

mat

MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

www.madrimasd.org

colec*ción* m i + d

IV PRICIT 2005-2008
Plan Regional de Ciencia y Tecnología
de la Comunidad de Madrid

MM
La Suma de Todos

CONSEJERÍA DE EDUCACIÓN
Comunidad de Madrid
www.madrid.org



instituto madrileño
de estudios avanzados
imdea
matemáticas

mat

MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

COORDINADORES

Manuel de León Rodríguez - CSIC
José Luis González Llanova - UCM
L. Alberto Ibort Latre - UC3M
Enrique Zuazua Iriondo - UAM

DIRECCIÓN TÉCNICA

Miguel Ángel Benítez

AUTORES

Daniel Azagra Rueda
Alfredo Bermúdez de Castro
Florentino Borondo Rodríguez
Santiago Carrillo Menéndez
Vicent Caselles Costa
Jorge Cortés Monforte
Oscar García Prada
Francisco Guinea López
Marco Antonio López Cerdá
Juan José López Velázquez
Ana María Mancho Sánchez
Froilán Martínez Dopico
Sonia Martínez Díaz
Alejandro Melle Hernández
Fernando Monge Gómez
Francisco Palacios Gutiérrez
Daniel Peña Sánchez de Rivera
David Ríos Insua
Juan Manuel Rodríguez Parrondo
Ángel Sánchez Sánchez
Rafael Sendra Pons
Carles Simó Torres
Ana Vargas Rey
Orlando Villamayor Uriburu



Biblioteca Virtual

CONSEJERÍA DE EDUCACIÓN
Comunidad de Madrid

Esta versión digital de la obra impresa forma parte de la Biblioteca Virtual de la Consejería de Educación de la Comunidad de Madrid y las condiciones de su distribución y difusión de encuentran amparadas por el marco legal de la misma.

www.madrid.org/edupubli

edupubli@madrid.org

EDITA

Comunidad de Madrid
Consejería de Educación
Dirección General de Universidades e Investigación

DISEÑO

base12 diseño y comunicación, s.l.

IMPRIME

Elecé Industria Gráfica, s.l.

DEPÓSITO LEGAL

M-40.580-2007

mat

MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

SUMARIO

PREFACIO.....	6
capítulo I INTERRELACIÓN ENTRE ANÁLISIS, GEOMETRÍA Y ECUACIONES EN DERIVADAS PARCIALES	10
Daniel Azagra Rueda - Universidad Complutense de Madrid	
capítulo II MATEMÁTICAS E INDUSTRIA	24
Alfredo Bermúdez de Castro - Universidade de Santiago de Compostela	
capítulo III QUÍMICA TEÓRICA Y COMPUTACIONAL.....	32
Florentino Borondo Rodríguez - Universidad Autónoma de Madrid	
capítulo IV MATEMÁTICA FINANCIERA	40
Santiago Carrillo Menéndez - Universidad Autónoma de Madrid	
capítulo V LAS MATEMÁTICAS Y EL PROCESAMIENTO DE IMÁGENES	48
Vicent Caselles Costa - Universitat Pompeu Fabra	
capítulo VI MATEMÁTICAS Y FÍSICA TEÓRICA	58
Oscar García Prada - Consejo Superior de Investigaciones Científicas	
capítulo VII MATEMÁTICAS Y MODELIZACIÓN EN CIENCIA DE MATERIALES	66
Francisco Guinea López - Consejo Superior de Investigaciones Científicas	
capítulo VIII LA OPTIMIZACIÓN Y EL MÉTODO CIENTÍFICO EN LA TOMA DECISIONES...	72
Marco Antonio López Cerdá - Universitat d'Alacant	
capítulo IX MATEMÁTICAS Y BIOLOGÍA: ALGUNOS PUNTOS DE CONTACTO.....	82
Juan José López Velázquez - Universidad Complutense de Madrid	
capítulo X MECÁNICA DE FLUIDOS COMPUTACIONAL Y APLICACIONES.....	94
Ana María Mancho Sánchez- Consejo Superior de Investigaciones Científicas	
capítulo XI MATEMÁTICA COMPUTACIONAL: UN NUEVO PILAR PARA EL DESARROLLO CIENTÍFICO Y TECNOLÓGICO	102
Froilán Martínez Dopico - Universidad Carlos III de Madrid	

capítulo XII		
MATEMÁTICAS, CONTROL Y ROBÓTICA	116	
		Sonia Martínez Díaz y Jorge Cortés Monforte -U. California, San Diego y Santa Cruz
capítulo XIII		
MATEMÁTICAS Y CRIPTOGRAFÍA	124	
		Alejandro Melle Hernández - Universidad Complutense de Madrid
capítulo XIV		
NECESIDADES FUTURAS EN INVESTIGACIÓN AERODINÁMICA	132	
		Fernando Monge Gómez y Francisco Palacios Gutiérrez Instituto Nacional de Técnica Aeroespacial
capítulo XV		
MATEMÁTICAS Y ECONOMÍA	150	
		Daniel Peña - Universidad Carlos III de Madrid
capítulo XVI		
MEJORES DECISIONES PARA UN MUNDO MEJOR.....	156	
		David Ríos Insua - Universidad Rey Juan Carlos y Real Academia de Ciencias
capítulo XVII		
MATEMÁTICAS Y NEUROCIENCIAS	162	
		Juan Manuel Rodríguez Parrondo - Universidad Complutense de Madrid
capítulo XVIII		
SISTEMAS COMPLEJOS: LA CIENCIA DEL SIGLO XXI	172	
		Ángel Sánchez Sánchez - Universidad Carlos III de Madrid
capítulo XIX		
CURVAS Y SUPERFICIES ALGEBRAICAS: COMPUTACIÓN Y APLICACIONES..	182	
		Rafael Sendra Pons - Universidad de Alcalá
capítulo XX		
SISTEMAS DINÁMICOS.....	192	
		Carles Simó Torres - Universitat de Barcelona
capítulo XXI		
ANÁLISIS ARMÓNICO: PERSPECTIVAS Y APLICACIONES	198	
		Ana Vargas Rey - Universidad Autónoma de Madrid
capítulo XXII		
MATEMÁTICAS Y SOCIEDAD DE LA INFORMACIÓN	208	
		Orlando Villamayor Uriburu - Universidad Autónoma de Madrid

Prefacio

LOS COORDINADORES

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Según Galileo Galilei, las Matemáticas son el lenguaje en el que está escrito el universo y, por tanto, constituyen el instrumento mediante el cual podemos comprender los fenómenos naturales, analizando y buscando pautas entre millones de datos, construyendo modelos que simulen la realidad permitiéndonos conocerla y aprender a controlarla a nuestro servicio. Podríamos decir que las Matemáticas son, en este sentido, una tecnología clave.

En un contexto histórico, la evolución de las Matemáticas ha seguido un camino lento y no lleno de dificultades, en ocasiones por la propia naturaleza de éstas y, en otras, por la necesidad de recursos que en la época eran inimaginables. De forma exponencial, las Matemáticas han experimentado un profundo cambio en las últimas décadas, debido fundamentalmente a la aparición de los ordenadores como instrumentos cotidianos. La actual potencia de cálculo de cualquier ordenador de mesa o portátil los ha convertido en aliados perfectos en las tareas de los matemáticos. Simulaciones que eran imposibles hace no muchos años están ahora al alcance de cualquiera, con multitud de paquetes informáticos accesibles. Para simulaciones mucho más complejas, los grandes ordenadores son la solución.

Estos cambios han colocado a las Matemáticas en un nuevo paradigma. Ya no se usa sólo el lápiz y el papel; a la capacidad del pensamiento añadimos ahora la capacidad de cálculo. Y son estas nuevas condiciones las que permiten la interacción de las Matemáticas con el sistema de I+D+i y sus aportaciones a la Sociedad de la Información.

Hasta hace muy poco, los matemáticos españoles y en particular los madrileños no han comenzado a adentrarse en este nuevo mundo que se abre por delante. Para conseguir una auténtica transferencia del conocimiento matemático a los sectores tecnológicos, productivos y financieros, los matemáticos necesitan ser conscientes de las nuevas oportunidades. Esta recopilación de artículos va en esa dirección: matemáticos de gran valía han reflexionado sobre sus respectivos campos de trabajo, intentando dar una panorámica de las Matemáticas del siglo XXI, identificando los grandes problemas y mostrando los campos que están esperando para su desarrollo. En definitiva, mostrando el futuro de esta apasionante ciencia que son las Matemáticas.

Esta colección de artículos acompaña al informe que la Comunidad de Madrid solicitó a la Universidad Autónoma de Madrid sobre la posible creación de un Instituto de Investigación en Matemáticas en la Región Madrileña, y en el que los editores de este documento han creado un Grupo de Trabajo encargado de su redacción. Este Instituto constituye un nuevo instrumento imprescindible para poder afrontar nuevos problemas en el campo de la transferencia de la investigación. Ambas iniciativas se complementan así de una manera natural.

Finalmente, queremos agradecer a los autores por su disponibilidad y excelente trabajo panorámico. También a la Comunidad de Madrid y a la Universidad Autónoma de Madrid por darnos la oportunidad de afrontar este ejercicio prospectivo. Esperamos que el mismo sea útil tanto para los gestores en su tarea de toma de decisiones, como para los matemáticos y científicos en general que quieran acercarse a las fronteras de las matemáticas y algunos de sus retos futuros más estimulantes en su interacción con otras ciencias y con el tejido I+D+i en general.



Capítulo I

INTERRELACIÓN ENTRE ANÁLISIS, GEOMETRÍA Y ECUACIONES EN DERIVADAS PARCIALES

POR DANIEL AZAGRA RUEDA
UNIVERSIDAD COMPLUTENSE DE MADRID
Departamento de Análisis Matemático,
Facultad de Matemáticas
28040 Madrid
Correo electrónico: azagra@mat.ucm.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

In this note we discuss several examples that will hopefully show the fruitful interplay that exists between three areas of mathematics, namely Partial Differential Equations, Differential Geometry, and Analysis. We first give a quick description of what the aims and methods of each area are, and next we present a few examples of how a problem arising in one of these areas can be solved by using results and tools from another area. Among the examples we discuss are the minimization of the area of surfaces with a given boundary, the Poincaré conjecture, the heat equation, and some problems of nonlinear analysis including part of my recent research.

Introducción

Una gran parte de los fenómenos de la física, la química, la biología y otras ciencias, así como multitud de problemas muy relevantes en ingeniería, pueden modelizarse y tratarse mediante el uso de ecuaciones diferenciales (ya sean éstas ecuaciones diferenciales ordinarias, EDOs, o en derivadas parciales, EDPs). Una ecuación en derivadas parciales es una ecuación que depende, en general, de varias variables numéricas x_1, x_2, \dots, t (que pueden representar, por ejemplo, dimensiones espaciales o temporales), así como de una función u y de sus derivadas parciales, hasta un orden m , respecto de dichas variables,

$$\frac{\partial^k u}{\partial x_{j_1} \dots \partial x_{j_k}}, \quad k \leq m.$$

Por ejemplo, la ecuación de Schrödinger que gobierna el movimiento de una partícula en la mecánica cuántica es la siguiente ecuación en derivadas parciales:

$$(1) \quad \frac{\partial u}{\partial t} = i\Delta u,$$

donde

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

(la función $|u(t, x, y, z)|^2$ nos da la densidad de probabilidad de encontrar a la partícula en el instante t y en la posición (x, y, z) ; es decir, si E es una región del espacio \mathbb{R}^3 entonces $\int_E |u(t, x, y, z)|^2 \partial x \partial y \partial z$ es la probabilidad, en el instante t , de encontrar a nuestra partícula en la región E).

Otro ejemplo muy familiar es la ecuación del calor

$$(2) \quad \frac{\partial u}{\partial t} = \Delta u$$

cuya solución $u(t, x, y, z)$ idealmente nos daría la temperatura en el instante de tiempo t y la posición (x, y, z) . A las ecuaciones de este tipo, para que la solución pueda estar determinada, se les suele añadir una condición inicial, que no es más que una función $u_0(x, y, z)$ que describe la situación en el instante inicial $t = 0$, es decir, se supone que $u(0, x, y, z) = u_0(x, y, z)$. Un tercer ejemplo clásico es el problema de Dirichlet

$$(3) \quad \Delta u = 0 \text{ en } \Omega, \text{ y } u = g \text{ en } \partial\Omega,$$

donde Ω es una región del espacio, y $\partial\Omega$ es su borde; esta ecuación aparece, por ejemplo, como el estado de equilibrio de la función de temperatura $u(t, x, y, z)$ de (2), esto es $u(x, y, z) = \lim_{t \rightarrow \infty} u(t, x, y, z)$, bajo la suposición de que en el borde de Ω la temperatura $u(t, x, y, z)$ se mantiene constante en el tiempo (por acción, por ejemplo, de un sistema calefactor y refrigerador). Conviene observar que en (3) no se especifica una condición inicial, sino una condición de frontera, es decir, se presupone conocida la solución del problema en el borde del recinto donde trabajamos y a partir de este dato y de la ecuación que gobierna el problema se espera poder encontrar la solución en cualquier punto del interior del recinto.

La geometría diferencial es la parte de la geometría que se ocupa de estudiar las propiedades locales y globales de las curvas, las superficies y, en general, las superficies de n dimensiones que son suaves, esto es, localmente aproximables por superficies *planas* de la misma dimensión. A estas superficies suaves se les llama *variedades diferenciables*. Cuando en cada espacio tangente a cada punto de una de estas variedades hay definido un producto escalar de modo que al variar el punto este producto escalar varía con regularidad, se dice que la variedad tiene una estructura riemanniana. En las variedades de Riemann la distancia entre dos puntos viene dada por la longitud de *curvas geodésicas*, que son las que tienen longitud mínima de entre todas las curvas que unen dichos puntos dentro de la variedad. Esta disciplina tiene también, por supuesto, muchas aplicaciones fuera de la propia matemática: quizás una de las más famosas sea proporcionar a los físicos el marco en el que la teoría general de la relatividad se desarrolla (en esta teoría se supone que el universo es una superficie tetradimensional curva, es decir, una variedad diferenciable de dimensión cuatro, con una métrica semiriemanniana).

Finalmente, el análisis se ocupa del estudio de las propiedades de las funciones y de los espacios de funciones, así como de las operaciones y operadores entre estos objetos. Por ejemplo, el análisis armónico estudia entre otras cosas cómo expresar una función dada como una suma infinita (o también una integral) de ondas de diversas frecuencias; una de sus muchas aplicaciones es por ejemplo el permitir resolver EDPs como la ecuación de Schrödinger (1). Por su parte, el análisis funcional contempla todas las funciones con alguna propiedad común no de manera aislada, sino dentro de un mismo espacio, y se ocupa de estudiar las propiedades de dicho espacio y de las *funciones de funciones* (esto es, funcionales y operadores) definidos sobre tal espacio. Las funciones pasan así a ser consideradas como *puntos* de un espacio vectorial más grande. Por ejemplo, si se considera el espacio de todas las funciones f definidas sobre un abierto Ω de \mathbb{R}^n cuyo cuadrado es integrable, se obtiene un espacio de Hilbert $L^2(\Omega)$ en el cual la *distancia* entre dos funciones f y g viene dada por $\sqrt{\int_{\Omega} |f - g|^2}$. Este espacio no puede identificarse a ningún \mathbb{R}^m , porque es un espacio vectorial de dimen-



sión infinita: no existe ningún subconjunto finito de funciones de $L^2(\Omega)$ con la propiedad de que cualquier otra función de $L^2(\Omega)$ pueda expresarse como combinación lineal de dichas funciones. Sin embargo, sí que existen bases infinitas: por ejemplo toda función de $L^2(0, 1)$ puede expresarse como una combinación lineal infinita de funciones ondulatorias de la forma $\sin \pi nx$ y $\cos \pi nx$, con $n \in \mathbb{Z}$ (lo que hace que este espacio sea uno de los marcos naturales en los que se desarrolla el análisis armónico). Un ejemplo de un problema típico del análisis funcional es el de encontrar el mínimo de un funcional (función de funciones) definido sobre un espacio de Hilbert, con ciertas restricciones (en efecto, puede demostrarse que resolver el problema (3) equivale a minimizar un funcional definido sobre un cierto espacio de Hilbert; ver el primer ejemplo de la siguiente sección).

Más en general, los espacios de Banach son espacios de funciones de dimensión infinita en los que la distancia entre dos funciones viene definida por una *norma* que no necesariamente proviene de un producto escalar, y los espacios de Fréchet (esenciales para desarrollar rigurosamente la *teoría de distribuciones*, que son *funciones generalizadas* que permiten resolver problemas en los que los datos o soluciones no son lo suficientemente regulares, y a su vez definir rigurosamente espacios como los de Sobolev que permiten tratar muchos problemas de EDPs), son espacios de dimensión infinita en los que la distancia entre dos de sus elementos viene dada por una familia infinita de *seminormas*. Tampoco los espacios de Hilbert y de Banach son objetos cuyo interés se reduzca al ámbito puramente matemático: por ejemplo, el espacio de Hilbert es una herramienta esencial para formular y tratar matemáticamente la teoría física de la mecánica cuántica, y los espacios de Banach tienen importantes aplicaciones en economía y otras ciencias.

El propósito de este artículo es mucho más modesto de lo que su título sugiere, y que daría para una obra de varios volúmenes. No pretendo dar una visión panorámica del estado actual de las investigaciones matemáticas que se encuentran en la frontera del análisis, la geometría diferencial y las EDPs. Sencillamente deseo poner de manifiesto, para un lector con una buena formación matemática, pero no necesariamente matemático profesional, la profunda interrelación que existe entre estas áreas de las matemáticas, para lo cual discutiremos varios ejemplos que muestran cómo los problemas de cada una de estas disciplinas motivan nuevos problemas relacionados en otra de las disciplinas y cómo también un avance en cualquiera de estas áreas significa una mejora en las otras. Por supuesto la selección de ejemplos es forzosamente muy limitada, y en la exposición incurriré a menudo en simplificaciones abusivas.

Ejemplos y discusión

Comencemos por explicar la relación entre dos de los ejemplos clásicos citados en la introducción.

El problema de Dirichlet. Una solución y una aplicación.

El problema de Dirichlet (3), que como ya comentamos más arriba puede describir el estado de equilibrio (en tiempo infinito) de una solución de la ecuación del calor que es independiente del

tiempo en la frontera del recinto Ω , puede resolverse con ayuda del análisis funcional como sigue. En primer lugar, se observa que el problema (3) equivale a resolver el siguiente problema:

$$(4) \quad \Delta u = f \text{ en } \Omega, \text{ y } u = 0 \text{ en } \partial\Omega.$$

En efecto, si sabemos resolver este problema, entonces elegimos G una función definida en $\overline{\Omega}$ tal que $G = g$ en $\partial\Omega$, definimos $f = -\Delta G$, y encontramos una (única) función u que resuelve (4). Entonces la función $u+G$ resuelve (3) (y es la única función que lo hace). En segundo lugar, no es difícil demostrar, usando integración por partes, que (4) equivale a resolver el problema de minimización

$$(5) \quad \min_{u=0 \text{ en } \partial\Omega} T(u)$$

donde $T: H \rightarrow \mathbb{R}$ se define por

$$T(u) = \int_{\Omega} \left(\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial u}{\partial z} \right)^2 + 2u(x,y,z)f(x,y,z) \right) dx dy dz$$

ya que $H = H^1(\mathbb{R}^3)$ es el espacio de Sobolev de las funciones cuyas derivadas parciales de primer orden están en $L^2(\mathbb{R}^3)$ (el espacio $H^1(\mathbb{R}^3)$ es otro espacio de Hilbert).

La restricción $u = 0$ en $\partial\Omega$ de los problemas (4) y (5) puede interpretarse del siguiente modo: $u \in H_0^1(\Omega)$, donde $H_0^1(\Omega)$ es el conjunto de todas las funciones que son límites en $H^1(\mathbb{R}^3)$ de funciones diferenciables de orden dos y que se anulan fuera de Ω . De esta manera $H_0^1(\Omega)$ es un subespacio cerrado del espacio $H^1(\mathbb{R}^3)$ y es por tanto también un espacio de Hilbert, y los problemas (4) y (5) son ambos equivalentes a

$$(6) \quad \min_{u \in H_0^1(\Omega)} T(u)$$

Puede verse fácilmente que el funcional T es estrictamente convexo, y está acotado inferiormente. Por su parte, el espacio $H_0^1(\Omega)$ es un espacio de Hilbert, y estos espacios tienen la propiedad de que todo funcional estrictamente convexo acotado inferiormente alcanza su mínimo en un único elemento. Se concluye así que existe una única solución al problema de Dirichlet (3). Ver [22] para más detalles.

En este ejemplo vemos cómo el análisis funcional resuelve un problema de EDPs que tiene una interpretación física concreta. El procedimiento seguido (relacionar un problema de EDPs con la minimización de un funcional) es típico del *cálculo de variaciones*, disciplina en la frontera del análisis funcional y las EDPs.

Pero, a continuación, veamos también cómo la solución de este problema de EDPs permite a su vez resolver un problema puramente matemático, de análisis complejo.



Sea Ω un recinto abierto y acotado del plano complejo $\mathbb{C} = \mathbb{R}^2$ con una frontera regular $\partial\Omega$. Supóngase que Ω es simplemente conexo, de modo que $\partial\Omega$ es difeomorfo a una circunferencia. El teorema de la aplicación de Riemann dice que, dado cualquier $p \in \Omega$, es posible encontrar una aplicación holomorfa Φ y con inversa holomorfa que transforma $\overline{\Omega} = \Omega \cup \partial\Omega$ en el disco unidad cerrado $D = \{z \in \mathbb{C} : |z| \leq 1\}$ y que lleva el punto p al punto 0. Este importante resultado de análisis complejo puede demostrarse apoyándose en el problema de Dirichlet como sigue. Por la discusión anterior sabemos que existe una única solución $u(z) = u(x, y)$ al problema de Dirichlet siguiente:

$$(7) \quad \Delta u = 0 \text{ en } \Omega, \text{ y } u = -\log |z - p| \text{ en } \partial\Omega,$$

Consideremos la función $v : \Omega \rightarrow \mathbb{R}$ definida por

$$v(z) = \int_p^z \left(-\frac{\partial u}{\partial y} dx + \frac{\partial u}{\partial x} dy \right)$$

donde la integral se realiza a lo largo de cualquier camino que una p con z dentro de Ω y sea independiente de dicho camino. La función de Riemann Φ se puede definir entonces por

$$\Phi(z) = (z - p) e^{u(z) + iv(z)};$$

puede comprobarse que Φ lleva $\overline{\Omega}$ en \overline{D} y es holomorfa con inversa holomorfa; ver [24, p. 324] para más detalles.

Pasamos ahora a considerar la solución, más reciente, de un problema también clásico, como es el de encontrar la superficie que minimiza el área de entre todas las superficies que tienen un mismo borde.

Las corrientes y el problema de Plateau.

El problema de Plateau es un clásico del cálculo de variaciones: dada una curva cerrada C en el espacio, se trata de encontrar la superficie S de menor área de entre todas las que tienen C como borde, y de saber si esta superficie S es regular (es decir, suave) o si puede tener puntos singulares (es decir, donde la superficie no es regular, porque tenga picos o crestas). La naturaleza proporciona una solución simple para este problema: si sumergimos C en una disolución de jabón, al sacarlo se obtiene una película de jabón cuyo borde es C y cuya área es mínima. Pero, ¿cómo puede describirse matemáticamente este problema?

Más en general puede considerarse el problema consistente en, dado un borde cerrado B de dimensión k (es decir, una *superficie* k -dimensional) cerrada, en un espacio de cualquier dimensión n , encontrar la *superficie* $(k+1)$ -dimensional S de menor volumen $(k+1)$ -dimensional de entre todas las que tienen B como borde y tienen forma parecida (técnicamente, tienen el mismo tipo de homología).

La solución a este problema general puede obtenerse combinando técnicas de geometría diferencial, topología algebraica, análisis funcional y teoría geométrica de la medida (rama del análisis que sirve entre otras cosas para tratar el volumen k -dimensional y otras propiedades de las superficies k -dimensionales que no necesariamente son regulares en todos los puntos). De esta manera puede asegurarse que realmente existe una superficie de área mínima para un borde dado, y que además ésta es regular en *casi todos* sus puntos. Uno de los puntos claves en esta solución es definir una noción de *corriente* (en inglés, *current*), que es una especie de *superficie generalizada* (similar e inspirada por el concepto de función generalizada, o *distribución*, que sirve para definir los espacios de Sobolev considerados más arriba).

Una distribución, o función generalizada, se define como un funcional lineal y continuo definido sobre el espacio vectorial topológico D de todas las funciones infinitamente diferenciables y con soporte compacto (es decir, que son cero fuera de un acotado) en un abierto de un espacio de dimensión n (por ejemplo la famosa *función* delta de Dirac, que suele describirse informalmente como una *función* que vale infinito en un punto x_0 , cero en todos los demás puntos, y tiene integral 1, puede definirse rigurosamente como la distribución que a cada función φ de D le asigna el número $\varphi(x_0)$). La razón para introducir este espacio más grande, y más incómodo, en ciertos sentidos, que el de todas las funciones es no solamente que a veces resulta útil poder manejar datos iniciales y soluciones que no necesariamente son regulares, y conceptos como el de *derivada generalizada*, sino también que, incluso aún cuando todos los datos del problema sean verdaderas funciones y muy regulares, el tratamiento matemático del problema puede resultar mucho más adecuado si se maneja un espacio mayor de funciones generalizadas (porque, por ejemplo, éste puede tener la propiedad, como en el primer ejemplo discutido, de que todo funcional convexo y acotado inferiormente alcance un mínimo, mientras que el más pequeño de las *verdaderas* funciones no tiene esta propiedad), si se encuentra la solución en este espacio, y si luego se ve que en realidad la solución así obtenida es una *verdadera* función.

Pues bien, las *corrientes*, o superficies generalizadas, usadas en la solución de este problema de minimizar áreas de superficies con bordes prefijados, son funcionales lineales y continuos en un espacio vectorial topológico de *formas diferenciales*, y las *corrientes rectificables* son aquellas corrientes a las que se les puede asignar un área (o en general volumen $(k + 1)$ -dimensional). El problema consiste entonces en minimizar el funcional que nos da el *área* sobre el conjunto de todas estas corrientes rectificables, sometido a ciertas restricciones (a saber, que las corrientes tengan el mismo borde y forma equivalente), y demostrar después que la corriente mínima así obtenida es una *verdadera* superficie y es regular en casi todo punto; ver [1, 17].

La conjetura de Poincaré y los trabajos de Perelman.

La conjetura de Poincaré, enunciada por este gran matemático a principios del siglo XX, dice que toda variedad M , compacta y sin borde, de dimensión 3, y que tenga la propiedad de que todo camino cerrado puede ser deformado continuamente (dentro de la propia variedad) a un punto, es home-



omorfa a la esfera de dimensión 3, $S^3 = \{(x, y, z, w) \in \mathbb{R}^4 : x^2 + y^2 + z^2 + w^2 = 1\}$. Muchos de los mejores matemáticos del siglo XX, y las técnicas más potentes desarrolladas por la topología diferencial y algebraica hasta el momento, fracasaron en sus intentos de demostrar la veracidad de esta conjetura, aunque sí pudieron probar que el enunciado análogo para variedades de dimensión mayor o igual que cuatro es verdadero (y el propio Poincaré demostró el enunciado análogo para variedades de dimensión 2). Este problema se considera como uno de los más importantes y difíciles de la matemática actual, y es uno de los siete problemas cuya solución será recompensada por el Instituto Clay de Matemáticas (<http://www.claymath.org/>) con un millón de dólares.

Hace tres años el matemático ruso G. Perelman sorprendió a la comunidad matemática al presentar tres artículos [19, 20, 21] en los que, en parte basándose en trabajos previos de Hamilton y otros matemáticos, afirma que demuestra no solamente la conjetura de Poincaré, sino la *conjetura de geometrización de Thurston*, lo que proporcionaría una clasificación completa de todas las variedades de dimensión tres. Aún no hay acuerdo unánime entre los especialistas acerca de si la demostración de Perelman es enteramente correcta o no, pero sí parece evidente que, incluso si no fuera totalmente correcta, estos artículos constituirían un avance muy importante hacia la resolución definitiva de este problema. Evidentemente, no voy a entrar yo a analizar una demostración de tal complejidad que tres años después aún da qué pensar a los mejores especialistas del tema, pero sí me gustaría señalar un aspecto relevante de ésta: la combinación de técnicas de ecuaciones en derivadas parciales y de geometría diferencial. En efecto, el uno de los ingredientes fundamentales de la prueba de Perelman es la consideración del flujo de Ricci,

$$\frac{d}{dt}g_{ij}(t) = -2R_{ij}$$

una ecuación diferencial que gobierna la evolución, respecto de un parámetro t , de una familia de métricas riemannianas $g_{ij}(t)$ en la variedad tridimensional M ; la idea (extremadamente simplificada) sería que, si M tiene la propiedad de deformación de caminos descrita más arriba, entonces, conforme aumenta t , la métrica evolucionaría hacia una estructura de curvatura constante que resultaría ser la esfera.

Algunos problemas y técnicas del análisis no lineal.

El análisis no lineal puede considerarse como un conjunto de técnicas, a medio camino entre análisis funcional, ecuaciones diferenciales, topología y geometría, que tienen en común el estudio de funciones, funcionales y ecuaciones no lineales. A menudo las ecuaciones diferenciales lineales son sólo una primera aproximación a fenómenos más complejos que en la realidad (entendida ésta, justificada o injustificadamente, como el modelo físico, biológico, económico, etc, que los describe) están gobernados por ecuaciones no lineales (es decir, que involucran operaciones no lineales entre las derivadas parciales de distintos órdenes). El análisis funcional lineal (o sea el que considera funcionales y operadores lineales definidos en y entre espacios de

funciones) resulta adecuado para tratar ecuaciones diferenciales lineales, pero insuficiente para un estudio profundo y más completo de las no lineales.

Muchos problemas de ecuaciones diferenciales admiten formulaciones equivalentes en términos de ecuaciones y problemas de punto fijo en espacios funcionales (a menudo espacios de Banach de dimensión infinita), es decir, problemas del tipo $G(f) = 0$, donde $G : U \subseteq X \rightarrow X$ es un operador no lineal de un subconjunto de un espacio funcional en dicho espacio, o del tipo $H(f) = f$; en realidad ambos son equivalentes en muchas ocasiones (al menos cuando G y H están definidos en todo X : considérese $G(f) = H(f) - f$). Para este tipo de problemas resulta muy útil desarrollar teoremas de existencia de puntos fijos y teoremas que permitan asegurar la existencia de ceros de funciones no lineales, así como teoremas de funciones inversas e implícitas, en el contexto de los espacios de funciones de dimensión infinita. La teoría del punto fijo (para aplicaciones contractivas y no expansivas definidas en espacios de dimensión infinita) y la teoría del grado topológico son dos de los campos del análisis no lineal que ayudan a despejar el camino en el tratamiento de estos problemas. Así, uno de los problemas fundamentales de la teoría del punto fijo es el de encontrar condiciones sobre un espacio X y un subconjunto suyo C para que toda función $H : C \rightarrow C$ no expansiva (o sea que $H(f, g) \leq d(f, g)$) posea un punto fijo (es decir, exista $f_0 \in C$ tal que $H(f_0) = f_0$); ver por ejemplo [25]. Por su parte, la teoría del grado topológico busca métodos para contar el número de ceros (y asegurar la existencia de éstos) de una función $G : U \subset X \rightarrow X$; ver [11]. La teoría del grado topológico en dimensión infinita (a menudo llamada teoría del grado de Brouwer-Leray-Schauder) es una generalización de la teoría del grado de Brouwer en espacios de dimensión finita (que permite dar demostraciones elegantes de teoremas tan importantes como el teorema del punto fijo de Brouwer o el teorema de la curva de Jordan), si bien presenta muchas más limitaciones en su aplicabilidad, porque las funciones entre espacios de dimensión infinita son mucho más difíciles de tratar y presentan propiedades generales bien diferentes de las de las funciones entre espacios finito-dimensionales: así, por ejemplo, sucede que la bola unidad de un espacio de dimensión infinita no tiene la propiedad del punto fijo (es decir, no es cierto que toda aplicación continua de la bola en sí misma tenga un punto fijo) e incluso existe una retracción Lipschitz y diferenciable de ella en su borde [6, 3], y también que cualquier hiperplano del espacio es difeomorfo a toda la esfera unidad [7, 2]. Debido a estas y otras peculiaridades topológicas de la dimensión infinita, la teoría del grado topológico no puede ocuparse de todas las funciones continuas, ni siquiera de las diferenciables, de un espacio en sí mismo, sino solamente de las perturbaciones compactas de la identidad (es decir funciones de la forma $G(f) = f - K(f)$, donde $K : U \subseteq X \rightarrow X$ es una función continua y tal que su imagen $K(U)$ es relativamente compacto); no obstante, en muchos problemas concretos esta limitación es irrelevante y la teoría del grado topológico es suficiente para obtener una solución satisfactoria del problema; ver [11].

Por otro lado, como ya constatamos en la discusión de los ejemplos 1 y 2, los problemas de minimización, es decir, de encontrar el mínimo de un funcional $F : U \subseteq X \rightarrow \mathbb{R}$, quizás con ciertas restricciones, aparecen en muchísimas ocasiones ligados al estudio (por ejemplo existencia y unicidad) de soluciones de ecuaciones diferenciales y problemas variacionales. En los ejemplos 1 y 2 los funcionales que aparecían eran estrictamente convexos y podía asegurarse que alcanza-



ban un único mínimo, ya que los espacios en donde estaban definidos eran espacios de Banach reflexivos. Sin embargo, problemas parecidos en espacios de Banach no reflexivos (espacios que pueden describirse precisamente, en virtud del teorema de James, como en los que *no toda* función convexa y acotada inferiormente en una bola cerrada alcanza su mínimo) no tendrían en general solución exacta.

Más en general, en el contexto de los problemas de minimización, una de las dificultades típicas del paso de los problemas lineales a los no lineales es que los funcionales a minimizar que aparecen ligados a éstos últimos no serán en general convexos ni siquiera cuando el espacio X sea un espacio de Hilbert, con lo que la esperanza de poder minimizarlos es bastante más remota, dado que en general una función $G : X \rightarrow \mathbb{R}$ definida en un espacio de Hilbert, acotada inferiormente, diferenciable, y con la propiedad de que $\lim_{\|f\| \rightarrow +\infty} G(f) = +\infty$, no tiene por qué alcanzar su ínfimo (ni siquiera tener ningún punto crítico, es decir un punto donde la derivada de G se anule, lo que sucedería ciertamente si G tuviera un mínimo). De hecho la topología de su dominio y la forma global de la función no tiene ninguna relevancia en esta carencia de mínimos y puntos críticos, dado que toda función continua $G : U \subset X \rightarrow \mathbb{R}$ definida en un abierto U de un espacio de Hilbert X (de dimensión infinita, por supuesto) puede aproximarse tanto como se quiera por funciones $H : U \rightarrow \mathbb{R}$ infinitamente diferenciables y que no tienen ningún punto crítico [4].

Con cierta frecuencia, en el estudio de este tipo de problemas, uno debe contentarse con encontrar *soluciones aproximadas* con ayuda de *principios variacionales* que, si bien no pueden garantizar la existencia de un mínimo para una función continua y acotada inferiormente $G : X \rightarrow \mathbb{R}$, sí que pueden asegurarla de la función $G - \varphi$, donde φ es una perturbación arbitrariamente *pequeña*, perteneciente a una clase de funciones *adecuadas* para el tratamiento del problema. Así por ejemplo, en el principio variacional de Ekeland (válido para todo espacio métrico completo, ver [16]) las funciones φ tienen forma de conos todo lo achatados que se deseen, mientras que en el principio variacional de Deville-Godefroy-Zizler [12, 13] las funciones φ son muy pequeñas y con derivada también todo lo pequeña que se desee.

Una de las aplicaciones del principio variacional de Ekeland es proporcionar una versión aproximada del teorema de Hopf-Rinow (que dice que cada dos puntos de una variedad riemanniana completa, conexa y de dimensión finita pueden conectarse por una geodésica que minimiza la distancia entre ellos); este teorema es falso en dimensión infinita, pero Ekeland demostró [16, 15] que en toda variedad riemanniana completa y conexa de dimensión infinita, *casi todo* par de puntos pueden conectarse por una geodésica minimizante. Por otro lado, el principio variacional de Deville-Godefroy-Zizler tiene muchas aplicaciones dentro de la geometría de espacios de Banach y permite también demostrar resultados de existencia y unicidad para EDPs del tipo de Hamilton-Jacobi en espacios de dimensión infinita [12, 13].

Otra posibilidad para solventar esta carencia de mínimos y de puntos críticos que en general pueden presentar las funciones en espacios o en variedades de dimensión infinita es la de añadir condiciones (como la de Palais-Smale) sobre las funciones a considerar que impongan la existencia de un mínimo en las situaciones análogas a las que permitirían garantizarlo si el

espacio fuera de dimensión finita. Para esta clase más restringida de funciones en variedades de dimensión infinita puede desarrollarse un análogo de la teoría de Morse en variedades de dimensión finita (la teoría de Morse trata de extraer información sobre la topología de los conjuntos de nivel de una función a partir del análisis de su segunda derivada en los puntos críticos de ésta). Usando esta teoría de Morse para funciones definidas en variedades infinito-dimensionales pudieron demostrarse teoremas como el de que toda variedad riemanniana compacta admite al menos una curva geodésica cerrada no trivial, ver [23].

Un último ejemplo tomado de mi investigación

Sin pretensión alguna de poner en un mismo nivel mi trabajo y el de los grandes matemáticos que participaron en los avances descritos más arriba, deseo terminar este artículo de una manera más personal con unas pocas palabras acerca de mi trabajo más reciente.

Ya hemos hablado de funciones generalizadas, derivadas generalizadas y soluciones generalizadas de EDPs, ligadas a espacios de distribuciones y de Sobolev. Estos espacios son adecuados para tratar multitud de problemas de EDPs, pero en otros (por ejemplo las ecuaciones en derivadas parciales no lineales de tipo Hamilton-Jacobi) ya no resultan ser los más apropiados. En efecto, consideremos la ecuación eikonal

$$(*) \quad \|\nabla u(x)\| = 1 \text{ si } x \in \Omega, \text{ y } u(x) = 0 \text{ si } x \in \partial\Omega,$$

donde Ω es un abierto acotado de \mathbb{R}^n . La ecuación eikonal aparece en relación con problemas de óptica geométrica, ver [24]. Esta ecuación no tiene ninguna solución clásica, ya que toda función u que se anule en el el borde de Ω debe tener un punto crítico x_0 en el interior de Ω y por tanto su derivada en este punto no puede cumplir que $\|\nabla u(x_0)\| = 1$. Recientemente, R. Deville y E. Mathéron han demostrado que, para la bola unidad abierta $\Omega = \{x \in \mathbb{R}^n : \|x\| < 1\}$ existe una función $u : \overline{\Omega} \rightarrow \mathbb{R}$ que es 1-Lipschitz y *diferenciable en todo punto*, y tal que $u(x) = 0$ para $x \in \partial\Omega$, $\|\nabla u(x)\| = 1$ en casi todo punto, y $\nabla u(0) = 1$. Esto indica que un concepto de solución tan ligado a las propiedades de las funciones en casi todo punto como es el de las distribuciones y las funciones de los espacios de Sobolev no es el más adecuado para tratar ecuaciones no lineales como la eikonal (en efecto, parece natural desear que si uno tiene una solución generalizada de esta ecuación que además es diferenciable en todos los puntos entonces satisfaga la ecuación en todos los puntos y no solamente en casi todos). En realidad la única solución *natural* de (*) es la función $u(x) = d(x, \partial\Omega)$ que nos da la distancia de x a la frontera de Ω (y que deja de ser diferenciable en todos los puntos de $\partial\Omega$ en que esta distancia es máxima).

Para las ecuaciones no lineales de Hamilton-Jacobi, Crandall y Lions [8, 9, 10] presentaron, a principio de los años 1980, un nuevo concepto de solución generalizada, el de solución de viscosidad. Este concepto está ligado al de subdiferencial y superdiferencial de viscosidad de una función en un punto, y por tanto a una parte del análisis no lineal que podría llamarse análisis no



regular (*nonsmooth analysis*, en inglés). La subdiferencial (de primer orden) de una función u en un punto x es el conjunto de las derivadas (de primer orden) de todas las funciones φ tales que $u - \varphi$ alcanza un mínimo en x (lo que significa que la gráfica de φ queda por debajo de la de u y ambas se tocan en el punto x); la superdiferencial se define cambiando máximo por mínimo. Se dice entonces que u es *subsolución de viscosidad* de la ecuación eikonal (*) si para toda superdiferencial ζ de u en x se tiene $\|\zeta\| \leq 1$. Se dice que u es *supersolución de viscosidad* de (*) si para toda subdiferencial ζ de u en x es $\|\zeta\| \geq 1$. La función u es *solución de viscosidad* de (*) si es a la vez subsolución y supersolución de viscosidad, y si satisface la condición de anularse en la frontera de Ω . Puede demostrarse que la función distancia al borde de Ω es la única solución de viscosidad de (*). Además, las soluciones de viscosidad sí que tienen la propiedad de que, en los puntos donde son diferenciables, satisfacen la ecuación en el sentido clásico.

En [5], hemos obtenido una extensión del análisis no regular de primer orden (hasta ese momento solamente estudiado para funciones definidas en \mathbb{R}^n y en espacios de Banach) al caso de funciones definidas en variedades riemannianas, tanto de dimensión finita como infinita, y hemos obtenido aplicaciones a la existencia y unicidad de soluciones de viscosidad de ecuaciones de Hamilton-Jacobi en variedades riemannianas que incluyen el resultado de que la función distancia a la frontera de un abierto de una variedad riemanniana es la única solución de viscosidad de la ecuación eikonal en la variedad. A su vez, este resultado, y el concepto de solución de viscosidad, han sido usados recientemente por Matthew J. Gursky y Jeff A. Viaclovsky como herramienta para probar teoremas de geometría diferencial sobre deformaciones de las métricas riemannianas y de EDPs que involucran a éstas, ver [18].

Bibliografía

- [1] Almgren, F. J., Jr.; *Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure*. Ann. of Math. (2) 87 (1968).
- [2] Azagra, D.; *Diffeomorphisms between spheres and hyperplanes in infinite-dimensional Banach spaces*. Studia Math. 125 (1997), no. 2, (179-186).
- [3] Azagra, D.; Cepedello, M.; *Smooth Lipschitz retractions of starlike bodies onto their boundaries in infinite-dimensional Banach spaces*. Bull. London Math. Soc. 33, no. 4, (2001), (443-453).
- [4] Azagra, D.; Cepedello, M.; *Uniform approximation of continuous mappings by smooth mappings with no critical points on Hilbert manifolds*. Duke Math. J. 124 (2004), no. 1, (47-66).
- [5] Azagra, D.; Ferrera, J.; López-Mesas, F.; *Nonsmooth analysis and Hamilton-Jacobi equations on Riemannian manifolds*. J. Funct. Anal. 220 (2005), no. 2, (304-361).
- [6] Benyamini, Y.; Sternfeld, Y.; *Spheres in infinite-dimensional normed spaces are Lipschitz contractible*. Proc. Amer. Math. Soc. 88 (1983), no. 3, (439-445).
- [7] Bessaga, C.; *Every infinite-dimensional Hilbert space is diffeomorphic with its unit sphere*. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 14, (1966), (27-31).

