

AULA DE CIÈNCIA
I CULTURA, 13

Models matemàtics en la ciència i la societat

Cicle Ferran Sunyer i Balaguer



Carles Casacuberta (ed.)

Salvador Benito

M. Teresa Brugal

Pere Caminal

Ildfonso Díaz

Daniel Devolder

Beatriz Giraldo

Eugenio Oñate

Pere Puig

Marta Sanz

Joan Solà-Morales

Montserrat Vallverdú

(AULA DE CIÈNCIA
I CULTURA, 13

Models matemàtics en la ciència i en la societat

Cicle Ferran Sunyer i Balaguer

Carles Casacuberta (ed.)

Salvador Benito

Daniel Devolder

Marta Sanz

M. Teresa Brugal

Beatriz Giraldo

Joan Solà-Morales

Pere Caminal

Eugenio Oñate

Montserrat Vallverdú

Ildefonso Díaz

Pere Puig

Aquesta obra recull les conferències pronunciades entre
el 7 de març i el 25 d'abril de 2001

Catalogació en publicació de la
FUNDACIÓ CAIXA DE SABADELL

Models matemàtics en la ciència i la societat. - (Aula de ciència i cultura ; 13)

Conté: Presentació del cicle / Carles Casacuberta ; Posibilidades de los métodos numéricos en el mundo industrial / Eugenio Oñate ; Modelización, análisis y control de sistemas climáticos / Ildefonso Díaz ; Captura-recaptura: un mètode per calcular el nombre de consumidors de drogues / Pere Puig, M. Teresa Brugal ; Anàlisi no lineal del patró ventilatori / Salvador Benito, Pere Caminal, Beatriz Giraldo, Montserrat Vallverdú ; Aprender a modelitzar amb equacions diferencials / Joan Solà- Morales ; Matemáticas y demografía. Del tiempo al espacio / Marta Sanz.

- I. Fundació Caixa de Sabadell
1. Matemàtiques - Conferències
2. Fundació Caixa de Sabadell - Conferències
51(042)
-

Edita: Fundació Caixa de Sabadell, 2002

carrer d'en Font, 25, 08201 - Sabadell

Tel.: (93) 725 95 22 A/e: fundacio@caixasabadell.org

© Carles Casacuberta (ed.), Salvador Benito, M. Teresa Brugal, Pere Caminal, Ildefonso Díaz, Daniel Devolder, Beatriz Giraldo, Eugenio Oñate, Pere Puig, Marta Sanz, Joan Solà-Morales i Montserrat Vallverdú.

Fotocomposició i producció: *Addenda*, s.c.c.l., Pau Claris, 92, Barcelona

Impressió: Winihard Gràfics, S.L., Polígon industrial, Av. del Prat s/n, Moià

Impressió sobre paper ecològic

ISBN: 84-95166-37-2

Dipòsit legal: B - 22.191 - 2002

Índex

Presentació del cicle	7
<i>Carles Casacuberta</i>	
POSIBILIDADES DE LOS MÉTODOS NUMÉRICOS EN EL MUNDO INDUSTRIAL	
<i>Eugenio Oñate</i> , CIMNE, Universitat Politècnica de Catalunya	9
MODELIZACIÓN, ANÁLISIS Y CONTROL DE SISTEMAS CLIMÁTICOS	
<i>Ildefonso Díaz</i> , Universidad Complutense de Madrid	35
CAPTURA-RECAPTURA: UN MÈTODE PER CALCULAR EL NOMBRE DE CONSUMIDORS DE DROGUES	
<i>Pere Puig</i> , Universitat Autònoma de Barcelona	
<i>M. Teresa Brugal</i> , Institut Municipal de la Salut, Barcelona	63
ANÀLISI NO LINEAL DEL PATRÓ VENTILATORI	
<i>Salvador Benito</i> , Hospital de Sant Pau, Barcelona	
<i>Pere Caminal</i> , CREB, Universitat Politècnica de Catalunya	
<i>Beatriz Giraldo</i> , CREB, Universitat Politècnica de Catalunya	
<i>Montserrat Vallverdú</i> , CREB, Universitat Politècnica de Catalunya . . .	79
APRENDRE A MODELITZAR AMB EQUACIONS DIFERENCIALS	
<i>Joan Solà-Morales</i> , Universitat Politècnica de Catalunya	93
MATEMÁTICAS Y DEMOGRAFÍA. DEL TIEMPO AL ESPACIO	
<i>Daniel Devolder</i> , CED, Universitat Autònoma de Barcelona	113
EVOLUCIONS ALEATÒRIES: ALGUNS PROBLEMES MATEMÀTICS DE LA SEVA MODELITZACIÓ	
<i>Marta Sanz</i> , Universitat de Barcelona	133

Presentació del cicle

MODELS I SIMULACIONS

Cada vegada més, els científics troben camins de col·laboració entre especialitats ben diverses. En matemàtiques, aquesta tendència ha anat unida a un interès creixent per fer-les més accessibles i implicar-les a gairebé tot arreu. Potser per una pressió social, o potser per l'afany de buscar objectius nous, el cas és que els matemàtics es comuniquen ara millor amb el seu entorn que en el passat recent.

Els models matemàtics són dissenys de teories i tècniques per a resoldre problemes nascuts en altres branques de la ciència o en la indústria. La base d'aquests models és sovint el càlcul diferencial, que va néixer al segle XVII com a eina auxiliar per a l'astronomia i la mecànica, i que encara avui és la tècnica de treball més important que els científics i els enginyers tenen al seu abast. Una equació diferencial o un conjunt d'equacions diferencials poden arribar a descriure perfectament l'evolució dels sistemes que es volen estudiar.

En molts casos, l'objectiu d'un model matemàtic és arribar a solucions efectives del problema que ha motivat el model, mitjançant mètodes numèrics i amb l'ajuda d'ordinadors. Moltes vegades no és possible (ni es pretén) obtenir solucions exactes. N'hi ha prou amb solucions aproximades que siguin útils als usuaris que han plantejat el problema. Per exemple, la resistència d'una columna en un edifici es pot descriure mitjançant unes equacions diferencials concretes, gràcies a les quals es pot preparar una simulació informàtica que permeti determinar amb molta precisió en quines condicions es trencarà la columna. Altres vegades, no es busca cap solució del model, sinó que es fa un estudi qualitatiu de l'evolució al llarg del temps del sistema que es considera. Així es poden descriure, per exemple, les variacions del clima de la Terra en el passat i en el futur.

Les matemàtiques discretes i la teoria de la probabilitat són altres especialitats més recents que ocupen un lloc preeminent en l'elaboració de models matemàtics. Cada vegada té menys pes el determinisme en els processos naturals o socials i s'utilitzen més els models probabilístics, que permeten descriure evolucions aleatòries molt ajustades a la realitat. Això es fa, per exemple, en el món financer o quan es volen tenir en compte elements imprevisibles en l'evolució del clima.

CONTINGUT DEL CICLE

Aquesta ha estat la segona sèrie de conferències del Cicle Ferran Sunyer i Balaguer, que es va iniciar l'any 1999 amb sis presentacions sobre les bases matemàtiques de la civilització tecnològica. Esperem que aquesta iniciativa cultural de la

Fundació Caixa de Sabadell i el Centre de Recerca Matemàtica tindrà continuïtat i permetrà anar mostrant diversos aspectes de la interacció de les matemàtiques amb la societat. Està resultant especialment útil per als estudiants de les universitats catalanes, que d'aquesta manera poden conèixer facetes de les matemàtiques més properes a àmbits de treball no necessàriament vinculats al món acadèmic. Per segona vegada, els estudiants han respost de manera ben positiva i han participat activament en el cicle, fent preguntes i redactant una col·lecció de treballs que han assolit una qualitat molt alta.

Tres de les conferències del cicle d'enguany (impartides per Eugenio Oñate, Ildefonso Díaz i Joan Solà-Morales) han dibuixat una bona part del que es pot fer amb equacions diferencials, tant a nivell de models teòrics com en diverses aplicacions a l'enginyeria, la climatologia, l'aeronàutica o la biologia (algunes d'elles realment impressionants). La conferència impartida conjuntament per Salvador Benito i Pere Caminal ha mostrat les tècniques matemàtiques que s'utilitzen per a l'anàlisi de la respiració assistida en alguns hospitals, en col·laboració amb laboratoris d'enginyeria biomèdica. La conferència de Daniel Devolder s'ha centrat en models matemàtics a la demografia, tant per a l'estudi de l'evolució de les poblacions al llarg del temps com per a la seva distribució territorial. La que han impartit Pere Puig i Teresa Brugal ha estat de caire estadístic i ha mostrat de quina manera es poden fer estimacions aproximades de les dimensions de col·lectius amagats, com ara els consumidors de drogues en una ciutat gran. La conferència de Marta Sanz ha tancat el cicle i ha ofert una visió molt escaient de la modelització d'evolucions aleatòries mitjançant equacions diferencials estocàstiques.

Amb aquest cicle s'ha volgut oferir una mostra variada i substancial d'allò que s'està fent en modelització al nostre entorn. No s'ha pretès reunir exemples de totes les línies d'investigació, que són moltes i molt diverses. Algunes conferències han descrit avenços recents i d'altres han estat de caire més general, fins i tot amb una clara intenció didàctica. Totes anaven adreçades a un públic ampli: als estudiants que volguessin descobrir aplicacions de les matemàtiques a temes actuals; als investigadors interessats en projectes innovadors dels seus col·legues, i a tots aquells que tinguessin ganes d'admirar la grandesa i les immenses possibilitats de les matemàtiques.

Carles Casacuberta

Professor Titular de la Universitat Autònoma de Barcelona

Director del cicle

Posibilidades de los métodos numéricos en el mundo industrial

Eugenio Oñate

Centre Internacional de Mètodes Numèrics en Enginyeria

Universitat Politècnica de Catalunya

INTRODUCCIÓN

Se denominan métodos numéricos las técnicas de la matemática que permiten expresar la solución de un problema en forma de números. Para muchos, los métodos numéricos son una rama de la matemática aplicada, es decir, aquella parte de la matemática interesada en la resolución de problemas que afectan directa o indirectamente a los intereses del hombre. El término «solución numérica» se utiliza a menudo frente al de «solución analítica» de un problema (también denominada «solución exacta»). La diferencia entre ambos tipos de solución es sustancial. Consideremos como ejemplo el estudio del comportamiento de un sistema de tipo físico o bien humanístico. Dicho sistema satisface unas leyes (ecuaciones) generales de tipo matemático (denominadas ecuaciones de gobierno); por ejemplo, unas ecuaciones o inecuaciones diferenciales o algebraicas en las que intervienen un conjunto de variables y parámetros físicos. La solución analítica es una expresión matemática que proporciona toda la información sobre el comportamiento del sistema, para cualquier valor de las variables y parámetros que intervienen en las ecuaciones de gobierno. Es, por tanto, la solución «universal» del problema a la que aspiraban los pitagóricos, Platón y tantos otros creyentes en la numerización del mundo. Por otro lado, la solución numérica expresa el comportamiento del sistema en función de números que se obtienen resolviendo las ecuaciones de gobierno para valores concretos de las variables y los parámetros del sistema. Como ejemplo, planteemos el problema de encontrar la ecuación que expresa todas las trayectorias posibles de un móvil que se desplaza en línea recta partiendo siempre del mismo punto. Obviamente, la solución general del problema es $y = kx$, donde y es la coordenada de un punto de la trayectoria, x el valor de la abscisa y k la pendiente de la recta. Dicha ecuación es universal y expresa el conjunto infinito de rectas que pasa por el origen. Dando valores a la pendiente k y a la abscisa x encontraremos las coordenadas de los puntos de las trayectorias seleccionadas, caracterizadas por los valores de x e y .

Un ejemplo clásico de solución numérica fue la obtenida por Arquímedes para el valor aproximado del número π a partir de la división de la circunferencia en polígonos, obtenidos incrementando progresivamente el número de lados y dividiendo el perímetro de cada polígono por el radio del círculo. Obviamente, al aumentar el número de lados

de cada polígono, aumentaba también la precisión en el valor del número π buscado. Tomando polígonos inscritos y circunscritos en la circunferencia de hasta 96 lados, Arquímedes logró acotar el valor de π entre 3,14084 y 3,14285 [1].

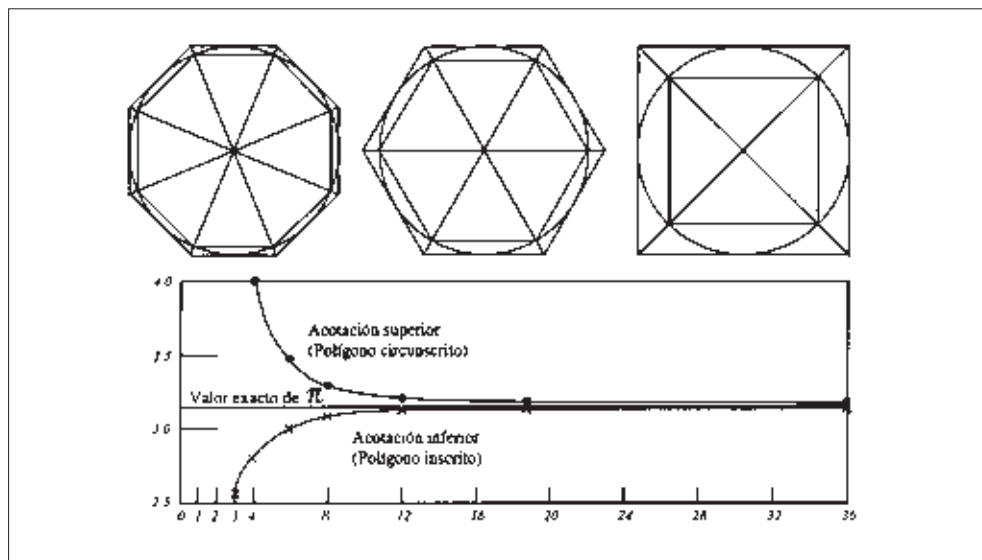


Figura 1. Método utilizado por Arquímedes para calcular el valor del número π a partir del perímetro de polígonos inscritos y circunscritos en una circunferencia.

La técnica utilizada por Arquímedes sintetiza la esencia de los métodos numéricos. La solución numérica se obtiene dividiendo el dominio que se estudia (la circunferencia) en formas geométricas sencillas (rectas), de las que se conocen todas sus propiedades (longitud). Se observa también que la solución numérica es aproximada y mejora (converge) al incrementar el número de divisiones del dominio. Finalmente, es muy importante advertir que la solución «exacta» del problema (el valor exacto de π , uno de los números inconmensurables) es desconocida y la solución numérica es la única alternativa, como sucede en la mayoría de los problemas prácticos.

Es interesante destacar que una técnica similar de «divide y vencerás» fue utilizada por los calculistas chinos en el siglo V d.C. para obtener una acotación del valor de π mediante divisiones del área del círculo en rectángulos inscritos y circunscritos a la circunferencia (Figura 2) [2].

Los métodos numéricos, por tanto, buscan números, mientras que los métodos analíticos buscan fórmulas matemáticas. Obviamente, la solución analítica, al ser universal, contiene todas las soluciones numéricas, mientras que de la solución numérica de un problema es imposible, en general, deducir la solución analítica.

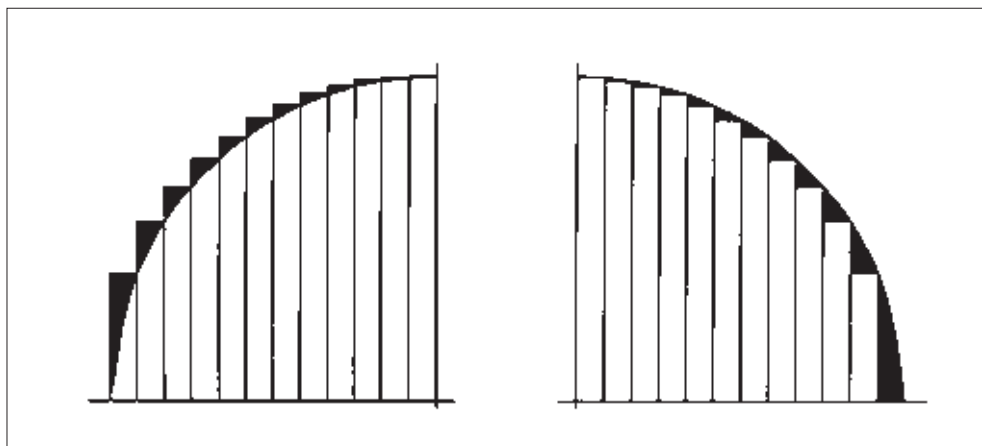


Figura 2. Aproximación del área del círculo mediante suma de áreas de rectángulos inscritos y circunscritos. Las zonas negras indican el error en el cálculo del área. Este método fue utilizado para evaluar el número π por matemáticos e ingenieros chinos en el siglo V d.C.

NECESIDAD Y LÍMITES DE LAS ECUACIONES DIFERENCIALES

Tanto los métodos numéricos como los analíticos comparten un punto inicial fundamental: la necesidad de plantear en forma matemática el problema a resolver. Durante miles de años, hasta el descubrimiento del cálculo infinitesimal por Newton y Leibniz, dichas formas matemáticas eran variantes de expresiones polinómicas, destinadas muchas veces a resolver problemas de índole geométrico (tal como encontrar el valor del número π). La revolución del cálculo infinitesimal estribó fundamentalmente en su capacidad de expresar cualquier problema de la naturaleza en forma matemática por medio de *ecuaciones diferenciales*. Al mismo tiempo, Newton, Leibniz y tantos otros insignes sucesores (Euler, Cauchy, Lagrange, Navier, Gauss, etc.) sentaron las bases para resolver dichas ecuaciones diferenciales de forma analítica [3-6].

El profundo significado que para las matemáticas representaron los nuevos métodos de cálculo diferencial e integral propuestos por Newton y Leibniz es comparable al que el descubrimiento del fuego tuvo para los hombres primitivos, o el de la electricidad para la revolución industrial.

Esta afirmación no es en absoluto exagerada. Antes de Newton y Leibniz no existía una metodología general para plantear en forma de ecuaciones matemáticas un problema concreto de la física, tal como, por ejemplo, la propagación del calor en un cuerpo, el flujo de un fluido o el equilibrio de un sólido elástico. Obviamente, al no poder plantearse el problema, su solución era imposible. Después de las aportaciones de

Newton y Leibniz, no sólo fue ya posible describir el comportamiento de cualquier sistema físico, fuera éste un sólido, un líquido o un gas, mediante ecuaciones diferenciales e integrales, sino que se dispuso de técnicas para resolverlas en muchos casos que, aunque usualmente eran simplificaciones del problema más general, permitieron avances significativos en el conocimiento científico y técnico. Así, mientras las matemáticas, como ciencia autónoma, exploraban nuevos campos de abstracción creciente, su aplicación a las demás ciencias se tornó cada vez más indispensable y eficaz. Esta aplicación se extendió, durante el siglo XVIII y el principio del XIX, de la mecánica y la astronomía a las restantes ramas de la física; más tarde a todas las ciencias naturales y, en el siglo XX, a todos los sectores del saber [7-9].

El optimismo que los primeros éxitos del cálculo infinitesimal infundió a la comunidad científica se vio matizado en posteriores aplicaciones por una desagradable evidencia: si bien todo problema podía plantearse en forma matemática por medio de ecuaciones diferenciales, la solución «exacta» de dichas ecuaciones sólo era posible para algunos casos particulares, que en ocasiones eran groseras simplificaciones de la realidad. Las dificultades que presentaba encontrar la fórmula matemática universal que diera la solución de problemas prácticos de la ciencia y la técnica hicieron patente la necesidad de encontrar formas alternativas de resolver las ecuaciones diferenciales.

Así, a principios del siglo XX, diversos científicos observaron que si las ecuaciones diferenciales para un problema concreto se discretizaban utilizando técnicas similares a las utilizadas por Arquímedes para calcular el valor del número π podían llegar a obtenerse los valores numéricos de la función incógnita. Habían nacido los denominados *métodos numéricos* [9-11].

La estrategia común de todos los métodos numéricos es la transformación de las ecuaciones diferenciales que gobiernan un problema en un sistema de ecuaciones algebraicas que dependen de un número finito de incógnitas. No obstante, puesto que este número de incógnitas es en la mayoría de los casos de muchos miles (e incluso millones), el sistema de ecuaciones final sólo puede resolverse con la ayuda del ordenador. Esto explica por qué, aunque muchos de los métodos numéricos eran conocidos desde el siglo XIX, su gran desarrollo y popularidad han sucedido paralelos al de los modernos ordenadores del siglo XX.

Los métodos numéricos representan, en definitiva, el retorno de los números como los auténticos protagonistas de la solución de un problema. El bucle iniciado por Pitágoras hace 2.500 años se ha cerrado, por tanto, en las últimas décadas, con la evidencia de que con ayuda de los métodos numéricos podemos obtener respuestas concretas a los enigmas del universo.

PLANTEAMIENTO GENERAL DE LA SOLUCIÓN NUMÉRICA DE UN PROBLEMA DE INGENIERÍA

Como se ha comentado en el apartado anterior, la mayor parte de los problemas de ingeniería se pueden expresar por medio de ecuaciones diferenciales en derivadas parciales con sus correspondientes condiciones de contorno. Dichas ecuaciones se obtienen usualmente por *reglas de balance* sobre un elemento diferencial en el interior del dominio de análisis y en sus contornos. Ejemplos concretos de esta técnica de balance son la ecuación diferencial de segundo orden que expresa el balance de flujo de calor en un dominio Ω (conocida en su forma genérica como ecuación de Poisson), las ecuaciones diferenciales de equilibrio interno y equilibrio en el contorno del dominio Ω en mecánica de sólidos, las ecuaciones de balance de cantidad de movimiento y de masa en dinámica de fluidos, etc.

De manera general, dichas ecuaciones diferenciales pueden expresarse como

$$\mathbf{A}\mathbf{u} - \mathbf{b} = 0 \quad \text{en } \Omega \quad (1a)$$

$$\mathbf{B}\mathbf{u} - \mathbf{q} = 0 \quad \text{en } \Gamma. \quad (1b)$$

En la expresión anterior \mathbf{A} y \mathbf{B} son operadores diferenciales que actúan sobre el vector de incógnitas \mathbf{u} ; \mathbf{b} y \mathbf{q} son vectores que contienen datos del problema. El operador \mathbf{A} se obtiene del balance de «flujos» sobre el dominio Ω , mientras que el operador \mathbf{B} resulta de las condiciones de balance en un dominio próximo al contorno Γ .

Como ejemplo de las ecs. (1) consideremos la ecuación de Poisson que gobierna la transmisión del calor por conducción en un dominio bidimensional:

$$\frac{\partial}{\partial x} \left(k_x \frac{\partial \phi}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_y \frac{\partial \phi}{\partial y} \right) + Q = 0 \quad \text{en } \Omega \quad (2)$$

$$k_x n_x \frac{\partial \phi}{\partial x} + k_y n_y \frac{\partial \phi}{\partial y} + \bar{q} = 0 \quad \text{en } \Gamma_q \quad (3)$$

$$\phi - \bar{\phi} = 0 \quad \text{en } \Gamma_\phi.$$

En las expresiones anteriores ϕ es la temperatura, k_x y k_y son las conductividades del material en las direcciones x y y , respectivamente, Q la fuente de calor sobre el dominio Ω , \bar{q} el calor que se pierde en dirección normal al contorno Γ_q , siendo n_x y n_y las direcciones del vector normal a dicho contorno y $\bar{\phi}$ la temperatura prescrita en el contorno Γ_ϕ . La unión de los contornos Γ_q y Γ_ϕ es el contorno total Γ del dominio Ω .

De las ecs. (2) y (3) se deducen los operadores y vectores de las ecs. (1) como

$$\mathbf{A} = \left[\frac{\partial}{\partial x} k_x \frac{\partial}{\partial x} + \frac{\partial}{\partial y} k_y \frac{\partial}{\partial y} \right] \quad (4)$$

$$\mathbf{B} = \begin{bmatrix} k_x n_x \frac{\partial}{\partial x} + k_y n_y \frac{\partial}{\partial y} \\ 1 \end{bmatrix} \quad (5)$$

$$\mathbf{u} = [\phi], \quad \mathbf{b} = [-Q] \quad \text{y} \quad \mathbf{q} = \left\{ \begin{array}{l} -\bar{q} \\ \bar{\phi} \end{array} \right\}. \quad (6)$$

En esencia, el papel de los métodos numéricos es resolver en forma de números las ecuaciones de gobierno para valores concretos de los parámetros físicos del problema (las conductividades k_x y k_y), de las fuentes exteriores (Q) y de las condiciones de contorno (el flujo de calor \bar{q} prescrito sobre el contorno Γ_q y el valor de la temperatura prescrita $\bar{\phi}$ sobre el contorno Γ_ϕ). En realidad, el técnico no se conforma con la mera solución de las ecs. (1), sino que además requiere que dicha solución se obtenga de la forma más *exacta, sencilla y económica* posible.

Esquema general de la solución numérica de un problema

La solución numérica de un problema gobernado por unas ecuaciones diferenciales con sus condiciones de contorno se basa en buscar una solución aproximada a las incógnitas del problema \mathbf{u} . Dicha solución aproximada se denomina $\hat{\mathbf{u}}$ y se expresa, en general, en términos de productos de funciones de aproximación conocidas $N_i(\mathbf{x})$ (generalmente de naturaleza polinómica) y coeficientes incógnitas \mathbf{a}_i . Así, puede escribirse

$$\mathbf{u} \cong \hat{\mathbf{u}} = \sum_{i=1}^N N_i(\mathbf{x}) \mathbf{a}_i. \quad (7)$$

La etapa anterior es conocida en el argot numérico como *discretización* de la solución del problema continuo. El problema se centra ahora en conocer los N vectores \mathbf{a}_i que son las incógnitas del problema.

El paso de las ecuaciones diferenciales de gobierno a un sistema de ecuaciones algebraico se efectúa usualmente a través del método de residuos ponderados (MRP) [12]. Este procedimiento puede entenderse fácilmente si se sustituye la aproximación (7) en las ecuaciones de gobierno (1).

Esto conduce a

$$\mathbf{A}\hat{\mathbf{u}} - \mathbf{b} = \mathbf{r}_\Omega \quad \text{en } \Omega \quad (8)$$

$$\mathbf{B}\hat{\mathbf{u}} - \mathbf{t} = \mathbf{r}_\Gamma \quad \text{en } \Gamma. \quad (9)$$

Los vectores \mathbf{r}_Ω y \mathbf{r}_Γ se denominan *residuos* de las ecuaciones de gobierno sobre el dominio y el contorno, respectivamente. El MRP se basa sencillamente en imponer que dichos residuos se anulan de forma ponderada sobre Ω y Γ . La expresión resultante del MRP es

$$\int_{\Omega} \mathbf{w}_i^T \mathbf{r}_{\Omega} d\Omega + \int_{\Gamma} \bar{\mathbf{w}}_i^T \mathbf{r}_{\Gamma} d\Gamma = 0. \quad (10)$$

En la ec. (10) \mathbf{w}_i y $\bar{\mathbf{w}}_i$ son vectores que contienen *funciones de peso* arbitrarias definidas sobre Ω y Γ , respectivamente. Sustituyendo en (10) las expresiones de \mathbf{r}_{Ω} y \mathbf{r}_{Γ} de (8) y (9) se obtiene

$$\int_{\Omega} \mathbf{w}_i^T [\mathbf{A}\hat{\mathbf{u}} - \mathbf{b}] d\Omega + \int_{\Gamma} \bar{\mathbf{w}}_i^T [\mathbf{B}\hat{\mathbf{u}} - \mathbf{t}] d\Gamma = 0. \quad (11)$$

Escogiendo ahora N vectores de funciones de peso \mathbf{w}_i y $\bar{\mathbf{w}}_i$ se obtiene un sistema algebraico de N ecuaciones con N incógnitas, que en problemas estacionarios se puede escribir de la forma

$$\mathbf{K} \mathbf{a} = \mathbf{f}. \quad (12)$$

En (12) \mathbf{K} es la denominada *matriz de rigidez* del sistema, \mathbf{a} es el vector de incógnitas que contiene los N vectores \mathbf{a}_i y \mathbf{f} es un vector que contiene términos conocidos provenientes de los vectores \mathbf{b} y \mathbf{t} en las ecuaciones de gobierno. La solución del sistema de ecuaciones algebraico (12) proporciona los N vectores \mathbf{a}_i . Sustituyendo éstos en la ec. (7) se obtienen los valores aproximados de las incógnitas \mathbf{u} en cada punto del dominio. Obviamente las expresiones de la matriz \mathbf{K} y del vector \mathbf{f} dependen de la selección de las funciones de peso. En el apartado siguiente se clasifican los métodos numéricos más comunes según el tipo de funciones de peso escogidas.

CLASIFICACIÓN DE LOS MÉTODOS NUMÉRICOS MÁS USUALES

Es muy interesante deducir los diferentes métodos numéricos que se utilizan en la práctica de la ingeniería a través de la selección de las funciones de peso \mathbf{w}_i y $\bar{\mathbf{w}}_i$. Seguidamente se efectúa una breve clasificación no exhaustiva.

Método de diferencias finitas

El método de diferencias finitas (MDF) se basa en escoger para \mathbf{w}_i y $\bar{\mathbf{w}}_i$ funciones *deltas de Dirac* en una serie de puntos seleccionados del dominio de análisis [12, 13] ordenados usualmente según una malla ortogonal. Esto equivale a *colocar* las ecuaciones diferenciales en dichos puntos. Las expresiones resultantes son

$$[\mathbf{A}(\hat{\mathbf{u}}) - \mathbf{b}]_i = 0 \quad i = 1, N_{\Omega} \quad (13a)$$

$$[\mathbf{B}(\hat{\mathbf{u}}) - \mathbf{t}]_j = 0 \quad j = 1, N_{\Gamma} \quad (13b)$$

donde N_{Ω} y N_{Γ} son el número de puntos de colocación en el dominio Ω y en el contorno Γ , respectivamente. Las derivadas que aparecen en los operadores \mathbf{A} y \mathbf{B} se evalúan generalmente por fórmulas en diferencias, en función de valores de las incógnitas en puntos adyacentes a cada punto donde se colocan las ecuaciones. Se destaca que la matriz \mathbf{K} resultante en el MDF no es, en general, simétrica.

Método de elementos finitos

En el método de elementos finitos (MEF) se divide el dominio de análisis en una malla (regular o irregular) formada por figuras geométricas sencillas (triángulos y cuadriláteros en 2D y tetraedros y hexaedros en 3D). La aproximación de las incógnitas (7) se efectúa ahora *en el interior* de cada elemento, de manera que los parámetros \mathbf{a}_i pasan a ser los valores de \mathbf{u} en puntos concretos del elemento (denominados nodos). En el MEF es usual escoger para las componentes de \mathbf{w}_i y $\bar{\mathbf{w}}_i$ las mismas funciones de aproximación N_i (método de Galerkin). Esto conduce en la mayoría de los casos, después de una adecuada integración por partes de algunos términos de \mathbf{A} , a una expresión simétrica de la matriz \mathbf{K} .

El carácter local de la aproximación del MEF trae también como consecuencia que la matriz \mathbf{K} y el vector \mathbf{f} de la ec. (12) pueden obtenerse fácilmente ensamblando las contribuciones individuales de los diferentes elementos, lo que simplifica notablemente el proceso de cálculo [12, 14].

En la Figura 3 se muestra de forma esquemática el proceso de análisis de la estructura de un puente por el MEF. A partir de la geometría inicial del puente se obtiene la malla de elementos finitos rectangulares que discretizan la losa superior y las vigas. Finalizado el proceso de cálculo, se obtiene una información completa del comportamiento del puente bajo las cargas actuantes, como, por ejemplo, el cambio de geometría bajo las cargas, las deformaciones y las tensiones en cada punto, etc. [14].

El proceso de «discretización» del puente es conceptualmente similar al utilizado por Arquímedes para calcular el perímetro de la circunferencia mediante su división en polígonos. Asimismo, los rectángulos de la Figura 3 juegan también un papel análogo a los utilizados por los calculistas chinos para evaluar el área del círculo en el siglo V d.C.

Desde el punto de vista del ingeniero de estructuras, el método de los elementos finitos puede considerarse como una extrapolación de los métodos de cálculo matricial para estructuras de barras al análisis de estructuras de tipo continuo [14]. De hecho, a principios de los años cuarenta surgen los primeros intentos de resolver problemas de elasticidad bidimensional con técnicas matriciales mediante la división del continuo en elementos de barra. Precisamente en 1943, R. Courant [15] introdujo por primera vez el concepto de «elemento continuo» al resolver problemas de elasticidad plana mediante la división del dominio de análisis en «elementos» triangulares sobre los que suponía una variación polinómica de la solución.

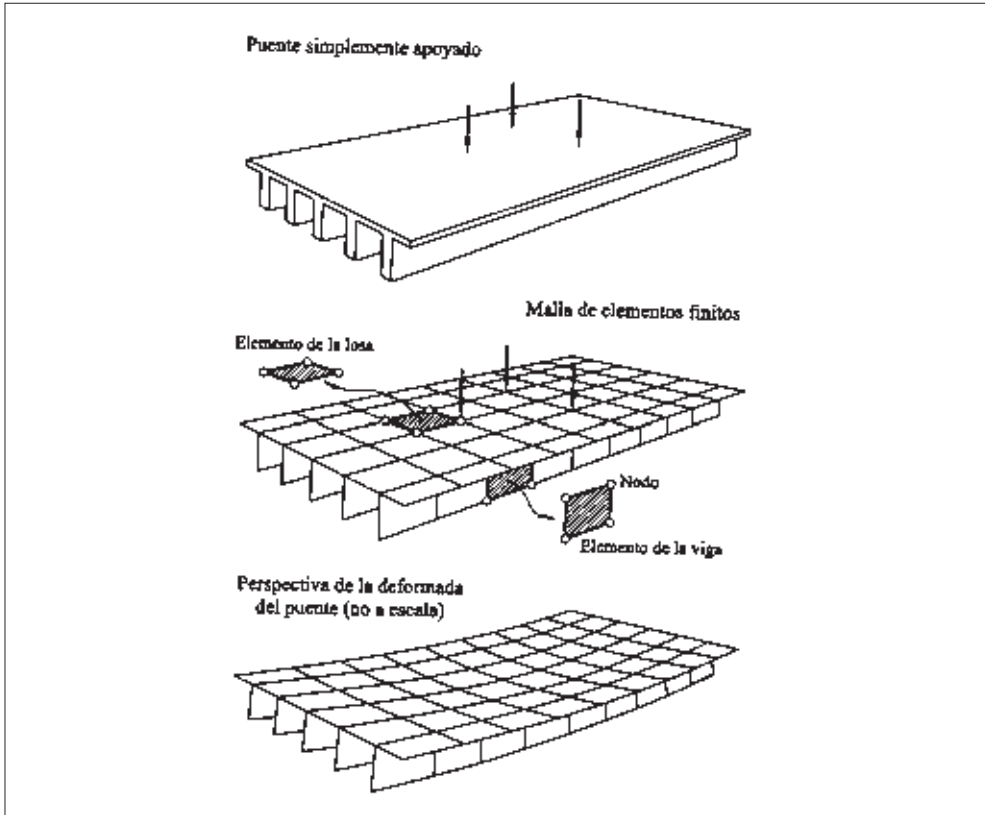


Figura 3. Análisis de la deformación de un puente por el método de elementos finitos.

La irrupción masiva de los ordenadores digitales en la década de los sesenta propició un avance espectacular de todos los métodos numéricos basados en técnicas matriciales, libres ya de las limitaciones que suponía hasta la fecha la solución de grandes sistemas de ecuaciones. Es en esta época cuando el método de los elementos finitos se consolida rápidamente como un procedimiento idóneo para la solución de toda una variedad de problemas de ingeniería y de física. Es importante advertir que, en este contexto, sus primeras aplicaciones están motivadas por problemas de cálculo de estructuras y, en particular, con aplicaciones estructurales en ingeniería aeronáutica [16, 17]. De hecho, fue R. Clough, de la Universidad de Berkeley, quien en 1960 y en relación con la solución de problemas de elasticidad plana sugirió por primera vez la denominación «elementos finitos» [18]. Desde esas fechas hasta la actualidad, el MEF ha tenido un desarrollo espectacular en su aplicación a otros campos. Así, apoyado por el avance de los ordenadores digitales, el MEF disfruta hoy en día de una posición única como una técnica de

solución potente de los problemas más diversos y complejos en innumerables campos de la ingeniería y la ciencia, desde el análisis estructural de edificios históricos hasta el cálculo de la resistencia de un hueso del cuerpo humano; desde el estudio de la aerodinámica de un avión hasta la evaluación del flujo de la sangre en las venas [12].

Método de elementos de contorno

En el método de elementos de contorno (MEC) se escogen las funciones de peso \mathbf{w}_i de manera que se anulen sobre el dominio Ω . La expresión integral resultante queda definida solamente sobre el contorno como

$$\int_{\Gamma} \bar{\mathbf{w}}_i [\mathbf{B}\hat{\mathbf{u}} - \mathbf{t}] d\Gamma = 0. \quad (14)$$

El MEC exige únicamente, por tanto, aproximar las incógnitas sobre los contornos del dominio de análisis, lo que conduce a una importante disminución del tamaño del sistema de ecuaciones resultante [19]. Lamentablemente la matriz de rigidez \mathbf{K} de dicho sistema no es simétrica, tal como sucede en el MDF.

El MEC permite, no obstante, un ahorro de cálculo substancial en la solución de problemas en los cuales la obtención de las ecuaciones integrales sobre el contorno es posible y sencilla, tales como problemas con propiedades lineales gobernados por la ecuación de Laplace (transmisión del calor, flujo en medio poroso, electromagnetismo, análisis elástico de sólidos, etc.). Pese a sus innegables ventajas, en algunos casos, el MEC no es todavía tan utilizado como el MEF para la solución de problemas industriales.

Métodos de puntos

No podemos concluir esta breve reseña sobre métodos numéricos sin mencionar el reciente auge de los métodos basados en discretizaciones utilizando únicamente un conjunto de puntos. Estos métodos, adjetivados comúnmente *sin malla* [20], de *partículas* o métodos de *puntos finitos* (MPF) [21], tienen la ventaja de no necesitar la usualmente costosa construcción de una malla sobre el dominio de análisis; basta con «rellenar» su interior con un gran número de puntos a los que se asocian los valores de las incógnitas del problema. El conjunto de puntos próximo a un punto concreto se denomina «nube». La variación de cada incógnita en el interior de una nube se expresa en función de las variables en cada punto de la misma utilizando técnicas de mínimos cuadrados ponderados. El paso final es imponer el cumplimiento de las ecuaciones diferenciales de gobierno del problema sobre cada nube (en forma integral), o bien directamente «colocando» las ecuaciones en cada uno de los puntos que discretizan el

dominio, de forma similar al MDF. En ambos casos se llega al sistema de ecuaciones algebraico buscado, cuya solución conduce a los valores numéricos de las incógnitas en cada punto [21].

TIPOLOGÍA DE LAS ECUACIONES EN PROBLEMAS DE INGENIERÍA

De forma esquemática, las ecuaciones que intervienen en los problemas de ingeniería más usuales pueden clasificarse en los tres grandes grupos siguientes.

Problemas de estática

Las ecuaciones tienen la forma de las expresiones (1), y en ellas todas las variables y parámetros son independientes del tiempo. El sistema de ecuaciones algebraicas después del proceso de discretización tiene la forma matricial expresada por la ec. (12).

Los ejemplos de aplicaciones en ingeniería de los problemas de estática son muy numerosos en el cálculo de estructuras y sistemas mecánicos, en la transmisión del calor en régimen estacionario, en problemas de filtración y electromagnetismo en régimen estacionario, en problemas estacionarios de dinámica de fluidos, etc.

Problemas dinámicos de primer orden

La forma general de las ecuaciones de gobierno en este tipo de problemas puede expresarse como

$$\mathbf{A}(\mathbf{u}, \dot{\mathbf{u}}) - \mathbf{b}(t) = 0 \quad \text{en } \Omega \quad (15a)$$

$$\mathbf{B}(\mathbf{u}, \dot{\mathbf{u}}) - \mathbf{q}(t) = 0 \quad \text{en } \Gamma \quad (15b)$$

donde $\dot{\mathbf{u}} = \frac{d}{dt} \mathbf{u}$, siendo t el tiempo. En este caso $\mathbf{u} \equiv \mathbf{u}(\mathbf{x}_i, t)$.

Un ejemplo de las expresiones anteriores es la ecuación de propagación del calor por conducción en régimen transitorio. En una dimensión,

$$\frac{\partial \phi}{\partial t} + \frac{\partial}{\partial x} \left(k \frac{\partial \phi}{\partial x} \right) + Q = 0 \quad \text{en } \Omega \quad (16a)$$

$$k \frac{\partial \phi}{\partial x} - \bar{q} = 0 \quad \text{en } \Gamma_q \quad (16b)$$

$$\phi - \bar{\phi} = 0 \quad \text{en} \quad \Gamma_{\phi} \tag{16b}$$

$$\phi(x, 0) = 0 \quad \text{para} \quad t = 0.$$

En (16a) ρ y c son la densidad y el calor específico, respectivamente. La forma del sistema de ecuaciones resultante después del proceso de discretización sobre el dominio espacial Ω se puede expresar como

$$\mathbf{Ca}' + \mathbf{Ka} = \mathbf{f}. \tag{17}$$

La ec. (17) es un sistema de ecuaciones algebraicas donde intervienen los parámetros incógnitas del problema \mathbf{a} y sus derivadas con respecto al tiempo \mathbf{a}' . La siguiente etapa en la solución numérica del problema es la integración temporal de la ec. (17). Ello puede hacerse utilizando diferentes técnicas para solución numérica de ecuaciones diferenciales parabólicas.

El procedimiento más popular para integrar en el tiempo la ec. (17) es aplicar el método de diferencias finitas. En cualquier caso, el procedimiento de integración temporal proporciona los valores de los parámetros incógnitas en el tiempo $n+1$ en función de valores de dichos parámetros conocidos en el tiempo n . En forma general

$$\mathbf{a}^{n+1} = \mathbf{g}(\mathbf{a}^n). \tag{18}$$

Existen diferentes técnicas para obtener el vector $\mathbf{g}(\mathbf{a}^n)$, bien resolviendo un sistema de ecuaciones algebraico en cada paso de tiempo (métodos implícitos), o de forma directa sin necesidad de resolver un sistema de ecuaciones (métodos explícitos) [12].

Las aplicaciones en ingeniería de esta clase de ecuaciones diferenciales son también muy numerosas. Entre los problemas más característicos podemos citar todos los de dinámica de fluidos en régimen transitorio; problemas no estacionarios de transmisión del calor, de filtración en medios porosos y de electromagnetismo; problemas de consolidación de suelos, etc.

Problemas dinámicos de segundo orden

Las ecuaciones de gobierno se escriben por

$$\mathbf{A}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}) - \mathbf{b}(t) = 0 \quad \text{en} \quad \Omega \tag{19a}$$

$$\mathbf{B}(\mathbf{u}, \dot{\mathbf{u}}, \ddot{\mathbf{u}}) - \mathbf{q}(t) = 0 \quad \text{en} \quad \Gamma. \tag{19b}$$

Un ejemplo típico es la ecuación de ondas. En una dimensión,

$$\rho \frac{\partial^2 \phi}{\partial t^2} + c \frac{\partial \phi}{\partial t} + \frac{\partial}{\partial x} \left(k \frac{\partial \phi}{\partial t} \right) + Q = 0 \quad \text{en} \quad \Omega \tag{20}$$

$$\begin{aligned}
 k \frac{\partial \phi}{\partial x} + \bar{q} &= 0 \quad \text{en} \quad \Gamma_q \\
 \phi - \bar{\phi} &= 0 \quad \text{en} \quad \Gamma_\phi \\
 \phi(x, 0) = \dot{\phi}(x, 0) &= 0 \quad \text{para} \quad t = 0.
 \end{aligned}
 \tag{21}$$

El proceso de discretización temporal conduce a un sistema algebraico de ecuaciones de la forma

$$\mathbf{M}\ddot{\mathbf{a}} + \mathbf{C}\dot{\mathbf{a}} + \mathbf{K}\mathbf{a} = \mathbf{f}.
 \tag{22}$$

La solución numérica en el tiempo por un algoritmo de diferencias finitas proporciona los valores de \mathbf{a} , $\dot{\mathbf{a}}$ y $\ddot{\mathbf{a}}$ en el tiempo $n + 1$ en función de sus valores en tiempos anteriores.

