiento posterior con otro tipo de entornos, como, por ejemplo, una hoja de cálculo. También se suele permitir la operación opuesta: importar datos de otro programa para utilizarlos como variables exógenas. Algunos entornos incorporan también funciones de análisis estadístico de los datos.

Una posibilidad muy interesante que tienen ciertos entornos es la de utilizar funciones de optimización como ya se comentó en la sección 6.4. De esta manera se pueden ajustar los parámetros del modelo resolviendo un problema de optimización; o también es posible encontrar los valores que hacen que el comportamiento del modelo sea óptimo (por ejemplo, que se maximicen los beneficios de una empresa).

Por último, suele existir la posibilidad de "jugar" con el modelo. Con este término (del inglés *game*) se alude a la característica de algunos entornos que permite que, en el transcurso de una simulación y en determinados instantes de tiempo, el usuario pueda cambiar los valores de ciertas variables. Estas variables suelen estar relacionadas con magnitudes sobre las que, en el sistema que se pretende modelar, un agente humano debe tomar decisiones. Estos juegos pueden tener dos objetivos diferenciados: por un lado, permiten entrenar al usuario para un posterior enfrentamiento al sistema real; por otro, permiten incluir al agente humano en el modelo, con lo que se pueden hacer simulaciones sin tener que modelar el proceso de toma de decisiones del pensamiento humano. En este último caso se suelen hacer experimentos variando la persona que juega con el modelo para, posteriormente, hacer un tratamiento estadístico de los resultados. En la sección 9.5 se comentó un ejemplo de uno de estos experimentos.

Bibliografía básica

- FORRESTER, J. W. (1961): *Industrial Dynamics*. Productivity Press.
- FORRESTER, J. W. (1968): *Principles of Systems*. Productivity Press.
- MEADOWS, D. L. (1970): *Dynamics of Commodity Production Cycles*. Productivity Press.
- FORRESTER, J. W. (1975): *Collected Papers of Jay W. Forrester*. Productivity Press.
- MASS, N. J. (1975): *Economics Cycles: An Analysis of Underlying Causes*. Productivity Press.
- MARTÍNEZ VICENTE, S. (1986): *Dinámica de sistemas* (2 volúmenes). Alianza Editorial.
- MARTÍNEZ VICENTE, S., y A. REQUENA (1988): *Simulación dinámica por ordenador*. Alianza Editorial.
- COYLE, R. G. (1996): *System dynamics modelling: a practical approach*. Chapman and Hall.

Dinámica urbana

- FORRESTER, J. W. (1969): *Urban Dynamics*. Productivity Press.
- ALFELD, L. E., y A. K. GRAHAM (1976): *Introduction to Urban Dynamics*. Productivity Press.

Modelos globales

- MEADOWS, D. H., D. L. MEADOWS, J. RANDERS y W. W. BEHRENS III (1972): *Los límites del crecimiento*. Fondo de Cultura económica.
- FORRESTER, J. W. (1973): *World Dynamics* (2.ª ed.). Productivity Press.
- MEADOWS, D. L., y D. H. MEADOWS (eds.) (1974): *Toward Global Equilibrium: Collected Papers*. Productivity Press.
- MEADOWS, D. H., D. L. MEADOWS y J. RANDERS (1994): *Más allá de los límites del crecimiento* (3.ª ed.). El País-Aguilar.

Aplicaciones en gestión

- COYLE, R. G. (1977): *Management System Dynamics*. John Wiley & Sons.
- ROBERTS, E. B. (ed.) (1978): *Managerial Applications of System Dynamics*. Productivity Press.

198 Dinámica de sistemas

- LYNEIS, J. M. (1980): *Corporate Planning and Policy Design*. Productivity Press.
SENGE, P. M. (1992): *La quinta disciplina*. Granica.
MORECROFT, D. W., y STERMAN, John (eds.) (1994): *Modelling for learning organizations*. Productivity Press.

Modelos económicos

- STERMAN, J. D. (1985): «A Behavioral Model of the Economic Long Wave». *Journal of Economic Behavior and Organization* 6 (1): 17-53.
FORRESTER, J. W. (1989): «The System Dynamics National Model: Macrobbehavior from Microstructure». En P. M. Milling y E. O. K. Zahn (eds.), *Computer-Based Management of Complex Systems: International System Dynamics Conference*. Springer-Verlag.

Metodología

- LEGASTO, A. A. Jr., J. W. FORRESTER y J. M. LYNEIS (eds.) (1980): *System Dynamics. TIMS Studies in the Management Sciences*, vol. 14. North-Holland.
RANDERS, J. (ed.) (1980): *Elements of the System Dynamics Method*. Productivity Press.
RICHARDSON, G. P., y A. L. PUGH, III (1981): *Introduction to System Dynamics Modeling with DYNAMO*. Productivity Press.
ROBERTS, N., D. F. ANDERSEN, R. M. DEAL, M. S. GRANT y W. A. SHAFFER (1983): *Introduction to Computer Simulation: A System Dynamics Modeling Approach*. Addison-Wesley.
WOLSTENHOLME, E. F. (1990): *System Enquiry - A System Dynamics Approach*. John Wiley.
VENNIX, J. A. M. (1996): *Group model building. Facilitating team learning using system dynamics*. John Wiley & Sons.



3418168

La *Dinámica de sistemas* nació durante los años cincuenta, de la mano de J. W. Forrester. Desde entonces su aplicación se ha extendido a disciplinas muy diferentes: además de haberse utilizado para los sistemas industriales y socioeconómicos de ámbito urbano o regional, se ha aplicado al área de los sistemas ecológicos y medioambientales, así como al suministro de recursos energéticos y a la defensa nacional entre otros. Basada en el comportamiento de sistemas mediante la construcción de un modelo de simulación informática que ponga de manifiesto las relaciones entre la estructura del sistema y su comportamiento, constituye en la actualidad una de las herramientas sistemáticas más

sólidamente desarrolladas y que mayor grado de aceptación e implantación han alcanzado. En esta obra se exponen los elementos básicos del lenguaje sistemático, utilizando ejemplos concretos y sirviéndose de técnicas complementarias que ayudan al análisis de los modelos. Así mismo, se analizan los problemas asociados a los sistemas complejos y se explican las técnicas de explotación de modelos. Esta obra parte de la *Introducción a la dinámica de sistemas* de Javier Aracil, que ha sido reescrita por completo —juntamente con Francisco Gordillo— para esta nueva edición, ampliándose y actualizándose sus contenidos para adaptarse mejor a las necesidades de estudiantes e investigadores.

Alianza Editorial

ISBN 978-84-206-8168-9



A standard linear barcode representing the ISBN 978-84-206-8168-9. Below the barcode, the numbers 9 788420 681689 are printed vertically.