- [8] Crandall, M.G.; Ishii, H.; Lions, P.L.; *User's guide to viscosity solutions of second order to fully nonlinear partial differential equations*. Bull. Amer. Math. Soc. 27, (1992), (1-67).
- [9] Crandall, H.; Lions, P.L.; Viscosity solutions of Hamilton-Jacobi equations, Trans. Amer. Math. Soc. 277, (1983), (1-42).
- [10] Crandall, H.; Lions, P.L.; *Hamilton-Jacobi equations in infinite dimensions, Part I: uniqueness of viscosity solutions*, J. Funct. Anal. 62, (1985), (379-396); *Hamilton-Jacobi equations in infinite dimensions, Part II: existence of viscosity solutions*, J. Funct. Anal. 65, (1986), (368-405); *Hamilton-Jacobi equations in infinite dimensions, Part III*, J. Funct. Anal. 68, (1986), (214-247); *Hamilton-Jacobi equations in infinite dimensions, Part IV: unbounded linear terms*, J. Funct. Anal. 90, (1990), (237-283); *Hamilton-Jacobi equations in infinite dimensions, Part V: B-continuous solutions*, J. Funct. Anal. 97, (1991), (417-465).
- [11] Deimling, K.; *Nonlinear functional analysis*. Springer-Verlag, Berlin, (1985).
- [12] Deville, R.; Godefroy, G.; Zizler, V.; *A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions*. J. Funct. Anal. 111 (1993), no. 1, (197-212).
- [13] Deville, R.; Godefroy, G.; Zizler, V.; *Smoothness and renormings in Banach spaces*. Pitman Monographs and Surveys in Pure and Applied Mathematics, 64. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, (1993).
- [14] Deville, R.; Mathéron, É.; *Infinite games, Banach space geometry and the eikonal equation*, preprint, (2005).
- [15] Ekeland, I.; *The Hopf-Rinow theorem in infinite dimension*. J. Differential Geom. 13, (1978), no. 2, (287-301).
- [16] Ekeland, I.; *On the variational principle*. J. Math. Anal. Appl. 47, (1974), (324-353).
- [17] Federer, H.; *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York, (1969).
- [18] Gursky, M.J.; Viaclovsky, J.A.; *Prescribing symmetric functions of the eigenvalues of the Ricci tensor*, Annals of Math., en vías de publicación.
- [19] Perelman, G.; *The entropy formula for the Ricci flow and its geometric applications*, <http://arxiv.org/abs/math.DG/0211159>.
- [20] Perelman, G.; *Ricci flow with surgery on three-manifolds*, <http://arxiv.org/abs/math.DG/030310>
- [21] Perelman, G.; *Finite extinction time for the solutions to the Ricci flow on certain threemanifolds*, <http://arxiv.org/abs/math.DG/0307245>.
- [22] Rauch, J.; *Partial differential equations*, Graduate Texts in Mathematics, 128. Springer-Verlag, New York, (1991).
- [23] Schwartz, J.T.; *Nonlinear functional analysis*. Notes by H. Fattorini, R. Nirenberg and H. Porta, with an additional chapter by Hermann Karcher. Notes on Mathematics and its Applications. Gordon and Breach Science Publishers, New York-London-Paris, (1969).
- [24] Taylor, M.E.; *Partial differential equations, I, II y III*. Applied Mathematical Sciences, vol. 115, 116 y 117. Springer-Verlag, New York, (1997).
- [25] Zeidler, E.; *Nonlinear functional analysis and its applications. I. Fixed-point theorems*. Translated from the German by Peter R. Wadsack. Springer-Verlag, New York, (1986).



Capítulo II

MATEMÁTICAS E INDUSTRIA

POR ALFREDO BERMÚDEZ DE CASTRO
UNIVERSIDAD DE SANTIAGO DE COMPOSTELA
Departamento de Matemática Aplicada,
Facultad de Matemáticas
CP 15782
Correo electrónico: mabermud@usc.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Modern computers provide us with the possibility of making enormous amounts of mathematical operations in a short time. This fact is now having a great impact in applied sciences and engineering, and thereby in industry.

An example of new technology where computers and mathematics plays a fundamental role is numerical simulation. Indeed, the physical world is governed by laws which can be written as mathematical models. Very often they consist of sets of differential equations which, when solved in a computer by using numerical algorithms, give us the relevant information about the phenomena under study. Thus, one can avoid, at least partially, costly experiments in the laboratory. While most of the models currently used in industrial applications are well known from the 19th century, it was necessary to wait for the arrival of electronic computers, in the middle of the last century, to be able to solve them in general realistic situations as those raised by industrial applications. Numerical simulations are now used to design aircrafts or cars, to create virtual accidents, to asses the environmental impact of human activities, to obtain the fair value of financial derivatives, etc. They are on the basis of so-called Computer Aided Engineering (CAE). In fact, Computer Aided Design (CAD) together with CAE are now unavoidable technologies in modern industry allowing to reduce the time and cost to launch new products onto the market.

In spite of this fact, the presence of computational subjects in the research areas cultivated by the Spanish mathematicians is rather small. While mathematical research in Spain has dramatically increased in the last ten years, its applications to industrial problems are not very significant. On the one hand, industrial managers are not aware of the utility that mathematics can have for their companies since mathematics are considered as an abstract interesting science but without any connection with real problems; generally speaking, society thinks that mathematicians are people trained to teach mathematics or to create new beautiful theories, but not to solve “real life” problems. On the other hand, this situation is obviously related to the lack of tradition in the collaboration between academia and industry in Spain, but also to the way of teaching mathematics in our universities and engineering schools, too far from the needs of applications.

Introducción

La posibilidad que proporcionan las actuales computadoras de hacer cálculos enormes y complicados en tiempos razonables está teniendo un fuerte impacto en las ciencias aplicadas y la ingeniería y, a través de ellas, también en la industria. Sin embargo, a pesar de que las matemáticas constituyen la base conceptual y metodológica de estos cálculos, posiblemente la falta de tradición de la investigación matemática en España sea la causa de que el interés por los temas computacionales haya sido escaso y, en todo caso, no acorde con el sorprendente pro-

greso que el conjunto de la disciplina ha experimentado en nuestro país en las últimas dos décadas. Aún así, cabría señalar que la situación ha mejorado en los últimos años y que en la actualidad existen grupos de investigación en cálculo numérico y modelización, de reconocido prestigio en el contexto internacional.

Aunque las matemáticas son una ciencia abstracta y los objetos que trata tienen este carácter, su desarrollo ha sido impulsado por el deseo de entender la naturaleza. Sin duda es ésta la razón para poder afirmar que las matemáticas son útiles para el sistema productivo, al menos en los dos sentidos siguientes. En primer lugar, el estudio de las matemáticas crea aptitudes genéricas como son la capacidad de análisis, la precisión en el lenguaje, el rigor, etc. En segundo lugar, las matemáticas sirven para resolver problemas tecnológicos, y para ayudar en la toma de decisiones en los ámbitos de la planificación económica y financiera, de la organización de la producción, etc.

Además, las matemáticas también son útiles para mejorar la calidad de vida de las personas y resolver problemas ligados a la conservación del medio ambiente. Para sustentar esta afirmación basta decir que constituyen la base de muchos dispositivos utilizados habitualmente en el diagnóstico médico o la cirugía, que permiten predecir el tiempo y analizar la evolución del clima, o que ayudan a evaluar el impacto ambiental producido por la actividad humana.

Sin embargo, la percepción común en nuestro país es que saber matemáticas sólo sirve para enseñarlas a otros o para crear nuevas matemáticas (investigación pura). Un reflejo de esta situación es que el número de matemáticos que trabajan en las empresas es muy escaso, y además la mayoría de ellos realizan tareas de informática de gestión, casi siempre bastante alejadas de las matemáticas. Algunas causas de esta situación son de carácter general y comunes a otras disciplinas: la falta de tradición científica y tecnológica, la escasa inversión de las empresas en investigación y desarrollo, el hecho de que la mayoría de la población desconoce la importancia que tienen en su bienestar los avances científicos, la todavía escasa colaboración entre los centros de investigación y el sistema productivo... No obstante, otras son particulares de las matemáticas: una orientación de los planes de estudio de las Facultades muy polarizada hacia la matemática pura y la abstracción excesiva, la imagen social de que la carrera de matemáticas es sólo para la docencia o la investigación básica, estarían entre ellas. A este respecto, resulta sintomático, además de muy negativo, que en las Escuelas Técnicas las matemáticas hayan sido, y todavía continúen siendo, un elemento de selección de los estudiantes en lugar de una herramienta fundamental para resolver los problemas de la ingeniería moderna.

La simulación numérica

El mundo físico se rige por leyes que se formulan con modelos matemáticos; por este motivo, los modelos permiten simular el comportamiento de dispositivos o procesos de interés industrial. Generalmente, están constituidos por complicados sistemas de ecuaciones cuyas incógnitas son



las magnitudes que caracterizan los fenómenos a simular. Si los ejemplos en la industria son innumerables (cálculo de las tensiones en una estructura, de la temperatura de una pieza, del comportamiento electromagnético de un dispositivo, del movimiento de los gases en el interior de una caldera, del ruido en un recinto, etc.), éstos en modo alguno agotan el universo de las aplicaciones: la predicción del tiempo, la evaluación del impacto ambiental, la simulación de un accidente natural,... o, en un terreno bien diferente, la valoración de los productos financieros, el análisis del riesgo y la gestión de carteras son tareas que hoy día se llevan a cabo basándose en modelos matemáticos que pueden resolverse con ayuda del ordenador.

Los modelos que rigen la mayoría de los procesos que interesan en ingeniería se conocen desde finales del siglo XIX. Sin embargo, hasta la aparición de los ordenadores, su resolución sólo era posible en casos muy particulares generalmente de carácter “académico”. Los ordenadores, junto con los métodos de cálculo numérico desarrollados en los últimos 50 años por los matemáticos, permiten hoy día resolver modelos complejos en tiempos de cálculo aceptables.

La metodología de la simulación numérica comprende diversas etapas. En primer lugar debe procederse a un análisis preciso y detallado de los fenómenos que caracterizan el proceso que se desea simular, para conocer su naturaleza. Después vendrá la construcción de uno o varios modelos; a continuación el análisis matemático de éstos y sus eventuales simplificaciones y, por último, la resolución mediante algoritmos de cálculo numérico que se completa con el pos-proceso y la visualización de los resultados obtenidos. La validación del modelo con medidas experimentales, cuando éstas sean factibles, es una etapa muy importante que permite dar fiabilidad a la simulación.

Algunos ejemplos de aplicación en la industria son el diseño global de aviones y automóviles y de las diferentes piezas que los componen, la optimización de procesos metalúrgicos como la obtención de metales por electrólisis o la solidificación de las coladas, el modelado de la combustión en la caldera de una central térmica, el diseño de electrodos para hornos de arco eléctrico, la construcción de hornos de inducción, la evaluación del impacto ambiental de vertidos contaminantes al mar, la reducción del ruido en recintos mediante la introducción de materiales aislantes o absorbentes, y/o la aplicación de técnicas de “control activo”, etc.

Un sector industrial de gran importancia en nuestro país es el que fabrica componentes para la industria aeronáutica o la del automóvil. Las grandes empresas ensambladoras, generalmente extranjeras, establecen unos requisitos de calidad muy exigentes que obligan a un diseño extraordinariamente preciso. Muchos de estos requisitos se refieren al comportamiento estructural, pero otros tiene que ver con aspectos térmicos, acústicos, de compatibilidad electromagnética o simplemente ergonómicos. Se trata de un sector que, de forma súbita, está utilizando herramientas de simulación numérica cada vez con mayor intensidad.

En resumen, los modelos matemáticos son una valiosa herramienta para la concepción y el diseño de dispositivos y procesos en la industria. La ingeniería moderna emplea cada vez más la simulación numérica (bajo lo que se denomina el “computer aided design/computer aided engi-

neering" (CAD/CAE), en la terminología inglesa). La razón es que el uso de modelos permite acortar y abaratar el proceso de diseño y salida al mercado de los productos, al reducir la construcción de prototipos y los ensayos en laboratorio. Por otra parte, como los ordenadores resultan cada vez más rápidos y menos costosos, estas tecnologías están ya al alcance de las pequeñas y medianas empresas.

El software comercial. Situación en España

En los primeros años tras la aparición de los ordenadores, la simulación numérica fue considerada de importancia estratégica fundamental. Por ello, en los países más avanzados algunos laboratorios de organismos públicos o de grandes empresas desarrollaron software propio, muchas veces en colaboración con grupos universitarios expertos en cálculo numérico. Sin embargo, el altísimo precio de los superordenadores requeridos para llevar a cabo los cálculos y los elevados costes de desarrollo del software, limitaron inicialmente el uso de la simulación numérica a la gran industria de los sectores energético, aeronáutico, de la automoción, del armamento, nuclear, etc. Pero en los últimos años la situación ha cambiado notablemente: por una parte el hardware ha experimentado un progreso espectacular en términos de la relación prestaciones/precio y, por otra, han aparecido en el mercado paquetes de software bien consolidados en la mayoría de los dominios de la física: mecánica de sólidos, mecánica de fluidos, combustión, electromagnetismo, medio ambiente, acústica, química computacional, etc.

Esta situación hace accesible y rentable la simulación numérica a empresas medianas y pequeñas de una gran variedad de sectores industriales, aunque con ciertas limitaciones, no sólo de carácter económico, sino también derivadas de la necesidad de contar con personal formado en el uso racional de este software. Como hemos señalado, el problema ahora ya no es el precio del hardware sino el del software y la situación se agrava porque en los últimos años estamos asistiendo a una concentración de los grandes paquetes en manos de muy pocas empresas que constituyen oligopolios a nivel mundial en un sector horizontal y estratégico para la industria moderna.

¿Cuál es el papel de España en este campo? Hasta el momento muy escaso, por no decir nulo: puede afirmarse con bastante seguridad que no existen programas de simulación numérica creados por empresas de nuestro país. La dependencia de empresas extranjeras, fundamentalmente norteamericanas, es casi total. La actividad en este terreno, además de escasa, se ha limitado al desarrollo de software de simulación de carácter específico para necesidades particulares de algunas industrias, pero no ha conducido a grandes paquetes comerciales de "propósito general". Sin embargo, el uso de la simulación numérica como herramienta de ingeniería asistida por ordenador sigue creciendo, lo que obliga a un elevado pago de royalties y a una fuerte dependencia tecnológica.

Llegados a este punto son pertinentes algunas preguntas: ¿debemos resignarnos a esta situación?, ¿el nivel científico y tecnológico de nuestro país hace posible generar desarrollos de software complementarios o competidores de los que están actualmente liderando el mercado mundial?, ¿sería



ésta una operación económicamente rentable? Yo me atrevo a responder afirmativamente a la segunda pregunta: en España hay investigadores en modelización, métodos numéricos y computación de altas prestaciones que convenientemente organizados garantizarían la viabilidad a medio plazo de proyectos para elaborar paquetes comerciales de simulación numérica para uso industrial.

Las otras preguntas tienen respuestas más difíciles. Naturalmente, puesto que estamos hablando de proyectos industriales, de la industria del software, habría que considerar todo el entramado de comercialización, soporte y mantenimiento, entendido este último como la puesta al día permanente de los productos para incorporar nuevos modelos, actualizar las técnicas numéricas o adaptarlos a las nuevas plataformas informáticas. Este entramado es algo esencial que habría que planificar desde el principio, incluyéndolo, al menos parcialmente, en la definición de los proyectos. En países avanzados de nuestro entorno, han existido excelentes desarrollos científicos en el ámbito de la simulación numérica que fracasaron porque no se previeron adecuadamente los mecanismos de transferencia al sector productivo. No sólo porque faltó la inversión necesaria para dotar al software desarrollado de un "envoltorio comercial" (interfaces, pre y pos-procesos, manuales, etc.) imprescindible de cara a su comercialización, sino, y sobre todo, porque no se previeron los mecanismos institucionales necesarios para facilitar su transferencia y explotación por el sector empresarial.

Cabe señalar un dato importante: en este momento la situación no es la misma en los diferentes campos de aplicación. Así como en mecánica de sólidos existe un gran número de paquetes comerciales, en temas como hidráulica, flujos medioambientales, combustión o electromagnetismo, la existencia de paquetes en el mercado es más reducida y parecen existir "nichos" donde todavía tendría sentido empresarial intentar una presencia española.

En definitiva, si el software de simulación numérica constituye un pilar estratégico para un desarrollo industrial moderno y competitivo, parece que deberíamos disponer a medio plazo de tecnología propia en este campo.

Téngase en cuenta que, como una tecnología que forma parte del área de las TIC en su vertiente de software, la elaboración de programas de simulación numérica no requiere la instalación de costosos laboratorios experimentales. Lo realmente imprescindible son los recursos humanos y, como hemos señalado más arriba, en este aspecto el nivel de nuestro país es suficiente.

Algunos ejemplos

En este apartado nos referiremos con algún detalle a un par de ejemplos que muestran el interés del cálculo para la industria. No voy a considerar ejemplos clásicos como son el uso del cálculo estructural con elementos finitos o de la mecánica de fluidos computacional (CFD) en la industria aeronáutica. Baste señalar que el grado de perfección alcanzado por estas tecnologías permite que los aviones modernos se diseñen casi completamente en el ordenador.

Comencemos por un problema clásico y de planteamiento sencillo, si bien su resolución requiere herramientas matemáticas no triviales como son los algoritmos de optimización discreta. Se trata de un ejemplo cuyo objetivo es mejorar un proceso industrial: el corte de bloques de piedra. En los últimos años el desarrollo del sector de la construcción ha propiciado la aparición de multitud de aserraderos de piedra que plantean el siguiente problema: a partir de grandes bloques extraídos en canteras, se desea satisfacer los pedidos de los clientes consistentes en unidades de diferentes tamaños, en general de forma paralelepípedica. Para ello se utilizan máquinas de corte, en las tres direcciones del espacio. El objetivo es establecer un proceso de corte que minimice los recortes inutilizables. Aquí tenemos un ejemplo para el que, frecuentemente, el software disponible en el mercado no se ajusta al particular sistema de corte, por lo que son necesarios desarrollos particulares.

Otro ejemplo. El cierre de minas a cielo abierto, generalmente de carbón, plantea el problema de ocupar el enorme hueco generado a lo largo de muchos años de extracción de mineral. Una alternativa usual consiste en llenarlo de agua para producir un lago artificial. El problema es prever la calidad de esta agua: la escorrentía de la cuenca provocada por la lluvias hace que el lago rebose una vez lleno, de manera que se van a producir vertidos a cauces cercanos que es preciso controlar. En efecto, los materiales de la escombrera generada por la explotación de la mina y las sustancias de la superficie del hueco pueden reaccionar con el agua, produciéndose especies contaminantes cuyo nivel es necesario controlar al objeto de evitar una concentración excesiva en los vertidos. Para que el proyecto pueda ser autorizado por las Administraciones es preciso hacer un estudio del impacto ambiental. Como no es posible "esperar y ver" se trata de un caso donde la simulación numérica puede y debe sustituir a la experimentación. El problema, aparentemente fácil, posee en realidad una gran dificultad matemática y numérica. En efecto, por una parte los modelos de hidráulica deberán permitir simular el llenado y, en particular, el avance del agua "mojando" la superficie del hueco. Por otra parte, las reacciones químicas que involucran las sustancias contaminantes son muchas y complejas. Estas dificultades generan otras de índole numérica, lo que se traduce en tiempos de cálculo muy elevados para simular procesos reales que pueden durar varios años. Entonces, el diseño de métodos numéricos eficientes y el uso de plataformas hardware de altas prestaciones resultan de interés crucial.

De la simulación a la optimización y el control

No queremos finalizar este breve artículo sin hacer una referencia al hecho de que la simulación es, en muchas ocasiones, un paso previo a la optimización o al control de un sistema. Por ejemplo, los criterios de eficiencia aerodinámica inspiran el diseño de las formas exteriores de un avión y, generalmente, estos criterios se reducen a la minimización de uno o varios funcionales matemáticos.

El control en tiempo real también necesita partir de un buen modelo de comportamiento: la reducción del ruido en un recinto mediante técnicas de "control activo" se analiza formulando problemas de control para la ecuación que rige la propagación del sonido. En este caso el criterio a



minimizar sería la presión media en el recinto o en alguna parte del mismo. Los controles a implementar serían sonidos producidos por altavoces o vibraciones de placas que cancelen las vibraciones indeseables. El diseño de estos dispositivos inteligentes que actúan como controladores y su ubicación óptima en el recinto, conduce a cálculos complejos que involucran tanto métodos de resolución numérica de ecuaciones en derivadas parciales como sofisticados algoritmos de optimización.

Bibliografía

- [1] *MATHEMATICS: Key for the European Knowledge-based Economy. A Roadmap for Mathematics in European Industry*, Publicación de la red MACSI-net, (2004).
(<http://www.macsinet.org/newsletter.htm>)
- [2] *Simulación Numérica y CAD en las Empresas Industriales de Galicia*. Informe Técnico 2005-002. CESGA. Santiago de Compostela, (2005).
- [3] Bermúdez, A.; Some applications of optimal control theory of distributed systems. A tribute to J. L. Lions. *ESAIM Control Optim. Calc. Var.* 8, (2002), (195-218).

Capítulo III

QUÍMICA TEÓRICA Y COMPUTACIONAL

POR FLORENTINO BORONDO RODRÍGUEZ
UNIVERSIDAD AUTÓNOMA DE MADRID
Departamento de Química, C-IX
Facultad de Ciencias
CANTOBLANCO-28049 Madrid
Correo electrónico: f.borondo@uam.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Computational Chemistry has become a very active field of research, that has gained a great deal of importance within the body of Theoretical Chemistry in the last decades. Just to give some data, the fraction of articles in the journals of the *American Chemical Society* making reference to any of the numerous Quantum Chemistry computer codes available raised from 1% to 12% from 1982 to 1992, and this has been a non-stopping tendency since then.

This effect has been due in part to the spectacular improvement of modern digital computers, but also to the development of new theoretical methods, allowing the accurate and efficient simulation of many processes of interest in Chemistry. This has made that today one can talk about *dry labs*, as a sensible alternative to generate data, whose experimental determination is problematic or just impossible in conventional laboratories. The prominent role played now by Computational Chemistry has been recently recognized by the scientific community with the two Nobel prizes awarded to Profs. Rudolph A. Marcus and John A. Pople in 1992 y 1998, respectively.

Starting from a brief description of the Born-Oppenheimer approximation, as the central tool in Computational Chemistry, we review in this paper the basis of the principal existing methods to study atomic and molecular electronic structure, paying always special attention to the corresponding relevant mathematical aspects.

In the second part, we describe the associated nuclear dynamics, whose complete understanding could lead some day to the development of a true Laser Selective Chemistry, in which be possible to design strategies to control efficiently chemical reactivity. The idea is to use laser monochromaticity as a 'surgical knife' to break selectively desired bonds and/or pump energy in selected moieties within molecules. In this respect, Mathematics, with all the tools and results obtained after the 1960's in the field of Dynamical Systems, play a prominent role providing an invaluable help. Just consider that from the vibrational point of view, molecules can be considered as a collection of coupled, anharmonic oscillators, in which the full richness of behavior described within nonlinear dynamics and chaos theory can be found at its full glory.

We conclude the paper by presenting some possible future developments in the field.

Introducción

En estos comienzos de siglo quedan muy lejos los días en que la Química se concebía como una disciplina restringida únicamente al entorno de un laboratorio. Hace dos siglos Comte, acercándose a esta cuestión, escribía en 1830 que *“cualquier intento de utilizar métodos matemáticos para el estudio de los fenómenos químicos debe ser considerado profundamente irracional y contrario al espíritu de la Química”*. El tiempo ha demostrado lo erróneo de esta aseveración, y hoy en día la Química Computacional es un campo emergente dentro de la Química Teórica, que

ha ganado en los últimos años gran importancia. Para constatar este cambio baste un dato. La fracción de artículos en las revistas de la *American Chemical Society* que hacían referencia a algún programa de Química Computacional pasó del 1% al 12% en la década de 1982 a 1992, y por supuesto esta tendencia ha seguido creciendo.

Este importante crecimiento se debe en gran parte al enorme desarrollo que han experimentado los modernos ordenadores digitales. Hoy en día las estaciones de trabajo más modestas y las “granjas” de PC's son capaces de proporcionar una potencia de cálculo comparable a la de los superordenadores de hace sólo algunos años. Otro factor importante ha sido la aparición de nuevos métodos teóricos, cada vez más sofisticados y precisos, que permiten el cálculo y simulación de complicados procesos y propiedades de sistemas químicos de relevancia. Así, se ha comenzado a hablar de los laboratorios “secos” (*dry labs*) como alternativa razonable para conocer datos del comportamiento de sistemas químicos, cuya medición en un laboratorio convencional sería muy prolija, complicada o incluso imposible. En palabras de J. A. Pople: “*En muchos aspectos los cálculos teóricos ya son más potentes que los experimentos. No están sujetos a consideraciones prácticas. Cualquier especie química puede ser escudriñada teóricamente; cálculos sobre cationes, aniones y otros intermedios reactivos, ..., no presentan, en principio, mayor problema que el cálculo de otras moléculas estables y más fácilmente observables...*”. El importante papel jugado por la Química Computacional en el panorama de la Química moderna ha sido refrendado por la comunidad científica con la concesión del Premio Nobel de Química a los Profs. Rudolph A. Marcus y John A. Pople en 1992 y 1998, respectivamente.

Concluiremos esta introducción poniendo de manifiesto que los métodos y enfoques de la Química Computacional llevan asociados interesantes y relevantes aspectos matemáticos [1] a los que se pasará revista en las secciones siguientes.

La Química Cuántica

La posibilidad de estudiar de forma teórica el enlace químico y las propiedades moleculares se abre en 1925 con la aparición de la ecuación de Schrödinger. Ésta sentó las bases matemáticas del concepto de cuantización, permitiendo explicar los espectros atómicos y otros hechos experimentales relacionados.

La ecuación de Schrödinger es una ecuación diferencial en derivadas parciales

$$(8) \quad \hat{H} \psi(r, R) = E \psi(r, R),$$

donde el operador Hamiltoniano, \hat{H} para una molécula compuesta de k núcleos y n electrones viene dado por la expresión

$$(9) \quad \hat{H} = -\sum_{\alpha} \frac{\hbar^2}{2M_{\alpha}} \nabla_{\alpha}^2 - \sum_i \frac{\hbar^2}{2m} \nabla_i^2 - \sum_{\alpha} \sum_{\beta > \alpha} \frac{Z_{\alpha} Z_{\beta} e^2}{R_{\alpha\beta}} - \sum_{\alpha} \sum_i \frac{Z_{\alpha} e^2}{r_{i\alpha}} + \sum_i \sum_{j > i} \frac{e^2}{r_{ij}}$$



donde los subíndices en letra griega se refieren a las coordenadas nucleares y los de letra latina a las electrónicas. Su solución nos proporciona los niveles de energía del sistema, así como las funciones de onda de los correspondientes estados. Los primeros nos permiten calcular los correspondientes espectros (electrónicos, vibrorrotacionales, etc.) y las segundas permiten el cómputo de cualquier propiedad fisicoquímica observable, con la ayuda de los correspondientes operadores. La ecuación de Schrödinger sólo tiene solución exacta en casos muy aislados, como son los átomos monoeléctricos, debidos a su simetría de campo central, o la molécula-ión de hidrógeno, H_2^+ . Por tanto, es necesario recurrir, en general, a métodos computacionales aproximados para su resolución, lo que ha dado lugar a la disciplina conocida como Química Cuántica [2].

El primer paso en la resolución de (8) es la aproximación de Born-Oppenheimer, que se basa en la gran diferencia de masa (al menos 1836 veces) que existe entre núcleos y electrones. Esto permite resolverla en dos pasos haciendo $\psi(r, R) = \psi_{el}(r; R) \psi_n(R)$. En el primero, se resuelve la parte correspondiente al movimiento de los electrones, suponiendo fijas las posiciones de los núcleos, mediante la ecuación de Schrödinger electrónica

$$(10) \quad \hat{H}_{el} \psi_{el}(r, R) = E_{el}(R) \psi_{el}(r; R),$$

donde

$$(11) \quad \hat{H}_{el} = -\sum_i \frac{\hbar^2}{2m} \nabla_i^2 - \sum_{\alpha} \sum_{\beta > \alpha} \frac{Z_{\alpha} Z_{\beta} e^2}{R_{\alpha\beta}} - \sum_{\alpha} \sum_i \frac{Z_{\alpha} e^2}{r_{i\alpha}} + \sum_i \sum_{j > i} \frac{e^2}{r_{ij}}.$$

En un segundo paso se considera la dinámica nuclear a través de la ecuación

$$(12) \quad \left[-\sum_{\alpha} \frac{\hbar^2}{2M_{\alpha}} \nabla_{\alpha}^2 + E_{el}(R) \right] \psi_n(R) = E_n \psi_n(R).$$

Dada la gran masa de los núcleos, también se puede estudiar alternativamente la correspondiente dinámica mediante aproximaciones clásicas o semiclásicas [3].

La separación de Born-Oppenheimer que acabamos de describir diseña un escenario que nos indica claramente que para realizar un estudio completo de la estructura, reactividad e interacciones entre moléculas, es necesario no olvidar ninguno de los dos aspectos complementarios (electrónico y nuclear), descritos por las dos ecuaciones de Schrödinger (10) y (12). Estos dos tipos de estudio dan lugar a líneas de investigación muy diferentes, que tradicionalmente se han llevado a cabo en grupos separados. A este respecto, hay que indicar que nuestra Comunidad de Madrid cuenta con grupos de investigación que gozan de un notable prestigio y reconocimiento internacional en ambos campos.

Cálculos de estructura electrónica

Las implementaciones prácticas más extendidas para la resolución de la ecuación de Schrödinger electrónica (10) se basan en el método autoconsistente (SCF¹) de Hartree-Fock (HF) con una función de onda expresada como un producto antisimetrizado o determinante de Slater

$$(13) \quad \psi_0 = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(n) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(n) \\ \dots & \dots & \dots & \dots \\ \phi_n(1) & \phi_n(2) & \dots & \phi_n(n) \end{vmatrix}$$

que considera a los electrones como partículas independientes descritas por los espinorbitales, ϕ_{ij} . La forma de estas funciones se obtiene aplicando el principio variacional, es decir minimizando el valor esperado del hamiltoniano

$$(14) \quad E = \frac{\int \psi_0^* \hat{H} \psi_0 d\tau}{\int \psi_0^* \psi_0 d\tau}$$

sujeta a la restricción de que los espinorbitales sean ortonormales. La aplicación de las ecuaciones SCF-HF implica la resolución iterativa de un conjunto de ecuaciones integro-diferenciales, lo cual prácticamente sólo es factible para sistemas atómicos o moleculares muy pequeños. En 1951 Roothaan, e independientemente Hall, propusieron un método para resolverlas que se basaba en desarrollar los orbitales ϕ_{ij} como una combinación lineal de funciones matemáticas conocidas o funciones de base (usualmente orbitales atómicos): $\phi_{ij} = \sum_j^N c_{ji} \chi_j$, ($N > n$), donde los coeficientes son las incógnitas a determinar variacionalmente. La resolución, de nuevo iterativa, del problema involucra la diagonalización de matrices grandes, pero la principal dificultad radicaba en el cálculo de las integrales de repulsión bielectrónica, cuyo número crece como N^4 , lo que ocasionaba problemas de tiempo de cálculo y espacio de almacenamiento en disco. Esto se resolvió en la década de los sesenta en que Boyd propuso el uso de funciones gaussianas de base, lo que constituyó la consolidación de los llamados métodos "ab initio". La siguiente mejora que se puede introducir en los métodos de tipo HF es la inclusión de la correlación entre electrones con distinto espín. Esto puede llevarse a cabo escribiendo la función de onda (13) como una combinación de determinantes generados mediante la promoción de electrones de orbitales ocupados a orbitales vacíos. Estos métodos son mucho más costosos computacionalmente, debido a que escalan como N^7 .

Una metodología alternativa para tratar la correlación electrónica es la teoría de Moller-Plesset, MPn, que desarrolla perturbativamente tanto energías como función de onda. La serie no siem-

¹ Estas siglas provienen del inglés: *Self Consistent Field*.



pre converge bien y el costo computacional es también alto: el nivel MP2 escala como N^5 y el MP4 como N^7 .

Los métodos anteriores se caracterizan porque tratan la correlación electrónica a través de la utilización de una función de onda expresada en una base de determinantes de Slater. Sin embargo, recientemente han surgido una serie de métodos alternativos basados en el funcional de la densidad, que también permiten calcular la energía de un sistema molecular incluyendo la correlación electrónica. La teoría del funcional de la densidad (DFT) se basa en el teorema de Hohenberg-Kohn, quienes en 1964 demostraron que la energía electrónica del estado fundamental de un sistema está completa y unívocamente determinada por la densidad electrónica, $\rho(r)$, del mismo. En este punto, hay que resaltar que mientras que las teorías “*ab initio*” de orbitales moleculares necesitan para el cálculo una función de onda que depende de $3N$ coordenadas, en la teoría DFT basta con conocer $\rho(r)$ que es una función de tres variables. Esto hace que los cálculos DFT sólo escalen como N^3 a pesar de tener en cuenta la correlación electrónica.

Movimiento de los núcleos. Dinámica no lineal

Una vez conocida la superficie de energía potencial del sistema, $E_{el}(R)$, y sus derivadas podemos proceder a estudiar el movimiento de los núcleos, a nivel cuántico, clásico o semiclásico. En particular, son de especial relevancia los movimientos de vibración altamente excitados que son de interés en espectroscopia de infrarrojo, procesos de transferencia intramolecular de energía y relajación vibracional, y por último, y más importante, la reactividad química. La comprensión detallada de los fenómenos anteriormente mencionados en términos de la dinámica subyacente podría concluir en un futuro al desarrollo de una verdadera Química Selectiva por Láser, en la que se combinen la monocromaticidad de los láseres y la comprensión teórica de los mecanismos a través de los que esta fluye dentro de los sistemas moleculares, para diseñar estrategias en las que se potencien (o eviten) selectivamente canales de reacción preferidos (o no deseados) [4]. Esto nos permite calcular y simular los diferentes espectros del sistema, si la energía de excitación es pequeña, o incluso entrar en el régimen en que se pueden producir reacciones químicas, si es alta y estamos en la zona espectral del continuo.

Desde este punto de vista se consideran a las moléculas como sistemas Hamiltonianos multidimensionales no integrables formadas por osciladores anarmónicos acoplados, cuya dinámica tiene lugar en superficies de potencial para los distintos estados electrónicos que se obtienen dentro de la aproximación de Born-Oppenheimer. Así, serían aplicables a su estudio las técnicas matemáticas puestas a nuestra disposición por la Dinámica No Lineal o Teoría del Caos. El correspondiente espacio de fases está formado por zonas regulares y caóticas entremezcladas, incrementándose la proporción de las regiones estocásticas a medida que la energía vibracional aumenta. Así, para energías de excitación muy bajas estaríamos en la zona armónica de la superficie de energía potencial y tendríamos los conocidos modos de vibración normales, que corresponden a movimientos colectivos en los que participan todos los átomos de la molécula (y son análogos a los fonones en estado sólido), siendo la dinámica correspondiente totalmente regular. Para ener-

gías de excitación superiores empezaríamos a entrar en las zonas no armónicas del potencial y a excitar los sobretonos, de forma que los acoplamientos entre modos comienzan a ser no despre- ciables. Desde el punto de vista teórico, en este régimen conviene abandonar la representación de modos normales para pasar a modos más localizados en los enlaces (o grupos de enlaces). Empe- zaríamos a estar en la región no lineal del movimiento. En esta zona las resonancias existentes en la molécula son operativas, de forma que la transferencia intramolecular de energía es impor- tante. Estos fenómenos son relevantes para la llamada relajación vibracional, de forma que si exci- tamos selectivamente zonas concretas de la molécula, esta puede usar canales eficientes de dis- persar la energía dentro de la misma. Por último si la excitación es elevada, la dinámica no line- al anterior tendría energía suficiente para vencer la relajación vibracional y depositar suficiente energía en un modo local que conduzca a la ruptura de alguno de los enlaces dando lugar a pro- cesos de reactividad química [6]. La reactividad química no debe entenderse solamente como ruptura de enlaces, sino también como desplazamiento de átomos dentro de la misma. Ejemplos típicos, serían la isomerización en la que la disposición espacial de los átomo cambia en la molécula, o las reacciones de transferencia de protón, intra o intermolecular, por efecto túnel que tanta importancia tienen en Bioquímica y Biología Molecular.

Todo lo anteriormente expuesto se refiere a reactividad en fase gas en la que las condiciones de vacío son elevadas y las moléculas pueden suponerse aisladas. Esto se logra hoy en día fácil- mente, ya que la detección de señales muy débiles no es ningún problema. Sin embargo la ver- dadera Química ocurre en disolución. Esto puede simularse añadiendo al modelo anterior ruido y "viscosidad" de forma que se obtenga una ecuación dinámica efectiva tipo Langevin.

Las técnicas para estudiar esta dinámica de los núcleos pueden ser la espectroscopia en las situaciones de baja (espectroscopia clásica) y media excitación (espectroscopia por bombeo de emisión estimulada) o de dinámica molecular (estudio de la reactividad mediante haces molecu- lares cruzados, que combinados con láser permiten preparar estados iniciales electrónicos espe- cíficos o reactivos orientados, etc.). En la actualidad, están bien comprendidas las bases del comportamiento de estos sistemas dinámicos en dos grados de libertad mediante los teoremas de Kolmogorov-Arnold-Moser (KAM) y de Poincaré-Birkhoff [7] a nivel de la Mecánica Clásica.

Sin embargo no hay que perder de vista el hecho de que el comportamiento atómico y molecular es eminentemente cuántico. Por ello una parte muy importante de estas investigaciones se dedi- ca al estudio de las manifestaciones cuánticas del caos o caos cuántico [8], en la que partiendo de las estructuras invariantes en el espacio de fases clásico se incluyen los efectos de interfe- rencia debidos a la naturaleza inherentemente ondulatoria de la materia [9].

Perspectivas futuras

Como resumen podemos decir que el estado actual de la Química Computacional es muy saluda- ble, de forma que es previsible que aparezcan en un futuro cercano resultados espectaculares y



desarrollos interesantes. Aventurando algunas de las líneas en las que se puede esperar que estos se produzcan, cabe mencionar las siguientes:

1. Avanzar en el desarrollo de métodos DFT y “coupled cluster” con escalaje del tiempo de cálculo lineal [10].
2. Encontrar el funcional de intercambio para DFT, así como desarrollar bases específicas para esta teoría.
3. Desarrollo de algoritmos eficaces para el cálculo de las derivadas de $E_{el}(R)$, así como para el cálculo de sus puntos estacionarios cuando hay uno o más autovalores negativos en el Hessiano.
4. Extensión de los métodos computacionales de la dinámica no lineal a sistemas con tres y más grados de libertad, donde tiene lugar la poco estudiada difusión de Arnold. Estudio de los procesos de transferencia intramolecular de energía asociados en modelos realistas.
5. Elucidar el papel de los invariantes clásicos (órbitas periódicas, órbitas homo y heteroclínicas) en la mecánica cuántica de sistemas dinámicos caóticos, teniendo en cuenta que se establece una competición entre la complejidad clásica inherente y los efectos de interferencia, constructiva y/o destructiva, propios del enfoque ondulatorio.
6. Modificación de la fórmula de la traza de Gutzwiller, única capaz hoy en día de cuantizar sistemas dinámicos caóticos, de forma que se evite el problema computacional derivado del crecimiento exponencial de número de órbitas periódicas necesarias con el tiempo.

Bibliografía

- [1] Porter, M.; Cvitanovic, P.; *Notices of the American Mathematical Society*, (2005).
- [2] Borondo, F.; Martín, F.; *Química Cuántica*, en *Química Física*, J. Bertrán y J. Núñez (eds.) (2002).
- [3] Child, M.S.; *Semiclassical Mechanics with Molecular Applications* (1991).
- [4] Rice, S.A.; Zhao, M.; *Optical Control of Molecular Dynamics* (2000).
- [5] Reichl, L.E.; *The Transition to Chaos* (2004).
- [6] Uzer, T.; *et al.*, *Phys. Rev. Lett.* 84, 610 (2000); *ibid.* 95, 058301 (2005).
- [7] de la Llave, R.A.; *Tutorial on KAM Theory*, Proc. Symp. Pure Math. (AMS, Providenc, 2001), vol. 69, (p.175).
- [8] Gutzwiller, M.C.; *Chaos in Classical and Quantum Mechanics* (1990).
- [9] Wisniacki, D.A.; Vergini, E.R.; Benito, M.; Borondo, F.; *Phys. Rev. Lett.* 94, 054101 (2005); *J. Chem. Phys.* 112, 111101 (2005).
- [10] Kudin, K.N.; Scuseria, G.E.; *Phys. Rev. B* 61, 16440 (2000).

Capítulo IV

MATEMÁTICA FINANCIERA

POR SANTIAGO CARRILLO MENÉNEZ
UNIVERSIDAD AUTÓNOMA DE MADRID
Departamento de Matemáticas,
Facultad de Ciencias y RiskLab-Madrid
CANTOBLANCO-28049 Madrid
Correo electrónico: santiago.carrillo@uam.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Mathematical finance is now a recognized part of applied mathematics. It began with derivative pricing (the well known Black-Scholes formula). Nowadays there are many open problems related with the pricing of complex derivatives, in particular related with credit risk, that require advanced mathematics (like Malliavin calculus, for example) and advanced computational methods. Nevertheless, risk management offers new and more complex challenges to mathematicians in quest for applications.

Introducción

A partir de 1973 con la célebre fórmula de Black-Scholes, la Matemática Financiera ha conocido un desarrollo espectacular del que dan fe las numerosas publicaciones especializadas que han eclosionado en los últimos años en paralelo con la expansión de los mercados. Entre otras razones este fenómeno se debe a la desregularización del dólar, y al desarrollo de las comunicaciones y de la informática que han aportado una eficacia (numérica) hasta entonces desconocida a la hora de pasar órdenes lo cual ha permitido una interconexión de los mercados.

En este contexto se afianzó la idea de gestionar el riesgo y se desarrollaron (debido a una demanda creciente) los instrumentos derivados correspondientes. Con el paso del tiempo, los productos han sido cada vez más complejos, precisando de un conocimiento matemático y de recursos computacionales cada vez mayores. Las inversiones en derivados se miden hoy en billones de euros.

Entre otros elementos importantes, la Matemática Financiera se caracteriza por un uso intensivo de:

- Cálculo estocástico.
- Procedimientos numéricos avanzados y uso intensivo de cálculo científico.
- Ecuaciones en derivadas parciales.

Todo ello explica por qué la Matemática Financiera se ha ido consolidando como una rama de las matemáticas (en general del Cálculo de Probabilidades) en la mayor parte de los países de nuestro entorno. Una primera consecuencia de este hecho ha sido un cambio drástico en los perfiles demandados en los departamentos de riesgo o en las mesas de derivados de las principales entidades financieras del planeta, mayormente matemáticos, físicos o ingenieros.

Sin pretender ser exhaustivo y procurando no ser excesivamente técnico, el objeto de estas líneas es dar una visión de las que, a juicio del autor, pueden ser las áreas más pujantes y más demandantes de métodos cuantitativos en las finanzas actuales.

Para ello, distinguiremos dos campos de aplicación de las matemáticas en finanzas: la valoración de derivados y la gestión de riesgos.

La selección de una bibliografía para un trabajo de este tipo siempre supone el riesgo de cierta arbitrariedad. Se ha procurado citar algunos de los trabajos seminales. Para referencias más recientes se remite al lector a la búsqueda por palabras clave en Internet o a contactar con el autor si necesita referencias más concretas.

Valoración de derivados

El punto de partida del éxito de las Matemática Financiera es la ecuación diferencial estocástica que describe la dinámica temporal de un activo (una acción) S_t :

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t \quad \text{es decir: } S_t = S_0 \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) t + \sigma dW_t \right\}$$

siendo r el rendimiento de la cuenta bancaria, W_t un movimiento browniano que modeliza la llegada de la información al mercado y determina el comportamiento aleatorio del activo y σ , la volatilidad, que representa la desviación típica de los rendimientos instantáneos del subyacente.

En este marco, la fórmula de Black-Scholes (1973) permite valorar el derivado más habitual: la *call europea* que da derecho a obtener una acción de dicho subyacente en un instante futuro T a un precio K (llamado de ejercicio) fijado ahora. El precio de dicho derivado es:

$$C (S_0, K, r, T, \sigma) = S_0 N (d_+) - Ke^{-rt} N (d_-)$$

siendo N la función de distribución de la normal tipificada y

$$d_{\pm} = \frac{1}{\sigma \sqrt{T}} \ln \frac{S_0}{Ke^{-rt}} \pm \frac{1}{2} \sigma \sqrt{T}$$

De los distintos parámetros del modelo, sólo σ no viene dada y hay que estimarla a partir de datos de mercado. Un procedimiento consiste en partir de los precios cotizados en los mercados y deducir cuál es la *volatilidad implícita* que, usada en la fórmula de Black-Scholes, la hace coincidir con el precio de mercado.

En la década de los ochenta quedó de manifiesto que el cálculo de la volatilidad implícita para distintos valores del precio de ejercicio, quedando fijos los demás parámetros, conducía a diferentes valores para σ .

Luego $\sigma = \sigma(K)$ es una función del precio de ejercicio y es indispensable tener en cuenta este hecho a la hora de valorar los posibles derivados obtenidos a partir del subyacente S_t . Este fenómeno se conoce como la *sonrisa de volatilidad*, debido a la forma que tenía en las primeras observaciones la gráfica de σ .



La cosa se complica cuando se quieren modelizar opciones en las que intervienen diversos subyacentes como, por ejemplo, en el caso de las opciones de la familia *mountain*: una clase de derivados introducida por Soci t  G n rale y cuya funci n de pago hace intervenir una cesta de numerosos (n) subyacentes de los cuales pueden ser tomados en consideraci n s lo una parte de los mismos (por ejemplo, se excluyen los n_1 peores y los n_2 mejores, obviamente con $n_1 + n_2 < n$).

Otro factor de complejidad radica en la posibilidad de cancelar el producto anticipadamente o en la posibilidad de que la fecha de ejercicio est  abierta (opciones americanas).

Todo ello requiere la capacidad de modelizar de forma apropiada la din mica de los subyacentes y de sus volatilidades para poder resolver el problema de valoraci n mediante simulaci n Monte-carlo. Es  ste un problema no resuelto de manera satisfactoria y en el cual existe un amplio campo de trabajo para los matem ticos.