Las aplicaciones prácticas de las ecuaciones de este tipo son muy usuales en análisis dinámico de sólidos y estructuras y en problemas de propagación de ondas (acústica, electromagnetismo, oleaje, etc.).

MÉTODOS NUMÉRICOS Y REALIDAD

Está claro que el objetivo de los métodos numéricos es reproducir lo más fielmente posible el comportamiento del mundo a través de números. Recordemos, no obstante, que el primer paso en ese proceso es establecer un *modelo matemático* de la realidad en cuestión. Los métodos numéricos permiten resolver en forma numérica con ayuda del ordenador las ecuaciones matemáticas de dicho modelo, generalmente expresadas por medio de ecuaciones o inecuaciones en derivadas parciales o algebraicas. Los números que resultan del cálculo, expresados por medio de gráficos dibujados por ordenador, representan la visión de la realidad que proporciona el *proceso de cálculo* escogido. Obviamente, la solución numérica sólo coincidirá con la *realidad* si: *a)* el modelo matemático incorpora todos los aspectos del mundo real; y *b)* el método numérico puede resolver exactamente las ecuaciones del modelo matemático.

En la práctica, ninguna de estas dos condiciones se cumple y hay que admitir que la predicción numérica no coincidirá con el comportamiento del mundo real. Se dice entonces que la solución numérica *aproxima* la realidad. Si conocemos la solución «real» del problema que se estudia, podemos compararla con la solución numérica y obtener el error de la predicción. En la práctica, sucede también que generalmente dicha solución real no existe, ya que todo modelo matemático expresa una idealización simplificada de la realidad. Así, pues, en el mejor de los casos es posible obtener soluciones

«exactas» de algunos modelos matemáticos que son aproximaciones de la realidad. Estas soluciones «exactas» (que denominábamos soluciones analíticas al inicio del artículo) sí que pueden compararse con las obtenidas resolviendo por métodos numéricos los mismos modelos matemáticos. Desgraciadamente, las soluciones exactas son también prácticamente imposibles de obtener para la mayoría de los modelos matemáticos que resuelven problemas de interés. Los pocos casos (generalmente de tipo académico) en que esta comparación es posible sirven para calibrar el método numérico. En el resto de situaciones, la única comparación factible se consigue con resultados de pruebas experimentales obtenidos para problemas concretos, en los cuales es posible realizar mediciones. Naturalmente, la validación experimental es útil para calibrar tanto el método numérico utilizado como el modelo matemático subyacente.

En resumen, nuestra visión de la realidad será siempre aproximada, tanto por las limitaciones de formalizar dicha realidad mediante un modelo matemático como por los errores inherentes en la aplicación de métodos numéricos a dicho modelo. Nuestra única referencia posible son las validaciones empíricas de los resultados numéricos, utilizando valores experimentales obtenidos en pruebas de laboratorio o de campo muy concretas. En la mayoría de los casos estamos solos frente al conjunto de números que resultan de la predicción de un problema cuya solución «real» es desconocida. Es en ese momento cuando toda la experiencia acumulada en la calibración y validación del modelo matemático y el método numérico escogidos debe utilizarse para aceptar o no los números que proporciona el cálculo.

Quizá ningún ejemplo como el de la estimación del número π ejemplifica las limitaciones de los métodos numéricos. Como es sabido, el número π es uno de los inconmensurables, es decir, tiene un número infinito de cifras y cualquier proceso de cálculo conduce a una solución numérica aproximada del valor de π . Tomemos, por ejemplo, el método utilizado por Arquímedes, a base de dividir la circunferencia en polígonos inscritos y circunscritos de un número de lados creciente. Refinando progresivamente dicho proceso de cálculo, es decir, tomando más y más polígonos inscritos y circunscritos a la circunferencia, se puede disminuir el error de la aproximación y acotar el valor de π entre dos números cada vez más próximos. En cualquier caso, la solución numérica será siempre una *estimación* del valor real de π , aunque deberemos aceptarla como útil para cualquier cálculo posterior en el que se necesite su valor (Figura 1).

Para poner de manifiesto las relaciones entre realidad, métodos numéricos, modelos matemáticos e informática, tendríamos que dar respuesta a muchas preguntas: ¿Qué hace veraz una solución numérica? ¿Por qué he de creer los valores numéricos obtenidos a través del ordenador? ¿Qué confiere a una solución numérica su utilidad? ¿Qué hace que sea buena o mala? ¿Qué la hace bella o fea? ¿Cuál ha sido la influencia de los métodos numéricos sobre las teorías del conocimiento y la existencia matemática, sobre la intuición matemática, sobre la educación matemática? ¿Qué relación existe entre el pen-

samiento matemático, los métodos numéricos y las capacidades potenciales del ordenador y la mente humana? ¿De qué modo contribuye la solución numérica de un problema a cambiar nuestra idea de la realidad, del conocimiento, del tiempo?

Una vez dada respuesta a las preguntas anteriores tendríamos bien andado el camino conducente a la creación de una filosofía de los métodos numéricos (o del cálculo). Así, al igual que la filosofía clásica se ha ocupado de lo verdadero, lo bueno y lo bello, así también la filosofía de los métodos numéricos podría ocuparse de la veracidad de los cálculos, de la bondad y belleza en la solución numérica de los problemas del universo [8]. Sólo a través de una reflexión profunda sobre estas ideas podremos aportar luz sobre una sociedad cada vez más dividida entre los que opinan que la informática y sus actividades subsiguientes son un mal irremediable, que degrada el espíritu y corrompe la inteligencia, y los que, en el polo opuesto, creen que los números que proporciona el ordenador, generados a través de modelos matemáticos de la realidad y métodos numéricos, nos ayudan a entender mejor el mundo que nos rodea y son un ingrediente más para alcanzar la justicia social y la pacificación del mundo.

APLICACIONES

Se presentan en este apartado diversas aplicaciones del método de elementos finitos a problemas de ingeniería.

La primera aplicación se refiere al análisis de estructuras históricas. En la Figura 4.1 se muestra la discretización en elementos hexaédricos de las cinco cúpulas de la



Figura 4.1. Discretización de las cúpulas de la Basílica de San Marcos con elementos finitos hexaédricos. Deformada de la estructura bajo su propio peso.

Basílica de San Marcos en Venecia. El análisis estructural no lineal consistió en calcular el coeficiente de seguridad a rotura de la estructura obtenido al aumentar progresivamente el peso hasta obtener la rotura de las cúpulas. La Figura 4.2 muestra los diferentes niveles de deterioro en la estructura en el momento de la ruina. La intensidad del deterioro se caracteriza por la gama de colores.

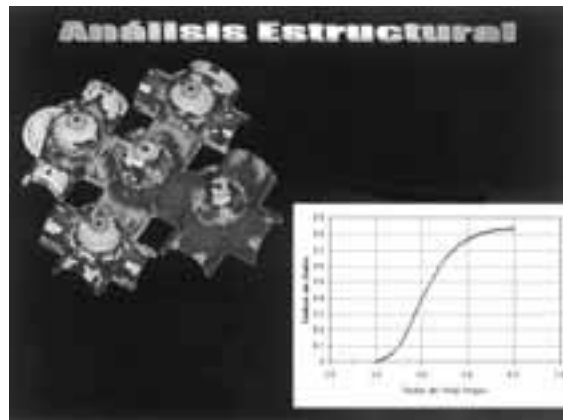


Figura 4.2. Basílica de San Marcos. Niveles de deterioro en la estructura para una carga equivalente a 7,11 veces el peso propio.



Figura 5. Análisis estructural de la Catedral de Barcelona.

En la Figura 5 se muestra un ejemplo de un estudio similar realizado sobre la Catedral de Barcelona. Para más detalle sobre análisis por el MEF de estructuras históricas consultar las referencias [22-24].

El siguiente ejemplo muestra diferentes resultados del estudio aerodinámico realizado sobre el nuevo edificio del Gran Telescopio de Canarias, situado en la isla de La Palma. El contorno del edificio que alberga el telescopio (Figura 6) se discretiza en elementos finitos. Tras ello se discretiza el volumen de aire en el interior y el exterior del edificio en tetraedros y se resuelven las ecuaciones de Navier-Stokes que gobiernan el flujo del aire en el interior de un dominio de control que incluye el edificio y su entorno. Como resultados se obtienen los campos de velocidades, presiones y temperaturas en todos los puntos del dominio. Estos resultados se utilizan para diseñar el sistema de ventilación del edificio, a través del estudio de la trayectoria de las partículas de aire, para evaluar las fuerzas sobre los diferentes instrumentos, para obtener la temperatura en las proximidades de los espejos del telescopio y para estimar la posible perturbación de la imagen por efecto de las turbulencias en el fluido [25].

En el siguiente ejemplo se presenta el análisis hidrodinámico y aerodinámico de un velero. En el estudio por el MEF se analiza el flujo de aire alrededor de las velas y del casco y la hidrodinámica de la parte sumergida de este último. En la Figura 7 se muestran diferentes resultados de este estudio [26].

El siguiente ejemplo es el de análisis del proceso de fabricación de un cigüeñal de un coche por fundición. El estudio implica simular con detalle el proceso de llenado del molde por el metal fundido y seguidamente analizar el proceso de solidificación y enfriamiento del metal, desde la temperatura de 1200° del acero fundido hasta la temperatura ambiente. Como resultado del estudio se obtiene también la distribución de deformaciones y tensiones en la pieza durante el proceso de solidificación y enfriamiento. La Figura 8 muestra los mapas de temperaturas en el cigüeñal en distintos instantes del proceso de enfriamiento. Para más información sobre este problema y otras aplicaciones del MEF a procesos de conformado de metales consultar las referencias [27, 28].

El último ejemplo es el estudio de la deformación de un automóvil en un choque contra una pared rígida. El análisis se realizó discretizando la estructura del coche por el MEF y resolviendo las ecuaciones de la dinámica de la estructura, teniendo en cuenta los efectos del contacto entre el vehículo y la pared durante el choque. En la Figura 9 se muestra la malla de elementos finitos utilizada y una vista de la deformada del vehículo después de 24 milisegundos de iniciarse el impacto. Se destaca que los resultados se han obtenido en un ordenador PC dual calculando en paralelo en los dos procesadores, mediante una partición de la estructura en dos dominios tal y como se muestra en la figura.

Los ejemplos anteriores son un botón de muestra de las posibilidades del MEF en su aplicación al mundo industrial. Para más detalles sobre ésta y otras aplicaciones puede consultarse la página web del Centro Internacional de Métodos Numéricos en Ingeniería en www.cimne.upc.es.



Figura 6. Análisis aerodinámico por el MEF del edificio del Gran Telescopio de Canarias. Discretización del edificio y el terreno circundante y diferentes resultados de la distribución de presiones y velocidades y de las trayectorias de las partículas de aire.

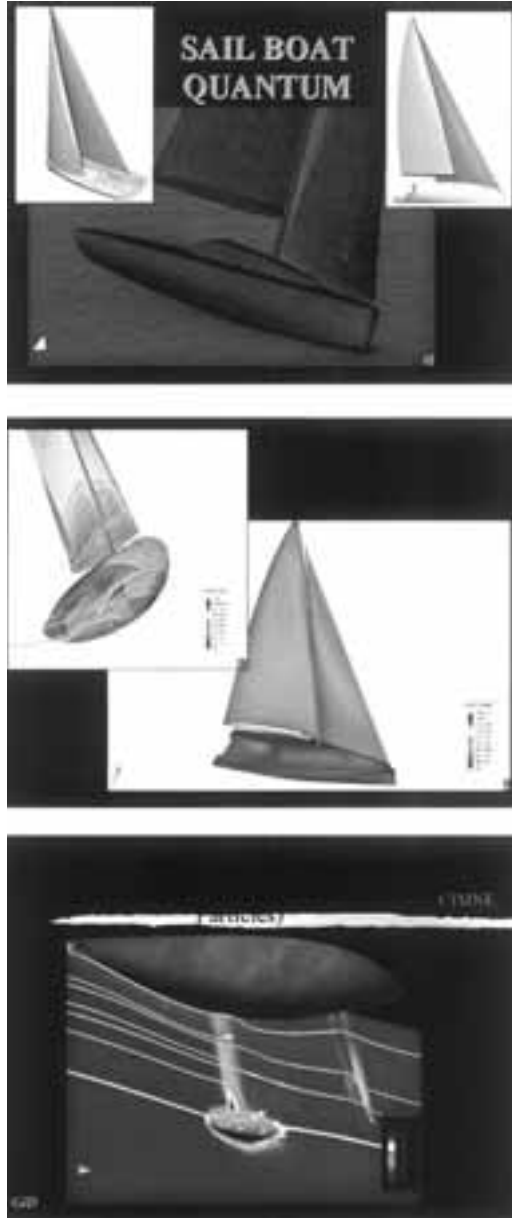


Figura 7. Análisis por el MEF de la aerodinámica e hidrodinámica de un velero de competición. Trayectorias de las partículas del aire y distribución de presiones del agua sobre el casco y apéndices.

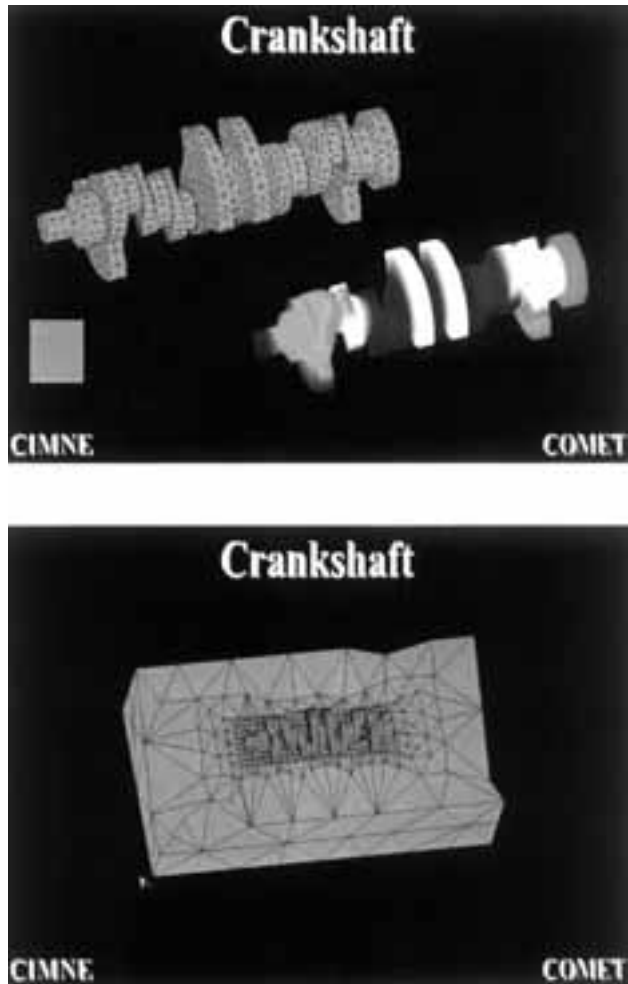


Figura 8. Análisis por el MEF del proceso de solidificación y enfriamiento de un cigüeñal fabricado por fundición.



Figura 9. Malla de elementos finitos, partición de la estructura en dos dominios para cálculo en paralelo y deformada del vehículo después de 24 milisegundos de iniciarse el impacto.

LÍMITES Y PERSPECTIVAS DE LOS MÉTODOS NUMÉRICOS

¿Es posible describir con la ayuda de las matemáticas y los métodos numéricos todos los aspectos del mundo? En realidad, los límites de los métodos numéricos, es decir, los límites de poder expresar en forma numérica cualquier problema del universo, están íntimamente ligados a la posibilidad de plantear (formalizar) dicho problema en forma matemática. Ante esta situación, surge inmediatamente la pregunta: ¿Es posible matematizarlo todo? ¿Habrá algo en el mundo que no pueda jamás llegar a ser descrito en lenguaje matemático? Como afirman P. J. Davies y R. Hersh [8], no parece que haya en el mundo físico nada no matematizable. Tenemos confianza, por tanto, en que la física es capaz de abarcar cualquier fenómeno físico y que es capaz de hacerlo por medio de un formalismo matemático.

El límite de lo claramente matematizable parece, por lo tanto, que coincide con el límite del mundo físico. ¿Qué otro mundo hay? Si nuestras convicciones son las de un materialista mecanicista puro, probablemente contestaremos que ninguno.

No obstante, es evidente que existen cosas tales como emociones, creencias, actitudes, sueños, intenciones, celos, envidias, ansias, pesares, sentimientos como la ira y la compasión y muchos otros. Estas cosas, que componen el mundo interior de la mente humana, y, más aún, todas aquéllas que abarcan la «vida interior» de la sociedad, de la civilización misma, por ejemplo, la literatura, la música, la política o las mareas y corrientes de la historia, ¿podrán ser matematizadas?

La respuesta no es sencilla, ni banal. A lo largo de la historia encontramos defensores de la creencia de que todo es matematizable, y por tanto numeralizable, y de lo contrario.

Tomando partido, podríamos afirmar que pueden describirse con modelos matemáticos todas las facetas del mundo que puedan encuadrarse mediante enunciados concretos. Naturalmente, si dicho enunciado es esencialmente de tipo humanista, la dificultad de establecer el modelo será mucho mayor. Tomemos, por ejemplo, el sentimiento de bienestar. En este caso, un experimentalista podría afirmar que el bienestar es mera función de los niveles de hormonas y azúcar en la sangre, con lo que se entrevé de nuevo la posibilidad de encontrar un modelo matemático/numérico para el bienestar. Un viejo conflicto entre científicos y humanistas proviene de que estos últimos sienten que debería existir una porción del mundo inmune a la matematización, mientras que los científicos son del sentir contrario, es decir, todo aspecto del mundo admite una descripción matemática y por consiguiente numérica. De nuevo, reaparece la vieja aspiración de Pitágoras y Platón, que va ganando adeptos a medida que se producen nuevos éxitos en la matematización del mundo.

En nuestros días, las bases de las ciencias de la naturaleza, la física, la astrofísica, la química y también todas las ingenierías son profundamente matemáticas. Los mode-

los matemáticos y numéricos son, en todas estas disciplinas, herramientas de uso cotidiano, tanto para la explicación de fenómenos del universo como para el diseño de productos y procesos. Más aún, para que una nueva teoría sea tenida como científica, y por tanto aceptada, es casi condición necesaria que sea expresable en lenguaje matemático, y es casi seguro que, si las matemáticas disponibles son inadecuadas para describir ciertos fenómenos observados, será posible idear y desarrollar las adecuadas.

Las ciencias de la vida, la biología y la medicina, están adquiriendo un carácter cada vez más matemático/numérico. Los mecanismos que controlan los procesos fisiológicos, la genética, la morfología, la dinámica de poblaciones, la epidemiología y la ecología han sido provistos de modelos matemáticos y numéricos.

No es posible comprender la teoría económica sin una sólida formación matemática y numérica. Las teorías de la competencia, de los ciclos y equilibrios comerciales exigen matemáticas y técnicas de cuantificación de tipo más profundo. La determinación de una política comercial o militar puede requerir teorías de decisión, teorías de juegos y estrategias de optimización.

Con independencia del problema que se resuelva, es importante recordar que el fin último de los métodos numéricos es proporcionar comprensión, no números. En suma, los métodos numéricos deben estar íntimamente ligados tanto a la fuente del problema como al uso que se pretenda hacer de los resultados numéricos; la aplicación de dichos métodos no es, por tanto, una etapa a considerar aisladamente de la realidad [8].

La palabra que quizá puede sintetizar el futuro más inmediato de las aplicaciones de los métodos numéricos es «multifísica». Los problemas no se abordarán más desde la perspectiva de un único medio físico e incorporarán todos los acoplamientos que caracterizan la complejidad de la realidad. Así, por ejemplo, el diseño de una pieza de un vehículo (un avión, un automóvil, etc.) se realizará teniendo en cuenta el proceso de fabricación y la función que dicha pieza ejercerá a lo largo de su «vida útil». Las estructuras en ingeniería civil se estudiarán teniendo en cuenta los efectos con el medio circundante (el terreno, el agua, el aire). Ejemplos similares pueden encontrarse en ingeniería naval y aeronáutica, entre otras, así como en bio-ingeniería, en economía, en demografía, en la modelización del comportamiento de personas y grupos y en prácticamente todas las áreas de la ciencia. La importancia de tener en cuenta el carácter no determinista de todos los datos será esencial para estimar la probabilidad de que los nuevos modelos matemáticos concebidos por el hombre se comporten de la forma prevista.

Sólo desde la perspectiva de una estrecha cooperación entre todas las partes del triángulo formado por el conocimiento profundo de las bases físicas y matemáticas de cada problema, los métodos numéricos y la informática podrán encontrarse soluciones efectivas a los megaproblemas del nuevo siglo. Esa cooperación deberá verse reflejada también en un mayor énfasis en la optimización de los recursos materiales y humanos

necesarios para afrontar con garantías el cambio de escala de los problemas a resolver y, sobre todo, en la puesta en marcha de acciones de formación innovadoras para preparar a las nuevas generaciones, que, con la ayuda de las matemáticas y los números, deberán abordar con éxito la solución de problemas multidisciplinares.

REFERENCIAS

- [1] P. Beckmann, *A History of π* , The Golem Press, 1971.
- [2] T. Kiang, «An old Chinese way of finding the volume of a sphere», *Math. Gazette*, LVI, 88-91, 1972.
- [3] M. E. Baron, *The Origins of Infinitesimal Calculus*, Dover Publications, 1987.
- [4] J. Rey Pastor y J. Babini, *Historia de la matemática* (dos volúmenes), Editorial Gedisa, 3ª ed., 1997.
- [5] C. B. Boyer, *Historia de la matemática*, Alianza Editorial, 1999.
- [6] E. T. Bell, *The Development of Mathematics*, Dover Publications, 1992.
- [7] G. Frey, *La matematización de nuestro universo*, G. del Toro Editor, 1972.
- [8] P. J. Davis y R. Hersch, *El sueño de Descartes*, Editorial Labor, 1989.
- [9] E. Oñate, *El aura de los números*, Reial Acadèmia de Doctors, Barcelona, 1998.
- [10] E. Oñate, «El bucle de los números», Publicación de Investigación CIMNE Núm. 192, 2000.
- [11] E. Oñate, «Límites de los métodos numéricos», Publicación de Investigación CIMNE Núm. 191, 2000.
- [12] O. C. Zienkiewicz y R. L. Taylor, *El método de los elementos finitos*, McGraw Hill-CIMNE, Barcelona, Vol. I-1993, Vol. II-1994.
- [13] R. V. Southwell, *Relaxation Methods in Theoretical Physics*, Clarendon Press, 1946.
- [14] E. Oñate, «Cálculo de estructuras por el método de los elementos finitos», CIMNE, Barcelona, 1992.
- [15] R. Courant, «Variational methods for the solution of problems of equilibrium and vibration», *Bull. Amer. Math. Soc.*, Vol. 49, pp. 1-23, 1943.
- [16] M. J. Turner, R. W. Clough, H. Martin y L. J. Topp, «Stiffness and deflection analysis of complex structures», *J. Aeron. Sci.*, Vol. 23, pp. 805-823, 1956.
- [17] J. J. Argyris y S. Kelsey, *Energy Theorems and Structural Analysis*, Butterworth, 1960.
- [18] R. W. Clough, «The finite element method in plane stress analysis», *Proc. 2nd A.S.C.E. Conf. in Electronic Computation*, Pittsburgh, Sept. 1960.
- [19] C. A. Brebbia y J. Domínguez, «Boundary elements. An introductory course», *Computational Mechanics Publ.*, McGraw Hill, 1989.

- [20] C. A. M. Duarte, «A review of some meshless methods to solve partial differential equations», *TICAM Report 95-06*, Univ. of Texas, Austin, May 1995.
- [21] E. Oñate, S. Idelsohn, O. C. Zienkiewicz y R. L. Taylor, «A finite point method in computational mechanics. Applications to convective transport and fluid flow», *Int. J. Num. Methods Eng.*, Vol. 39, pp. 3839-3866, 1996.
- [22] E. Oñate, A. Hanganu, A. Barbat, S. Oller, R. Vitaliani, A. Saetta y R. Scotta, «Structural analysis and durability assessment of historical construction using a finite element damage model», *Structural Analysis of Historical Construction I*, P. Roca *et al.* (Eds.), CIMNE, Barcelona, 1997.
- [23] P. Roca, L. Pellegrini, E. Oñate y A. Hanganu, «Analysis of the structure of Gothic cathedrals. Applications to Barcelona Cathedral», *Structural Analysis of Historical Construction II*, P. Roca *et al.* (Eds.), CIMNE, Barcelona, 1998.
- [24] P. Roca *et al.* (Eds.), *Structural Analysis of Historical Construction*, CIMNE, Barcelona, Vol.1 (1997), Vol. 2 (1998).
- [25] R. Codina, C. Morton, E. Oñate y O. Soto, «Numerical aerodynamic analysis of large buildings using a finite element model with application to a telescope building», *Int. Journal of Numerical Methods for Heat and Fluid Flow*, Vol. 10, Núm. 6, 616-633, 2000.
- [26] Integrated Design and Production of Advanced Sails (ID-PASS), Proyecto CRAFT de la Comisión Europea, Informe Final, CIMNE, Barcelona, 2001.
- [27] M. Chiumenti, C. Agelet de Saracibar y M. Cervera, «Constitutive modelling and numerical analysis of thermo-mechanical phase-change systems», Monografía CIMNE Núm. 48, 1999.
- [28] E. Oñate, «Perspectivas de modelos constitutivos y técnicas de elementos finitos para análisis de procesos de conformado de metales», Publicación CIMNE Núm. 94, 1996.
- [29] X. Plana, «Análisis por el método de elementos finitos del impacto de un vehículo con el programa SIMPACT», Proyecto de Fin de Carrera, Escuela Técnica Superior de Ingenieros Industriales de Barcelona, 1999.

Modelización, análisis y control de sistemas climáticos

Ildefonso Díaz

Universidad Complutense de Madrid

1. MODELIZACIÓN, ANÁLISIS Y CONTROL: UNA TRILOGÍA UNIVERSAL

En 1986, el Consejo Internacional de Uniones Científicas decidió poner en marcha el IGBP: Programa Internacional Geosfera-Biosfera, llamado comúnmente *Cambio Global* (*Global Change* en inglés). En la descripción de sus objetivos principales se indicaba:

«... describir y comprender los procesos interactivos de origen físico, químico o biológico que regulan el sistema de la Tierra, los cambios que se producen en el sistema y la forma en que los factores humanos influyen sobre estos cambios».

El programa inicial ha ido tomando cuerpo desde entonces, completando decisiones anteriores tomadas en el marco de la Organización Meteorológica Mundial de las Naciones Unidas y de la UNESCO. Pero, ¿hasta qué punto es posible desarrollar este programa?

Los progresos en el conocimiento de nuestro planeta han dejado patente, de modo cada vez más incuestionable, cómo las acciones humanas pueden llegar a modificar el clima a largo plazo e incluso las condiciones meteorológicas en un plazo más corto de tiempo. Las respuestas a las necesidades energéticas de la sociedad tienen, en nuestros días, múltiples lecturas y dejarán de ser evidentes en siglos venideros. Cada vez hay una mayor atención a las causas e intentos de mitigación de las catástrofes llamadas «naturales»: lluvias torrenciales, tornados, tifones, etc. Lo mismo sucede con los peligros para la salud derivados de la polución en nuestras ciudades.

El estudio del conjunto de todas esas cuestiones desborda claramente el ámbito propio de cada una de las disciplinas científicas *clásicas* consideradas aisladamente. Sin embargo, el estudio de esa problemática utiliza, cada vez en mayor grado, una «herramienta», que podríamos catalogar de universal en atención a su gran versatilidad para su aplicación, constituida por las Matemáticas, los superordenadores y el control.

De manera general, la coordinación entre las Matemáticas y los superordenadores es imprescindible. Hoy día no cabe la menor duda que una Matemática que prescindiera de la ayuda de los máximos recursos de cálculo de su tiempo no podrá ofrecer las respuestas cuantitativas que demandan sistemas tan complejos como el del sistema climático.

Siendo esto obvio, no lo es menos que esas capacidades de computación y la enorme cantidad de datos hoy disponibles serían absolutamente inútiles sin la ayuda de modelos matemáticos. La simbiosis se asemeja a la típica entre *software* y *hardware*. Pero, si esto es así, ¿qué contribuciones puede aportar esa simbiosis al estudio global del sistema climático?

La capacidad de memoria de los superordenadores permite almacenar y analizar cantidades crecientes de datos resultados de medidas, operaciones que sin los ordenadores serían imposibles de realizar. Una vez accesibles estos datos y gracias al progreso en el conocimiento fundamental de los fenómenos involucrados, es posible la modelización matemática y, en una fase ulterior, la necesaria validación de los modelos.

El estudio de sistemas complejos es posible, en nuestros días, gracias a una metodología global que podríamos denominar *trilogía universal: la modelización matemática, el análisis y la simulación* mediante superordenadores, y las acciones correctivas, destinadas a «mejorar la situación», en resumen, *el control* (Díaz y Lions [15]).

Es imprescindible comenzar por la modelización de los elementos más simples, los «bloques básicos». Después, conviene estudiar las distintas correlaciones, las reacciones de componer unos elementos sobre otros, lo que llamaremos aquí *interacciones* (en inglés *feedbacks*: terminología que resulta algo ambigua cuando nos interesamos por el control del sistema). Se trata de interacciones que conducen a la articulación de esos bloques básicos y con ello a la construcción de grandes modelos y sofisticados códigos informáticos.

Además, el progreso de los conocimientos científicos y la *abundancia extraordinaria de datos*, en particular de los datos obtenidos por satélites espaciales, permiten ir más allá en la elaboración de los modelos. Cada vez son más numerosos los fenómenos que pueden ser tenidos en consideración, en especial los que conciernen a la *biosfera* y la *criosfera*. Modelos clásicos para la atmósfera y el océano líquido son cada vez más fielmente *validados*.

Como es natural, en ese camino también aparecen dificultades. Las escalas espaciales son muy variadas, del centímetro (o menos) a decenas de miles de kilómetros, y las *escalas temporales* involucradas son también muy diferentes: de las de los casquetes polares a las de la evolución de las plantas. Es de esta manera, haciendo frente paulatinamente a estas dificultades, como se construye, bloque a bloque, una jerarquía de modelos matemáticos cada vez más completa con la idea de que los modelos sean cada vez más fielmente válidos.

Pero volvamos al clima. Su evolución se rige por un apilamiento, un edificio, donde muchos de los ladrillos de base son caóticos y con unas condiciones iniciales que, cualquiera que sea el horizonte en el que uno se sitúe, no son conocidas más que parcialmente. ¿Será entonces impredecible?

Para empezar, la cuestión no es nueva. La turbulencia, especie de caos a la vez temporal y espacial, preocupaba ya a Lucrecio (entre el 95 y el 55 a. C.). La turbulencia

no es más fácil de definir que el caos ni que el propio concepto de complejidad. Sin embargo, no hay ninguna ambigüedad al afirmar que la turbulencia es un fenómeno complejo; de hecho, uno de los grandes problemas del final del siglo xx. La consideración de la *sensibilidad* de los modelos ante los datos y las escalas es, pues, de una importancia capital.

Se nos presenta una jerarquía de modelos; numerosos centros de investigación y de previsión van a comparar, en los años venideros, métodos y resultados. ¿Qué tipo de decisiones tendrán que tomar?

Los distintos modelos actuales no dan todavía resultados totalmente similares. A veces, incluso sus consecuencias son opuestas. Pero todos indican cambios: la mayoría un aumento medio de la temperatura. Y no existe ninguna prueba de lo contrario, es decir, «que no pasa nada» o «que lo que pasa es irrelevante». El sentido común obliga, por tanto, a tomar medidas de «seguridad». Medidas conservadoras destinadas a parar, o al menos a ralentizar, todo lo que es claramente nocivo. Éste es precisamente el comienzo de decisiones afortunadamente tomadas, por ejemplo, en el protocolo de Montreal en septiembre de 1987.

Es más que probable que los cambios que se produzcan se notarán positiva o negativamente según las regiones, los estados o las economías. «Conflicto de intereses» clásico en Economía. Entramos así en un tema nuevo, inmenso, un tema mixto que mezcla los modelos climatológicos y los modelos económicos.

Para precisar algo más esa interacción hace falta una relación de las variables de control: impuestos, normas de emisión de gas con efecto invernadero, reglamentación sobre la contaminación, etc. También conviene no «congelar» en análisis demasiado rápidos variables que de hecho deben ser consideradas como controles. A continuación es preciso construir las funciones de coste, los criterios que hay que optimizar y que hacen intervenir restricciones físicas y económicas. Estos criterios son locales, casi individuales, nacionales, regionales o mundiales. Hará falta incluir «costes ecológicos», que habrá que definir. No hay ningún motivo para que, en estas condiciones, exista una solución óptima. Estamos ante una situación de criterios múltiples donde la única esperanza de llegar a políticas (o controles) aceptables pasa por la cooperación para los que resulta pertinente utilizar la noción de equilibrio introducida por W. Pareto (1848-1923).

La aplicación de la metodología de la *trilogía universal* podrá mostrar tendencias y permitirá realizar algunos escenarios globales. Su aplicación a problemas «locales» ya ha dado lugar a numerosos éxitos: gestión óptima de la energía hidráulica, redes de transporte de energía, etc. Es por esto que nos parece muy probable que se pueda ir más lejos en temas aún pendientes: gestión del «sistema de aguas», control de la evolución de los sistemas vivos en océanos, lagos, estuarios, lagunas, aguas subterráneas, etc.

Es evidente que la aplicación de la *trilogía universal* no basta, por sí misma, para resolver el conjunto inmenso de problemas planteados en torno al sistema climático.

Muchas otras disciplinas están también implicadas en ese reto. Pero si la aplicación de la *trilogía* no resulta *suficiente* para el estudio y la comprensión del sistema del planeta Tierra, hoy día es ya un hecho constatado que su aplicación sí es poco menos que *necesaria*.

2. UN EJEMPLO: MODELOS CLIMÁTICOS DE BALANCE DE ENERGÍA

Aunque en la conferencia se abordaron diferentes modelos atmosféricos y oceánicos, la presente exposición se limitará tan sólo a una clase de modelos climáticos: los denominados de *balance de energía*.

Antes de entrar en detalles conviene apelar a la importante diferencia entre Climatología y Meteorología en atención a la muy distinta escala temporal que los caracteriza. En el segundo caso la predicción se pretende con gran exactitud, por lo que el periodo en consideración se suele limitar a días y, a lo sumo, un par de semanas. La predicción suele requerir métodos computacionales que contrastan con los de naturaleza más cualitativa utilizados en Climatología para el análisis y diagnóstico, sobre modelos simplificados, a grandes escalas temporales que van desde la década a las decenas de siglos.

No viene mal acudir a la definición que un experto de reconocido prestigio (Schneider [38]) propone sobre la noción de clima: Estado promediado de la atmósfera observado como tiempo meteorológico sobre un periodo finito de tiempo a lo largo de los años. En los estudios climáticos, la incógnita no es tanto la temperatura puntual e instantánea, $T(x, t)$, que viene determinada por los principios de la Termodinámica, sino los promedios espaciales y temporales definidos a partir de tal temperatura:

$$u(x, t) = \frac{1}{2\tau|B(x)|} \int_{t-\tau}^{t+\tau} \int_{B(x)} T(y, s) dy ds.$$

Una de las clases de modelos climáticos más elementales, pero de gran valor de diagnóstico, son los llamados de *balance de radiación de energía*. Aunque los antecedentes se puedan remontar a un trabajo de S. Arrhenius de 1896, se podría decir que tal tipo de modelos fue propuesto independiente pero simultáneamente, en 1969, por M. I. Budyko [6] y W. D. Sellers [39].

2.1. Modelización

Los modelos obedecen a un sencillo balance entre las distintas energías (la parte absorbida de la emitida por el Sol, R_a , y la emitida por la Tierra como cuerpo caliente, R_e) y los términos que expresan la variación de energía calorífica en términos de la capacidad calorífica y de la derivada de la temperatura promediada y la redistribución espacial, D ;

$$c \frac{\partial u}{\partial t} = R_a - R_e + D.$$

Una deducción más cuidada puede ser llevada a cabo partiendo de la ecuación de conservación de la energía interna y aplicando los operadores de promedio involucrados en la definición de $u(x, t)$ (véase, por ejemplo, Kiehl [27]).

En todo caso, es obvio que se requiere más información para poder «cerrar» el modelo de manera que quede expresado únicamente en términos de la temperatura promediada $u(x, t)$. Esa información adicional puede ser entendida como adecuadas *leyes de estado* que describiremos a continuación. La fracción de radiación absorbida se expresa como $R_a(u) = Q \beta(u)$, donde $\beta(u)$ es el coalbedo planetario y Q es la constante solar cuyo valor actual es de $Q = 342,5 \text{ W/m}^2$. La función coalbedo toma valores comprendidos entre 0 y 1, y representa la fracción de energía recibida que es absorbida por la superficie: en otras palabras, el cociente entre la energía absorbida y la energía incidente. En zonas cubiertas de hielo refleja más la luz solar que, por ejemplo, en los océanos y, por tanto, el coalbedo es mayor en estas últimas. Se observa que existen zonas muy próximas con coalbedos muy diferentes. Es los modelos de balance de energía se considera una variación rápida del coalbedo en un entorno de una temperatura crítica, que usualmente se toma como $u = -10 \text{ }^\circ\text{C}$, y que corresponde a la temperatura en la que el hielo pasa de ser transparente a blanco. En el modelo propuesto por Budyko, $\beta(u)$ es discontinua.

$$\beta(u) = \begin{cases} \beta_i & u < -10, \\ [\beta_i, \beta_\omega] & u = -10, \\ \beta_\omega & u > -10, \end{cases} \quad (1)$$

En el modelo propuesto por Sellers, se supone que $\beta(u)$ es una función más regular (Lipschitziana, al menos), como por ejemplo

$$\beta(u) = \begin{cases} \beta_i & u < u_i, \\ \beta_i - \left(\frac{u - u_i}{u_\omega - u_i} \right) (\beta_i - \beta_\omega) & u_i \leq u \leq u_\omega, \\ \beta_\omega & u > u_\omega, \end{cases} \quad (2)$$

con u_i y u_ω temperaturas fijadas y cercanas a $-10 \text{ }^\circ\text{C}$.

El otro ingrediente en el balance es la radiación que por enfriamiento emite la Tierra. Esta energía emitida, $R_e(u)$, es representada de dos modos distintos según los autores antes mencionados. En el modelo de Sellers $R_e(u)$ se expresa escribiendo u en grados Kelvin (por tanto $u > 0$) y utilizando la ley de Stefan-Boltzmann

$$R_e(u) = \sigma(u)u^4 \quad (3)$$

donde σ es una función regular, positiva y acotada que representa la *emisividad*. En el modelo de Budyko se argumenta que, dado que las variaciones observadas son pequeñas, basta considerar una linealización de esa ley en torno a la temperatura media terrestre del momento (aproximadamente 15 °C), por lo que se toma una expresión lineal con coeficientes obtenidos mediante observación por satélite

$$R_e(u) = Bu + C \quad (4)$$

con B y C parámetros positivos (en la actualidad $A = 210 \text{ W/m}^2$, $C = 1,9 \text{ W/m}^2$ y que incluyen fenómenos empíricos tales como el efecto invernadero, cambios antropogénicos, etc. Son las llamadas *variables internas*.

Pero pasemos al primer escalón de modelos distribuidos de balance de energía. La necesidad de tomar modelos más sofisticados que el cero dimensional es obvia, pues todos sabemos que en ciertas zonas del planeta la vida es más agradable que en otras debido a la diferencia de climas. Una primera precisión a lo anteriormente expuesto consiste, pues, en suponer ahora que

$$R_a = R_a(x, u) = QS(x)\beta(u)$$

con $S(x)$ la llamada *función de insolación anual*, que es una función regular y con valores positivos (cuando se toman escalas temporales menores se ha de suponer $S = S(x, t)$ tomando valores nulos durante los periodos de «noche polar»). Ésa es la función que M. Milankovitch calculó minuciosamente, en 1920, para tiempos pasados por medio de la Mecánica Celeste y que le permitió ofrecer una justificación de las glaciaciones pasadas. Por otra parte, en lo concerniente a la energía de emisión, es natural suponer que

$$R_e = R_e(x, u)$$

con $R_e(x, u)$ dada como antes salvo que ahora $\sigma = \sigma(x, u)$ y $B = B(x)$, $C = C(x)$.

Un primer modelo en esta dirección se obtiene al considerar la temperatura superficial promediada en el tiempo anual o estacionalmente y en el espacio sobre paralelos de anchura infinitesimal. La razón de considerar tales promedios espaciales se debe a que las observaciones disponibles muestran que la convención meridional es considerablemente pequeña frente a la longitudinal. La consideración de la dependencia espacial

de la temperatura introduce en el balance un nuevo término, la redistribución de calor, que viene dada mediante un operador de difusión de segundo orden. Como es acostumbrado en Mecánica Celeste, se suele introducir una nueva variable $x \in (-1, 1)$ dada por $x = \sin \lambda$ con λ la latitud. Incorporando al anterior balance un término de difusión, la evolución de la temperatura superficial $u(x, t)$ aparece gobernada por la siguiente ecuación parabólica semilineal,

$$c(x)u_t - (k(1 - x^2) u_x)_x = R_a(x, u) - R_e(x, u), \quad x \in (-1, 1), \quad t > 0, \quad (5)$$

donde $c(x)$ es la capacidad calorífica y k el coeficiente de difusión (que a veces se supone dependiente de la posición x , de u o de la variación de la temperatura u_x), tomando c y k valores positivos (véase, por ejemplo, North [37]).

Nótese que la ecuación (5) es *degenerada* pues el coeficiente del gradiente u_x se anula en los extremos (es decir, en los polos). De hecho, otro tipo de degeneración puede aparecer en ciertas variantes de estos modelos. Así, P. H. Stone [41] propuso un coeficiente de difusión de la forma $k = k(x, u, \nabla u)$, en particular $k = b(x) |\nabla u|$, con el fin de incluir los efectos negativos producidos por las corrientes atmosféricas de gran escala. En este caso, la ecuación semilineal (así llamada pues el término no lineal no afecta a las derivadas de la incógnita) (5) pasa a ser una ecuación cuasilineal (en la que ahora el término no lineal involucra a las derivadas espaciales de la incógnita)

$$c(x)u_t - ((1 - x^2)^{\frac{p}{2}} |u_x|^{p-2} u_x)_x = R_a(x, u) - R_e(x, u) \quad (6)$$

con $p = 3$ (el caso genérico $1 < p < \infty$ tiene la virtud de incluir también en su seno a la ecuación semilineal cuando se toma $p = 2$).

Pasemos ahora al caso bidimensional. Existen interesantes variantes y generalizaciones de los modelos unidimensionales (5) y (6). Por ejemplo, el estudio de la distribución de la temperatura superficial promediada como función del tiempo, de la latitud y también de la longitud da lugar a una ecuación no lineal en derivadas parciales de tipo parabólico

$$c(x)u_t - \operatorname{div}(k\nabla u) = R_a - R_e \quad (7)$$

que ahora tiene lugar sobre una superficie que representa a la superficie terrestre. En una primera aproximación se puede identificar esta superficie con la de la esfera unidad (tomando como unidad el radio de la Tierra) y que se suele denotar como la superficie \mathbb{S}^2 de \mathbb{R}^3 . En ese caso, usando coordenadas esféricas, $x = (\cos \varphi \sin \lambda, \cos \varphi \cos \lambda, \sin \lambda)$ se tiene la identificación con los puntos de un rectángulo del plano introducida en 1568 y comúnmente denominada como proyección de G. Mercator (1512-1594).