Los derivados sobre acciones (renta variable) representan s lo una peque a parte del mercado global de derivados. Otra parte la constituyen los derivados de renta fija y divisa y los de riesgo de cr dito.

Los derivados de renta fija tienen como finalidad la gesti n de la incertidumbre derivada de los movimientos de los tipos de inter s y los derivados de divisas aquella asociada a los tipos de cambio. Su cuota de mercado es muy superior a la de los derivados de renta variable.

Los modelos usados en renta fija son m s complejos que los anteriores al tener que integrar la informaci n relativa a distintos plazos. Por ejemplo, la siguiente ecuaci n se corresponde con una versi n simplificada del Libor Market Model, uno de los modelos m s usados en este momento en los mercados para la din mica de los forward (tipos a plazo):

$$\frac{dF_i(t)}{F_i(t)} = \sum_{j=m(t)}^i \frac{\partial_j F_j(t) \sum_{q=1}^p \lambda_{j-m(t),q} \lambda_{i-m(t),q}}{1 + \partial_j F_j(t)} dt + \sum_{q=1}^p \lambda_{i-m(t),q} dW_t^q$$

Aqu  las $\lambda_{j,q}$ son volatilidades observadas para los tipos forward a diferentes plazos y los W_t^q son movimientos brownianos. El n mero p de factores tiene que determinarse para proceder a los ajustes.

Las observaciones hechas m s arriba, en particular la existencia de una sonrisa din mica que hay que tener en cuenta en la valoraci n, tambi n son de aplicaci n en este contexto, agravadas por la complejidad a adida. Esto explica por qu  la valoraci n de estos derivados moviliza hoy en d a una gran cantidad de *quants*.

Los derivados de riesgo de cr dito, que permiten cubrirse del incumplimiento (*default*) de una contrapartida, son el otro gran foco de inter s en este momento de los mercados financieros.

En general descansan en la modelizaci n conjunta de los subyacentes susceptibles de incumplimiento, una cesta m s o menos grande, con el  nfasis puesto en estos posibles incumplimientos (first to default, second to default, etc.).

El uso de este tipo de derivados ha supuesto la introducción de lo que hasta hace poco se habría considerado como supercomputación, el uso de decenas o centenares de procesadores en paralelo, en no pocas instituciones financieras. En muchos casos, la simulación Montecarlo tiene la dificultad añadida de una dimensionalidad alta. Son muchos los desafíos, tanto matemáticos como computacionales, existentes en este campo.

Gestión de riesgos

Ya hemos señalado la relación entre derivados y gestión del riesgo, pero esta tiene otra faceta: el capital económico o regulatorio. Las entidades financieras están obligadas a tener parte de su capital invertido en instrumentos muy líquidos con el fin de poder hacer frente a posibles pérdidas por riesgo de mercado, riesgo de crédito o riesgo operacional.

El Nuevo Acuerdo de Capital (Basilea II) y su próxima plasmación en normativa europea define la manera de calcular dicho capital. Estos instrumentos más líquidos son menos rentables por lo que el cálculo eficiente del capital económico es crítico en la gestión de una entidad financiera.

A la hora de calcular este capital económico se trata de hallar un determinado percentil de la distribución de pérdidas, previamente determinada. Éste es otro ámbito de trabajo para un matemático.

De los tres tipos de riesgos mencionados, el de mercado puede ser el más conocido aunque el menos relevante en términos de capital. Sin embargo queda mucho por hacer en cuanto a su modelización, especialmente en cuanto a colas pesadas se refiere: las distribuciones empíricas presentan unas colas con un peso mayor que el estimado a partir de una normal lo cual implica una probabilidad mayor de sucesos extremos que la que cabe inferir a partir de la normal.

El riesgo operacional ha sido el último en ser incorporado por el Comité de Basilea y está suponiendo el uso de modelos actuariales (*loss distribution approach*) con la dificultad añadida de tener que calcular un percentil elevado (99,9) por lo que la implementación de este tipo de modelos exige combinar procedimientos estadísticos avanzados, métodos numéricos y soluciones informáticas innovadoras (*grid*).

El riesgo de crédito plantea una situación algo distinta: aquí el regulador ha impuesto un modelo unifactorial gaussiano a partir del cual se obtiene el capital económico. Sin embargo cabe esperar que en los próximos años se vayan abriendo camino los modelos multifactoriales que permiten recoger el *efecto diversificación* de las carteras de crédito y éste es otro campo en el cual los matemáticos tendrán algo que decir.

Las entidades financieras están trabajando ya muy activamente en la implantación del acuerdo de Basilea. No cabe duda que en este proceso surgirán no pocas oportunidades de colaboración universidad-empresa en cuanto a matemática aplicada y computación se refiere.



Otros aspectos

Aún sin tener la pretensión de exhaustividad, es obligado mencionar, aunque sea brevemente, otros aspectos de los posibles campos de aplicación de las matemáticas en finanzas.

En primer lugar, conviene señalar que el modelo de Black-Scholes antes mencionado tiene varios inconvenientes entre los que destaca su carácter simétrico y el hecho que las pérdidas realmente observadas sean mayores y más frecuentes que las previstas por los mismos. De ahí la existencia de un esfuerzo por la búsqueda de modelos que describan mejor el comportamiento de los mercados. Señalemos en esa línea los trabajos destinados a sustituir el movimiento browniano por procesos de Levy. También cabe destacar los esfuerzos realizados para la aplicación del Cálculo de Malliavin a las finanzas o una nueva línea de trabajo en lo que se ha denominado *optimización robusta*.

En no pocas situaciones surgen problemas de optimización complejos. La asignación eficiente de capital, la determinación de activos para titularizar, la optimización de carteras que aparece en muy diversas situaciones: fondos de pensiones o de inversión, fondos de gestión alternativa (hedge funds) y otros, son algunos ejemplos de este tipo de problemas. Son muchas las cosas todavía por hacer en estos campos.

Es interesante señalar que la gestión de riesgos ha saltado del mundo de las finanzas a otros sectores: el energético (derivados de energía), el medio ambiente (emisiones de dióxido de carbono), o el de los seguros (bonos catástrofe o derivados de meteorología). Todos ellos plantean nuevos retos relacionados con la modelización de nuevos subyacentes, en general no gaussianos, con productos derivados poco líquidos, pero que son, sin lugar a dudas, una de las nuevas fronteras para las matemáticas del riesgo.

Por último diremos que el mundo de los seguros está en vísperas de una revolución similar a la que se ha producido en finanzas. La irrupción del cálculo estocástico en seguros es ya una realidad y los trabajos de *Solvencia II* anuncian un marco regulatorio parecido al de Basilea II. Sin lugar a dudas un nuevo campo de acción para enfoques cuantitativos.

Bibliografía

- [1] Bachelier, L.; *Théorie de la Spéculation*. Ann. Sci. Ecole Norm Sup. 17 21-86. Reprinted in The random character of stock market prices MIT Press, Cambridge, Mass, (1964), (17-78).
- [2] Black, F.; Scholes, M.; *The pricing of options and corporate liabilities*. Journal of Political Economy, 81, (1973), (637-59).
- [3] Bones J.; *Elements of a Theory of Stock-Options Value*. Journal of political Economy 72, (April 1964), (163-175).
- [4] Brace, A.; Gatarek D.; Musiela M.; *The market model of interest rate dynamics*. Mathematical Finance, 7, (1997), (127-154).
- [5] Carr, P.; Ellis, K.; Gupta, V.; *Static hedging of exotic options*. Journal of Finance 53, (1998).

- [6] Duffie, D.; Singleton, K.; *Modeling term structures of defaultable bonds*. The Review of Financial Studies, 12-4 (1999), (687-720).
- [7] Ho, T.S.Y.; Lee, S.B.; *Term structure movements and pricing interest rate contingent claims*. Jour. Fin. 41, (1986), (1011-1028).
- [8] Harrison, J.M.; Kreps, D.M.; *Martingales and arbitrage in multiperiod security markets*. J. Econom. Theory, 20, (1979), (381-408).
- [9] Harrison, J.M.; Pliska, S.R.; *Martingales and Stochastic Integrals in the theory of continuous trading*. Stochastic processes and their applications. 11, (1981), (215-260).
- [10] El Karoui, N.; Karatzas, I.; *The optimal stopping problem for a general American put-option*. Mathematical Finance IMA vol. 65, (1995), (77-88). Springer-Verlag, New York.
- [11] El Karoui, N.; Geman, H.; Rochet, Y.; *Changes of numeraire, changes of probability measure and option pricing*, Journal of Applied Probability 32, (1995), (443-458).
- [12] Mc Kean, H.P. (Jr); *A free boundary problem for the heat equation arising from a problem in mathematical economics*, Industrial Management Review 6, (1965), (13-21).
- [13] Longstaff, F.A.; Schwartz, E.S.; *Interest rate volatility and the term structure: a two-factor general equilibrium model*. Jour. Fin. 47, (1992), (1259-1282).
- [14] Mandelbrot, B.; *The variation of certain speculative prices*. Journal of Business 36, (1963), (394-419).
- [15] Markovitz, H.; *Portfolio selection*. J. Finance 8, (1952), (77-91).
- [17] Osborne, M.F.M.; *Brownian Motion in the stock market*. Operation Research 7, (1959), (145-173).
- [18] Ross, S.A.; *The Arbitrage Theory of Capital Asset Pricing*. Journal of Economic Theory 13, (1976), (341-360).
- [19] Samuelson, P.A.; *Rational theory of warrant pricing*. Industrial Management Review 6, (1965), (13-21).
- [20] Samuelson, P.A.; *Mathematics of speculative price*. SIAM Review a 15, (1973), (1-42).
- [21] Sharpe, W.; *Capital Asset Prices: A Theory of Market Equilibrium under conditions of risk*. J. of Finance September, (1964), (425-442).
- [22] Vasicek, O.A.; *An equilibrium characterization of the term structure*. Journal of Financial Economics, 5, (1977), (177-88).



Capítulo V

LAS MATEMÁTICAS Y EL PROCESAMIENTO DE IMÁGENES

POR VICENT CASELLES COSTA
UNIVERSITAT POMPEU FABRA
Passeig Circumvalació, 8
08003 Barcelona
Departamento de Tecnología
e-mail: vicent.caselles@upf.edu

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Image processing and computer vision have become an active area of research for mathematicians due to its impact in the development of new technologies. Its applications range from industrial vision to medical imaging, satellite images, video processing, the postproduction of digital cinema and arts. This area of research is mostly directed by its immediate applications but it requires a discussion of its mathematical foundations and the organization of its numerous contributions and algorithms to understand its universality and the limits of its applicability. Mathematical research can contribute to all aspects involved in the current activity in image processing: from modeling to the development and implementation of new algorithms without forgetting its mathematical analysis. To cover this set of activities we require the collaboration of engineers, computer scientists, physicists and mathematicians. Mathematicians may contribute also in all aspects and we want to stress the idea that its training is well adapted to discuss modeling questions, a task that has been traditionally object of more attention by other communities.

The areas of application of image processing are numerous and without the intention of being exhaustive, we could mention some of them: video processing and analysis (with its applications to surveillance, traffic control, tracking objects in motion, etc.), the creation of tools for digital cinema post-production, the vast field of medical imaging (image reconstruction, interpretation and aid to diagnostic), digital photography, stereo vision and 3D reconstruction from image sequences, the restoration and interpretation of satellite images, shape recognition, the search of image in the web, image compression, surface processing and its numerous applications, image synthesis and its applications in cinema and video games, and a large etcetera both in basic research and applications.

These fields of research and application do not fit exactly in any of the traditional mathematical disciplines but require the ideas and contributions of many, if not all, of them. This is a great advantage of image processing and computer vision related to its multi-disciplinary character, most —if not all— of mathematical research areas have something to say: variational calculus, physical modeling based on partial differential equations, geometry (differential, topological, discrete, ...), topology, data structures, stochastic processes, harmonic analysis and sampling theory, coding and information theory, numerical analysis and optimization, optics, color theory, physics in and computer science. Besides these areas we could mention all areas in which image processing applies.

By all these reasons and by its many applications image processing is probably going to attract more attention in the mathematics community and it will be interesting to observe his evolution both in our country and in the rest of the world concerning the number of researches devoted to it, and its connection with the industrial world.

As an illustration, in the expanded version of this text we described some of the application fields mentioned above. We selected some fields without going into detail in some others by reasons of space:

- a) Video processing and digital cinema.
- b) Digital photography.
- c) Satellite image processing.
- d) Three-dimensional image reconstruction from stereo pairs or video sequences.
- e) Shape recognition and search of images in the web.
- f) Medical Imaging.
- g) Image coding and compression.
- h) Computer graphics and surface processing.

Introducción

El procesamiento de imágenes y la visión por ordenador se han convertido en un área de investigación importante debido al rápido desarrollo de las nuevas tecnologías. Sus aplicaciones se extienden desde la visión industrial a las imágenes médicas, las imágenes satelitales, el vídeo y el cine digitales y el arte. Esta investigación está dirigida por las aplicaciones pero requiere de una discusión de sus fundamentos matemáticos para poder organizar el cuerpo de las numerosas contribuciones y algoritmos, comprender su universalidad y los límites de su aplicabilidad. La investigación matemática puede y debe contribuir a todos los aspectos involucrados en el procesamiento de imágenes: tanto a la modelización como al desarrollo e implementación numérica de algoritmos, sin olvidar su análisis matemático. Para cubrirlos se requiere de auténticos equipos multidisciplinares que incorporen a ingenieros, informáticos, físicos y matemáticos. Los matemáticos pueden aportar su experiencia en todos los aspectos; quisiéramos señalar expresamente que su formación les permite abordar los problemas de modelización que tradicionalmente han sido objeto de mayor atención por parte de otras comunidades.

Los campos de aplicación del procesamiento de imágenes son numerosos y sin pretender ser exhaustivos podemos mencionar algunos de ellos : procesamiento de vídeo (con sus múltiples aplicaciones: vigilancia, control de tráfico, seguimiento de objetos en movimiento, etcétera) y la creación de herramientas para la postproducción de cine digital, el ámbito de las imágenes médicas (reconstrucción, interpretación y ayuda al diagnóstico), la fotografía digital, la visión estéreo y la reconstrucción tridimensional a partir de secuencias de vídeo, la restauración e interpretación de las imágenes tomadas por satélites, el reconocimiento de formas y la búsqueda de imágenes en la web, la compresión de imágenes, el procesamiento de superficies, la síntesis de imágenes y la simulación para videojuegos y un largo etcétera tanto en los temas de investigación básica como en las aplicaciones.

Estos campos de aplicación y desarrollo de la actividad investigadora no se encuadran dentro de ninguna de las disciplinas tradicionales de las matemáticas pero requieren de todas ellas. Ésta es la gran ventaja del procesamiento de imágenes ligada a su carácter multidisciplinar, muchas por no decir todas las áreas de las matemáticas tienen algo que decir: el cálculo de variaciones, la modelización física apoyada en ecuaciones en derivadas parciales, la geometría (dife-



rencial, topológica, discreta ...), la topología, las estructuras de datos, la estadística, la teoría de procesos estocásticos, el análisis armónico y la teoría del muestreo, la teoría de la información, codificación, el análisis numérico y la optimización, la óptica, la teoría del color, y en general, la física, sin olvidar los aspectos de ingeniería e informática. Hemos mencionado estas áreas, pero podríamos sumar muchos ámbitos de experiencia en los que las imágenes se aplican.

Por todo ello y sobre todo por sus numerosas aplicaciones este área va a adquirir un gran desarrollo en el futuro próximo y es interesante observar los movimientos dentro de la comunidad matemática que se producen en los países de nuestro entorno que afectan tanto al número creciente de investigadores dedicados a ello, como a la relación con el mundo industrial y la creación de nuevas empresas.

A título ilustrativo, vamos a describir en lo que sigue algunos de los campos de aplicación mencionados más arriba sin que los temas elegidos signifiquen más que otros sobre los que pasaremos brevemente por razones de espacio.

Algunos campos de aplicación del procesamiento de imágenes

a) Procesamiento de vídeo y cine digital

El tránsito desde el cine y la televisión analógicas a la tecnología digital plantea una serie de problemas tecnológicos de interés práctico. Los sistemas analógicos se caracterizan por tener más baja resolución, niveles más elevados de ruido y artefactos y menor flexibilidad, los sistemas digitales permiten imágenes más nítidas con menores niveles de ruido y menos artefactos, siendo más robustas a la transmisión.

La coexistencia de sistemas analógicos y digitales en TV y vídeo requiere el poder intercambiar los formatos de ambos sistemas según diferentes niveles de calidad y complejidad. A los formatos de producción hay que añadir los formatos del receptor (*sistemas de display*). Los receptores han de ser capaces de convertir el formato recibido en el formato del dispositivo de display, a ser posible en tiempo real. Estos formatos se definen por su muestreo espacio-temporal y su relación de aspecto [3]. La televisión usa un formato de muestreo llamado entrelazado mientras que los ordenadores personales usan el llamado muestreo progresivo. El uso del ordenador como pantalla de visualización plantea el debate entre el desarrollo algoritmos de des-entrelazado de vídeo o la producción de vídeo progresivo. La primera estrategia es de menor coste y permite la compatibilidad con los formatos tradicionales de la televisión. Se necesitan pues algoritmos eficientes que permitan la conversión de vídeo entrelazado a vídeo progresivo (*IPC, Interlaced to Progressive Conversion*), el aumento de la resolución espacial (*scan-rate conversion*) y el aumento de la resolución temporal (*frame-rate conversion*).

Por otra parte, en el contexto del cine, si bien la adquisición y exhibición siguen siendo primordialmente analógicas, en película de 35mm, la etapa intermedia de postproducción es cien por cien digital en la mayoría de los casos. Y la tendencia además es a migrar de película de 35mm a vídeo digital de alta resolución, por su menor costo, manteniendo la calidad de imagen y la apariencia del *film*.

Des-entrelazado: El formato entrelazado es una forma de muestreo espacio-temporal. Para evitar el parpadeo el ojo humano requiere una frecuencia temporal de unos 60 Hz lo que, con la resolución espacial de 625 líneas y el ancho de banda de la TV analógica original, permitiría tan solo el display de unos 25 o 30 cuadros de imagen por segundo. Para resolver este problema cada cuadro se divide en dos campos, en el primero se muestrean las líneas pares y en el segundo las impares (o al revés). Dos campos consecutivos han de estar adquiridos en diferentes instantes de tiempo. El proceso de des-entrelazado es la operación que permite pasar del formato de vídeo entrelazado al formato progresivo duplicando la densidad de muestreo en la dimensión vertical de la imagen y eliminando el aliasing espacio temporal. Mencionemos un hecho importante: el muestreo de vídeo no cumple las demandas del teorema de Shannon ya que no se ha procedido a un filtrado antes del sub-muestreo - lo que permitiría evitar el aliasing. La razón es que la degradación del filtrado (*blurring*) introducido sería peor que el aliasing. Esta característica dificulta el problema de la conversión de formatos que no puede abordarse con métodos lineales de interpolación. Los métodos usados en la actualidad incorporan tanto la interpolación espacial (que tiene en cuenta los contornos de la imagen) como la compensación o adaptación al movimiento y las restricciones dadas por el teorema generalizado del muestreo. Éste es un ámbito perfecto para el trabajo de un matemático, pudiendo abordarlo en dos frentes: el análisis de las restricciones físicas impuestas por el proceso de adquisición de las imágenes con el objetivo de mejorar los modelos y determinar los márgenes de calidad que pueden obtenerse, y el desarrollo de algoritmos que incorporen las restricciones de funcionamiento en tiempo real. El trabajo [1] contiene una revisión excelente sobre la relevancia del problema y las técnicas más comunes empleadas en los algoritmos de des-entrelazado.

Aumento de resolución espacial: El aumento de resolución espacial de una imagen a partir de varias de ellas tiene muchas aplicaciones además de la conversión de formatos de vídeo [3], desde las imágenes médicas a las imágenes satelitales, pasando por la mejora de la calidad de impresión o las aplicaciones de vídeo-vigilancia donde interesa magnificar objetos (como matrículas de coche o caras) para reconocerlos. El aumento de resolución es posible cuando se dispone de varias imágenes diferentes de la misma escena u objeto que cuando son registradas (puestas en el mismo sistema de referencia) presentan desplazamientos con precisión por debajo del píxel. Si estos desplazamientos existen y si además hay aliasing presente, no podemos deducir una imagen a partir de la otra, lo que permite construir una imagen de mayor resolución [16, 17]. Las secuencias



de vídeo, las imágenes adquiridas por satélites orbitando alrededor de la tierra, o diversas vistas de una escena adquiridas con una cámara proporcionan las diferentes imágenes con el movimiento relativo necesario para construir una imagen de mayor resolución. La estimación del desplazamiento relativo entre las diversas imágenes (digamos k de ellas de tamaño $N \times M$) permite compensar ese desplazamiento, lo que proporciona $k \times N \times M$ valores en una retícula irregular que hay que transformar en una imagen definida en una retícula regular cuadrada. De lo dicho se deduce que este problema está ligado a muchos otros: el cálculo del desplazamiento relativo entre dos imágenes, problemas de restauración, de muestreo y de interpolación, por destacar algunos de los ingredientes principales.

Aumento de resolución temporal: Otro caso relevante en las aplicaciones es el del aumento de resolución temporal de una secuencia de vídeo, ya sea para conversión de formatos (pasar de los 25 cuadros por segundo del sistema PAL a los 30 de NTSC, por ejemplo), o en postproducción para generar efectos como el de cámara lenta [16, 17]. Problemas análogos se plantean si queremos cambiar el aspect-ratio de la secuencia.

Efectos en post-producción digital de cine: Como mencionamos antes, la práctica habitual es que la post-producción de cine se haga digitalmente aún cuando tanto el material original como el de exhibición sea película analógica 35mm. El metraje filmado se escanea a alta resolución, se procesa digitalmente, y se transfiere de nuevo a film para su proyección en cines. Tanto el escaneado de alta resolución como la transferencia a film son procesos muy costosos, pero que se justifican por la versatilidad que permite la post-producción digital frente a la clásica postproducción de laboratorio óptico: procesos como el dosificado de color, zooms, re-encadres, se pueden hacer digitalmente en mucho menos tiempo, con mucho mejores resultados y a un coste virtualmente independiente de la cantidad de pruebas (mientras que en efectos de laboratorio cada prueba tiene el costo -nada despreciable- de la impresión de una copia). Por otro lado hay efectos que sólo se pueden hacer bien digitalmente: noche americana [3], eliminación de objetos, dosificado no lineal de color, dosificado de color para objetos particulares de la secuencia, modificación de la profundidad de campo, eliminación de ruido y restauración de películas rayadas (*scratches*), restauración de películas con color desvaído, etc. Ello plantea una necesidad real de buenos y eficientes algoritmos de segmentación, tracking, restauración (*inpainting*), estimación de la profundidad de la escena (*depth from shading*), corrección de color, etc. ([18]).

b) La fotografía digital

El auge de la fotografía digital (con cámaras o teléfonos móviles) y las posibilidades que con ella se abren es un campo fértil para el desarrollo de herramientas de software dirigidas a la mejora de la calidad de las imágenes, a la mejora del contraste de color, la corrección de distorsiones, la mejora de la nitidez, el aumento de resolución, etc [4]. Con estas herramientas técnicas los usuarios pueden conseguir con el teléfono móvil fotografías de gran calidad. Este campo permite desarrollar y hacer efectivos muchos de los algoritmos básicos de procesamiento de imágenes.

c) Procesamiento de imágenes satelitales

Ésta es una de las áreas tradicionales del procesamiento de imágenes para la que se han desarrollado y se siguen desarrollando muchos algoritmos: deconvolución y eliminación de ruido, fusión de imágenes multispectrales, reconstrucción de modelos digitales de elevación de terreno y de topografía urbana. Por otro lado, el uso de imágenes satelitales en agricultura o teledetección plantea problemas de segmentación y de reconocimiento de patrones, sean éstos ríos, carreteras, zonas urbanas, o rurales, por no hablar de las numerosas aplicaciones militares [19, 20].

d) Reconstrucción tridimensional a partir de imágenes estéreo o vídeo

Ésta es otra de las áreas clásicas de la visión por ordenador, el de la reconstrucción tridimensional de una escena a partir de pares de imágenes estéreo o de secuencias de imágenes.

La base de estas aplicaciones es la geometría proyectiva y la calibración de cámaras y sus desarrollos han sido espectaculares en los últimos 15 años [5]. Entre otras múltiples aplicaciones mencionemos la de la reconstrucción tridimensional de sitios históricos o cascos urbanos a partir de vídeo lo que ofrecería posibilidades de navegación mucho más sofisticadas que las que podemos tener actualmente con Google Earth [6]. Finalmente y en el ámbito de las aplicaciones médicas, podemos mencionar a título de ejemplo la reconstrucción - y visualización - del tracto digestivo a partir de las imágenes obtenidas por endoscopia [7].

e) Reconocimiento de formas y búsqueda de imágenes en bases de datos

El reconocimiento de formas es uno de los objetivos fundamentales del análisis de imágenes y, a pesar de los desarrollos espectaculares que se han producido en tiempos recientes, aún estamos lejos de poseer algoritmos industriales, por ejemplo, para la detección de logos en la publicidad [8]. Las numerosas aplicaciones industriales van a ser un motor de desarrollo de este campo: piénsese en la necesidad de buscar imágenes por su contenido en la web (compárese con la facilidad para buscar textos) u otras bases de datos (piénsese en la posibilidad de identificar un objeto



concreto en una base de datos de objetos robados) [9]. Entre otras posibles aplicaciones industriales mencionemos las ligadas al control de calidad, piénsese en el control automático de la calidad en industrias de azulejos por análisis de los patrones de textura de los mismos.

Por otra parte, el reconocimiento de formas o patrones es el hecho clave de la visión humana y su estudio está relacionado con la psicología y la neurociencia [21]. La interacción de estas disciplinas es fundamental para el avance de la visión computacional.

f) El campo del procesamiento y análisis de las imágenes médicas

Éste es uno de los campos más vastos dentro del procesamiento de imágenes y abarca desde su adquisición hasta su procesamiento e interpretación como ayuda al diagnóstico. Aunque existen numerosos equipos de investigación dentro de las empresas que construyen los aparatos de adquisición, el análisis de estas imágenes como ayuda al diagnóstico es un campo en el que queda mucho camino por recorrer. Existen grupos de investigadores especializados en el estudio de imágenes de cada uno de los órganos del cuerpo humano: cerebro, corazón, hígado, sistema vascular, etc. En muchos de estos casos, el primer paso es la segmentación para la reconstrucción tridimensional del órgano en cuestión a partir de las imágenes dadas. Para mencionar un ejemplo que ilustra las múltiples implicaciones del problema, mencionemos el caso del estudio del sistema vascular: este estudio requiere la obtención de modelos anatómicos individualizados, el estudio y caracterización morfológica del dicho sistema o alguna de sus partes donde haya un problema, y eventuales simulaciones de dinámica de fluidos para comprender las presiones a que se ve sometida la estructura en cuestión [10, 11]. Hemos mencionado este caso para dar a entender la complejidad del problema y la necesidad de una interacción multidisciplinar que sólo está al alcance de equipos muy potentes y estructurados.

g) Codificación y compresión de imágenes

Este área es una de las que han tenido un desarrollo más espectacular en los últimos 20 años llevando al desarrollo de los estándares actuales de compresión tipo JPEG o MPEG [13, 14]. Los desarrollos en compresión de imágenes han estado ligados a la palabra clave "wavelets" (ondículas, en castellano) y el actual JPEG 2000 [12] está basado en esta herramienta de compresión.

Las wavelets se desarrollaron a partir de una interacción fructífera entre ingenieros, físicos y matemáticos, y podemos decir en este caso, que los matemáticos jugaron un papel fundamental [15]. Los nombres de Jean Morlet, Alex Grossman, Yves Meyer, Stéphane Mallat, Ingrid Daubechies son nombres de actores principales en este capítulo. Actualmente, la compresión sigue atrayendo la atención de muchos investigadores (como se refleja en las numerosas publicaciones en revistas y congresos) debido a la cantidad inmensa de datos que almacenar.

h) Gráficos por ordenador y procesamiento de superficies

Ésta es un área muy desarrollada con fuertes conexiones industriales como por ejemplo con la industria del automóvil, la industria del cine o los videojuegos. Existen ya multiples herramientas para crear imágenes sintéticas o representar y visualizar superficies de objetos industriales con la iluminación adecuada [7]. Algunos de los problemas arriba mencionados, como la reconstrucción del tracto digestivo a partir de endoscopias, o la reconstrucción de zonas urbanas o sitios históricos a partir de vídeos reales, plantean numerosos problemas de reconstrucción y visualización de superficies, que atraen la atención de los investigadores [22], y que merecen ser impulsados en el contexto nacional.

Digamos, finalmente, que estos ejemplos son sólo eso, ejemplos en un vasto mar de aplicaciones en las cuales interviene el procesamiento de imágenes y la visión por ordenador, y esto no es más que un ejemplo de procesamiento de datos siendo éstos datos con estructura geométrica bidimensional (como es el caso de las fotografías), tridimensional (caso de imágenes 3D o secuencias de vídeo), o incluso tetra-dimensional (para el caso de imágenes 3D en movimiento).

Bibliografía

- [1] De Haan, G.; Bellers, E.B; *De-interlacing: an overview*. Proc. IEEE 86(9), (1998), (1839-1857).
- [2] De Haan, G.; *Progress in Motion Estimation for consumer video Format Conversion*, IEE Trans. Consumer Elect. 46(3), (2000), (449-459).
- [3] Haro, G.; Bertalmío, M.; Caselles, V.; *Visual Acuity in Day for Night*. Int. J. Comp. Vision 2006, 69, (2006), (109-117).
- [4] DXO: *Photography, D-SLR and high end digicam automatic image quality enhancement*, <http://www.dxo.com/en/photo/home/default.php>
- [5] Hartley, R.; Zisserman, A.; *Multiple View Geometry in Computer Vision*, Cambridge University press, (2004).
- [6] Dick, A.R.; Torr, P.H.S.; Cipolla, R.; *Modelling and Interpretation of Architecture from Several Images*. International Journal of Computer Vision archive 60(2), (2004), (111-134).
- [7] Greiner, G.; *Personal communication*. <http://www9.informatik.uni-erlangen.de>.
- [8] Doermann, D.S.; Rivlin, E.; Weiss, I.; *Logo recognition using geometric invariants Document Analysis and Recognition, 1993.*, Proceedings of the Second International Conference on (1993), (894-897).
- [9] Sagarmay, D.; Zhang, Y.; *An overview of content-based image retrieval techniques, Advanced Information Networking and Applications, 2004. AINA 2004. 18th International Conference, (2004), (59-64).*
- [10] Frangi, A.F.; Niessen, W.J.; Viergever, M.A.; *Three-dimensional modeling for functional analysis of cardiac images, a review*, Medical Imaging, IEEE Transactions on Volume 20, Issue 1, (Jan 2001), (2-5).
- [11] Frangi, A.F.; Laclaustra, M.; Lamata, P.A.; *A registration-based approach to quantify flowmediated dilation (FMD) of the brachial artery in ultrasound image sequences*; Medical Imaging, IEEE Transactions on Volume 22, Issue 11, (2003), (1458-1469).



- [12] Rao, K.R.; Huh, Y.; *JPEG 2000, Video/Image Processing and Multimedia Communications* 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom 16-19, (2002), (1-6).
- [13] Wallace, G.K.; *The JPEG still picture compression standard*, Consumer Electronics, IEEE Transactions on Volume 38, Issue 1, (Feb. 1992), (27-34).
- [14] Sikora, T.; *MPEG digital video-coding standards*, Signal Processing Magazine, IEEE Volume 14, Issue 5, (1997), (82-100).
- [15] Mallat, S.; *A Wavelet Tutorial of Signal Processing*, Academic Press, (1999).
- [16] Shechtman, E.; Caspi, Y.; Irani, M.; *Space-time super-resolution*, Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 27, Issue 4, (2005), (531-545).
- [17] Irani, M.; Peleg, S.; *Super resolution from image sequences*, Pattern Recognition, 1990. Proceedings., 10th International Conference on Volume II, (1990), (115-120).
- [18] Dettmer, R.; *Digital cinema: a slow revolution*, IEE Review Volume 49, Issue 10, (2003), (46-50).
- [19] *Remote Sensing and Image Processing*, <http://www.landcareresearch.co.nz/services/remotesensing/>
- [20] *TTI Production, Earth Sciences and GIS, Satellite Mapping*. <http://www.tti.fr/en/>
- [21] Deco, G.; Rolls, E.; *Computational Neuroscience of Vision*. Oxford University Press, (2001).
- [22] Rumpf, M.; <http://numod.ins.uni-bonn.de/people/rumpf/rumpf.shtml>.



Capítulo VI

MATEMÁTICAS Y FÍSICA TEÓRICA

POR OSCAR GARCÍA PRADA
CONSEJO SUPERIOR
DE INVESTIGACIONES CIENTÍFICAS (CSIC)
Departamento de Matemáticas
Serrano 121
28006 Madrid
Correo electrónico: oscar.garcia-prada@uam.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Galileo said that “the book of nature is written in mathematical language”. This idea is totally confirmed by the development of modern science and especially by physics. It is surprising to see how the different mathematical theories always end up describing some phenomenon in nature. Nothing expresses this better than the much-quoted phrase of Eugene Wigner, who refers to “the unreasonable effectiveness of mathematics in physics”.

Mathematics and physics have a long and fruitful history of interaction. In the classical era, associated with names such as Newton, Laplace and Hamilton, little distinction was made between the disciplines of mathematics and physics. Geometry and analysis, especially the study of differential equations, played a unifying role. The modern era, however, saw a growing separation between mathematics and physics, although of course the two still influenced each other. An important example of interaction is that between Einstein's general relativity and Riemannian geometry. Einstein's theory, which replaces Newton's theory of gravitation, is based on radically different ideas, linked to non-Euclidean geometries introduced by Riemann in the 19th century, when nobody suspected that so abstract mathematics had anything to do with reality! Another crucial interaction takes place with the appearance of quantum mechanics in the 1920s. When at the beginning of the 20th century the theory of Hilbert spaces was being developed in mathematics, nobody imagined that this would supply the basic language to express the laws that govern the quantum phenomena in nature.

The fact is that the interaction between mathematics and physics is a two-way path, producing benefits for both disciplines. On the one hand, the theories developed by mathematicians are being used by the physicists more and more frequently in the current developments in physics. But what is really surprising is that ideas coming from physics are giving rise to new and important results in mathematics. In this regard, in the last thirty years, the interaction between geometry and physics has produced one of the most refreshing events in 20th century mathematics, with enormous ramifications and a scope so vast that can dominate the mathematics of the 21st century.

After the successful interaction between Einstein's theory of relativity and differential geometry there was a gradual divergence between geometry and physics. But all this changed abruptly in the middle 1970s after the emergence of gauge theories with their differential-geometric background, as the favoured framework for the quantum field theory of elementary particles. From a mathematical point of view, gauge theory has produced spectacular results in the study of the topology and the geometry of low-dimensional manifolds.

In spite of the great success of quantum field theory with its surprising physical predictions, and its mathematical applications, this is an incomplete theory since it does not incorporate the theory of gravitation —an unresolved central problem in theoretical physics. However, in the 1980s string theory has emerged as a serious candidate to solve this problem. From a mathematical point of view, string theory has motivated tremendous developments not only in differential geometry and topology, but most surprisingly, in algebraic geometry. The impact of physics on geometry has

been such that, as Michael Atiyah suggests, one is tempted to reverse the Wigner's dictum and wonder at “the unexpected effectiveness of physics in mathematics”.

It is difficult to predict the future, but judging from the activity developed in last three decades and its enormous implications, everything indicates that the interaction between mathematics and physics has a long and bright future.

Introducción

Todo el mundo sabe que los físicos, químicos, biólogos y otros científicos utilizan constantemente las matemáticas en su trabajo. Lo que es menos sabido, sin embargo, es hasta qué punto las matemáticas están inextrecablemente unidas a las teorías científicas que describen las leyes de la naturaleza. Como dice Galileo “el libro de la naturaleza está escrito en lenguaje matemático”. Esta idea se ve plenamente confirmada por el desarrollo de la ciencia moderna y especialmente por la física. Resulta sorprendente cómo las diversas teorías matemáticas acaban siempre por describir algún fenómeno de la naturaleza. Nada expresa mejor este hecho que la muy citada referencia de Eugene Wigner a la “irrazonable eficacia de las matemáticas en la física”.

En este artículo, intentaremos ilustrar con algunos ejemplos esta extraordinaria eficacia de las matemáticas en la física. Pero lo cierto es que la interacción entre las matemáticas y la física da lugar a un camino de ida y vuelta con beneficios para ambas disciplinas. Los matemáticos se sienten muy satisfechos de ver que teorías que ellos han desarrollado resultan ser útiles a los físicos y están siendo utilizadas con mayor frecuencia en los más recientes desarrollos de la física. Sin embargo, lo que resulta más sorprendente es que ideas provenientes de la física hayan dado lugar a nuevos e importantes resultados en matemáticas. En este sentido, en los últimos treinta años la interacción entre geometría y física ha dado lugar a uno de los eventos más refrescantes en las matemáticas del siglo XX, con enormes ramificaciones y un alcance tan vasto que puede llegar a dominar las matemáticas del siglo XXI. Como sugiere Michael Atiyah, uno se siente tentado de invertir la frase de Wigner y reflexionar sobre la “inesperada eficacia de la física en las matemáticas”.

Matemáticas y Física: una larga historia de intercambios

En el periodo “clásico”, asociado a nombres tales como Newton, Laplace y Hamilton, se hacía poca distinción entre las disciplinas de matemáticas y física. La geometría y el análisis, principalmente el estudio de ecuaciones diferenciales, jugaron un papel unificador fundamental. Es evidente que la física hizo progresos espectaculares gracias a la creación por Newton y Leibniz del cálculo diferencial e integral. Otro ejemplo importante de interacción, entre los muchos que se pueden mencionar en este periodo, lo constituye la creación de las series de Fourier para describir la propagación del calor. Se trata de sumas infinitas de funciones trigonométricas que han jugado posteriormente un papel fundamental en las ciencias y en las tecnologías.



La era moderna ha visto, sin embargo, una creciente separación entre las matemáticas y la física. Por supuesto, ambas se han influido mutuamente. Por ejemplo, la teoría del electromagnetismo desarrollada por Maxwell a finales del siglo XIX, está en los orígenes de la teoría de Hodge de formas armónicas desarrollada en los años 1930. Otro ejemplo aún más conocido lo proporciona la interacción entre la teoría de la relatividad de Einstein y la geometría Riemanniana. La teoría de Einstein, que viene a suplantarse la teoría de la gravitación de Newton, reposa en conceptos radicalmente distintos, ligados a geometrías no euclídeas introducidas por Riemann en el siglo XIX, cuando nadie sospechaba que unas matemáticas tan abstractas tuvieran algo que ver con la realidad! No cabe duda de que la teoría de la relatividad ha impulsado a su vez el desarrollo de la geometría Riemanniana. También en el siglo XIX, Hamilton desarrolló la mecánica clásica, introduciendo el denominado formalismo hamiltoniano. La mecánica clásica ha dado lugar a la geometría simpléctica, cuyo pleno desarrollo sólo se ha producido en tiempos recientes.

Es en la teoría de la relatividad de Einstein donde por primera vez se aplica de forma consciente el principio de simetría. Este principio, que expresa la regularidad de ciertas propiedades o la conservación de determinadas cantidades, jugará un papel fundamental en toda la física teórica del siglo XX. Pues bien, detrás de la idea de simetría está la noción de grupo de Lie. La teoría de grupos de Lie iniciada por Lie en el siglo XIX, y posteriormente desarrollada por Cartan y otros, proporcionará uno de los objetos más unificadores dentro de las matemáticas y de la física teórica. La simetría de la teoría de la relatividad no es más que la invariancia de las ecuaciones de Einstein bajo la acción del grupo de Lie de transformaciones de Lorentz. La geometría Riemanniana, la geometría de Kähler de variedades compleja y la geometría simpléctica son simplemente geometrías asociadas a los grupos ortogonal, unitario y simpléctico, respectivamente —los denominados grupos de Lie clásicos, que junto con algunos grupos de Lie excepcionales han jugado un papel muy importante en la historia de las matemáticas del siglo XX—. Este modo de ver la geometría es ya preconizado por Klein en su “Programa de Erlangen”.

En una línea más analítica, la teoría de grupos de Lie ha dado lugar al denominado análisis armónico no conmutativo. Se trata de una generalización de la teoría de Fourier, donde las series de Fourier o integrales de Fourier corresponden esencialmente a los grupos de Lie conmutativos definidos por el círculo y la recta real. Cuando uno reemplaza estos grupos por grupos de Lie más complicados se obtiene una bella y elaborada teoría en la que se combinan el análisis y la teoría de representaciones de grupos de Lie.

En teoría de números, el denominado “Programa de Langlands” se desarrolla dentro del contexto de la teoría de grupos de Lie. Todo grupo de Lie lleva asociada una teoría de números y un programa de Langlands, desarrollado en gran medida. Este programa ha influido en gran parte del trabajo realizado en teoría algebraica de números durante la segunda mitad del siglo XX. El estudio de formas modulares, incluido el trabajo de Wiles sobre el último teorema de Fermat, entra dentro de este contexto.

Otra interacción crucial entre matemáticas y física se produce con la aparición de la mecánica cuántica en los años 1920. Cuando a principios del siglo XX se desarrolla en matemáticas la teoría de espacios de Hilbert, nadie imaginaba que ésta proporcionaría el lenguaje básico en el que

se expresan las leyes que rigen los fenómenos cuánticos de la naturaleza. Este papel de los espacios de Hilbert en la mecánica cuántica dio a su vez un tremendo impulso a la teoría espectral de operadores y la teoría de representaciones de grupos de Lie.

Hay otros desarrollos importantes en los que las matemáticas han ido por delante de la física. Entre estos, cabe mencionar la teoría de espinores, cuyos orígenes se encuentran en Hamilton y Clifford y que fue desarrollada posteriormente por Cartan y Weyl. Su introducción por Dirac en la física ha resultado absolutamente básica ya que estos objetos permiten describir los fermiones de las teorías físicas. En matemáticas los espinores se entiende bien algebraicamente y su papel en la teoría de representaciones del grupo ortogonal proporciona la conexión con la física. No obstante el significado geométrico de los espinores sigue siendo un verdadero misterio.

Recientes interacciones entre geometría y física

Tras la exitosa y fructífera interacción entre la teoría de la relatividad de Einstein y la geometría diferencial, se fue produciendo un gradual alejamiento entre la física y la geometría. La plétora de nuevas partículas surgidas con la introducción de los grandes aceleradores después de la segunda guerra mundial, y la lucha de los teóricos por manipular los infinitos que plagaban las teorías cuánticas de campos, estaban lejos de interesar a los matemáticos. Por supuesto, hubo matemáticos que desesperadamente trataron de sentar desde la retaguardia los fundamentos teóricos, mientras que los físicos adquirían gran virtuosismo en el manejo de los diagramas de Feynman y en el uso de las simetrías de los grupos de Lie, que gradualmente pusieron cierto orden en el panorama. Pero todos estos esfuerzos deben poco a la comunidad matemática en general.

Todo esto cambió abruptamente en la mitad de los años 1970 con el surgimiento de las teorías gauge, en particular la teoría de Yang-Mills, introducida veinte años atrás por los físicos Yang y Mills. Fundamentadas en la geometría diferencial, estas teorías proporcionan la herramienta apropiada para describir la teoría cuántica de campos de partículas elementales. No sólo se disponía ya de un lenguaje común sino que progresivamente se fue viendo que los aspectos más delicados en un frente y otro estaban profundamente relacionados. Esto lo ilustra maravillosamente la relación entre las “anomalías” de la teoría cuántica de campos —la violación de determinadas simetrías de una teoría clásica al pasar a la teoría cuántica— y el teorema del índice de Atiyah-Singer por el que ambos autores han recibido recientemente el premio Abel (que se añade a la medalla Fields recibida por Atiyah en 1966).

La base geométrica de las teorías de Yang-Mills se encuentra en la geometría diferencial de fibrados cuyo objeto fundamental es el estudio de conexiones y su correspondiente curvatura. Uno de los ingredientes fundamentales es el grupo de gauge —un grupo de Lie que describe cierto tipo de simetrías internas, denominadas simetrías de *gauge*—. El modelo estándar, por el que Glashow, Salam y Weinberg recibieron el premio Nobel en 1979, consigue unificar en una sola teoría, con una elección conveniente de grupo de gauge, tres de las cuatro interacciones fundamentales



de la naturaleza: la interacción electromagnética, la interacción débil y la interacción fuerte. Estas dos últimas fuerzas son relevantes en los fenómenos cuánticos a pequeña escala, que tienen lugar en los núcleos de los átomos y son responsables en gran medida de la estabilidad de la materia.