Cuando, por ejemplo, $k = 1$, aparece el llamado *operador de Laplace-Beltrami*, que admite una expresión directa en términos de las derivadas parciales de u con respecto a λ y a φ . De hecho, si se supone que $u(\lambda, \varphi, t) = u(\lambda, t)$, es decir, u es independiente de φ , entonces encontramos el operador en derivadas parciales de (5). En esta formulación se hace evidente que no se ha de añadir condición de contorno ninguna. Sin embargo, la anterior modelización mediante coordenadas esféricas no es totalmente satisfactoria por dos razones distintas. En primer lugar, la parametrización de la latitud provoca singularidades en los polos (la derivabilidad de la incógnita en esos puntos está defectuosamente contemplada). Pero además, es bien sabido que la superficie terrestre no es exactamente una esfera. Por estas razones, el marco matemático adecuado para formular más correctamente el modelo bidimensional es el de las *variedades diferenciables*. La idea es recubrir la superficie en estudio, que ahora denotaremos por M , por un conjunto de abiertos U tales que posean una «buena» proyección (un *difeomorfismo*) sobre una parte del plano $\phi: U \rightarrow \mathbb{R}^2$ y de tal manera que exista una buena transición de una «carta» (así es como se denomina al par (U, ϕ)) a otra $((V, \gamma))$ si es que existe un solapamiento entre ellas. Las distancias y los ángulos, correctamente definidos por el producto euclídeo de vectores de \mathbb{R}^2 , pueden ser extendidos a la variedad M generándose así una métrica (se dice entonces que M es una *variedad Riemanniana bidimensional*). En nuestro caso pediremos también que esa variedad sea «compacta y sin borde». Finalmente, los operadores diferenciales, gradiente de una función escalar y divergencia de una función vectorial, pueden extenderse también al caso de funciones definidas sobre M , con lo que se puede dar perfecto sentido a la ecuación (7) (véase, por ejemplo, Díaz y Tello [18]).

También han sido formuladas numerosas variantes a los modelos citados anteriormente. Entre ellas figuran modelos de balance de energía con un término estocástico de *ruido blanco* para incorporar las repercusiones por la acción de los volcanes (véase North [37] y el tratamiento matemático en G. Díaz y J. I. Díaz [7]). Otro aspecto que también ha sido estudiado en la literatura se refiere al carácter periódico en tiempo de la función de insolación $S = S(x, t)$. Este hecho conduce a los llamados *modelos estacionales*, en los que la unidad de la escala temporal es una estación y en los que se supone que S es anualmente periódica. Mencionemos también un tipo de modelos unidimensionales formulados mediante ecuaciones integro-diferenciales (véase Budyko [6] y Held y Suarez [23])

$$c(x) \frac{\partial u}{\partial t} = R_a - R_e + \gamma \left(\int_{-1}^1 u(x, t) dx - u(x, t) \right).$$

Modelos climatológicos de balance de energía más sofisticados son formulados como un sistema de ecuaciones en derivadas parciales para la temperatura atmosférica, la temperatura superficial y la humedad (véase Hetzer, Jarausch y Mackens [26] y su bibliografía).

Pese a las sofisticaciones mencionadas en la última observación, es claro que si «completamos» un modelo simple, como es el de balance de energía, con muchos otros factores, éste puede llegar a perder el valor de su simplicidad y convertirse en un modelo complejo a semejanza de los de Circulación General.

2.2. Sobre el análisis matemático y numérico de los modelos

Comencemos considerando el modelo cero-dimensional en el que, tras renormalización de la escala temporal, podemos suponer siempre que $c = 1$, es decir, se trata del problema (de los denominados de Cauchy)

$$\begin{cases} \frac{du}{dt} = R_a(u) - R_e(u), \\ u(0) = u_0. \end{cases} \quad (8)$$

La existencia de una única *solución clásica local*, es decir, tal que $u \in C^0([0, \tau]) \cap C^1(0, \tau)$ para algún $\tau > 0$ y verifica la ecuación en todo punto, se sitúa dentro del marco de los trabajos pioneros de Auguste Cauchy (1789-1857), Giuseppe Peano (1858-1932) y Auguste Picard (1884-1962), cuando se supone que $R_a(u)$ y $R_e(u)$ son funciones «regulares» (de hecho basta que sean Lipschitz continuas) como es el caso de las hipótesis estructurales propuestas por Sellers. Además, como también se supone que R_a es una función acotada y R_e creciente, esa solución local se puede extender, de manera única, a una *solución clásica global* definida en todo el intervalo $[0, \infty)$.

El estudio del problema bajo las hipótesis propuestas por Budyko, con R_a discontinua, es notablemente diferente. Los problemas de Cauchy para funciones discontinuas han sido ampliamente estudiados en la literatura, desde las contribuciones de Constantin Carathéodory (1873-1950). En ese caso, no cabe esperar que exista una solución local clásica $u(t)$. Sin embargo, es posible mostrar la existencia de, al menos, una *solución local fuerte*, es decir tal que $u \in C([0, \tau])$, du/dt existe (y es, al menos, una función localmente integrable) y la ecuación se verifica en todos los puntos de $[0, \tau)$ *salvo, a lo sumo, un subconjunto de medida nula*.

El paso de solución local débil a *solución global fuerte* tampoco presenta una especial dificultad en virtud de las hipótesis supuestas sobre R_a y R_e .

Como se ha indicado, la demostración de la unicidad de soluciones es estándar en el caso Lipschitz (modelo de tipo Sellers), pero, sin embargo, pasa a ser una cuestión mucho más delicada cuando $R_a(u)$ es discontinua en u (modelo de tipo Budyko). De hecho, la consideración de ese tipo de cuestiones conduce a reformular el problema (8) como un *problema de tipo multívoco*. En efecto, el valor de la función $R_a(u)$ en el punto de discontinuidad $u = -10$ (recuérdese que $R_a(u) = Q\beta(u)$ con β dada por (1))

no puede ser definido con precisión, por lo que es conveniente suponer que a ese punto se le asigna todo el intervalo, es decir, $\beta(-10) = [\beta_i, \beta_w]$. Ahora se suele identificar β con su *grafo* (como subconjunto de \mathbb{R}^2) y, dado que $R_a(u)$ se supone no decreciente en u , ese grafo pertenece a la clase de los llamados *grafos maximales monótonos* (la condición de maximal le viene por la imposibilidad de incluir ese grafo en otro monótono mayor: véase Brezis [5]). En este contexto, el problema (8) debería ser escrito más propiamente como

$$\begin{cases} \frac{du}{dt} \in R_a(u) - R_e(u) & \text{en } (0, \infty), \\ u(0) = u_0, \end{cases} \quad (9)$$

y entenderlo en el sentido de que existen un par de funciones u y b , con

$$b(t) \in \beta(u(t)) \text{ para casi todo } t \in (0, \infty), \quad (10)$$

tales que

$$\begin{cases} \frac{du}{dt}(t) = Qb(t) - R_e(u(t)) & \text{para casi todo } t \in (0, \infty), \\ u(0) = u_0. \end{cases} \quad (11)$$

Una vez hechas estas precisiones, en lo sucesivo, no haremos ninguna distinción en los símbolos utilizados para describir las ecuaciones asociadas a una función $R_a(u)$ (o β) que unas veces pueda ser supuesta regular y en otras un grafo maximal monótono, entendiéndose que en el último caso se han de introducir matizaciones del estilo de las anteriormente expuestas.

Pero volvamos a la cuestión de la unicidad de soluciones para el caso de $R_a(u)$ (o β) multívoco. Para ilustrar globalmente la situación que se presenta ante esa disyuntiva es útil comenzar estudiando el conjunto de soluciones estacionarias u_∞ de (8), es decir, $u_\infty \in \mathbb{R}$ satisfaciendo

$$0 = R_a(u) - R_e(u). \quad (12)$$

El conjunto de soluciones de (12) depende obviamente de la elección de R_a y R_e , así como de los diferentes valores de los parámetros. Es especialmente relevante analizar las distintas soluciones que se presentan cuando se varía la constante solar Q . Si, por simplicidad, suponemos $R_e(u)$ lineal, el efecto similar a variar Q es el de variar las rectas generadas por $R_a(u)$, con lo que se llega a diferentes casos que son esquematizados en la Figura 1 (para el caso de una $R_a(u)$ de tipo Sellers véase (2)).

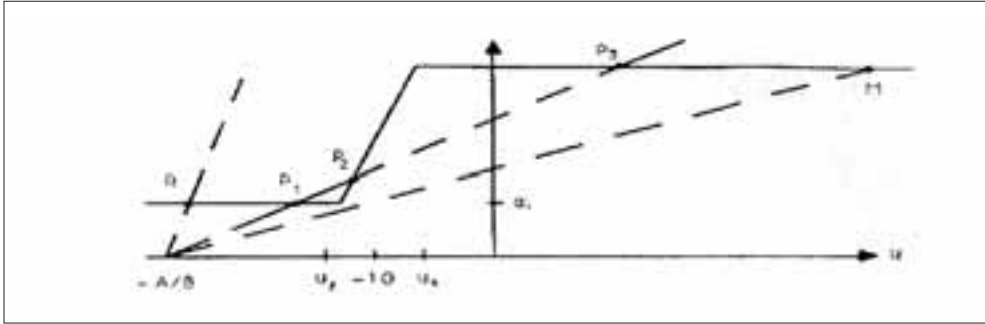


Figura 1. Intersección entre dos gráficas obtenidas para diferentes valores de Q .

Se obtiene un *diagrama de bifurcación*, como el de la Figura 2, que permite ver el número de soluciones para un valor dado de Q .

Volviendo al caso del problema de Cauchy para $R_\alpha(u)$ (o β) multívoco, es sencillo ver que si, por ejemplo, tomamos Q tal que $u_\infty^1 = -10$ sea una de las tres soluciones de equilibrio $u_\infty^2 < u_\infty^1 < u_\infty^3$, entonces el problema de Cauchy (8) (o, más propiamente, (9)) con dato inicial $u_0 = -10$ tiene infinitas soluciones. En efecto, $u_1(t) = -10$ es una solución, en ese caso tomaríamos $b = (-10B + C)/Q \in \beta(-10)$. Dos soluciones distintas vienen dadas por

$$\begin{cases} c \frac{du_2}{dt}(t) = Q\beta_i - Bu_2(t) - C, \\ u_2(t_0) = -10, \end{cases}$$

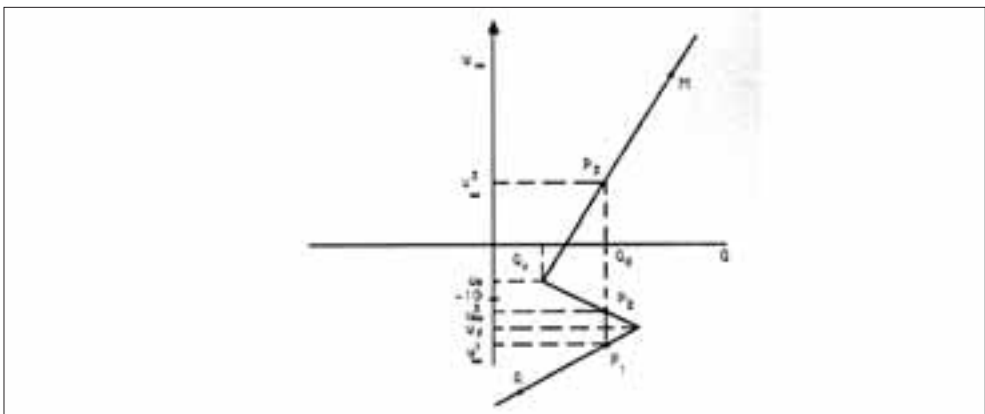


Figura 2. Diagrama de bifurcación en función de la constante solar Q para el modelo de tipo Sellers.

y

$$\begin{cases} c \frac{du_3}{dt}(t) = Q\beta_\omega - Bu_3(t) - C, \\ u_3(t_0) = -10. \end{cases}$$

Además, como la ecuación de (8) es autónoma (las funciones $R_a(u)$ y $R_e(u)$ no dependen explícitamente de t) es claro que las funciones

$$u_\sigma = \begin{cases} -10 & t \in [t_0, t_0 + \sigma], \\ u_2(t - \sigma) & t \in [t_0 + \sigma, \infty), \end{cases}$$

y

$$\hat{u}_\sigma = \begin{cases} -10 & t \in [t_0, t_0 + \sigma], \\ u_3(t - \sigma) & t \in [t_0 + \sigma, \infty), \end{cases}$$

son también soluciones de (8), independientemente del valor de $\sigma > 0$. Finalmente, estudiando el signo de $R_a - R_e$ es fácil obtener una descripción dinámica del comportamiento cualitativo de las soluciones.

El estudio del comportamiento cuando $t \rightarrow \infty$ puede ser llevado a cabo fácilmente bajo hipótesis mucho más generales. Si comenzamos por abordar el caso en el que R_a y R_e son funciones regulares monótonas (con R_e acotada) encontraríamos primero que el diagrama de bifurcación sigue siendo, en sentido amplio, una curva en forma de «ese» con al menos dos *puntos de retorno*. Analizando el *problema linealizado* podríamos comprobar que las ramas crecientes de la curva de bifurcación están formadas por *estados* (puntos en nuestro caso) *estables* (ante «pequeñas» variaciones del dato inicial) y las decrecientes por *estados inestables*. Si tomásemos para Q el valor estimado en nuestros días, veríamos que la solución u_∞^3 representa la temperatura globalmente promediada del planeta (cercana a 15 °C). La solución u_∞^2 es una solución extraña inestable y la u_∞^1 corresponde a un posible estado de glaciación (nótese que es inferior a -10 °C) que es también estable.

De hecho, el estudio de la estabilidad puede ser fácilmente extendido al caso de perturbaciones iniciales arbitrarias (no necesariamente «pequeñas»). En efecto, es fácil ver que la *función potencial*

$$F(u) = -\int_{u^*}^u (R_a(v) - R_e(v)) dv, \tag{13}$$

(con u^* arbitrariamente fijado) es una función de Lyapunov pues

$$\frac{dF}{dt}(u(t)) = - \left(\frac{du}{dt}(t) \right)^2$$

Así, los mínimos relativos de F corresponden a soluciones estables y los máximos relativos a soluciones inestables. Del diagrama de bifurcación es fácil deducir que si Q decrece desde su valor en nuestros días (por ejemplo debido a una polución atmosférica generalizada, por la acción simultánea de numerosos volcanes, por la caída de un enorme asteroide sobre la Tierra, etc.) la temperatura estacionaria asociada $u_3^\infty(Q)$ tendría que comenzar a disminuir (o, lo que es lo mismo, el punto $(Q, u_3^\infty(Q))$ bajaría por la izquierda del diagrama). Si Q atravesase un cierto valor crítico (Q_c , el correspondiente al punto de retorno de la curva) la temperatura tendría que decrecer dramáticamente hasta buscar otra temperatura promediada estable $u_1^\infty(Q_1)$ que correspondería a una glaciación. El camino inverso también sería abrupto, pues partiendo de $u_1^\infty(Q_1)$ si ahora Q aumentase (debido al cese de las causas que llevaron a su disminución) tendríamos que ascender por otro camino, no coincidente con el anterior, hasta alcanzar una temperatura estable $u_3^\infty(Q_1)$. Es el fenómeno conocido como *histéresis* que ocurre con gran frecuencia en la naturaleza y en muchos otros sistemas (véase, por ejemplo, Krasnoselskii y Pokrovskii [28] y sus referencias). Procesos de ese tipo son también típicos de la *Teoría de Catástrofes* (Thom [43]).

El interés de los modelos de balance de energía está motivado precisamente por este tipo de resultados que muestran la sensibilidad de la solución con respecto a parámetros de gran importancia, como es la constante solar Q . Pero, ¿se mantendrán ese tipo de resultados para modelos de balance de energía no homogéneos más realistas en los que se incorpore la difusión espacial de temperaturas?

Escalando en complejidad, consideremos ahora el modelo de balance de energía bidimensional asociado a una representación de la superficie terrestre como una variedad Riemanniana bidimensional compacta y sin borde (como, por ejemplo, $M = \mathbb{S}^2$). Recordemos que una formulación posible es la de encontrar $u = u(x, t)$ con $(x, t) \in M \times [0, \infty)$ tal que

$$(P) \begin{cases} c(x)u_t - \operatorname{div}(k\nabla u) = R_a(x, u) - R_e(x, u) & \text{en } M \times (0, \infty), \\ u(x, 0) = u_0(x) & \text{en } M. \end{cases} \quad (14)$$

Observemos que si suponemos $M = \mathbb{S}^2$ y si las funciones $c(x)$, $R_a(x, u)$ y $R_e(x, u)$ sólo dependen de la latitud entonces cada solución $u_1(x, t)$ del problema unidimensional (5) genera una solución $u_2(x, y, t)$ del modelo (14) mediante rotación alrededor del eje terrestre, es decir, $u_2(x, y, t) = u_1(\operatorname{sen} \varphi, t)$ donde $(x, y) \in \mathbb{S}^2$ y φ es la latitud.

Como en el caso del modelo cero-dimensional, los resultados de existencia y unicidad de soluciones de (5) son de diferente naturaleza según que se suponga la función R_a discontinua (o multivaluada) o no. La existencia de, al menos una, solución, supuesto el dato inicial regular, puede probarse sin dificultad mediante diferentes métodos. La cuestión de la unicidad de soluciones cuando R_a es discontinua es mucho más delicada (en contraste con el caso en el que R_a es Lipschitz continua). En primer lugar, es posible mostrar mediante contraejemplos que la unicidad se vulnera cuando, por ejemplo, el dato inicial toma su máximo o mínimo al nivel $u_0 = -10$. Sin embargo, si el dato inicial u_0 atraviesa el nivel $u_0 = -10$ de manera «no degenerada» ($\nabla u_0(x_0) \neq 0$ en los puntos $x_0 \in M$ donde $u_0(x_0) = -10$) entonces se mantiene la unicidad de soluciones (Díaz y Tello [18]).

Antes de analizar la cuestión de la estabilidad y bifurcación de los estados estacionarios, merece la pena entretenerse en mencionar un aspecto notable que, a diferencia del caso del modelo cero-dimensional, cobra ahora sentido. Nos referimos a la simulación de las *fronteras libres* separando las zonas de la Tierra cubiertas por hielo y nieve. Recordando los argumentos aludidos para la parametrización del distinto albedo terrestre, tales curvas vendrían definidas como las curvas para las que se produce una transición brusca en el co-albedo, es decir, sobre las que se ha de tener que $u = -10$. El modelo mejor adaptado para tal simulación es el de tipo Budyko, en el que se supone que el co-albedo se representa por una función discontinua.

Si, por simplicidad, retornamos al caso unidimensional (véase (5)), las fronteras libres ahora pasan a ser puntos (desconocidos *a priori*), separando la zona cubierta de hielo y nieve $\{x \in [-1, 1] : u(x, t) < -10\}$ del resto. Supongamos que esos puntos se limitan a dos $S_+(t)$, $S_-(t)$ que corresponden a las latitudes que «representan» los promedios de las fronteras de los dos casquetes polares. Se tiene pues que $u(S_{\pm}(t), t) = -10$. El intervalo temporal de definición de las interfases $S_+(t)$, $S_-(t)$ no es siempre $[0, \infty)$ sino que depende fuertemente de los datos Q , $S(x)$, de la especificación de $R_a(x, u)$ y del dato inicial $u_0(x)$. Puede probarse (Xu [45]) que $S_+(t)$ y $S_-(t)$ son funciones C^∞ (al menos mientras no colapsen) que evolucionan según la ecuación diferencial

$$S'_+(t) = \frac{k[u_{tx}(S_+(t)-, t) - u_{tx}(S_+(t)+, t)]}{QS(S_+(t))(\beta_f - \beta_i)}$$

y la análoga para $S'_-(t)$, donde se ha utilizado la notación genérica de $f(t)_+ = \lim_{h \downarrow 0} f(t+h)$. El caso bidimensional es más delicado.

En ocasiones, la representación del borde de los casquetes polares por medio de una curva sobre la superficie M no es enteramente satisfactoria, pues imágenes emitidas desde los satélites muestran extensiones, que pueden llegar a tener un espesor de 40 km, en donde es difícil, sino imposible, decidir si se trata de agua o hielo. Una formulación matemática de ese tipo de situaciones equivaldría a suponer que el conjunto

$M(t) = \{x \in M : u(x, t) = -10\}$ tuviese medida positiva. A veces se denomina a tal conjunto como *zona pastosa* (*mushy región*, en inglés) por su similitud con regiones similares provenientes de la modelización de problemas de Stefan de cambio de fase. Es posible mostrar que si u es solución del problema (14) entonces tal conjunto se reduce a un conjunto de curvas que no puede tener medida positiva. Por el contrario, en el caso del modelo cuasilineal (6), propuesto por Stone, en el que $k = |\nabla u|^{p-2}$, se tiene que el conjunto $M(t)$ puede tener medida positiva si $p > 2$ (Díaz [13]).

Los estados de equilibrio $u_\infty(x)$ del problema (14) son ahora bastante más complejos que las constantes que aparecían en el modelo cero-dimensional. Ahora vienen dados como soluciones del problema estacionario

$$-\operatorname{div}(k \nabla u_\infty) = R_a(x, u_\infty) - R_e(x, u_\infty), \text{ en } M. \tag{15}$$

Como $R_e(\cdot, u_\infty)$ es creciente en u , la ecuación (15) tiene carácter de *problema no lineal del tipo de autovalor* y, así, la multiplicidad de soluciones depende fuertemente de los diferentes valores del parámetro solar Q . Usando técnicas clásicas de análisis no lineal (tales como el *método de continuación* y el *Teorema de la función implícita* (véase, por ejemplo, Zeidler [45]), se prueba que, bajo adecuadas condiciones, al igual que en el diagrama de bifurcación del modelo cero-dimensional, hay una curva de bifurcación que comienza en un estado u_∞^0 cuando $Q = 0$ y tiende al infinito si $Q \rightarrow +\infty$ teniendo además un número par de puntos de retorno correspondientes a $Q = Q_c^1$ y $Q = Q_c^2$. De esta manera, si $Q < Q_c^1$ o $Q > Q_c^2$ hay unicidad de soluciones y si $Q \in (Q_c^1, Q_c^2)$ existen al menos tres soluciones (un resultado más fino puede hallarse en Arcoya, Díaz y Tello [1]).

El estudio de la estabilidad de los estados estacionarios puede llevarse a cabo por medio de distintos métodos. Un primer resultado en esa dirección es el que muestra que el llamado *conjunto ω -límite*,

$$\omega(u) = \{u_\infty : M \rightarrow \mathbb{R} : \exists t_n \rightarrow +\infty \text{ para los que } u(t_n, \cdot) \rightarrow u_\infty\},$$

está formado únicamente por soluciones de (15) (Díaz, Hernández y Tello [14]). La estabilidad ante «pequeñas variaciones» del dato inicial se puede analizar por medio del *principio de estabilidad linealizada* cuando se supone que las funciones R_a (o, lo que es lo mismo, β) y R_e son derivables respecto de u . El problema linealizado en torno a un estado estacionario u_∞ es también de tipo de autovalores

$$-\operatorname{div}(k \nabla v) - [QS\beta'(u_\infty) - \frac{\partial}{\partial u} R_e(x, u_\infty)] v = \lambda v, \text{ en } M. \tag{16}$$

Es bien conocido que los autovalores de este problema constituyen una sucesión no decreciente $\{\lambda_j\}_{j \in \mathbb{Z}}$ que tiende a infinito y, además, se tiene que $\lambda_0(Q, u) < \lambda_1(Q, u)$. Según el principio de estabilidad linealizada, dado $Q > 0$, una solución u_∞ de (15) es asintóticamente estable (respectivamente, inestable), en $C([-1, 1])$, si el punto de equilibrio del problema parabólico verifica $\lambda_0(Q, u_\infty) < 0$ (respectivamente, $\lambda_0(Q, u_\infty) > 0$) (para más detalles sobre la aplicación al problema (15) véase Hetzer [25]).

La estabilidad no lineal (es decir, ante perturbaciones no necesariamente «pequeñas») para el problema (14) puede estudiarse también, y sin necesidad de suponer la derivabilidad de R_a y R_e , con ayuda del funcional

$$J(u) = \int_M \{k |\nabla u(x)|^2 + F(x, u(x))\} dx,$$

donde

$$F(x, u) = -\int_0^u \{R_a(x, v) - R_e(x, v)\} dv.$$

Se tiene que

$$\frac{d}{dt} J(u(\cdot, t)) = -\int_{-1}^1 c(x) \left(\frac{\partial u}{\partial t}(x, t) \right)^2 dx \leq 0$$

por lo que J es un funcional de Lyapunov y las soluciones estables se corresponden con los mínimos de J y las no linealmente inestables con otros puntos críticos como máximos o puntos de silla. La búsqueda de estados estacionarios no se debe limitar, pues, al estudio de los puntos extremales (mínimos y máximos) locales del funcional J . En esta dirección es de señalar la importancia del resultado conocido como el *Lema del paso de montaña* (véase, por ejemplo, Nirenberg [36]) que bajo circunstancias muy generales asegura que entre dos mínimos ha de existir necesariamente un punto de silla (para una aplicación de este lema a modelos climáticos véase Ghil y Childress [22]).

La persistencia o no, para $t \in (0, \infty)$, de las fronteras libres ($S_+(t)$ y $S_-(t)$ en el problema unidireccional) asociadas a estados de evolución $u(x, t)$ es de gran importancia en el estudio del conjunto de los estados de equilibrio $u_\infty(x)$ del problema (15): si por ejemplo $S_-(T_0) = S_+(T_0)$ para algún T_0 y ese punto interior de $(-1, 1)$ entonces se puede mostrar que el estado estacionario $u_\infty(x)$ representa el caso de una superficie terrestre cubierta de hielo y nieve, pero si el punto $S_-(T_0) = S_+(T_0)$ es $x = 1$ o $x = -1$ entonces $u_\infty(x)$ representa la ausencia de superficies heladas.

El estudio numérico de los modelos de balance de energía ha sido llevado a cabo por numerosos autores para elecciones concretas de R_a y R_e (son de citar los trabajos de Held y Suarez, North y Hetzer, Jarausch y Mackens antes citados: véanse también Bermejo [2] y Bermejo, Díaz y Tello [4]). Un método frecuentemente utilizado es el de Galerkin (Mengel, Short y North [34], Lin y North [30], Díaz y Tello [3], Tello [42]). Consideremos el problema

$$(P^*) \begin{cases} u_t - (\rho(x)u_x)_x + Bu + C \in QS(x) \beta(u), & (t, x) \in (0, T) \times (-1, 1), \\ u(0, x) = u_0(x) & x \in (-1, 1) \end{cases}$$

donde $\rho(x) = 1 - x^2$ con $u_0 \in V = \{v \in L^2(\Omega) : v_x \in L^2_p(\Omega)\}$. Se sabe (Díaz [9]) que la solución u está en el espacio $L^2(0, T; V)$. Veamos cómo aproximarla. Construiremos «soluciones aproximadas» pertenecientes a espacios de dimensión finita $V_m \subset V$ generados por las autofunciones w_i del operador diferencial $(\rho(x)u_x)_x$. Gracias a la compacidad del operador inverso del operador diferencial, podemos suponer que $V_m = [w_1, w_2, \dots, w_m]$ espacio vectorial de dimensión m , donde: $w_i \in V$ es una autofunción de autovalor λ_i , para todo m , w_1, \dots, w_m son linealmente independientes y las combinaciones lineales de las funciones w_i , $i \in \mathbb{N}$ son densas en V . En nuestro caso, es bien conocido (véase por ejemplo Legendre [29], Simmons [40]) que las autofunciones del operador de difusión son los polinomios de Legendre, dados por $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$, $P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$ y en general,

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n(x^2 - 1)^n}{dx^n},$$

correspondientes a los autovalores $\lambda_n = n(n + 1)$. Además

$$\int_{-1}^1 P_i(x)P_j(x) dx = \frac{2}{2j + 1} \delta_{ij}.$$

Llamaremos «solución aproximada» del problema a $u_m = u_m(t, x)$ dada por

$$u_m = \sum_{i=0}^m \alpha_i(t)P_i(\cdot)$$

donde α_i queda determinado por

$$(P_m) \begin{cases} \langle u_{m,t}, P_i \rangle + \langle (1 - x^2)(u_m)_x, (P_i)_x \rangle + \langle Bu_m, P_i \rangle + \langle C, P_i \rangle = \\ = \langle QS(x)z_m, P_i \rangle & i = 0, \dots, m \\ u_m(0, x) = u_{0m}(x), \end{cases}$$

para algún $z_m \in L^\infty((0, T) \times (-1, 1))$, $z_m \in \beta(u_m)$ para todo $(t, x) \in (0, T) \times (-1, 1)$ y siendo $u_{0m}(x) = \sum_{i=1}^m u_0^i P_i(x)$.

Aquí el corchete \langle, \rangle representa el producto escalar en $L^2(\Omega)$. Sustituyendo u_m por su expresión en la base de V_m se obtiene que

$$\begin{cases} \frac{2}{2i+1} \alpha_i'(t) + \frac{2\lambda_i}{2i+1} \alpha_i(t) + \frac{2B}{2i+1} \alpha_i(t) = \Gamma_i(\alpha_1, \dots, \alpha_m) & \text{si } i \neq 0 \\ 2\alpha_0'(t) + 2\lambda_0\alpha_0(t) + 2B\alpha_0(t) + C = \Gamma_0(\alpha_1, \dots, \alpha_m) \\ \alpha_i(0) = u_0^i \end{cases}$$

con $\Gamma_i(\alpha_1, \dots, \alpha_m) = \langle QS(x) z_m, P_i \rangle$. Así, hemos obtenido un sistema de $m + 1$ ecuaciones diferenciales ordinarias con $m + 1$ incógnitas $\alpha_i(t)$ y $m + 1$ datos iniciales u_0^i , $i = 0, \dots, m$. La existencia de solución para este sistema está garantizada mediante resultados conocidos (véase, por ejemplo, Filippov [21]).

Para mostrar la convergencia se deben obtener estimaciones a priori (independientes de m), para lo cual se toma u_m como función test y se llega a que

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} |u_m|^2 + \int_{\Omega} \rho(x) |(u_m)_x|^2 + \left(B - \frac{\varepsilon}{2}\right) \int_{\Omega} |(u_m)|^2 \leq K$$

donde K es una cierta constante positiva independiente de m . Finalmente, tomando ahora $(u_m)_t$ como función test se obtiene

$$\left(1 - \frac{\varepsilon}{2}\right) \int_{\Omega} |u_m'|^2 + \frac{1}{2} \frac{d}{dt} \int_{\Omega} \left(\rho(x) |(u_m)_x|^2\right) + \frac{B}{2} \frac{d}{dt} \int_{\Omega} |(u_m)|^2 \leq K,$$

lo que permite pasar al límite $u_m \rightarrow u$ en V y $u_m' \rightarrow \partial u / \partial t$ en $L^2(0, T; L^2(\Omega))$ y mostrar que u satisface la formulación débil del problema. Numerosas experiencias numéricas son asequibles en la literatura (algunas de ellas utilizadas para analizar previsiones ante escenarios con duplicación del CO_2 , etc.). Una representación (estática) de una de esas simulaciones aparece en la Figura 3.



Figura 3. Predicción de temperatura de hace 115.000 años en el estudio de glaciaciones según Mengel, Short y North (loc. cit.).

2.3. Control

En 1955, John von Neumann escribía [35]:

«Probably intervention in atmospheric and climate matters will come in a few decades, and will unfold on a scale difficult to imagine at present»

Hoy, una parte de su sueño es casi una realidad y así, por ejemplo, los protocolos internacionales sobre emisión de gases de efecto invernadero son los ejemplos más ilustrativos de controles planteados a una escala global. También lo son los protocolos mundiales y europeos sobre emisión de aerosoles (respondiendo a un intento de frenar el crecimiento del agujero, en la estratosfera, de la capa de ozono de la Antártida (beneficioso por proteger a la Tierra de una radiación excesiva de los rayos ultravioleta) o a una altura más cercana a nosotros, la nociva acción de la contaminación (típicamente urbana) del ozono troposférico, sobre las personas y la agricultura.

Si nos centramos en un modesto fin adaptado a nuestros modelos sencillos, nos podemos preguntar si es posible «conducir» una distribución «inicial» de temperaturas (que ahora denotamos por $y(0, x)$) pues en la Teoría de Control la incógnita no es el estado del sistema, que se suele denotar por y , sino el control v hasta una distribución deseada $y_d(x)$ transcurrido un periodo dado T y todo ello «actuando» sólo desde una subvariedad ω de M . Como se ha comentado, la incógnita ahora es el control $v : (0, T) \times \omega \rightarrow R$ que permita que $y(T : v) = y_d$ donde $y(\cdot : v)$ denota la solución de (P) reemplazando $R_e(t, x, y)$ por $g(y) - f(t, x) - v(t, x)\chi_\omega$ con χ_ω la función característica de ω . Cuando la respuesta es afirmativa se dice que (P) es controlable. Sin embargo, el carácter parabólico de la ecuación de (P) conlleva ciertos efectos regularizantes que hacen imposible tal tipo de respuestas para funciones genéricas y_d . Un concepto alternativo es el de la *controlabilidad aproximada*: dado $\varepsilon > 0$ buscamos ahora un control v_ε (definido de nuevo sobre $(0, T) \times \omega$) tal que $d(y(T, v_\varepsilon), y_d) \leq \varepsilon$, donde $d(\cdot, \cdot)$ representa la distancia en algún espacio de funciones sobre M (normalmente $L^2(M)$, o, más en general, $L^p(M)$ con $1 \leq p \leq \infty$).

La naturaleza de nuestro dominio espacial M aporta algunas dificultades adicionales al estudio. Una formulación más sencilla pero que conserva las dificultades esenciales de nuestra meta corresponde al caso en el que sustituimos M por un abierto regular y acotado Ω de \mathbb{R}^2 (aquí \mathbb{R}^2 puede ser sustituido por \mathbb{R}^N con $N \geq 1$). En ese caso, conviene tomar como condición de contorno sobre $(0, T) \times \partial\Omega$ una de tipo de Neumann pues las funciones test asociadas a la formulación débil son similares a las que corresponden al caso en el que M es una variedad Riemanniana sin borde. Otra simplificación irrelevante es suponer $f \equiv 0$. Así, la nueva formulación del problema de controlabilidad aproximada sería la siguiente: dado ω , abierto acotado de Ω , y dados $y_0, y_d : \Omega \rightarrow \mathbb{R}$ y $\varepsilon > 0$ hallar $v_\varepsilon : (0, T) \times \omega \rightarrow \mathbb{R}$ tal que $d(y(T : v_\varepsilon), y_d) \leq \varepsilon$ donde, en general, $y(T : v)$ representa la solución del problema

$$(P_\omega) \begin{cases} y_t - \Delta y + g(y) \in \mathcal{QS}(x)\beta(y) + v\chi_\omega & \text{en } (0, T) \times \Omega \\ \frac{\partial y}{\partial n} = 0 & \text{sobre } (0, T) \times \partial\Omega \\ y(0, \cdot) = y_0(\cdot) & \text{en } \Omega, \end{cases}$$

donde n es el vector unitario exterior a $\partial\Omega$.

Un primer resultado sobre el tema (Díaz [10]) consiste en mostrar que la controlabilidad aproximada depende, de manera esencial, del comportamiento asintótico de los términos no lineales de la ecuación (y no de su regularidad). Así, si se supone que $y_0, y_d \in L^2(\Omega)$, con β multivaluado y acotado y g una función no-decreciente tal que

$$|g(s)| \leq C_1 + C_2 |s| \quad \forall s \in \mathbb{R}, |s| > \bar{M} \tag{17}$$

para algunas constantes C_1, C_2 y \bar{M} , entonces el problema (P_ω) es controlable aproximadamente en $L^2(\Omega)$, es decir existe $v_\varepsilon \in L^2((0, T) \times \omega)$ tal que

$$\|y(T : v_\varepsilon) - y_d\|_{L^2(\Omega)} \leq \varepsilon.$$

El resultado anterior puede ser extendido al caso en el que se reemplaza $L^2(\Omega)$ por $L^p(\Omega)$ con $1 \leq p < \infty$ o $C(\bar{\Omega})$. La idea principal es la aplicación del Teorema de punto fijo de Kakutani (véase Lions [31], [33], Henry [24], Fabré, Puel y Zuazua [19], Díaz [8], Díaz y Ramos [16] para otros resultados relativos). Nótese que este resultado es aplicable al caso particular del modelo de Budyko en el que $g(y) = By$ pero sin embargo la condición (17) no es satisfecha para el modelo de Sellers. Aún más preciso, es posible mostrar (Díaz [8], [10]) que si suponemos

$$g(y) = \lambda |y|^{p-1} y \quad \text{para } y \in \mathbb{R} \text{ y para algún } \lambda > 0 \text{ y } p > 1 \tag{18}$$

entonces aparece un *fenómeno de obstrucción*: supongamos que $\partial\omega$ verifica la condición de la esfera tangente exterior e interior. Sea $y_0 \in L^\infty(\Omega)$. Entonces existe una función $Y_\infty \in C([0, T] \times (\Omega - \bar{\omega}))$ tal que para todo $v \in L^2((0, T) \times \omega)$ la solución correspondiente $y(t, x; v)$ verifica que

$$|y(t, x; v)| \leq Y_\infty(t, x) \quad \text{para } (t, x) \in (0, T] \times (\Omega - \bar{\omega}).$$

La función de obstrucción Y_∞ puede ser construida tal que

$$Y_\infty(t, x) = +\infty \quad \text{sobre } (0, T) \times \partial\omega$$

$$\frac{\partial Y_\infty}{\partial n}(t, x) = 0 \quad \text{sobre } (0, T) \times \partial\Omega.$$

En consecuencia, la condición (18) implica que el problema (P_ω) (en general) no puede ser controlable aproximadamente pues si $|y_d(x)| > Y_\infty(T, x)$ en casi todo punto x sobre un subconjunto de medida positiva D de $\Omega - \bar{\omega}$ entonces, para todo $v \in L^2((0, T) \times \omega)$

$$\|y(T; v) - y_d\|_{L^2(\Omega)} \geq \|Y_\infty(T, \cdot) - y_d\|_{L^2(\Omega)}$$

y así, si $\varepsilon > 0$ es suficientemente pequeño es imposible elegir v de manera que se satisfaga la propiedad requerida. Debido a la relevancia del modelo de Sellers, se plantea una cuestión natural: ¿es posible mostrar la controlabilidad aproximada para el problema P_ω en una clase más pequeña de estados deseados y_d ?

La respuesta es positiva (Díaz [11], [12]). Para explicitar el resultado es cómodo simplificar aún más el problema en consideración limitándonos al problema

$$(P_p) \begin{cases} y_t - \Delta y + \lambda |y|^{p-2} y = v\chi_\omega & \text{en } (0, T) \times \Omega, \\ \frac{\partial y}{\partial n} = 0 & \text{sobre } (0, T) \times \partial\Omega, \\ y(0, \cdot) = y_0(\cdot) & \text{en } \Omega. \end{cases}$$

En primer lugar, conviene introducir las *funciones de obstrucción mejoradas* $\underline{Y}_\infty, \bar{Y}_\infty \in C((0, T] \times \Omega - \bar{\omega})$. La función \underline{Y}_∞ es obtenida como la solución débil (supuesto $y_0 \in L^1(\Omega)$) del problema

$$\begin{cases} \underline{Y}_t - \Delta \underline{Y} + \lambda |\underline{Y}|^{p-2} \underline{Y} = 0 & \text{en } (0, T) \times (\Omega - \bar{\omega}), \\ \underline{Y}_\infty = -\infty & \text{sobre } (0, T) \times \partial\omega, \\ \frac{\partial \underline{Y}_\infty}{\partial n} = 0 & \text{sobre } (0, T) \times \partial\Omega, \\ \underline{Y}_\infty(0, \cdot) = y_0(\cdot) & \text{sobre } \Omega, \end{cases}$$

e \bar{Y}_∞ satisface las mismas condiciones salvo que ahora $\bar{Y}_\infty = +\infty$ sobre $(0, T) \times \partial\omega$. Tras esto ya se puede enunciar el resultado de controlabilidad aproximada restringida. Sean $y_d \in C(\bar{\Omega})$ y $\varepsilon > 0$ dados. Sea $y_d \in C(\bar{\Omega})$ tal que

$$\underline{Y}_\infty(T, x) - \varepsilon < y_d(x) < \bar{Y}_\infty(T, x) + \varepsilon \quad \forall x \in \Omega - \bar{\omega}. \quad (19)$$

Entonces, para todo ε existe $v_\varepsilon \in C([0, T] \times \bar{\omega})$ tal que si $y(t : v)$ es la solución correspondiente de (P_p) se tiene que

$$\|y(T : v_\varepsilon) - y_d\|_{C(\bar{\Omega})} \leq \varepsilon. \quad (20)$$

Nótese que la hipótesis (19) es óptima en el sentido de que si v_ε es tal que se verifica (20), entonces por el principio de comparación

$$\underline{Y}_\infty(t, x) < y(t, x : v_\varepsilon) < \bar{Y}_\infty(t, x) \quad \forall (t, x) \in [0, T] \times (\Omega - \bar{\omega})$$

con lo que

$$\underline{Y}_\infty(T, x) - \varepsilon < y(T, x : v_\varepsilon) - \varepsilon \leq y_d(x) \leq y(T, x : v_\varepsilon) + \varepsilon < \bar{Y}_\infty(T, x) + \varepsilon$$

y necesariamente se tiene (19).

Para concluir, haremos alusión a unas recientes experiencias numéricas sobre la aproximación numérica del control v_ε antes mencionado (Díaz y Ramos [17]). Para no complicar técnicamente la exposición nos limitaremos al caso del problema unidimensional

$$P(u) \begin{cases} y_t - y_{xx} + g(y) = v(t)\delta(0) & \text{en } (0, T) \times (-1, 1), \\ y(\pm 1, t) = 0 & t \in (0, T), \\ y(0, x) = y^0(x) & \text{sobre } (-1, 1), \end{cases} \quad (21)$$

donde g es una función continua no-decreciente, $v(t)$ es el control buscado y $\delta(0)$ es la distribución Delta de Dirac en $x = 0$ (en realidad, para evitar complicaciones adicionales, conviene remplazar la localización en el origen por el caso de un pequeño intervalo ω a su alrededor). Aquí, T está arbitrariamente fijado e y^0 es una función dada ($y^0 \in C^0([-1, 1])$). Para el análisis numérico del control, se aproximará (cuando $k \rightarrow \infty$) éste por una sucesión de soluciones u_k del *problema de control óptimo penalizado* en el que los controles v_k se obtienen minimizando el funcional

$$J_k(v) = \|v\|_{L^2(0, T)} + k \|y(T : v) - y_d\|_{L^2(-1, 1)} \quad (22)$$

(la convergencia rigurosa no es totalmente conocida: véanse los resultados parciales de Lions [32] y Fernández y Zuazua [20]). Las *funciones de obstrucción mejoradas*, $Y_{\pm\infty}$, satisfacen ahora

$$P(\pm\infty) \begin{cases} y_t - y_{xx} + g(y) = 0 & \text{en } (0, T) \times (-1, 1), \\ y(\pm 1, t) = 0, y(0, t) = \pm\infty & t \in (0, T), \\ y(0, x) = y^0(x) & \text{sobre } (-1, 1). \end{cases} \quad (23)$$

Su existencia está asegurada sólo si g es superlineal, pese a que, curiosamente, la solución del problema de control óptimo (22) posea siempre una solución v_k para toda función continua no-decreciente g , independientemente de si g es superlineal o no.

Nuestros resultados numéricos muestran que, para k fijo, el coste «minimal» $J_k(v)$ (y la norma del control óptimo v_k) para una función superlineal g se hacen mucho mayores si (19) no es satisfecha. Para más detalles véase Díaz y Ramos [16].

Numerosas cuestiones matemáticas para esta clase sencilla de modelos esperan aún respuesta, pero el reto más ilusionador es el de analizar si los resultados anteriormente expuestos siguen siendo válidos para modelos más complejos.

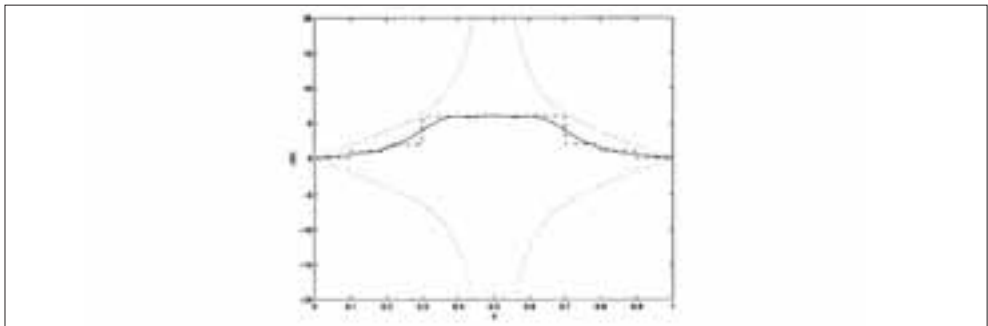


Figura 4. Estado deseado y funciones de obstrucción.

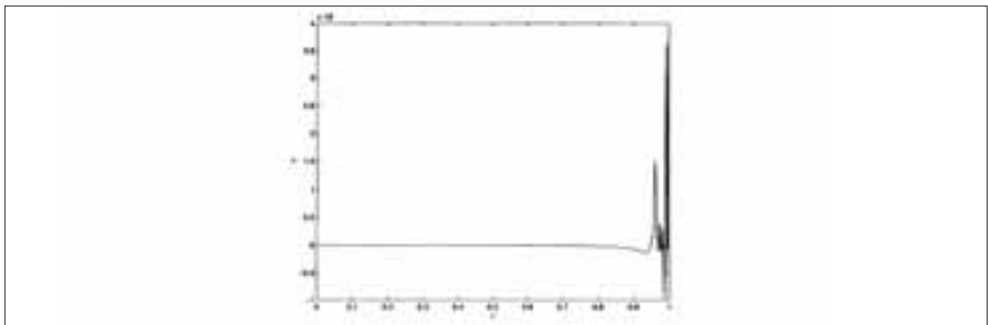


Figura 5. Control óptimo para altos valores de k .

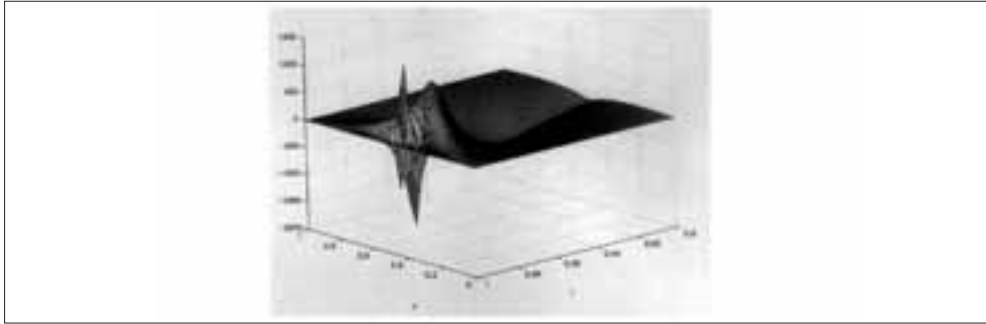


Figura 6. Aproximación del estado del sistema correspondiente al control óptimo.

REFERENCIAS

- [1] Arcoya, D., Díaz, J. I. y Tello, L.: «S-Shaped bifurcation branch in a quasilinear multivalued model arising in Climatology», *Journal of Differential Equations*, 149, 1998, pág. 215-225.
- [2] Bermejo, R.: «Numerical solution to a two-dimensional diffusive climate model», *Modelado de Sistemas en Oceanografía, Climatología y Ciencias Medio-ambientales: Aspectos Matemáticos y Numéricos*, A. Valle y C. Parés (eds.), Universidad de Málaga (1994), pág. 15-30.
- [3] Bermejo, R., Díaz J. I. y Tello, L.: Artículo en preparación (véase el Apéndice A de [43]).
- [4] Bermejo, R., Díaz, J. I. y Tello, L.: Artículo en preparación.
- [5] Brezis, H.: *Operateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert*, North-Holland, Amsterdam, 1973.
- [6] Budyko, M. I.: «The effects of solar radiation variations on the climate of the Earth», *Tellus*, 21, 1969, pág. 611-619.
- [7] Díaz, G. y Díaz, J. I.: «On a nonlinear discontinuous stochastic parabolic equation arising in Climatology». Artículo en preparación.
- [8] Díaz, J. I.: «Sur la contrôlabilité approchée des inéquations variationnelles et d'autres problèmes paraboliques non linéaires», *Comptes Rendus de l'Académie des Sciences de Paris*, 312, Série I, 1991, pág. 519-522.
- [9] Díaz, J. I.: «Mathematical analysis of some diffusive energy balance climate models», *Mathematics, Climate and Environment*, J. I. Díaz y J. L. Lions (eds.), Masson, París, 1993, pág. 28-56.

- [10] Díaz, J. I.: «On the controllability of some simple climate models», *Environment, Economics and their Mathematical Models*, J. I. Díaz, J. L. Lions (eds.), Masson, París, 1994, pág. 29-44.
- [11] Díaz, J. I.: «Controllability and obstruction for some non linear parabolic problems in Climatology», *Modelado de Sistemas en Oceanografía, Climatología y Ciencias Medio-ambientales* (C. Parés y A. Valle, eds.), Universidad de Málaga, 1994, pág. 43-58.
- [12] Díaz, J. I.: «On the Mathematical Treatment of Energy Balance Climate Models», *The Mathematics of Models in Climatology and Environment* (J. I. Díaz, ed.), ASI NATO Global Change Series I, n.º 48, Springer-Verlag, Heidelberg, 1996, pág. 217-252.
- [13] Díaz, J. I.: «On a free boundary problem arising in Climatology», *Free Boundary Problems: Theory and Applications*, (N. Kenmochi, ed.), Vol. II, Gakkotosho Co., Tokyo, 2000, pág. 92-109.
- [14] Díaz, J. I., Hernández, J. y Tello, L.: «On the multiplicity of equilibrium solutions to a nonlinear diffusion equation on a manifold arising in Climatology», *J. Math. An. Appl.*, 216, 1997, pág. 593-613.
- [15] Díaz, J. I. y Lions, J. L.: *Matemáticas, superordenadores y control para el planeta Tierra*, Editorial Complutense, Madrid, 2001.
- [16] Díaz, J. I. y Ramos, A. M.: «Positive and negative approximate controllability results for semilinear parabolic equations», *Revista de la Real Academia de Ciencias Exactas, Físicas y Nat. de Madrid*, Tomo LXXXIX, 1995, pág. 11-30.
- [17] Díaz, J. I. y Ramos, A. M.: «Numerical experiences regarding the localized control of nonlinear parabolic problems», CD-Rom *Proceedings of the European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2000)*, Barcelona, 11-14 septiembre 2000.
- [18] Díaz, J. I. y Tello, L.: A nonlinear parabolic problem on a Riemannian manifold without boundary arising in Climatology, *Collect. Math.*, 50, 1999, pág. 19-51.
- [19] Fabre, C., Puel, J. P. y Zuazua, E.: Approximate controlability of the semilinear heat equations, *Proc. Royal Soc. Edinburgh, Sect. A.*, 125, 1995, pág. 31-61.
- [20] Fernández, L. A. y Zuazua, E.: *Approximate Controllability of the Semilinear Heat Equation via Optimal Control*, preprint de la Universidad de Cantabria, 1996.
- [21] Filippov, A. F.: *Differential equations with discontinuous right hand sides*, Kluwer Academic Publishers, 1988.
- [22] Ghil, M. y Childress, S.: *Topics in Geophysical Fluid Dynamics: Atmospheric, Dynamics, Dynamo Theory and Climate Dynamics*, Springer-Verlag, Nueva York, 1987.
- [23] Held, I. M. y Suarez, M. J.: «Simple albedo feedback models of the ice caps», *Tellus*, 36, 1974.

- [24] Henry, J.: *Étude de la contrôlabilité de certaines équations paraboliques*, Thèse d'Etat, Université de Paris VI, 1978.
- [25] Hetzer, G.: «S-Shapedness for Energy Balance Climate Models of Sellers-Type», *The Mathematics of Models for Climatology and Environment* (J. I. Díaz, ed.) NATO ASI Series, Springer-Verlag, 1997, pág. 253-288.
- [26] Hetzer, G., Jaraus, H. y Mackens, W. A.: «Multiparameter sensitivity analysis of a 2D diffusive climate model», *Impact and Computing in Science and Engineering*, 1, 1989, pág. 327-393.
- [27] Kiehl, J. T.: «Atmospheric general circulation modeling», *Climate System Modeling*, K. E. Trenberth (ed.), Cambridge University Press, 1992, pág. 319-370.
- [28] Krasnoselskii, M. A. y Pokrovskii, A. V.: *Systems with Hysteresis*, Springer-Verlag, Berlin, 1989.
- [29] Legendre, A.: «Recherches sur l'attraction des spheroides», *Mem. des sav. étrangers*, 1785, pág. 411-434.
- [30] Lin, R. Q. y North, G. R.: «A study of abrupt climate change in a simple nonlinear climate model», *Climate Dynamics*, 4, 1990, pág. 253-261.
- [31] Lions, J. L.: *Contrôle Optimal des Systèmes Gouvernés par les Équations aux Dérivées Partielles*, Dunod, París, 1968.
- [32] Lions, J. L.: «Remarques sur la contrôlabilité approchée», *Actas de las jornadas Hispano-Francesas sobre control de sistemas distribuidos*, Univ. de Málaga, 1991, pág. 43-58.
- [33] Lions, J. L.: «Exact controllability for distributed systems. Some trends and some problems», *Applied and Industrial Mathematics*, R. Sigler (ed.), Kluwer, 1991, pág. 59-84.
- [34] Mengel, J. G., Short, D. A. y North, G. R.: «Seasonal snowline instability in an energy balance model», *Climate Dynamics*, 2, 1988, pág. 127-131.
- [35] von Neumann, J.: *Can we Survive Technology?* Nature, 1955. (También en *John von Neumann: Collected Works*, Vol. VI, Pergamon, 1966.)
- [36] Nirenberg, L.: «Variational and topological methods in nonlinear problems», *Bull. Amer. Math. Soc.*, 4, 1981, pág. 267-302.
- [37] North, G.R.: «Introduction to simple climate model», *Mathematics, Climate and Environment*, (J. I. Díaz y J. L. Lions, eds.), Masson, París, 1993, pág. 139-159.
- [38] Schneider, S. H.: «Introduction to climate models». *Climate System Modelling*, (K. E. Trenberth, ed.), Cambridge Univ. Press, Cambridge, 1992, pág. 3-26.
- [39] Sellers, W. D.: «A global climatic model based on the energy balance of the earth-atmosphere system», *J. Appl. Meteorol.*, 8, 1969, 392-400.
- [40] Simmons, G. F.: *Ecuaciones diferenciales con aplicaciones y notas históricas*, McGraw-Hill, 1993.
- [41] Stone, P. H.: «A simplified radiative-dynamical model for the static stability of rotating atmospheres», *J. Atmos. Sci.*, 29, 1972, pág. 405-418.

- [42] Tello, L.: *Tratamiento matemático de algunos modelos no lineales en Climatología*, Tesis Doctoral, Universidad Complutense de Madrid, 1996.
- [43] Thom, R.: *Structural Stability and Morphogenesis*, Benjamin, 1975.
- [44] Xu, X.: «Existence and regularity theorems for a free boundary problem governing a simple climate model», *Applicable Anal.*, 42, 1991, pág. 33-59.
- [45] Zeidler, E.: *Functional Analysis and Applications*, Vols. I-V, Springer-Verlag, Berlín, 1988.

Captura-recaptura: un mètode per calcular el nombre de consumidors de drogues

Pere Puig

Universitat Autònoma de Barcelona

M. Teresa Brugal

Institut Municipal de la Salut, Barcelona

L'ESTIMADOR DE LINCOLN-PETERSEN

Per il·lustrar en què consisteixen els mètodes de captura i recaptura començarem amb un exemple. Suposem que estem interessats en conèixer el nombre de peixos d'una certa espècie que hi ha en un estany. Òbviament no volem utilitzar cap mètode agressiu que pugui perjudicar la vida a l'estany o el seu medi ambient. Un possible sistema per estimar aquest nombre de peixos pot ser el següent:

1. Capturem un nombre determinat de peixos, els marquem i els deixem anar una altra vegada a l'estany. A aquesta quantitat de peixos que hem capturat inicialment li direm n_1 .

2. Passat un temps prudencial per tal que els peixos marcats es barregin amb els altres, tornem a fer-ne una segona captura. Anomenarem n_2 a aquest nombre de peixos que tenim ara. Amb aquesta segona captura trobarem que uns peixos ja estan marcats (recapturats) i d'altres no. Al nombre de peixos que ja estan marcats el denotarem com m .

Aquesta informació ens permet calcular l'abundància de la població de manera aproximada, a partir del següent raonament intuïtiu:

Si N és el nombre total de peixos que hi ha a l'estany, es verifica que

$$\frac{n_1}{N} \approx \frac{m}{n_2}$$

i, per tant,

$$N \approx \frac{n_1 n_2}{m}.$$

A partir d'aquest raonament tan «naïve» obtenim el que es coneix com l'estimador de Lincoln-Petersen de l'abundància de la població:

$$\hat{N} = \frac{n_1 n_2}{m}.$$

El nom d'aquest estimador prové de dos dels seus precursors, Frederick C. Lincoln i Carl G. J. Petersen. El primer d'ells va néixer a Denver, Colorado, el 1892, i va ser un dels principals desenvolupadors dels mètodes de registre, anellatge i seguiment d'aus migratòries als E.U.A.



G. J. Petersen



F. C. Lincoln



P. S. Laplace

Petersen va néixer a Dinamarca el 1860 i va estudiar l'abundància i migracions de poblacions de peixos. El mètode d'estimació que porta el seu nom va ser publicat el 1896.

Tot i que aquest estimador ha passat a la història de l'Estadística amb el nom d'aquests dos relativament recents precursors, de fet va ser Laplace el 1783 qui va utilitzar per primera vegada aquest mètode per estimar el nombre d'habitants de França. Podríem descriure el que va fer de la manera següent:

- Un cens incomplet fet per les autoritats de l'època va ser la primera captura. El nombre total d'individus d'aquest cens és la quantitat n_1 .
- Una enquesta feta a nivell de tot el país va ser la segona captura. El nombre total d'individus enquestats és la quantitat n_2 . Els individus que van aparèixer tant en el cens com en l'enquesta configuren la quantitat m .

Laplace, d'aquesta forma, analitza per primera vegada la informació procedent del creuament de dues llistes, que és la manera com actualment es fa servir la metodologia captura-recaptura en el món de la Medicina i de les Ciències Socials.

LA CORRECCIÓ DE CHAPMAN. ESTIMACIÓ PER INTERVAL

L'estudi d'aquests estimadors es basa en el fet que, fixats n_1 i n_2 , el nombre de peixos de la segona captura m segueix una distribució hipergeomètrica. És a dir, és una variable aleatòria amb funció de probabilitat

$$P(m = k) = \frac{\binom{n_1}{k} \binom{N - n_1}{n_2 - k}}{\binom{N}{n_2}}.$$

Els valors que pot prendre m es troben a l'interval $[\max(0, n_2 + n_1 - N), \min(n_1, n_2)]$. Feta una observació d'aquesta variable, l'estimador de màxima versemblança del paràmetre N és precisament l'estimador de Lincoln-Petersen. Aquest estimador, malgrat la seva senzillesa, té bastant biaix i no té sentit quan m val zero. Per aquesta raó, Chapman, el 1951, va introduir-hi una petita modificació i va proposar l'estimador

$$\hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)}{(m + 1)} - 1$$

que en redueix notablement el biaix i està definit per $m = 0$. Els detalls d'això, juntament amb altres propietats de la distribució hipergeomètrica, es poden trobar a Johnson i Kotz (1976). La variància de l'estimador de Chapman es pot calcular aproximadament mitjançant l'expressió

$$V(\hat{N}_c) = \frac{(n_1 + 1)(n_2 + 1)(n_1 - m)(n_2 - m)}{(m + 2)(m + 1)^2}$$

i, tenint en compte que els estimadors de màxima versemblança segueixen asimptòticament una distribució normal, podem calcular-ne intervals de confiança aproximats a partir de l'expressió

$$N = \hat{N}_c \pm z \sqrt{V(\hat{N}_c)}$$

on z és el percentil de la distribució normal estàndard corresponent al nivell de confiança triat ($z = 1,96$ si el nivell de confiança és del 95%).

Exemple (àrea de Casale Monferrato, 1988)

Tenim una llista de pacients de diabetis de centres hospitalaris públics i privats de la regió de Casale Monferrato. El total de pacients és 452. D'altra banda també tenim una llista computeritzada de prescripcions d'insulina amb un total de 1135. El nombre d'individus coincidents en ambdues llistes és de 249. Volem fer una estimació de la població de diabètics d'aquesta àrea. Per aquest exemple, $n_1 = 452$, $n_2 = 1135$, $m = 249$.

Per tant, podem fer aquesta estimació i calcular un interval de confiança del 95% de la manera següent:

$$\hat{N}_c = \frac{(452 + 1)(1135 + 1)}{(249 + 1)} - 1 = 2057$$

$$V(\hat{N}_c) = 5900.$$

Finalment, obtenim:

$$N = 2057 \pm 1,96 \sqrt{5900} = 2057 \pm 151$$

on, com podem observar, la precisió és molt bona.

CONDICIONS BÀSIQUES PER APLICAR EL MODEL

No sempre les situacions reals permetran que l'estimador de Lincoln-Petersen (o el de Chapman) ens doni una bona estimació de N . Per això és necessari assegurar-nos que es compleixen certes condicions, que descriurem a continuació:

1. La població ha de ser tancada. Vol dir que N és constant en el període de captura-recaptura. Per tant, no poden haver-hi morts, naixements o migracions en aquest període.

2. Tots els individus han de tenir la mateixa probabilitat de ser capturats a la primera captura. No poden haver-hi diferències degudes, per exemple, a l'edat o al sexe. Quan es produeix aquesta anomalia, cas bastant freqüent en aplicacions a grups humans, es pot corregir si es fan anàlisis estratificades, és a dir, fent els càlculs per grups d'edat, sexe, classe social, etc.

3. El marcatge no ha d'afectar la probabilitat de captura de l'individu.

$\binom{N}{n_2}$ 4. La segona captura ha de ser una mostra aleatòria simple, és a dir, cadascuna de les possibles mostres ha de tenir la mateixa probabilitat de ser agafada.

5. Els individus no han de perdre les marques entre la primera i la segona captura. En poblacions humanes, cada individu ha de ser perfectament identificable i distingible d'una llista a l'altra.

A vegades l'interval de confiança que podem calcular pot resultar massa gran per als nostres propòsits. Hem de tenir present que la precisió de l'estimació es pot augmentar si s'agafen n_1 i n_2 més grans.

Una altra opció és tenir present que el procés de captura i marcatge es pot repetir

unes quantes vegades més. A cada pas es verificaria si els individus capturats estan marcats i, en cas contrari, es marcarien abans de deixar-los anar. Aquesta manera de procedir dóna lloc a l'anomenat estimador de Schnabel. Els detalls d'aquest estimador i d'altres procediments es poden trobar al llibre de Seber (1982).

MODELS DE CREUAMENT AMB DIVERSES LLISTES

En els problemes pràctics de càlcul de l'abundància de poblacions humanes, freqüentment hem de tenir en compte la informació procedent de més de dues llistes. A més, l'heterogeneïtat dels individus sovint ens obliga a considerar estratificacions, les quals s'han d'estudiar amb compte per no tenir més estrats o grups dels estrictament necessaris.

Un altre problema és la dependència entre llistes, és a dir, la situació en la qual el fet que un individu estigui en una llista afecti la probabilitat que estigui en una altra. Per tal d'afrontar aquesta situació, una possible solució és utilitzar la metodologia dels models log-lineals. Per tenir una idea intuïtiva d'en què es basen aquests models, ens limitarem, de moment, a un cas en què tenim dues llistes. Observem que la informació procedent de dues llistes la podem representar en una taula de contingència:

		Llista 2	
		Sí	No
Llista 1	Sí	n_{11}	n_{10}
	No	n_{01}	n_{00}

La quantitat n_{11} representa el nombre d'individus que coincideixen en ambdues llistes, n_{10} és el nombre d'individus que només estan a la llista 1 i n_{01} els que només estan a la llista 2. La quantitat n_{00} no és observable i per tant estariem davant d'una taula de contingència amb dades perdudes o «missing». La grandària de la població vindria donada per $N = n_{11} + n_{10} + n_{01} + n_{00}$. Aleshores l'estimació de N es redueix a l'estimació de n_{00} . Per exemple, en el cas dels pacients de diabetis tindríem la taula següent:

		Prescripcions Insulina	
		Sí	No
Centres Hospitalaris	Sí	249	203
	No	886	n_{00}

La grandària de la població seria $N = 1338 + n_{00}$. L'estimador de Chapman que hem utilitzat abans ens permet també estimar n_{00} .

La idea dels models log-lineals es basa en suposar que el logaritme dels valors esperats a cada cel·la de la taula és una funció lineal de les variables indicadores de les llistes, és a dir,

$$\log(E(n_{ij})) = a + b L_1 + c L_2 + d L_1 L_2$$

on L_1 i L_2 són les variables indicadores de cadascuna de les llistes i a, b, c, d són paràmetres que cal estimar. Donada aquesta representació, tindríem que

$$\begin{aligned} \log(E(n_{11})) &= a + b + c + d \\ \log(E(n_{10})) &= a + b \\ \log(E(n_{01})) &= a + c \\ \log(E(n_{00})) &= a. \end{aligned}$$

El paràmetre més interessant és a perquè ens estima n_{00} . El coeficient d és el que habitualment s'anomena «terme d'interacció» i mesura la dependència existent entre ambdues llistes. Malauradament aquest model no és estimable perquè hi ha massa paràmetres (tenim 4 equacions i 5 incògnites). Cal eliminar-hi un paràmetre i, per tant, quan tenim dues llistes només s'utilitza el model en què se suposa independència, és a dir,

$$\log(E(n_{ij})) = a + b L_1 + c L_2.$$

Per estimar els paràmetres es treballa com si les observacions de les cel·les seguisin una distribució de Poisson, on el logaritme de les seves esperances satisfà la relació esmentada. L'estimador del paràmetre d'interès a és el mateix que el que obtindríem si féssim servir l'estimador de Lincoln-Petersen per calcular la grandària poblacional.

Exemple: model amb tres llistes

El model log-lineal més general que podem considerar en aquest cas és

$$\log(E(n_{ij})) = a + b L_1 + c L_2 + d L_3 + e L_1 L_2 + f L_1 L_3 + g L_2 L_3$$

on L_1 , L_2 i L_3 són les variables indicadores de cadascuna de les llistes. Com abans, el paràmetre d'interès és a . Observem que no podem incloure-hi una interacció d'ordre 3 ($L_1 L_2 L_3$) perquè el model estaria sobre-parametritzat.

Com a exemple hipotètic, suposem que volem conèixer la quantitat d'atacs de gossos que hi ha hagut en una determinada ciutat en un cert període de temps. Per a això disposem de tres llistes de registres tal i com ho podem observar a continuació:

C. Animal	Hospital	Policia	N. atacs
1	1	1	1
0	1	1	7
1	0	1	15
0	0	1	326
1	1	0	27
0	1	0	323
1	0	0	91
0	0	0	?

Notem com, per exemple, només un atac apareix registrat en els tres registres alhora i 326 atacs només han estat registrats per la policia. Aquests 0 i 1 serien els valors de les variables indicadores L_1 , L_2 i L_3 .

Per estimar els paràmetres del model i eliminar els que són superflus, cal utilitzar un paquet estadístic adient com SAS, GLIM, etc. Més endavant veurem un exemple de model amb tres llistes en l'àmbit del consum de drogues.

APLICACIONS A LA MEDICINA

Des de l'inici de la història de la Medicina, un dels principals objectius d'aquesta disciplina va ser conèixer la grandària i les característiques d'una determinada malaltia. Per dur a terme aquest objectiu es va desenvolupar el mètode epidemiològic, que consistia a respondre preguntes sobre la causa de la malaltia i la mesura dels problemes de salut, elaborar possibles explicacions i verificant de forma empírica les hipòtesis realitzades. Amb l'arribada de l'estadística al segle XVIII, aquesta disciplina arribà al seu màxim desenvolupament.

Per mesurar la freqüència (nombre de casos) d'una determinada malaltia, s'utilitzen els conceptes d'incidència i prevalença. La incidència és el nombre de casos nous d'una determinada malaltia en un temps determinat i la prevalença són tots els casos existents en un temps determinat.

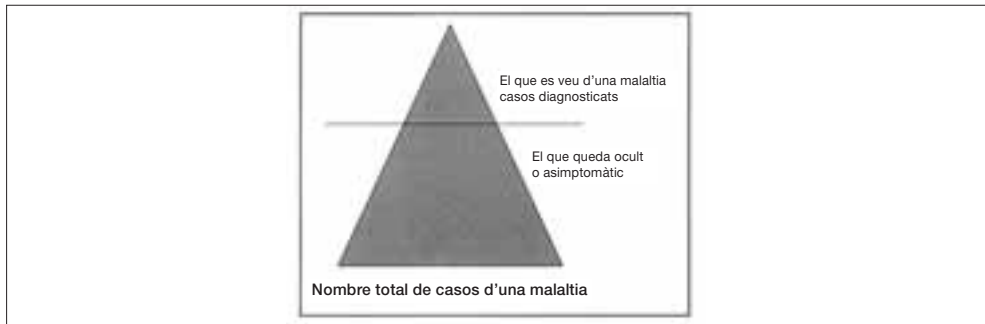


Figura 1. Representació gràfica de l'iceberg de la malaltia.

Una de les principals tècniques emprades per estimar la incidència i/o prevalença d'una malaltia és la tècnica de captura-recaptura.

Les àrees de la salut on més s'ha utilitzat aquesta tècnica han estat: la monitorització del càncer, malalties infeccioses, consum de drogues, accidents de trànsit, defectes congènits, diabetis, sense sostre, prostitució i immigrants.

Tal i com s'ha descrit abans, la tècnica consisteix a estimar el nombre d'afectats d'una determinada malaltia a partir del solapament de diferents registres. És a dir, els individus que apareixen en un registre es consideren una captura i, d'aquests, els que tornen a aparèixer en una altra llista o registre seran els recapturats (duplicats).

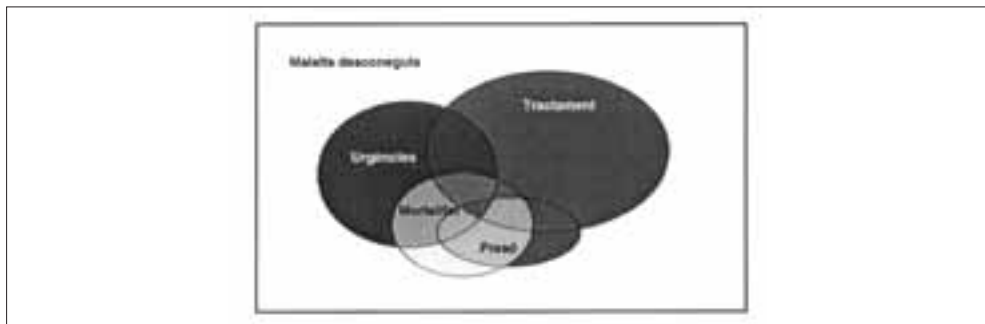


Figura 2. Representació gràfica del solapament de diferents registres.

COM ES FA A LA PRÀCTICA?

En primer lloc, cal identificar diferents llistes o registres d'individus on estigui disponible la informació que busquem. Normalment s'utilitzen registres socio-sanitaris ja existents, que recullen de forma sistemàtica informació sobre la malaltia, els estils de

vida, etc. Normalment aquests solen estar a l'abast de forma informatitzada. Els registres més utilitzats són els procedents de: centres de salut, hospitals, escoles, registres laborals, registres específics de malalties, etc.

El segon pas que cal fer és mirar si les mostres escollides compleixen les assumpcions bàsiques explicades amb anterioritat. És a dir, cal tenir present que han de ser poblacions tancades; per tant, els registres de mortalitat no són útils ja que vulneren aquest principi. També s'ha de tenir en compte que les llistes han de ser independents unes de les altres, tot i que aquesta assumpció pot ser obviada amb els models log-lineals, ja que es poden modelar les possibles dependències entre mostres. I finalment cal recordar que tots els membres de la població han de tenir la mateixa probabilitat de ser capturats. Aquest tercer supòsit pot estar influenciat per factors com: la diferent accessibilitat als serveis socio-sanitaris, la territorialització del fenomen estudiat, etc. Per tant, potser caldrà estratificar per tal d'obtenir l'homogeneïtat requerida.

Una vegada segurs que es compleixen els diferents supòsits, hem de crear un identificador únic amb el qual aparellem els diferents individus de les distintes llistes. Normalment s'utilitza com a identificador les dues primeres lletres d'ambdós cognoms, la data de naixement i el sexe. Aquest identificador pot variar d'un estudi a un altre i es poden fer diferents combinacions amb el nom, el cognom, el DNI, la data de naixement, etc.

Finalment, una vegada aparellades les llistes i coneguts el nombre de casos que apareixen en una o més llistes, s'utilitza l'estimador de Chapman explicat anteriorment o s'ajusta un model log-lineal per estimar el nombre de casos desconeguts.

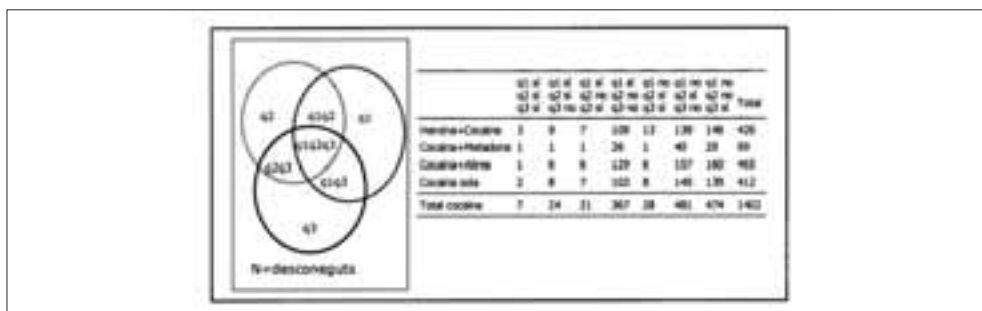


Figura 3. Exemple d'aparellament entre llistes.

QUINS PROBLEMES PODEM TROBAR-HI?

El primer problema amb què ens solem trobar és la pobra qualitat dels registres, que fa que sigui difícil crear un identificador únic igual en tots els registres. Això supo-

sa que, actualment, força sovint una mateixa persona tingui identificadors diferents entre un registre i un altre. Per exemple: hospitals que posen de forma sistemàtica com a data de naixement el 15 de juny i l'any de naixement del pacient, etc. Un altre problema són els noms estrangers, ja que poden estar escrits de diferents formes en distints registres i, per tant, els identificadors seran diferents tot i que en realitat l'individu és el mateix.

Amb la nova llei de protecció de dades (Ley Orgánica 15/1999, sobre Protecció de Dades de Caràcter Personal, LOPD), en el camp socio-sanitari cada vegada serà més difícil fer encreuament de registres i, per tant, utilitzar aquest tipus de tècnica per fer estimacions de la prevalença.

El segon problema que té aquesta tècnica són els biaixos que podem fer a l'hora d'estimar prevalences, a causa de possibles captures diferencials. És a dir, poblacions difícils de ser capturades per les seves característiques; per exemple, una població benestant que no utilitza els serveis sanitaris públics, individus que no poden ser entrevistats perquè no són mai a casa, etc.

Finalment, el tercer problema del qual hem de parlar és la dependència entre les diferents fonts d'informació (registres). Aquest és un problema solucionable, ja que es poden considerar dins del model de regressió log-lineal les possibles dependències entre les mostres, però que cal estudiar per no introduir en el model més paràmetres dels realment necessaris.

EXEMPLES: ESTIMACIÓ DE LA PREVALENCIA DEL CONSUM DE DROGUES IL·LEGALS

En la dècada dels vuitanta i principis dels noranta, el consum d'heroïna va provocar una seriosa preocupació socio-sanitària, sobretot pels problemes sanitaris associats a aquest consum (sida, tuberculosi, etc.) i pel comportament antisocial que s'hi relaciona. Per aquests motius, i amb la finalitat de conèixer la magnitud, característiques i distribució territorial del consum regular d'heroïna a la ciutat de Barcelona, ens vàrem plantejar a l'any 1993 realitzar una estimació de la prevalença del nombre de consumidors d'heroïna mitjançant el mètode de captura-recaptura.

La metodologia emprada va ésser la d'utilitzar les tres fonts d'informació de recollida sistemàtica disponibles: inicis de tractament per heroïna, urgències en consumidors d'heroïna i ingressats a la presó que referien consumir heroïna. La definició de cas fou homogènia per als tres indicadors: persones que referien un consum diari o quasi diari d'heroïna.

Un cop creat l'identificador, els casos es varen aparellar segons el grau de solapa-

ment entre les mostres i això es va fer segons la presència o absència del cas en cada una de les fonts (Taula 1). Per minimitzar l'heterogeneïtat de la captura es va estratificar per edat i sexe.

Taula 1. Distribució dels pacients segons edat, sexe i presència en els registres. Barcelona 1993

Població	Urgèn. Sí	Urgèn. Sí	Urgèn. Sí	Urgèn. Sí	Urgèn. No	Urgèn. No	Urgèn. No
	Tract. Sí	Tract. Sí	Tract. No	Tract. No	Tract. Sí	Tract. Sí	Tract. No
	Presó Sí	Presó No	Presó Sí	Presó No	Presó Sí	Presó No	Presó Sí
Total BCN	19	164	69	781*	37	857	300
Homes joves (15-29 anys)	6	64	37	373	15	363	165
Dones joves (15-29 anys)	7	46	17	139	8	154	32
Homes grans (30-44 anys)	6	46	13	208	10	255	81
Dones grans (30-44 anys)	0	8	4	59	4	85	22

Després, per cada estrat, es va ajustar un model de regressió log-lineal amb les interaccions necessàries per a un bon ajust. Per realitzar aquest model de captura-recaptura amb múltiples fonts, es va utilitzar el programa GLIM. Es considerava que el model ajustava bé quan l'estadístic χ^2 del test de la raó de versemblança no era estadísticament significatiu, en contrastar respecte el model següent jeràrquicament més gran. L'interval de confiança a 95% va ser calculat segons el mètode de la raó de versemblança. A l'apèndix es pot veure com a exemple un programa de SAS per ajustar el model per al cas de la població total.

Els resultats obtinguts es poden veure a la Taula 2. Veiem que la taxa que vàrem estimar per a tota la ciutat va ser de 12,9 (10,1-17,2) per 1.000 habitants de 15-44 anys; és a dir, per a tota la ciutat de Barcelona es varen estimar un total de 9.176 addictes actius a heroïna.

Taula 2. Distribució del nombre estimat d'usuaris d'heroïna segons edat i sexe. Barcelona 1993.

Població	Nombre de coneguts	Nombre de desconeguts (IC 95%)	Població (IC 95%)	Taxes per 1.000 hbs. (IC 95%)
Total BCN	2.227	6.949 (4.961-9.995)	9.176 (7.188-12.222)	12,9 (10,1-17,2)
Total homes (15-44 anys)	1.641	6.081 (4.069-9.475)	7.722 (5.710-11.116)	21,9 (16,2-31,5)
Total dones (15-44 anys)	586	897 (669-1.209)	1.483 (1.255-1.795)	4,2 (3,5-5,0)
Homes joves (15-29 anys)	1.021	3.993 (2.402-7.161)	5.014 (3.423-8.182)	25,9 (17,7-42,2)
Dones joves (15-29 anys)	403	475 (337-669)	878 (740-1.072)	4,7 (3,9-5,7)
Homes grans (30-44 anys)	619	2.066 (1.105-4.303)	2.685 (1.724-4.922)	16,9 (10,8-30,9)
Dones grans (30-44 anys)	182	546 (314-996)	728 (496-1.178)	4,3 (2,9-6,9)

Les taxes no es distribuïen de forma uniforme entre ambdós sexes, ni en els diferents grups d'edat. El grup més afectat va ésser el dels homes i, dins d'aquests, el grup d'homes joves.

Les estimacions fetes per àrea de residència (distribució territorial) es poden veure a la Figura 4. Aquí es pot veure que les zones més afectades eren la zona de Ciutat Vella, la portuària i la nord-est de la ciutat. Totes elles corresponen a les zones de Barcelona amb els indicadors socio-econòmics més baixos.



Figura 4. Distribució de les taxes estimades per barri de residència. Barcelona 1993.

Un segon exemple de captura-recaptura en la investigació sobre drogues és el desenvolupat l'any 1999 a Barcelona per conèixer les característiques i magnitud del consum problemàtic de cocaïna.



Figura 5. Anunci de gotes de cocaïna contra el dolor dental (1885).

A Barcelona, en el darrer trienni es va posar de manifest un clar increment dels problemes relacionats amb la cocaïna i es va observar que tant les urgències relacionades amb el consum de drogues com els inicis de tractament per aquesta substància o les morts per sobredosi presentaven un constant increment dels problemes relacionats amb la cocaïna.

En aquest cas, l'indicador analitzat fou el d'urgències relacionades amb el consum de cocaïna durant l'any 1999. En utilitzar una sola llista, vàrem dividir-la en quadrimestres i per minimitzar l'heterogeneïtat de la captura es varen estratificar les mostres segons drogues consumides a banda de la cocaïna.

Taula 3. Distribució dels diferents perfils segons grau de solapament entre mostres. Barcelona 1999.

	q1 sí	q1 sí	q1 sí	q1 sí	q1 no	q1 no	q1 no	
	q2 sí	q2 sí	q2 no	q2 no	q2 sí	q2 sí	q2 no	
	q3 sí	q3 no	q3 sí	q3 no	q3 sí	q3 no	q3 sí	Total
Heroïna + Cocaïna	3	9	7	109	13	139	146	426
Cocaïna + Metadona	1	1	1	26	1	40	29	99

El solapament entre registres el podem veure a la Taula 3. En aquest cas, s'observa que el grau de solapament entre quadrimestres en algunes substàncies és molt petit. Per aquesta raó no es van considerar models jeràrquics com s'havia fet amb l'heroïna. En aquest cas, la tècnica utilitzada va ser la de calcular la població desconeguda i ponderar

els estimadors dels diversos models mitjançant l'AIC i el BIC (AICW i BICW, respectivament), tècnica proposada per Regal i Hook per als casos en què el nombre d'efectius en les diferents caselles és baix.

Amb aquest sistema, es va calcular una taxa estimada de consum problemàtic de cocaïna del 3,1% de la població resident a Barcelona, de 15 a 54 anys. Les estimacions no es varen distribuir de forma uniforme entre els diferents perfils de consumidors: el grup amb més usuaris era el de cocaïna sola, seguit pel de cocaïna més heroïna i finalment pel grup de cocaïna més altres.

HA ESTAT ÚTIL AQUEST TIPUS DE RECERCA?

Sens dubte sí, ja que aquests estudis han permès realitzar una estimació de la prevalença del consum problemàtic d'heroïna i/o cocaïna i caracteritzar aquests consumidors segons territori i perfils diferenciats.

En el camp de la prevenció, de les polítiques sanitàries, de la gestió de recursos i de la intervenció sanitària, aquest tipus de recerca és particularment interessant. Gràcies a ella es pot mapificar i caracteritzar la malaltia i es pot determinar quins són els grups més prioritaris per fer la intervenció o la distribució de recursos.

APÈNDIX

El programa següent, fet en llenguatge SAS, ajusta les dades de la població total de Barcelona per a l'exemple de l'heroïna:

```
data a;  
input PR TR UR n;  
PRTR=PR*TR;PRUR=PR*UR;TRUR=TR*UR;  
cards;  
1 1 1      19  
0 1 1      164  
1 0 1       69  
0 0 1      781  
1 1 0       37  
0 1 0      857  
1 0 0      300  
;  
proc genmod;  
model n=PR TR UR PRUR TRUR/d=poisson;  
estimate 'missing' intercept 1/exp;  
run;
```

El model ajustat no utilitza la interacció entre Presó i Tractament (PR*TR), ja que no resultaria significativa.

El programa SAS calcula els intervals de confiança d'una manera diferent que el GLIM i per això, si fem servir aquest programa, l'interval de confiança per al nombre de desconeguts és lleugerament diferent que el de la Taula 2: 4.906-9.841.

REFERÈNCIES

- [1] Johnson NL, Kotz S. *Discrete Distributions*. John Wiley & Sons. Nova York, 1970.
- [2] Seber, GAF. *The estimation of animal abundance and related parameters*. Griffin, Londres, 1973.
- [3] Bishop YMM, Fienberg SE, Holland PW. *Discrete multivariate analysis: theory and practice*. Cambridge, MA: MIT Press, 1995.
- [4] Hook EB, Regal RR. «Capture-recapture methods in Epidemiology: Methods and limitations». *Epidemiol Rev* 1995; 17:243-64.
- [5] Regal RR, Hook EB. «Goodness-of-fit based confidence intervals for estimates of size of closed population». *Stat Med* 1984; 3:287-91.
- [6] Brugal MT, Domingo-Salvany A, Maguire A, Caylà JA, Villalbí JR, Hartnoll R. «A small area analysis estimating the prevalence of addiction to opioids in Barcelona», 1993. *J Epidemiol Community Health* 1999; 53:488-94.
- [7] Domingo-Salvany A, Hartnoll RL, Maguire A, Brugal MT, Albertín P et al. «Analytical considerations with capture-recapture prevalence estimation: Case studies of estimating opiate use in Barcelona Metropolitan Area». *Am J Epidemiol* 1998; 148:732-40.

Anàlisi no lineal del patró ventilatori

Salvador Benito

Hospital de la Santa Creu i Sant Pau, Barcelona

Pere Caminal

CREB, Universitat Politècnica de Catalunya

Beatriz Giraldo

CREB, Universitat Politècnica de Catalunya

Montserrat Vallverdú

CREB, Universitat Politècnica de Catalunya

1. INTRODUCCIÓ

El pulmó és l'òrgan encarregat de dur a terme l'intercanvi gasós amb l'aire atmosfèric. La seva funció principal és prendre l'oxigen de l'aire per portar-lo a la sang i eliminar-ne l'anhidrid carbònic. Les característiques anatòmiques del pulmó li permeten augmentar i disminuir de volum. Aquesta funció es coneix amb el nom de ventilació. Estrictament, el pulmó és un òrgan passiu a les variacions de volum que li indueixen les estructures de la seva contigüitat. La precisa i peculiar disposició anatòmica i funcional de la pleura, de la graella costal i dels músculs respiratoris permet que es puguin produir les modificacions de la caixa toràctica, que condicionen les variacions de volum del pulmó.

La ventilació permet, de forma inconscient, fer funcions tan precises com parlar, cridar, sospirar i, molt especialment, riure. Cap d'aquestes funcions no es podria realitzar si el pulmó no variés de volum, no fos capaç de distendre's, no permetés l'entrada i sortida de l'aire, i també, i més important, si el pulmó no ventilés no es podria renovar l'aire alveolar, fet imprescindible per a la vida. La ventilació alveolar permet eliminar l'anhidrid carbònic, que arriba a la sang del capil·lar pulmonar i l'enriqueix amb l'oxigen que arriba de l'aire atmosfèric.

La ventilació pulmonar consisteix en el moviment d'un volum de gas des de l'atmosfera a l'interior del pulmó i viceversa. Aquest fet té com a objecte renovar el gas pulmonar. Perquè existeixi moviment d'un fluid s'ha de produir una diferència de pressió entre els dos punts per on es desplaça aquest fluid. Perquè existeixi flux s'ha de generar abans una diferència de pressió. En el cas del sistema respiratori, la contracció dels músculs inspiratoris genera una depressió que es transmet al pulmó gràcies a les característiques i disposició de l'espai pleural. La pressió generada pels músculs, que superen la pressió elàstica del pulmó, serà pressió apta per generar flux. Aquest flux condicionarà

un augment de volum de gas en el pulmó. Un cop acabat l'esforç muscular inspiratori, l'augment de volum produirà un augment de pressió intrapulmonar, per sobre de la pressió atmosfèrica, que invertirà el sentit del flux i completarà el cicle respiratori amb l'expiració. Aquest és el mecanisme físic de forma simplificada, pel qual el pulmó permet l'entrada i sortida de gas.

La regulació d'aquest intercanvi de gasos és possible perquè existeix un conjunt de neurones al cervell, que s'anomena centre respiratori, que regula els mecanismes descrits i són les responsables d'aquest exhaustiu control.

2. PATRÓ VENTILATORI

La freqüència respiratòria és un dels paràmetres més sensibles, encara que inespecífics, del sistema respiratori. La freqüència respiratòria defineix el nombre de cicles ventilatoris que es realitzen en un minut, essent la freqüència normal entre 12-20 per minut. El cicle respiratori es divideix en dues fases: la d'increment de volum pulmonar, o inspiratòria, i la de sortida d'aquest volum, o espiratòria. Entre les dues situacions s'aprecien dues fases sense canvi de volum: la pausa inspiratòria i la pausa espiratòria (Figura 1). Els paràmetres més freqüentment utilitzats que caracteritzen el cicle respiratori en relació al temps són: la duració de la inspiració (T_I), que inclou la pausa inspiratòria; la duració de l'expiració (T_E), que inclou la pausa espiratòria; i la duració del cicle complet ($T_{TOT} = T_I + T_E$).