Desde el punto de vista matemático, las teorías de Yang-Mills fueron utilizadas de modo espectacular por Donaldson (medalla Fields 1986) en los años 1980 en el estudio de la clasificación de las variedades diferenciables de dimensión cuatro —precisamente la dimensión del espacio-tiempo de Einstein—. Haciendo uso del espacio de soluciones a las ecuaciones de Yang-Mills —ecuaciones en derivadas parciales no lineales—, Donaldson fue capaz de penetrar de manera profunda en la estructura de las variedades de dimensión cuatro. Otra sorprendente aplicación de la teoría cuántica de campos es al estudio de nudos en variedades de dimensión tres. Los trabajos de Witten en este campo le valieron la medalla Fields en 1990. Más recientemente, la teoría de Seiberg-Witten, que involucra espinores además de conexiones ha permitido obtener algunos de los resultados de Donaldson de un modo más simple.

A pesar del éxito de la teoría cuántica de campos con sus sorprendentes predicciones físicas, y sus aplicaciones a problemas básicos de matemáticas, ésta resulta ser una teoría incompleta al no incorporar la fuerza gravitatoria —la otra interacción fundamental de la naturaleza—. Éste es uno de los problemas centrales de la física teórica que está aún sin resolver. En los años 1980 ha surgido una teoría cuya creciente coherencia a lo largo de los últimos años la está convirtiendo en seria candidata a la resolución del problema: la teoría de cuerdas. De modo muy aproximado, se puede decir que en esta teoría los objetos fundamentales no son partículas asimilables a puntos, como se considera en las teorías cuánticas de campos tradicionales, sino una especie de cuerdas minúsculas. Las diferentes partículas conocidas corresponderían a distintos modos de vibración de estas cuerdas, algo análogo a como las diferentes vibraciones de una cuerda de violín corresponden a las distintas notas musicales. La teoría de cuerdas incorpora además una simetría que los físicos consideran imprescindible: la supersimetría. Así se habla de teoría de supercuerdas. La supersimetría intercambia bosones (partículas con espín entero, como los fotones) con fermiones (partículas con espín semientero, como los electrones).

Una de las consecuencias más dramáticas de la teoría de cuerdas es que, para ser una teoría coherente, el espacio-tiempo debe tener más dimensiones que las cuatro usuales. En la teoría actual el número de dimensiones es diez (aunque en la más reciente denominada teoría M, directamente relacionada con la teoría de cuerdas, esta dimensión es once). Se podría decir que en cada punto del espacio-tiempo usual hay un espacio compacto de dimensión seis de tamaño extremadamente pequeño. Esta es una idea que tiene sus orígenes en la teoría de Kaluza-Klein de los años 1920 —cuando los geómetras llevaban ya tiempo explorando espacios de dimensión superior—. Estos “pequeños” espacios son de gran interés y conectan con teorías geométricas de enorme actualidad (variedades de Calabi-Yau, variedades con holonomía especial, etc.).

Uno de los hechos más notables en la moderna interacción entre geometría y física es el papel tan preponderante que juega la topología. Retrospectivamente, el origen de esto se encuentra

ya en Dirac, cuya explicación de la cuantización de la carga eléctrica es esencialmente topológica. Pero todavía más notable resulta el papel que desde hace unos años y en la actualidad juega la geometría algebraica. Resulta absolutamente sorprendente que dualidades básicas de la teoría de cuerdas y de la teoría M se expresen en términos de objetos tan abstractos como las categorías derivadas, y que los espacios de moduli que parametrizan estructuras geométricas, extensamente estudiados por los geométricos algebraicos, formen hoy parte de la jerga habitual de la física teórica. Esta interacción está dando lugar en estos momentos a uno de los periodos más fructíferos de la historia de la geometría algebraica.

Perspectivas futuras

Es difícil predecir el futuro, pero a juzgar por la actividad desarrollada en las últimas tres décadas y sus enormes implicaciones, todo parece indicar que la interacción entre matemáticas y física tiene un largo y brillante porvenir. No es en vano que este campo de investigación ha atraído y sigue atrayendo a algunos de los más brillantes investigadores, lo que se refleja, por ejemplo, en el número considerable de medallas Fields concedidas en los últimos años en esta área. Uno de los aspectos más interesantes y ricos de esta interacción es su carácter marcadamente multidisciplinar, involucrando un gran número de áreas fundamentales de la física teórica y de las matemáticas. Es absolutamente fascinante y sorprendente constatar a lo largo de la historia como incluso las matemáticas más básicas y abstractas juegan un papel fundamental en la descripción de teorías físicas, que posteriormente dan lugar a nuevas y avanzadas tecnologías. Parece claro que esto seguirá sucediendo no sólo en relación con la física, sino con otras ciencias, como la biología, para las que muy probablemente nuevas y fundamentales teorías matemáticas deberán ser desarrolladas.

Bibliografía

- [1] Atiyah, M.; *Collected Works*, Volume 6, Oxford University Press, (2004).
- [2] García-Prada, O.; *El teorema del índice de Atiyah-Singer*, EL PAIS, (26 de mayo de 2004).
- [3] Greene, B.; *The Elegant Universe*, Jonathan Cape, London, (1999).



Capítulo VII

MATEMÁTICAS Y MODELIZACIÓN
EN CIENCIA DE MATERIALES

POR FRANCISCO GUINEA LÓPEZ
CONSEJO SUPERIOR
DE INVESTIGACIONES CIENTÍFICAS (CSIC)
Departamento de Teoría
Instituto de Ciencia de Materiales de Madrid
Cantoblanco. 28049 Madrid
Correo electrónico: paco.guinea@icmm.csic.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Models play a major role in Materials Science. They help to understand, and even predict, the behavior of novel materials. Models allow us to put bounds on the values of parameters where a given material may show interesting features. They are also of great help in the design of devices based on new materials.

Models in Material Science come with different levels of mathematical rigor. The ultimate goal is a quantitative understanding of a system where many competing interactions play a role. A researcher working on the modelling of new materials requires a significant mathematical background, acquired during a Ph. D. period fully devoted to mathematical models.

The tools used in this field are also very varied. Some problems are dealt with using methods of numerical analysis. The existence of simplified versions of some models for which a full analytical solution can be obtained is of great help. Finally, many useful results are obtained by means of perturbative calculations and asymptotic analysis.

There are a number of topics in Materials Science where a long and fruitful experience of collaboration between mathematicians, physicists and chemists exists, such as heat diffusion, turbulence, or continuum elasticity. The greater complexity of the problems addressed now implies that the number of active researchers, and techniques used is much greater than in the past. Among the many interesting problems being studied today, we mention I) nanomechanic oscillators [1, 2], II) magnetic materials whose electronic properties can be finely tuned [3, 4], and III) pattern formation out of equilibrium.

Introducción

Uno de los aspectos de la investigación en Ciencia de Materiales es la modelización del comportamiento y propiedades de nuevos materiales. Ello permite, por ejemplo, acotar el tipo de sistemas y parámetros a estudiar experimentalmente para una funcionalidad dada, entender resultados que se apartan de lo esperado, o incluso predecir nuevos comportamientos aún no observados.

Los modelos en Ciencia de Materiales tiene un nivel y rigor matemático muy variado, aunque se pretende una comprensión lo más cuantitativa posible del fenómeno a estudiar. En general, los investigadores que trabajan en este campo han necesitado ampliar su formación matemática en unos estudios de doctorado.

Las herramientas matemáticas que se utilizan son también muy variadas. Así, el estudio de los materiales cristalinos o de las transiciones de fase requiere unos conocimientos extensos de simetrías y teoría de grupos. La simulación del comportamiento de materiales se trata con téc-

nicas de análisis numérico. Es de gran utilidad la definición de modelos con solución exacta, que sirven de referencia para el estudio de procesos más complejos pero más próximos a los materiales de interés. Finalmente, muchos resultados útiles se consiguen mediante técnicas perturbativas o desarrollos asintóticos en las proximidades de situaciones bien entendidas.

Existen temas en Ciencia de Materiales con una gran tradición en interacciones fructíferas entre la matemática avanzada y las investigaciones prácticas. Aspectos clásicos son los estudios de difusión del calor, de la turbulencia, o la teoría de la elasticidad. La mayor complejidad de los problemas estudiados ha conducido, en general, a una mayor sofisticación de las técnicas matemáticas utilizadas, y a un aumento del número relativo de investigadores dedicados a los aspectos teóricos de la Ciencia de Materiales. A continuación se describen algunos ejemplos muy actuales que ilustran el tipo de temas donde se aplican conocimientos matemáticos avanzados para la comprensión de fenómenos de gran interés por sus aspectos de ciencia experimental básica, y por sus posibles aplicaciones prácticas.

Resonadores de muy pequeña dimensión

Un grupo de sistemas muy recientes que están siendo intensamente estudiados son los resonadores mecánicos de pequeña dimensión [1, 2] (nanoelectromechanical resonators, NEMs). Se trata de elementos mecánicos, como vigas o cables, de dimensiones menores de una micra. Los modos de vibración posibles tienen energías relativamente grandes, debido a sus pequeñas dimensiones y a las restricciones sobre las posibles longitudes de onda que ello impone.

Las técnicas actuales de investigación a bajas temperaturas permiten llegar a situaciones donde la energía térmica de estos osciladores es comparable con el cuanto de energía asociado a sus modos normales. Por lo tanto, se trata de elementos mecánicos, en muchos aspectos similares a los estudiados en muchas ramas de la ingeniería, pero que necesitan de una descripción que tenga en cuenta efectos cuánticos. Los sistemas con un gran número de átomos, como estos, pero que presentan propiedades asociadas a la mecánica cuántica, constituyen un tema de gran interés en investigación básica, dadas las paradojas que se producen al extrapolar las leyes de la mecánica cuántica al mundo macroscópico.

Desde un punto de vista más práctico, las pequeñas dimensiones de estos sistemas hacen que que sus frecuencias normales sean muy sensibles a pequeñas modificaciones. Ello hace que estos sistemas puedan ser utilizados como detectores de sustancias con gran precisión. La adhesión de una sola molécula orgánica de grandes dimensiones induce cambios en las propiedades mecánicas del sistema que pueden ser observados experimentalmente.

La modelización previa facilita considerablemente el diseño y estudio experimental de estos sistemas. Esta modelización requiere de técnicas matemáticas diversas y complejas. Son necesarias ecuaciones del continuo (elasticidad e hidrodinámica) que pueden ser tratadas analí-



tica o numéricamente. El efecto de las imperfecciones y el desorden se ve amplificado por las pequeñas dimensiones de los dispositivos, lo cual conlleva la necesidad de técnicas estadísticas sofisticadas. Finalmente, los aspectos cuánticos del problema implican la necesidad de relacionar las descripciones continuas, válidas a grandes escalas, con los modelos cuánticos a escalas atómicas.

Materiales magnéticos con propiedades electrónicas variables

Los ordenadores actuales contienen elementos magnéticos de almacenamiento, estables, de poco consumo energético, pero con tiempos de respuesta lentos, y circuitos electrónicos basados en semiconductores para la realización de operaciones, rápidos, pero de alto consumo, y volátiles. Un tema de gran interés en Ciencia de Materiales es el estudio de sistemas magnéticos conductores en los que las propiedades electrónicas sean altamente dependientes de las propiedades magnéticas.

Los materiales en estudio son especialmente complejos, dado que su composición ha de incluir elementos magnéticos, elementos que den lugar a electrones deslocalizados que puedan transmitir señales eléctricas, y elementos que den estabilidad al conjunto. Muchos esfuerzos se centran en el estudio de óxidos de metales magnéticos [3, 4].

La investigación en este tema se realiza de forma interdisciplinar, combinando la síntesis química, la caracterización de sus propiedades físicas, y la modelización del comportamiento esperable en nuevos compuestos aún por producir. La modelización es especialmente importante dado que, debido al elevado número de elementos a combinar, el número de compuestos posibles es muy grande. Se utiliza el análisis numérico a gran escala, para estudiar las propiedades globales de estabilidad de los compuestos. El estudio del posible orden magnético, y de las transiciones de fase asociadas, requiere de técnicas estadísticas. Las propiedades electrónicas fuera del equilibrio se pueden tratar mediante análisis perturbativos alrededor de situaciones de equilibrio sencillas. Finalmente, la complejidad de estos sistemas se ve aumentada por el papel importante que juegan diversos tipos de desorden, cuya modelización también requiere de técnicas estadísticas avanzadas.

Formación de patrones fuera del equilibrio

El estudio de la formación de patrones en procesos fuera del equilibrio es un campo muy amplio de la Ciencia de Materiales, además de tratarse de un tema de interés en otras áreas, como la Astronomía o la Biología. La síntesis de muchos materiales se produce en situaciones muy lejanas al equilibrio térmico. Se puede utilizar la tendencia de estos procesos a formar patrones característicos para la creación de estructuras autoorganizadas. Desde un punto de vista más fundamental, los procesos fuera del equilibrio constituyen un desafío teórico y experimental, en contraste con el conocimiento acumulado sobre los diagramas de fases de los materiales en equilibrio.

Una clase ampliamente estudiada de estos procesos es la formación de estructuras en los procesos de solidificación, como la generación de dendritas, o de formas hexagonales en la formación de hielo y nieve. Los modelos teóricos son fundamentales para aislar las causas más importantes de estos fenómenos. Un avance muy importante en este sentido, ha sido la comprensión de las estrechas relaciones entre el crecimiento de dendritas durante la formación de un cristal, la formación de dedos viscosos en la difusión de líquidos inmiscibles, la agregación que genera los granos de polvo que flotan en la atmósfera, o la estructura que forman las grietas durante la fractura de un material [5, 6].

Una parte importante de los avances en este campo se ha producido gracias al desarrollo de modelos matemáticos sofisticados. Estos modelos combinan el análisis numérico, técnicas de ecuaciones diferenciales en problemas con fronteras móviles, transformaciones conformes de unas formas a otras, y desarrollos asintóticos válidos cuando determinados parámetros de control son pequeños. Estos avances, a su vez, son aplicables a problemas en otras áreas, como la formación de grandes estructuras en el universo, o la aparición de las mismas formas en diversos seres vivos.

Conclusiones

Se han descrito diversos problemas, muy generales, en Ciencia de Materiales donde modelos matemáticos sofisticados son muy necesarios para el avance de en su comprensión. El tipo de modelos es muy amplio, y se utilizan muchos aspectos de la Matemática Aplicada.

Los modelos utilizados en Ciencia de Materiales sirven, además, para subrayar las analogías entre procesos observados en otras ramas de la Ciencia, facilitando el avance en temas de carácter interdisciplinar.

Bibliografía

- [1] Craighead, H.G.; *Nanoelectromechanical systems*, Science 290, (1532-1535), (2000).
- [2] Knobel, R.G.; Cleland, A.N.; *Nanometre-scale displacement sensing using a single electron transistor*, Nature 424, (291-293), (2003).
- [3] Coey, J.M.D.; Viret, M.; von Molnar, S.; *Mixed valence manganites*, Adv. in Phys. 48, (167-293), (1999).
- [4] Dagotto, E.; *Complexity in strongly correlated electronic systems*, Science 309, (257-262), (2005).
- [5] Witten, T.A.; Sander, L.M.; *Diffusion limited aggregation, a kinetic critical phenomenon*, Phys. Rev. Lett. 47, (1400-1403), (1981).
- [6] Vicsek, T.; *Fractal Growth Phenomena*, World Scientific, (Singapore, 1992).



Capítulo VIII

LA OPTIMIZACIÓN Y EL MÉTODO CIENTÍFICO EN LA TOMA DE DECISIONES

POR MARCO ANTONIO LÓPEZ CERDÁ
UNIVERSITAT D'ALACANT
Dpto. de Estadística e Investigación Operativa
Apdo. Correos 99
03080 Alicante
Correo electrónico: marco.antonio@ua.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Optimization is behind any planned human activity. Airlines companies schedule crews and aircraft to minimize cost. Investors seek to design portfolios in order to avoid excessive risks while they try to achieve high returns. Firms aim for a maximum of efficiency in planning their production processes.

Moreover, nature also optimizes, and physical systems tend to a minimum energy state. The molecules in an isolated chemical system react with each other until the total potential energy of their electrons is minimized. Finally, rays of light follow paths that minimize their travel time.

To use optimization one starts by identifying some *objective*, a quantitative measure of the performance of the systems which are being analyzed. This objective might be profit, time, potential energy, or any magnitude that can be represented by a single number. The objective will depend on certain features of the system, called *variables*. The goal is to find values of the variables that optimize the objective. Often the variables are *constrained*, in some way. The process of identifying objective, variables, and constraints is known as *modeling*. Constructing an appropriate model is the first step, sometimes the most important, in the optimization process. Once the model has been formulated, an optimization algorithm will be used to find its solution. There is no universal optimization algorithm, and each one is specially tailored to a particular type of optimization problems. Finally, the model may be improved by means of some *sensitivity analysis*, which reveals how much sensitive is the solution to changes in the data and in the model.

Nowadays, there are many research groups working in optimization in different Spanish universities. Some of these groups are more involved in theoretical research, and others are more interested in applications. These groups have achieved a high degree of international recognition, they publish frequently in the most prestigious scientific journals, and they are responsible of the organization of relevant scientific events. Some of these teams also practice R&D and act as consultants of important firms, organizations, etc. The optimization fields in which these groups have more relevant contributions are the following: global, multiobjective, stochastic, dynamics, combinatorial, data mining, routing, location, interior point methods, semi-infinite, and game theory.

La optimización, el hombre y la naturaleza

La optimización está presente en cualquier actividad planificada del ser humano. Las compañías aéreas planifican sus vuelos y la rotación de las tripulaciones con el afán de minimizar los costes o, lo que es equivalente, de maximizar sus beneficios. Los inversores orientan sus decisiones de forma que se minimicen los riesgos a la vez que se garanticen niveles de rentabilidad satisfactorios. En general, las industrias aspiran a una eficiencia máxima a la hora de diseñar sus productos y de organizar sus propios procesos productivos.

Por su parte la naturaleza también optimiza, y los sistemas físicos tienden a un estado de mínima energía. Las moléculas en un sistema químico aislado reaccionan entre ellas hasta que la energía potencial de sus electrones alcanza su mínimo valor. Los rayos de luz siguen aquellas trayectorias que minimizan la duración de su viaje.

Raíces históricas de la optimización

Desde los tiempos críticos de la Segunda Guerra Mundial, la investigación operativa ha venido aplicando el método científico y su principal instrumento, las matemáticas, al análisis y resolución de problemas complejos que surgen en diferentes ámbitos, como la industria, la administración, los negocios, la gestión, la salud pública, la política, la ecología, etc. La investigación operativa se ocupa de modelar, simular y/o optimizar problemas reales asociados al funcionamiento de sistemas de gran complejidad. Cabe afirmar que la investigación operativa trata de resolver estos problemas de una "forma óptima", y al servicio de este objetivo básico se desarrolló la optimización, también llamada programación matemática (aunque este término equívoco, ya que puede conducir a la impresión errónea de que el objetivo de esta disciplina consiste en "elaborar programas informáticos desde un enfoque matemático"). La optimización, junto con otras áreas de la investigación operativa, proporciona una base rigurosa para el análisis científico de las alternativas a considerar ante un problema de decisión complejo.

La optimización hunde sus raíces en el cálculo de variaciones y en los trabajos de Euler y Lagrange. El desarrollo de la programación lineal y del *método simplex*, por George Dantzig en los años 40 del siglo pasado, constituyó el estímulo fundamental para el progreso de la optimización moderna en la segunda mitad del siglo pasado. Otro de los eventos decisivos en los orígenes de la optimización fue la publicación del histórico artículo de Kuhn y Tucker en 1951 en que se establecían las condiciones necesarias de optimalidad para problemas de optimización generales, incluyendo restricciones en forma de desigualdad: eran las llamadas *condiciones de Kuhn y Tucker* (también conocidas como de Karush-Kuhn-Tucker -KKT, en breve- en reconocimiento de la existencia de una versión previa, más débil, debida a Karush en 1939).

La relación entre la programación lineal y la teoría de juegos estuvo presente en los primeros análisis de la teoría de la dualidad y su relación con el teorema del minimax de von Neumann.

En los años 50 del siglo último la investigación estuvo centrada en algunas subclases importantes de problemas que poseen una estructura particular, proponiéndose para su resolución algoritmos específicos que explotan la estructura del problema en beneficio de la eficiencia computacional del método. Mencionemos, por ejemplo, los problemas de *transporte*, *transbordo*, *asignación*, *flujo en redes*, *secuenciación*, etc., problemas todos muy presentes en una gran variedad de aplicaciones en campos tales como la planificación de proyectos, las técnicas de localización, el diseño de itinerarios de distribución y, en un contexto más general, la planificación de grandes sistemas logísticos.



A mitad de los años 50 aparecen aplicaciones del método simplex a ciertos problemas no lineales. Tal es el caso de la *programación separable* consistente en una ligera variante del método simplex que permite abordar un problema no-lineal reemplazando las funciones no-lineales presentes en el problema mediante poligonales convenientes.

También es ese periodo se asiste al desarrollo de diferentes algoritmos eficientes para resolver el problema de *programación cuadrática*, que es el problema no-lineal más sencillo, y para el que las condiciones de KKT proporcionan información muy valiosa para la resolución exacta del problema. Precisamente en las condiciones de KKT se basaron los métodos más importantes para resolver un problema no-lineal, propuestos en el periodo 1955-1970 (métodos de direcciones factibles, métodos de penalización, métodos de programación secuencial, etc.). No obstante, hasta los años 90, con el desarrollo pleno de la informática, no fue posible su aplicación eficiente a problemas reales.

El modelo de optimización

La optimización arranca con la identificación de un *objetivo*, o medida cuantitativa de la realización del proceso o sistema estudiado. Este objetivo puede ser beneficio, tiempo, energía potencial, o cualquier cantidad o combinación de cantidades que puedan ser representadas numéricamente. El objetivo dependerá de ciertas características del sistema, llamada *variables*. Nuestro propósito es determinar valores de las variables que optimicen el objetivo. A menudo las variables están *restringidas*, de alguna manera. Por ejemplo, cantidades tales como la densidad de un electrón en una molécula, o la tasa de interés de un préstamo no podrán tomar valores negativos.

El proceso de identificar el objetivo, las variables, y las restricciones, en relación con un problema dado, es conocida como fase de *modelación*. La construcción de un modelo adecuado es la primera etapa, a veces la más importante, en un proceso de optimización. Si el modelo es demasiado simplista no proporcionará la suficiente información sobre el problema real investigado, pero si es complejo en exceso, puede resultar demasiado difícil de resolver, es decir de abordar numéricamente.

Una vez que el modelo ha sido formulado, un algoritmo de optimización será aplicado para encontrar una solución. Usualmente, los modelos y los algoritmos son lo suficientemente complejos como para requerir la ayuda del ordenador en la implementación de los cálculos. No existe un algoritmo de optimización de validez universal. Más bien existen numerosos algoritmos, cada cual especialmente diseñado para resolver determinado tipo de problemas. Es responsabilidad del usuario elegir el método más adecuado a su aplicación específica. Esta elección es de gran trascendencia, siendo la clave de si el problema es resuelto de forma rápida o lenta o si, lo que es más grave, no se llega al alcanzar nunca solución alguna.

Después de que un algoritmo ha sido aplicado, tenemos que ser capaces de reconocer si ha conducido a una solución óptima o si, por el contrario, ha proporcionado una solución que no lo es.

En muchos casos se dispone de elegantes expresiones matemáticas, conocidas como *condiciones de optimalidad*, para comprobar que la solución suministrada por el algoritmo es ciertamente óptima. Cuando las condiciones de optimalidad no son satisfechas en un determinado punto generado por el algoritmo, suelen proporcionar por defecto información muy útil acerca de cómo podemos mejorar la solución actual, y aproximarnos de forma secuencial a un óptimo. Finalmente, el modelo puede ser perfeccionado aplicando técnicas tales como el *análisis de sensibilidad*, que revelan la sensibilidad de la solución a los cambios en el modelo y en los datos.

Formalmente, el problema de optimización puede ser formulado en los siguientes términos:

$$\text{mín } f(x) \text{ sujeto a } h_i(x) = 0, i = 1, 2, \dots, p, : \quad g_j(x) \leq 0, j = 1, 2, \dots, m, : \quad x \in X \subset \mathbb{R}^n$$

Todas las funciones que intervienen en el modelo anterior son escalares (real-valoradas). Los vectores $x \in X$ que satisfacen adicionalmente todas las restricciones se llaman *soluciones factibles*, y aquellas soluciones factibles en las que la función objetivo alcanza su valor mínimo se llaman *soluciones óptimas*.

Habitualmente tendremos que practicar algunas transformaciones sencillas para expresar nuestro problema de optimización en el formato anterior, pero estas transformaciones son realizadas de forma automática por la mayor parte de los programas y del software al uso.

Un caso especial de singular importancia es el modelo de *programación lineal*, en el que todas las funciones que intervienen en el problema (la función objetivo, y las que aparecen en las restricciones) son lineales. En multitud de ocasiones las variables serán, por propia naturaleza, enteras, dando lugar a la llamada programación matemática con *variables enteras*.

Como se ha comentado, no podemos aspirar a elaborar una teoría general sobre "cómo plantear y resolver problemas" pues, en la mayoría de ocasiones, cada modelo asociado a un problema concreto requiere una teoría propia. A continuación describimos, a título de ejemplo, uno de los modelos simples más sugerentes, cuya formulación, en primera instancia, resulta aparentemente muy distante del modelo general presentado más arriba. Se trata del *problema de localización*:

Sea $I = \{ 1, 2, \dots, N \}$ un conjunto de potenciales ubicaciones de plantas para la producción de producto diferenciado. Una planta puede ser creada en cada posible localización $i \in I$, lo cual comporta un coste c_i . Asumiremos que cada una de las plantas construida puede proporcionar una cantidad ilimitada del producto.

Sea un conjunto $J = \{ 1, \dots, M \}$ de clientes que tienen una cierta demanda del producto en cuestión. Para cada par (i, j) el coste global del proceso y de transporte viene dado por $g_{ij} \geq 0$. El objetivo será determinar un subconjunto de ubicaciones potenciales $S \subset I$ y abrir las plantas correspondientes con el criterio de minimizar los costes totales. El problema puede ser modelizado como sigue:



$$F(S) = \sum_{i \in S} c_i + \sum_{i \in S} \min_{j \in S} g_{ij} \longrightarrow \min_{S \in I}$$

El problema recién formulado es la generalización del bien conocido *problema del cubrimiento* y, en consecuencia, es un problema de gran complejidad computacional (NP-hard). Métodos exactos, algoritmos aproximados con niveles garantizados, heurísticas lagrangianas, algoritmos de iteración aleatoria o de búsqueda local han sido propuestos para resolver este problema. Algunas clases de problemas resolubles en tiempo polinomial fueron identificadas. Un survey interesante puede encontrarse en [MF90].

Presente y futuro de la optimización

Hoy en día existen grupos de investigación dedicados a la optimización en muchas universidades españolas, y en particular en las universidades de la Comunidad Autónoma de Madrid. Algunos de estos grupos desarrollan una investigación de carácter teórico, mientras que otros ponen un mayor énfasis en las aplicaciones, o en cuestiones algorítmicas y metodológicas. Un indicio común del alto nivel de tales grupos, y de su proyección internacional, es su activa presencia en foros internacionales como los grupos de trabajo de EURO (Association of European Operational Research Societies), su participación activa en la organización de eventos científicos internacionales, las numerosas publicaciones en las revistas científicas de mayor prestigio, y su participación destacada en proyectos I+D+i (en colaboración con empresas e instituciones).

A continuación se describen las *líneas de investigación* de estos grupos, comenzando con aquellos temas de carácter más teórico, y siguiendo con los temas más algorítmicos y con mayor orientación hacia la modelación y las aplicaciones. Todas las líneas de investigación que se describen a continuación se caracterizan por su importante potencial de transferencia tecnológica.

- *Optimización estructurada no-diferenciable* y su interacción con la geometría. A destacar sus aplicaciones al análisis asintótico de las ecuaciones en derivadas parciales (EDP) de tipo parabólico.
- Técnicas de análisis y resolución numérica de problemas de *optimización global*. Una línea emergente de investigación se basa en explorar nuevas técnicas de descomposición en *diferencia de funciones convexas* (d.c.), y su aplicación a los algoritmos numéricos de ramificación y acotación. Existen importantes aplicaciones a la Estadística y la Ingeniería.
- *Optimización vectorial* y de multifunciones. En campos como la política, la economía, los negocios, las ciencias sociales, la ingeniería o la industria es habitual considerar múltiples aspiraciones u objetivos, a veces enfrentados, con lo que se hace necesario el estudio de técnicas de decisión basadas en un número finito de objetivos o criterios (programación multiobjetivo o decisión multicriterio), en un número no finito (programación vectorial), o incluso problemas en los que hay que optimizar una multifunción. Las aplicaciones más importantes se encuentran dentro de la propia optimización y de la economía.

- *Estabilidad y mal-condicionamiento* en optimización. El estudio de la cualificación de restricciones, de las propiedades pseudo-Lipschitz de los principales elementos del problema (el conjunto factible, el valor óptimo y el conjunto de soluciones óptimas), la caracterización de diferentes tipos de mal-condicionamiento, la obtención de estimaciones de la distancia de un problema dado al mal-condicionamiento, y las implicaciones numéricas de estas cuestiones son temas de considerable proyección de futuro.
- *Programación estocástica con variables enteras*. La programación estocástica combina las ventajas de incorporar la incertidumbre de los modelos (mediante la representación en forma de árbol representativo de escenarios) y las posibilidades de modelización de la programación matemática con variables enteras. Los modelos que resultan son de una dificultad extrema y no existen métodos eficientes de resolución. En esta línea de trabajo existen varios grupos de investigación (URJC, UCM, etc.) con aportaciones notables.
- *Modelos de optimización dinámica, estocástica y combinatoria* de sistemas productivos, logísticos y financieros. La investigación en este tema se centra en el desarrollo de nuevos métodos, formulaciones y algoritmos para la resolución de modelos de optimización dinámica, estocástica y combinatoria, motivados por problemas de planificación y control en aplicaciones que incluyen sistemas productivos, logísticos y financieros. Un grupo de la Universidad Carlos III ha desarrollado, con éxito, modelos de programación dinámica (orientados al diseño de políticas de asignación dinámica de recursos en sistemas productivos y de telecomunicación), modelos de programación estocástica (orientados a la planificación de decisiones en ingeniería financiera y en mercados eléctricos), modelos de optimización combinatoria (correspondientes a problemas de planificación de sistemas logísticos), y modelos competitivos (relacionados con problemas de imputación de costes/ beneficios en sistemas cooperativos multi-agente).
- *Problemas combinatorios* difíciles, para los que no se conocen algoritmos eficientes (polinomiales en el tamaño de los datos). La ampliación de la clase de problemas difíciles que pueden ser resueltos eficazmente es uno de los retos de las matemáticas en la actualidad. Entre los más importantes se encuentran los de agregación de preferencias, gestión de bases de datos, minería de datos, compresión de imágenes y datos, diseño o expansión de redes óptimas, planificación de producción y decisión multicriterio. Esta línea de trabajo distingue claramente dos áreas complementarias: investigación en métodos generales para resolver problemas combinatorios complejos (algorítmica, combinatoria, geometría discreta, geometría computacional, programación lineal y no lineal, optimización multiobjetivo, etc) e investigación en problemas concretos (particionamiento o coloración en grafos, tarificación en redes, expansión de líneas de transporte urbano, competición y cooperación en mercados, teoría de localización, etc). Avances en cualquiera de estas áreas aumentarían la presencia de la matemática en la sociedad y supondría extender las fronteras de aplicación de la misma.
- *Programación matemática y minería de datos*. La minería de datos es un área emergente, a medio camino entre la informática, la inteligencia artificial, y la estadística y la investigación operativa, que diseña algoritmos con los que extraer de los datos patrones comprensibles que gene-



ren conocimiento útil o interesante, y con importantes aplicaciones en genómica, medicina, telecomunicaciones, informática, finanzas, etc. Una línea dentro de la minería de datos es la utilización de métodos de optimización, fundamentalmente en el campo de la clasificación (no supervisada). Ejemplos paradigmáticos son las *máquinas de vector soporte* y los *métodos de vecino más cercano*.

- Modelización y *optimización de problemas de gran dimensión*, y su aplicación a problemas reales de nuestro entorno social. Los métodos más idóneos en la optimización de problemas de gran dimensión son los *algoritmos de punto interior* cuya complejidad es polinómica. En la actualidad algún grupo de investigación español colabora INE de Alemania, en una técnica de protección de datos.
- Diseño de *rutas óptimas* de vehículos. Comporta nuevos retos matemáticos a la vez que resuelve importantes problemas reales. En la actualidad es posible resolver óptimamente problemas de grandes dimensiones (con grafos asociados de más de veinte mil nodos), gracias no tanto a los avances informáticos (que sin duda han contribuido de forma importante) como a las investigaciones matemáticas del poliedro asociado a las soluciones factibles. La *combinatoria poliédrica* se ha revelado como una herramienta fundamental para la resolución de complejos problemas de optimización con recursos limitados.
- *Teoría de juegos*. Algunas líneas de investigación en las que algunos grupos españoles han realizado contribuciones significativas son, entre otras, el estudio de los equilibrios, de la competencia y la cooperación en situaciones en las que interaccionan varios agentes (en particular en modelos de teoría de colas, de gestión de inventarios, de secuenciación, de redes de flujo y de planificación de proyectos), la determinación de tarifas, el diseño y análisis de estructuras de votación, la conciliación y el arbitraje, el estudio de redes sociales y económicas aportando el enfoque de la teoría de juegos a conceptos clásicos en sociología, etc.

Bibliografía

- [1] Bazaraa, Shetty and Sherali, *Nonlinear Programming: Theory & Applications*. Wiley, (1994).
- [2] Bertsekas, Dimitri P., *Dynamic Programming and Optimal Control*. Belmont, MA, Athena Scientific, (1995).
- [3] Bertsekas, Dimitri P., *Nonlinear Programming*, second edition. Athena Scientific, (1999).
- [4] Dennis and Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, (1983).
- [5] Fiacco and McCormick, *Sequential Unconstrained Minimization Techniques*. SIAM Books, (1990).
- [6] Fletcher, R., *Practical Methods of Optimization*. Wiley, (1987).
- [7] Floudas and Pardalos, *Recent Advances in Global Optimization*. Princeton University Press, (1992).
- [8] Gill, Murray and Wright, *Practical Optimization*. Academic Press, (1981).
- [9] Horst and Pardalos, *Handbook of Global Optimization*. Kluwer, (1995).

- [10] Luenberger, *Introduction to Linear and Nonlinear Programming*. Addison Wesley, (1984).
- [11] Mirchandani P.B., Francis R.L. *Discrete Location Theory*. John Wiley & Sons, (1990).
- [12] Nash, S. and Sofer, A., *Linear and Nonlinear Programming*. McGraw-Hill, (1996).
- [13] Nocedal and Wright, *Numerical Optimization*. Springer Verlag, (1999).
- [14] Pinter, *Global Optimization in Action: Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications*. Kluwer, (1996).



Capítulo IX

MATEMÁTICAS Y BIOLOGÍA: ALGUNOS PUNTOS DE CONTACTO

POR JUAN JOSÉ LÓPEZ VELÁZQUEZ
UNIVERSIDAD COMPLUTENSE DE MADRID
Departamento de Matemática Aplicada
Facultad de Matemáticas
Madrid 28040
Correo electrónico: JJ_Velazquez@mat.ucm.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

The interaction between mathematics and biology has a very long history. However, in recent years there has been an increasing interest in the use of mathematical methods to provide new insights in the understanding of biological problems. In this article I describe some recent developments in this direction, including in particular the use of stochastic and kinetic equations to describe cell aggregates, ideas of the theory of homogenization and others.

Introducción

Durante la última década se han incrementado considerablemente las interacciones entre biólogos y matemáticos. El número de participantes en las reuniones científicas dedicadas a la biología teórica o a biología matemática ha crecido exponencialmente durante los últimos años. No existe una única causa que explique el desarrollo repentino de este área. Uno de los factores que ha jugado un papel esencial en este hecho fue la culminación con éxito del proyecto que pretendía obtener la secuencia del genoma humano. Durante la realización de este proyecto la comunidad bioquímica comenzó a obtener secuencias gigantescas de símbolos que describen la sucesión de ácidos nucleicos en la cadena del ADN. Es completamente natural pensar en dichas cadenas de ácidos nucleicos como fórmulas matemáticas o sucesiones de números. Obviamente no fue necesario que transcurriese mucho tiempo antes de que los matemáticos e informáticos se interesasen en dichas "fórmulas" y trataran de explorar las posibles regularidades que tienen lugar en las mismas, y que por otra parte los biólogos trataran de obtener la cooperación de matemáticos y físicos para tratar de almacenar y desentrañar la información contenida en las secuencias de ADN.

De todas formas, la interacción que está teniendo lugar en la actualidad en la interfase situada entre la biología y las matemáticas no está teniendo lugar solamente en cuestiones relacionadas con el genoma. De hecho, el creciente interés por la biología matemática de los últimos años ha dado lugar a nuevos resultados en algunas áreas más tradicionales de este campo, así como a una tendencia general a la matematización de numerosos problemas biológicos que, al menos hasta ahora, no se habían considerado desde este punto de vista.

Un área menos nueva de lo que a veces se dice...

Aunque la reciente explosión de resultados en biología matemática puede hacer pensar lo contrario, la idea de aplicar métodos matemáticos al estudio de problemas biológicos no es nueva en absoluto.

Por ejemplo, Galileo formuló algunos argumentos de lo que hoy denominaríamos "análisis dimensional" en los que mostraba que las dimensiones de los organismos biológicos no pueden crecer de forma arbitraria debido a los límites impuestos por la resistencia de los materiales de los que están formados los seres vivos.

Protocientíficos como Leonardo da Vinci o Goethe se interesaron por el problema de la filotaxis, es decir, en el estudio de la disposición geométrica de las algunas estructuras de las plantas, tales como ramas, semillas y otras semejantes. Hace siglos que se observó que en muchas plantas dichas estructuras se sitúan en ciertos puntos que están contenidos en unas curvas espirales, y que el número de dichas curvas, que puede variar de unas partes de la planta a otras, en muchas plantas es un número de la sucesión de Fibonacci $\{1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots\}$.

Estas espirales son particularmente visibles, por ejemplo en los girasoles o en las piñas. Una explicación convincente de este fenómeno se ha obtenido muy recientemente (cf. [11]). Los autores de este trabajo demostraron que la sucesión de Fibonacci aparece espontáneamente en algunos sistemas dinámicos que simulan la forma en la que se sitúan, por ejemplo, las semillas de las plantas. Más aún, dichos autores obtuvieron sistemas puramente mecánicos, sin relación ninguna con la biología en las que un conjunto de gotas fluidas se situaban en estructuras espirales análogas a las observadas en las plantas.

Una recopilación de muchas de las investigaciones realizadas hasta el primer tercio del siglo XX tratando de explicar las regularidades biológicas con ayuda de argumentos matemáticos se puede encontrar en el famoso libro [9]. Algunos de los muchos problemas que se estudian en dicho libro son, por ejemplo, el estudio de las formas de algunos organismos sencillos (y pequeños) en los que las fuerzas de tensión superficial juegan un papel dominante, la descripción de las formas espirales que aparecen en numerosos organismos tales como los moluscos, la comparación de las diferentes teorías postuladas para explicar la formación de las colmenas de las abejas. En este interesante ejemplo se forman estructuras con una simetría hexagonal global que aparece a partir de la interacción de una gran cantidad de dinámicas individuales que por sí mismas no darían lugar a dicha simetría. En [9] se estudia también el empaquetamiento óptimo de células individuales sometidas a fuerzas de presión externa, etc, la ley de Murray que describe la ley de potencias a la que deben ajustarse los radios de las arterias y capilares para obtener la máxima eficiencia en el transporte del flujo sanguíneo en los animales, etc.

El libro [9] es una interesante muestra de la gran cantidad de problemas biológicos que habían sido estudiados usando matemáticas ya en el siglo XIX y comienzos del XX. Una gran parte del libro examina el efecto que tienen las leyes de la mecánica en las formas de los organismos biológicos. No es por ello casual que el área de la biología en donde posiblemente existe un mayor grado de matematización sea la biomecánica, un tema al que me referiré con más detalle posteriormente.

Hay otras muchas áreas de biología que han tenido interacciones relevantes con la matemática en el pasado. En particular, no ha sido inusual que el estudio de problemas matemáticos motivados por la biología haya producido progresos relevantes en matemáticas. El estudio de la genética de poblaciones estuvo en el origen de muchos importantes avances en estadística obtenidos por R. A. Fisher. El estudio de la dinámica de poblaciones ha producido interesantes avances en la teoría de ecuaciones diferenciales ordinarias. A. M. Turing descubrió una famosa inestabilidad que lleva su nombre (cf. [15]) que puede tener lugar en los sistemas de ecuaciones diferenciales que describen reacciones químicas combinadas con procesos de difusión tal y como



sucede frecuentemente en sistemas biológicos. El estudio del sistema de ecuaciones formulado por Hodgkin-Huxley para describir la propagación de señales químicas y eléctricas en el axón del calamar gigante ha dado lugar al desarrollo de diferentes técnicas para el estudio de ecuaciones en derivadas parciales,...

Aunque posiblemente existen pocas actividades en las que el riesgo de equivocarse sea mayor que en la de hacer predicciones sobre el futuro desarrollo de las ciencias, mencionaré a continuación algunos temas en los que los desarrollos actuales podrían dar lugar a progresos interesantes en el futuro, tanto desde el punto de vista de la creación de nuevas matemáticas como del desarrollo de nuevas técnicas que podrían mejorar la comprensión de los sistemas biológicos. En un área con la extensión que tiene en la actualidad la biología matemática será inevitable la ausencia de temas importantes. La elección de temas será necesariamente arbitraria y estará basada en gran parte en mis preferencias y experiencia personal.

Estocasticidad: El efecto de los números pequeños

Numerosos procesos biológicos tienen lugar en regiones de dimensiones tan pequeñas que los efectos de las moléculas individuales juegan un papel esencial.

Consideremos por ejemplo el caso de *E. coli*. Esta bacteria es uno de los "organismos modelo" que han sido estudiados por los biólogos con gran detalle. En particular, un proceso biológico para el que se puede obtener información a partir del estudio de este organismo es el de la movilidad celular.

La bacteria *E. coli* se mueve mediante la propulsión que le suministra un conjunto de flagelos (entre 4 y 10) situados en la parte posterior de la célula. Entender la forma detallada en la que el movimiento flagelar produce la fuerza de propulsión que mueve a la bacteria es una cuestión interesante de mecánica de fluidos que ha recibido una considerable atención en el pasado y que fue resuelta hace varias décadas. Un aspecto esencial de este problema es que, debido a la pequeñez de la bacteria, su movimiento tiene lugar a números de Reynolds muy pequeños, o dicho en términos menos técnicos, las fuerzas de viscosidad del fluido son mucho más importantes que la inercia del movimiento. Resultados clásicos de mecánica de fluidos, que pueden encontrarse en los libros de texto, establecen que la fuerza de fricción, cuando el número de Reynolds es muy pequeño, es proporcional a la velocidad, y es mayor para un cilindro que se mueve perpendicularmente a su eje que para un cilindro que se mueve paralelamente a su eje. Cuando los flagelos producen el desplazamiento de la bacteria tienen una forma aproximadamente helicoidal en la parte posterior de la bacteria, y realizan un movimiento de rotación en torno al eje de dicha hélice. Es fácil ver que en ese movimiento los flagelos no se mueven con relación al fluido ni paralela ni perpendicularmente a su eje, sino con una combinación de ambos movimientos. La diferencia de fuerzas de fricción para estos tipos de movimiento que fue mencionada anteriormente tiene como consecuencia que la fuerza neta experimentada por el flagelo no sea opuesta al movimiento de rota-

ción del flagelo, sino que tiene una componente adicional en la dirección hacia la que está orientada la célula. Dicha componente adicional es la que proporciona la propulsión celular. Una descripción más detallada de este mecanismo se encuentra en [5], [16].

El efecto de la estocasticidad que se mencionaba en el título de esta sección no aparece, sin embargo, en el estudio de la propulsión mediante flagelos, que es un problema que puede estudiarse en gran parte mediante las ecuaciones clásicas de la mecánica de fluidos, sino en otra parte del problema, concretamente en la detección de las señales que indican a la célula la dirección hacia la que debe moverse. Los flagelos tienen dos formas distintas de moverse. Puede ocurrir que todos los flagelos roten de forma coordinada, o por el contrario puede suceder que los flagelos se muevan de forma independiente y descoordinada. En el primer caso, la célula se mueve en forma aproximadamente rectilínea y se suele decir que la bacteria "corre". En el segundo caso la bacteria se mueve de forma esencialmente aleatoria, sin ninguna orientación definida y se suele decir que "da tumbos". El que la bacteria se mueva de una u otra forma depende del grado de ocupación de unos receptores químicos situados en la superficie de la célula. Dichos receptores miden la concentración externa a la célula de ciertas sustancias químicas que la bacteria necesita para sobrevivir. La más estudiada de estas sustancias se denomina aspartato. Es precisamente en este proceso de medición donde aparecen los efectos de estocasticidad. En efecto, las bacterias de *E. coli* son muy pequeñas (su longitud es de 2×10^{-4} cm) y por ello el número de receptores no es demasiado grande, a lo sumo de unos pocos centenares. La bacteria tiene tendencia a moverse hacia las regiones con mayor concentración de aspartato y para ello la estrategia que sigue es la de incrementar el número de "carreras" frente a "tumbos" cuando percibe que los receptores incrementan su grado de ocupación. Sin embargo, debido a la pequeñez del número de receptores, el grado de ocupación de los receptores no es una variable determinista, sino una variable aleatoria, en la que las fluctuaciones en torno al valor medio no son despreciables. Esta estocasticidad se manifiesta en el hecho de que muchas de las variables que harían falta para describir la dinámica de las bacterias de *E. coli*, como por ejemplo la distancia recorrida en cada "carrera" o la duración de las mismas, son variables aleatorias.

Esta estocasticidad, inherente a la dinámica de numerosos organismos, es bastante habitual en los procesos biológicos y se debe a las pequeñas dimensiones de estos sistemas. El mundo en que viven muchos de estos organismos no es un mundo determinista, newtoniano, que pueda describirse mediante ecuaciones diferenciales. Por el contrario, las bacterias y muchas otras estructuras biológicas han aprendido a usar la aleatoriedad para sobrevivir. Desde el punto de vista del matemático esto plantea interesantes cuestiones y retos a la hora de obtener modelos y describir la dinámica de estos organismos.

Homogenización: o de cómo las células calculan límites

Son innumerables las estructuras biológicas que están formadas por un elemento que se repite un gran número de veces, pero en donde la estructura actúa de forma colectiva mediante la acción coordinada de las diferentes piezas individuales. Piénsese por ejemplo en los estomas de



las hojas de las plantas, los conos y bastones que actúan como receptores de la luz en la retina, la estructura celular de los tejidos, los cilios que propulsan a muchas células,...

Durante los últimos cincuenta años los matemáticos han desarrollado un conjunto de técnicas para estudiar el comportamiento colectivo de sistemas formados por múltiples piezas individuales que se conocen con el nombre de métodos de homogenización. Dichos métodos se han aplicado con éxito en el estudio de las propiedades de medios materiales, y recientemente han comenzado a ser usados en el estudio de alguno de los sistemas biológicos mencionados arriba (cf. [3]).

Un ejemplo de sistema biológico que puede estudiarse empleando las ideas de la teoría de homogenización, que es particularmente atractivo por varias razones, es el estudio de la forma en la que los receptores de una célula capturan una sustancia química. Es bastante frecuente que una célula que precisa obtener un tipo específico de molécula que se encuentra distribuida en su entorno disponga de un conjunto de receptores, que típicamente son otra molécula o un grupo molecular, y que tienen una gran afinidad química por la molécula que se pretende capturar.

¿Cuál es la disposición y el número óptimo de receptores que permiten atrapar la sustancia química externa de la forma más eficaz posible? La respuesta a este problema tal y como está planteado sería la de recubrir toda la superficie celular de receptores, pero en ese caso quedaría poco espacio para otras estructuras. ¿Es posible conseguir un grado de eficiencia razonable en la captura de sustancias externas sin recubrir toda la superficie celular? La solución de este problema puede encontrarse en [5] en donde se estudia esta cuestión suponiendo que la célula es esférica. La geometría no esférica puede estudiarse de forma similar, incrementando algo las dificultades técnicas, pero no el resultado esencial. El análisis de [5] llega a una conclusión interesante. Supongamos que el radio de la célula es R y que el tamaño de cada receptor es a . Típicamente a es mucho menor que R . Supongamos también que se tienen N receptores en la superficie de la célula. En principio podría pensarse que para conseguir que la captura de la sustancia externa ocurra con la misma rapidez con la que sucedería si toda la superficie externa estuviese llena de receptores haría falta que una fracción importante del área de la célula estuviese llena de receptores, es decir que se tuviese $N (a/R)^2 \approx 1$. Lo sorprendente es que dicha tasa óptima de captura puede obtenerse con un número mucho menor de receptores, concretamente

$$(15) \quad N (a/R) \approx 1.$$

La razón por la que sucede esto es porque, antes de ser capturadas por los receptores, las moléculas de la sustancia externa se mueven de forma browniana por el exterior de la célula, lo que es un método muy eficaz de explorar la superficie celular.

Algo que es particularmente llamativo en este ejemplo es el hecho de que el límite (15) es precisamente el que da lugar a problemas matemáticos más interesantes cuando el problema se formula en términos de ecuaciones diferenciales. Por interesante se entiende aquí que en ese régimen los receptores comienzan a actuar de forma colectiva y no puede despreciarse la interacción entre ellos. Esta capacidad de los sistemas biológicos para "identificar" límites matemáticos relevantes

ocurre con alguna frecuencia. Otro ejemplo en el que también sucede esto es en el problema del movimiento de la bacteria *E. coli* que se mencionó en el apartado anterior. En ese problema el tiempo medio de duración de las "carreras" es exactamente del orden de magnitud necesario para que los efectos del movimiento browniano de la bacteria comiencen a ser importantes. Si dicho tiempo hubiese sido mayor las bacterias tendrían un movimiento puramente difusivo y serían incapaces de dirigirse hacia las regiones con mayor concentración de aspartato. Un tiempo menor tendría como consecuencia que la velocidad media hacia las regiones con altas concentraciones de aspartato sería menor. Presumiblemente la actuación de la evolución darwiniana tratando de lograr una mayor eficiencia ha sido la responsable de estos delicados ajustes que posiblemente van a plantear interesantes problemas a los matemáticos en los próximos años.

Otro tipo de problemas biológicos en los que los métodos de la teoría de homogenización y otros relacionados podrían jugar un papel interesante es en el estudio de las ecuaciones efectivas de los medios aleatorios. Un ejemplo reciente en el que se esbozan nuevas herramientas matemáticas para estudiar problemas biológicos se puede ver en [6]. Numerosas estructuras biológicas están formadas por células que tienen un importante grado de variabilidad y que han de ser descritas mediante variables estocásticas. La teoría de los procesos difusivos en medios aleatorios ha sido desarrollada en gran parte por matemáticos puros interesados en la teoría de la probabilidad, y por físicos interesados en cuestiones tales como difusión de neutrones o problemas análogos. Sin embargo, con bastante frecuencia los problemas que se han abordado en dichas áreas no son exactamente las cuestiones que surgen en los problemas biológicos. En otros casos, en particular en el caso de los resultados obtenidos por los matemáticos más teóricos, el lenguaje en el que han sido formulados los resultados plantea serias dificultades a los científicos que se dedican a la biología.

Hacia una teoría cinética de los sistemas biológicos

Un tema en el que en los últimos años ha habido una gran actividad por parte de varios grupos de matemáticos y físicos ha sido en el estudio de ecuaciones cinéticas para describir agregados celulares. Matemáticamente dichas ecuaciones son bastante similares a las que se emplean en la dinámica de gases enrarecidos. La idea que subyace bajo estos modelos es la de considerar a las células como si fuesen las "moléculas" de un gas. En los modelos dichas células cambian su dirección de movimiento o su velocidad o bien de forma aleatoria, o bien debido a la interacción con otras células. La idea de describir de esta forma los agregados celulares no es nueva (cf. [1], [14]), aunque en los últimos años se han realizado progresos en el estudio matemático de alguno de estos modelos. Una de las aplicaciones más habituales de estos estudios es la de obtener "coeficientes de difusión" para las células de la misma forma en que las ecuaciones cinéticas de la dinámica de gases permiten obtener valores de los coeficientes de difusión, viscosidad y otros análogos (cf. [10]).

En la mayor parte de los modelos cinéticos de células que se han estudiado hasta ahora se suele suponer que las células no interaccionan con su entorno, excepto por la interacción indirecta que las células puedan tener con una sustancia química que ellas mismas producen. Uno de los



pocos modelos cinéticos en donde se tienen en cuenta las interacciones entre células es el estudiado en [12]. La ecuación cinética considerada en dicho artículo describe la evolución de un agregado celular bidimensional formado por células que se caracterizan mediante dos variables individuales, que son su posición y su orientación. Cuando dos células del agregado se encuentran, se reorientan y tratan de alinearse en la misma dirección. Desde el punto de vista matemático estas ecuaciones tienen considerables analogías con la ecuación de Boltzmann que se emplea en el estudio de la dinámica de gases, aunque existen también algunas diferencias. En efecto, en la ecuación de Boltzmann dos partículas que se encuentran no tratan de alinear su velocidad, sino que colisionan elásticamente. En concreto las soluciones de la ecuación de Boltzmann satisfacen principios de conservación de la energía y de la cantidad de movimiento, que no se cumplen para las ecuaciones de [12]. Como consecuencia, el tipo de dinámicas que se obtienen para estas últimas ecuaciones es bastante distinto del que se obtiene en la dinámica de gases. En el caso de la ecuación considerada en [12] existen soluciones en las que las células se alinean en dos direcciones opuestas, un tipo de comportamiento que se ha observado en unos organismos llamados myxobacterias, que se han estudiado de forma bastante detallada con el fin de obtener información sobre los mecanismos de movilidad celular (cf. [13]).

La principal objeción que se puede plantear a los modelos cinéticos de agregados celulares es que, en muchos casos, las densidades celulares son demasiado altas para que estas ecuaciones puedan proporcionar una descripción realista de dichos agregados. En el caso de la dinámica de gases se conocen desde hace más de cincuenta años las hipótesis que deben satisfacer las interacciones moleculares y las densidades de partículas para que dicha dinámica se pueda describir empleando ecuaciones cinéticas. Por ejemplo, la ecuación de Boltzmann es válida si el camino libre medio entre colisiones es mucho mayor que la distancia de interacción mutua entre moléculas. La tarea de precisar matemáticamente las hipótesis que garanticen la validez de los límites cinéticos para agregados celulares no ha sido realizada aún. Uno de los factores que dificultan este problema es el hecho de que muchas de estas interacciones tienen lugar mediante la emisión de sustancias químicas que pueden tener sus propias longitudes y tiempos característicos asociados tales como el tiempo medio de degradación, la longitud característica de difusión y otros. Posiblemente, uno de los pocos ejemplos experimentales en donde la distancia entre las bacterias es lo suficientemente grande como para que puedan usarse modelos cinéticos que den información cuantitativa sobre las densidades de agregados celulares es el considerado en [8].

Movilidad celular y biomecánica

Muchos de los problemas mencionados hasta ahora están relacionados con la movilidad celular. El movimiento de las células juega un papel importante en una gran cantidad de procesos biológicos tales como embriogénesis, inflamación, curación de heridas,... Por otra parte el estudio de los sistemas que emplean las células para desplazarse plantea una gran cantidad de cuestiones físicas y matemáticas de interés.

Se han descrito anteriormente algunos de los problemas de mecánica de fluidos que plantea el desplazamiento empleando flagelos de algunos organismos tales como *E. coli*. Tal y como se indicó allí, el origen de la fuerza producida por los flagelos se entiende bien. Algo que sin embargo no se comprende de forma tan satisfactoria es la forma en la que los flagelos se reorganizan con el fin de rotar conjuntamente en forma coordinada. Se sabe que dicha reorganización sucede como respuesta a la saturación de los receptores que miden la concentración de aspartato en la superficie celular, pero no se conoce el mecanismo detallado, que posiblemente involucra, además de a las señales químicas, las propiedades mecánicas del flagelo y del medio fluido que lo rodea.

Otro tipo de propagación celular, diferente de la producida por los flagelos, es la propagación ameboide, debida a la expansión y retracción de pseudópodos que tiene lugar en algunas células eucariotas como por ejemplo en el "organismo modelo" *Dictyostelium discoideum* (Dd). Las fuerzas que dan lugar al movimiento de este tipo de organismos se deben principalmente a la dinámica de una estructura interna a la célula denominada citoesqueleto que además de estas funciones motrices desempeña otras muchas funciones importantes para la vida celular, tales como transportar sustancias químicas de unas partes de la célula a otras. El citoesqueleto es una estructura muy dinámica y cuya estructura puede variar dependiendo del grado de ocupación de algunos receptores externos a la célula que permiten medir, por ejemplo, gradientes de sustancias químicas. Otro factor que también juega un papel importante en el movimiento de este tipo de células son las moléculas que adhieren el citoesqueleto con el substrato externo en el que se apoya la célula. Uno de los intentos de obtener un modelo mecánico para este tipo de sistemas se puede encontrar en [2].

En todos los ejemplos mencionados anteriormente se han descrito ejemplos de movimiento celular de células individuales. Sin embargo, un problema que tiene particular relevancia es el del estudio de sistemas biológicos en los que muchas células dan lugar a movimientos colectivos, tal y como sucede por ejemplo en los procesos de embriogénesis. Desde el punto de vista matemático el problema sería el de obtener "ecuaciones efectivas" para el movimiento de agregados celulares que individualmente se mueven con los mecanismos indicados anteriormente. En este punto es relevante mencionar, aunque sea de forma fugaz, el área de la biomecánica. Este área, que está considerablemente desarrollada, ha abordado una gran cantidad de problemas que van desde el estudio de los flujos sanguíneos en el sistema circulatorio hasta las ecuaciones constitutivas de los tejidos pasando por los mecanismos de vuelo empleados por los insectos o las aves. En la búsqueda de ecuaciones constitutivas para agregados celulares es muy posible que los matemáticos familiarizados en el uso de técnicas de homogenización pudiesen realizar alguna contribución de interés.

Patrones en biología: "Arte celular" vs. funcionalidad

Uno de los temas clásicos de biología matemática ha sido el estudio de los diversos patrones que pueden aparecer en los sistemas biológicos. En biología abundan patrones geométricos tales como las espirales, patrones en forma de diana, agregados de puntos, franjas, cuadrados y otros muchos. Algunos de



los problemas más conocidos que han sido estudiados por muchos de los fundadores de la moderna biología matemática son, entre otros, el estudio de las manchas de los leopardos o las cebras, el estudio de la pigmentación de los peces, el estudio de los patrones que aparecen en la superficie de las conchas marinas o los patrones de ondas espirales que forman las señales químicas emitidas por el Dd.

Todas estas investigaciones han incrementado considerablemente la perspectiva de los matemáticos que han aprendido así la inmensa cantidad de comportamientos distintos a los que pueden dar lugar aparentemente sencillas ecuaciones diferenciales en derivadas parciales.

Sin embargo, en los últimos tiempos se ha ido abriendo paso la idea de que independientemente de su valor estético lo que realmente resulta relevante de los variados patrones que aparecen en sistemas biológicos es el determinar su funcionalidad, es decir, determinar si realmente los organismos los crean porque dichos patrones son más eficaces para obtener algún beneficio desde el punto de vista biológico, o si por el contrario son simplemente un "resto de la evolución". Una segunda razón por la que resulta interesante el estudio de los patrones es porque en ocasiones proporcionan información sobre los mecanismos subyacentes de interacción celular. En muchas ocasiones los organismos muestran sorprendentes patrones geométricos pero que sólo se manifiestan en condiciones muy especiales en el laboratorio. En dichos casos los patrones no tienen ningún valor biológico para el organismo pero permiten inferir, por ejemplo, que las células individuales emiten alguna sustancia que les permite dar lugar a un comportamiento colectivo. Un ejemplo interesante de esto se encuentra en [7], donde los autores obtienen tal tipo de conclusión a partir de los patrones que muestran ciertos agregados de *E. coli*.

Algunas reflexiones finales

En las páginas anteriores he mencionado algunos problemas que han dado, y en otros casos podrían dar lugar a interacciones fructíferas entre biología y matemáticas. Hay otras muchas áreas de contacto entre estas dos ciencias que no he mencionado ya que las anteriores reflexiones se han orientado hacia los temas que conozco mejor o me resultan más próximos. No he hablado por ejemplo de tomografía, que es un área muy desarrollada matemáticamente. Tampoco he hablado de la búsqueda de regularidades en el genoma, una tarea a la que hoy se están dedicando abundantes recursos en muchas partes. En estos problemas, así como en muchos de los mencionados en las secciones anteriores, se están realizando numerosas simulaciones numéricas con el fin de obtener información. Por otra parte es también bastante evidente que uno de los motores que están fomentando la actividad en muchas de estas cuestiones son las posibles aplicaciones biotecnológicas que pudiesen surgir de estas investigaciones.

Bastantes de las investigaciones que se podrían realizar en la interfase entre biología y matemáticas, y esto incluye muchas de las cuestiones mencionadas arriba, son posiblemente parte de lo que podría denominarse "ciencia normal", entendiendo por esto la combinación acertada de los métodos usuales de trabajo de matemáticos, físicos y biólogos. Los métodos experimentales dis-

ponibles en biología junto con las herramientas de modelización y estudio de ecuaciones, así como los métodos numéricos y analíticos de los que disponen los físicos y matemáticos, podrían proporcionar una mejor comprensión de muchos fenómenos biológicos.

Sin embargo, es difícil que el matemático que interaccione con los biólogos no perciba que éstos están comenzando a esbozar unos conceptos que hoy por hoy no tienen una definición matemática precisa, pero que podrían suponer cambios tan radicales en la forma de entender muchos problemas de matemáticas como los que ha supuesto la física en el pasado. A modo de ejemplo se puede señalar la cuestión de entender las "redes bioquímicas", es decir el conjunto de conexiones que existen mediante reacciones químicas de las diferentes sustancias que toman parte en un proceso biológico. La primera vez que un matemático observa uno de estos diagramas la impresión suele ser de desasosiego, debido a la complejidad de los mismos. Por otra parte, los biólogos hace tiempo que llevan hablando de algunos conceptos tales como "robustez" con lo que quieren indicar que estas "redes bioquímicas" suelen ser capaces de seguir funcionando eficazmente aunque se modifique alguna de las sustancias involucradas. Más aún, dicha robustez le confiere al sistema la capacidad de evolucionar hacia configuraciones distintas, es decir, tiene dos de las características deseables para un sistema biológico: robustez y flexibilidad.

Creo que los matemáticos son conscientes de que hoy no disponen de herramientas para describir de forma precisa el tipo de conceptos sugeridos arriba. En términos más matemáticos, hoy no existen métodos para describir la dinámica de las ecuaciones diferenciales que describen esas reacciones químicas. Es sabido que sistemas de ecuaciones con pocas variables (pero más de dos) son capaces de generar dinámicas caóticas, y en sistemas con muchas variables que interaccionan débilmente se pueden emplear los métodos de la física estadística que permiten describir el sistema mediante un conjunto reducido de variables, pero ¿cómo estudiar sistemas con muchas variables fuertemente interconectadas? El intento por tratar de comprender algo mejor este problema es uno de los que ha dado lugar recientemente a un aumento de interés en la teoría de grafos y ha dado lugar a hallazgos (cf. [4]) que comienzan a esbozar lo que podría ser esa futura "biomatemática" que acabarán percibiendo en su totalidad las generaciones futuras.

Bibliografía

- [1] Alt, W.; *Biased random walks models for chemotaxis and related diffusion approximations*, J. Math. Biol. 9, (1980), (147-177).
- [2] Alt, W.; Dembo, M.; *Cytoplasm dynamics and cell motion: Two phase flow models*. Mathematical Biosciences 156, (1999), (207-228).
- [3] Andreucci, D.; Bisegna, P.; Caruso, G.; Hamm, H.E.; DiBenedetto, E.; *Mathematical models of the spatio-temporal dynamics of second messengers in visual transduction*. Biophysical J. 85, (2003), (1358-1376).
- [4] Barabási, A.L.; *Linked: The New Science of Networks*. Perseus, Cambridge, MA, (2002).
- [5] Berg, H.C.; *Random walks in biology*, Princeton Univ. Press, (1993).



- [6] Bollenbach, T.; Kruse, K.; Pantazis, P.; González-Gaitán, M.; Jülicher, F.; *Robust formation of morphogen gradients*, Phys. Rev. Lett. 94, (2005).
- [7] Budrene, E.O.; Berg, H.C.; *Dynamics of formation of symmetrical patterns by chemotactic bacteria*, Nature 376, (1995), (49-53).
- [8] Mittal, N.; Budrene, E.O.; Brenner, M.P.; van Oudenaarden, A.; *Motility of Escherichia coli cells in clusters formed by chemotactic aggregation*, Proc. Nat. Acad. Sci. 100, 23, (2003), (13259-13263).
- [9] Thompson, D.; *On Growth and Form*. Cambridge University Press, (1942).
- [10] Dickinson, R.B.; Tranquillo, R.T; *Transport equations and indices for random and biased cell migration based on single cell properties*, SIAM J. Appl. Math. 55, 5, (1995), (1419-1454).
- [11] Douady, S.; Couder, Y.; *Phyllotaxis as a physical self-organized growth process*. Phys. Rev. Lett. 68, (1992), (2098-2101).
- [12] Geigant, E.; Stoll, M.; *Bifurcation analysis of an orientational aggregation model*, J. Math. Biol. 46 (6), (2003), (537-563).
- [13] Mogilner, A.; Edelstein-Keshet, L.; *Spatio-angular order in populations of self-aligning objects: formation of oriented patches*, Phys. D 89, 3-4, (1996), (346-367).
- [14] Patlak, C.S.; *Random walk with persistence and external bias*, Bull. Math. Biophysics 15, (1957), (311-338).
- [15] Turing, A.M.; *The chemical basis of morphogenesis*. Phil. Trans. Roy. Soc. B 237, (1952), (37-72).
- [16] Vogel, S.; *Life in Moving Fluids: The Physical Biology of Flow*, Princeton Univ. Press, (1996).



Capítulo X

MECÁNICA DE FLUIDOS
COMPUTACIONAL Y APLICACIONES

POR ANA MARÍA MANCHO SÁNCHEZ
CONSEJO SUPERIOR
DE INVESTIGACIONES CIENTÍFICAS (CSIC)
Departamento de Matemáticas,
Instituto de Matemáticas y Física Fundamental
Serrano 121. Madrid CP 28006
Correo electrónico: a.m.mancho@mat.csic.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Fluid Mechanics is an old discipline that has fascinated for centuries to mathematicians, physicists and engineers. The sinusoidal motion of a fluid changing abruptly in space and time suggests an extremely complex mathematical problem. A remarkable step forwards in the history of fluid mechanics was given by Euler who in 1775 wrote for the first time the equations of motion of an inviscid flow. Some years later, in 1825 Navier and Stokes introduced separately the viscous term in what today we call Navier-Stokes equations, which are extremely difficult to solve. Modern computers are an important tool to deal with hard problems expressed in terms of fluid mechanics equations and therefore a new discipline called Computational Fluid Dynamics (CFD) has been developed. A recent study outlined by an expert committee of the Division of Fluid Mechanics of the American Physical Society describes important applications of fluids and their influence in our daily life [3]. To explain the impact of the mathematical and computational research on some applied examples of fluid mechanics is the purpose of this summary.

Wafers made from hyperpure silicon form the basis for the production of semiconductor elements and are therefore indispensable for practically all electronic components. The silicon wafers are characterized by their diameter, their crystal orientation, the doping additive and their surface finish. Mathematically the core problem is described by the incompressible Navier-Stokes equations coupled with an advection-diffusion equations for heat and concentrations. Efficient and reliable CFD methods are the crucial technologies for a thorough understanding, prediction and optimization of the production of new and state of the art semiconductor single crystals. For instance instabilities may be predicted with accurate simulations [6] and they need to be avoided to have dopants homogeneously distributed.

Microfluidics is a key component of established and developing technologies ranging from lab-on-a-chip biotech devices to inkjet printing. The lack of mixing is often a key obstacle to the effective functioning of microfluidic devices as viscous effects dominate at small scales, and inducing turbulence to improve mixing is impractical. Dynamical systems theory provides a suitable paradigm for deterministic mixing. A powerful idea emerging from pure mathematics is the so-called linked twisted map (LTM) which provides conditions for good mixing properties. Fluid particle motion in micromixers can be described in terms of LTM (or its generalizations) and this show that microfluidic applications can benefit by a closer linkage and use of basic theory.

Introducción

La Mecánica de Fluidos es una disciplina que durante siglos ha fascinado a Matemáticos, Físicos e Ingenieros. La razón quizás sea la complejidad de su dinámica que no deja de asombrarnos: el movimiento del agua del mar, o de una torrentera, el humo de un cigarrillo, etc. Las formas sinuosas que cambian bruscamente y se retuercen sobre sí mismas dejan entrever un problema matemático muy complejo. En la historia de los fluidos es destacable el avance de L. Euler quien

en 1755 formalizó su descripción escribiendo por primera vez las ecuaciones diferenciales que rigen el movimiento de un fluido no viscoso. Años después, en 1825, C. Navier y G. Stokes introdujeron el término de viscosidad en las ecuaciones que hoy denominamos de Navier-Stokes. Estas ecuaciones describen y cuantifican el comportamiento de los fluidos, pero en multitud de ocasiones su tratamiento exacto no es factible lo que requiere abordar los problemas desde la perspectiva computacional. La potencia de cálculo de los ordenadores modernos ha abierto enormes posibilidades en esta dirección, y se ha constituido en una rama de la Mecánica de Fluidos llamada Mecánica de Fluidos Computacional (CFD) que tiene muchas aplicaciones. Típicamente éstas se dan en el ámbito del procesado de materiales en industria, o en procesos geofísicos, astrofísicos o biológicos. Tampoco faltan aplicaciones que explotan las posibilidades visuales de los fluidos desarrollando software para la creación de efectos especiales en la industria del cine o del entretenimiento [1]. Un vistazo al índice de sesiones de la reunión anual de la División de Dinámica de Fluidos [2] de la American Physical Society (DFD-APS) es suficiente para percibir la diversidad de problemas en que los fluidos intervienen: biofluidos, hemodinámica, microfluidos, gotas, ruptura y coalescencia en fluidos, procesado de materiales, transporte en nanotubos y nanocanales, nanotecnologías, capas delgadas, espumas, fluidos no Newtonianos y disoluciones poliméricas, ondas, fenómenos de interfases, flujos reactivos, acústica, turbulencia, turbulencia escalar, procesos de mezcla, suspensiones, fluidos multifásicos y estratificados, aerodinámica, flujos en medios porosos, convección térmica, inestabilidades, geofluidos, astrofluidos, dinámica atmosférica, dinámica oceánica, etc. Un reciente estudio elaborado por un comité de expertos de esta división (DFD-APS) detalla interesantes aplicaciones de los fluidos y su influencia en nuestra vida cotidiana [3].

Es imposible abordar en esta presentación todos los tópicos mencionados, por esta razón en los próximos apartados detallamos cuál es el impacto de la investigación matemática y computacional en una selección de ellos.

Chips, fluidos e inestabilidades

Los discos u obleas de silicio muy puro son la base de la producción de los semiconductores o chips, indispensables para la fabricación de prácticamente todos los componentes electrónicos. Son el núcleo de la moderna micro y nanoelectrónica presente en ordenadores, teléfonos móviles, reproductores de DVD, pantallas ultraplanas, sistemas de navegación, control de aeronaves, etc. La industria de las obleas de silicio representaba en 2004 un volumen de negocio de casi 6000 millones de dólares con tendencia creciente. Los elementos que caracterizan las obleas de silicio son: su diámetro (actualmente los más grandes alcanzan los 300mm), la estructura cristalina, la concentración de dopantes (que le proporcionan sus propiedades semiconductoras) y su acabado (pulidos, no pulidos o recubiertos). El control de cada uno de estos elementos en la fase de producción industrial requiere gran precisión tecnológica. Típicamente el crecimiento de un lingote de silicio se realiza con el método de Czochralski (CZ). El silicio virgen policristalino se funde y se le añaden dopantes que le proporcionan las propiedades eléctricas deseadas. El material pro-



cesado, al fundirse entra en convección por efectos de tensión superficial (efecto Marangoni) y por empuje de Arquímedes (efecto Rayleigh-Bénard). Un conocimiento profundo de cómo es la dinámica del fluido en estas condiciones es necesario ya que la convección puede originar una distribución no uniforme del material dopante, de forma que al cristalizar se producen defectos en el cristal [4]. Por esta razón se han publicado muchos trabajos en esta línea que intentan clarificar aspectos fundamentales. Artículos de revisión sobre las inestabilidades en flujos termocapilares pueden encontrarse en [10, 11, 12]. Métodos eficientes y fiables de la Mecánica de Fluidos Computacional son cruciales para entender, predecir y optimizar la tecnología de producción de semiconductores monocristalinos.

Matemáticamente este problema se describe por las ecuaciones de Navier-Stokes:

$$(16) \quad \partial_t u + (u \cdot \nabla) u = Pr (-\nabla p + \nabla^2 u + R \Theta e_z),$$

la condición de incompresibilidad

$$(17) \quad \nabla \cdot u = 0,$$

y la ecuación del calor,

$$(18) \quad \partial_t \Theta + u \cdot \nabla \Theta = \nabla^2 \Theta.$$

Las condiciones de contorno, que dependen del recipiente de fundido, consideran los efectos de tensión superficial (efecto Marangoni) que acoplan el campo de velocidades y de la temperatura, así como otros parámetros externos de tipo térmico. El sistema (16)-(18) junto con sus condiciones de contorno específicas, presenta soluciones estacionarias con alta simetría (azimutal o traslacional según sea la geometría del recipiente). Sin embargo estas soluciones para ciertos valores de los parámetros externos pueden bifurcar hacia otras más complejas bien estacionarias u oscilatorias [5, 6, 7] que rompen la simetría y que en el proceso de producción es conveniente evitar pues introducen inhomogeneidades en el cristal. El análisis de estabilidad lineal determina las condiciones en que las bifurcaciones o inestabilidades se producen. La teoría de bifurcaciones, aquí esbozada muy sucintamente para un problema concreto de las ecuaciones de los fluidos, es en realidad un problema básico en análisis funcional con muchas aplicaciones en ondas de agua, en el combamiento de vigas, etc. [8, 9].

Las inestabilidades en fluidos que se han descrito son protagonistas no sólo de procesos de manufactura de materiales en la industria sino también de multitud de fenómenos en el ámbito de la astrofísica y geofísica: el manto terrestre y su dinámica se describen desde la perspectiva de un fluido muy viscoso que en convección causa el movimiento de las placas continentales. El núcleo de la Tierra es un fluido cargado que en movimiento origina el campo magnético terrestre. La simulación numérica directa (DNS) de estos procesos ayuda indudablemente al avance de estas ciencias.

Transporte en fluidos: de los océanos a los “lab-on-a-chip”

Un fluido en movimiento se caracteriza por su campo de velocidades que en ocasiones puede llegar a ser muy complejo. Una pregunta de indudable interés es ¿cómo son arrastradas las partículas en suspensión por un fluido en movimiento? Quizás hayamos formulado esta pregunta de una manera abstracta, sin embargo es posible reconocerla en otras muchas que nos resultan más cercanas e interesantes. En el ámbito de la oceanografía, de la zoología marina, de la tecnología este problema -que tiene un fuerte trasfondo matemático- cuenta con muchas aplicaciones. Por ejemplo los oceanógrafos miden, caracterizan y describen las corrientes marinas y oceánicas. Los nutrientes (fitoplancton, zooplancton) son arrastrados por el mar en movimiento. ¿Podrían atravesar un flujo de corriente dado y proporcionar alimento a las especies marinas al otro lado de la barrera? ¿Cómo se dispersan en el océano los contaminantes vertidos por un petrolero hundido? Las plantas desaladoras de agua marina situadas en la costa tienen un indudable impacto ecológico porque devuelven al mar grandes cantidades de agua con concentraciones salinas muy superiores a las que se ha adaptado el habitat marino del entorno. ¿Cómo se dispersa el agua que la planta desaladora vierte en el litoral? ¿Pueden las corrientes marinas mantener confinados esos vertidos? [13].

En el proceso de mezclado y transporte de partículas en suspensión en un fluido intervienen los procesos difusivos, pero si éstos no son importantes aquellos se describen simplemente con las ecuaciones de advección:

$$(19) \quad \frac{dx}{dt} = v(x, t), \quad x \in \mathbb{R}^n$$

donde $x(t)$ representa la trayectoria de una partícula y $v(x, t)$ representa el campo de velocidades del fluido. La ecuación (19) es un sistema dinámico no autónomo cuando $v(x, t)$ depende explícitamente del tiempo t . Un paso básico en el análisis de los sistemas dinámicos es identificar en el espacio de fases regiones de dinámica cualitativamente diferente y las fronteras que las separan. Esto, desde el punto de vista matemático, es todavía hoy un reto para los sistemas no autónomos aperiódicos. Numerosos trabajos [19, 20] introducen conceptos tales como “estructuras lagrangianas coherentes” (en inglés *Lagrangian coherent structures* LCS) o “variedades invariantes” cuyo cálculo numérico y aplicabilidad se discute en [21, 22]. Estos conceptos son de gran ayuda para el análisis de problemas reales [13] y por ello un avance en la teoría básica supondrá sin duda un progreso en las aplicaciones mencionadas.

El transporte de partículas en fluidos también es importante para la microfluídica. Actualmente los dispositivos de microfluidos son parte de una ciencia en ebullición [14] que registra una enorme actividad con miles de artículos publicados y cientos de patentes registradas [16, 17]. Los dispositivos de microfluidos son aparatos minúsculos (por debajo de los milímetros) que contienen fluido cuya dinámica se describe en la aproximación continua de las ecuaciones en derivadas parciales y esto a pesar de tratarse de escalas muy pequeñas. Muchos aspectos de los microfluidos son todavía una incógnita [15]. A grandes rasgos puede decirse que su número de



Reynolds es muy pequeño, los efectos de tensión superficial son muy importantes y la viscosidad es una propiedad dominante. El interés en estos miniaparatos ha crecido ya que han encontrado muchas aplicaciones en campos tan diversos como el enfriamiento de dispositivos electrónicos, las impresoras de chorro de tinta, la industria aeroespacial o la industria biomédica. En concreto en esta última área nuevos dispositivos denominados “laboratorio en un chip” (en inglés lab-on-a-chip) se han desarrollado para obtener sistemas de análisis integrados y móviles. El objetivo a largo plazo sería eliminar pruebas clínicas en laboratorios clásicos que consumen mucho tiempo. Un aspecto común a muchas de estas tecnologías es el mezclado. Es decir, ¿cómo se homogeneizan o mezclan los elementos inmersos en el fluido de estos micro-artefactos? Responder a esta pregunta es importante porque en muchos de estos dispositivos la ausencia de mezclado obstaculiza su correcto funcionamiento. Recientemente una serie de artículos referenciados en [18] han tratado este tema. Allí se muestra que conceptos y teoremas relacionados con el Linked Twisted Map, propios de la matemática pura, son clave para el diseño óptimo de estas unidades. En concreto el primer artículo es una revisión matemática rigurosa de definiciones y elementos teóricos de los sistemas dinámicos necesarios para tratar estos problemas. Este campo indudablemente se beneficiaría de una colaboración más estrecha entre aplicaciones y teoría.

Bibliografía

- [1] Para el desarrollo de software especializado en fluidos para efectos especiales en cine veáse por ejemplo: <http://www.nextlimit.com>.
Sobre la descripción de otros algoritmos matemáticos (fuera del ámbito de los fluidos) y sus aplicaciones en la animación puede consultarse por ejemplo: <http://www.cds.caltech.edu/marsden/research/demos/droppingbunnies.ph>
- [2] DFD-APS Meeting ver <http://meetings.aps.org/Meeting/DFD05/sessionindex2>
- [3] http://www7.nationalacademies.org/usnctam/Fluid_Mechanics_II.html
- [4] Hurle, D.T.J.; “*Temperature oscillations in molten metals and their relationship to growth striae in melt grown crystal*”. *Phil. Mag.*, 13, (305-310), (1966).
- [5] Hoyas, S.; Herrero, H.; Mancho, A.M.; “*Thermal convection in a cylindrical annulus heated laterally*”. *J. Phys. A: Math and Gen.* 35, (4067-4083) (2002).
- [6] Hoyas, S.; Herrero, H.; Mancho, A.M.; “*Bifurcation diversity in dynamic thermocapillary liquid layers*”. *Phys. Rev. E* 66, (057301-1 - 057301-4) (2002).
- [7] Hoyas, S.; Mancho, A.M.; Herrero, H.; Garnier, N.; Chiffaudel, A.; “*Benard-Marangoni convection in a differentially heated cylindrical cavity*”. *Physics of Fluids* 17, (054104-1 - 054104-2) (2005).
- [8] Stakgold, I.; “*Branching of solutions of nonlinear Equations*”. *SIAM Rev.* 13, (289-332) (1971).
- [9] Crandall, M.G.; Rabinovitz, P.H.; “*Bifurcations from simple eigenvalues*”. *Funct. Anal.* 8, (321-356) (1971).
- [10] Levich, H.G.; Krylov, V.S.; “*Surface tension-driven phenomena*”. *Ann. Rev. Fluid Mech.* 1, (293-316) (1969).
- [11] Davis, S.H.; “*Thermocapillary instabilities*”. *Ann. Rev. Fluid Mech.* 19, (403-435) (1987).
- [12] Schatz, M.F.; Neitzel, G.P.; “*Experiments on thermocapillary instabilities*”. *Ann. Rev. Fluid Mech.* 33, (93-127) (2001).

- [13] Lekien, F.; Coulliette, C.; Mariano, A.J.; Ryan, E.H.; Shay, L.K.; Haller, G.; Marsden, J.; "Pollution release tied to invariant manifolds: A case study for the coast of Florida". *Physica D* 210 (1-2), 1 (2005).
- [14] Ottino, J.M.; Wiggins, S.; "Designing Optimal Micromixers". *Science* 305, (485-486), (2004).
- [15] Bayraktar, T.; Pidugu, S.B.; "Characterization of liquid flows in microfluidic systems". *Int. J. of Heat and Mass Transfer* 49 815 (2006).
- [16] Kamholz, A.E.; "Proliferation of microfluidics in literature and intellectual property". *Lab Chip* 4 (2), 16N (2004).
- [17] Stone, H.A.; Stroock, A.D.; Adjari, A.; "Engineering flows in small devices: Microfluidics toward a Lab-on-a-Chip". *Annu. Rev. Fluid Mech.* 36 381 (2004).
- [18] Wiggins, A.; Ottino, J.M.; Eds. *Philos. Trans. R. Soc. London Ser. A* 362 (15 May 2004).
- [19] Haller, G.; Yuan, G.; "Lagrangian coherent structures and mixing in two-dimensional turbulence" *Physica D* 147, 352 (2000).
- [20] Mancho, A.M.; Small, D.; Wiggins, S.; Ide, K.; "Computation of Stable and Unstable Manifolds of Hyperbolic Trajectories in Two-Dimensional, Aperiodically Time-Dependent Vector Fields". *Physica D* 182, 188 (2003).
- [21] Sobre la aplicabilidad de las LCS a los mecanismos de alimentación de las medusas en zoología marina ver: <http://www.cds.caltech.edu/marsden/research/demos/fluidtransport.php>
- [22] Coulliette, C.; Wiggins, S.; "Intergyre transport in a wind-driven, quasigeostrophic double gyre: An application of lobe dynamics". *Nonlinear Processes in geophysics* 8 (1-2), 69 (2001).



Capítulo XI

MATEMÁTICA COMPUTACIONAL: UN NUEVO PILAR PARA EL DESARROLLO CIENTÍFICO Y TECNOLÓGICO

POR FROILÁN MARTÍNEZ DOPICO
UNIVERSIDAD CARLOS III DE MADRID
Departamento de Matemáticas
28911 Leganés
Correo electrónico: dopico@math.uc3m.es

*"The mathematical sciences identify and study structures; they formulate powerful concepts that can unify and clarify phenomena in the natural and technological worlds, in the other sciences as well as within mathematics itself; **they organize and design efficient ways to compute**; they form the very language of science"*

NSF Mathematical Sciences Report, 2003.

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

A change in the Scientific Method

The Scientific Method has two classical pillars: experiments and theory. A third pillar has arisen in the last decades as a consequence of the widespread use of computers: *computational simulation*. Computational simulation is a method of scientific research that employs a computer system to study physical, biological or social systems according to laws derived from theory and experiments. Computational simulation is a very recent methodology. However, there is a general agreement among scientists of very different areas on the fact that most of the scientific and technological discoveries in the future will be influenced, motivated or related to modern computational methods.

A consequence of computation: science and technology are becoming more mathematical

Nowadays, the overwhelming majority of scientific and technological disciplines use computers very often. This forces scientists and technologists of a large variety of fields to interact with computers. Computers only work on mathematical objects -numbers and letters-, by performing on them mathematical operations-computations, sorting and classifications. As a consequence many sciences that in the past were disconnected of Mathematics are becoming more and more mathematical. This is specially significative in health and social sciences.

Today, for instance, the mathematical models used in Genomic can be more complicated than those used in a mathematically oriented science as Physics. Besides, many of the modern scientific models are specifically developed to be studied with computers. Their goal is very often to obtain precise quantitative information of phenomena that depend on many parameters. Therefore, modern scientific models require new and powerful techniques of Computational Mathematics that can be very different from the classical tools of Numerical Analysis. These techniques have to be developed by interdisciplinary teams of mathematicians and scientists or engineers of other disciplines.

A mean to do impossible experiments

Computational simulation is a method to perform impossible or prohibited experiments. For instance, it is impossible to do experiments to see how a large pandemic spreads. Only computational simulation can estimate its behavior. The same applies to natural catastrophes as earthquakes, tsunamis, fires, droughts, etc.

A mean to increase profits in Industry and Finances

Computational simulation of new designs in industry, as for instance, cars, aircrafts, satellites, etc, can save many millions of euros compared to building a series of prototypes. Similar considerations remain valid for the design of new computers architectures involving high risk innovations, or for the design of new drugs in pharmaceutical industry. Computer simulation of financial markets is at present a powerful tool in the design of market strategies, still under development.

What is Computational Mathematics?

We have shown that Computational Mathematics is an important discipline that appears in many scientific, technological and industrial scenarios. It is time to present a definition of this discipline. We believe that Computational Mathematics has to be an interdisciplinary subject. We propose the following wide definition: *Computational Mathematics is the collection of computer programs, algorithms, techniques and theories required to solve on a computer mathematical models of problems in science and technology.* Here, we understand the word science in its most ample sense, including from the classical sciences as Physics, Chemistry, Biology, Geology, Medicine, etc, to Social or Health sciences. We insist that Computational Mathematics must solve problems. It is not enough to develop Theorems or Algorithms, also the practical computer implementation of the algorithms, i.e., the *software*, has to be developed.

New algorithms in Teraflops age

Today, the most efficient computers are able to perform 10^{12} floating point operations per second, i.e., one *Teraflop*. This spectacular number refers to the so-called *peak performance* of the computers, i.e., the maximum speed they are able to reach. This super-fast computers are sophisticated machines and to obtain their maximum performance is a very difficult task that requires to develop new generations of algorithms. These algorithms have to fill the gap between computer architectures that are available and the applications that must be executed.

Un cambio en el método científico

Los dos pilares clásicos del método científico son los experimentos y la teoría. La teoría tiene por objetivo explicar los experimentos, predecir el resultado de nuevos experimentos y contribuir al diseño y la construcción de "máquinas" que pongan los fenómenos naturales al servicio del bienestar de la humanidad. La teoría y experimentación son dos aspectos plenamente establecidos dentro del desarrollo científico y tecnológico. Su utilidad está refrendada por la ingente cantidad y la gran calidad de los logros científicos y tecnológicos alcanzados en los últimos siglos en campos tan diversos como las ciencias espaciales, las ciencias de la vida, la producción



de energía, el mundo atómico y subatómico, la invención y desarrollo de los ordenadores, las ciencias de la computación, etc. Precisamente la generalización del uso de los ordenadores ha hecho que en las dos últimas décadas haya aparecido **un tercer pilar del método científico: la simulación computacional**.

Podemos definir informalmente la simulación computacional como un método de investigación científica que emplea los ordenadores para simular sistemas físicos, biológicos o sociales de acuerdo a las leyes derivadas de la teoría y los experimentos. Comparada con la teoría y la experimentación, la simulación computacional es una disciplina muy joven, aún no totalmente establecida, cuyos métodos específicos y sus conexiones con la teoría y la experimentación se encuentran en desarrollo. No sabemos cuánto puede dar de sí, pero sospechamos que mucho, es más, casi sospechamos que la mayor parte de los grandes logros científicos y tecnológicos del futuro estarán influidos por, motivados por o relacionados con los modernos métodos de computación.

Una consecuencia de la computación: la matematización de las ciencias y las tecnologías

En consonancia con lo descrito en la sección anterior es difícil imaginar en la actualidad alguna disciplina científica o tecnológica que no haga un uso frecuente de los ordenadores. Eso obliga a los científicos y tecnólogos de ámbitos muy diversos como medicina, biología, ciencias sociales, etc, a comunicarse con un ordenador. Los ordenadores sólo trabajan sobre objetos matemáticos -números y letras- y sólo pueden realizar sobre ellos operaciones matemáticas -cálculos, ordenaciones y clasificaciones (que no corresponden a otra cosa que al concepto algebraico de relación). Ello ha provocado de manera natural que muchas ciencias que tradicionalmente tenían poco que ver con las matemáticas se hayan *matematizado* enormemente en la última década. Este proceso es especialmente llamativo en las ciencias sociales y de la vida.