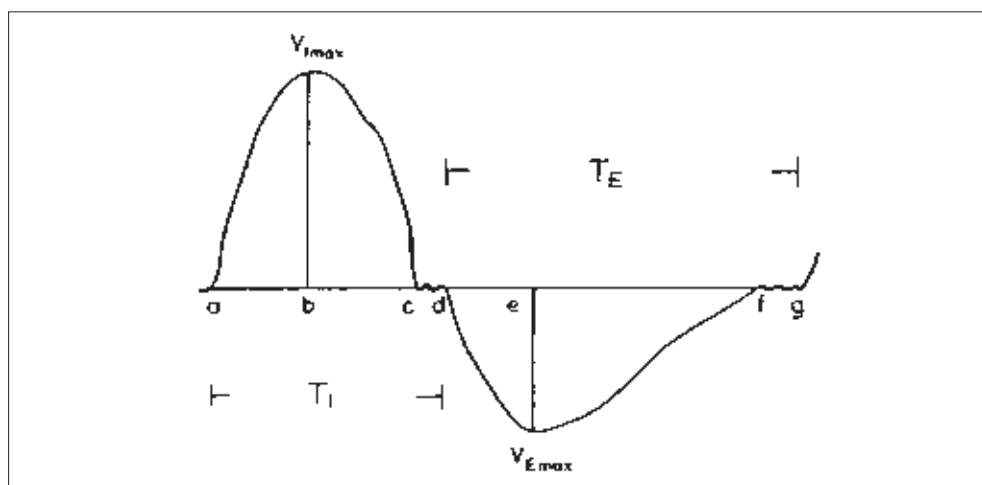


Figura 1. Senyal de volum corresponent a un cicle respiratori, amb indicació de les duracions d'inspiració (T_I) i espiració (T_E).

L'estudi de l'activitat del centre respiratori ha permès dividir-lo funcionalment en un controlador de temps i un controlador de l'impuls inspiratori. Mitjançant unes operacions matemàtiques sobre el càlcul de volum per minut espirat (V_E) es poden separar aquests dos components [1]:

$$\begin{aligned} V_E &= V_T \cdot f \\ f &= 1/T_{TOT} \\ V_E &= V_T \cdot (1/T_{TOT}) = (V_T / T_I) \cdot (T_I / T_{TOT}). \end{aligned}$$

Aquests dos factors, que defineixen bé la intensitat de l'estímul inspiratori o flux mig inspiratori (V_T / T_I) i el temps que consumeix la fase inspiratòria sobre el temps total (T_I / T_{TOT}) tenen el gran inconvenient de la variabilitat intraindividu i interindividu. No se solen donar valors de normalitat, per la gran variació que es produeix respiració a respiració. Aquesta peculiar variació del patró ventilatori de les persones s'ha denominat personalitat ventilatòria [2].

Quan s'estudia el control del sistema respiratori, s'aprecia una gran variabilitat del patró ventilatori analitzat respiració a respiració. Tobin i col. [3], a més de suggerir sistemes no invasius per estudiar el patró respiratori, troben una considerable variabilitat dels volums i de l'impuls inspiratori (V_T / T_I), així com de les relacions dels temps del cicle respiratori (T_I / T_{TOT}). Mostren com per als subjectes joves la variació de V_T / T_I és del 32% i del 42% per als avis, i del 18% i el 23%, respectivament, quan s'estudia T_I / T_{TOT} . Aquesta dispersió apareix tant en l'anàlisi intraindividu com en la interindividu quan s'estudien unes 250 respiracions. En els estudis intraindividu dia a dia es troba una resposta més constant per als valors mitjans del patró ventilatori, essent del 9% per a V_T / T_I i del 3% per a T_I / T_{TOT} . Aquesta petita variabilitat interindividu i intraindividu del patró ventilatori suggereix que és possible construir un marge estret de normalitat que permet detectar anormalitats en un subjecte concret. A més, suggereix que aquesta és la forma adequada d'analitzar el control de la respiració davant de formes més integrades, com la freqüència respiratòria o el volum per minut. En relació a l'anàlisi respiració a respiració, suggereix, per la gran variabilitat, que es realitzin registres molt perllongats per al càlcul, en lloc d'intentar caracteritzar un cicle respiratori «representatiu».

Les variacions temporals del patró ventilatori i la seva anàlisi no lineal són el subjecte de la revisió d'E. N. Bruce [4]. La seva revisió analitza una sèrie de treballs que estudien l'origen de la variabilitat respiratòria via neuronal i els efectes que en la variabilitat tenen els estímuls neuronals i químics, per entendre la fisiologia del control de la respiració. En el seu article recomana estudis de complexitat del patró ventilatori. Hi destaca que, en analitzar d'una forma conjunta i lineal el cicle respiratori respiració a respiració, no s'és conclouent perquè se superposen cicles respiratoris periòdics amb aleatoris correlacionats, amb no lineals deterministes i amb aleatoris. L'anàlisi multiperiòdica de la

variabilitat, concretament l'estudi de fases i els plànols de retorn, podria donar una informació visual útil que expliqués el comportament del patró ventilatori. A més, els estudis de dimensió de capacitat, exponents de Hurst i Lyapunov, entre altres, permetran la quantificació de la dinàmica multidimensional.

En un estudi, Brack i col. [5] suggereixen que els canvis que observen en la variabilitat del patró ventilatori quan es genera una càrrega elàstica a subjectes sans poden expressar un intent del sistema de control per compensar la càrrega i minimitzar la dispnea induïda per aquesta càrrega. En un altre treball [6], el mateix grup afirma que les disrítmies de la respiració es presenten en clínica; els seus mecanismes són desconeguts i és precís, per conèixer-ne la fisiopatologia, trobar els factors responsables de la variabilitat. Troben que l'estímul hipercàpnic en hiperòxia augmenta la fracció no aleatòria de l'activitat variable de la respiració i que aquesta depèn de l'increment que es produeix de CO_2 . En ambdós estudis utilitzen per a l'anàlisi de variabilitat l'autocorrelació i l'anàlisi espectral. Aquestes eines matemàtiques van ser recomanades pel grup de Bruce [7], però no són les que s'han proposat posteriorment [4] per a l'anàlisi de la informació clínica oculta i que ja s'han comentat.

Donaldson [8], mitjançant el càlcul de l'exponent de Lyapunov dels paràmetres respiratoris obtinguts en adults normals, afirma que la trajectòria respiratòria no és aleatòria sinó caòtica, o no lineal, i que això és de gran interès en la interpretació del comportament respiratori. En un model simulat, Sammon [9] dissenya una distribució de representació geomètrica del generador central del patró ventilatori, a fi d'explorar: com és de sensible a l'homogeneïtat del sistema; de què depenen les bifurcacions que apareixen en les formes geomètriques, és a dir, els atractors d'aquestes distribucions geomètriques; i com les variables de temps i d'amplitud modifiquen aquest comportament geomètric del patró respiratori. Aquest model de simulació ho fa en base a dades experimentals. Concretament, en un estudi experimental [10], les rates amb *vagus indemnes* mostren un patró respiratori amb majors graus de llibertat dels que s'aprecia després de la vagotomia, però més interessants són els resultats que hi troben en modificar la capacitat residual funcional, mitjançant modificacions en la pressió traqueal amb CPAP i CNAP. A mesura que disminueix la pressió traqueal, augmenta de forma progressiva la dimensió de correlació del patró respiratori. Aquestes variables, quan s'analitzen de forma complexa, tenen un comportament diferenciable en modificar-s'hi el volum pulmonar.

3. INTERÈS CLÍNIC DE L'ANÀLISI DEL PATRÓ VENTILATORI

La insuficiència respiratòria aguda és la impossibilitat de realitzar l'intercanvi de gasos, de captar oxigen i eliminar anhídrid carbònic. En un percentatge elevat d'aquestes situacions clíniques s'ha de substituir la funció pulmonar de forma artificial mitjan-

çant un ventilador mecànic. El ventilador artificial es connecta a un tub que, introduït a través de la boca, arriba a la tràquea. Aquesta ventilació artificial s'utilitza mentre no millora la malaltia que condiciona la insuficiència respiratòria aguda.

Un dels objectius de la ventilació mecànica (VM) és recuperar, el més aviat possible, la respiració espontània. La retirada de la ventilació suposa una dificultat en un nombre reduït de pacients, pròxim al 30% en una població de pacients medicoquirúrgics. En patologies concretes, com és l'agudització d'una insuficiència respiratòria crònica que requereixi VM, la necessitat d'utilitzar tècniques de suport parcial, que permeten una recuperació progressiva de la respiració, s'aproxima al 80% d'aquests pacients. El pas següent, en aquest grup de pacients que han estat ventilats, és retirar el tub endotraqueal. Aquesta maniobra no està exempta de risc: el més important és la necessitat de precisar una reintubació abans de les primeres 48 hores. Entre el 15 i el 20% del grup de pacients que requereixen una recuperació progressiva de la respiració fracassen en mantenir la respiració espontània i requereixen ésser reintubats [11, 12]. Aquesta reintubació s'associa a un augment de mortalitat d'aquests pacients [13].

S'han proposat diversos paràmetres predictors de l'èxit en l'extubació. La majoria d'ells, amb excepció de l'adequat intercanvi d'oxigen, estan relacionats amb paràmetres de la ventilació. Concretament, s'utilitzen la freqüència respiratòria (f), el volum circulant (V_T) i el volum per minut. També es recomanen quocients entre ells (f / V_T) [14] i paràmetres relacionats amb l'activitat del centre respiratori ($P_{0.1}$) [15]. La utilització clínica d'aquests predictors no ha disminuït la necessitat de reintubació.

S'ha tingut l'oportunitat d'estudiar abans d'ésser extubats [16] alguns paràmetres funcionals respiratoris dels pacients que superen amb èxit l'extubació i de comparar-los amb els dels pacients que no són capaços de mantenir la respiració espontània i requereixen ésser intubats. En aquest estudi es destaca que, si bé en alguns paràmetres es mostraven diferències significatives, és difícil amb aquests valors fer prediccions en clínica. Quan es varen estudiar, els millors indicadors mostraven una bona sensibilitat però una baixa especificitat, amb un poder predictiu i una precisió diagnòstica entre el 60 i el 70%. Tot això fa que, igual que han trobat altres autors, els paràmetres respiratoris analitzats d'aquesta forma són de poca ajuda en la predicció del resultat de l'extubació.

4. ADQUISICIÓ DE SENYALS I PACIENTS ESTUDIATS

S'ha estudiat un grup de 18 pacients amb insuficiència respiratòria aguda que van ser assistits amb ventilació mecànica. Aquests pacients varen estar ingressats a l'Hospital de la Santa Creu i Sant Pau. Cada pacient va ser sotmès a dos nivells diferents

de pressió suport en la ventilació assistida (PSV). Per a cada cas es van adquirir els senyals de volum respiratori mitjançant pletismografia respiratòria per inductància.

La pletismografia respiratòria per inductància [17, 18] és una tècnica per determinar el volum d'aire toràcic i abdominal dels individus. La mesura es fa mitjançant un transductor inductiu. El sensor consta de dues bandes elàstiques amb espines que es col·loquen entorn de l'abdomen i del tòrax, respectivament. Aquesta és una tècnica no invasiva, a diferència de l'espirometria i de la pneumotocografia. Els seus resultats són força precisos en pacients intubats i més quan no és necessària la col·laboració del pacient.

Els senyals de volum respiratori corresponents a cada nivell de PSV es van enregistrar durant 30 minuts, amb una freqüència de mostreig de 250 Hz. Aquests senyals han estat degudament processats per obtenir tant el senyal corresponent al flux respiratori com les sèries temporals de duració del cicle respiratori (T_{TOT}), el temps d'inspiració (T_I) i el volum circulant (V_T). Amb aquestes sèries s'han pogut obtenir també els valors corresponents al flux mig inspiratori (V_T / T_I) i a la fracció inspiratòria del cicle respiratori (T_I / T_{TOT}).

La Taula 1 presenta les característiques clíniques dels pacients analitzats. Per a cada pacient és té la freqüència respiratòria (f) (respiracions/minut) abans d'adquirir el senyal de volum respiratori, la ventilació minut (V_e) (litres/minut), el valor de la pressió positiva al final de l'expiració (PEEP) (cmH_2O) i el mode ventilatori amb el qual ha estat assistit el pacient abans del registre, que pot ser pressió assistida controlada (ACV) o pressió de suport (PSV).

Pacient	f	V_e	PEEP	Mode
1	22	11,5	8	PS20
2	32	15,6	8	PS10
3	24	11,0	7	PS14
4	18	10,0	8	ACV
5	30	11,4	8	PS8
6	22	17,9	6	PS12
7	15	10,0	8	ACV
8	24	12,0	5	PS10
9	26	11,0	8	PS12
10	30	9,6	6	PS12
11	23	11,2	6	PS15
12	26	14,0	10	PS14
13	18	9,5	4	PS8
14	27	10,2	6	PS8
15	28	14,0	8	PS14
16	20	9,0	6	ACV
17	26	12,0	6	PS10
18	26	11,5	6	PS10

Taula 1. Relació de pacients analitzats: f (respiracions/minut) - freqüència respiratòria, V_e (litres/minut) - nivell de ventilació per minut del pacient, PEEP (cmH_2O) - valor de pressió positiva al final de l'expiració i Mode - tipus de ventilació.

Amb la informació clínica i els registres corresponents per a cada pacient es té el comportament del patró ventilatori per a dos nivells diferents de pressió suport. Aquesta informació permetrà analitzar amb diferents tècniques la complexitat del sistema respiratori, segons la variabilitat del seu patró respiratori.

5. MÈTODES

A. Introducció a la dinàmica no lineal

Les tècniques tradicionals de l'anàlisi de senyals en el domini temporal i el freqüencial no són sempre suficients per caracteritzar la complexa dinàmica respiratòria. Diferents treballs publicats presenten l'aplicació dels conceptes de la dinàmica no lineal a l'anàlisi de complexitat de diferents sistemes fisiològics [19-22]. Aquestes tècniques permeten diferenciar els sistemes que provenen de processos estocàstics d'aquells que, tot i no ser predictibles, es pot afirmar que són deterministes.

Les mesures del comportament caòtic dels sistemes, com la dimensió de correlació, mostren que el sistema respiratori té un comportament dinàmic irregular durant els estats de vigília i aquesta estructura dinàmica complexa disminueix durant els estats del son [23]. Altres estudis han mostrat que la dinàmica respiratòria en nens durant el son pot descriure's com un sistema caòtic [24].

En aquest estudi s'ha analitzat la complexitat del patró respiratori a diferents nivells de pressió de suport, utilitzant la dinàmica simbòlica. L'estudi del patró respiratori ofereix la possibilitat d'augmentar el coneixement de les alteracions funcionals provocades per diferents processos, com són la limitació crònica del flux aeri, la insuficiència respiratòria i diferents patologies restrictives. Amb la utilització d'una metodologia de l'anàlisi de la dinàmica no lineal es pretén aportar nous índexs per al pronòstic i tractament terapèutic de les malalties respiratòries, així com també intentar determinar el moment òptim de la desconnexió de la ventilació mecànica.

B. Dinàmica simbòlica

La dinàmica simbòlica es basa en la transformació de les dades a un patró descrit per elements que són uns pocs símbols. En aquest sentit se simplifica l'estudi de la dinàmica en descriure els senyals mitjançant seqüències de símbols. Aquesta descripció dels sistemes produeix una pèrdua de la informació variable mentre que roman la part més robusta i invariant [25].

El primer pas consisteix a trobar una descripció simbòlica adequada de les sèries

temporals [26]. Per tant, per a aquest propòsit s’ha buscat un codi senzill que permeti la descripció dels senyals mitjançant símbols. S’han proposat 4 símbols diferents per codificar les sèries temporals [27]:

$$S_n = \begin{cases} 0, & t_n > (1 + a)\mu \\ 1, & \mu < t_n \leq (1 + a)\mu \\ 2, & (1 - a)\mu < t_n \leq \mu \\ 3, & t_n \leq (1 - a)\mu \end{cases}$$

on μ és el valor mig de la sèrie temporal i a és un paràmetre adimensional que quantifica la desviació típica. A continuació es presenten diferents aspectes del comportament de les paraules construïdes a partir dels símbols S_n .

La primera aproximació consisteix a calcular la freqüència d’ocurrència dels símbols [28]. La distribució d’aquestes freqüències permet diferenciar dinàmiques més uniformes d’aquelles que presenten una més gran complexitat. Per obtenir un contingut ampli de la dinàmica d’un sistema cal considerar paraules de longitud elevada. No obstant això, cal arribar a un compromís entre la longitud de la paraula i la longitud de la sèrie temporal en estudi. El nombre de possibles paraules de tamany l que poden construir-se amb l’alfabet $\{0, 1, 2, 3\}$ és 4^l . Les sèries temporals en estudi tenen aproximadament 500 mostres. Per tot això, el compromís considerat per caracteritzar el contingut dinàmic del sistema i la freqüència d’aparició de les paraules ha estat construir paraules de longitud 3 símbols. Així s’ha obtingut un total de 64 possibles paraules.

A les figures 2a i 2b es presenta la distribució de probabilitat de les 64 paraules d’una sèrie temporal T_{TOT} d’un pacient amb PSV baixa i alta, respectivament, i s’ha observat un comportament diferent en la dinàmica del sistema.

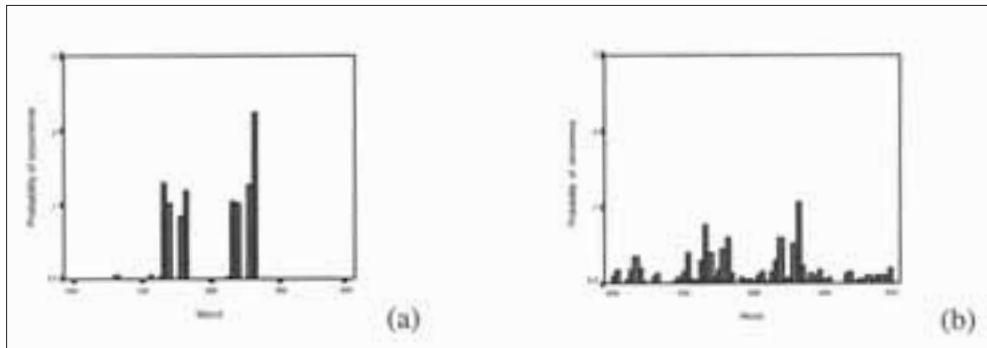


Figura 2. Probabilitat d’ocurrència de cadascuna de les 64 paraules a la sèrie temporal T_{TOT} d’un pacient amb nivell de pressió de suport ventilatòria baixa (a) i alta (b).

Aquesta distribució de probabilitat d'ocurrència de les paraules ofereix nous índexs capaços de caracteritzar la dinàmica del sistema. Aquests índexs s'han obtingut comptabilitzant el nombre de paraules amb una probabilitat d'ocurrència superior a un llindar. Els llindars considerats han estat 1%, 2% i 5%. Altres índexs considerats en aquest treball per mesurar la complexitat han estat la probabilitat d'ocurrència de les paraules més significatives i el nombre de paraules prohibides. S'ha definit com a paraula significativa aquella amb probabilitat d'ocurrència estadísticament significativa en comparar els dos grups amb nivells PSV diferents. S'han definit com paraules prohibides aquelles amb una probabilitat d'ocurrència menor que un llindar, essent el llindar considerat de 0,1%.

A més de la longitud seleccionada per a la construcció de les paraules, $l = 3$, també s'ha considerat el nombre de símbols que se solapen (τ) entre una paraula i la següent ($\tau = 0, 1$ i 2 símbols). Els valors estudiats del paràmetre adimensional α que quantifica la desviació típica de la sèrie temporal han estat 0,01, 0,025, 0,05, 0,1, 0,2 i 0,5. Un altre paràmetre considerat en aquest estudi ha estat la determinació del nombre de mostres n mínim necessari per caracteritzar la dinàmica del sistema. Per aquest motiu s'ha realitzat l'estudi de la incidència de considerar 100, 200, 300 i 400 mostres de cada sèrie. L'anàlisi de la significació estadística dels resultats s'ha fet mitjançant un test no paramètric de la variància (test de Mann-Whitney) entre els índexs obtinguts per als dos nivells de pressió de suport en la ventilació assistida. S'ha considerat un nivell de significació estadística de $p < 0,05$ per a l'anàlisi.

6. RESULTATS

A la Taula 2 es presenten els resultats de l'estudi realitzat per a l'obtenció dels valors òptims dels paràmetres α , τ i n . En l'anàlisi de la selecció dels valors òptims dels paràmetres es va calcular, per a cada valor, el nombre de paraules la probabilitat d'ocurrència de les quals donava diferències estadísticament significatives en comparar les sèries temporals en els nivells de PSV alta i baixa. Es van considerar diferents nivells de significació estadística p , tal com indica la taula. Es presenten els resultats corresponents a les sèries T_{TOT} i T_I / T_{TOT} , perquè aquesta darrera ja inclou l'efecte de T_I .

$p <$	T_{TOT}			T_I / T_{TOT}		
	0,01	0,035	0,05	0,01	0,035	0,05
$a = 0,025$	3	11	14	0	2	7
$a = 0,2$	7	16	22	7	11	30
$a = 0,5$	2	5	6	1	4	25
$\tau = 0$	7	16	22	7	11	18
$\tau = 1$	8	18	26	6	16	17
$\tau = 2$	10	24	30	8	17	23
$n = 100$	9	14	16	1	5	5
$n = 200$	6	20	25	5	14	16
$n = 300$	7	20	25	8	19	25
$n = 400$	10	24	30	8	17	23

Taula 2. Anàlisi dels paràmetres característics per al càlcul de la dinàmica simbòlica. Nombre de paraules amb nivell de significació estadística menor a l'1%, al 3,5% i al 5% obtingut en comparar la PSV baixa i alta a les sèries T_{TOT} i T_I / T_{TOT} .

El nombre més gran de paraules que presentaven diferències significatives, entre els nivells de PSV alta i baixa, es va obtenir per als valors dels paràmetres $a = 0,2$, $\tau = 2$ i $n = 400$, a totes les sèries temporals analitzades. Aquests resultats es van aconseguir independentment del nivell de significació estadística p considerat. En l'estudi del darrer paràmetre n es van analitzar els registres i es van considerar subgrups de les 100, 200 i 400 primeres mostres de cada sèrie temporal.

Un cop determinats els valors òptims dels paràmetres a , τ i n , es va analitzar el nombre de paraules amb probabilitat d'ocurrència superior als llindars 1%, 2% i 5% tant en les sèries temporals T_{TOT} , T_I , com T_I / T_{TOT} per als dos nivells de PSV alta i baixa. La Taula 3 en presenta els resultats obtinguts. Les diferències estadísticament més significatives es van aconseguir amb un llindar de probabilitat d'ocurrència superior a l'1% per a les sèries temporals T_{TOT} . El nombre de paraules amb probabilitat d'ocurrència superior a un llindar també va presentar diferències significatives en analitzar les sèries temporals T_I i considerar el llindar 2%.

A l'estudi realitzat en els senyals T_{TOT} , amb un llindar de probabilitat d'ocurrència superior a l'1%, es van trobar les paraules següents que presentaven una significació estadística de $p < 0,05$ en diferenciar entre els dos nivells de PSV: «012», «112», «122», «200», «211», «221», «223», «232», «233», «323», «332», «333». En un estudi similar fet amb les sèries T_I / T_{TOT} es van trobar les paraules següents: «022», «112», «122», «211», «220», «221», «223», «233», «322», «332», «333». Per a les sèries T_I no es van trobar paraules que permetessin discriminar entre alta i baixa pressió, ja que hi ha poques paraules amb un bon nivell de significació p i no es mantenen en variar el llindar de probabilitat.

T_{TOT}	PSV Baixa	PSV Alta	p
Prob > 1%	13,9 ± 8,5	23,7 ± 9,2	0,004
Prob > 2%	9,4 ± 2,9	12,5 ± 3,4	0,006
Prob > 5%	6,1 ± 2,1	4,5 ± 2,4	0,033
T_I	PSV Baixa	PSV Alta	p
Prob > 1%	17,0 ± 7,9	21,7 ± 11,9	n.s.
Prob > 2%	9,2 ± 2,1	12,3 ± 4,1	0,008
Prob > 5%	5,7 ± 2,4	5,5 ± 2,4	n.s.
T_I / T_{TOT}	PSV Baixa	PSV Alta	p
Prob > 1%	17,3 ± 10,5	27,6 ± 9,8	0,008
Prob > 2%	10,4 ± 3,9	14,4 ± 4,1	0,008
Prob > 5%	6,1 ± 2,1	4,7 ± 2,4	n.s.

Taula 3. Valor mitjà i desviació estàndard del nombre de paraules amb probabilitat d'ocurrència més gran que 1%, 2% i 5% per a les sèries temporals T_{TOT} , T_I i T_I / T_{TOT} segons els nivells de PSV baixa i alta.

En aplicar dinàmica simbòlica s'acostuma a anomenar paraula prohibida aquella paraula que té una probabilitat d'aparèixer molt baixa. En aquest estudi es consideren paraules prohibides aquelles que presenten una probabilitat d'ocurrència per sota del llindar 0,1%. La Taula 4 presenta el nombre de paraules prohibides, amb probabilitat d'ocurrència menor al 0,1%, trobades en l'anàlisi de les sèries temporals T_{TOT} i T_I . En el cas de les sèries T_{TOT} , el nombre de paraules prohibides del grup de baixa pressió (31 ± 15) és superior al grup d'alta pressió ($17,7 \pm 13,6$) i s'obté un nivell de significació suficientment bo ($p = 0,008$) per poder discriminar entre PSV alta i baixa. En el cas de considerar les sèries temporals del temps d'inspiració T_I , aquest estadístic no ha permès discriminar entre ambdós nivells.

	PSV Baixa	PSV Alta	p
T_{TOT}	31 ± 15	17,7 ± 13,6	0,008
T_I	26,7 ± 13,9	19,2 ± 14,5	n.s.

Taula 4. Valor mitjà i desviació estàndard del nombre de paraules prohibides (probabilitat d'ocurrència < 0,1%) per a les sèries temporals T_{TOT} i T_I , segons els nivells de PSV baixa i alta.

7. DISCUSSIÓ I CONCLUSIONS

Els pacients afectes d'una insuficiència respiratòria aguda, en el curs de la seva malaltia, precisen sistemes per augmentar el volum pulmonar, o se substitueix la ventilació temporalment per després recuperar-la. En aquestes situacions, si bé es modifica el patró ventilatori, l'anàlisi lineal dels seus components no és capaç de mostrar diferències. Això es deu a la variabilitat que presenta, però, a més, i molt important, no s'utilitza en

clínica aquesta determinació per la dificultat de quantificar de forma manual molts cicles respiratoris. Mitjançant una forma automatitzada, que permetés mesurar el patró ventilatori d'un gran nombre de cicles respiratoris, i l'adequada anàlisi no lineal de les seves variables, se'n podrien quantificar les diferències.

S'ha aplicat una metodologia basada en l'estudi de la dinàmica no lineal de sistemes, la dinàmica simbòlica, que ha permès caracteritzar el patró respiratori mitjançant l'anàlisi de les sèries temporals T_{TOT} , T_I , T_I / T_{TOT} en considerar dos nivells diferents de pressió de suport de la ventilació assistida (PSV). S'han proposat uns índexs per a dur a terme aquesta caracterització: nombre de paraules amb probabilitat d'ocurrència superior a un llinar, nombre de paraules prohibides i paraules més significatives.

S'han determinat els valors òptims dels paràmetres que intervenen en el càlcul de la dinàmica simbòlica: $a = 0,2$, $\tau = 2$ símbols i $n = 400$ mostres.

Per a les sèries temporals corresponents a la durada del cicle respiratori T_{TOT} , el nombre de paraules de longitud 3, amb probabilitat d'ocurrència superior a l'1%, s'ha concentrat en $13,9 \pm 8,5$ paraules, del total de 64 paraules possibles, per als registres amb pressió de suport ventilatòria (PSV) baixa, mentre que els registres amb PSV alta s'han caracteritzat per tenir una distribució més complexa amb $23,7 \pm 9,2$ paraules. Aquesta anàlisi presenta un nivell de significació estadística $p = 0,004$. En l'anàlisi de les sèries temporals corresponents al temps d'inspiració T_I , les diferències significatives entre els dos grups es van presentar en la distribució de paraules amb probabilitat d'ocurrència superior al 2% ($p = 0,008$).

Mitjançant l'anàlisi de les paraules prohibides (probabilitat d'ocurrència inferior al 0,1%) s'ha obtingut un nou índex capaç de caracteritzar PSV baixa i alta en les sèries T_{TOT} , amb un nivell de significació $p = 0,008$, mentre que les sèries T_I no han presentat diferències significatives.

S'ha definit un conjunt de paraules que permeten discriminar, tant a les sèries T_{TOT} com a les sèries T_I / T_{TOT} , els dos nivells de pressió: «112», «122», «211», «221», «223», «233», «332» i «333».

Per tant, es pot concloure que la dinàmica simbòlica permet caracteritzar la complexitat del sistema respiratori a més de relacionar el grau de complexitat i el nivell de pressió de suport. La variabilitat del patró respiratori depèn del nivell de pressió de suport en la ventilació assistida al qual s'ha sotmès cada pacient, i ha estat la sèrie temporal T_{TOT} la que millor ha caracteritzat la seva dinàmica.

8. AGRAÏMENTS

Aquest treball ha estat finançat parcialment per la CICYT (2FD97-1197-C02-02) i pel FIS (99/0226).

9. REFERÈNCIES

- [1] J. Milic-Emili, M. M. Runstein. «Driving and timing components of ventilation». *Chest*; 70s:131-133. (1976).
- [2] S. A. Shea, R. L. Horner, G. Benchetrit, A. Guz. «The persistence of a respiratory ‘personality’ into stage IV sleep in man». *Respir Physiol*, 80:33-44. (1990).
- [3] M. J. Tobin, M. J. Mador, S. M. Guenther, R. F. Lodato, M. A. Sackner. «Variability of resting respiratory drive and timing in healthy subjects». *J Appl Physiol*, 65:309-317. (1988).
- [4] E. N. Bruce. «Temporal variations in the pattern of breathing». *J Appl Physiol*, 80:1079-1087. (1996).
- [5] T. Brack, A. Jubran, M. J. Tobin. «Effect of elastic loading on variational activity of breathing». *Am J Respir Crit Care Med*; 155:1341-1348. (1997).
- [6] A. Jubran, B. J. B. Grant, M. J. Tobin. «Effect of hyperoxic hypercapnia on variational activity of breathing». *Am J Respir Crit Care Med*; 156:1129-1139. (1997).
- [7] M. Modarreszadeh, E. N. Bruce, B. Gothe. «Nonrandom variability in respiratory cycle parameters of humans during stage 2 sleep». *J Appl Physiol*; 69:630-639. (1990).
- [8] G. C. Donaldson. «The chaotic behaviour of resting human respiration». *Respir Physiol*; 88:313-321. (1991).
- [9] M. Sammon. «Symmetry, bifurcations, and chaos in a distributed respiratory control system». *J Appl Physiol*; 77:2481-2495. (1994).
- [10] M. Sammon, J. R. Romaniuk, E. N. Bruce. «Bifurcations of the respiratory pattern associated with reduced lung volume in the rat». *J Appl Physiol*; 75:887-901. (1993).
- [11] L. Brochard, A. Rauss, S. Benito, G. Conti, J. Mancebo, N. Rekik, A. Gaspareto, F. Lemaire. «Comparison of 3 methods of gradual withdrawal from ventilatory support during weaning from mechanical ventilation». *Am J Respir Crit Care Med*; 150:896-903. (1994).
- [12] A. Esteban, F. Frutos, M. J. Tobin, I. Alia, J. F. Solsona, I. Valverdú, R. Fernández, M. A. de la Cal, S. Benito, R. Tomás. «A comparison of four methods of weaning patients from mechanical ventilation». *N Engl J Med*; 332:345-350. (1995).
- [13] A. Torres, J. M. Gatell, E. Aznar, M. El-Ebiary, J. Puig de la Bellacasa, J. González, M. Ferrer, R. Rodríguez-Roisin. «Re-intubation increases the risk of nosocomial pneumonia in patients needing mechanical ventilation». *Am J Respir Crit Care Med*; 152:137-141. (1995).
- [14] K. L. Yang, M. J. Tobin. «A prospective study of indexes predicting the outcome of trials of weaning from mechanical ventilation». *N Engl J Med*; 324:1145-50. (1991).
- [15] C. S. Sassoan, K. Mahutte. «Airway occlusion pressure and breathing pattern as predictors of weaning outcome». *Am Rev Respir Dis*; 148:860. (1993).
- [16] I. Vallverdú, N. Calaf, M. Subirana, A. Net, S. Benito, J. Mancebo. «Clinical cha-

- racteristics, respiratory functional parameters, and outcome of a two-hour T-piece trial in patients weaning from mechanical ventilation». *Am J Respi Crit Care Med*; 158:1855-62. (1998).
- [17] A. Net, S. Benito. *Función pulmonar en el paciente ventilado*. Ed. DOYMA. (1990).
- [18] R. M. Dacmarek, D. Hess, J. K. Stoller. *Monitoring in respiratory care*. Ed. Mosby. (1993).
- [19] D. R. Rigney, W. C. Ocasio, K. P. Clark, J. Y. Wei and A. L. Goldberger. «Deterministic mechanism for chaos and oscillations in heart rate and blood pressure». *Circulation*; 1-659. (1992).
- [20] R. G. Turcott, M. C. Teich. «Fractal character of the electrocardiogram: distinguishing heart failure and normal patients». *Ann Biomed Eng*; 24:269-293. (1996).
- [21] P. Achermann, R. Hartmann, A. Gunzinger, W. Guggenbuhl, A. A. Brobely. «All-night sleep EEG and artificial stochastic control signals have similar correlation dimensions». *Electroencephalogr Clin Neurophysiol*; 90:384-387. (1994).
- [22] A. Voss, R. Dietz, H. Fiehring, H. J. Kleiner, J. Kurths, P. Saperin, H. J. Vossing, A. Witt. «High Resolution ECG, Heart Rate Variability and Nonlinear Dynamics: Tools for High Risk Stratification». *Computers in Cardiology*, IEEE Computer Society Press, Los Alamitos: 261-264. (1993).
- [23] C. González, P. Caminal, S. Benito. «Complexity analysis of the respiratory system». *Proc. 5th Conf. European Society for Engineering and Medicine ESEM99*: 565-566. (1999).
- [24] M. Small, K. Judd, M. Lowe, S. Stick. «Is breathing in infants chaotic? Dimension estimates for respiratory patterns during quiet sleep». *J Appl Physiol*; 86(1):359-376. (1999).
- [25] B. L. Hao. «Symbolic dynamics and characterization of complexity». *Physica D*; 51: 161-176. (1991).
- [26] J. Kurths, A. Witt, H. Atmanspacher, F. Feudel, H. Scheingraber, R. Wackerbauer. «General remarks on complexity». *Inside versus Outside*. Springer-Verlag: 219-234. (1994).
- [27] A. Voss, J. Kurths, H. J. Kleiner, A. Witt, N. Wessel, P. Saperin, K. J. Osterziel, R. Schurath, R. Dietz. «The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death». *Cardiovascular Research*; 31: 419-433. (1996).
- [28] J. Kurths, A. Voss, P. Saperin, A. Witt, H. J. Kleiner, N. Wessel. «Quantitative analysis of heart rate variability». *Chaos*; 5 (1):88-94. (1995).

Aprendre a modelitzar amb equacions diferencials

Joan Solà-Morales

Universitat Politècnica de Catalunya

Part 1. Què és la modelització matemàtica?

PLANTEJAR PROBLEMES O MODELITZAR?

Tots estem d'acord que la modelització consisteix a convertir un problema *real* en un problema *matemàtic*. En altres paraules, molt més col·loquials, podríem dir que modelitzar consisteix a escriure les equacions que corresponen a un problema. Usar les informacions descriptives i convertir-les en relacions entre objectes numèrics, geomètrics o topològics. Convertir una situació real plena d'ambigüitats i d'apreciacions qualitatives en un problema precís entre objectes abstractes unívocament definits. Potser afegint-hi hipòtesis, si el problema no és prou precís, i discutint si són adequades.

Des del meu punt de vista això no coincideix exactament amb allò que tradicionalment en l'ensenyament de les Matemàtiques s'ha anomenat *plantejament de problemes*. Això segon és molt més concret que no pas *modelitzar*. Per exemple, a l'ensenyament secundari obligatori tradicionalment s'anomena *plantejar problemes* al que cal per respondre a qüestions com ara: *com hem de descompondre el nombre 48 en dues quantitats de manera que en dividir la primera quantitat per la segona ens doni de quocient 3 i de resta 4?* A nivell universitari això correspondria per exemple al problema, menys elemental, de *quines són les corbes el radi de curvatura de les quals és proporcional al quadrat de la distància a una recta donada?* Per plantejar les equacions corresponents a aquests dos problemes no ens cal fer cap suposició ni discutir cap hipòtesi. Es tracta d'un sistema d'equacions lineals en el primer cas i d'una equació diferencial en el segon cas, que probablement tots nosaltres l'escriuríem fins i tot de la mateixa manera. També entraria en aquesta categoria la pregunta següent (també a nivell de l'ESO): *Una aixeta omple un dipòsit en 15 minuts i una altra l'omple en 10 minuts; en quant de temps l'ompliran les dues alhora?* En aquest darrer cas el problema no és sobre objectes matemàtics, però em sembla que el seu plantejament no seria discutible (tot i que hi ha hipòtesis implícites, com ara que les dues aixetes són independents...).

En canvi hi ha problemes en què, per a plantejar-los, necessitem conèixer i discutir el fenomen físic que representen. I això és el que jo en dic *modelitzar*, i ho diferencio una mica del que n'he dit *plantejament d'equacions*. Per exemple, el famós problema de la màquina llevaneus: *Està nevant amb regularitat. A les dotze surt una màquina lleva-*

neus que recorre 2 km en la primera hora i només 1 km en la segona. A quina hora va començar a nevar?, o bé l'altre: *La població d'un país es constipa en promig una vegada a l'any. Quina fracció de la població es constipa més d'una vegada a l'any?* Per contestar aquestes preguntes necessitem hipòtesis addicionals, però que poden ser més o menys raonables. En el primer cas seria que la llevaneus elimina una quantitat constant de neu per unitat de temps. En el segon cas, que la probabilitat de tornar-se a constipar abans d'un cert temps té una funció de densitat exponencial. Totes dues hipòtesis són plausibles i raonables, però és evident que també s'ha de fer la feina de posar-les de manifest, de discutir-les i de raonar-les.

Per posar un altre exemple us demano que considereu el problema següent, que només aparentment s'assembla a un dels quals hem parlat anteriorment: *un desguàs buida un dipòsit en 15 minuts i un altre en 10. En quant de temps es buidarà si els obrim tots dos alhora?* Aquest és un problema molt més complex que els considerats anteriorment. Per començar cal conèixer la forma del dipòsit. Si mirem els llibres d'hidràulica, veurem que per resoldre problemes d'aquest tipus s'utilitza l'anomenat Teorema de Bernoulli. Però si som curiosos veurem que el Teorema de Bernoulli només es dedueix del balanç de massa i del balanç de quantitat de moviment si el fluid és ideal, o sigui incompressible i no viscos. A més, s'utilitza la hipòtesi, de vegades anomenada de règim *quasi estàtic*, que el moviment del fluid és tan lent que podem suposar que la pressió és la pressió hidrostàtica, la pressió que correspon al fet que el fluid estigui en repòs. Prescindir d'aquestes dues hipòtesis seria, doncs, també una actitud possible i fins i tot raonable en certs casos, però faria créixer molt més la complicació del problema.

Torno a dir, doncs, que almenys per a mi modelitzar és plantejar problemes en termes matemàtics, però incloent-hi, quan cal, hipòtesis addicionals, discutint-les i raonant-les.

MODELITZACIÓ, ANÀLISI I SIMULACIÓ

També cal que tinguem present que *modelitzar* no és una activitat aïllada. En primer lloc està motivada per interessos relacionats amb el problema real i segons quins siguin aquests interessos caldrà tenir en compte més o menys hipòtesis a l'hora de modelitzar. Però, des del punt de vista més matemàtic, modelitzar està lligat amb dues activitats matemàtiques molt importants: l'anàlisi matemàtica del model i la seva simulació numèrica. L'anàlisi matemàtica del model consisteix a deduir lògicament les seves propietats, com ara quantes solucions té i quins paràmetres descriuen el conjunt de totes les solucions possibles, les propietats qualitatives i geomètriques de les solucions i també com aquestes estan relacionades amb els problemes límit (*paràmetres grans i petits*, com agrada dir a alguns: temps grans, per exemple, donen lloc a les anàlisis de comporta-

ments dinàmics asimptòtics, i temps petits, a les anàlisis de *capes inicials*, *initial layers*, important en les pertorbacions singulars).

També el model haurà de ser utilitzat per a simulació numèrica. Quan dic *simulació numèrica* vull dir essencialment el mateix que *càlcul numèric*. O sigui que el model haurà de ser utilitzat per realitzar càlculs efectius, que moltes vegades seran la resposta a la motivació inicial del problema, especialment si el problema prové d'una aplicació pràctica, més que no pas si el que es pretén principalment és millorar el coneixement qualitatiu d'algun fenomen.

Aquestes dues conseqüències de la modelització, anàlisi i càlcul, també han d'influir en el procés mateix de la modelització. Dit clarament: no és gaire prudent plantejar-se problemes que hom no sabrà resoldre. Si hom es planteja el problema dels dos desguassos del que he parlat anteriorment, pot incloure-hi efectes viscosos, capilaritat i altres fenòmens a les interfases, teories ben sofisticades sobre la contracció de la vena líquida a la sortida o moltes altres coses. Tot això seria possible, però no sé si seria prudent. Tindriem un problema enorme, del qual possiblement no sabríem ni tan sols si té solució, ni quantes en té, i probablement la possibilitat de fer càlculs numèrics amb ell seria més que discutible. Per tant, a l'hora de fer models també hem de tenir en compte allò que no sabem resoldre.

MODELITZACIÓ I ORDINADORS

Parlar de la simulació numèrica dels models ens porta de la mà al tema del paper dels ordinadors. Aquest tema és extraordinàriament important, i de fet no ha fet més que començar. Des de fa, diguem, dues dècades, l'ús extens i popular d'ordinadors personals afecta a totes les branques de l'activitat humana. I pel que fa al món de la tecnologia, el seu paper ha estat i ha de seguir sent el de substituir alguns experiments costosos (encara que no tots) per simulacions amb ordinador.

Però que ningú no s'enganyi. L'ordinador només pot treballar amb objectes abstractes, que poden anomenar-se objectes matemàtics sense abús de llenguatge. Per molt que els resultats que ens mostri la pantalla d'un ordinador siguin dibuixos colorejats que semblin fotografies d'una realitat aparent, el que hi ha darrera són centenars i centenars d'equacions matemàtiques, potser senzilles o potser molt complicades, i que la màquina és capaç de resoldre de manera molt eficient.

En tot cas, doncs, en ciència i en tecnologia la modelització, és a dir, convertir els fenòmens en equacions o en relacions matemàtiques, és de plena actualitat i ho serà fins i tot de forma més intensa en els propers anys, especialment en aquelles activitats on la innovació i la recerca creativa siguin importants.

ÀREES DE LA MATEMÀTICA ON LA MODELITZACIÓ ÉS IMPORTANT

No en totes les àrees de les matemàtiques les activitats de modelització són igualment importants. Igualment, no totes les àrees de l'activitat humana admeten models matemàtics eficients. Em permeto donar temptativament una llista, fins i tot ordenada, de les activitats de modelització que jo considero més importants. És a dir, més útils i transcendents en el moment actual. La llista seria la següent: 1. Models estocàstics. 2. Models amb equacions diferencials. 3. Optimització. 4. Modelització discreta. 5. Modelització geomètrica.

Dic, doncs, que els models estocàstics constitueixen potser l'àrea on la modelització és més imprescindible i, moltes vegades, més discutible. Les diferents lleis de probabilitat són eines que poden aplicar-se per explicar certs fenòmens, però és clar que no totes expliquen tots els fenòmens. Naturalment la llei Normal té una importància especial, explicada pel Teorema Central del Límit. Però hi ha tota la casuística d'altres lleis, més o menys estàndard, més o menys *ad hoc*, que cal conèixer i saber aplicar, especialment a l'Estadística. En particular, una eina tan important i útil com la teoria de cues i la simulació de processos que se'n deriva no ha d'aplicar-se sense discutir les hipòtesis que tot sovint s'amaguen darrera dels càlculs voluminosos.

Per a mi, la segona en importància és la modelització amb equacions diferencials. A ella dedicaré les pàgines que segueixen. Naturalment, com a «professional» que sóc de les equacions diferencials, potser podria ser criticat pels meus col·legues per haver-la posat en segon lloc. Si ho he fet així no és pas per la seva importància abstracta en el conjunt de la ciència. Possiblement, les equacions diferencials han tingut un paper més important que la teoria de la probabilitat en definir això que actualment entenem per ciència i tecnologia. Més aviat ho he fet per la seva importància pràctica, en un sentit molt elemental: exagerant una mica, diré que crec que al menys ara i aquí és una mica més fàcil de guanyar diners fent modelització estocàstica que fent modelització amb equacions diferencials.

També són molt importants avui dia els problemes d'optimització. I són molt pràctics, i les empreses s'hi juguen molts diners en ells. Però la meua impressió és que en ells el model no és tan important o, si més no, tan discutible. La funció «cost» que s'ha de minimitzar normalment no se sotmet a discussió. Potser la importància gran de la modelització apareix, en canvi, quan s'han de fer simplificacions dels problemes mirant de mantenir les principals característiques del problema inicial: els problemes d'optimització sovint resulten de dimensions computacionalment inabordables, i les successives simplificacions en resulten imprescindibles.

Parlar de modelització discreta com un problema a part potser pot semblar exagerat. Però la importància creixent de les eines combinatòries i de la Teoria de Grafs en tantes activitats humanes i tecnològiques em fa creure que, sobretot pensant en el futur, és una eina de modelització que ha de ser cada cop més important.

El mateix pot dir-se de la modelització geomètrica. La importància tan gran de l'anomenada informàtica gràfica ha fet també créixer en importància tots els temes relacionats amb la representació gràfica, especialment de cossos de tres dimensions i de les relacions entre ells.

MATEMÀTICA I NO MATEMÀTICA

Jo penso que podem admetre que potser la modelització matemàtica no és exclusivament una part de la matemàtica. Però hem d'admetre també que no és una part exclusivament de fora de la matemàtica. Podrien dir que està a cavall entre la matemàtica i la no matemàtica.

Aleshores, quines persones hem de considerar que són les que han de saber modelitzar? Matemàtics o no matemàtics? Naturalment, aquesta pregunta té més interès acadèmic que no pas real. En el món de la realitat, si un model està ben fet ningú no preguntarà si l'ha fet un matemàtic, un físic, un enginyer o un economista. En canvi, des del punt de vista de la formació universitària, per exemple, sí que és raonable que ens preguntem fins a quin punt és important ensenyar a modelitzar als matemàtics, als físics o als enginyers, per exemple.

La meua opinió sobre aquesta qüestió és una mica eclèctica, perquè crec que la modelització és al mateix temps una activitat «matemàtica» i una activitat «no matemàtica». Opino que uns i altres han de saber una mica de modelització. Segurament, la modelització és un terreny en el qual s'ha de saber quant més millor del problema real i quant més millor del model matemàtic. En altre cas ens trobarem o bé amb models de laboratori, que no tenen res de matemàtics i que poden ser poc eficients, o bé amb models purament especulatius, inacceptables per un tecnòleg.

Possiblement, el terreny de la modelització sigui un dels terrenys on de manera més útil matemàtics i no matemàtics han de col·laborar.

Part 2. Aprendre a modelitzar amb equacions diferencials

DIDÀCTICA DE LA MODELITZACIÓ?

No m'atreveixo a posar per escrit consells o opinions sobre com s'ha d'ensenyar a modelitzar amb equacions diferencials, però sí que goso escriure què s'ha de fer per *aprendre* a modelitzar amb equacions diferencials. D'alguna manera, però, crec que el

problema més important de la didàctica, o sigui de l'art d'ensenyar, és decidir què és el que s'ha d'ensenyar, i potser no tant com s'ha d'ensenyar. Si admetem això, veurem que els problemes d'ensenyar i els d'aprendre són fonamentalment els mateixos: decidir què és l'important, i, si pot ser, no equivocar-se.

Jo penso que la millor manera d'aprendre a modelitzar amb equacions diferencials és estudiar i conèixer els principals models que històricament han aparegut i que han mantingut fins als nostres dies la seva vigència i el seu interès. Veure com altres han resolt els problemes amb els que es van trobar i mirar de fer el mateix em sembla una bona recepta.

Amb aquest punt de vista, doncs, que podríem qualificar de «clàssic», em permeto fer una llista dels que jo considero els principals problemes d'equacions diferencials. Les set principals equacions o sistemes, podríem dir. Tot i que això mereix alguna matisació: jo els considero els principals perquè ara me'ls miro des del punt de vista de la modelització. Considero que són importants perquè la majoria dels altres models acaben derivant-se d'aquests o derivant-se de manera similar a aquests.

La llista completa és la següent: 1) Mecànica de partícules i de sòlids rígids. 2) Equacions de la biologia matemàtica. 3) Conducció de la calor o difusió de densitats. 4) Elasticitat. 5) Mecànica de fluids. 6) Electromagnetisme. 7) Mecànica quàntica.

Ara dedicaré algunes pàgines a parlar d'ells. De cadascun m'agradarà d'escriure les equacions, en alguna forma més o menys particular, i donar almenys una referència bibliogràfica que pugui servir al lector per iniciar-se en el tema esmentat. En aquest aspecte convé dir que en el primer volum de l'excel·lent obra enciclopèdica de R. Dautray i J.-L. Lions [1] hi ha excel·lents resums dels temes 3) a 7).

1) *Mecànica de partícules i de sòlids rígids*

El moviment mecànic de partícules puntuals o de sòlids rígids és un sistema mecànic amb un nombre finit de graus de llibertat. Dóna lloc a sistemes d'equacions diferencials ordinàries, en els que l'única variable contínua és el temps, i les equacions s'obtenen sempre fent balanç de la quantitat de moviment, és a dir utilitzant la segona llei de Newton.

No cal dir que aquestes equacions han tingut una importància històrica enorme i que els problemes lligats al seu plantejament o resolució estan completament lligats al naixement del càlcul infinitesimal. La seva importància pràctica era aleshores que permetien fer càlculs relacionats amb l'astronomia i la determinació de calendaris i horaris, però també a la vegada perquè permetien el coneixement aprofundit de moviments d'aparells senzills, com ara el pèndol, de molta importància també en la construcció de rellotges.

El principal d'aquests models és, sense dubte, l'anomenat problema gravitatori de

n cossos, l'estudi del qual constitueix l'anomenada *Mecànica Celeste*. En aquest problema es tenen n cossos puntuals de masses m_1, m_2, \dots, m_n que es mouen en l'espai ordinar de tres dimensions i s'atreuen mútuament segons la llei d'atracció gravitatòria de Newton. Les incògnites en aquest cas són les trajectòries d'aquests cossos, és a dir, les funcions del temps $x_i(t), y_i(t)$ i $z_i(t)$ que ens donen la posició del cos i en cada instant de temps. El sistema de $3n$ equacions diferencials ordinàries que compleixen, com és ben sabut, és, escrit sense resumir,

$$m_i \frac{d^2}{dt^2} x_i(t) = \sum_{j \neq i} k m_i m_j \frac{x_j(t) - x_i(t)}{[(x_j(t) - x_i(t))^2 + (y_j(t) - y_i(t))^2 + (z_j(t) - z_i(t))^2]^{3/2}},$$

$$m_i \frac{d^2}{dt^2} y_i(t) = \sum_{j \neq i} k m_i m_j \frac{y_j(t) - y_i(t)}{[(x_j(t) - x_i(t))^2 + (y_j(t) - y_i(t))^2 + (z_j(t) - z_i(t))^2]^{3/2}},$$

$$m_i \frac{d^2}{dt^2} z_i(t) = \sum_{j \neq i} k m_i m_j \frac{z_j(t) - z_i(t)}{[(x_j(t) - x_i(t))^2 + (y_j(t) - y_i(t))^2 + (z_j(t) - z_i(t))^2]^{3/2}},$$

per a $i = 1, 2, \dots, n$, on k és la constant de la gravitació universal.

Aquest sistema, que com veiem s'escriu de manera força elemental, resulta contenir una estructura extraordinàriament rica de fenòmens dinàmics quan $n > 2$.

Una bona referència per acostar-se als problemes de la mecànica de partícules i sòlids rígids, i que inclou una introducció a la Mecànica Celeste, podria ser el llibre de R. Abraham i J. Marsden [2].

2) Equacions de la biologia matemàtica

Un altre món important del qual provenen moltes equacions i sistemes diferencials és el dels problemes de la biologia matemàtica. El sistema potser més famós en aquest camp és el sistema de Lotka-Volterra, que modela l'evolució al llarg del temps del nombre d'individus de dues espècies biològiques, de les quals es suposa que la segona (els depredadors) s'alimenta només de la primera (les preses) i que aquesta obté el seu aliment d'algun recurs extern que se suposa il·limitat. Si diem $x(t)$ al nombre de preses a l'instant t i $y(t)$ al nombre de depredadors, aquest sistema, també ben conegut, és

$$\frac{d}{dt} x(t) = ax(t) - bx(t)y(t),$$

$$\frac{d}{dt} y(t) = cx(t)y(t) - dy(t),$$

ón a , b , c i d són paràmetres positius. Malgrat que aquest sistema tampoc no pot resoldre's elementalment, l'anàlisi dels fenòmens que hi apareixen és força senzill. Tot i això, he volgut reproduir-lo aquí a causa de la seva enorme influència en la modelització de molts fenòmens biològics, especialment de dinàmica de poblacions quan aquestes interactuen les unes sobre les altres.

És important que fem notar que no tots els models, ni molts menys, de la biologia matemàtica són equacions diferencials ordinàries. En molts d'ells apareixen noves variables contínues que no són el temps, com poden ser el tamany, l'edat o d'altres, i donen lloc a equacions en derivades parcials o a equacions integro-diferencials, com veurem més endavant.

Una referència important per a la modelització matemàtica de fenòmens biològics és el llibre de K. Haderler [3].

3) *Conducció de la calor o difusió de densitats*

Les variacions de temperatura en un cos a causa de la conducció interna de la calor es modelitzen amb una de les equacions en derivades parcials més importants: l'equació de la calor. Les hipòtesis que sustenten aquesta modelització són les mateixes que les que serveixen per modelitzar la difusió d'una densitat, per exemple dins d'un fluid.

Suposem, doncs, que la incògnita és la funció $u(x, y, z, t)$ que representa la densitat que estem estudiant en el punt (x, y, z) i en l'instant t . Una equació força completa per a l'evolució d'aquesta densitat seria la següent. En ella la densitat u varia a causa de la difusió, però també degut al fet que és transportada perquè el fluid en el qual es difon està en moviment i també perquè reacciona amb el medi que l'envolta.

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} + b \frac{\partial u}{\partial y} + c \frac{\partial u}{\partial z} + \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z} \right) u = k \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) + F(u)$$

on $(a, b, c) = (a(x, y, z, t), b(x, y, z, t), c(x, y, z, t))$ representa el camp de velocitats (que pot ser conegut o pot estar acoblat a una altra equació) del fluid que transporta la densitat, k és la difusivitat i $F(u)$ és el terme de reacció.

Aquesta equació és més freqüent de veure-la escrita amb la notació clàssica del càlcul vectorial, que usa l'operador «nabla» de Hamilton, $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)^T$. Si es diu $V = (a, b, c)$ aleshores seria

$$\frac{\partial u}{\partial t} + \nabla \cdot (uV)k\nabla^2 u + F(u)$$

però jo crec que aquesta segona forma pot inspirar als no especialistes sensacions de misteri completament injustificades. A la literatura típicament matemàtica és més freqüent posar Δ en lloc de ∇^2 per referir-se a l'operador Laplaciana.

Aquestes equacions s'han d'acompanyar de condicions de contorn per a les variables d'espai. Les condicions de contorn que tenen sentit no són úniques i poden tenir una importància enorme en el comportament de les solucions. Aquesta és una part molt important en el procés de modelització, però aquí no podem pas entretenir-nos-hi. Les condicions més simples i naturals per a una densitat u dins d'un recinte (p. ex. un reactor químic) Ω són les que diuen que la densitat no es difon a través de les parets del contenidor, és a dir, que aquest és aïllant o impermeable:

$$\frac{\partial u(x, y, z, t)}{\partial \vec{n}} = 0 \quad \text{si } (x, y, z) \in \partial\Omega$$

on \vec{n} és el vector normal exterior unitari a $\partial\Omega$.

Per entendre els fenòmens lligats a la conducció de la calor, jo recomano per exemple el llibre de N. Ozisik [4].

4) Elasticitat

L'elasticitat i la mecànica de fluids constitueixen el que es coneix normalment com la Mecànica dels Medis Continus, per a la qual una referència introductòria general podria ser el llibre de S. C. Hunter [5]. L'elasticitat estudia els moviments interns dels sòlids elàstics i les incògnites principals són els desplaçaments en les tres direccions dels eixos, $u(x, y, z, t)$, $v(x, y, z, t)$ i $w(x, y, z, t)$ o, en altres paraules, el vector de desplaçaments $U = (u, v, w)$, que es defineix com la diferència entre la posició a temps t i la posició en equilibri de la partícula que en equilibri ocupa la posició (x, y, z) .

La teoria de l'elasticitat que podem qualificar de «pràctica» és la que també es coneix com *elasticitat lineal* o *elasticitat linealitzada* i per simplificar podria dir-se que correspon a petits desplaçaments. Tot i això, la relació entre l'elasticitat lineal i la no-lineal o general és un tema que no és pas trivial.

Les anomenades *equacions de Navier* per a les incògnites $u(x, y, z, t)$, $v(x, y, z, t)$ i $w(x, y, z, t)$ són les següents:

$$\rho \frac{\partial^2 u}{\partial t^2} = (\lambda + \mu) \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 v}{\partial x \partial y} + \frac{\partial^2 w}{\partial x \partial z} \right) + \mu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right)$$

$$\rho \frac{\partial^2 v}{\partial t^2} = (\lambda + \mu) \left(\frac{\partial^2 u}{\partial x \partial y} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 w}{\partial y \partial z} \right) + \mu \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \right)$$

$$\rho \frac{\partial^2 w}{\partial t^2} = (\lambda + \mu) \left(\frac{\partial^2 u}{\partial x \partial z} + \frac{\partial^2 v}{\partial y \partial z} + \frac{\partial^2 w}{\partial z^2} \right) + \mu \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \right)$$

on ρ és la densitat i λ i μ són els coeficients de Lamé.

Naturalment, també poden escriure's en notació vectorial, i queda

$$\rho \frac{\partial^2}{\partial t^2} U = (\lambda + \mu) \nabla (\nabla \cdot U) + \mu \nabla^2 U.$$

Han d'acompanyar-se de condicions de contorn, com ara

$$(u(x, y, z, t), v(x, y, z, t), w(x, y, z, t)) = (0, 0, 0) \quad \text{si } (x, y, z) \in \partial\Omega$$

per exemple, si les parets del nostre sòlid elàstic estan constituïdes per un sòlid rígid immòbil.

Un bon lloc per estudiar les equacions de l'elasticitat és en el llibre de T. Hughes i J. Marsden [6] o el de P. Ciarlet [7].

5) Mecànica de fluids

En la formulació més freqüent de fluids, les incògnites principals són les tres components $u(x, y, z, t)$, $v(x, y, z, t)$ i $w(x, y, z, t)$ de la velocitat $V(x, y, z, t) = (u, v, w)$ de la partícula que a l' instant t ocupa la posició (x, y, z) . Naturalment, hi ha diversos models per als fluids, que poden ser incompressibles o compressibles, viscosos o no viscosos. Per diverses raons, el sistema més conegut és el que correspon al cas viscos i incompressible, que són les anomenades equacions de Navier-Stokes:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + \nu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial y} + \nu \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \right)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial z} + v \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \right)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0$$

on ρ és la densitat, v la velocitat cinemàtica i $p(x, y, z, t)$ és una nova incògnita, la pressió, que es determina gràcies al fet que tenim una equació més.

El sistema escrit en forma vectorial és:

$$\frac{\partial}{\partial t} V + (V \cdot \nabla)V = -\frac{1}{\rho} \nabla p + \nabla^2 V$$

$$\nabla \cdot V = 0.$$

Les condicions de contorn típiques dels fluids viscosos són:

$$(u, v, w) = (0, 0, 0) \quad \text{per a } (x, y, z) \in \partial\Omega.$$

Un bon llibre per iniciar-se en la teoria matemàtica de la mecànica de fluids és, per exemple, el d'A. J. Chorin i J. Marsden [8].

6) *Electromagnetisme*

En l'estudi dels fenòmens electromagnètics (anomenats macroscòpics) a l'interior dels medis continus, les magnituds involucrades són en general les següents: ρ la densitat de càrrega, J la densitat de corrent, E el camp elèctric, D el desplaçament elèctric, B la inducció magnètica i H el camp magnètic. Totes aquestes magnituds depenen de (x, y, z, t) . La primera és un escalar, i totes les altres són vectors. Per tant, en total es tracta de setze magnituds escalars.

Les equacions de Maxwell poden escriure's com

$$-\frac{\partial D}{\partial t} + \nabla \times H = J,$$

$$\nabla \cdot D = \rho,$$

$$-\frac{\partial B}{\partial t} + \nabla \times E = 0,$$

$$\nabla \cdot B = 0.$$

Òbviament, aquestes vuit equacions són insuficients per a la determinació de les setze magnituds esmentades. En cada situació concreta, algunes de les magnituds poden considerar-se conegudes i el problema es redueix a trobar les altres. També es poden establir relacions més o menys senzilles, anomenades *constitutives*, entre E i D i entre B i H , amb el qual D i H desapareixen del sistema. Tal com les hem escrit aquí, que és la forma més habitual, semblen indicar que ρ i J són conegudes, encara que han de complir una certa equació de compatibilitat (equació de continuïtat) i que també es coneixen les relacions entre E i D i entre B i H . Altres possibilitats són tenir en compte el moviment mecànic de les càrregues, amb l'anomenada equació de la força de Lorenz i la segona llei de Newton, o bé una relació lineal entre E i J (llei d'Ohm).

El lector interessat pot utilitzar, per exemple, el llibre de J. D. Jackson [9].

7) Mecànica quàntica

En l'equació de Schrödinger de la mecànica quàntica per al moviment d'una única partícula de massa m sotmesa a un potencial (de força) $V(x, y, z, t)$, que suposem conegut, la incògnita $\Phi(x, y, z, t)$ és una funció complexa, la *funció d'ona*, el mòdul $|\Phi|^2$ de la qual representa la densitat de probabilitat que a l'instant t la partícula ocupi el lloc (x, y, z) . L'equació que ha de satisfer és

$$i\hbar \frac{\partial \Phi}{\partial t} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} \right) + V\Phi.$$

L'interès en els models quàntics des d'àrees tecnològiques i fins i tot industrials ha augmentat molt en els darrers temps, especialment per la possibilitat de modelar molècules complicades, interessants per a la indústria química o farmacèutica.

Una referència clàssica en aquest tema és l'obra de L. Landau i E. Lifschitz [10].

És evident que amb aquesta llista dels set principals models que usen equacions diferencials me'n deixo molts, i alguns d'ells molt importants. En particular, no he considerat les equacions que provenen de problemes estrictament geomètrics. Ho he fet perquè he volgut fixar-me en les que tenen més interès per a la ciència i la tecnologia, i especialment per a l'enginyeria. Això potser no hauria d'excloure necessàriament els problemes d'origen geomètric. Els problemes de processat d'imatges i la informàtica

gràfica tenen enorme interès tecnològic i fins i tot industrial, i molts d'ells acaben convertint-se en problemes relacionats amb equacions diferencials. El lector sabrà disculpar que jo hagi pres un punt de vista una mica més clàssic. Si més no, haurà servit per a no allargar massa aquest escrit.

Part 3. Per a què serveix modelitzar amb equacions diferencials?

Amb els set models que he descrit de manera molt imprecisa i resumida, en cert sentit he «pres partit» a favor d'unes equacions i sistemes una mica clàssics, podria pensar-se que fins i tot una mica «antiquats». El més jove dels sistemes que he escrit aviat farà cent anys, i el més antic almenys en té tres-cents. En bona mesura podríem qualificar-los de «problemes del segle dinou». Aquesta opció a favor dels models clàssics és plenament conscient per la meua part, i vull dedicar les pàgines finals d'aquest escrit a posar de manifest el seu interès i la seva utilitat actuals.

En cert sentit no ens ha d'estranyar que aquests problemes plantejats, per exemple, al segle dinou encara siguin d'actualitat. Tant des del punt de vista de l'anàlisi matemàtica com des del càlcul numèric, els instruments per resoldre aquests problemes no han aparegut fins a la segona meitat del segle vint. Possiblement les equacions ja podien plantejar-se molt abans, però segur que els enginyers i els tecnòlegs no ho feien gaire sovint, ja que un cop plantejades resultaven inabordables. Només fins fa relativament poc temps (vint anys, he dit més amunt) l'esquema de treball que davant d'un problema inclou *modelitzar-analitzar-simular* s'ha tornat una cosa possible, assequible i eficient.

També vull que la meua defensa de l'actualitat d'aquests models tingui un caràcter local. Barcelona, Catalunya i Espanya no han estat precisament, almenys fins fa poc, llocs de gran desenvolupament de la matemàtica. I encara és molta la feina que ens queda per fer. Jo crec que especialment ens queda molta feina a fer en posar la matemàtica en relació amb altres ciències i amb la tecnologia i la indústria. Per això he triat quatre exemples d'utilització dels models dels quals he parlat abans que s'hagin produït en el meu entorn geogràfic, entre gent que conec i amb la qual tinc relació.

Amb això, el missatge que pretenc transmetre, especialment a la gent jove, és que val la pena de conèixer els models dels quals he parlat, i que altres persones, exactament del nostre entorn, s'han obert camí professional i científic treballant en aquestes qüestions. Els quatre exemples són els següents: i) Un problema de navegació celeste estudiat per l'equip del Prof. Carles Simó (UB): Transferència a òrbites Halo. ii) L'anàlisi d'un model biològic realitzat pel Prof. Àngel Calsina (UAB i UdG) i col·laboradors: La reproducció dels rotífers. iii) Un model en derivades parcials per a sistemes de suspen-

sió-amortiment, que hem començat a considerar un petit grup de persones a la UPC.
 iv) El càlcul de potencials elèctrics en preses de terra de grans instal·lacions elèctriques, realitzat pel Prof. Manuel Casteleiro i el seu equip, primer a la UPC i després a la UdC.

Es tracta de quatre problemes que formen part de quatre dels models que hem considerat anteriorment i que tots quatre tenen en comú que no s'estudien partint de necessitats internes de la matemàtica sinó lligats a necessitats o preguntes que vénen de fora del món de la matemàtica. Tampoc no són tots quatre igualment motivats: les seves diferències van des de ser un problema d'interès principalment industrial a ser d'interès principalment científic.

i) *Transferència a òrbites halo*

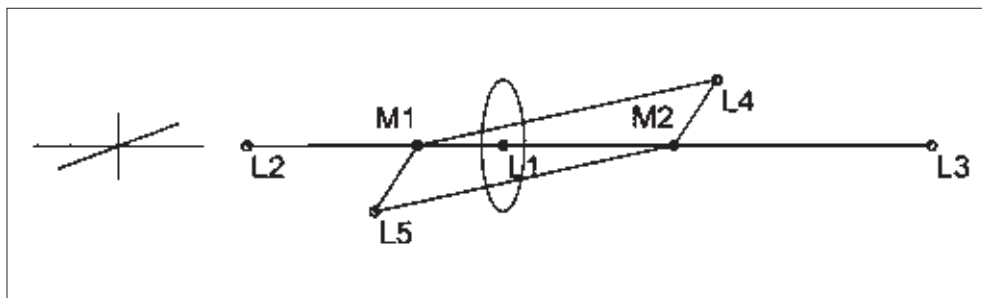
Aquest és un problema lligat al model de la Mecànica Celeste. Des dels temps de Lagrange (1772) és sabut que en un sistema gravitatori format per dos cossos en rotació (el Sol i la Terra en aquest cas, en general anomenats cossos primaris) hi ha cinc posicions d'equilibri (*punts de llibració o punts de Lagrange*) per a un tercer cos que tingui una massa menyspreable en comparació amb els dos primers (per exemple, una nau espacial). El primer d'aquests punts, anomenat L_1 , se situa sobre el segment que uneix els dos cossos primaris; L_2 i L_3 se situen també sobre aquesta recta però a una i altra banda dels dos primaris, i L_4 i L_5 se situen fora d'aquesta recta formant un triangle equilàter amb els dos primaris.

És ben sabut que L_1 , L_2 i L_3 són inestables i que L_4 i L_5 són pràcticament estables. Per això, per mantenir una nau de manera permanent a les proximitats de L_1 hauria de ser necessari realitzar algun tipus de control o maniobres per al manteniment en estació. Però resulta que el punt L_1 del sistema Sol-Terra és una ubicació excel·lent per a situar-hi un observatori de fenòmens solars, perquè minimitza la suma de les distàncies sol-observatori i observatori-terra.

D'altra banda, situar l'observatori exactament sobre L_1 té un greu inconvenient i és que la transmissió de les observacions i altres missatgers de la nau a la Terra vindria enormement distorsionada per les emissions electromagnètiques del mateix Sol i seria impossible dirigir les antenes terrestres cap a la nau sense captar també el Sol.

Per això, el projecte d'observació solar SOHO de l'Agència Espacial Europea (ESA) conjuntament amb la NASA, iniciat a la dècada de 1980, es va proposar de situar una nau no exactament sobre L_1 sinó circulant en una trajectòria periòdica propera a L_1 al voltant de l'eix terra-sol. Aquestes òrbites, anomenades òrbites halo, tenen una amplitud petita (principalment en una direcció), però prou gran perquè quan són observades des de la Terra puguin direccionar-se les antenes cap a l'estació observadora sense rebre interferències del Sol.

La figura següent mostra esquemàticament els primaris, els punts de llibració i una òrbita tipus halo al voltant de L_1 .



L'observatori SOHO (Observatori Heliosfèric i Solar) va ser dissenyat amb la finalitat d'estudiar l'estructura interna del Sol, la corona solar i l'origen del vent solar. Tot això mitjançant una observació llarga i ininterrompuda que va començar l'any 1996. Molta més informació i imatges poden veure's a la pàgina <http://sohowww.nascom.nasa.gov>.

El càlcul de les òrbites de transferència i del manteniment en estació del SOHO en una òrbita halo va ser encarregat per la ESA a l'equip de matemàtics catalans format pels professors G. Gómez, A. Jorba, J. Llibre, R. Martínez, J. Masdemont i C. Simó. Part del seu treball pot veure's descrit a [11]. El model matemàtic bàsic és el del problema gravitatori de n cossos del qual hem parlat més amunt, però al llarg dels càlculs ells van adoptar altres models, de complexitat variable, segons cada situació.

Cal dir que la família d'òrbites halo té el seu origen en la continuació de la família d'òrbites periòdiques prop de L_1 , dita de Liapunov, ben coneguda dins de la tradició matemàtica.

ii) Model matemàtic de la reproducció dels rotífers

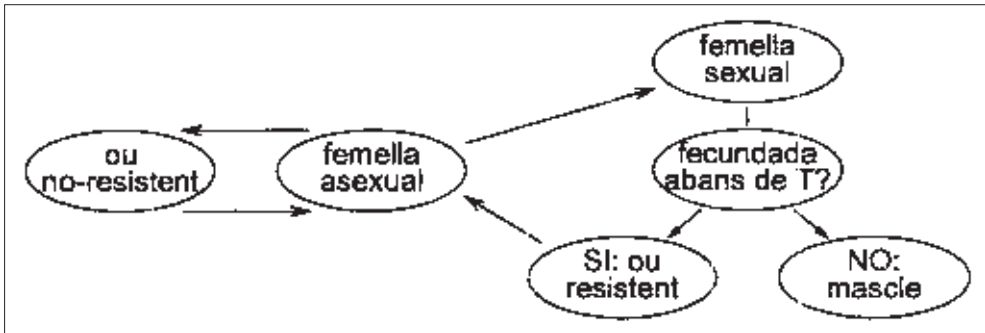
Els *rotífers monogononts* són petits invertebrats que habiten en medis aquàtics amb fortes variacions estacionals. Aquesta espècie de rotífers té mascles i femelles, les quals produeixen dos tipus d'ous en un mecanisme una mica complicat, que descrivim a continuació.

El seu cicle reproductiu té dues fases: una asexual i l'altra sexual. La fase asexual és la manera usual en què la població creix, en base a un nombre repetit de generacions de noves femelles asexuals que produeixen uns ous, que anomenem de tipus *no-resistant*, dels quals neixen noves femelles asexuals. A més d'aquest mode de reproducció asexual, i paral·lelament a ell, hi ha també períodes, de durada variable, de reproducció sexual. Aquests períodes estan lligats a l'aparició de condicions adverses, ja sigui per excessiva densitat d'individus en un medi de recursos decreixents o per condicions climatològiques negatives. Com a resultat de la fase sexual, es produiran ous del tipus que anomenem *resis-*

tent, que podran sobreviure a llargs períodes adversos, eclosionar en el moment apropiat i donar naixement a noves femelles asexuals que reinicien la fase asexual.

Quan s'inicia la fase sexual, les femelles asexuals comencen a produir al mateix temps filles asexuals i filles sexuals, amb una certa proporció. Al principi de la fase sexual, no hi ha mascles presents a la població, i per tant les femelles sexuals no poden ser fecundades. En aquest cas, aquestes produeixen ous dels quals només neixen mascles. En cas que aquestes femelles sexuals siguin fertilitzades, aleshores produeixen els ous de tipus resistent dels quals ja hem parlat. Però, més precisament, el que passa és que hi ha una edat de bloqueig T de les femelles sexuals, de manera que si no són fecundades abans de l'edat T aleshores produeixen ous dels quals només poden néixer mascles i si són fecundades abans de T produeixen ous només de tipus resistent.

El cicle asexual i el sexual estan esquematitzats a la figura següent:



Les variables del model són $u(x, t)$, $v(x, t)$ i $w(x, t)$, les densitats respecte a l'edat (x) de femelles sexuals no fecundades, de fecundades i de mascles, respectivament. Aquestes densitats de població satisfan el sistema següent d'equacions integro-diferencials en derivades parcials:

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + \mu u = -E\left(\int_0^\infty w(x, t) dt\right) u \chi_{[0, T]}(x),$$

$$\frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} + \mu v = E\left(\int_0^\infty w(x, t) dt\right) v \chi_{[0, T]}(x),$$

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial x} + \delta w = 0,$$

$$u(0, t) = B, \quad v(0, t) = 0, \quad w(0, t) = b \int_M^\infty u(x, t) dx$$

on μ i δ són les taxes de mortalitat de femelles i mascles respectivament, E la taxa de contactes mascle-femella, $\chi_{[0, T]}$ la funció característica de l'interval $[0, T]$, b la fertilitat de les femelles no fecundades i M l'edat de maduració de les femelles ($0 < T \leq M$).

Aquest model matemàtic, proposat pel Prof. A. Calsina i col·laboradors, ha estat utilitzat per explicar diversos fenòmens relacionats amb aquesta forma de reproducció. En particular, aquest model pot ser utilitzat per mostrar que, contràriament al que pot semblar, no és estable des del punt de vista evolutiu l'estratègia de prendre $T = M$, encara que aquesta maximitzaria la producció d'ous de tipus resistent ([12]).

iii) *Un model en derivades parcials per a sistemes de suspensió-amortiment*

Els sistemes de suspensió-amortiment usualment utilitzats, per exemple en els automòbils, consten d'un sistema elàstic (una molla) acoblat en paral·lel amb un amortidor. L'amortiment és una força que s'oposa a la velocitat de desplaçament i, encara que és poc realista suposar que aquesta força depèn linealment de la velocitat, el model matemàtic que s'acostuma a usar és l'equació diferencial ordinària

$$m \frac{d^2}{dt^2} u(t) = -ku(t) - r \frac{d}{dt} u(t) + \phi(t)$$

en la qual m és la massa, k el coeficient de recuperació, r el d'amortiment i $\phi(t)$ és la força exterior aplicada. Aquesta és segurament l'equació diferencial ordinària més senzilla i més important que hi ha. És l'equació de les oscil·lacions forçades esmorteïdes, i no cal dir que les seves propietats són estudiades a tots els llibres.

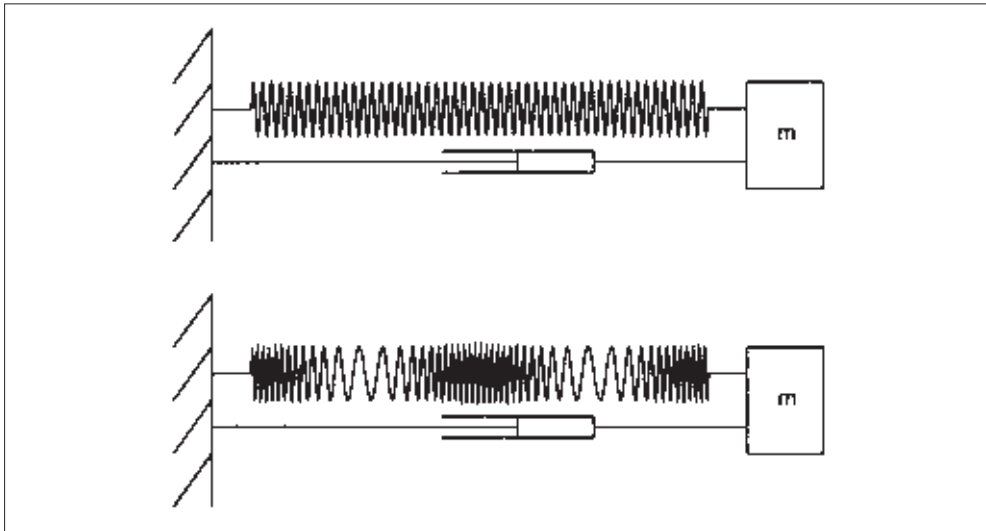
En canvi, el model que nosaltres hem començat a estudiar [13], considera la molla com un medi elàstic continu, en el qual el desplaçament de cada part de la molla no és proporcional a la distància a l'origen fix sinó que està descrit per la funció incògnita $u(x, t)$. Suposem doncs que a $x = 0$ la molla està soldada a una paret immòbil, i que a $x = L$ està soldada a una massa rígida m capaç de moure's. Admetrem que el moviment de m produeix un frec amb el medi que l'envolta i es genera una força sobre m proporcional a la velocitat $\partial u(t, L)/\partial t$, que, com abans, escriurem $-r\partial u(t, L)/\partial t$. Les equacions del moviment són

$$\rho \frac{\partial^2 u(t, x)}{\partial t^2} = k \frac{\partial^2 u(t, x)}{\partial x^2} + \alpha \frac{\partial^3 u(t, x)}{\partial t \partial x^2}, \quad \text{per a } 0 < x < L$$

$$m \frac{\partial^2 u(t, L)}{\partial t^2} = -k \frac{\partial u(t, L)}{\partial x} - \alpha \frac{\partial^2 u(t, L)}{\partial t \partial x} - r \frac{\partial u(t, L)}{\partial t} + \phi(t),$$

on ρ és la densitat lineal de la molla i α el coeficient de dissipació interna viscosa de la molla ($0 \leq \alpha \ll 1$).

La diferència entre els dos models queda representada a la figura següent:



Es tracta, doncs, d'una equació d'ones amb condicions de contorn dinàmiques. Encara que l'estudi està en fase preliminar, és força clar que serà important posar de manifest en quines situacions el segon model donarà informacions rellevants que el primer no dona. Concretament, la dissipació de les freqüències altes de vibració sembla molt menor en el segon model que no pas en el primer.

iv) Potencials elèctrics creats per grans postes a terra

El disseny de postes a terra per a grans instal·lacions elèctriques suposa el càlcul dels potencials elèctrics creats en el terreny en cas que es produeixi alguna incidència en la instal·lació. Especialment important és el coneixement d'aquest potencial sobre la superfície mateixa del terreny, ja que és la part que pot afectar la seguretat de les persones i del material. En particular, aquest potencial ha de complir certes restriccions de màxima tolerància pel que fa a diferències de potencial entre punts propers.

La posta a terra està constituïda per un conjunt d'electrodes enterrats, que ocupen una regió que anomenarem Ω , i Γ a la seva superfície. Usant les equacions de Maxwell, el càlcul de la funció potencial $V(x, y, z)$ es redueix a la resolució del següent problema exterior de Dirichlet per a l'equació de Laplace:

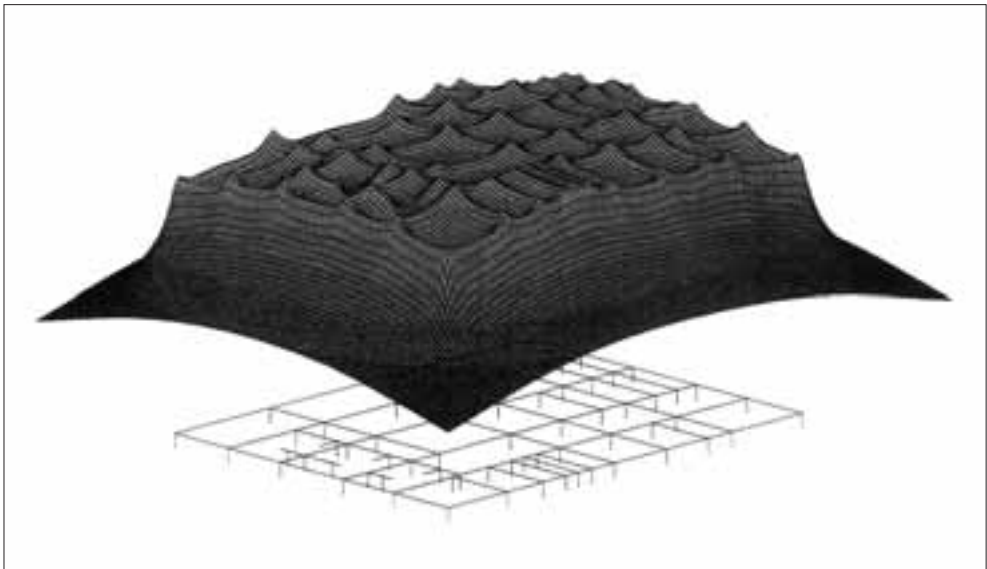
$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \quad \text{per a } (x, y, z) \in \{z < 0\} \setminus \Omega$$

$$V = V_\Gamma \quad \text{per a } (x, y, z) \in \Gamma, \quad \frac{\partial V}{\partial z} = 0 \quad \text{per a } z = 0$$

on V_Γ és el potencial elèctric (constant) que se suposa que es crea a Γ a causa de la incidència.

Aquest càlcul, que es complica a causa de la geometria complexa de Γ , que acostuma a tenir forma de xarxa, ha estat objecte d'estudi i desenvolupament en els darrers anys pel grup constituït pels Prof. Casteleiro, Colominas i Navarrina, primer a la Universitat Politècnica de Catalunya i posteriorment a la Universidade da Coruña, amb el suport de les empreses FECSA i Unión Fenosa (vegi's [14], per exemple).

La figura següent, extreta de [15] per cortesia de l'autor, mostra una xarxa de posta a terra i el gràfic de la funció potencial sobre la superfície.



REFERÈNCIES

- [1] R. Dautray i J.-L. Lions: *Mathematical Analysis and Numerical Methods for Science and Technology* (6 vol.). Springer-Verlag, 1990.
- [2] R. Abraham i J. Marsden: *Foundations of Mechanics*. Benjamin/Cummings Publishing Co., Inc., Reading, Mass., 1978.
- [3] K. Hadeler: *Mathematik für Biologen*. Springer-Verlag, Berlin-New York, 1970. (N'hi ha traducció castellana.)
- [4] M. Necati Ozisik: *Heat Conduction*. John Wiley & Sons, 1993.
- [5] S. C. Hunter: *Mechanics of Continuous Media*. Ellis Horwood Ltd., Chichester; Halsted Press [John Wiley & Sons, Inc.], New York, 1983.
- [6] J. Marsden i T. Hughes: *Mathematical Foundations of Elasticity*. Dover Publications, Inc., New York, 1994.
- [7] P. Ciarlet: *Mathematical Elasticity. Vol I. Three-Dimensional Elasticity*. North-Holland Publishing Co., Amsterdam, 1988.
- [8] A. J. Chorin i J. Marsden: *A Mathematical Introduction to Fluid Mechanics*. Springer-Verlag, New York, 1993.
- [9] J. D. Jackson: *Classical Electrodynamics*. Wiley, New York, 1975. (N'hi ha traducció castellana.)
- [10] L. Landau i E. Lifschitz: *Quantum Mechanics: Non-Relativistic Theory*. Pergamon, Oxford, 1958. (N'hi ha traducció castellana.)
- [11] G. Gómez, A. Jorba, J. Llibre, R. Martínez, J. Masdeimon i C. Simó: *Dynamics and Mission Design near Libration Points* (4 vol.). World Scientific, Singapore, 2001.
- [12] A. Calsina, J. M. Mazón i M. Serra: «A mathematical model for the phase of sexual reproduction in monogonont rotifers». *J. Math. Biol.* 40 (2000), núm. 5, pàg. 451-471.
- [13] M. Pellicer i J. Solà-Morales (en preparació).
- [14] I. Colominas, F. Navarrina i M. Casteleiro: «A boundary element numerical approach for Earthing grid computation». *Computer Methods in Applied Mechanics and Engineering*, 174 (1999), pàg. 73-90.
- [15] I. Colominas: «Cálculo y diseño asistido por ordenador de tomas de tierra en instalaciones eléctricas». Tesis Doctoral. Universidade da Coruña, 1995.

Matemáticas y demografía. Del tiempo al espacio

Daniel Devolder

Centre d'Estudis Demogràfics, Universitat Autònoma de Barcelona

INTRODUCCIÓN

En este trabajo, haré una panorámica del uso de leyes y modelos matemáticos en el campo de la demografía, ciencia del estudio de la población. La demografía y la estadística fueron desde su origen en el siglo XVII estrechamente asociadas con las ciencias matemáticas y condicionadas por el estado de su desarrollo. En concreto, y a partir de los siglos XVII y XVIII, fueron importantes para la demografía sobre todo la evolución del tratamiento matemático de la progresión geométrica, la teoría de las probabilidades y la función exponencial, que dieron a la demografía el impulso para desarrollar sus modelos fundadores, el modelo de la tabla de mortalidad y luego los modelos de crecimiento en el tiempo y de composición por edad, que presentaré en los puntos 1 y 2. Estos modelos consideran la población en su conjunto, pensando normalmente en la población de un país entero. Se desarrollaron en paralelo con la aparición de las monarquías absolutistas de Francia e Inglaterra y recibieron un gran impulso en el siglo XIX y principios del XX con los modernos sistemas de administración pública centralizada. El estado centralizado moderno encontró en la demografía la ciencia que justificó el desarrollo de los sistemas de recuento de la población, tales como los censos y los registros continuos de nacimientos, defunciones y migraciones. Pero esta óptica centralizada, considerando la población como un conjunto de individuos en un país, un espacio político esencialmente unidimensional, ha frenado el desarrollo de métodos de análisis del grupo familiar y del parentesco, así como de las características del poblamiento, que necesita como mínimo de las 2 dimensiones del plano geográfico. Es solamente de forma muy reciente que la demografía se ha ocupado de desarrollar modelos de la familia y de la variación de la densidad en el espacio, que presentaré en los puntos 3 y 4.