Ya no son sólo las disciplinas tradicionalmente más matemáticas como la física, la química y las ingenierías las que requieren el uso de técnicas matemáticas sofisticadas. Ahora los modelos matemáticos requeridos, por ejemplo, por la genética pueden ser tanto o más complicados que los de la física y la ingeniería, y además se hallan menos desarrollados y comprendidos. Una característica de los nuevos modelos es que con frecuencia nacen por y para ser estudiados computacionalmente con ordenadores. Su objetivo es obtener información cuantitativa precisa de fenómenos muy complicados que involucran muchos parámetros, frente a los modelos más simples de la ciencia de hace treinta años que pretendían, o bien obtener información cualitativa de la naturaleza a costa de despreciar muchos de los factores que intervenían en un fenómeno, o bien describir cuantitativamente fenómenos dependientes de un número reducido de parámetros.

Todo esto hace distintos a los modelos matemáticos que actualmente aparecen en la ciencia y la tecnología: son usualmente más complejos y, por ello, más exigentes y retadores respecto de las técnicas de matemática computacional que requieren.

En el año 2002 asistimos a uno de los logros científicos más importantes de los últimos años: la descripción completa de la secuencia del genoma humano [14, 8]. Menos conocido es que detrás de este logro se encuentra un algoritmo matemático que redujo drásticamente el tiempo necesario para completar la secuencia. Dicho algoritmo, conocido como *whole genome shotgun sequencing* [15, 7], fue considerado visionario y condenado al fracaso por muchos grupos de investigación involucrados en la descripción del genoma humano. Dichos grupos utilizaban algoritmos de tipo *divide y conquista* -que necesitaban un tiempo prohibitivo de computación- y mantenían que la aproximación *whole genome shotgun sequencing* no podría utilizarse para hallar secuencias de los genomas de mamíferos. Otro detalle muy significativo de esta historia es que la persona que desarrolló este exitoso algoritmo fue un matemático, Eugene Myers, que obtuvo su licenciatura en matemáticas en 1975 en el California Institute of Technology, y su doctorado en *Computer Science* en 1981 en la University of Colorado, con una tesis doctoral sobre un problema eminentemente matemático titulada *A Depth-First Search Characterization of k-Connectivity and Its Application to Connectivity Testing*.

De esta pequeña historia sobre un hito científico tan importante debemos sacar algunas conclusiones: la matemática computacional fue la clave de la resolución del problema; el tipo de matemática computacional que se utilizó fue muy distinta de lo que clásicamente se entiende en matemáticas como análisis numérico; y la interdisciplinariedad jugó un papel fundamental. Pero sobre todo, ¿quién podría pensar hace 20 años en que las palabras algoritmo y biología aparecieran juntas en un logro científico de primera línea?

Terminamos esta sección con una observación fundamental: muchas ciencias y tecnologías vuelven en la actualidad sus ojos hacia las matemáticas porque las matemáticas pueden realizar, gracias a la existencia de ordenadores, cálculos y tareas que eran inimaginables hace unos pocos años. En este contexto recae sobre la *matemática computacional la responsabilidad de llenar, mediante el desarrollo de algoritmos innovadores, el hueco existente entre las arquitecturas de los ordenadores actualmente disponibles y las aplicaciones que deben ser ejecutadas*. Hoy más que nunca es cierta la famosa frase de Galileo: *El gran libro de la naturaleza está escrito en símbolos matemáticos*, pues cada vez se aplica a más disciplinas. Muchas de estas disciplinas no han usado hasta hace pocos años los resultados de las matemáticas y por lo tanto plantean nuevas exigencias a las matemáticas. La matemática computacional tiene que contribuir también a hacer accesibles *computacionalmente* a otras ciencias el mayor número posible de resultados de las matemáticas puras y abstractas. Si son accesibles, muchos de ellos serán utilizados.

Un medio para realizar experimentos irrealizables

La simulación computacional constituye un método para realizar experimentos irrealizables. Por citar un problema de actualidad: no se puede hacer un experimento sobre cómo se extiende una epidemia de grandes magnitudes, menos aún si dicha epidemia afecta a seres humanos. Sin embargo, desafortunadamente, las epidemias aún existen y debemos saber cómo se propagan antes de



que aparezcan. Sólo la simulación computacional de modelos matemáticos para estos problemas puede darnos una idea de su comportamiento. Otros problemas de este tipo incluyen las catástrofes naturales -terremotos, maremotos, incendios, inundaciones, sequías, etc.- y posibles ataques terroristas con armas biológicas.

Predecir los efectos de procesos como los citados en el párrafo anterior sólo será posible con modelos matemáticos masivamente computacionales, que requieren el desarrollo de nuevos algoritmos para el uso inteligente de ordenadores de gran potencia. Decimos que estos modelos son masivamente computacionales porque deben incorporar multitud de términos y efectos que los hacen intratables analíticamente. No se trata sólo de tener una idea cualitativa de cómo se propaga una epidemia, sino información cuantitativa precisa que permita tomar decisiones reales que involucran intervalos de tiempo reales y vidas humanas. No estamos aún en ese nivel de predicción, pero los avances computacionales en ciencias predictivas como la meteorología nos hacen pensar que podemos no estar lejos, siempre que se hagan los esfuerzos científicos y presupuestarios adecuados.

Un medio para aumentar beneficios en la industria y las finanzas

La simulación computacional de nuevos diseños en la industria, como por ejemplo de aviones, satélites, coches y barcos, puede contribuir a ahorrar muchos millones de euros en industrias como la automovilística y la aero-espacial, si comparamos sus costes con la construcción de prototipos reales de coches, aviones, barcos o satélites. Es obvio que la simulación computacional no puede, ni debe, evitar la construcción de ciertos prototipos finales sobre los que analizar el comportamiento real, pero evita la construcción de muchos diseños que fracasarían, aumenta la seguridad de los diseños construidos y contribuye a mejorar la calidad final de los productos.

Todo ello redunda en una jerarquía de beneficios: para los usuarios, en cuanto a calidad y precio de los productos obtenidos, y financieros para las empresas involucradas al favorecer la reducción de costes.

Consideraciones similares a las presentadas en el párrafo anterior se aplican en muchos otros ámbitos como el propio diseño de nuevos ordenadores con arquitecturas que involucren innovaciones de alto riesgo, el desarrollo de nuevos fármacos, la planificación de misiones espaciales, la simulación del comportamiento de los mercados financieros, etc. De nuevo en el mundo de los mercados financieros nos encontramos con la imposibilidad de realizar experimentos, o estos pueden ser muy peligrosos y causar crisis financieras de distintos niveles. Sólo la simulación computacional puede proporcionarnos información.

¿Qué es la matemática computacional?

Hemos evitado hasta el momento dar una definición concreta de lo que entendemos por matemática computacional. Nuestro objetivo ha sido mostrar su importancia y ubicuidad en la mayor par-

te de los desarrollos científicos y tecnológicos actuales antes de entrar en una descripción más específica de esta disciplina. Desde nuestra perspectiva la matemática computacional debe entenderse como una actividad integradora y multidisciplinar. Por ello proponemos una definición amplia: *la matemática computacional es el conjunto de programas de ordenador, algoritmos, técnicas y teorías necesario para resolver en un ordenador modelos matemáticos de problemas que surgen en la ciencia y la tecnología* [4]. Queremos insistir en que entendemos ciencia en el sentido más amplio posible, englobando en este término tanto la economía, las ciencias sociales y las ciencias de la salud, como las ciencias clásicas: física, química, biología, geología, etc.

Remarquemos una palabra que aparece en la definición anterior: *resolver*. La matemática computacional debe proporcionar soluciones prácticas y concretas a los problemas. En ese sentido se basa en disciplinas matemáticas clásicas y, por supuesto, en el análisis numérico, pero va más allá de ellas. A la matemática computacional no le basta con desarrollar y analizar nuevos algoritmos, debe proporcionar los programas optimizados que los implementen, resolviendo también todos los problemas prácticos que esto requiere.

Las matemáticas puras pagan, por su belleza y abstracción, un alto precio: desarrollan un conocimiento estructural, pero no siempre poseen una comprensión profunda de los objetos individuales que desean estudiar. Las aplicaciones, por el contrario, estudian objetos individuales concretos. La matemática computacional también estudia problemas concretos, se ve forzada a ello por su unión con los ordenadores, que sólo pueden trabajar de una manera concreta, no abstracta. Como ya hemos comentado, uno de los objetivos prioritarios de la matemática computacional es también contribuir a hacer útiles, computables, y posiblemente aplicadas muchas ideas poderosas de las matemáticas que habitan aún en el mundo de la abstracción [3].

Nuevos algoritmos para la supercomputación en la era de los Teraflops

Estamos en la era de los Teraflops: los ordenadores actualmente más potentes son capaces de realizar más de un Teraflop (10^{12} = un millón de millones) de operaciones por segundo. Ser capaces no significa que siempre lo hagan. De hecho estos números espectaculares se refieren a lo que se conoce como *peak performance*, es decir, la velocidad máxima de cálculo que son capaces de alcanzar [1].

Estos supercomputadores son máquinas complicadas que trabajan en paralelo. Conseguir que alcancen su máxima eficiencia al resolver un determinado problema requiere el desarrollo de nuevos y sofisticados algoritmos. Queremos dejar claro que no se trata de reprogramar los algoritmos existentes, sino de desarrollar nuevos métodos numéricos que se adapten a las nuevas arquitecturas. Ello exige, simultáneamente, nuevos análisis de convergencia, errores y estabilidad. Se trata de hacer nuevas matemáticas que estén desde un principio en perfecta comunión con las arquitecturas disponibles [6]. Sin estas matemáticas la potencia de los ordenadores más modernos queda seriamente limitada.



La velocidad y prestaciones de los ordenadores han aumentado espectacularmente en las últimas décadas, y seguirán haciéndolo, pero con seguridad no al mismo ritmo. En cualquier caso no a un ritmo suficiente como para enfrentarse con los algoritmos actualmente existentes a los complejos modelos que surgen hoy en día en las aplicaciones científicas y tecnológicas. *Estamos necesariamente en una era de nuevos algoritmos.* Además no hay que olvidar que en nuestra breve era de simulación computacional -a lo más de unas cinco décadas de duración- los algoritmos por sí solos han logrado aumentar la velocidad de cálculo en muchos problemas científicos y tecnológicos en un factor comparable al resultante del aumento de prestaciones de los ordenadores. Un ejemplo famoso es la evolución de los algoritmos para resolver la ecuación de Poisson en tres dimensiones, un problema que surge en multitud de aplicaciones, y en particular en el cálculo de potenciales electrostáticos [6]. Para este problema el desarrollo de nuevos algoritmos ha logrado, por sí solo, que los cálculos sean dieciséis millones de veces más rápidos (;16.000.000!).

Concluimos con una observación: el desarrollo de nuevos algoritmos sólo tiene como límite el ingenio y la inteligencia de los hombres, mientras que la velocidad de los ordenadores tienen límites físicos que no tardarán en alcanzarse.

Nuevos horizontes en Matemática Computacional

Terminamos esta breve panorámica sobre la matemática computacional enumerando algunos problemas actuales en los que existe una intensa actividad investigadora. No pretendemos, ni podemos en tan breve espacio, ser exhaustivos. La selección depende de nuestro conocimiento y preferencias personales, y sin duda dejamos temas importantes sin mencionar.

Investigación básica en métodos numéricos y algoritmos para problemas matemáticos clásicos

Cuando un investigador o un profesional se enfrenta a la tarea de resolver en un ordenador un modelo matemático para un problema aplicado, lo primero que hace es buscar en la literatura y en las librerías de software existentes antes de inventar un nuevo algoritmo. Muchas veces esto es suficiente y, en cualquier caso, casi siempre los algoritmos existentes para tareas básicas se utilizan como componentes en los nuevos algoritmos. Hay aún muchos problemas básicos para los que los algoritmos existentes no son satisfactorios, o son susceptibles de mejoras que pueden incrementar significativamente su precisión, velocidad o rango de aplicabilidad. Esta investigación comprende los problemas clásicos del análisis numérico y es un aspecto esencial para el avance de la matemática computacional. Entre las disciplinas más importantes citamos las siguientes.

- *Algoritmos para resolver numéricamente ecuaciones diferenciales en derivadas parciales:* Estas ecuaciones son la base de muchos de los modelos matemáticos que se utilizan en la simula-

ción computacional de aquellos problemas de ciencia e ingeniería que son continuos. Citamos algunos temas actuales de investigación: técnicas de mallado adaptativo y multimalla para optimizar los métodos de elementos finitos; nuevas perspectivas en métodos espectrales de alto orden; descomposición en dominios, etc.

- *Algoritmos para resolver numéricamente ecuaciones diferenciales ordinarias:* Estas ecuaciones son fundamentales en los problemas de mecánica clásica y por lo tanto aparecen en multitud de aplicaciones en ingeniería. La investigación actual se centra en el diseño de métodos específicos adaptados a las características particulares de los problemas a tratar, en particular, métodos exponencialmente adaptados para problemas de gran dimensión, métodos adaptados para problemas con discontinuidades, métodos Runge-Kutta multirresolución, fórmulas para sistemas Hamiltonianos, etc.
- *Algoritmos en geometría computacional y diseño geométrico asistido por ordenador (CAGD):* con sus aplicaciones al reconocimiento de formas, la morfología computacional, el diseño de circuitos integrados a muy alta escala (VLSI), la visión artificial, los sistemas de información geográfica, las redes de transporte y comunicación, y la robótica. Ello comporta el estudio de representación de curvas y superficies, problemas de intersección, representación y manipulación de figuras y objetos, interpolación en varias variables, estructuras de datos geométricos, complejidad algorítmica, etc.
- *Algoritmos de álgebra lineal numérica:* Se trata de un campo que proporciona herramientas básicas para la resolución matemática y numérica de numerosos problemas científicos y tecnológicos con gran impacto potencial. Los algoritmos del álgebra lineal numérica constituyen un cuello de botella para la eficiencia y precisión de muchos de los algoritmos que aparecen en el resto de las disciplinas numéricas. Es por ello que son los algoritmos más sofisticados y optimizados desde el punto de vista de precisión y eficiencia. Hay una gran actividad investigadora en la obtención de métodos adaptados a estructuras particulares y destacan por su interés en las aplicaciones métodos de aproximación de valores propios y valores singulares con alta precisión relativa. Investigaciones en curso como los métodos super-rápidos ($O(n^2)$) para la multiplicación de dos matrices cualesquiera pueden cambiar drásticamente la eficiencia de multitud de algoritmos numéricos de disciplinas diversas.

Algoritmos paralelizables para problemas matemáticos

Destacamos por separado la importancia de la paralelización. Podemos decir que entran aquí muchos de los problemas matemáticos aplicados, incluidos los de la sección anterior. La paralelización constituye un problema muy difícil, pues requiere un conocimiento profundo tanto del problema matemático en consideración como de la arquitectura del ordenador a utilizar. Exigen con frecuencia el desarrollo de nuevos métodos numéricos, pues los algoritmos secuenciales pueden no ser utilizables en paralelo. Sin embargo, de este tipo de algoritmos depen-



derá la resolución de muchos problemas científicos y tecnológicos de gran complejidad que aparecen en disciplinas como: el diseño de fármacos, la genética, problemas medioambientales de carácter global (calentamiento global, contaminación), pruebas de accidentes en aerodinámica, etc. [6].

Nuevas disciplinas en matemática computacional

Incluimos aquí, como mero botón de muestra, tres disciplinas no tratadas clásicamente en análisis numérico y que sin embargo tienen actualmente un gran impacto tecnológico y aplicado, a la par de dar lugar a problemas matemáticos y computacionales de gran envergadura. Su denominador común es que tratan problemas discretos, bien geométricos o combinatorios. Debe quedar claro que existen muchos otros problemas que están recibiendo actualmente una gran atención y que no mencionamos por concisión.

- *Procesamiento de imágenes (Imaging Science)*: Es quizá una de las disciplinas modernas que plantea problemas matemáticos más novedosos y retadores y a la que más esfuerzo y financiación se le dedica en Estados Unidos. Una de sus aplicaciones más importantes la constituye el *procesamiento de imágenes de distintos métodos de diagnóstico en medicina*. Existen revistas especializadas exclusivamente en ello como *IEEE Transactions on Medical Imaging* y ya se ha acuñado el término inglés de *Imaging Science*. Como problemas matemáticos específicos citamos: las matemáticas del reconocimiento de caras, el uso de productos tensoriales o de Kronecker en restauración de imágenes, el problema del *inpainting* o interpolación de imágenes.
- *Generación de mallas*: Es un problema motivado originalmente por los métodos numéricos de elementos finitos para resolver ecuaciones en derivadas parciales, pero tiene en la actualidad muchas más aplicaciones como, por ejemplo, en gráficos computacionales. Las técnicas de elementos finitos necesitan dividir una región del plano o del espacio en “pedazos” finitos que tengan buenas propiedades. Hace unos pocos años esto se hacía de manera ad-hoc para cada problema en consideración pero en la actualidad este problema surge con tanta frecuencia que ha aparecido una gran demanda de algoritmos que realicen esta tarea de forma óptima y automática. El problema puede considerarse resuelto en dos dimensiones gracias al algoritmo implementado en el celebrado código *triangle* [11, 12] que ha supuesto un gran reconocimiento y diversos premios para su creador Jonathan Shewchuk [10]. Este tipo de problemas sigue abierto en más de dos dimensiones y en dos dimensiones para otro tipo de mallas. En la Figura 1 mostramos el uso de *triangle* en la simulación de la propagación de corrientes eléctricas en el miocardio [9].

FIGURA 1

Las mallas triangulares producidas por *triangle* han sido usadas en la simulación de la propagación de corrientes eléctricas en el miocardio [9]

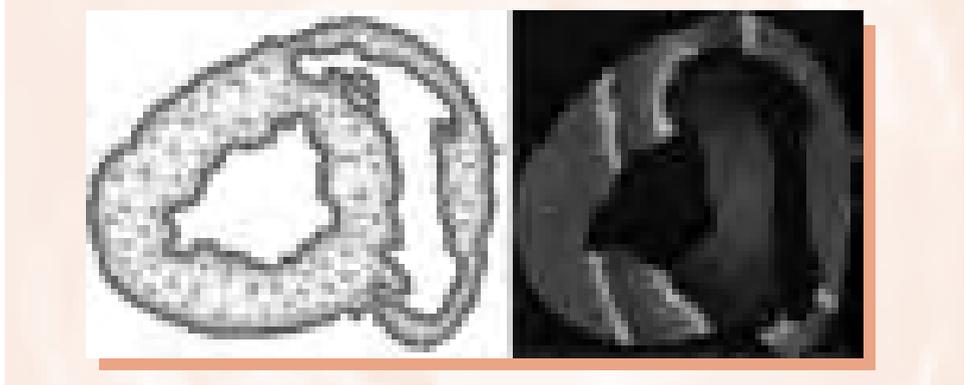
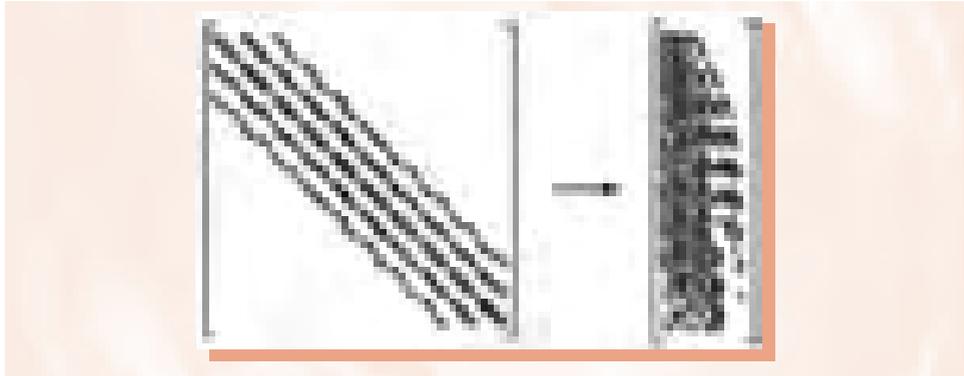


FIGURA 2

Matriz Jacobiana y su representación comprimida que da lugar al desarrollo mediante grafos de algoritmos eficientes para su cálculo [5]



- *Computación científica y problemas combinatorios*: Esta moderna disciplina trata sobre el desarrollo, análisis y aplicación de algoritmos combinatorios para resolver problemas computacionales científicos y tecnológicos. Algunos de los temas actuales en este campo son: desarrollo de preconditionadores para métodos iterativos mediante el estudio de inmersiones de grafos [13]; y el uso de técnicas combinatorias de grafos en problemas no lineales para el cálculo eficiente de Jacobianos y Hessianos [5]. Ver Figura 2.

Estas técnicas están relacionadas con uno de los problemas computacionales abiertos más interesantes y aplicados: la búsqueda de algoritmos eficientes y estables para la *diferenciación automática*. Los algoritmos de diferenciación automática condicionan la eficiencia y estabilidad de los métodos de Newton para resolver sistemas de ecuaciones no-lineales con muchas incógnitas.



Conclusión y referencias

Es un hecho cierto que muchos de los avances científicos y tecnológicos de las últimas décadas descansan en los métodos y capacidades de la matemática computacional. Las perspectivas de futuro son que esta tendencia se incrementará cada vez más. Por ello hay que apostar decididamente por la matemática computacional, tanto a nivel de investigación como de educación. Esto implicaría que muchos de los matemáticos de las nuevas generaciones estarían entrenados para resolver problemas desde el principio hasta el final, lo que implica una formación multidisciplinar que haga natural la colaboración con otros científicos.

He reunido en este documento ideas y formas de pensar que he encontrado dispersas en muchos lugares. Mis opiniones se han visto especialmente influidas por lecturas de artículos en *SIAM News* y por varios de los excelentes documentos que SIAM (Society for Industrial and Applied Mathematics) proporciona en la página web <http://www.siam.org/about/science/>.

Bibliografía

- [1] Cipra, B.A.; *SC2002: A Terable Time for Supercomputing*, SIAM News, 36, 2 (2003), (1-3).
- [2] Cohn, H.; Kleinberg, R.; Szegedy, B.; Umans, C.; *Group-theoretic Algorithms for Matrix Multiplication*, Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, (2005), (379-388). (Invited to FOCS 2005 Special Issue)
- [3] Edelman, A.; Elmroth, E.; Kagström, B.; *A geometric approach to perturbation theory of matrices and matrix pencils. Part I: versal deformations*, SIAM J. Matrix Anal. Appl., 18 (1997), (653-692).
- [4] Golub, G.; Ortega, J.M.; *Scientific Computing: An Introduction with Parallel Computing*, Academic Press, San Diego, (1993).
- [5] Gebremedhin, A.; Manne, F.; Pothen, A.; *What color is your Jacobian? Graph coloring for computing derivatives*, SIAM Review, 47 (4), (2005), (629-705).
- [6] Keyes, D.; *Computational and Applied Mathematics in Scientific Discovery*, <http://www.siam.org/about/science/sci-news.php>, (2004).
- [7] Myers, E.; *Whole-Genome DNA Sequencing*, IEEE Computational Engineering and Science 3, 1, (1999), (33-43).
- [8] Myers, E.W.; Sutton, G.G.; Smith, H.O.; Adams, M.D.; Venter, J.C.; *On the Sequencing and Assembly of the Human Genome*, Proc. Natl. Acad. Sciences 99, 7, (2002), (1661-1671).
- [9] Pormann, J.B.; Board, J.A.; Rose, D.J.; Henriquez, C.S.; *Large-scale modeling of cardiac electrophysiology*, Computers in Cardiology 2002. Volume 29 (IEEE) (Cat. No.02CH37421) (2002), (259-62).
- [10] Robinson, S.; *Hooked on meshing, Researcher Creates Award-Winning Triangulation Program*, SIAM News 36, 9, (2003), (1-8).
- [11] Shewchuk, J.; *Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator*, in Applied Computational Geometry: Towards Geometric Engineering (Ming C. Lin and Dinesh Manocha, editors), volume 1148 of Lecture Notes in Computer Science, Springer-Verlag (Berlin), (May 1996), (203-222).

- [12] Shewchuk, J.; *Delaunay Refinement Algorithms for Triangular Mesh Generation*, Computational Geometry: Theory and Applications 22, (2002), (21-74).
- [13] Spielman, D.; Teng, S.; *Nearly linear time algorithms for graph partitioning, graph sparsification, and solving linear systems*, Proceedings of the 36th Annual ACM Symposium on Theory of Computing, (2004), (81-90).
- [14] Venter, J.C.; Adams, M.D.; Myers, E.W.; et al., *The Sequence of the Humane Genome*, Science 291 (2002), (1304-1351).
- [15] Weber, J.; Myers, E.; *Human Whole Genome Shotgun Sequencing*, Genome Research 7 (1997), (401-409).



Capítulo XII

MATEMÁTICAS, CONTROL Y ROBÓTICA

POR SONIA MARTÍNEZ DÍAZ
UNIVERSITY OF CALIFORNIA, SAN DIEGO
Mechanical and Aerospace Engineering
Jacobs School of Engineering
La Jolla, CA 92093, USA
Correo electrónico: soniamd@ucsd.edu
Página web: <http://flyingv.ucsd.edu/sonia>

JORGE CORTÉS MONFORTE
UNIVERSITY OF CALIFORNIA, SANTA CRUZ
Applied Mathematics and Statistics
Baskin School of Engineering
Santa Cruz, CA 95064, USA
Correo electrónico: jcortes@ucsc.edu
Página web: <http://www.ams.ucsc.edu/~jcortes>

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

In the near future, new generations of robots and autonomous vehicles will play a fundamental role in scientific and commercial applications of great social impact. Some of their uses include distributed monitoring and surveying (e.g., crop monitoring and management), rapid deployment in disaster relief and situation awareness operations, environmental monitoring (e.g., study of oceanographic and atmospheric interactions), health monitoring of civil infrastructure (bridges, buildings, oil pipes, etc); and distant planet image generation (e.g., mobile satellites equipped with interferometers). In these contexts, coordinated, sensor-equipped vehicles will carry out a variety of search and rescue, data gathering and fusion, detection and estimation tasks.

The complexity of these multi-robot systems presents new challenges that lie at the confluence of communication, computation and control. Although technology provides the physical components to build such sensor networks, the potential benefits of such systems have not been realized yet. As of today, there is a lack of understanding on how to coordinate and assemble the individual devices together. In other words, there are not systematic methodologies that allow to control large-scale distributed systems like these. As a consequence, there is a great necessity to expand the currently available set of tools and paradigms to design and manage these systems in an efficient manner. In particular, it is envisioned that new substantial contributions will be made from the Mathematical Sciences. In this article we enumerate some of the challenges that are associated with robotic systems of this nature.

Introducción

En un futuro cercano, nuevas generaciones de robots y vehículos autónomos jugarán un papel crítico en aplicaciones científicas y comerciales de gran impacto social. Algunas de estas aplicaciones incluyen sistemas de observación y supervisión del medio ambiente (por ejemplo, grupos de sensores químicos portátiles capaces de detectar componentes tóxicos), sistemas de despliegue rápido en situaciones catastróficas (por ejemplo, redes de robots móviles para prevención de desastres naturales y atención de emergencia), redes autónomas de muestreo en estudios biológicos y oceanográficos (por ejemplo, observación de especies amenazadas, validación de modelos micro-climáticos), supervisión de infraestructuras críticas (puentes, edificios, gaseoductos, etc) o estudio y obtención de imágenes de planetas distantes (por ejemplo, por medio de grupos de vehículos espaciales equipados con interferómetros navegando en formación).

En estos contextos, sensores y vehículos coordinados desarrollarán operaciones de búsqueda y rescate, reconocimiento, recogida y fusión de datos, detección y estimación de polución, etc. La Figura 3 muestra algunas plataformas en desarrollo:

FIGURA 3

De izquierda a derecha, ejemplos tomados de UIUC Multi Rover Laboratory, NASA Terrestrial Planet Finder y Sandia National Laboratory



Las ventajas potenciales que resultarán de utilizar colecciones de sensores son numerosas. Por ejemplo, ciertas tareas son difíciles, si no imposibles, cuando son realizadas por un solo agente. Además, un grupo de agentes ofrece de manera inherente una resistencia más robusta ante fallos de los agentes individuales o de los canales de comunicación. En las aplicaciones mencionadas, se emplearán sistemas autónomos tales como robots móviles; vehículos anfibios y todo-terreno; robots espaciales para construcción y servicio; robots con patas, robots-serpiente, robots-pezuña y otros robots biomiméticos. Estos sistemas robóticos estarán gobernados por dinámicas híbridas no lineales y se moverán en medios cambiantes y probablemente inseguros. Interaccionarán con su entorno vía ligaduras de rodamiento, impactos y fuerzas viscosas. Los avances tecnológicos en computación, procesos de fabricación y comunicaciones harán posible que estos dispositivos puedan operar con niveles cada vez mayores de autonomía y destreza.

Tales sistemas estarán controlados mediante esquemas jerárquicos que les permitirán operar autónomamente, interactuar y cooperar con otros robots, y recibir órdenes de un operador humano. Una de las capacidades más básicas e importantes de un robot es la habilidad para planificar y ejecutar con seguridad sus propios movimientos. Esta capacidad es crucial para poder desarrollar lenguajes de programación robótica más sofisticados y poder realizar comportamientos grupales tales como la creación de formaciones y la manipulación cooperativa.

Desde una perspectiva matemática, numerosas áreas de investigación, tanto fundamental como aplicada, tienen una relevancia directa en los problemas de control y robótica a los que nos enfrentamos para hacer posible estos avances. En la actualidad, conceptos, herramientas y métodos de Mecánica Geométrica, Teoría de Sistemas Dinámicos, Teoría de Estabilidad, Análisis No-diferenciable, Teoría de Sistemas, Investigación Operativa, Optimización Distribuida, Ciencias de la Computación, Geometría Computacional, Algoritmos Distribuidos, Teoría de Grafos, Control de Topologías, Estimación y un largo etcétera, están jugando un papel fundamental en el desarrollo de estas capacidades. *El papel de las Matemáticas es crítico a la hora de hacer realidad los escenarios descritos anteriormente.*



Los futuros escenarios y las novedosas aplicaciones de robots autónomos hacen necesario el desarrollo del conjunto de herramientas de control actualmente disponibles. En este informe, nos centramos en dos clases de problemas en los que los avances matemáticos prometen tener un gran impacto:

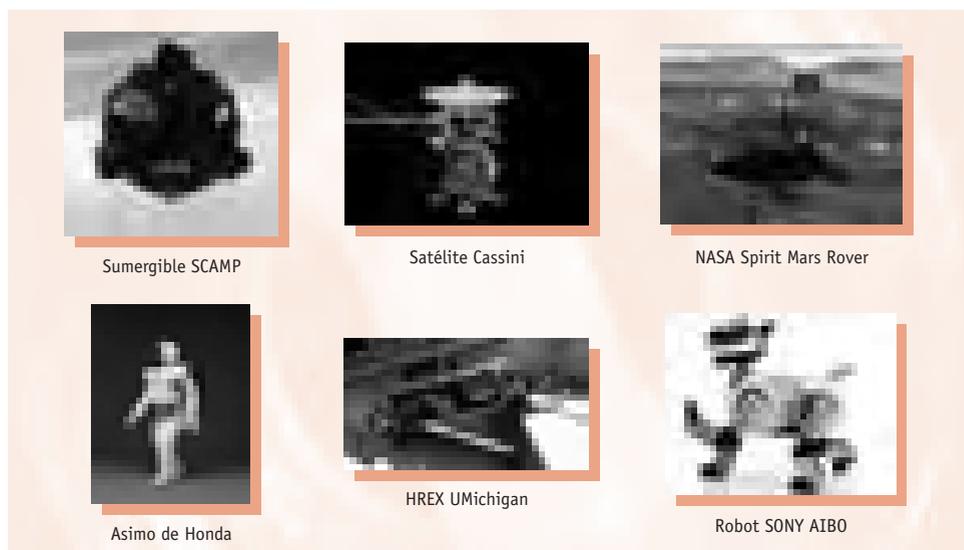
- (i) el control de movimientos de sistemas autónomos (Sección 2).
- (ii) la coordinación de grupos de sistemas autónomos para realizar cooperativamente tareas específicas (Sección 3).

Control de movimientos de sistemas autónomos

Hoy en día disponemos en gran variedad de ambientes de toda clase de sistemas -desde satélites en el espacio hasta sumergibles en el océano- que operan con cada vez mayores niveles de autonomía (ver Figura 4). El funcionamiento correcto de estos sistemas se basa en la selección y ejecución de algoritmos que les permiten realizar sus tareas y adaptarse a las condiciones imprevistas de su entorno. En particular, todo sistema que requiera algún tipo de movilidad precisa resolver el problema fundamental del *control de movimientos*.

FIGURA 4

Vehículos autónomos y robots biomiméticos



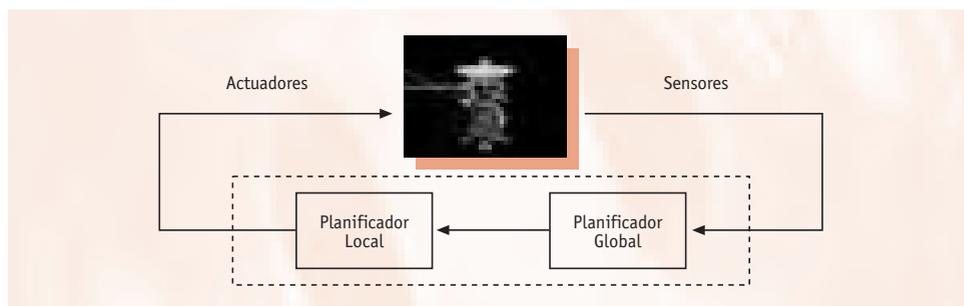
Por esta razón, uno de los temas principales en robótica es el diseño de mecanismos novedosos de locomoción, tanto físicos como de software, que permitan la fabricación de prototipos más ágiles y hábiles. Así por ejemplo, se han desarrollado múltiples variaciones de herramientas básicas, tales como la rueda. En el *Spirit*, uno de los dos rovers enviado a Marte por la NASA en el 2004, cada

rueda posee un motor individual para lograr mayor versatilidad sobre terrenos rugosos [8]. Otra alternativa viene dada por los robots “con patas” (ver Figura 4), que estarían mejor adaptados a suelos con escaleras y hoyos. Esta solución está claramente inspirada en la Naturaleza y es parte de un esfuerzo más amplio por reproducir toda una clase de sistemas “biomiméticos” o “biomorfos”. La motivación tras esta línea de investigación va desde la búsqueda de nuevas formas de locomoción, pasando por la posible explotación comercial (robots de entretenimiento), hasta su utilización como herramienta para entender y reproducir el funcionamiento de los seres vivos (lo que se ha dado en llamar “integrative biology” [7]). En cualquier caso, la fabricación del robot es uno de los pasos necesarios para la realización de un sistema autónomo. Otro aspecto fundamental es el desarrollo de los algoritmos que nos permitan obtener el máximo partido de la tecnología disponible.

Una tendencia generalizada en la búsqueda de algoritmos para el control de movimientos de un robot es la definición de métodos ad hoc, que funcionan muy bien para un sistema robótico concreto, pero no pueden ser extrapolados a otros sistemas similares. Sin embargo, se puede observar que gran número de prototipos poseen características comunes que hacen posible su tratamiento de forma unificada. *Esta perspectiva ha generado algoritmos válidos para una amplia clase de sistemas, que hacen un uso extensivo de la estructura matemática de sus modelos (ver, por ejemplo, [1, 2, 5]).* Estos algoritmos sirven para definir primitivas del movimiento, es decir, movimientos sencillos que pueden combinarse en un comportamiento global de sistema. El control de los sistemas robóticos mediante primitivas se emplea en gran número de prototipos y se conoce como *control de bajo nivel* de un sistema embebido. Tal y como se ilustra en la Figura 5, el Planificador Global o *control de alto nivel* del sistema autónomo es el encargado de elegir la combinación adecuada de primitivas que el Planificador Local o *control de bajo nivel* genera en respuesta a las condiciones externas del sistema. Esto da lugar a un control por realimentación o por “feedback”, garantizando robustez frente a perturbaciones externas. Las características deseables de las primitivas son: que sean robustas, implementables en tiempo real y poco costosas desde el punto de vista computacional y sensorial. En este sentido, el tipo de algoritmos que están desarrollando desde esta perspectiva unificadora son aplicables a sistemas con un número limitado de medios de acción o “actuadores” y, por tanto, no tan costosos desde el punto de vista del hardware. Como consecuencia, estas primitivas pueden ser útiles en la definición de mecanismos de seguridad para sistemas en los que puedan fallar algunos de sus actuadores.

FIGURA 5

Sistema moderno de control



Coordinación de sistemas autónomos

La coordinación de sistemas autónomos está adquiriendo una gran importancia en numerosas áreas de la ingeniería. El despliegue de grandes grupos de sensores móviles y vehículos autónomos está convirtiéndose en una realidad gracias a los avances tecnológicos en redes y en miniaturización de sistemas electro-mecánicos. Estas y otras tecnologías emergentes han ido llenando rápidamente el vacío que existía para poder fabricar grandes cantidades de pequeños dispositivos, autónomos, baratos y energéticamente eficientes, equipados con las capacidades de comunicarse entre sí, medir y estimar fenómenos físicos en su entorno, procesar los datos que adquieren y decidir su movimiento con la información procesada.

La complejidad de tales sistemas móviles presenta nuevos desafíos en la frontera entre la *comunicación*, la *computación* y el *control*. Aunque la tecnología suministra los componentes físicos de tales redes de sensores, y aunque estos sistemas tendrían un impacto positivo en numerosas aplicaciones, los beneficios potenciales de tales sistemas no se han materializado todavía.

A día de hoy, la limitación fundamental es la falta de entendimiento para coordinar los dispositivos individuales en un todo coherente.

Dicho de otra manera, no existen metodologías sistemáticas para controlar sistemas distribuidos y fiables a gran escala, tales como una red de sensores desarrollando tareas complejas.

Como consecuencia de estas limitaciones tecnológicas, existe una fuerte necesidad por expandir el conjunto de paradigmas y herramientas actualmente disponibles para el diseño de redes escalables de sensores y la verificación formal de su rendimiento. *Y en esta tarea, varias áreas de las Matemáticas tienen mucho que aportar.* Algunos desafíos a la hora de diseñar algoritmos de coordinación para realizar tareas cooperativas son los siguientes: los algoritmos deben ser:

Adaptativos: los algoritmos de coordinación deben dotar a la red de la habilidad para responder y adaptarse a diversas tareas, ambientes dinámicos y topologías variables (debido a llegadas, salidas o fallos de agentes individuales). La red ha de ser capaz de reaccionar a gran variedad de condiciones cambiantes.

Distribuidos: los algoritmos de coordinación deben depender de la menor cantidad posible de información global, y en cambio, es deseable que el comportamiento de cada agente dependa sólo de información local. Los algoritmos distribuidos tienen la ventaja de ser escalables (es decir, funcionan en redes formadas por un número arbitrario de agentes y satisfacen las limitaciones de ancho de banda en las comunicaciones), y robustos ante los fallos en la comunicación o el mal funcionamiento de los agentes individuales. A medida que la red evoluciona, las relaciones de vecindad cambian. Por tanto, los algoritmos distribuidos para grafos de comunicación fijos (habituales, por ejemplo, en computación paralela y en ciertos tipos de aplicaciones en internet) no son aplicables a este tipo de situaciones dinámicas.

Asíncronos: los algoritmos de coordinación deben ser implementables de manera asíncrona. Esto significa que los agentes de la red pueden estar evolucionando con distintas velocidades, y con diferentes capacidades de computación y comunicación. Además, no existe un *reloj global* para toda la red, sino que cada agente evoluciona de acuerdo a su propia noción de tiempo. Casi en cualquier escenario real, la información se propaga entre los agentes con cierto retraso, y además éste nunca es el mismo. Estas características hacen a la vez crucial y complicado establecer garantías formales acerca del comportamiento de la red.

Corrección asintótica verificable: mientras se completa la tarea asignada, los algoritmos de coordinación han de garantizar una convergencia segura frente a posibles fallos en el sistema (colisiones, atascos de comunicación, fallos de los dispositivos en cada robot, etc.). La importancia de la verificación formal crece con la dimensión y la complejidad de la red de robots considerada.

Motivados por estos problemas (ver también [3, 4]), en la actualidad, se está desarrollando un gran esfuerzo interdisciplinar en esta dirección en el que están implicados investigadores de diferentes departamentos. El trabajo [6] recoge un tratamiento más técnico del progreso hecho hasta la fecha, y de algunas perspectivas futuras.

Conclusiones

En términos generales, los retos fundamentales a los que nos enfrentamos para hacer posible este nuevo horizonte de aplicaciones tecnológicas, son los de 1) optimizar y desarrollar cada uno de los aspectos que definen a un sistema autónomo (movilidad, comunicación, computación...) para hacerlos más eficaces y 2) saber integrar de forma óptima las múltiples componentes de distinta índole que forman parte de cada sistema y que caracterizan la relación de sistema a sistema.

En particular, estos objetivos no podrán realizarse sin el correspondiente desarrollo de las herramientas matemáticas que nos permitan analizar todas las facetas involucradas. A grandes rasgos se hace necesario:

- El desarrollo de herramientas matemáticas para la modelización y el análisis de los sistemas autónomos y robóticos. Esto es fundamental para generar sistemas con mayores niveles de autonomía, más versátiles y precisos.
- El análisis pormenorizado de los límites fundamentales que la tecnología impone y, teniendo esto en cuenta, el diseño de mecanismos de control y herramientas matemáticas que sepan explotar las características particulares de dicha tecnología.
- El análisis matemático se puede y se debe complementar con técnicas de simulación avanzadas que permitan un mayor entendimiento de cómo integrar los distintos aspectos de comunicación, computación, sensores y actuadores de estos sistemas complejos.



El estudio sistemático de la estructura matemática de los problemas descritos en este informe (control de movimientos y coordinación de sistemas autónomos) promete ser de una gran utilidad en numerosos dominios científicos y comerciales. En nuestra opinión, esta afirmación está ampliamente refrendada por la realidad investigadora actual en estos campos, y, al menos en los EEUU, existe un consenso generalizado al respecto (ver, por ejemplo, [4, 7]).

Bibliografía

- [1] Bloch, A.M.; *Nonholonomic Mechanics and Control*. Number 24 in Interdisciplinary Texts in Mathematics. Springer Verlag, (2003).
- [2] Bullo, F.; Lewis, A.D.; *Geometric Control of Mechanical Systems*, volume 49 of Texts in Applied Mathematics. Springer Verlag, New York, (2004).
- [3] Uny Cao, Y.; Fukunaga, A.S.; Kahng, A.; *Cooperative mobile robotics: Antecedents and directions*. *Autonomous Robots*, 4(1): (7-27), (1997).
- [4] Committee on Networked Systems of Embedded Computers. *Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers*. National Academy Press, (2001).
- [5] Martínez, S.; *Geometric Methods in Nonlinear Control Theory with Applications to Dynamic Robotic Systems*. PhD thesis, University Carlos III, Madrid, Spain, (2002).
- [6] Martínez, S.; Cortés, J.; Bullo, F.; *Motion coordination with distributed information*. *IEEE Control Systems Magazine*, November 2005. Vol 27, (2007), nº4.
- [7] Murray, R.M.; editor. *Control in an Information Rich World*. SIAM, Philadelphia, (2003). Report of the Panel on Future Directions in Control, Dynamics and Systems.
- [8] NASA rover website. <http://marsrovers.jpl.nasa.gov>.

Capítulo XIII

MATEMÁTICAS Y CRIPTOGRAFÍA

POR ALEJANDRO MELLE HERNÁNDEZ
UNIVERSIDAD COMPLUTENSE DE MADRID
Departamento de Álgebra,
Facultad de Matemáticas
CP 28040
Correo electrónico: amelle@mat.ucm.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

We live in an information-based society. The techniques to keep secret the information are the field of study of cryptography. The cryptography are the mathematical tools related to confidentiality, data integrity, entity authentication and data origin authentication.

There are two forms of encryption: Symmetric Key Encryption which is fast but both parties need to know a shared secret key and Public Key Encryption which is slow but only one party needs to keep a key private.

The first and most popular public key exchange algorithm is RSA. The security of RSA is based on the intractability of the integer factorization problem. Diffie-Hellman key exchange relies on difficulty of computing discrete logarithms. There are a few other key exchange schemes that are used in practice, for example, the Digital Signature Algorithm (DSA) and the Elliptic Curve Digital Signature Algorithm (ECDSA). The security of those schemes is based on the discrete logarithm problem in the multiplicative group of a prime field or in the group of points of an elliptic curve over a finite field.