1. EL ANÁLISIS MATEMÁTICO DE LA MORTALIDAD

1.1. Orígenes de la tabla de mortalidad

El origen de la demografía y de la estadística modernas se remonta a un libro publicado en el año 1662, *Observaciones naturales y políticas sobre los boletines de mortalidad de la ciudad de Londres*. Después de una larga controversia, parece ahora

probado que el autor principal de este libro fue una persona con formación científica, William Petty, y que el autor único que aparece en la cubierta, el comerciante de telas John Graunt, sólo lo fue de los capítulos más descriptivos.¹ Este libro incluye la primera tabla de mortalidad moderna, basada de hecho sobre un modelo matemático más que sobre datos reales. Como lo vemos en la figura siguiente, que reproduce un gráfico original del siglo XVII construido por Christian Huygens, uno de los fundadores de la teoría de las probabilidades, esta tabla de mortalidad está construida a partir de una hipótesis de riesgo de morir constante entre 6 y 76 años. La curva indica para cada edad el número de supervivientes a partir de una cifra inicial de 100 nacimientos. Para nosotros este gráfico representa una curva exponencial, de razón positiva e inferior a 1. Pero en el siglo XVII todavía no se conocía esta función.² Los astrónomos utilizaban los resultados de Lord Napier sobre la construcción de la función logarítmica, pero habría que esperar al siglo XVIII para el estudio de su función inversa. De hecho, en la construcción de la tabla, Petty no utilizó tampoco los logaritmos, sino una progresión geométrica de razón 0,64, con redondeo al entero inferior a cada paso del cálculo, tal como se indica en la figura, un original del año 1669, dibujado por Christian Huygens, a partir de los datos de supervivientes a cada edad de Petty.³

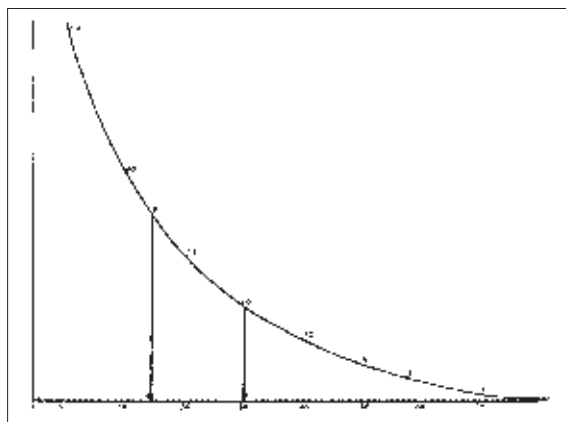


Figura 1. Gráfica teórica de mortalidad en el siglo XVII (número de supervivientes en cada edad a partir de 100 nacimientos).

1. Lo prueba de manera definitiva Hervé Le Bras, *Naissance de la mortalité*, Gallimard-Le Seuil, Paris, 2000. William Petty es conocido como fundador de la «aritmética política», precursora de la economía política moderna.

2. Los primeros trabajos que abordan el problema de forma moderna se publican en los 10 últimos años del siglo XVII (Bernoulli y Leibniz).

3. El proceso de construcción de esta tabla fue tan misterioso como su autoría. Grandes estadísticos como Karl Pearson o epidemiólogos como el Major Greenwood no consiguieron reproducirlo. Hervé Le Bras, en el libro citado, encuentra finalmente el algoritmo exacto, tan ciegamente sencillo y evidente, tal como lo hubiera hecho el Caballero Dupin, de «La carta robada» de Edgar A. Poe.

Regla de construcción de la función de superviviente de Petty: progresión geométrica de razón 0,64, con redondeo a cada etapa del cálculo:

Nacimiento	6 años	16 años	26 años	...
$\left(\frac{64}{100}\right)^0$	$\left(\frac{64}{100}\right)^1$	$\left(\frac{64}{100}\right)^2$	$\left(\frac{64}{100}\right)^3$...

A partir de esta tabla modelo, interpretada en un principio de manera probabilística, Christian Huygens y su hermano Louis derivaron, en 1669, los conceptos de probabilidad de sobrevivir entre 2 edades:

$$p(x, x + n) = \frac{S(x + n)}{S(x)}$$

y la esperanza de vida a partir de una edad (la esperanza matemática de la función de supervivencia reducida a un solo nacimiento inicial):

$$e(x) = \frac{1}{S(x)} \int_x^\infty S(y) dy$$

donde $S(x)$ son los supervivientes a la edad x de la tabla de Petty.

Huygens utilizaba para la población el lenguaje de las apuestas en juegos de azar. Esta interpretación probabilística domina todavía hoy en día, y se sigue utilizando por ejemplo el término de «esperanza» en vez de «vida media».

Los progresos posteriores en la elaboración de tablas de mortalidad, por parte sobre todo del astrónomo Edmund Halley y del estadístico Johann Süssmilch, se centraron en utilizar datos reales y no un simple modelo matemático como lo hizo Petty, lo que permitió mostrar que la hipótesis de riesgo constante con la edad no era correcta.

Otro aspecto interesante de este primer estudio de la mortalidad en los siglos XVII y XVIII es la creencia en la existencia de un riesgo de mortalidad único, fuera de los años de crisis, guerras o epidemias. Incluso después de reconocer que el riesgo de morir no es constante con la edad, los primeros demógrafos asimilaban la mortalidad a un juego con probabilidades iguales para todos y invariable en el tiempo. Buscaban regularidades que podían ser útiles para las aplicaciones en el campo actuarial, de determinación de pagos anuales en contratos de seguro de vida o rentas vitalicias. Esto explica por qué intentaban construir una tabla universal, única, válida en cualquier lugar y momento. En este sentido buscaban realmente una Ley de mortalidad, con mayúscula. Hoy en día, y muy al contrario, los demógrafos calculan tablas para todos los lugares y en cada momento posible, resaltando los aspectos diferenciales de cada una.

En el siglo XIX, cuando la obtención de datos no representaba ya un problema, pero también cuando se hizo evidente que el riesgo de morir no era igual, más alto por ejemplo en las ciudades que en el campo, y también más bajo que a finales del siglo XVII, los esfuerzos se centraron en la resolución de dos problemas: la modelización de la variación del riesgo de morir con la edad, por una parte, y el análisis de los factores de variación del nivel de este riesgo a cada edad, por otra.

1.2. Modelización de la mortalidad por edad

A partir del siglo XIX, el problema inicial de Petty y de Halley está formulado de otra manera: la mortalidad por edad obedece a dos clases de parámetros: los que afectan su nivel general y los de forma, que afectan a su progresión con la edad. Pero sigue la preocupación por encontrar una Ley universal. Así, en 1825, el matemático Benjamin Gompertz elaboró la primera función que respondía a este propósito de encontrar una Ley universal de mortalidad capaz de tener en cuenta estos atributos de nivel y forma. Se trata de una función exponencial de la edad, en la que de la manera más sencilla posible, se multiplican los dos parámetros de nivel y de forma con la edad. La Ley de Gompertz es válida a partir de los 30 años, es decir intenta describir el aumento del riesgo de morir con la edad como resultado del proceso de envejecimiento. Vemos en la figura para Estados Unidos, donde está representado este riesgo en una escala logarítmica, que la fórmula es bastante razonable, puesto que el aumento con la edad de este riesgo es casi log-lineal. Pero la Ley de Gompertz tiene la misma ambición que la Ley de la gravitación de Newton: describir totalmente un fenómeno sin explicarlo. A partir del momento en el que se utiliza para buscar explicaciones, demuestra sus limitaciones. Con dos ejemplos vamos a ver que cuando se empieza con esta Ley como instrumento para el estudio

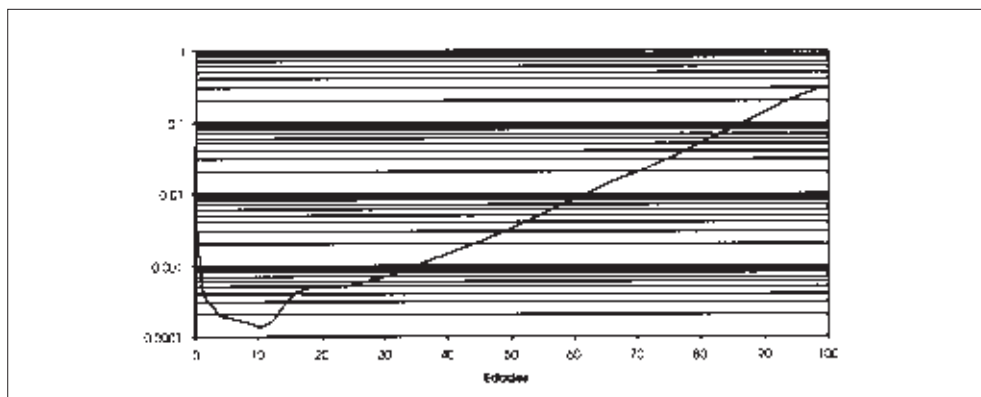


Figura 2. Logaritmo del riesgo de morir por edad, Estados Unidos, 1997.

de la mortalidad, podemos llegar a problemas o paradojas que muestran sus limitaciones, fuerzan un replanteamiento de las hipótesis de partida y nos permiten mejorar nuestro conocimiento de los determinantes biológicos de la mortalidad.

El primer problema es saber si la forma y el nivel del riesgo de morir a edades muy avanzadas están correctamente aproximados con la fórmula exponencial de Gompertz. Con el aumento del tamaño de las poblaciones, y sobre todo el perfeccionamiento de los métodos de recogida de datos, podemos ahora calcular los valores del riesgo de morir a edades superiores a 80 años⁴. En general se constata que la progresión de este riesgo deja de ser log-lineal y el crecimiento se reduce con la edad (Figura 3). Entonces la fórmula de Gompertz no es satisfactoria. Se han buscado muchas fórmulas alternativas a la de Gompertz, basadas o bien en el aumento del número de parámetros, o bien en modelos tan parsimoniosos como el de Gompertz, que hacen generalmente uso de hipótesis nacidas en el campo de las ciencias biológicas o de campos más alejados (teorías físicas del movimiento de las moléculas en un gas, teoría de la fiabilidad de los sistemas físicos, etc.). Pero una solución aceptable, que tiene el interés de conservar la sencillez de la fórmula de Gompertz, es buscar una formulación de tipo logística.

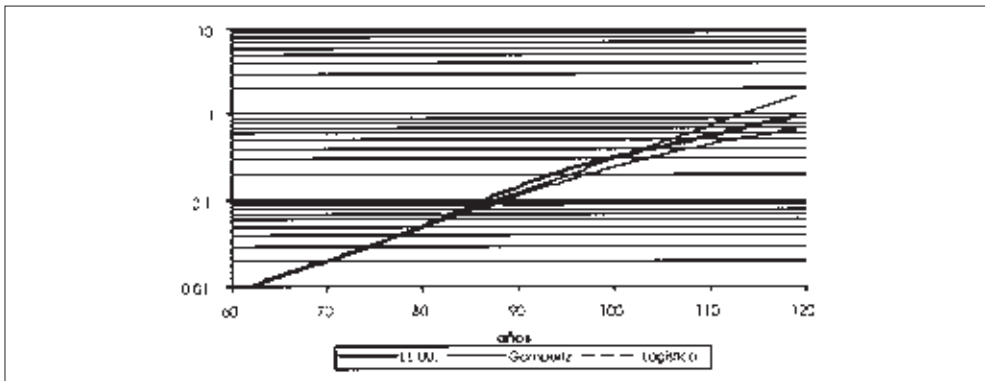


Figura 3. Mortalidad a edades avanzadas, Estados Unidos, 1997.

De las muchas fórmulas probadas que dan un mejor ajuste a partir de 80 años, las más sencillas y más utilizadas son las logísticas⁵. La forma general es la siguiente:

4. Una revisión reciente de modelos y resultados en A. R. Thatcher, V. Kannisto, y J. W. Vaupel, *The Force of Mortality at Ages 80 to 120*, Odense, 1998, Odense University Press.

5. Perks, W. «On some experiments in the graduation of mortality statistics». *Journal of the Institute of Actuaries*, 63:12.

$$\frac{\mu(x)}{1 - \mu(x)} = Ae^{bx} \Leftrightarrow \mu(x) = \frac{Ae^{bx}}{1 + Ae^{bx}} .$$

Esta nueva formulación presenta también el interés de poner un límite de tipo tendencial al valor del riesgo de morir, puesto que, tratándose de una probabilidad, no puede exceder 1, cuando en la fórmula de Gompertz los valores del riesgo tienden hacia el infinito con la edad⁶. Pero la introducción de una logística no fue solamente una solución formal. El hecho mismo de la disminución de la tasa de crecimiento de la mortalidad con la edad encuentra una explicación razonable si consideramos que las personas no son iguales delante del riesgo de morir y que hay diferencias debido a factores genéticos. La forma logística parece dar cuenta de un proceso de selección, en el cual hay distintos grupos de personas, cada uno con una ley de mortalidad tipo Gompertz, pero con un nivel distinto a cada edad. El cambio de forma de la curva y la reducción de la tasa de crecimiento de este riesgo con la edad probablemente se explican porque a edades avanzadas disminuye el peso de los grupos con mayor riesgo genético de mortalidad y aumenta de forma inversa el peso de los grupos con bajo riesgo genético de mortalidad.

Un segundo problema, la paradoja de Gumbel⁷, viene a confirmar esta conclusión. Este autor predijo, en base al ajuste de la ley de Gompertz, que en poblaciones con una alta mortalidad general, el riesgo de morir a edades avanzadas tendía a igualar o incluso a estar por debajo de este riesgo en poblaciones con una baja mortalidad general. Es lo que se puede hoy en día constatar con las tablas de mortalidad, con los datos del cuadro siguiente. Debido al crecimiento más lento del riesgo de morir con la edad, en poblaciones con una mortalidad general más alta llegan incluso en vida más personas a edades muy avanzadas que en poblaciones con una mortalidad general baja.

País	Hombres supervivientes a los 85 años (por 100.000)	Esperanza de vida al nacimiento (años)
Austria	9.947	66,6
Japón	10.287	67,7
Colombia	12.588	58,2
El Salvador	16.525	56,4
Argelia	28.546	63,2

6. El riesgo de morir, llamado en demografía cociente instantáneo o fuerza de mortalidad, no puede sobrepasar 1, pero tampoco puede lógicamente tomar el valor de 1 en una edad finita, puesto que si una persona llega a esta edad menos un segundo, ¡no tendría ya derecho a vivir!

7. Gumbel, E. J., *La durée extrême de la vie humaine*, Paris, 1937. Ver Le Bras, H., «Lois de mortalité et âge limite», *Population*, 3, 1976.

La explicación de esta paradoja utiliza el mismo tipo de argumento basado sobre un efecto de selección en el cual mueren antes los individuos con el riesgo de morir más elevado, debido a las diferencias intrínsecas o genéticas. Pero incluso en este caso, se tendría que admitir que los ancianos de los países con una mortalidad general más alta son más resistentes o tienen un riesgo de morir más bajo a edades avanzadas que los ancianos de los países con una mortalidad general más baja. La explicación más probable estaría en el nivel de la respuesta inmunológica de los organismos en relación inversa con el grado de uso de antibióticos, en general mayor en los países de baja mortalidad general.

Tanto la forma de la mortalidad a edades avanzadas como la paradoja de Gumbel se explicarían entonces por la *heterogeneidad* de la mortalidad entre individuos, debido seguramente a factores genéticos. Una parte importante de la investigación en mortalidad está actualmente dedicada al estudio de esta heterogeneidad, tanto para entenderla como para tenerla en cuenta en los estudios comparativos. El uso de una modelización matemática de la forma de la curva de riesgos por edad ha permitido, pues, plantear este problema y dar pistas sobre las causas de la inflexión de la curva normalmente log-lineal a partir de los 80 años.

1.3. Mortalidad normal y accidental. Mortalidad endógena-exógena

Para los autores de los siglos XVII y XVIII, la mortalidad tenía solamente dos niveles: el de los años normales, tal como la recogían sus tablas de mortalidad, y el nivel de crisis, por ejemplo debido a la peste bubónica. Hoy en día se suelen distinguir también otros dos tipos de mortalidad: la mortalidad debida a factores endógenos o genéticos y la mortalidad exógena, que se explica por factores del entorno, como las enfermedades infecciosas. La separación entre estos dos tipos puede basarse sobre el análisis de las causas de muerte. Pero aquí de nuevo el análisis matemático permite ordenar los datos y modelizar los procesos subyacentes. Jean Bourgeois-Pichat elaboró el primer método de este tipo, que permitió separar el nivel de la mortalidad infantil (del primer año) entre el componente endógeno (principalmente malformaciones genéticas difícilmente curables) y el componente exógeno o evitable. El razonamiento de Bourgeois-Pichat es muy interesante, y aunque el modelo está ahora superado, sigue siendo un buen ejemplo de los procedimientos que se utilizan hoy en día en las modelizaciones biomédicas⁸.

8. Bourgeois-Pichat, J., «De la mesure de la mortalité infantile», *Population*, 1946. Un trabajo actual es Manton, K. G. y Yashin, A. *Mechanisms of Aging and Mortality: Searches for New Paradigms*, Odense, 2000, Odense University Press.

2. MODELOS DE POBLACIÓN, EN EL TIEMPO Y POR EDAD

2.1. Modelo exponencial de Euler

Leonard Euler se adelantó un siglo y medio a su tiempo, publicando un trabajo redescubierto en el siglo xx en el cual anticipaba resultados importantes de los modelos modernos de las poblaciones⁹. A partir de una hipótesis de mortalidad constante en el tiempo y suponiendo que los nacimientos siguen una progresión geométrica de razón constante, mostró por ejemplo que la estructura por edad de la población (la proporción de población en cada grupo de edad) es constante en el tiempo. Estableció también otros resultados acerca de las poblaciones exponenciales que fueron redescubiertos en el siglo xx.

2.2. Modelo de las poblaciones estables de Lotka

Alfred Lotka, un matemático norteamericano, no conocía los resultados de Euler cuando fundó entre los años 1910 y 1930 la teoría matemática de las poblaciones, estableciendo los principales resultados de lo que llamó la teoría de las poblaciones estables¹⁰. Lotka empezó con el mismo modelo de población exponencial desarrollado por Euler, llamándolas «poblaciones malthusianas», haciendo referencia en esto a la oposición que hacía Malthus entre las poblaciones humanas que tienden a crecer de forma geométrica y los alimentos que lo hacen de forma aritmética. Pero Lotka introdujo una hipótesis nueva que le permitió ir más allá que el simple modelo exponencial. Sustituyó la hipótesis de nacimientos en progresión geométrica de Euler por una hipótesis de fecundidad por edad constante en el tiempo. Este cambio pequeño en apariencia fue decisivo a la hora de estudiar los procesos de convergencia de las poblaciones reales hacia poblaciones modelo. Las poblaciones modelo con fecundidad constante en el tiempo son lo que Lotka llamó las poblaciones estables. Forman un subconjunto de la clase de las poblaciones malthusianas o exponenciales.

Lotka estableció fórmulas muy útiles para el demógrafo. Por ejemplo, para el conjunto malthusiano redescubre la fórmula de Euler acerca de la proporción de personas de edad x en relación con la población total, proporción constante en el tiempo:

$$c(x) = be^{-rx}S(x)$$

9. Euler, L. *A general investigation into the mortality and multiplication of the human species*, Académie Royale des Sciences et Belles-Lettres, 1760.

10. Lotka, A. J., «Relation between birth rates and death rates», *Science*, N.S., 26:21-22 es el primer trabajo. Tratamiento completo en: Lotka, A. J., *Théorie analytique des associations biologiques*, Paris, Herman, 1936.

donde $c(x)$ es esta proporción a la edad x , b es la tasa de natalidad, r la tasa de crecimiento de los nacimientos y $S(x)$ la proporción de sobrevivientes a la edad x en la tabla de mortalidad.

Pero sobre todo muestra cómo calcular el valor de la tasa de crecimiento r de la población en el subconjunto de las poblaciones estables, resolviendo una ecuación fundamental, llamada desde entonces *ecuación de Lotka*:

$$\int_{\alpha}^{\beta} e^{-rx} m(x) S(x) dx = 1$$

con α y β los límites del intervalo de edad fértil (15 a 49 años para las mujeres) y $m(x)$ la tasa (o riesgo) de fecundidad a la edad de x años, que se calcula dividiendo los nacimientos de madres de edad x por el efectivo de mujeres de edad x . Esta ecuación tiene una única raíz real, la *tasa r de crecimiento intrínseca* de la población estable. Lotka mostró también que las raíces complejas de esta ecuación describían las condiciones del proceso de convergencia de una población real hacia la población estable definida a partir de niveles constantes en el tiempo de la mortalidad y la fecundidad. Lotka desarrolló el estudio de este proceso de convergencia a partir de una ecuación de recurrencia sobre los nacimientos que Euler también había considerado:

$$B(t) = \int_{\alpha}^{\beta} B(t-x) m(x) S(x) dx$$

donde $B(t)$ son los nacimientos en el momento t . Lotka estableció los principales teoremas de la convergencia, pero sin llegar a probarlos de forma rigurosa. Habrá que esperar los trabajos de Feller sobre los conjuntos auto-renovados¹¹, poblaciones humanas entre otras, y el estudio general de las condiciones de convergencia de Lopez¹² para establecer de forma definitiva estos teoremas.

2.3. La población estable como población equilibrada

El modelo matemático de la población exponencial de Euler y su extensión en el siglo XX, a partir de los trabajos de Lotka sobre las poblaciones estables, proporciona un instrumento de análisis que aplican de forma habitual los demógrafos. Se utiliza por ejemplo para completar y rectificar datos defectuosos o parciales, sobre todo en los paí-

11. Feller, W., «On the integral equation of renewal theory», *Annals of Mathematical Statistics*.

12. Lopez, A., *Problems in Stable Population Theory*, Princeton, 1961.

ses desarrollados. Pero las poblaciones modelo proporcionan una medida para los desequilibrios de las poblaciones humanas reales. Desde el siglo XVIII, la mortalidad y la fecundidad no se han mantenido nunca constantes, dividiéndose por tres hasta la segunda guerra mundial. También la fecundidad tiene tendencia a oscilar en el tiempo desde la segunda guerra mundial. Entonces las poblaciones estables son un modelo muy aproximado de las poblaciones reales contemporáneas. Pero pueden servir para definir lo que podría ser la población en condiciones de equilibrio, en el sentido de un crecimiento en el tiempo a una tasa constante, ¿lo que sería el ideal de todo planificador!

Así a cada población real podemos asociarle una población estable que tiene la misma tabla de mortalidad y una serie de nacimientos en progresión geométrica, con una razón igual a la de la tasa de crecimiento tendencial de la serie real. La población real se puede comparar a la estable asociada, lo que permite visualizar los desequilibrios calculando las desviaciones relativas a cada edad, como está hecho con el gráfico para Estados Unidos¹³.

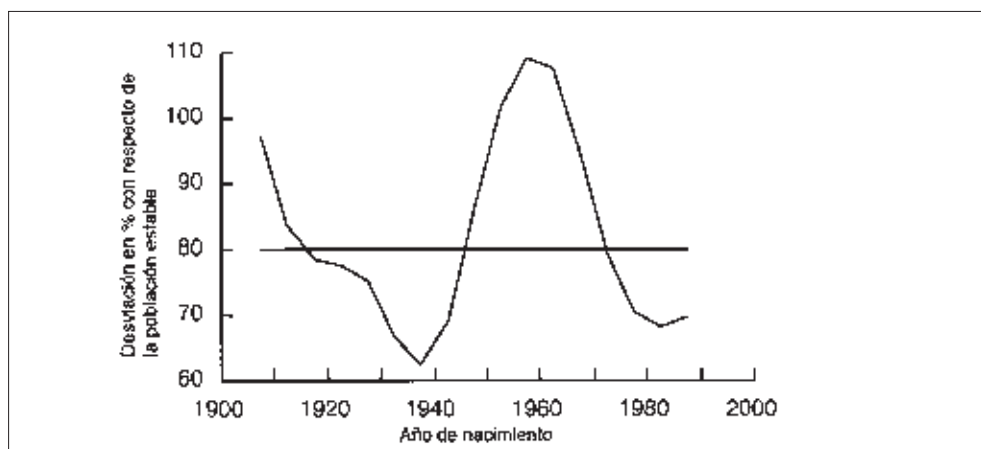


Figura 4. Estados Unidos, 1991. Comparación de la población femenina con una población estable próxima.

2.4. La convergencia débil y fuerte

El proceso de convergencia de una población real hacia una población estable supone que a partir de un momento los riesgos por edad de la mortalidad y de la fecundidad son constantes en el tiempo. En estas condiciones, la estructura por edad de una

13. Devolder, D., «Les types d'instabilités des populations du passé», *Cahiers des Annales de Démographie Historique*, 2, 2000.

población real se acerca cada vez más a la estructura de la población estable definida por estos niveles constantes de la mortalidad y la fecundidad. Este proceso de convergencia, al que se refería Lotka, se conoce ahora como convergencia *débil*. Sus características dependen de la estructura por edad inicial de la población real y de la forma de las curvas de riesgo por edad de la mortalidad y la fecundidad (el análisis de estos factores de forma se conduce a partir del estudio de las raíces complejas de la ecuación de Lotka). En general se observa que bastaría con unos 150 años de constancia de la mortalidad y la fecundidad para que una población real «olvidara» su estructura por edad inicial y se volviera estable. Pero este proceso de convergencia débil supone que se cumplan estas hipótesis irrealistas de constancia en el tiempo. Para llegar a resultados más generales, se estudia también el llamado proceso de convergencia *fuerte*, en el que se elimina la hipótesis de constancia en el tiempo de la fecundidad y la mortalidad. Por ejemplo, Coale¹⁴ formuló una conjetura según la cual si a partir de un momento sometemos dos poblaciones distintas a la misma serie de valores de los riesgos por edad de mortalidad y fecundidad, variable en el tiempo, estas dos poblaciones tenderán al cabo de un cierto tiempo hacia la misma estructura por edad (proporción de las personas a cada edad en relación con la población total). Esta conjetura fue demostrada por Lopez¹⁵. Es un resultado importante en la medida en la que establece de forma definitiva esta tendencia de las poblaciones a «olvidar» las estructuras por edad del pasado.

Pero a la inversa, como este proceso de convergencia es largo (un siglo o más), subraya también la fuerza de inercia contenida en la estructura por edad, que determina durante largo tiempo la tendencia de la evolución. Este factor de inercia se estudia bajo el nombre de *momentum* de la población¹⁶.

2.5. El proceso de la convergencia como restricción al crecimiento

El proceso de convergencia, y sobre todo de convergencia débil, se puede interpretar como la imposición de una restricción a una población¹⁷. Así la hipótesis de fecundidad constante de las poblaciones estables significa que se impone una relación constante en el tiempo entre el efectivo de los nacimientos y el efectivo de las mujeres en edad de tener hijos. En el marco más amplio de las poblaciones exponenciales de

14. Coale, A. J., «How the age distribution of a human population is determined». *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 1957.

15. op. cit. También una demostración más sencilla, a partir de la idea de procesos de media móviles, en Arthur, W. B., «The ergodic theorems of demography: a simple proof», *Demography*, 1982.

16. Keyfitz, N., «On the momentum of population growth», *Demography*, 8. También Li, N. y Tuljapurkar, S., «Population momentum for gradual demographic transitions», *Population Studies*, 53, 1999.

17. Le Bras, H., «Fluctuations et croissance des populations soumises à une contrainte», *Population*, 2, 1983.

Euler, la hipótesis de evolución exponencial de los nacimientos es también una restricción en el sentido que los nacimientos de dos años consecutivos tienen que estar en una relación constante. Otro tipo de restricción que se ha estudiado es la de forzar una población a tener un crecimiento nulo, es decir, que se igualen los nacimientos con las defunciones, un objetivo aparentemente razonable para una política demográfica. Obviamente estas restricciones son difíciles de hacer cumplir para la población de un país entero, pero en algunos casos se pueden aplicar, por ejemplo para la población de un cuerpo profesional (maestros, policías, ...) o la plantilla de una gran empresa. El problema que plantean estas restricciones es que la relación de constancia que se tiene que cumplir fuerza el ajuste de otros componentes de la evolución, y este ajuste puede tener consecuencias imprevistas y en algunos casos peores que la situación a la que se pensaba remediar con la restricción. En el gráfico siguiente se aprecia como una fecundidad constante, y aun más una tasa de crecimiento constante, son restricciones que provocan un ajuste de la población en forma de fluctuaciones en el tiempo de los nacimientos.

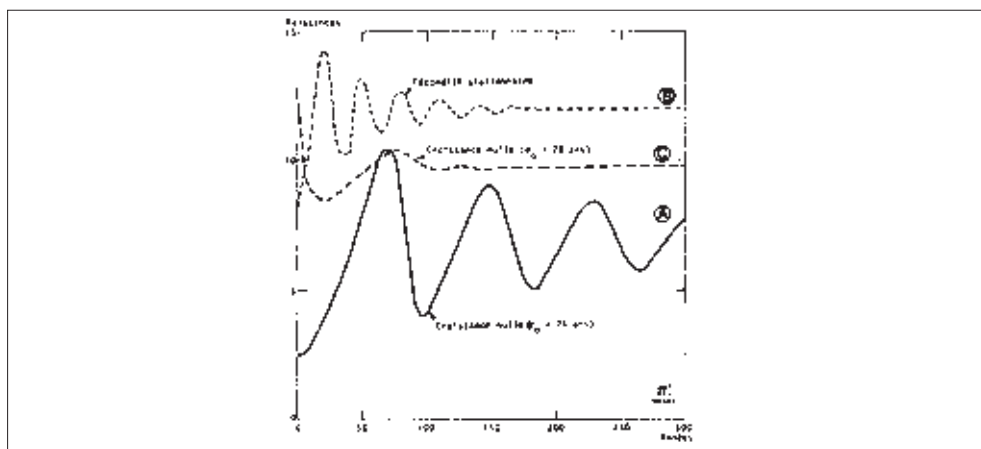


Figura 5. Evolución comparada de los nacimientos en el caso de crecimiento nulo y en el caso de fecundidad constante.

En este contexto el estudio matemático de la convergencia tiene un especial interés, sobre todo en la medida en la que puede orientar a un planificador hacia el tipo de restricción que sea menos peligroso o más suave en términos de estos ajustes dinámicos.

2.6. Modelo de población estable generalizado

El modelo de Lotka es una aproximación: las poblaciones estables son siempre hipotéticas, y solamente en casos muy peculiares se podría llegar a las condiciones de

constancia necesarias. Esto explica por qué de forma más reciente se han establecido fórmulas más *generales* que las del modelo estable, que permiten conducir cálculos directos, sin aproximaciones, a partir de datos de poblaciones reales¹⁸. En este sentido, más que un modelo, basado normalmente sobre hipótesis simplificadoras, tenemos que hablar de una contabilidad, puesto que las nuevas relaciones son siempre exactas.

La idea de partida de esta generalización es considerar las tres dimensiones de edad, tiempo y generación de las poblaciones como caminos que recorren los grupos de personas de la misma edad. Y la relación entre variables se puede deducir de la velocidad de los desplazamientos sobre estas tres dimensiones demográficas, sin privilegiar ninguna de ellas. Esto se puede apreciar a partir del gráfico siguiente, que representa lo que llaman los demógrafos una superficie de Lexis. El tamaño de la población por edad en cada momento es la elevación de la superficie en este gráfico, pero los desplazamientos se producen sobre esta última. Este tamaño se reduce a medida que pasa el tiempo, la primera dimensión. Pero el paso del tiempo es también un aumento según la dimensión de la edad para las personas. Este aumento conjunto del tiempo y de la edad son las dos fuerzas elementales que, combinadas, dan como resultado un desplazamiento en diagonal sobre la dimensión de generación (aumento a la vez del tiempo y de la edad para personas nacidas el mismo año).

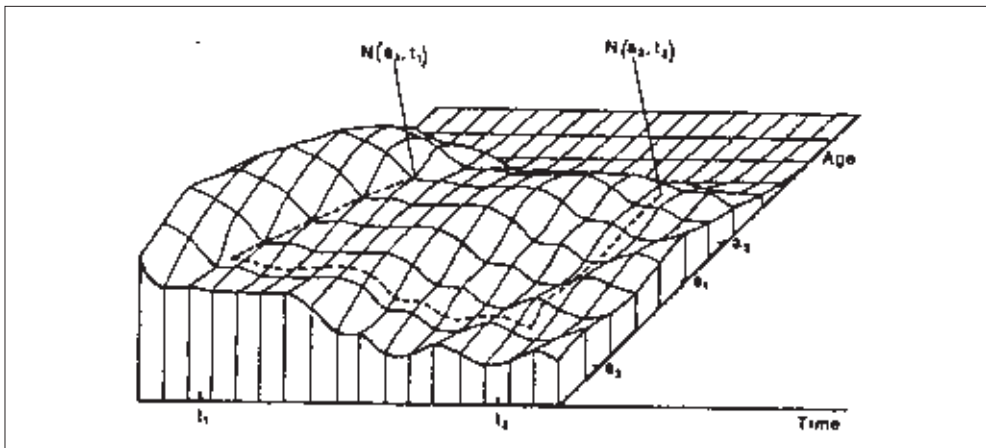


Figura 6. Un desplazamiento sobre una superficie de Lexis.

18. Preston S. H. y Coale, A. J., «Age structure, growth, attrition and accession: a new synthesis», *Population Index*, 1982. También Arthur, W. B. y Vaupel, J. W., «Some general relationships in population dynamics», *Population Index*, 1984.

Este razonamiento conduce a la definición de tres tasas de cambio fundamentales que permiten «viajar» sobre el gráfico de tres dimensiones de la población:

$$r(a, t) = \frac{\partial \ln N(x, t)}{\partial t},$$

la tasa de crecimiento en el tiempo de la población de edad x ,

$$v(a, t) = - \frac{\partial \ln N(x, t)}{\partial x},$$

la tasa de crecimiento con la edad de la población de edad x ,

$$\mu(a, t) = - \frac{\partial \ln N(x + y, t + y)}{\partial y},$$

la tasa de crecimiento con la edad y el tiempo, que a su vez es el riesgo o la fuerza de la mortalidad de la generación a la edad x , si no hay migraciones.

Estas tres tasas están unidas por la identidad fundamental de este modelo generalizado:

$$\mu(x, t) = v(x, t) - r(x, t).$$

Sobre esta base se pueden reformular todas las fórmulas del modelo estable a partir de estas tasas. Por ejemplo, la proporción de población a la edad x ,

$$c(x) = b e^{-\int_0^x r(y, t) dy} S(x),$$

la ecuación de Lotka,

$$\int_{\alpha}^{\beta} e^{-\int_0^{\infty} r(y, t) dy} m(x) S(x) dx = 1.$$

Esta formulación generalizada permite también establecer nuevas relaciones, que no podían aparecer en el modelo estable, por ejemplo relativas a los factores del envejecimiento¹⁹, las relaciones entre tablas de mortalidad del momento y tablas de generación²⁰, etc.

19. Preston, S. H., Himes, C. y Eggers, M., «Demographic conditions responsible for population aging», *Demography*, 1989.

20. Horiuchi S. y Preston, S., «Age-specific growth rates: the legacy of past population dynamics», *Demography*, 1988.

3. MODELOS DE LA REPRODUCCIÓN Y DEL PARENTESCO

Los modelos demográficos tradicionales empiezan con la población como objeto de estudio. Pero son los individuos los elementos generadores del cambio de las poblaciones, porque tienen determinados comportamientos de procreación y están sujetos a riesgos de morir o de migrar. La demografía desde sus orígenes se define toda como una transición incesante entre estos dos niveles, macro para la población y micro para los individuos. Esto explica por qué hay una verdadera dificultad de la demografía para situarse a niveles intermedios, como son el nivel familiar y en general el de las redes que relacionan los individuos entre sí, tanto el parentesco como los grupos sociales. Para conseguir estudiar esta dimensión nueva, las relaciones entre personas, se han desarrollado de forma mucho más reciente modelos de la reproducción y del parentesco que empiezan con los individuos en vez de con la población en su conjunto. La idea es reconstruir poco a poco el nivel familiar elemental (pareja, hijos) y luego llegar hasta todo el parentesco, y eventualmente hasta el nivel de las redes de sociabilidad. Hay varios tipos de modelos de la familia y del parentesco. Pero los más comunes y útiles son modelos que empiezan con un individuo central (ego), permiten construir genealogías y también calcular las probabilidades de tener familiares en vida, a cada edad del ego.

Los modelos de parentesco más utilizados hoy en día son *micro-simulaciones*. Complementan y reemplazan los modelos matemáticos iniciales que llevaban a fórmulas para el cálculo de relaciones de parentesco que incorporaban integrales múltiples, lo que hace difícil el cálculo de resultados. Como ejemplo de fórmula matemática para determinar el número esperado de primas hermanas, se tiene que calcular el valor de fórmulas con integrales múltiples como:

$$\int_{\alpha}^{\beta} \left[\int_{\alpha}^{\beta} \left\{ \int_y^{a+x+y} \left(\int_{\alpha}^{a+x+y+z} S(w)m(w)S(a+x+y-z-w)dw \right) \frac{S(z)}{S(y)} m(z)dz \right\} W(y)dy W(x)dx \right]$$

Se necesitan métodos de cálculo numérico, en los cuales es necesario utilizar números aleatorios. Entonces es más rápido y más interesante incorporar los factores estocásticos al principio mismo de la modelización, lo que es el planteamiento de una micro-simulación. Éstas ofrecen además la ventaja de permitir incorporar de forma sencilla nuevas hipótesis sobre los comportamientos, sin aumento del nivel de complejidad de los cálculos.

La micro-simulación empieza normalmente con un individuo en vida a cada edad y calcula a partir de distribuciones de probabilidad el número de parientes en vida, progresando paso a paso a partir de la madre, el padre, el cónyuge, etc. Estas micro-simulaciones se llaman a menudo de Monte-Carlo porque utilizan procedimientos de tiraje de números

al azar que sirven para seleccionar el valor de una probabilidad dentro de distribuciones preestablecidas (de mortalidad, fecundidad, nupcialidad, migración, etc.)²¹.

4. MODELOS ESPACIALES: SIMULACIÓN DE AGENTE Y DENSIDAD FRACTAL

El estudio de la difusión de la población en el espacio o el problema general de los tipos de poblamiento son temas que tradicionalmente los demógrafos no han considerado como suyos sino del dominio de la geografía. La única dimensión territorial relevante que se toma en cuenta en demografía son las migraciones. Pero las migraciones se definen en relación con el espacio político unidimensional de un país u otras unidades administrativas. Para estudiar los movimientos humanos dentro de un espacio continuo y bidimensional, y no segmentado desde una visión política, es necesario abandonar el concepto mismo de migración y interesarse por la dinámica de la difusión en el espacio. Hay dos formas de estudiar esta dinámica: desde una visión de los procesos individuos elementales o bien a partir de la distribución espacial en un momento dado.

4.1. *El juego de la difusión*

La modelización de los procesos de difusión en el espacio encuentra un formalismo cómodo en lo que se conoce como el «juego de la vida». Lo que empezó como una curiosidad matemática se ha convertido hoy en día en algo muy útil para la experimentación en biología, en física o en informática²². La idea es trabajar con una estructura de celdas, normalmente en dos dimensiones, ocupadas por individuos de una población. En la versión original del juego, introducido por el matemático John Conway, la reglas son relativas a la ocupación de las celdas por nacimiento o defunción, en función de la presencia de individuos en las celdas vecinas. Se pueden adaptar estas reglas al caso de las migraciones, con una población en este espacio de celdas que se desplazan de una celda libre a otra en función de un parámetro de atracción (los hombres aborrecen la soledad) y otro de repulsión (una densidad local demasiado alta puede ser un problema). Como ejemplo de resultados, podemos tomar dos poblaciones distribuidas inicialmente al azar a las que se le han aplicado dos series distintas de

21. Una síntesis y resultados en Devolder, D. *Effects of the European late marriage pattern on kinship. A study using a microsimulation model*, en prensa, 2001. Disponible en la web en la dirección: <http://www.ced.uab.es/pdfs/paperspdf/text135.pdf>.

22. Por ejemplo Sigmund, K., *Games of Life. Explorations in ecology, evolution and behaviour*, Oxford, 1993.

valores de estos parámetros de atracción y repulsión²³. La aplicación iterativa de estas reglas conduce a la situación de equilibrio representada en los dos dibujos de Conway siguientes:

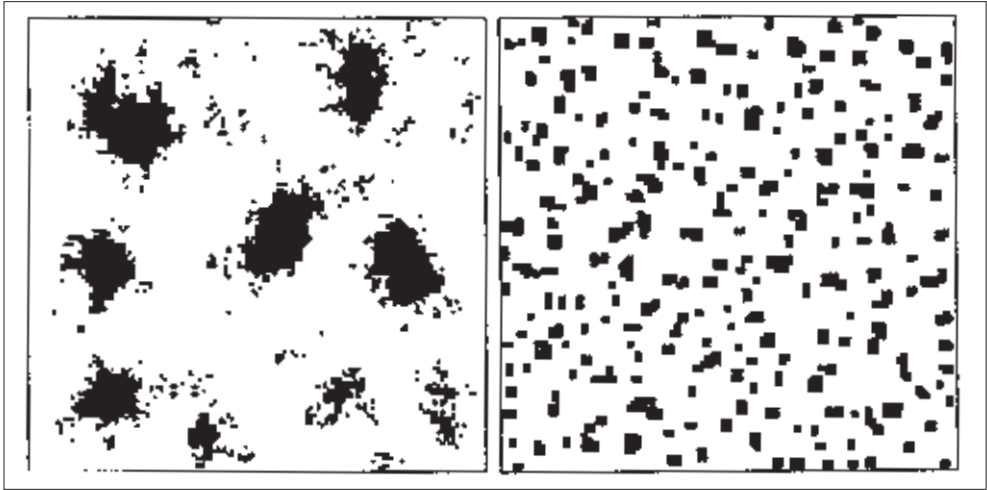


Figura 7. Dos situaciones de equilibrio en el juego de la vida, según Conway.

Estas dos formas de ocupación del territorio son representativas de la división histórica de las zonas rurales en Europa occidental, entre poblaciones dispersas y poblaciones agrupadas. Este tipo de modelización sugiere que esta oposición entre dos tipos fundamentales de poblamiento se puede explicar por pequeñas alteraciones de las reglas básicas de atracción y repulsión.

4.2. La densidad fractal

Otro procedimiento para estudiar estas reglas elementales es empezar con el tipo de poblamiento observado en vez de intentar reconstruirlo, como en el enfoque anterior²⁴. La ocupación humana del territorio normalmente no es continua y se suelen encontrar repeticiones de las conformaciones espaciales. Estas repeticiones se

23. Le Bras, H., *Essai de géométrie sociale*, Paris, 2000.

24. Frankhauser, P., «The fractal approach. A new tool for the spatial analysis of urban agglomerations», *Population: An English Selection*, 1998.

producen en varias partes del territorio, pero también en una misma zona, pasando de una escala a otra más pequeña. Esta ocupación del territorio con un patrón repetido sugiere una analogía con las figuras de la geometría fractal, como lo podemos ver con el mapa de la ciudad francesa de Besançon, comparado con una alfombra de Sierpinski.

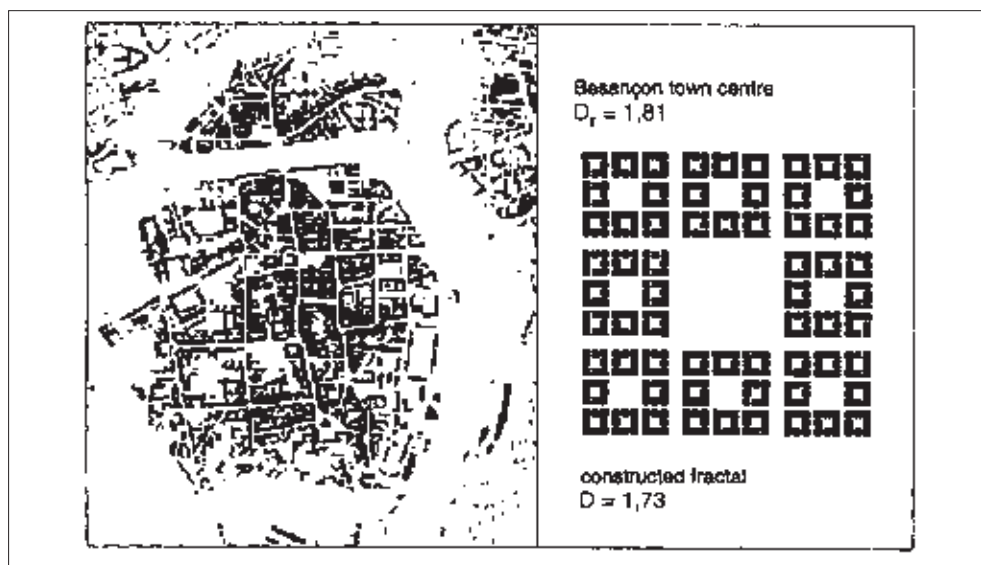


Figura 8. El centro de Besançon ($D_f = 1,81$) comparado con un fractal ($D = 1,73$). La distribución de los espacios libres en el fractal se ha ajustado a las condiciones reales, que parecen caracterizarse por la presencia de bloques amplios construidos alrededor de patios interiores.