One of the fundamental tools used in information security is the signature. From practical view point the main ingredient of the signature in cryptology are the hash functions. The main role of a cryptographic hash function is in the provision of digital signatures. Since hash functions are generally faster than digital signature algorithms, it is typical to compute the digital signature to some document by computing the signature on the document's hash value, which is small compared to the document itself. It was a big surprise that a vulnerability of standard hash functions was announced in Feb. 2005. The attack primarily affects some digital signature applications, including timestamping and certificate signing operations, where one party prepares a message for the generation of a digital signature by a second party, and third parties then verify the signature.

When quantum computers reach approximately 30 to 40 q-bits they will start to have the speed (parallelism) needed to attack the methods society uses to protect data and processes, including encryption, digital signatures, random number generators, key transmission, and other security algorithms. In particular all the standards used nowadays will become obsolete and one should be prepared to have ready several other cryptosystems that are conjectured to resist quantum computers, how efficient are they, in theory and in practice?

To know something else: <http://www.cryptomathic.com/labs/ellipticcurves.html>

Second Cryptography Hash Workshop: <http://www.csrc.nist.gov/pki/HashWorkshop/index.html>

Mathematics and Internet security: <http://www.mathaware.org/mam/06/>

Introducción

A nadie parece extrañar la frase "... hoy vivimos en la sociedad de la información..", cada día cualquiera de nosotros almacena, procesa e incluso transmite información en forma digital. Sin duda, desde los años ochenta, el uso de los ordenadores, de las redes locales, de Internet, del comercio electrónico, de la telefonía inalámbrica, del *wireless*... es parte de nuestra vida diaria.

La contrapartida al uso de toda esa información sobre redes abiertas son los problemas de *privacidad* y *seguridad*. Se deben pues buscar mecanismos tanto legales como científico-técnicos para garantizar ambas. Un ejemplo de plena actualidad es la puesta en marcha, en un futuro cercano, del Documento Nacional de Identidad electrónico.

La *Criptografía* es la parte científico-técnica que trata de garantizar los fundamentos de la seguridad en la transmisión de información. Los objetivos básicos que persigue la Criptografía son:

- *Confidencialidad*, se debe mantener la información en secreto para toda entidad excepto para los que tengan autorizado el acceso a ella.
- *Integridad de los datos*, se debe garantizar que el mensaje no ha sido modificado durante la transmisión.
- *Autenticidad*, el receptor del mensaje debe ser capaz de verificar quién es el emisor y, por otro lado, ambos deben ser capaces de identificarse mutuamente.
- *No repudio*, el emisor no debe ser capaz de negar que ha transmitido el mensaje.

Para mantener la confidencialidad de la información, el *texto en claro* (que puede ser texto, datos, ...) debe ser codificado antes de ser transmitido, obteniéndose el *texto cifrado*. Una vez enviado el texto cifrado el receptor debe *descifrar* el mensaje. La idea básica de *cifrar* el texto es garantizar que no haya una tercera parte no autorizada que sea capaz de acceder a la información sin conocer la clave.

El *Criptoanálisis* es la ciencia que estudia los ataques de agentes extraños contra los criptosistemas. El objetivo de los ataques es variable y depende de las habilidades del atacante. Puede ser que se pretenda recuperar texto en claro del texto cifrado, o bien que se pretenda substituir parte del texto en claro original, o bien falsificar firmas digitales... Uno de los principios del Criptoanálisis es el *principio de Kerckhoff* que establece que el criptosistema debe ser de conocimiento público, incluyendo los algoritmos y su implementación, y que lo único que se debe mantener en secreto es la clave con la que se cifra. Lo que distingue la criptografía moderna de la clásica es que en la criptografía moderna se sigue a rajatabla este principio.

En esta nota quiero presentar y discutir las componentes primitivas y los criptosistemas más usados; *primitivos* se refiere a que son las piezas básicas de otros criptosistemas.



Criptosistemas de clave simétrica

Un método que se viene usando desde la antigüedad para garantizar la confidencialidad es que tanto el emisor como el receptor se pongan de acuerdo previamente en una *clave* k que les permita tanto cifrar como descifrar sus mensajes, manteniendo por supuesto la clave en secreto. Este tipo de sistemas criptográficos basado en compartir una clave secreta se llaman sistemas *simétricos*. Una vez fijada la clave k se tiene un algoritmo de cifrado e_k y un algoritmo de descifrado d_k con la condición evidente de que para cualquier texto en claro m , se verifique que $d_k(e_k(m)) = m$. Obsérvese que para ponerse de acuerdo en la clave se debe usar un canal seguro de comunicación.

Se usan esencialmente dos tipos básicos de cifrado simétrico: el cifrado en *flujo* y el cifrado en *bloque*. Una vez generada la clave se transmite a los usuarios mediante un medio seguro, por ejemplo usando los criptosistemas de clave pública que se verán después. En el cifrado en flujo se cifra cada carácter combinándolo con una sucesión obtenida mediante un proceso *pseudoaleatorio*.

En el cifrado en bloque se trabaja con bloques de texto en claro y texto cifrado de longitud dada n . En general se usan tamaños de clave $n = 64, 128$ o 256 bits. En el año 2000, el cifrado simétrico en bloques Rijndael fue seleccionado (en un concurso abierto sin precedentes) como el AES (*Advanced Encryption Standard*) que substituyó, después de veinte años, al DES (*Data Encryption Standard*) propuesto en su día por el NIST (*National Institute of Standards and Technology*) de EEUU. El Rijndael cifra el texto en claro en bloques de tamaño 128 bits y la longitud de su clave puede variar entre 128 y 256 bits. El Rijndael ha sido diseñado para resistir el criptoanálisis clásico (criptoanálisis lineal, diferencial, ...). Sin embargo el Rijndael posee una estructura muy algebraica y ataques usando estructuras algebraicas (bases de Gröbner, aproximaciones cuadráticas...), no han sido estudiados con todo detalle (aunque de momento los intentos de romperlo han fracasado).

Criptosistemas de clave pública

El principal problema de los criptosistemas de clave simétrica es la necesidad de tener un canal de comunicación seguro para intercambiar la clave usada. En 1976 Diffie-Hellman [3] introdujeron el concepto de *criptografía de clave pública*. Cada entidad tiene su clave $k = (pk, sk)$ que consiste en una clave privada sk y otra pública pk . La clave privada se debe mantener en secreto y la pública se debe hacer accesible a cualquier persona o entidad con la que se quiera establecer comunicación. La potencia de este sistema de claves reside en que cualquier cosa que se cifre con una de ellas se descifra con la otra. Si **A** quiere transmitir información m , de forma confidencial, a **B** lo que hace es usar la clave pública pk_B de **B** para cifrar la información $pk_B(m)$ y transmitir el texto cifrado a **B**. Para descifrar el mensaje, **B** usa su clave privada sk_B : $sk_B(pk_B(m)) = m$. No es necesario por tanto ningún contacto a priori entre **A** y **B**.

Si por otra parte **B** quiere estar seguro de que es **A** quien le transmite la información lo que se debe hacer es que **A** cifre el texto en claro con su clave privada $sk_A(m)$ el receptor **B** debe obtener la clave pública de **A** (de cualquier servidor en el que esté publicada) y descifra el mensaje $pk_A(sk_A(m)) = m$. De esta forma se garantiza la autenticidad del emisor.

El sistema de clave pública más usado es el RSA [13] que está basado en la dificultad computacional de factorizar un número n , suficientemente grande, en sus factores primos. En el RSA se elige $n = pq$ siendo p y q primos muy grandes. Un primer ataque al RSA consiste en factorizar el entero n . Los algoritmos de factorización de enteros y los tests de primalidad (basados en teoría de números, en geometría algebraica,...) son, y han sido, aplicaciones de resultados matemáticos teóricos. Desde el punto de vista de la implementación, los algoritmos más eficientes de factorización son la criba cuadrática, las cribas de cuerpos de números (basado en cálculos en anillos de enteros algebraicos) y el algoritmo basado en curvas elípticas introducido por Lenstra. El último entero factorizado usado en RSA es de 576 bits. Este es el motivo de que en la actualidad se recomiende usar p y q de tamaño de 576 bits y n de 1024 bits como módulo (incluso algunas agencias estatales proponen que n tenga 2048 bits). Desde el punto de vista teórico, cabe resaltar, que el algoritmo AKS permite saber si un entero es primo usando un algoritmo de complejidad polinomial. Sin embargo, por lo general, los ataques más exitosos se deben al mal uso de los parámetros del RSA [8]. Para su implementación práctica matemáticos e ingenieros deben trabajar juntos en resolver el problema de encontrar la relación entre efectividad y seguridad.

FIGURA 6

Esquema de confidencialidad



Otro de los problemas computacionalmente difíciles que se usan en criptosistemas de clave pública es el *problema del logaritmo discreto*. En general, si (G, \cdot) es un grupo cíclico finito de orden n generado por a , i.e. $G = \{a, a^2, \dots, a^{n-1}, a^n = 1\}$, y $x \in G$, el problema consiste en encontrar k tal que $x = a^k$. Para que un grupo (G, \cdot) se pueda usar en criptografía debe verificar las siguientes propiedades: I) los elementos del grupo deben ser expresados de un modo compacto, II) la operación del grupo debe ser computada eficientemente, III) el problema del logaritmo discreto debe ser intratable en G . En el año 1976, Diffie-Hellman establecieron un protocolo de intercambio de claves basado en el problema del logaritmo discreto sobre el grupo de unidades del cuerpo finito \mathbb{F}_p . En la actualidad y conociendo los resultados matemáticos actuales, los grupos que se usan en la práctica son los siguientes:

- 1) $(G, \cdot) = (\mathbb{F}_q^*, \cdot)$, donde \mathbb{F}_q^* es el grupo de unidades del cuerpo finito \mathbb{F}_q con $q = p$ con p primo o $q = 2^n$.
- 2) $(G, \cdot) = (E(\mathbb{F}_q), +)$, donde $E(\mathbb{F}_q)$ es el conjunto de puntos racionales de una curva elíptica sobre el cuerpo finito \mathbb{F}_q , con $q = p$ con p primo o $q = 2^n$.

El algoritmo más eficiente para resolver el problema del logaritmo discreto en los grupos de tipo 1) es el algoritmo del cálculo del índice, que es sub-exponencial. En general para los de tipo 2) no existen tales algoritmos y la complejidad de los algoritmos conocidos es exponencial. Esto hace que los sistemas basados en curvas elípticas proporcionen la misma seguridad con tamaños de clave mucho menores; por ejemplo, una clave de curva elíptica de 163 bits proporciona la misma seguridad que otra de 1024 bits de RSA o de tipo 1). Tamaños de clave menores implican menor tiempo de proceso y menor espacio de almacenamiento, factores fundamentales para la aplicación práctica de la criptografía (por ejemplo en wireless, tarjetas electrónicas, ...) [7].

Otros grupos algebraicos (variedades abelianas sobre cuerpos finitos) donde el problema del logaritmo discreto es difícil y que se pueden usar en criptografía son los grupos de las jacobianas de curvas hiper-elípticas de género 2. De hecho se prueba que los criptosistemas son igual de eficientes que los de las curvas elípticas. El estudio, como herramientas de criptoanálisis, de formas bilineales no-degeneradas sobre curvas elípticas ha llevado a diseñar criptosistemas de clave pública basados en el problema de Diffie-Hellman usando formas bilineales no-degeneradas.

FIGURA 7



Esquema de firma digital



Junto a la confidencialidad y la autenticidad, la criptografía debe garantizar la integridad y el no-repudio de los datos transmitidos. Tanto en los criptosistemas de clave simétrica como de clave pública hay métodos para garantizar la integridad (por ejemplo usando funciones *hash*). Los sistemas de clave pública, a través de los protocolos de *firma digital*, son capaces de garantizar la

autenticidad, la integridad y el no-repudio. Este protocolo es de gran importancia en el mundo actual, tiene ámbito jurídico y legal y se usa para firmar contratos, realizar compras, hacer gestiones con la administración pública, entre otras muchas cosas. Como se puede ver en el esquema anterior, para realizar una firma de un texto en claro m , se procede de la siguiente manera: el firmante **A** hace un *hash* o resumen $h(m)$ de m (integridad), a continuación cifra el resultado con su clave privada (autenticidad, no-repudio) $sk_A(h(m))$, lo añade al final del mensaje y transmite $(m, sk_A(h(m)))$. Cuando **B** lo recibe, descifra $pk_A(sk_A(h(m)))$, calcula $h(m)$ y comprueba que ambos coinciden; a esta fase se le denomina verificación de la firma. La seguridad de los algoritmos de firma digital es crucial para asegurar el futuro de las transacciones electrónicas.

Existen al menos dos problemas que comprometen seriamente tanto a los algoritmos de cifrado como al protocolo de firma digital y al nuevo DNI electrónico, mencionado al principio:

- *Análisis criptográfico de las funciones hash*: En la práctica los algoritmos de firma primero comprimen el texto en claro y luego cifran el texto comprimido. La finalidad de las funciones hash es limitar el tamaño de la firma independientemente de la longitud del texto en claro. Se pueden hacer protocolos probabilísticos de firma digital sin más que modificar la función hash usando sucesiones pseudoaleatorias de bits. Una función *hash* $h : \{0, 1\}^* \rightarrow \{0, 1\}^n$ es una función que comprime una cadena arbitraria de bits a una cadena de bits de longitud fija n . Las funciones *hash* no son inyectivas y se llaman colisiones a los $m, m' \in \{0, 1\}^*$ tales que $h(m) = h(m')$. Para que una función hash se pueda usar en criptografía debe verificar: 1) ha de ser fácil calcular $h(m)$ pero computacionalmente difícil calcular las preimágenes $h^{-1}(y)$, 2) debe ser computacionalmente difícil que fijado m se encuentre m' tal que $h(m) = h(m')$ y 3) debe ser computacionalmente difícil encontrar colisiones. Las colisiones se pueden usar para alterar contenidos de documentos firmados. En el año 2005 se encontraron colisiones a la función hash SHA-1. Esta función fue desarrollada por la NSA de EEUU y en la actualidad se usa como standard tanto a nivel oficial (gobiernos,...) como comercial (Microsoft, Sun, IBM...), [9].

El RSA fue el primer criptosistema de clave pública con protocolo de firma digital. El Gamal propuso un protocolo de firma digital no determinista (basado en el logaritmo discreto) que fue modificado ligeramente por el NIST dando lugar al DSA (Digital Standard Algorithm). Desde el año 2000 se está usando un algoritmo de firma digital ECDSA basado en curvas elípticas, que se está aplicando en el ámbito comercial y en las gestiones con las administraciones públicas. Todos estos protocolos usan el SHA-1 como función hash.

- *Criptografía cuántica y post-cuántica*: En el año 1994 Shor [14] encontró un algoritmo de factorización que en un "ordenador cuántico" tiene complejidad polinomial. Por tanto, si se llega a construir un ordenador cuántico de tamaño industrial todos los criptosistemas y protocolos basados en el problema de la factorización o en el problema logaritmo discreto serían inseguros. En la actualidad no se sabe si se llegará a construir un ordenador cuántico y cuándo, pero dado el interés del tema hay muchos grupos de investigación que incluyen matemáticos trabajando en ello. Hay que decir a este respecto que ya se usa criptografía basada en fenómenos cuánticos, incluso a nivel comercial, para la transmisión de información aunque no existan todavía ordenadores cuánticos.



Por todo lo anterior se deben proponer otros sistemas de clave pública no basados en estos problemas. Entre ellos cabe destacar: criptosistemas basados en problemas relacionados con retículos (por ejemplo NTRU, [6]), criptosistemas basados en teoría de códigos, criptosistemas basados en el problema de la conjugación en el grupo de trenzas [1], criptosistemas basados sistemas de ecuaciones polinomiales multivariables de grado dos sobre cuerpos finitos que usan algebra conmutativa y geometría algebraica, (por ejemplo HFE [12]...). Para que cualquiera de estos protocolos puedan ser admitidos comercialmente es necesario un extenso trabajo de análisis y criptoanálisis en el que las matemáticas juegan un papel fundamental.

El autor quiere agradecer a Ignacio Luengo Velasco y a Julia Sánchez Meneses por sus discusiones muy interesantes y esclarecedoras sobre el tema.

Bibliografía

- [1] Anshel, I.; Anshel, M.; Fisher, B.; Goldfeld, D.; *New key agreement protocols in braid group cryptography*. Topics in cryptology-CT-RSA 2001 (San Francisco, CA), (13-27), Lecture Notes in Comput. Sci., 2020, Springer, Berlin, (2001).
- [2] Churchhouse, R.F.; *Codes and ciphers : Julius Caesar, the enigma, and the Internet Cambridge*: Cambridge University Press, (2002).
- [3] Diffie, W.; Hellman, M.; *New directions in cryptography*, IEEE Transactions on Information Theory 22, (1976), (644-654).
- [4] ElGamal, T.; *A public key cryptosystem and a signature scheme based on discrete logarithms*, IEEE Transactions on Information Theory 31 (1985), (469-472).
- [5] Fúster Sabater, A.; de la Guía, D.; Hernández, L.; Montoya, F.; Muñoz-Masqué, J.; *Técnicas criptográficas de protección de datos*, Ed. Ra-ma, cop., (2004), Madrid.
- [6] Hoffstein, J.; Pipher, J.; Silverman, J.H.; *NSS: an NTRU lattice-based signature scheme*. Advances in cryptology-EUROCRYPT 2001 (Innsbruck), Lecture Notes in Comput. Sci., 2045, Springer, Berlin, (2001), (211-228).
- [7] Koblitz, N.; *Algebraic aspects of cryptography* Berlin, Springer, (1999) (2.ª ed.)
- [8] Koblitz, N.; Menezes, A.; *A survey of public-key cryptosystems*, SIAM Rev. 46, no. 4, (2004), (599-634).
- [9] Landau, S.; *Find me a hash*. Notices Amer. Math. Soc. 53, no. 3, (2006), (330-332).
- [10] Menezes, A.; van Oorschot, P.C.; Vanstone, S.A.; *Handbook of applied cryptography*, With a foreword by Ronald L. Rivest. CRC Press Series on Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, (1997).
- [11] Mollin, R.; *An introduction to cryptography* Boca Raton : Chapman-Hall/CRC, cop., (2001).
- [12] Patarin, J.; *Asymmetric cryptography with a hidden monomial and a candidate algorithm for $_64$ bits asymmetric signatures*. Advances in cryptology-CRYPTO '96 (Santa Barbara, CA), Lecture Notes in Comput. Sci., 1109, Springer, Berlin, (1996), (45-60).
- [13] Rivest, R.L.; Shamir, A.; Adleman, L.; *A method for obtaining digital signature and public key cryptosystem*, Communication of the ACM 21, (1978), (120-126).
- [14] Shor, P.W.; *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer* SIAM J. Comput. 26, no. 5, (1997), (1484-1509).
- [15] Stinson, D.; *Cryptography : theory and practice* Boca Raton, Chapman-Hall/CRC Press, cop., (2002).

Capítulo XIV

NECESIDADES FUTURAS EN INVESTIGACIÓN AERODINÁMICA

POR FERNANDO MONGE
INSTITUTO NACIONAL DE TÉCNICA AEROESPACIAL
Área de Dinámica de Fluidos,
Departamento de Aerodinámica y Propulsión
28850 Torrejón de Ardoz, Madrid
Correo electrónico: mongef@inta.es
<http://www.inta.es>

FRANCISCO PALACIOS
INSTITUTO NACIONAL DE TÉCNICA AEROESPACIAL
Área de Dinámica de Fluidos,
Departamento de Aerodinámica y Propulsión
28850 Torrejón de Ardoz, Madrid
Correo electrónico: fpalacios@inta.es
<http://www.inta.es>

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

The goal of this paper is to summarize the current needs in aerodynamic research for aeronautics. Most of these needs require a careful development of different mathematical tools as aerodynamics is a field strongly dependent on a solid mathematical ground. The text is divided into the following sections: first, a general introduction, followed by a study of the different contributions of CFD to the design of aerodynamics in aircrafts. There will also be an analysis of the current European Framework as well as of the different priorities set by the Spanish Framework for aerodynamics research. Then, the priorities in aerodynamics research of the aeronautical industry within the European and Spanish framework will be stated. Finally, several examples of on-going projects or currently under development will be presented along with a summary of recommended future work.

Resumen

El presente artículo está encaminado a sintetizar las actuales necesidades de investigación en aerodinámica aeronáutica. Gran parte de estas necesidades requieren un meticuloso trabajo matemático pues no en vano la aeronáutica es una disciplina sustentada en una compleja base matemática. El presente texto se articula sobre los siguiente puntos: primero se realizará una introducción general, posteriormente se desarrollaran las contribuciones de la CFD¹ al diseño aerodinámico de aeronaves, una vez finalizada esta parte más técnica se comentarán de manera resumida el actual Marco Europeo y cuáles son las prioridades que establece el Marco Español de trabajo en aerodinámica. Por último, se indicarán las prioridades de la industria aeronáutica en aerodinámica que se derivan de los marcos Europeos y Español. Finalmente, se presentarán a modo de ejemplo algunos proyectos en desarrollo o en fase de desarrollo, finalizando con un resumen de recomendaciones de trabajo futuro.

Introducción

Una gran parte de la aeronáutica moderna se desarrolla en torno a las fundamentales ecuaciones de Navier-Stokes, que fueron formuladas en el siglo XIX, y que hasta el desarrollo de la informática y de la computación no pudieron ser abordadas de manera útil para la Industria. Sin embargo, pese a los grandísimos avances que se han producido en el último tercio del siglo XX y en paralelo con el desarrollo de la computación, todavía queda un camino complejo y largo por recorrer.

Muy brevemente, las ecuaciones de Navier-Stokes son las que modelizan el comportamiento de los fluidos mediante el planteamiento de sistemas de conservación de masa, cantidad de movimiento y energía, además de las ecuaciones constitutivas de los gases. El interés de estas ecua-

¹ CFD: Computational Fluid Dynamics

ciones reside en el hecho de que, una vez resueltas, es posible obtener el campo fluido alrededor de un objeto (típicamente un avión) y partiendo de ese campo fluido, se calculan valores integrales como son la fuerza o momentos que ejerce el aire sobre el avión.

Lamentablemente, por lo general, debido a su complicación y exigencia de cálculo, no es posible abordar de manera directa este sistema de ecuaciones en derivadas parciales, lo que limita en gran medida la aplicación real. Es por ello que en la ingeniería se buscan atajos y simplificaciones que permitan el obtener resultados con un valor ingenieril bien definido. A raíz de esto se desarrollan los modelos de cálculo aerodinámico en función de los diferentes niveles de aproximación que el ingeniero desee emplear.

Si se abordan directamente las ecuaciones de Navier-Stokes empleamos métodos Directos (DNS) o parcialmente directos (LES y DES). Si por el contrario evitamos el atacar directamente la resolución de la turbulencia nos encontraremos con métodos que se caracterizan por modelizarla. Dentro de esta familia encontramos los métodos de ecuaciones promediadas de Navier-Stokes (RANS). Si además, como es lo habitual, estamos trabajando con altos números de Reynolds y corriente adherida, será posible desacoplar los efectos viscosos, resolviendo por una parte las ecuaciones de capa límite y por otra el modelo no viscoso (ecuaciones de Euler).

Si, además, el número de Mach es suficientemente pequeño, entonces existe potencial de velocidad, y el sistema de ecuaciones de Euler se reduce a una sola ecuación (potencial). Por último, en régimen incompresible dicha ecuación potencial puede linealizarse (ecuación de Laplace) y resolverse mediante métodos de superposición de singularidades (métodos de paneles).

Por otro lado, acoplado al problema de la resolución numérica de las ecuaciones diferenciales correspondientes, el cómo se discretiza el espacio (mallado) adquiere una importancia fundamental, pues en la práctica es lo que necesita más tiempo de ingeniería. Intentar conseguir una automatización de dicho mallado es, por tanto, esencial.

Contribuciones de la CFD al diseño aerodinámico de aeronaves

La CFD ha sido aplicada con éxito a multitud de actividades dentro del campo aeronáutico. En algunas de ellas resulta actualmente imprescindible y su aplicación ha sido masiva. Por otro lado, en otras situaciones su aplicación ha sido algo más discreta pero relevante. En cualquier caso, a medida que los métodos de diseño y la computación avanzan, en igual medida se desarrollan nuevas y más trascendentes aplicaciones de la CFD. En la actualidad, las facetas del diseño aerodinámico donde la CFD ha sido más relevante son:

- Diseño de la forma en planta del ala.
- Diseño del ala para la velocidad de crucero.
- Diseño de dispositivos de punta de ala.
- Diseño de las superficies de cola.



- Diseño del fuselaje.
- Corrección de resultados de Ensayos de Túnel.
- Integración célula - grupo moto-propulsor.
- Diseño de tomas de motor.

De manera moderada, el CFD ha sido aplicado al:

- Diseño de dispositivos hiper-sustentadores.
- Formación de hielo.
- Estudio del flujo de aire en el interior de la cabina.

Por último, de manera ligera y, por lo tanto, con mayores necesidades de desarrollo:

- Generadores de torbellinos
- Límites de bataneo
- Aeroelasticidad
- Diseño para estabilidad y control
- Ruido en cabina y ruido exterior

Todas estas incorporaciones de la CFD al diseño aerodinámico tienen como último fin la mejora de los actuales aviones. A modo ilustrativo son muy destacables las grandes mejoras en eficiencia debidas a cada una de las principales disciplinas técnicas durante el periodo 1980 - 2005:

TABLA

Aerodinámica	25%
Materiales	20%
Motores	40%
Sistemas	15%

A su vez, y en gran medida gracias a los desarrollos matemáticos efectuados, el tiempo empleado en los túneles de Ensayos Aerodinámicos, y en consecuencia el coste de diseño, ha sido reducido de manera muy significativa:

TABLA

1980	100%
1990	75%
2005	50%

FIGURA 8

Ensayos en túnel aerodinámico con pinturas sensibles a la presión (DLR)



FIGURA 9

Distribución de presión sobre el ala LANN: condiciones no estacionarias



Marco Europeo

El marco europeo para principios del siglo XXI viene marcado por las metas fijadas en las reuniones Aeronautics 2001 que establecieron de una manera realista las previsiones y los objetivos a cumplir en los próximos 20 años, hasta el 2020: estas metas, acordadas tras la creación de ACARE (Consejo Asesor de Investigación Aeronáutica en Europa) se plasmaron en la SRA (Agenda Estratégica de Investigación), actualizada cada 2 años, proporcionando a su vez un programa de trabajo que se utiliza como referencia en la elaboración de objetivos del Programa Marco.

Las recomendaciones del “Aeronautics - 2020” en aspectos aerodinámicos son las siguientes:

- Reducción a la mitad del tiempo de diseño, lo que justifica la necesidad del desarrollo continuo de herramientas CFD.
- Disminución del 50% en las emisiones de CO₂ (la misma disminución en consumo de combustible) y del 80% en las de NO_x, lo que impone el diseño de aeronaves más eficientes en cruce-



ro y, por lo tanto, la disminución de la resistencia aerodinámica (D) y el aumento de la relación sustentación / resistencia (L/D).

- Reducción a la mitad del ruido percibido: diseño de mejores dispositivos hipersustentadores durante las fases de despegue y ascenso (máxima potencia) y de aproximación y aterrizaje (máximas deflexiones).
- Capacidad de volar en cualquier condición meteorológica: efectos de ráfagas, lluvia intensa, formación de hielo.

FIGURA 10

Malla híbrida sobre la configuración ala-fuselaje



FIGURA 11

Dispositivos hiper-sustentadores: problemas simultáneos



Todas estas prioridades aerodinámicas se articularon alrededor de un programa de trabajo dividido en una serie de áreas temáticas.

Área 1: Refuerzo de la competitividad.

- Diseño integrado y desarrollo del producto.
- Aerodinámica. Herramientas analíticas y experimentales avanzadas. Herramientas de diseño para disminuir los DOC (Costes Directos de Operación). Disminución de resistencia aerodinámica: conceptos, tecnologías y sistemas. Flujo de aire en cabina. Nuevas configuraciones de aeronaves y tecnologías innovadoras.

Área 2: Impacto medioambiental (emisiones de gases y ruido).

Área 3: Mejora de la seguridad. Prevención de accidentes. Desarrollo de mecanismos de protección ante ráfagas, lluvia intensa y formación de hielo.

Área 4: Aumento de la capacidad operativa

- Disminución de los estándares de separación en despegue y aterrizaje. Caracterización y control de los torbellinos de estela.

Marco Español

Para abordar el marco español en aeronáutica es muy conveniente realizar un pequeño recorrido histórico sobre los últimos 15 años de la aeronáutica en España con objeto de terminar en el actual subprograma nacional de transporte aéreo:

- 1992, Ayudas gubernamentales a la antigua CASA para participar en programas de desarrollo de AIRBUS.
- 1993 - 1998, Primer Programa Aeronáutico Nacional.
- 1998 - 2003, Segundo Programa Aeronáutico Nacional.
- 2004 - 2007, Integración del Programa Aeronáutico en el PROFIT: Subprograma Nacional de Transporte Aéreo.

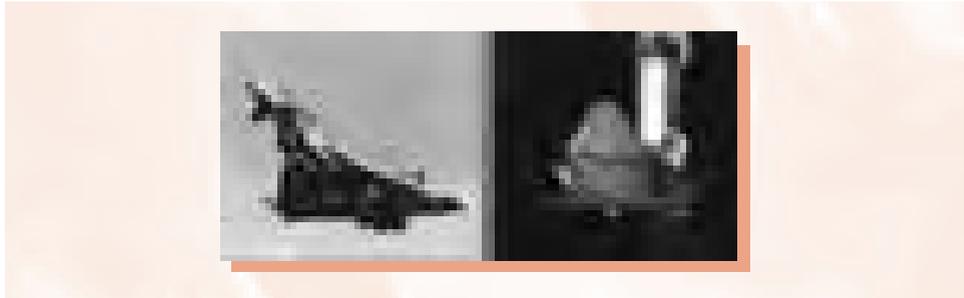
FIGURA 12

Cálculo de dispositivos hiper-sustentadores



FIGURA 13

Ejemplos de configuraciones de tomas de motor



De esta manera se ha llegado al actual Subprograma Nacional de Transporte Aéreo, que es el que financia de manera gubernamental los desarrollos aeronáuticos en España. Este Subprograma tiene como objetivo el contribuir al incremento del conocimiento científico-tecnológico de las Empresas Aeronáuticas Españolas, buscando especialmente el afianzar la especialización tecnológica de la industria aeronáutica española, facilitar la participación de las empresas españolas en los programas y consorcios aeronáuticos internacionales (Programa Marco de la UE) y finalmente el difundir el conocimiento científico-tecnológico.

Las prioridades temáticas que se establecen en este subprograma Nacional, pasan por el estudios de I+D tecnológico de configuraciones de aeronaves y las disciplinas y tecnologías específicas que contribuyen a la definición detallada de la aeronave, incluyendo específicamente a la aerodinámica entre ellas, en cuyo desarrollo las matemáticas tienen un papel cada vez más importante.

Prioridades de la Industria Aeronáutica en Aerodinámica

En base a la descripción del marco y Europeo y Español, es posible la descripción de las prioridades de la industria aeronáutica:

- Control de Flujo: control y disminución de la turbulencia y del desprendimiento de la corriente; diseño en presencia de control de flujo.
- Control Activo de Cargas.
- Herramientas de Simulación Avanzadas: mejora de la velocidad de cálculo, eficiencia y precisión (especialmente en aerodinámica no estacionaria y en aero-elasticidad); gestión de grandes volúmenes de datos de distintas fuentes; desarrollo de métodos y herramientas de diseño óptimo para toda la envolvente de vuelo; tratamiento de las incertidumbres en sistemas no-lineales multi-disciplinares; fusión de datos numéricos y experimentales.

FIGURA 14

Configuraciones esbeltas cálculo RANS a Mach 2 Alpha = 10°



FIGURA 15

Avión supers. A baja velocidad: distribución de presión ($\alpha = 10^\circ$, $Re = 25 \times 10^6$)



- Diseño y Optimización: previsión del efecto de un aumento sustancial en la potencia de cálculo (10^3 - 10^6 veces) en el diseño aeronáutico.
- Apoyo a la Certificación de las aeronaves: modelización de las condiciones extremas de funcionamiento de las aeronaves.
- Configuraciones Avanzadas: frenos aerodinámicos: incremento de la resistencia con cambio mínimo de sustentación y mínimo ruido; aerodinámica del avión de tres superficies; aerodinámica de las superficies de cola no convencionales; influencia aerodinámica de las tolerancias de fabricación, del acabado superficial y de formación de hielo.

Por situar la investigación actual dentro de las prioridades anteriores, se enumeran a continuación algunos de los proyectos europeos en curso o recientemente finalizados:



- 3AS: Active Aero-elastic Aircraft Structures.
- AEROMEMS II: Advanced aerodynamic flow control using MEMS.
- EUROLIFT-II: European High Lift Programme II.
- WALLTURB: A European Synergy for the assessment of wall turbulence.
- REMFI: Rear Fuselage and Empennage Flow Investigation.
- The Design of Morphing Aircraft.
- FAR-WAKE: Fundamental Research on Aircraft Wake Phenomena.
- DESIDER: Detached Eddy Simulation for Industrial aerodynamics.

(se han subrayado los que cuentan con participación española; además, el proyecto REMFI, en particular, está coordinado por AIRBUS-España, en línea con su responsabilidad en la parte posterior del fuselaje y de las superficies de control de la cola).

FIGURA 16

Avión supersónico a baja velocidad: distribución de presión ($\alpha = 10^\circ$, $Re = 25 \times 10^6$)

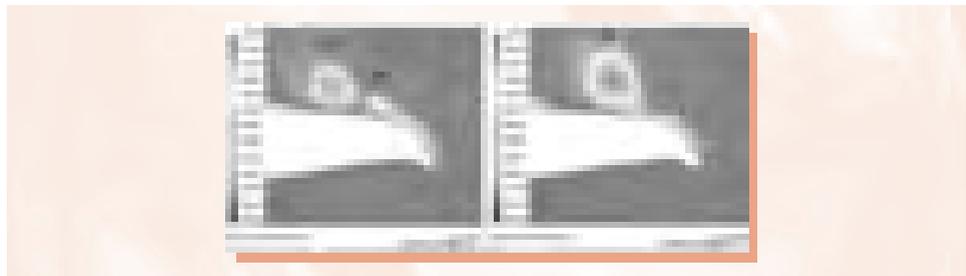


FIGURA 17

Avión supersónico a baja velocidad: sustentación y resistencia
 $M = 0,25$, $Re = 6,9 \times 10^6$. Comparación de cálculos y ensayos en túnel



Como complemento a esos proyectos puede ser de interés también mencionar las actividades recientes sobre temas aerodinámicos desarrolladas en GARTEUR ², lo que añade un mayor visión global a las necesidades de la Industria Aeronáutica en materia aerodinámica:

² Group for the Aeronautical Research and Technology in Europe (www.garteur.org)

- Implementación de criterios de transición en cálculos de Navier-Stokes.
- Desarrollo de métodos para cálculo aerodinámico no estacionario.
- Predicción de la resistencia aerodinámica.
- Predicción de la máxima sustentación.
- Validación del cálculo de configuraciones ala-fuselaje en régimen transónico.
- Cálculo aerodinámico de configuraciones hiper-sustentadoras.
- Aerodinámica de tomas integradas de motor.
- Cálculos de Navier-Stokes para configuraciones esbeltas.
- Actuaciones de avión supersónico a baja velocidad.
- Predicción de la degradación de actuaciones debida a la formación de hielo.
- Predicción de la formación y crecimiento de hielo en componentes aeronáuticos.

De entre todos estos proyectos consideramos de interés, teniendo en cuenta el contexto de este trabajo, destacar algunos aspectos concretos:

Si bien el cálculo aerodinámico ha progresado considerablemente durante las últimas décadas, también lo han hecho los **medios de ensayo**. Como ejemplo, en la fig. 9 se muestra la distribución de presión obtenida mediante la técnica PSP (Pinturas Sensibles a la Presión) sobre un avión de combate en uno de los túneles aerodinámicos del DLR alemán. Más que una competencia con los métodos de cálculo, estos desarrollos tienen que contemplarse como un estímulo y un complemento a las posibilidades que ofrecen los cálculos.

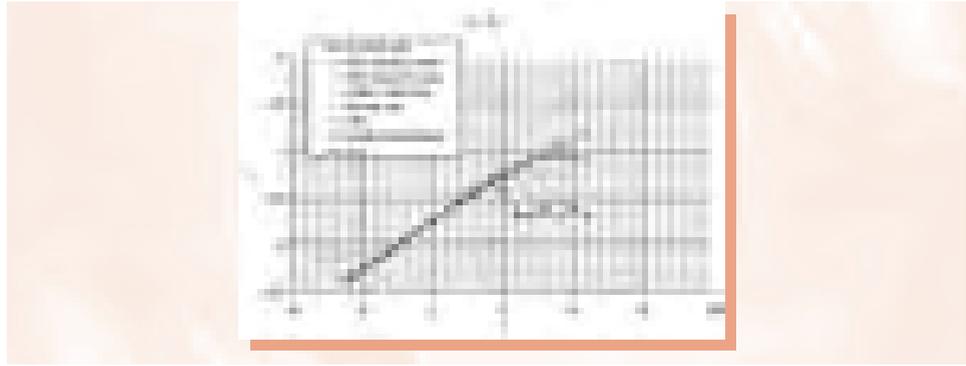
FIGURA 18

Comparación de formas de hielo calculadas



FIGURA 19

Comparación de coeficientes de sustentación de ala con hielo



Precisamente una de las limitaciones principales del cálculo, y por ello una de las cuestiones a las que tendría que dedicarse más esfuerzo, por los enormes tiempos de cálculo necesario en casos de interés industrial, es el estudio de los **fenómenos no estacionarios**. En la fig .9 puede verse el resultado parcial de uno de tales casos, el de un ala LANN en movimiento uniformemente acelerado. Un cálculo completo puede durar varias semanas de ordenador.

Más arriba se mencionó someramente la importancia de la generación de mallados. Una de las razones es que se trata de la tarea que suele demandar más tiempo de ingeniería, especialmente si se pretende que las mallas sean portables, robustas y capaces de capturar los fenómenos físicos en estudio. Puede llevar varios meses generar una malla de calidad alrededor de una configuración de avión completo. En la fig.10 se muestra un ejemplo de malla híbrida con varios millones de celdillas, que aprovecha simultáneamente las ventajas de las mallas no estructuradas (para el campo lejano) y estructuradas (en las proximidades de la aeronave).

Desde un punto de vista físico, y en consecuencia también matemático, si pretende resolverse numéricamente, uno de los problemas más exigentes es el de la aerodinámica de los dispositivos hipersustentadores, por el elevado número de características que se presentan simultáneamente: desprendimiento laminar, desprendimiento turbulento, ondas de choque, transición laminar-turbulenta, capas límites confluentes, etc. (fig.11); fenómenos muy dependientes del número de Reynolds local, lo que hace también especialmente complicada la extrapolación de ensayos en túnel a las condiciones reales. Para el calculista se trata de un reto de especial complejidad, como se muestra por las discrepancias existentes en los distintos resultados obtenidos en sustentación y resistencia para una configuración típica (fig.12), discrepancias que son mayores a medida que progresa el desprendimiento de la corriente.

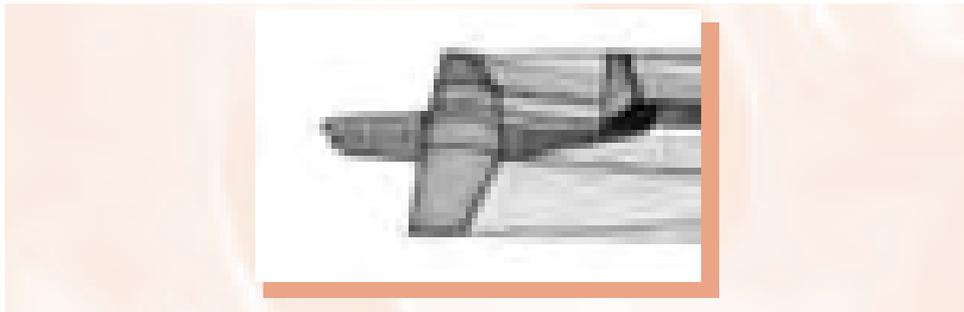
FIGURA 20

Efecto de la formación de hielo en el comportamiento aerodinámico



FIGURA 21

Mallado de superficie para el cálculo aerodinámico mediante el método de paneles



Un problema específico de la aviación militar lo constituye el diseño de las tomas de motor, pues al ser necesaria la realización de maniobras más exigentes también lo es el cumplimiento de las condiciones de funcionamiento del motor en un rango mayor de velocidades y ángulos de ataque que en el de la aviación civil. A su vez ello conduce a formas más complejas (fig.13) con la correspondiente dificultad adicional en la realización de mallados, tanto volumétricos como de superficie.

Casi cualquier cálculo aerodinámico suele tener como objetivo último la disminución de la resistencia aerodinámica, para lo que es decisivo en primer lugar que su determinación sea lo más precisa posible. En general es ahí donde los métodos de cálculo presentan la limitación principal, siendo habituales discrepancias de hasta del orden del 10% (fig. 14).

Durante los últimos años, tras la retirada del **avión supersónico** civil Concorde, el estudio renovado del vuelo supersónico ha adquirido un nuevo impulso, siendo el problema crítico el comportamiento a baja velocidad de una aeronave que se diseña para un funcionamiento óptimo a



alta velocidad. Los dispositivos hiper-sustentadores adquieren así una importancia crucial. Las figuras 15, 16 y 17 recogen algunos de los resultados más importantes obtenidos hasta el momento en la evaluación y selección de los métodos más apropiados.

Se mencionó anteriormente también la necesidad de diseñar aeronaves que puedan operar en cualquier condición meteorológica, lo que incluye el vuelo con **formación de hielo** en las alas. En la fig.18 puede verse la enorme dispersión existente entre los distintos métodos numéricos disponibles, lo que justifica el trabajo necesario por realizar todavía en este terreno.

FIGURA 22

Cálculo aerodinámico mediante el método de paneles



FIGURA 23

Diseño de aeronaves no tripuladas utilizando métodos de paneles



Si lo que se comparan son las características aerodinámicas de formas con hielo, la comparación no es mucho más favorable (fig. 19), especialmente más allá de la máxima sustentación.

También, incluso cuando por medios técnicos se elimina el hielo formado, en ocasiones permanece un hielo residual que da lugar a comportamientos no estacionarios indeseables (fig.20), que hay que intentar evitar.

Cuando se observa que con los mejores métodos de cálculo disponibles, que suelen ser también los más costosos en tiempo, persisten incertidumbres no fácilmente cuantificables, se justifica la realización de **cálculos más tradicionales** (métodos de paneles), especialmente en las etapas preliminares de diseño. Los mallados son más sencillos de realizar (fig.21), pero requieren por parte del usuario en ocasiones un mayor conocimiento físico del problema que con los métodos más sofisticados, como por ejemplo en la localización y discretización de la estela.

Ello hace necesario que la Universidad no abandone la enseñanza de estos métodos, que tanto servicio continúan prestando a la Industria (figuras 22 y 23).

Una prueba adicional de que no siempre está justificado utilizar los métodos más complejos, es en el caso del **cálculo de estelas** (fig.24). Si lo que interesa es determinar su posición, las ecuaciones de Euler proporcionan prácticamente idénticos resultados que las de Navier-Stokes, con mucho menor tiempo de cálculo. No ocurre lo mismo, evidentemente, cuando lo que interesa es determinar su intensidad (fig.25). Es necesario por lo tanto una evaluación previa del problema para decidir las herramientas de cálculo más apropiadas a cada caso, lo que también tendría que ser tema prioritario de enseñanza en la Universidad, para evitar retrasos difícilmente justificables en los departamentos de diseño por un empleo no óptimo de las herramientas disponibles.

Para finalizar, y continuando con la cuestión de la elección de las herramientas de cálculo a utilizar en cada caso, ante el reto de tener que abordar diseños no convencionales (figuras 26 y 27) para poder alcanzar los saltos cuantitativos que requiere la Industria se hace necesario la **validación** de dichas herramientas en unas condiciones para las que muchas veces originalmente no se han desarrollado, siendo así necesario un esfuerzo adicional no siempre tenido en cuenta.

FIGURA 24

Estudios de estelas: interferencias de despegue y aterrizaje



FIGURA 25

Intensidad y posición de los torbellinos de estela



Recomendaciones

Tanto los requisitos planteados por la Industria como la experiencia adquirida en los proyectos abordados durante los últimos años nos permiten resumir los que consideramos que serán problemas abiertos durante bastante tiempo aún, todos ellos con una contribución necesaria de las matemáticas:

- Automatización de la generación de mallados.
- Aceleración de la convergencia.
- Modelización de la turbulencia.
- Transición laminar-turbulenta.
- Determinación de la resistencia aerodinámica.
- Determinación de la máxima sustentación.
- Altos ángulos de ataque (corriente desprendida).
- Cuantificación de las incertidumbres.
- Validación (configuraciones no convencionales).
- Diseño y optimización multidisciplinar.