Este objeto matemático podría ser un modelo de las reglas de segmentación del territorio, seguidas por la población de esta ciudad. La alfombra se genera en base a una cuadrícula inicial con un perímetro de longitud l , según una regla de partición en 8 cuadrículas de perímetro $lr = l/3$. Esta operación se reproduce a la escala siguiente y el efecto de invariancia de escala se puede medir a través de la dimensión fractal, el parámetro D del gráfico calculado por:

$$D = \frac{\ln N}{\ln(1/r)}$$

donde N es el número de cuadrículas creadas a partir de una anterior.

Este parámetro de invariancia de escala es una medida más adecuada que la densidad de la población (efectivo por unidad de superficie). Toma valores entre 1 (desierto) y

2 (superficie totalmente ocupada) y permite una cuantificación basada sobre la observación de que la ocupación del territorio es un fenómeno discreto (en cada punto hay población o no la hay) y no continuo, como lo sugiere una métrica como la densidad.

Queda entonces como problema la estimación de la dimensión fractal asociada con cada territorio, lo que en general se puede resolver con procedimientos de tratamiento informático de imágenes. La dimensión fractal observada puede servir entonces para construir un modelo del proceso elemental de fragmentación del espacio seguido por la población del territorio analizado.

Evolucions aleatòries: alguns problemes matemàtics de la seva modelització

Marta Sanz i Solé

Universitat de Barcelona

1. INTRODUCCIÓ

Els matemàtics, i potser també la majoria de científics, tenen una percepció de l'atzar ben diferent de la que té la majoria de la gent. Em ve a la memòria un article del Sr. Oriol Pi de Cabanyes, publicat a *La Vanguardia* l'any 2000, que duia per títol «Noticias del azar». Era un article fresc, dels que al final de la lectura et deixen un somriure de satisfacció. Parlava d'atzar com a sinònim de casualitat, com la paraula que utilitzem per cobrir la nostra ignorància i incertesa; acabava dient que, sense casualitats, la vida seria molt avorrida.

Analitzant aquestes afirmacions des del punt de vista d'una matemàtica, sense concessions poètiques, hi trobo un fons de veritat i un embolcall de ficció. En efecte, la incorporació de l'atzar en l'anàlisi matemàtica de fenòmens científics, com l'evolució de poblacions, la transmissió de senyals, les variacions dels mercats financers, fenòmens de turbulència, etc., es fa, parcialment, per poder donar entrada a un cúmul de causes i condicions que no podem descriure amb precisió. En aquest sentit, utilitzem models aleatoris quan tenim una certa ignorància o incertesa.

Ara bé, també ho podem presentar des d'un punt de vista més optimista i, de fet, més científic: Utilitzem models aleatoris quan, pensant en termes de relació causa-efecte, no hi ha una expressió funcional *simple* entre el que observem i la causa que ho ha produït. Expressat d'aquesta manera, no parlem d'ignorància ni d'incertesa, sinó d'inaadequació del model determinista per explicar la realitat.

Ambdós tipus de models, aleatoris i deterministes, es fonamenten en teories matemàtiques rigoroses i, per tant, exclouen les conclusions *per casualitat* o *per sorpresa*.

2. BREU RESUM HISTÒRIC

La part de les matemàtiques que tracta de l'atzar s'anomena *teoria de la probabilitat*. A diferència d'altres branques, el desenvolupament de la probabilitat és relativament recent: se situa al segle XVII. Una correspondència entre els matemàtics Blaise

Pascal (1623-1662) i Pierre de Fermat (1601-1665) va donar lloc a l'anàlisi d'uns quants jocs d'atzar, a l'aplicació de tècniques de combinatòria per resoldre'ls i a l'aparició de paradoxes que a la vegada provocaren nous desenvolupaments.

Ben aviat es van posar de manifest els problemes de fonamentació lògica, molt importants, que la noció i el tractament de l'atzar suposaven. Cal destacar que molts d'aquests no van quedar resolts fins a començaments del segle xx, amb el desenvolupament de la teoria de la mesura i de la integració per Borel i Lebesgue, entre altres. Malgrat això, de les discussions sobre jocs es va passar a la consideració de problemes matemàtics que van portar a construir les bases de la teoria. Cal citar, per exemple, les contribucions de Jakob Bernoulli (1654-1705), Abraham de Moivre (1667-1754), Leonhard Euler (1707-1783), Pierre-Simon de Laplace (1749-1827), Karl-Friedrich Gauss (1777-1855), Simeon-Denis Poisson (1781-1840), entre d'altres. Molt aviat, la teoria es va obrir a aplicacions diverses. Citem, per exemple, els problemes de demografia i assegurances, de riscos econòmics i socials o d'estimació d'errors en mesures físiques. Recíprocament, com resulta freqüent en els avenços científics, les aplicacions van motivar la creació d'un cos teòric amb entitat pròpia.

Un salt qualitatiu notable es va donar, com hem citat anteriorment, a començaments del segle xx. Les contribucions més importants van venir de Rússia (Moscou, Sant Petersburg), començant per l'axiomatització de la teoria feta per A. N. Kolmogorov el 1933; de França i, una mica més tard, dels Estats Units. Es va començar a treballar amb les nocions de *procés estocàstic* o *aleatori*, d'*estructures estocàstiques* i en la seva projecció com a models matemàtics per a problemes diversos del propi camp de les matemàtiques o d'altres àrees científiques.

3. PROCESSOS I ESTRUCTURES ESTOCÀSTIQUES

Un procés estocàstic és una família de variables aleatòries, $\{X_t, t \in T\}$, definides en un espai de probabilitat (Ω, \mathcal{F}, P) , indexades per un espai de paràmetres T . Molt sovint el conjunt T representa el temps, és a dir $T = [0, \infty)$ o $T = \mathbb{N}$. Més endavant, però, veurem que és també necessari considerar conjunts més generals. L'anàleg determinista dels processos estocàstics són les funcions.

Una estructura estocàstica és una propietat de les distribucions de probabilitat d'un procés estocàstic. L'estructura estocàstica més senzilla correspon a la noció d'*independència*, però és també la menys interessant en les aplicacions a la modelització d'evolucions aleatòries. Per estudiar aspectes evolutius, interessa considerar relacions de *dependència* entre les diferents variables aleatòries X_t . Introduïm-ne algunes.

a) *Estacionarietat*. Aquesta noció està motivada per resultats de la Física sobre l'existència de mesura invariant en el temps per a l'evolució de sistemes dinàmics conservatius en l'espai de fases. Un procés estocàstic és estacionari si les seves distribucions en dimensió finita són invariants per translació. És a dir, si per a qualssevol t_1, t_2, \dots, t_n de T i $s > 0$, la distribució de probabilitat del vector aleatori $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ és la mateixa que la de $(X_{t_1+s}, X_{t_2+s}, \dots, X_{t_n+s})$.

b) *Dependència markoviana*. Va ser introduïda per A. A. Markov (1856-1922) en un intent d'estendre la llei dels grans nombres a variables aleatòries dependents. Després la va aplicar a l'estudi estadístic de la successió de vocals i consonants en el text de Pushkin, E. Onegin. Actualment la propietat de Markov és àmpliament utilitzada en problemes d'organització de serveis (cues), de gestió d'estocs, en models genètics i d'evolució de poblacions.

La propietat de Markov s'expressa dient que el futur i el passat són condicionalment independents del present. La independència condicional és un concepte que es formula utilitzant σ -àlgebres d'esdeveniments. Donades tres σ -àlgebres, \mathcal{F}_i , $i = 1, 2, 3$, diem que \mathcal{F}_3 és condicionalment independent de \mathcal{F}_1 donada \mathcal{F}_2 , si, per a tot $B \in \mathcal{F}_3$,

$$P(B | F_1 \vee F_2) = P(B | F_2).$$

En el cas dels processos estocàstics, fixat un instant $t > 0$, les σ -àlgebres que cal considerar per expressar la propietat de Markov estan relacionades amb els esdeveniments que podem descriure en funció de X_s , $0 \leq s \leq t$ (passat), X_s , $s > t$ (futur) i X_t (present).

c) *Martingala*. Aquesta noció va ser introduïda per Ville per a la descripció de l'evolució temporal dels guanys d'un jugador. Aquesta és la definició: El procés $\{X_t, t \in [0, \infty)\}$ és una martingala, si per a tot $0 \leq s \leq t$,

$$E(X_t | X_r, 0 \leq r \leq s) = X_s.$$

Expressa la idea de joc equitatiu o just, en el sentit que el valor mig dels guanys es manté constant.

Les trajectòries d'un procés estocàstic són les funcions reals obtingudes en fixar el component aleatori. Més precisament, tot procés estocàstic dóna lloc a una aplicació $X: \Omega \times T \rightarrow \mathbb{R}$. Fixat $t \in T$ tenim una variable aleatòria. Podem, però, també fixar $\omega \in \Omega$ i, aleshores, obtenim una funció real definida sobre T . Aquesta funció s'anomena *trajectòria* del procés. Les trajectòries del procés corresponen a les observacions dels fenòmens aleatoris. La majoria d'exemples importants de processos aleatoris tenen trajectòries molt irregulars. Això fa que el seu tractament analític no sigui, en general, possible amb les eines de l'anàlisi matemàtica clàssica. Aquesta dificultat ha

estat l'origen del desenvolupament de l'anomenat càlcul estocàstic del qual parlarem breument en la secció 5.

4. EXEMPLES DE PROCESSOS ESTOCÀSTICS

Potser el més popular dels processos estocàstics és el moviment brownià o procés de Wiener. La manera més simple, però a la vegada potser més abstracta, de definir el moviment brownià és com un procés $\{B_t, t \in [0, \infty)\}$ gaussià, és a dir, tal que per a qualsevol t_1, t_2, \dots, t_n , la llei del vector aleatori $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$ és una normal multidimensional, amb $E(B_t) = 0$ i

$$E(B_{t_1}B_{t_2}) = \min(t_1, t_2).$$

L'origen d'aquest procés cal trobar-lo en la descripció de les observacions que el botànic anglès R. Brown va fer de partícules de pol·len en suspensió en un líquid. Més tard, aquest model va captar l'interès d'Albert Einstein, en relació a fenòmens físics de difusió i, posteriorment, va ser utilitzat per Norbert Wiener en el desenvolupament d'un càlcul diferencial en dimensió infinita que, en cert sentit, va ser el precursor del càlcul estocàstic que hem esmenat més amunt.

A partir del model de Brown, no és difícil imaginar que les trajectòries del moviment brownià són irregulars. Un resultat de Dvoretzki, Erdős i Kakutani prova que, amb probabilitat 1, no són diferenciables en cap punt. Les trajectòries del moviment brownià proporcionen, doncs, un exemple de funció real que és contínua en qualsevol punt, però no n'és derivable en cap. Un camí ple de punxes! La següent aproximació al moviment brownià fa més creïble aquesta afirmació.

Considerem un individu que es desplaça una unitat d'espai a la dreta o a l'esquerra depenent del resultat del llançament d'una moneda. Suposem que parteix en el temps $n = 0$ d'una posició $S_0 = 0$; en el temps $n = 1$ es trobarà en $S_1 = 1$ o $S_1 = -1$, etc. La posició en un temps $n > 0$, enter qualsevol, pot expressar-se així:

$$S_n = S_0 + \sum_{j=1}^n X_j,$$

on $\{X_j, j \geq 1\}$ és una successió de variables aleatòries independents que prenen els valors 1, -1 amb probabilitat 1/2. Equivalentment, $S_n = S_{n-1} + X_n$. La família $\{S_n, n \geq 0\}$ és també un exemple de procés estocàstic. S'anomena la passejada aleatòria de Bernoulli, simètrica, sobre els enters.

Per obtenir el moviment brownià, forcem el viatjant a llançar la moneda i, per

tant, decidir el sentit dels seus desplaçaments, molt més sovint. Canviem, per tant, l'escala del temps: la nova unitat de temps serà δ . Simultàniament l'obliguen a fer desplaçaments més petits. Canviem, doncs, les variables aleatòries X_n . Concretament, considerem una successió de variables aleatòries independents, idènticament distribuïdes, $\{Y_n, n \geq 1\}$, que prenen valors $\sqrt{\delta}$ i $-\sqrt{\delta}$ amb probabilitat 1/2. Aleshores, si posem

$$W_t^\delta = \sum_{k=1}^{\lfloor t/\delta \rfloor} Y_k,$$

es pot demostrar que hi ha convergència en llei, quan δ tendeix a zero, cap al moviment brownià.

És suficient considerar processos estocàstics indexats per $T = [0, \infty)$? Hem associat la noció de procés estocàstic a una evolució temporal. Ara bé, en problemes de modelització, pot ser necessari considerar evolucions en funció de diversos paràmetres, per exemple, la situació a la superfície de la terra, la situació a l'espai, etc. Si el conjunt dels paràmetres T és multidimensional (fins i tot pot ser un espai de dimensió infinita), parlem de *camp aleatori*. Per exemple, com a generalització del moviment brownià, podem introduir el *llançol brownià* o procés de Wiener en dos paràmetres. És un procés gaussià, $\{B_{s,t}, (s, t) \in [0, \infty)^2\}$, amb $E(B_{s,t}) = 0$ i

$$E(B_{s_1,t_1} B_{s_2,t_2}) = \min(s_1, s_2) \min(t_1, t_2). \quad (1)$$

Es pot fer una construcció basada en consideracions geomètriques anàlogues a les que hem especificat abans per al moviment brownià. Les trajectòries són ara superfícies molt irregulars.

En el *llançol brownià* els dos paràmetres s i t juguen un paper simètric. Podem trobar-nos amb problemes on sigui més realista diferenciar-los; per exemple, si considerem evolucions en el temps de fenòmens que tenen lloc en un context espacial. Pot ser també que ens interessi diferenciar el grau de dependència estocàstica en els diversos paràmetres. Per a processos gaussians, la manera de fer-ho és a través del funcional de covariància. Això ens porta a introduir una classe de processos més àmplia, com expliquem tot seguit.

Considerem un procés estocàstic $\{F(\phi)\}$ indexat per un conjunt de funcions D definides sobre $[0, \infty) \times \mathbb{R}^d$, prou regulars. Suposem que per a qualssevol ϕ_1, \dots, ϕ_d de D , la llei del vector aleatori $(F(\phi_1), \dots, F(\phi_d))$ és una normal multidimensional, amb $E(F(\phi)) = 0$ per a tot $\phi \in D$, i

$$E(F(\phi_1) F(\phi_2)) = \int_0^\infty dr \int_{\mathbb{R}^d} dx \int_{\mathbb{R}^d} dy \phi_1(r, x) f(x - y) \phi_2(r, y), \quad (2)$$

on f és una funció positiva i compleix determinades propietats que no precisem aquí. És a dir, es tracta d'un procés gaussià, centrat, amb funció de covariància definida per (2).

En la variable temporal r ; aquesta ve donada mitjançant la mesura de Lebesgue, dr ; en la variable espacial x s'expressa mitjançant una operació de convolució. Diem que $\{F(\phi)\}$ és un soroll blanc en temps, o incorrelacionat, i correlacionat en espai.

Observem que el llançol brownià es pot obtenir com a cas particular d'un camp aleatori com el que acabem de descriure. En efecte, posem $B_{s,t} = F(1_{[0,s]} \times 1_{[0,t]})$ i suposem que la funció f és una distribució de Dirac en zero. Aleshores, prenent $\phi_i = 1_{[0,s_i]} \times 1_{[0,t_i]}$, $i = 1, 2$ resulta que la integral del segon membre de (2) esdevé

$$\int_0^\infty dr 1_{[0,s_1]}(r) 1_{[0,s_2]}(r) \int_{\mathbb{R}} dx 1_{[0,t_1]}(x) 1_{[0,t_2]}(x) = \min(s_1, s_2) \min(t_1, t_2).$$

El llançol brownià és un soroll incorrelacionat en les dues coordenades.

L'elecció de camps aleatoris, correlacionats o no, es fa en base a dos criteris. En primer lloc, tenint present el problema que volem modelitzar; en segon lloc, tenint en compte les possibilitats matemàtiques que tenim a l'abast per a l'anàlisi del model. Insistirem en aquest punt en l'apartat 6 d'aquest article.

5. CÀLCUL ESTOCÀSTIC

Les tècniques del càlcul diferencial i integral per a funcions no poden aplicar-se als processos estocàstics, ja que aquests, en general, presenten un comportament molt més irregular. En el decenni 1950-1960, el matemàtic japonès K. Itô va proposar una noció d'integració (estocàstica) respecte del moviment brownià, va estudiar les propietats del procés estocàstic donat per aquesta integral, va establir regles de càlcul, com la fórmula del canvi de variables, i va aplicar els resultats a l'estudi d'equacions diferencials estocàstiques. Avui coneixem aquesta teoria com a càlcul estocàstic.

Els treballs d'Itô tenen com a precedent les aportacions de Wiener. En un article de l'any 1938, Wiener manifesta la necessitat de donar un fonament teòric al procediment de la integració estocàstica per motivacions de la mecànica estadística.

... Nevertheless, it (Statistical mechanics) developed without an adequate harmony of concepts and mathematical technique, which is only now in the process of development at the hands of the modern school of students of integral theory. [Wi]

A continuació fem una breu introducció a la construcció de la integral estocàstica d'Itô. L'objectiu que ens proposem és de donar sentit a la integral d'un procés aleatori, $\{X_t, t \in [0, \infty)\}$, respecte del moviment brownià $\{B_t, t \in [0, \infty)\}$. Per a tot $t \geq 0$, denotem per \mathcal{F}_t la σ -àlgebra generada per B_s , $0 \leq s \leq t$. Suposem primer que

$$X_t(\omega) = Z_a(\omega)1_{[a, b]}(t), \quad (3)$$

on Z_a és una variable aleatòria \mathcal{F}_a -mesurable i $0 \leq a \leq b$. Definim

$$\left(\int_0^\infty X_s dB_s \right) (\omega) = Z_a(\omega)(B_b - B_a)(\omega). \quad (4)$$

El resultat és, doncs, una variable aleatòria. A partir d'ara, no explicitarem la dependència en ω , per simplificar les notacions. Clarament, podem estendre per linealitat la definició (4) a processos estocàstics que són combinacions lineals amb coeficients reals de processos com els de (3). La qüestió fonamental és: Fins on podem estendre la definició (4)? És a dir, quin tipus de processos podem integrar? Es poden caracteritzar? La resposta donada per Itô és la següent:

La integral estocàstica respecte del moviment brownià es pot definir per a la classe de processos estocàstics $X = \{X_t, t \in [0, \infty)\}$,

- mesurables respecte a la σ -àlgebra producte $\mathcal{F} \otimes B([0, \infty))$, on $B([0, \infty))$ denota la σ -àlgebra de Borel,
- adaptats, és a dir, tals que per a cada $t \geq 0$, X_t és \mathcal{F}_t -mesurable,
- $E\left(\int_0^\infty X_s^2 ds\right) < \infty$.

Denotarem per $L_a^2([0, \infty) \otimes \Omega)$ el conjunt d'aquest processos.

El procediment de construcció de la integral estocàstica es basa en l'extensió, per continuïtat, d'un funcional lineal entre espais de dimensió infinita. La igualtat

$$E\left(\int_0^\infty X_s dB_s\right)^2 = E\left(\int_0^\infty X_s^2 ds\right),$$

fàcilment comprovable utilitzant la propietat de martingala del moviment brownià, és la clau de l'extensió. Aquesta igualtat expressa que la integral estocàstica és una isometria entre l'espai de les variables aleatòries de quadrat integrable i l'espai $L_a^2([0, \infty) \otimes \Omega)$ que hem introduït més amunt. No entrem aquí en el detall dels resultats d'anàlisi funcional que es necessiten per justificar el mètode.

El procés *integral estocàstica* es defineix mitjançant $\int_0^\infty 1_{[0, t]}(s) X_s dB_s$ i el denotarem per $Z_t = \int_0^t X_s dB_s$, $t \geq 0$. Es pot demostrar que $\{Z_t, t \geq 0\}$ té la propietat de martingala i totes les variables aleatòries del procés tenen esperança zero i moment de segon ordre donat per

$$EZ_t^2 = E\left(\int_0^t X_s^2 ds\right).$$

No és difícil imaginar generalitzacions d'aquesta integral a processos amb paràmetre multidimensional o a processos que tenen com a factor una funció determinista. Per exemple, podem integrar respecte del llançol brownià processos $\{X_{s,t}, (s, t) \in [0, \infty)^2\}$ mesurables respecte a la σ -àlgebra producte $\mathcal{F} \otimes B([0, \infty)^2)$, adaptats a la família de σ -àlgebres $\mathcal{F}_{s,t}$ generades per les variables $B_{u,v}$, $0 \leq u \leq s$, $0 \leq v \leq t$ i tals que

$$E\left(\int_0^\infty \int_0^\infty X_{s,t}^2 ds dt\right) < \infty.$$

Com pot veure's, per exemple, en [W2], la justificació de tals generalitzacions no és, en absolut, trivial.

6. EQUACIONS EN DERIVADES PARCIALS ESTOCÀSTIQUES

Com hem dit a l'apartat 5, l'aplicació immediata i, també, la motivació de la teoria d'integració estocàstica d'Itô van ser les equacions diferencials estocàstiques.

Voldríem introduir aquí un tema relacionat, però potser menys conegut pel públic matemàtic, que és una font de problemes molt interessants i susceptible d'aplicacions en diversos camps científics: *Les equacions en derivades parcials estocàstiques*. Per tal d'evitar massa tecnicismes, ens centrarem en exemples i, sobre ells, discutirem en un to divulgatiu alguns problemes.

6.1. Model estocàstic de resposta neuronal ([W1])

L'objecte d'estudi d'aquest exemple són les cèl·lules nervioses, les neurones. Com en totes les cèl·lules, existeix una distribució desigual d'ions, en aquest cas de sodi, a ambdues bandes de la membrana citoplasmàtica; això produeix una diferència de potencial elèctric, que s'anomena *potencial de repòs*. Les neurones reben estímuls que provoquen variacions brusques del potencial, degut a la inversió de la polaritat (*despolarització*) de la membrana neuronal; això produeix el *potencial d'acció*. Els estímuls arriben a les neurones per les *dendrites*, que tenen una configuració arbòria. Hi arriben a través d'unes connexions funcionals que s'anomenen sinapsis, que embolcallen la zona dendrítica. Com a resultat d'aquest procés, els estímuls arriben de forma incontrolada i en un gran nombre en un petit interval de temps.

L'estudi de l'evolució del potencial d'acció al llarg del temps t , i en les diferents parts de la neurona, està òbviament relacionat amb la seva resposta als estímuls. Volem donar un model d'aquesta evolució. L'evolució d'un potencial és un fenomen que els físics han estudiat abastament; el model simplificat que en donen és el de *l'equació del*

cabla. Aquí el cable serà el cos de la neurona, idealitzat per un cilindre prim o, més idealitzat encara, per un segment de recta. En el nostre problema, l'evolució ha de tenir en compte les entrades produïdes pels estímuls que hem descrit anteriorment. Per tant, en l'equació que ens expressa la difusió del corrent, hi haurem d'afegir algun terme que modelitzi els estímuls persistents.

Això ens porta a parlar d'equacions diferencials o, en aquest exemple, en derivades parcials estocàstiques. Formalment, es tracta d'equacions que descriuen fenòmens deterministes (com l'equació de la calor o l'equació de les ones) però a les quals afegim un terme *aleatori*. No cal dir que haurem de precisar què entenem per aquesta descripció tan ambigua.

Concretem una mica més. El model que proposem per a l'evolució del *potencial d'acció* es pot descriure de forma no rigorosa de la manera següent. Representem la neurona per l'interval $[0, L]$. Denotem per t el temps, que pot ser qualsevol nombre positiu, i per $x \in [0, L]$, l'espai. $V(t, x)$ designa el *potencial d'acció* a l'instant t en el punt x . Posem

$$\begin{aligned} \frac{\partial V}{\partial t}(t, x) &= \frac{\partial^2 V}{\partial x^2}(t, x) - V(t, x) + \sigma((V(t, x))\dot{F}(t, x)) \\ V(x, 0) &= V_0(x), \\ \frac{\partial V}{\partial x}(t, 0) &= \frac{\partial V}{\partial x}(t, L) = 0, \quad t > 0, \end{aligned} \tag{5}$$

on σ és una funció real de variable real. La segona de les equacions dóna la condició inicial, això és, per a $t = 0$; la tercera dóna les condicions a la vora, és a dir, en els dos extrems del cable.

L'equació (5) s'ha obtingut modificant l'equació en derivades parcials

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x) &= \frac{\partial^2 u}{\partial x^2}(t, x) - u(t, x) = 0 \\ u(x, 0) &= u_0(x) \\ \frac{\partial u}{\partial x}(t, 0) &= \frac{\partial u}{\partial x}(t, L) = 0, \quad t > 0. \end{aligned} \tag{6}$$

Hi hem afegit un terme, $\sigma((V(t, x))\dot{F}(t, x))$, no lineal, que inclou l'atzar, el qual representem aquí formalment per $\dot{F}(t, x)$, la derivada d'un procés estocàstic. Expressem així que la resposta neuronal depèn del potencial local i de les entrades d'estímuls múltiples i desordenades.

No cal dir que l'expressió (5) és purament formal. Per donar-li un sentit rigorós caldrà triar el tipus de procés estocàstic que, en forma diferencial, hem escrit com $F(t, x)$; després interpretarem (5) en forma integral i analitzarem alguns problemes derivats de tal interpretació.

La forma integral de (5) està totalment inspirada en l'expressió de la solució de (6) mitjançant la seva funció de Green $G(t; x, y)$.

Pel al problema (6), suposant, més generalment, que el segon membre de la primera equació és una funció $g(t, x)$, és conegut que

$$u(t, x) = \int_0^L G(t; x, y)u_0(x) + \int_0^t \int_0^L G(t - s; x, y)g(s, y) ds dy,$$

amb

$$G(t; x, y) = \frac{\exp(-t)}{\sqrt{4\pi t}} \sum_{n=-\infty}^{\infty} \left[\exp\left(-\frac{(y-x-2nL)^2}{4t}\right) + \exp\left(-\frac{(y+x-2nL)^2}{4t}\right) \right]. \quad (7)$$

Fixem $t_0 > 0$. Per analogia amb el que acabem d'exposar, direm que el procés estocàstic $\{V(t, x), (t, x) \in [0, t_0] \times [0, L]\}$ és solució de l'equació (5) si es compleix que

$$V(t, x) = \int_0^L G(t; x, y)V_0(y) dy + \int_0^t \int_0^L G(t - s; x, y)\sigma(V(s, y))F(ds, dy). \quad (8)$$

Clarament, d'acord amb el que hem dit a l'apartat 5, per donar sentit a la integral estocàstica de (8) haurem d'imposar condicions sobre la funció σ . D'altra banda, com ara veurem, el comportament de la funció de Green, dit altrament, el tipus d'equació en derivades parcials que considerem, també té molta importància. Finalment, l'altre ingredient fonamental és el procés F .

Situem-nos en el cas particular més senzill en què la funció σ sigui constant o fitada i el procés F és un llançol brownià $\{F(t, x), (t, x) \in [0, t_0] \times [0, L]\}$. Una anàlisi acurada de la funció de Green definida en (7) proporciona la fitació següent:

$$G(t; x, y) \leq C(t)^{-\frac{1}{2}} \exp\left(-\frac{|y-x|^2}{4t}\right), \quad (9)$$

on C és una constant positiva. Aleshores, tenint en compte les condicions per a l'existència de la integral estocàstica, resulta que $\int_0^t \int_0^L G(t - s; x, y) \sigma(V(s, y)) F(ds, dy)$ està ben definida si, i només si,

$$\int_0^t \int_0^L (G(s; x, y))^2 ds dy < \infty.$$

Això es compleix, ja que

$$\int_{\mathbb{R}} s^{-\frac{1}{2}} \exp\left(-\frac{|y-x|^2}{2s}\right) = \sqrt{2\pi}.$$

Considerem ara la variant del model anterior donada per

$$\begin{aligned} \frac{\partial V}{\partial t}(t, x) &= \frac{\partial^2 V}{\partial x^2}(t, x) + \sigma(V(t, x))\dot{F}(t, x) \\ V(0, x) &= V_0(x), \end{aligned} \tag{10}$$

$t \geq 0, x \in \mathbb{R}^d$. Si la funció σ és zero, tenim l'equació clàssica de la difusió de la calor, en un espai d -dimensional, provenint d'una font situada en un punt. Va ser deduïda per Joseph Fourier a partir de la seva llei sobre el flux de l'energia calorífica i el principi físic de la conservació de l'energia. Anomenarem l'equació (10) *equació de la calor estocàstica*.

La funció de Green per a aquest problema és

$$G_1(t, x) = (2\pi t)^{-\frac{d}{2}} \exp\left(-\frac{|x|^2}{2t}\right).$$

Per analogia amb l'equació (5), expressarem la solució de (10) mitjançant l'equació integral següent:

$$V(t, x) = \int_{\mathbb{R}^d} G(t; x-y) V_0(y) dy + \int_0^t \int_0^L G(t-s; x-y) \sigma(V(s, y)) F(ds, dy). \tag{11}$$

Retornant al problema de l'existència de la integral estocàstica que apareix en (11), observem que, per a tota $d \geq 2$,

$$\int_0^t \int_{\mathbb{R}^d} (G(t-s; x-y))^2 ds dy = \infty.$$

En efecte,

$$\int_0^t \int_{\mathbb{R}^d} (G(t-s; x-y))^2 ds dy = C \int_0^t s^{-\frac{d}{2}} ds = \infty.$$

La conclusió d'aquest exemple és la següent. En dimensió espacial més gran que 1, l'equació de la calor estocàstica amb un soroll aleatori incorrelacionat en temps i també en cadascun dels components espacials no és tractable amb tècniques de l'anàlisi estocàstica, aplicables a processos estocàstics reals. Ara bé, si prenem un soroll correlacionat en espai, com hem definit en (2), tenim la possibilitat d'imposar condicions sobre la funció de correlació f per tal d'aconseguir que la integral estocàstica tingui moment de segon ordre finit. Precisem aquesta afirmació. L'expressió de la funció de covariància donada en (2) pot escriure's en funció de l'operació de convolució $*$ de la manera següent:

$$E(F(\phi_1)F(\phi_2)) = \int_{\mathbb{R}^d} f(x)(\phi_1 * \tilde{\phi}_2)(x) dx, \tag{12}$$

on $\tilde{\phi}_2(x) = \phi_2(-x)$. Utilitzant resultats ben coneguts d'anàlisi harmònica, el segon membre de (12) coincideix amb la integral $\int_{\mathbb{R}^d} \mu(dx) \mathcal{F}\phi_1(x) \overline{\mathcal{F}\phi_2(x)}$, on \mathcal{F} denota la transformada de Fourier i μ és la mesura espectral associada a f . Per exemple, per a la funció $G_1(t, \cdot)$, que és la densitat d'una llei normal d -dimensional, la transformada de Fourier (funció característica, en llenguatge probabilístic) val

$$\mathcal{F}G_1(t, \cdot)(x) = \exp(-4\pi^2 t |x|^2).$$

Càlculs senzills demostren que, en aquest cas, el segon membre de (12) és finit si, i només si

$$\int_{\mathbb{R}^d} \frac{\mu(dx)}{1 + |x|^2} < \infty.$$

En certa manera aquest és un punt decebedor de la teoria, però com hem esmenat abans, hi ha fenòmens per als quals resulta més realista una modelització mitjançant processos aleatoris correlacionats en la coordenada espacial. El següent n'és un exemple.

6.2. Un model de creixement de poblacions ([D-S])

Considerem una població, distribuïda espacialment, que evoluciona seguint l'equació en derivades parcials

$$\begin{aligned} \frac{\partial X}{\partial t}(t, x) &= \frac{1}{2} \frac{\partial^2 X}{\partial x^2}(t, x) + \sigma X(t, x) \\ X(x, 0) &= X_0(x), \end{aligned} \tag{13}$$

$t \geq 0$, $x \in \mathbb{R}^d$, on el valor de $\sigma > 0$ està relacionat amb el medi en el qual es produeix l'evolució.

Aquest és un model determinista. Una manera d'introduir la incertesa en la descripció del medi consisteix a substituir el paràmetre determinista σ per un ingredient aleatori.

Per exemple, en [D-S] es proposa modificar (13) de la manera següent:

$$\begin{aligned} \frac{\partial X}{\partial t}(t, x) &= \frac{1}{2} \frac{\partial^2 X}{\partial x^2}(t, x) + \sigma X(t, x) \dot{W}(t, x) \\ X(x, 0) &= X_0(x), \end{aligned} \tag{14}$$

on s'ha introduït un paràmetre aleatori en la forma d'un múltiple ($\sigma > 0$) d'un soroll aleatori blanc en temps i correlacionat en espai. La correlació en espai es presenta aquí com una condició natural, ja que les propietats en punts propers del medi han d'estar relacionades.

Aquest model és una versió més complexa de propostes fetes per altres autors, vegeu [T], basats en equacions diferencials estocàstiques que aleatoritzen el model competitiu clàssic de Lotka-Volterra. Més precisament, hom considera l'equació diferencial

$$dX_i(t) = [u_i(X_1(t), \dots, X_n(t)) + \gamma_i v_i(X_1(t), \dots, X_n(t))] dt,$$

$i = 1, 2, \dots, n$, on γ_i són paràmetres associats al medi. Si se substitueixen aquests paràmetres per un soroll correlacionat, s'obté una equació diferencial estocàstica com

$$\begin{aligned} dX_i(t) &= [u_i(X_1(t), \dots, X_n(t)) + \bar{\gamma}_i v_i(X_1(t), \dots, X_n(t))] dt \\ &+ \sigma v_i(X_1(t), \dots, X_n(t)) dW_i(t), \end{aligned} \tag{15}$$

on W_i , $i = 1, 2, \dots, n$ és un procés gaussià correlacionat i $\bar{\gamma}_i$ representa el valor mig d'aquest procés.

Com s'explica en el treball citat de Dawson i Salehi, l'avantatge del model definit per (14), respecte del definit per (15), és que permet de trobar una distribució estacionària (situació d'equilibri), quan la dimensió d és més gran o igual a 3. L'existència d'aquesta situació d'equilibri concorda amb els resultats empírics, però el model (14) no el dona.

6.3. Ones estocàstiques

L'estudi matemàtic de l'equació té el seu origen al segle XVIII, amb els treballs de Johann Bernoulli sobre el moviment de la corda vibrant. Curiosament, aquest problema

va produir debats científics de molta profunditat i riquesa motivats per arguments contraposats de Daniel Bernoulli, fill de Johann, Euler, d'Alembert, Lagrange i, finalment, Fourier. De fet, el concepte de *funció*, potser el concepte més crucial de les matemàtiques, va néixer, en part, provocat per aquest debat entorn de les ones (vegeu [L]).

L'equació de la corda vibrant és la següent:

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(t, x) - c \frac{\partial^2 u}{\partial x^2}(t, x) &= \phi(t, x), \\ u(0, x) &= f(x), \\ \frac{\partial u}{\partial t}(0, x) &= g(x), \\ u(t, 0) = u(t, L) &= 0. \end{aligned} \tag{16}$$

on $0, L$ representen les abscisses dels punts on hem fixat la corda, x l'abscissa d'un punt arbitrari de la corda i $u(t, x)$ el desplaçament vertical a l'instant t ; la funció f dona la posició inicial i g la velocitat inicial. Altres fenòmens físics, com el moviment de sòlids elàstics, la propagació d'ones sonores en diversos medis, admeten descripcions similars, adaptant les condicions inicials i/o a la vora.

Com descriure la vibració d'una corda de guitarra d'un *hippy* tocant en el desert enmig d'una tempesta de sorra? Lògicament, la vibració es veurà afectada pels impactes constants dels grans de sorra damunt les cordes. Aquests impactes, com els impulsos en les neurones, no arriben d'una forma *previsible*, ben determinada. Arriben a *l'atzar*. Tenim, doncs, un altre exemple en el qual un model com el (16), en el qual afegirem una pertorbació de caràcter aleatori, sembla més convenient. L'equació d'ones estocàstica en \mathbb{R}^d (sense condicions a la vora) és del tipus següent:

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2}(t, x) - \frac{\partial^2 u}{\partial x^2}(t, x) &= \sigma(u(t, x))\dot{F}(t, x) + b(u(t, x)), \\ u(0, x) &= f(x), \\ \frac{\partial u}{\partial t}(0, x) &= g(x), \end{aligned} \tag{17}$$

on σ, b són funcions reals i $\dot{F}(t, x)$ representa la derivada formal d'un procés aleatori com el que hem definit en l'apartat 6.1.

Per simplificar l'exposició, suposarem $f, g \equiv 0$. Seguint l'aproximació de D'Alembert (vegeu també (8) per a l'equació de la calor) es proposa com a solució de (17) un procés estocàstic $\{u(t, x), t \geq 0, x \in \mathbb{R}^d\}$ que compleix

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} S(t-s, x-y) [\sigma(u(s, y)) F(ds, dy) + b(u(s, y)) ds dy], \quad (18)$$

on $S(t, x)$ és la solució fonamental de l'equació d'ona determinista. També aquí l'existència de la integral estocàstica en (18), com una variable aleatòria real, depèn de la correlació espacial del procés $F(t, x)$ i de la dimensió. En efecte, la funció $S(t, x)$ té una expressió que varia amb la dimensió. Concretament,

$$\begin{aligned} S(t, x) &= \frac{1}{2} 1_{\{|x| \leq t\}}, \quad d = 1, \\ S(t, x) &= C_d \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{\frac{d-2}{2}} (t^2 - |x|^2)^{\frac{1}{2}}, \quad d \geq 2, \quad d \text{ parell}, \\ S(t, x) &= C_d \left(\frac{1}{t} \frac{\partial}{\partial t} \right)^{\frac{d-3}{2}} \frac{\sigma_t^d}{t}, \quad d \geq 3, \quad d \text{ imparell}, \end{aligned} \quad (19)$$

on σ_t^d denota la mesura uniforme sobre la superfície de l'esfera d -dimensional de radi t . Suposem, per simplificar, que σ és una constant. La condició

$$\int_0^t \int_{\mathbb{R}^d} S(t-s, x-y)^2 ds dy < \infty,$$

es compleix si, i només si, $d = 1$. La conclusió és la mateixa que per a l'equació de la calor. Podrem utilitzar un soroll aleatori incorrelacionat en temps i en espai (soroll blanc en temps i espai) si, i únicament si, $d = 1$. En dimensions superiors haurem de considerar processos aleatoris correlacionats en espai amb una funció de correlació f (o, equivalent, una mesura espectral μ que, d'acord amb (12), compleixi

$$\int_{\mathbb{R}^d} \mu(d\xi) (\mathcal{F}S(t, \cdot)(\xi))^2 < \infty. \quad (20)$$

La transformada de Fourier de $S(t, x)$ té una expressió independent de la dimensió:

$$\mathcal{F}S(t, \cdot)(\xi) = \frac{\sin(2\pi t |\xi|)}{2\pi |\xi|}.$$

Es pot demostrar fàcilment que (20) es compleix si, i només si,

$$\int_{\mathbb{R}^d} \frac{\mu(d\xi)}{1 + |\xi|^2} < \infty,$$

que és la mateixa condició que havíem trobat per a l'equació de la calor.

El lector interessat en exemples més seriosos de l'equació d'ones estocàstica que els que hem presentat més amunt, i on es posa en evidència l'interès de tractar el cas d -dimensional, pot consultar el treball d'oceanografia de [M]. L'objectiu del treball de Miller és proporcionar un model matemàtic per simular la superfície marina en zones tropicals de l'oceà Pacífic. La necessitat d'acudir a la modelització es deu al fet que l'obtenció de dades suficients per a la descripció de tal superfície necessitaria la presència de moltes estacions de recollida de dades situades estratègicament, la qual cosa és costosa i difícil. El model proposat és el d'una equació com (18) amb el coeficient σ constant. Els resultats de l'anàlisi estadística dels errors comesos amb el model, comparats amb les dades reals, en els casos en què es pot disposar d'aquestes, resulten ser molt satisfactoris.

7. CONCLUSIONS

En comparació amb els models deterministes, mitjançant els models d'evolució aleatoris tenim la possibilitat d'incorporar més informació sobre les causes que produeixen el fenomen que estudiem i també sobre les condicions en què es realitza l'evolució. Això els dona un caràcter més realista i, segons es desprèn d'una anàlisi acurada comparativa, permet arribar a conclusions més raonables i menys discrepants amb la realitat que observem.

El fonament matemàtic d'aquests models el constitueix una teoria, la de la probabilitat, que, inicialment, es va desenvolupar motivada per problemes de pura curiositat científica. Un pur joc: l'anàlisi dels jocs!

Actualment, en l'aplicació de tècniques probabilistes i en la pròpia recerca en probabilitats hi ha un fort component d'altres matèries de la matemàtica. D'altra banda, les dues vessants, creació i aplicació, es mouen en un context molt interdisciplinar. Aquest procés de creació i projecció científica s'ha repetit moltes vegades en les ciències fonamentals. Tornant a citar l'article periodístic de Pi de Cabanyes,

... A veces se dan en la vida casualidades que dejan a uno maravillado.

Han estat casualitats? En un entorn social en el qual hi ha una clara tendència a primar els resultats tangibles, cal no menystenir l'atzar.

REFERÈNCIES

- [D-Z] G. Da Prato; J. Zabczyk: *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992.
- [D-S] D. A. Dawson; H. Salehi: «Spatially homogeneous random evolutions». *Journal of Multivariate Analysis* 10, 141-180 (1980).
- [D] R. Dalang: «Extending the martingale measure stochastic integral with applications to spatially homogeneous spde's». *Electronic Journal of Probability*, 4, 1-29 (1999). <http://www.math.washington.edu/ejpecp/>.
- [I1] K. Itô: «Stochastic integrals». *Proc. Imp. Acad. Tokyo* 20, 519-524 (1944).
- [I2] K. Itô: «Stochastic differential equations». *Memoirs A.M.S.* 4, 1951.
- [K-S] I. Karatzas; S. E. Shreve: *Brownian Motion and Stochastic Calculus*. Springer-Verlag, Berlin, Heidelberg, New York, 1988.
- [L] N. Luzin: «The evolution of the notion of function: Part I». *Amer. Math. Monthly*, 59-67 (1998).
- [M] L. E. Maistrov: *Probability Theory: A Historical Sketch*. Academic Press, New York, London, 1974.
- [M] R. N. Miller: «Tropical data assimilation experiments with simulated data: The impact of the tropical ocean and global atmosphere thermal array for the ocean». *Journal of Geophysical Research* 95, C7, 11461-11482 (1990).
- [R-Y] D. Revuz; M. Yor: *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, Heidelberg, New York, 1999.
- [T] M. Turelli: «A reexamination of stability in randomly varying versus deterministic environments with comments on the stochastic theory of limiting similarity». *Theor. Popul. Biology* 13, 244-267 (1978).
- [W1] J. B. Walsh: «A stochastic model of neural response». *Adv. Appl. Prob.* 13, 231-281 (1981).
- [W2] J. B. Walsh: «An introduction to stochastic partial differential equations». In: *École d'Été de Probabilités de Saint-Flour XIV-1984*. Lecture Notes in Mathematics 1180. Springer-Verlag, Berlin, Heidelberg, New York, 1986.
- [Wi] N. Wiener: «The homogeneous Chaos». *Amer. J. Math.* 60, 897-936 (1930).



Cada vegada més, els científics troben camins de col·laboració entre especialitats ben diverses. En matemàtiques, aquesta tendència ha anat unida a un interès creixent per fer-les més accessibles i implicar-les a tot arreu on sigui necessari. Quan es parla de models matemàtics, cal pensar en solucions ben estructurades de problemes nascuts en altres branques de la ciència o en la indústria, o bé en eines creades per fer front a aquests problemes. Sovint les tècniques de càlcul numèric o de simulació basada en equacions diferencials hi juguen un paper important.

Amb la publicació d'aquestes conferències s'ha volgut oferir una mostra variada i substancial d'allò que s'està fent en modelització. No s'ha pretès (ni es podia pretendre) reunir exemples de totes les línies d'investigació, que són moltes i molt diverses. Alguns articles descriuen avenços recents i d'altres són de caire més general, fins i tot amb una clara intenció didàctica. Tots s'adrecen a un públic ampli i, en general, a tots aquells que saben admirar la grandesa i les immenses possibilitats de les matemàtiques.