FIGURA 26

Dispositivos hiper-sustentadores no convencionales



FIGURA 27

Configuraciones no convencionales



Conviene recordar que, aunque esta lista parezca muy ambiciosa, no se han mencionado otras cuestiones relacionadas con la vertiente matemática de la aerodinámica, como los métodos SPH, espectrales, lattice-Boltzmann, con distintas aplicaciones, o los cálculos a bajos números de Reynolds, de interés cada vez mayor en el diseño de micro-UAV³s, todo lo cual puede ser objeto de futuros trabajos de colaboración entre la Ingeniería y las Matemáticas.

³ UAV: Unmanned Air Vehicle



Agradecimientos

Agradecemos al comité organizador de la Jornada la solicitud para presentar este breve resumen de necesidades aerodinámicas en la industria aeronáutica. Estamos convencidos de que la participación de la comunidad matemática contribuirá a mejorar la situación en algunos de los numerosos problemas todavía susceptibles de mejora, siendo necesario para ello un esfuerzo conjunto de la Industria, Centros de Investigación y Universidades, en paralelo a lo que ocurre ya actualmente en la mayoría de países de nuestro entorno.

Bibliografía

- [1] European Aeronautics: a vision for 2020.
<http://europa.eu.int/comm/research/growth/aeronautics2020/en/index.html>
- [2] CEAS (Council of European Aerospace Societies) / KATnet Conference of Key Aerodynamic Technologies, Bremen, Alemania 20-22 de junio de 2005
www.kat-net.net y www.ceas.org
- [3] Work Programme 2002-2006. Thematic Priority: 1.4 Aeronautics and Space. Revisión de Junio de 2005.
ftp://ftp.cordis.lu/pub/fp6/docs/wp/sp1/d_wp_200216_en.pdf
- [4] GARTEUR Annual Report 2005. Mayo 2006 Información general sobre las actividades de GARTEUR.
web www.garteur.org
- [5] <http://www.cordis.lu/fp6/projects.htm>

Capítulo XV

MATEMÁTICAS Y ECONOMÍA

POR DANIEL PEÑA
UNIVERSIDAD CARLOS III DE MADRID
Dpto. de Estadística
c/ Madrid, 126
28903 Getafe (Madrid)
Correo electrónico: dpena@est-econ.uc3m.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

This article reviews the key role of Mathematics in the economic sciences and presents some examples of how mathematics has contributed to shape economic thought. Mathematics are used in economics to build models which describe the relationship among the economic variables, to obtain optimal policies and decisions and to forecast. Section 2 of this article discusses some of the contributions to Economic Theory from Mathematics and Mathematical economics: setting and solving equilibrium problems, the development of game theory, dynamic optimization theory, decision theory and social choice models. These areas form the core of modern economic theory and the research in these topics have been recognized by many Nobel Prizes in Economics. Section 3 discusses Statistics and Econometrics and their contribution to understand economic relationships, to model the dependency among economic variables and to build time series models for forecasting and control. Section 4 includes some concluding remarks.

Introducción

Este trabajo resume el papel central de la Matemática en las Ciencias Económicas (que incluyen la economía propiamente dicha, pero también la administración de empresas), ilustra mediante algunos ejemplos su contribución y presenta algunos de los temas de investigación que pueden ser importantes en el futuro.

En las Ciencias económicas la matemática se utiliza para construir modelos que establezcan relaciones entre las variables de interés y permitan establecer reglas generales, comparar política alternativas y hacer predicciones. Para construir estos modelos es necesario entender las propiedades de las variables económicas y las relaciones entre ellas, formular y resolver los problemas de optimización que son el núcleo de la teoría económica, tanto estáticos como dinámicos, analizar sistemas dinámicos y sus propiedades (convergencia, estabilidad, etc.) y resolver los problemas de equilibrio que con el centro de la teoría económica. En segundo lugar la matemática, y en particular la Estadística, se utiliza para validar estos modelos mediante los datos observados. La primera actividad corresponde al campo de la Economía Matemática y la segunda al campo de la Estadística y la Econometría. Esta clasificación puede verse también con la aplicación de la Matemática a problemas deterministas o probabilísticos. Es ilustrativo que los dos primeros premios Nobel en Economía se hayan adjudicado a representantes de estas dos líneas de trabajo: el primer año del premio, 1969, se adjudica a los estadísticos/económetras R. Fish y J. Tinbergen por su contribución al desarrollo de los modelos econométricos. El año siguiente, 1970, lo recibe P. Samuelson que aplica con éxito el Análisis matemático al desarrollo de la teoría económica.

Esta nota se estructura como sigue. En la sección 2 presentamos de forma sucinta algunos de los problemas principales abordados en la Economía matemática. En la sección 3 comentamos algunas de las contribuciones de la Estadística a la Economía. La sección 4 incluye algunos comentarios finales.

La Economía Matemática

Un modelo económico está en equilibrio cuando las decisiones de los agentes económicos son compatibles y coherentes con sus objetivos. Por ejemplo, un mercado donde se intercambia un bien está en equilibrio si el precio es tal que la cantidad ofertada es igual a la cantidad demandada. El problema del equilibrio es pues central a la teoría económica y la teoría del equilibrio general (que se refiere a toda la economía) constituye un cuerpo teórico consolidado en el que distintas ramas de la matemática juegan un papel esencial (Debreu, 1987).

Durante la mitad del siglo XX y utilizando como herramienta matemática el teorema del punto fijo de Kakutani, se obtuvieron dos resultados fundamentales:

1. La existencia de equilibrio competitivo (walrasiano) en una economía de intercambio, resultado obtenido por K. Arrow y G. Debreu, con lo cual se resolvía la vieja conjetura de L. Walras, enunciada en la segunda mitad del siglo XIX, acerca de la bondad como mecanismo de un sistema de precios para realizar asignaciones en una economía.
2. La existencia de equilibrio (de Nash) en un juego estático no-cooperativo (es decir, donde no hay imposiciones externas) de tipo finito (es decir, cuando hay un número finito tanto de agentes como de estrategias factibles), si ampliamos el espacio de estrategias, permitiendo que los agentes puedan ocultar sus acciones eligiendo su estrategia de forma aleatoria.

En la actualidad las técnicas matemáticas son fundamentales para analizar el efecto de las políticas económicas, la evolución de la riqueza y la distribución de la renta y la variabilidad de los precios. Muchos de los fenómenos económicos se desarrollan a lo largo de un periodo de tiempo, finito o infinito, de forma que las decisiones presentes influyen notablemente en las futuras. De esta forma, se hace imprescindible utilizar las técnicas de optimización dinámica para atacar de forma eficiente dichos problemas.

Las herramientas utilizadas son la teoría de espacios métricos y de Banach, la teoría de la medida e integración, los procesos de Markov, los teoremas de punto fijo de Banach, Brower, Schauder, Kakutani y Fan, así como el cálculo de variaciones y el principio del máximo de Pontryagin. Por otra parte, para estudiar propiedades de continuidad del equilibrio con respecto a los datos iniciales (recursos y preferencias) se manejan topologías no triviales en el espacio de las características de los agentes. La programación dinámica, creada por Bellman, permite determinar soluciones robustas de los problemas de optimización, lo que permite reaccionar de forma óptima frente a desviaciones inesperadas. Esta propiedad va asociada a grandes problemas de cálculo por lo que el cálculo numérico juega un papel destacado en la resolución de estos problemas.

Aunque el concepto de equilibrio de Nash fue criticado inicialmente, en aquellos años el campo de estudio primordial era la teoría de juegos cooperativa introducida por Von Neumann, su enfoque ha sido seguido en la teoría de juegos con más aplicaciones prácticas en la economía: la teoría no cooperativa. El concepto de equilibrio de Nash ha sido refinado para juegos dinámicos



(equilibrio de Nash perfecto en subjugos), extendiéndolo a situaciones en los que se presentan asimetrías de información para los diversos jugadores (equilibrio de Nash bayesiano) y, en general, en una nueva rama a caballo entre la teoría de juegos y la lógica matemática, el estudio del comportamiento de los agentes que no son puramente racionales.

Un tercer área de investigación importante es la teoría de la elección social. Arrow demostró la inexistencia de un mecanismo de agregación de preferencias que cumpliera ciertas condiciones “razonables”. Este resultado, conocido como teorema de la imposibilidad de Arrow, conduce al problema de determinar aquellas reglas sociales en las que deberían basarse las decisiones sociales. Este problema forma parte de otro más general, conocido como agregación de preferencias. Esencialmente, se trata de determinar las propiedades generales agregadas (propiedades macroeconómicas) de la sociedad, partiendo del comportamiento individual de cada uno de los agentes. Para modelar estas ideas se parte de una función (la función de elección social) cuyo dominio y conjunto imagen están constituidos por todas las relaciones de preferencias que podrían tener los agentes que componen la sociedad. Sobre esta función se imponen condiciones matemáticas, cuyo contenido refleja las nociones de justicia de la sociedad. Por lo general, estas propiedades determinan una forma única para la función de elección social. Fijada una sociedad particular, esta función transforma el conjunto de preferencias de los agentes que la constituyen en una única relación agregada de preferencias. Consecuentemente, las decisiones que se toman son aquellas que maximizan las preferencias agregadas por la función de elección social. Gran parte de la investigación en este área se ha dirigido a encontrar propiedades razonables de la función de elección social que no conlleven el resultado dictatorial (D. Black, A. Sen) o a encontrar otros contextos donde las conclusiones del Teorema de Arrow sigan siendo válidas (Teorema de Gibbard-Satterthwaite). Durante la década de 1990, el área de Economía Política ha experimentado un crecimiento importante. En los modelos que se consideran en este campo, los espacios tienen, de forma natural, una topología más complicada que la que aparece en Teoría de Juegos o Equilibrio General. En consecuencia, ha sido necesario recurrir a la Teoría de Puntos fijos de Lefschetz, utilizada, por ejemplo, en modelos de formación endógena de partidos.

Por último un campo reciente de investigación es el de equilibrio intertemporal en mercados financieros. En esta situación puede no existir precio de equilibrio ya que la no acotación permite oportunidades de arbitraje. A partir del trabajo de pionero de Hart en 1974 se han estudiado condiciones que limitan las oportunidades de arbitraje y que se traducen en una acotación endógena del conjunto de las estrategias. Los desarrollos en el análisis de portafolios en tiempo continuo de Black y Scholes (1973) y otros han motivado el estudio del problema de la existencia de equilibrio en economías con conjuntos de estrategias no acotados e infinitas mercancías. Se requieren ahora condiciones más fuertes que utilizan técnicas específicas del análisis funcional.

La Estadística y la Econometría

Las relaciones entre variables que aparecen en cualquier libro de introducción a la Economía: función de producción, curva de oferta, curva de demanda, etc. se refieren a relaciones intrínsecamente estadísticas, es decir, que solamente pueden interpretarse en términos de valores espe-

rados o promedios. Por ejemplo, una función de producción relaciona el número de unidades que pueden fabricarse de un producto con los factores de producción, por ejemplo, horas/hombre y horas/máquina, pero esta función sólo tiene sentido si se interpreta como una relación ideal que podría estimarse a partir de los datos mediante un modelo de regresión. La necesidad de contrastar empíricamente las leyes económicas impulsó el desarrollo de métodos estadísticos dando lugar a la Econometría. Aunque las herramientas básicas para relacionar variables estadísticas, los métodos de regresión, fueron desarrollados por Gauss, Galton y Pearson en el siglo XIX, su aplicación a la medición de relaciones económicas planteó nuevos problemas como consecuencia de la imposibilidad de realizar experimentos controlados de medición como en el mundo físico. Los problemas de autocorrelación, heterocedasticidad, errores en las variables, endogeneidad y otros fueron estudiados inicialmente ligados a las aplicaciones económicas, aunque luego han sido utilizados y extendidos en otras áreas de aplicación de la estadística.

Un campo de especial importancia en las aplicaciones de los modelos estocásticos a la economía es el análisis de series temporales y la predicción económica. Hasta los años 70 la mayoría de los modelos econométricos utilizados se basaban en los modelos de regresión clásicos, concebidos para situaciones estáticas. En estos años Box y Jenkins (1970) desarrollan una metodología para construir modelos lineales de series temporales, los modelos ARIMA (autoregressive integrated moving average) que suponen un avance fundamental en la modelización de series temporales en cualquier área, y, en particular, en las ciencias económicas. La aplicación de estos modelos para medir la relación entre variables económicas lleva al concepto de cointegración por el que Granger y Engle han recibido en 2003 el premio Nobel de Economía. Engle (1982), ha sido además pionero en el estudio de los modelos condicionalmente heterocedásticos (ARCH) de gran importancia en la modelización de series temporales financieras.

En el campo de la administración de empresas, las herramientas estadísticas han tenido un impacto especial en el área de producción, con el desarrollo de los métodos de control de calidad; en marketing, con las aplicaciones del diseño de experimentos y de las técnicas de análisis multivariante para clasificar y segmentar a los clientes, encontrar grupos homogéneos y reducir la dimensión de grandes masas de datos; y en las finanzas y los seguros, con la modelización y gestión del riesgo, la selección de carteras y las teorías de valoración y cobertura de riesgos. Un campo aparte merecen las herramientas de Investigación Operativa, en las que no entraremos por falta de espacio y porque son objeto de otra viñeta en este volumen.

Los problemas de investigación abiertos en la modelización estocástica en las ciencias económicas son numerosos. Las técnicas estadísticas clásicas se han desarrollado para aprovechar óptimamente una información escasa, con tamaños de muestra pequeños y medianos y con pocas variables, pero en la actualidad, cada vez con más frecuencia, se dispone de datos económicos y empresariales enormes con muchas variables. Por ejemplo, un banco de datos de los gastos con tarjetas de crédito emitidas por una entidad financiera puede incluir miles de variables y centenares de miles de observaciones que tienen además una dimensión temporal. Las técnicas estadísticas para analizar estos bancos de datos son insuficientes, ya que: (1) es previsible un alto grado de heterogeneidad, siendo la relación entre las variables distinta en distintos elementos de la muestra; y (2) el



número de variables disponible puede ser muy alto e incluso mayor que el de observaciones. En estos casos los procedimientos habituales para encontrar grupos y relacionar variables son poco efectivos. Por ejemplo, construir un modelo de series temporales con cien variables es muy complejo (y en la práctica inviable) con los métodos actuales diseñados para unas pocas series. Esto ha dado lugar a un desarrollo de la estadística multivariante que se conoce también como herramientas de minería de datos (Data mining, véase Hastie *et al*, 2001). Un segundo problema es la combinación de información de distintas fuentes. La facilidad para acceder a muchos datos distintos tomados en condiciones heterogéneas ha puesto en un lugar central los llamados problemas de meta-análisis donde debemos combinar información diversa, frecuentemente con métodos Bayesianos. Un campo donde se utiliza de forma efectiva esta combinación de información es en la predicción mediante el llamado promedio Bayesiano de modelos: en lugar de seleccionar un modelo para los datos y generar predicciones con él, podemos construir muchos modelos distintos y combinar sus predicciones con pesos proporcionales a sus probabilidades a posteriori.

Conclusiones

La Matemática ha jugado un papel central en el desarrollo de las Ciencias económicas, que la utilizan para construir sus modelos explicativos y para contrastarlos con los datos reales. Es previsible que la creciente capacidad de computación conduzca a modelos matemáticos más complejos y a métodos más sofisticados de estimación y predicción. Para ello son necesarios avances en las propias matemáticas para aportar las herramientas conceptuales y analíticas que permitan abordar estos nuevos problemas.

Agradecimientos

La exposición sobre Economía Matemática se basa en las ideas de Carlos Hervés, Francisco Maruhenda, Carmelo Núñez y Juan Pablo Rincón. La contribución que puede haber en esta sección es suya, y si hay errores son de mi responsabilidad. Alejandro Balbás y Diego Moreno me han enviado también sus opiniones sobre problemas abiertos de investigación que me han ayudado mucho. Quiero agradecer a estos ejemplares colegas su generosa aportación a este proyecto.

Bibliografía

- [1] Black, F.; Scholes, M.; The pricing of options and corporate liabilities, *Journal of Political Economy*, 81, (1973), (637-659).
- [2] Box, G.E.P.; Jenkins, G.; *Time series Analysis, Forecasting and Control*, Holden Day, (1970).
- [3] Debreu, G.; *Theory of value: an axiomatic analysis of economic equilibrium*. New Haven: Yale University Press, (1987).
- [4] Engle, R.F.; "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, (1982), (987-1007).
- [5] Hastie, T.; Tibshirani, R.; Friedman, J.; *The elements of statistical learning*. Wiley, (2001).

Capítulo XVI

MEJORES DECISIONES PARA UN MUNDO MEJOR

POR DAVID RÍOS INSUA
UNIVERSIDAD REY JUAN CARLOS
Y REAL ACADEMIA DE CIENCIAS
Departamento de Estadística e Investigación Operativa
c/ Tulipán, s/n
28933 Móstoles (Madrid)
Correo electrónico: david.rios@urjc.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

*/*Better Decisions for a Better World */* is the vision of the EDEMOCRACIA-CM program of the Government of Madrid, with which we try to benefit from the ubiquitous presence of Internet to deploy tools over the web to facilitate rational support to individuals and/or groups to undertake improved decisions. This note identifies some of the technological and mathematical challenges around such concept, both when devising personal assistants and group decision support tools, specifically in the context of public policy decision making, a.k.a. electronic democracy.

Introducción

Mejores decisiones para un mundo mejor es la visión que asociamos al programa *EDEMOCRACIA-CM, Conceptos y Sistemas de apoyo a la Democracia Electrónica*, recientemente seleccionado por la Comunidad de Madrid entre sus líneas estratégicas, formado por un grupo interdisciplinario (matemáticos, estadísticos, investigadores operativos, informáticos, telemáticos, juristas, politólogos y economistas) de la UNED, UC3M, UPM y URJC.

En el mismo se parte de dos premisas básicas:

- La presencia casi ubicua de Internet permite poner a disposición de las personas herramientas que facilitan la toma racional de decisiones, en cualquier situación.
- Involucrar, desde un principio, a las personas con interés en un problema de toma de decisiones, permite mejorar la calidad de las decisiones tomadas, haciéndolas más consensuadas y transparentes.

A lo largo de muchos años, la investigación en Ciencias de la Decisión, ver por ejemplo French y Ríos Insua (2000), ha aportado comprensión sobre cómo las personas deben tomar decisiones, sobre cómo las toman y sobre cómo podríamos ayudarlas a tomar mejores decisiones. En particular, ha habido numerosos desarrollos en las metodologías bayesianas, que han aportado herramientas analíticas clave en el apoyo de la toma de decisiones: árboles de decisión, diagramas de influencia, la síntesis de juicios de expertos y datos, la modelización de utilidades multiatributo y las herramientas del análisis de sensibilidad, por ejemplo. Se consigue así que el objetivo del apoyo a la toma de decisiones sea el descubrimiento y la exploración de las implicaciones de los juicios del decisor. El análisis de decisiones y las herramientas de apoyo a la toma de decisiones han avanzado tanto que es realista apoyar la toma de decisiones estratégicas por parte de un grupo de decisores con objetivos razonablemente comunes.

Al mismo tiempo que se han producido estos desarrollos metodológicos, matemáticos y de software, se viene reconociendo cada vez más la importancia de una comunicación clara y sensible entre el público y el gobierno en asuntos que afectan al riesgo, y, más generalmente, a una mayor implicación de todos los grupos que tengan interés en un proceso de toma de decisiones.

Creemos, así, que el crecimiento en potencia de cálculo y entornos gráficos basados en la web, junto con el mayor acceso a Internet aportan una vía para sacar el entendimiento proporcionado por las herramientas de ayuda a la decisión fuera de las habitaciones cerradas de los decisores, para, al menos, transmitir el razonamiento que sustentan las decisiones a los grupos de interés y, en el mejor de los casos, posibilitar una mayor implicación de los mismos en la propia toma de decisiones.

De esta manera, resulta concebible el desarrollo de una infraestructura de apoyo a la toma de decisiones y de comunicaciones que permita realizar análisis múltiples que aporten diversas perspectivas sobre cuestiones a ser exploradas y comparadas, identificando, por tanto, los acuerdos y diferencias en las percepciones y valores de los distintos grupos de interés. Podríamos, entonces, diseñar dispositivos que permitan:

- A un individuo razonar sobre un problema de toma de decisiones moderadamente complejo, para encontrar la alternativa óptima.
- A un grupo distribuido razonar y negociar sobre un problema moderadamente complejo de toma de decisiones en el sitio, hasta acordar una alternativa.

Asistentes personales

La primera clase de dispositivos se referiría a asistentes personales, por ejemplo integrados en nuestro PDA o móvil 3G, que nos ayudasen a tomar decisiones en distintos contextos, teniendo en cuenta nuestras preferencias y creencias.

Para ello, siguiendo el ciclo del análisis de decisiones, tendríamos que:

1. Diseñar un sistema adecuado (ligero, eficiente y riguroso) de estructuración de problemas de decisión.
2. Diseñar un sistema adecuado de modelización de preferencias.
3. Diseñar un solver y un analizador de sensibilidad adecuados.
4. Diseñar un mecanismo adecuado de comunicación con un sistema web.
5. Implantar el sistema.

Además de las herramientas computacionales y de modelización mencionadas, se podría desarrollar toda una biblioteca para problemas estándar de toma de decisiones (compra de casa, compra de coches...) para facilitar su resolución en el sitio, en forma similar al sistema MOIRA (Ríos Insua et al, 2000) que facilita la solución de problemas de adopción de contramedidas, después de un accidente nuclear.



Hacia la democracia electrónica

La siguiente clase de dispositivos permitiría a un grupo de personas tomar una decisión de forma más participativa y consensuada, promoviendo una aproximación a la democracia electrónica, ver Ríos Insua et al (2006), más innovadora y transformadora que los conceptos actualmente manejados, que, esencialmente, consisten en usos políticos de siglos pasados facilitados por el empleo de las TIC (p.ej., voto electrónico en lugar de voto en papel o flashmobs en lugar de manifestaciones convocadas mediante pasquines).

En este caso, apoyaríamos de nuevo las fases del ciclo del análisis de decisiones antes mencionadas, permitiendo a cada participante explorar sus juicios y determinar su alternativa óptima.

Además, puesto que, típicamente, los participantes llegarían a soluciones óptimas diferentes, deberíamos añadir a nuestro dispositivo:

1. Un mecanismo adecuado de negociación implementado en web, para el caso, habitual, de que haya discrepancias en las preferencias de los distintos participantes.
2. Un mecanismo de votación, para el caso en el que las negociaciones fracasasen;
3. Un mecanismo de renegociación para el caso, no tan inusual, de que la solución votada sea subóptima.

De nuevo, sería conveniente crear una biblioteca de problemas estándar, tomando, por ejemplo, como modelo el sistema PARBUD, Ríos *et al* (2005), que hemos diseñado para apoyar la elaboración de presupuestos participativos.

Algunos retos para las Matemáticas de la Decisión

En definitiva, el proceso general de toma de decisiones por grupos que proponemos implantar pasaría por:

- la estructuración del problema de toma de decisiones;
- la modelización de preferencias y creencias para cada participante, comprobando sus posibles inconsistencias;
- la determinación de la alternativa óptima para cada participante, indicando la robustez de la misma;
- la negociación, en caso de que los participantes no obtuviesen la misma solución óptima;
- la votación, en caso de que no se llegase a una solución negociada;
- la postnegociación, en caso de que la solución votada no fuese suficientemente buena;
- la comunicación de resultados.

En esta sección, formulamos muy brevemente algunas de las cuestiones de tipo matemático que habría que resolver para la correcta implantación de este paradigma.

Estructuración de problemas: Se trataría, esencialmente, de determinar los métodos de estructuración de problemas de toma de decisiones más adecuados para implementar en web, teniendo en cuenta la poca sofisticación de los usuarios al otro lado de la red. Para ello, será necesario realizar comparaciones entre métodos de estructuración para asegurar su relevancia para el tipo de decisiones complejas que consideramos. Otras dos cuestiones importantes apenas tratadas en la literatura, serían la investigación de métodos para evaluar la sensibilidad de la solución respecto de la estructura del problema y la exploración de la selección y la mezcla de modelos de decisión para un mismo problema de toma de decisiones.

Modelización de incertidumbre: Desde nuestra perspectiva, la cuestión principal en este campo sería estudiar métodos de asignación de creencias adecuados para su implementación en web, con énfasis en la previsible poca sofisticación de los usuarios.

Modelización de preferencias: Dada la asignación de preferencias no asistida por un analista a través de la web, deberían desarrollarse nuevos modelos de asignación de preferencias de decisores, con énfasis en el análisis de modelos de errores en preferencias.

Cálculo de alternativas óptimas: Se debería promover el desarrollo de métodos de optimización adecuados a nuestro contexto de decisores múltiples, a través de la web. Así, entre otros, se deberían explorar extensiones de los métodos de simulación de probabilidades aumentadas, aproximaciones miopes penalizadas, y su combinación con métodos de programación neurodinámica, para el cálculo de alternativas óptimas para cada decisor participante en problemas complejos de toma de decisiones. Además, puesto que habrá varios participantes con preferencia y creencias dispares, y posiblemente imprecisas, se deberían desarrollar nuevos métodos y algoritmos de cálculo de soluciones no dominadas.

Análisis de sensibilidad: La presencia de múltiples decisores exigirá herramientas más potentes de análisis de sensibilidad. En particular, sería interesante estudiar cómo emplearlas para explorar posibles causas de conflicto entre decisores y promover consenso, posiblemente con métodos multi-agente.

Análisis de negociaciones: Lógicamente, dada la previsible disparidad de opiniones de los participantes, se requerirán herramientas de análisis de negociaciones. Para ello, debería realizarse un estudio unificado de métodos de arbitraje y de negociación adecuados a las estructuras de problemas considerados y relacionarlos con métodos de votación.

Conclusiones

Se ha descrito cómo los desarrollos recientes en TIC permiten poner a disposición del gran público herramientas que facilitan la toma de decisiones racionales. Esta actividad, a su vez, sugiere



interesantes retos matemáticos con inmediata aplicabilidad práctica, que podrían tener una buena acogida en un Instituto de Matemáticas.

Agradecimientos: Nota preparada con ayudas del programa Towards Electronic Democracy de la European Science Foundation, del programa EDEMOCRACIA-CM de la DGU de la Comunidad de Madrid, del Laboratorio URJC-Fundación DMR Consulting de Ingeniería de la Decisión y del proyecto EDEMO del Ministerio de Educación y Ciencia.

Bibliografía

- [1] French, S.; Ríos Insua, D.; *Statistical Decision Theory*, Edward Arnold, (2000), (301 pgs).
- [2] Ríos, J.; Ríos Insua, D.; Fernandez, E.; Rivero, J.; Participatory budget formation through the web, en *Lecture Notes in Notes in Computer Science*, 3146, (2005), (225-234).
- [3] Ríos Insua, D.; Gallego, E.; Ríos Insua, S.; Mateos, A.; MOIRA: A DSS for restoration of aquatic ecosystems contaminated by radionuclides, *Annals of Operations Research*, 95, (2000), (341-364).
- [4] Ríos Insua, D.; Ríos, J.; Fernández, E.; *Hacia la Democracia Electrónica*, manuscrito, (2006), (210 pgs).

Capítulo XVII

MATEMÁTICAS Y NEUROCIENCIAS

POR JUAN MANUEL RODRÍGUEZ PARRONDO
UNIVERSIDAD COMPLUTENSE DE MADRID
Departamento de Física Atómica, Molecular y Nuclear,
Facultad de Física
CP 28040
Correo electrónico: parrondo@fis.ucm.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

We have explored the rich and complex interplay between Mathematics and Neuroscience following three main branches of research. The first one is the use of mathematical models to understand the functioning of neurons, originated in the seminal work by Hodgkin and Huxley, back in 1952. The dynamics of single neurons and ensembles of neurons is still a very active field of research, and new properties such as synchronization, criticality or excitability help to understand the behaviour and function of the building blocks of our brain. The second branch is devoted to the logical aspects of cognition. Gödel's theorem and the theory of computability can be considered as the first steps toward a "mathematics of cognition", which has been further developed specially in the context of computer science and artificial intelligence. Finally, in the third branch new mathematical concepts, such as chaos, complex networks or stochastic resonance, are being applied to the analysis of neurological and psychological data and to the design of models for actual cognitive processes in animals and humans. These three research lines are continuously exchanging ideas and concepts in order to solve one of the most exciting problems of modern science: how our brain works.

Introducción

Para los Estados Unidos los 90 fueron oficialmente la "década del cerebro" y el siglo XXI lo será para Japón. Entender cómo funciona nuestro cerebro no es sólo una de las más ambiciosas aspiraciones intelectuales de la ciencia, sino que tiene también aplicaciones prácticas en dos ámbitos: el tratamiento de enfermedades neurológicas e incluso psíquicas, y el desarrollo de la robótica y la inteligencia artificial, mediante diseños inspirados en el cerebro de humanos y animales.

En los últimos años nuestro conocimiento del cerebro ha aumentado de forma dramática pero al mismo tiempo tenemos la sensación de encontrarnos tan solo en el comienzo: ignoramos más de lo que hemos llegado a comprender. El cerebro es uno de los sistemas más complejos a los que se enfrenta la ciencia. Es por ello que las ciencias positivas se acercan a él desde todos los flancos: psicología, bioquímica, biología celular, informática y modelización matemática.

Nuestro cerebro es esencialmente un conjunto de neuronas conectadas entre sí y nuestras capacidades cognitivas son fruto de la actividad colectiva de estas neuronas. El cerebro es un sistema complejo, tal y como se entiende en la actualidad este término: está formado por muchas subunidades relativamente simples y cuya interacción da lugar a propiedades emergentes complejas (véase el artículo de Ángel Sánchez en este mismo volumen). Por ello, prácticamente todas las disciplinas que se agrupan bajo la llamada Ciencia de los Sistemas Complejos tienen aplicación en el cerebro: física estadística, teoría del caos y de los sistemas dinámicos, teoría de redes complejas, probabilidad y procesos estocásticos. Es más, puesto que la conciencia y las funciones cognitivas no son más que propiedades emergentes de nuestro cerebro, sólo podrán ser comprendidas utilizando estas Ciencias de la Complejidad.

Las neurociencias indagan acerca del cerebro y el sistema nervioso en todos sus niveles: el funcionamiento de las neuronas, el diseño de circuitos neuronales, la organización funcional del cerebro y la relación entre esta organización y la conducta o los procesos cognitivos. Y lo han hecho tanto en sus aspectos físico-químicos y fisiológicos, como en sus aspectos lógicos. En todos estos niveles y en todos estos aspectos, las matemáticas han sido indispensables.

Es difícil resumir en unas pocas páginas los enormes logros que las neurociencias han alcanzado y la importancia de las matemáticas en los mismos. Hemos optado por distinguir tres grandes líneas de investigación. La primera, la *línea biológica*, se ha centrado en el estudio de las neuronas y su capacidad de procesamiento de información. La segunda, la *línea lógico-computacional*, parte de la lógica matemática e intenta dilucidar cuáles son los elementos fundamentales de los procesos cognitivos. Esta línea culmina con las distintas estrategias utilizadas en inteligencia artificial. Finalmente, la *línea de la neurología y la psicología cuantitativa* estudia el comportamiento del cerebro de forma global y su relación con la conducta o con la respuesta a ciertos estímulos.

No existen fronteras nítidas entre estas líneas. Unas y otras se funden y comparten conceptos, técnicas y resultados experimentales. Sin embargo, nos serán útiles para esbozar el extenso y variado panorama de las neurociencias.

La línea biológica

Nuestro cerebro procesa información a través de los impulsos nerviosos que atraviesan las neuronas. Una neurona típica de nuestra corteza cerebral consta de un cuerpo central, o *soma*, del que nace un filamento, el *axón*, y una serie de filamentos más pequeños en forma de árbol, las *dendritas* (ver figura 28). El impulso nervioso se transmite desde las dendritas hasta el axón.

En 1939, Hodgkin y Huxley realizaron las primeras mediciones de la actividad eléctrica del llamado axón gigante del calamar y años más tarde, en 1952 [1], presentaron el primer modelo matemático para describir el impulso nervioso, el *modelo de Hodgkin-Huxley*. Por sus trabajos experimentales recibieron el Premio Nobel de Medicina, junto con John Eccles, en 1963.

El modelo de Hodgkin-Huxley es un sistema de ecuaciones diferenciales con un punto estacionario estable, que corresponde al estado de reposo de la neurona y tiene el siguiente comportamiento peculiar: si una pequeña perturbación aparta al sistema del punto estacionario, éste debe realizar una “excursión” relativamente larga para volver a dicho punto. El resultado es que el sistema responde de forma abrupta ante pequeñas perturbaciones, pero la respuesta no depende de la intensidad de la perturbación (ver figura 29). Es una respuesta todo o nada, 0 ó 1, es decir, una respuesta digital. Además, la neurona, una vez que “dispara”, necesita un cierto tiempo para recuperarse y poder ser excitada de nuevo.



FIGURA 28

Representación esquemática de una neurona



Éstas son las características principales del impulso nervioso que atraviesa nuestras neuronas y definen un tipo de sistema dinámico que se denomina sistema excitable y que no es exclusivo de las células nerviosas sino que se da en muchos otros ámbitos de la biología, como en la excitación de las células del tejido cardíaco [2] o en las redes que regulan la expresión de genes de importancia en el estudio del cáncer [3].

Existen muchas variantes y simplificaciones del modelo de Hodgkin-Huxley, como el modelo de una variable integrate-and-fire o las ecuaciones de Fitzhugh-Nagumo, con sólo dos grados de libertad.

Todos estos modelos se han aplicado a neuronas individuales [4, 5] y a redes. Pero tanto el comportamiento de las neuronas individuales, que es más variado de lo esbozado aquí, como el modo en que éstas procesan la información, son objeto de la investigación actual. Hasta mediados de los años 80 se creía que la información transmitida a través de la neurona estaba únicamente codificada en el número de disparos por segundo, o ritmo de disparo, y que el patrón detallado de disparos era irrelevante. El hecho de que este patrón variara en experimentos en donde el sujeto está expuesto al mismo estímulo sustentaba esta idea (ver figura 30). Sin embargo, más recientemente se ha podido comprobar la relevancia de estos patrones temporales: en la famosa teoría de la atención de Crick y Koch [6], basada en la sincronización de patrones de disparo; en la mayor reproducibilidad de dichos patrones cuando los estímulos están afectados por ruido [7, 8]; o en la utilización de desfases de los patrones de disparo en el procesamiento de señales auditivas [4, 9]. Todo ello ha despertado el interés por los modelos de tiempo continuo de neuronas, como el de Hodgkin-Huxley, y, en especial, por las propiedades de sincronización [10] y su respuesta ante el ruido [11]. A raíz de estos resultados, se espera el desarrollo de una nueva *teoría de la información* que incorpore la codificación en patrones temporales de disparo [9].

FIGURA 29

El potencial de acción y el modelo de Hodgkin-Huxley



En la figura superior se puede observar el potencial resultante del modelo como respuesta a un estímulo, mientras que en la inferior se muestran potenciales obtenidos experimentalmente en distintas neuronas: a) axón gigante del calamar, b) neurona del nodo de Ranvier de una rana y c) neurona del cortex visual de un gato.

Finalmente, se ha estudiado también la estadística de disparos tanto en experimentos como en modelos matemáticos [12] y responde a una ley de potencias similar a la que se da en otros sistemas complejos, como terremotos o pilas de arena.

FIGURA 30

Disparos de una neurona individual en el cortex de un mono ante el mismo estímulo



La línea lógico-computacional

La lógica matemática experimentó una auténtica revolución a principios del siglo XX gracias al teorema de Gödel. El teorema, que resolvía uno de los famosos problemas de Hilbert, evidenció



las limitaciones de los procedimientos finitos para la fundamentación de la aritmética. Pero también puso en marcha un programa de investigación sobre estos procedimientos finitos que dio lugar a una definición formal -y universal- de *algoritmo* a través de la llamada *máquina de Turing*.

En 1952 McCulloch y Pitts publicaron un artículo fundamental para la neurociencia [13]. En él definieron un modelo extremadamente simplificado de neurona, la *neurona formal de McCulloch-Pitts*. Se trata de un simple dispositivo con un gran número de entradas (las dendritas) y una salida (el axón). El dispositivo suma las entradas y si el resultado supera un cierto umbral emite una señal a través de su axón. Con estas neuronas pueden construirse redes en donde los axones alimentan las entradas de otras neuronas. Estas conexiones o *sinapsis* forman la *red neuronal* pero, como ocurre en el cerebro, pueden ser más o menos intensas. El *peso sináptico* es un coeficiente que mide esta intensidad de conexión y puede ser incluso negativo para dar cuenta de conexiones inhibitorias, en las que una neurona inhibe o dificulta la actividad de otra. McCulloch y Pitts [13] no sólo definieron su modelo de neurona, sino que probaron que las redes de estas neuronas formales eran capaces de realizar cualquier algoritmo, es decir, probaron que una red neuronal era capaz de realizar las mismas tareas de procesamiento de información que una máquina de Turing y, por tanto, que cualquier ordenador digital.

Pero una de las diferencias fundamentales entre una máquina de Turing o un ordenador convencional y el cerebro es que en este último no existe un “programador” que diseñe las distintas redes neuronales que lo forman. En 1949, el psicólogo Donald Hebb, formuló una regla de ajuste de los pesos sinápticos que permite un diseño autónomo o “aprendizaje” [14]. La regla de Hebb establece que “cuando un axón de la célula A está suficientemente cerca como para excitar la célula B y de forma repetida o consistente toma parte en su disparo, tiene lugar algún proceso de crecimiento o cambio metabólico en una o ambas células que hace que la eficacia de A, en tanto que célula que activa el disparo de B, se incremente” [14].

Las redes de neuronas McCulloch-Pitts junto con la regla de Hebb y alguna de sus variantes, fueron el punto de partida de toda una disciplina en computación y neurociencia: las redes neuronales. El campo experimentó un auge extraordinario en los años 80, gracias en gran parte al trabajo de John Hopfield [15], quien estableció una relación matemática entre estas redes y ciertos modelos de materiales magnéticos desarrollados en el seno de la física estadística.

Este auge se centró en el llamado *conexionismo*: el estudio del procesamiento de información mediante redes neuronales [16]. Surgieron intentos de modelizar procesos cerebrales mediante redes neuronales así como aplicaciones prácticas de las mismas en inteligencia artificial. Es hoy todavía un campo de investigación muy activo la aplicación de redes neuronales a una gran variedad de problemas, como reconocimiento de caras [17], objetos o escritura [18], control de sistemas [19], diagnóstico de enfermedades [20] y predicción de toda clase de fenómenos: plegamiento de proteínas [21], tiempo atmosférico [22], mercados financieros [23], crisis económicas [24], etc.

Las redes neuronales se han tratado de forma exhaustiva desde el punto de vista matemático [25] y este tratamiento las ha relacionado con distintos campos de la física y la matemática. Hemos

mencionado su relación con modelos de la física estadística. Pero una red neuronal no sólo almacena patrones o ejemplos, sino que también generaliza la información contenida en dichos patrones y es capaz de clasificar o reconocer estímulos nuevos. Esta capacidad de generalización es un caso particular de un viejo problema en estadística y tratamiento de datos. Por ello la teoría del aprendizaje en redes neuronales se ha fundamentado en la estadística matemática a través de la llamada *dimensión de Vapnik-Chervonenkis*, un concepto abstracto y muy general que mide la capacidad de un dispositivo para generalizar a partir de ejemplos [26, 27].

Finalmente, otro de los aspectos que puede en un futuro ser relevante en la elaboración de una *matemática de la cognición* es la teoría de control. En los últimos años, los trabajos del neurólogo Antonio R. Damasio han puesto de manifiesto el papel que el cuerpo y su representación en el cerebro juegan en procesos cognitivos superiores como las emociones y la conciencia [28].

Por ello, es muy posible que este tipo de procesos no puedan ser explicados mediante modelos puramente neuronales, sino que sea necesario incluir un sistema físico y su representación, como ocurre en robótica, en donde la teoría de control es una herramienta básica [29].

La línea de la neurología y la psicología cuantitativa

Finalmente, las matemáticas juegan un papel importante también en el estudio global del cerebro y la conducta realizado por neurólogos y psicólogos.

Existe una gran diversidad de técnicas de medición de la actividad cerebral: electroencefalografía, resonancia magnética, tomografía por emisión de positrones, etc., y muchas de ellas permiten obtener datos incluso cuando los sujetos están conscientes y realizando ciertas tareas.

El tratamiento e interpretación de todos estos datos está requiriendo nuevas herramientas matemáticas, como la teoría del caos determinista [30] o la de redes complejas [31].

Por otro lado, la psicología cuantitativa ha utilizado siempre las matemáticas para cuantificar y diseñar modelos de la conducta. En los últimos años, este campo se ha beneficiado de los resultados de la Ciencia de los Sistemas Complejos. Uno de los ejemplos más estudiados es la influencia del ruido en la percepción. Existe un fenómeno denominado *resonancia estocástica* en el que la adición de ruido a un estímulo puede ayudar a reconocerlo y se ha observado experimentalmente en animales [32] y en humanos [33].

Son muchos otros los ejemplos de aplicación de nuevas técnicas matemáticas al comportamiento humano: la teoría de colas [34], el estudio de fluctuaciones [35], etc. Todos ellos, en mayor o menor medida, nos dan información acerca de cómo trabaja el cerebro y pueden ayudar a descubrir las estrategias de diseño de nuestros circuitos neuronales.



Conclusiones: las neurociencias en Madrid

Madrid fue la ciudad donde Santiago Ramón y Cajal descubrió, en 1887, que el tejido nervioso está formado por neuronas y que éstas son células individuales. En los congresos de neurociencias celebrados en Madrid, los mayores expertos del mundo peregrinan hacia el Instituto Cajal en la calle Serrano para ver sus preparaciones originales. La tradición de Cajal ha hecho de Madrid un lugar muy activo en neurociencia: el Instituto Cajal, dirigido por Alberto Ferrús, se centra en aspectos biológicos y experimentales; mientras que el grupo de Neurocomputación Biológica, dirigido por Pablo Varona, y el Neural Processing Lab, dirigido por Gonzalo G. de Polavieja, ambos en la Universidad Autónoma de Madrid, están en la vanguardia de los estudios teóricos y el diseño de modelos de neuronas. No es ésta una lista exhaustiva, puesto que, en nuestra Comunidad Autónoma, son más los investigadores y grupos que se dedican al estudio del cerebro, de la conducta o a las aplicaciones de las redes neuronales, pero sólo los tres mencionados dan una idea de lo activa, destacada y multidisciplinar que es la comunidad neurocientífica madrileña.

Hemos visto cómo en el estudio del cerebro han confluído neurobiología, computación y un gran número de disciplinas matemáticas: lógica, teoría de los sistemas dinámicos, complejidad, caos, teoría de control, estadística y probabilidad, teoría de la información. Esta interacción entre neurociencia, matemática y computación ha sido muy fructífera, especialmente en las últimas tres décadas, y todo indica que el cerebro seguirá siendo una fuente de inspiración para el desarrollo de nuevas aplicaciones informáticas. No sólo la medicina, sino también la industria, va a ser beneficiaria de esta interacción. Madrid se encuentra en un excelente momento para continuar destacando en este campo, siempre que asuma la necesidad de afrontar la investigación de modo marcadamente multidisciplinar, integrando equipos de biólogos, médicos, psicólogos, informáticos, matemáticos y físicos.

Bibliografía

- [1] Hodgkin, A.L.; Huxley, A.F.; *A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve*. Journal of Physiology-London 117, (1952), (500-544).
- [2] Karma, A.; *Spiral Breakup in Model-Equations of Action-Potential Propagation in Cardiac Tissue*. Physical Review Letters 71, (1993), (1103-1106).
- [3] Ma, L.; Wagner, J.; Rice, J.J.; Hu, W.W.; Levine, A.J.; Stolovitzky, G.A.; *A plausible model for the digital response of p53 to DNA damage*. Proceedings of the National Academy of Sciences 102, (2005), (14266-14271).
- [4] Koch, C.; *Computation and the single neuron*. Nature 385, (1997), (207-210).
- [5] Koch, C.; *Biophysics of Computation: Information Processing in Single Neurons* (Oxford University Press, New York, (1999).
- [6] Crick, F.; Koch, C.; *A framework for consciousness*. Nature Neuroscience 6, (2003), (119-126).
- [7] van Steveninck, R.R.D.; Lewen, G.D.; Strong, S.P.; Koberle, R.; Bialek, W.; *Reproducibility and variability in neural spike trains*. Science 275, (1997), (1805-1808).

- [8] Cecchi, G.A.; Sigman, M.; Alonso, J.M.; Martinez, L.; Chialvo, D.R.; Magnasco, M.O.; *Noise in neurons is message dependent*. Proceedings of the National Academy of Sciences 97, (2000), (5557-5561).
- [9] Rieke, F.; Warland, D.; van Steveninck, R.R.D.; Bialek, W.; *Spikes : Exploring the Neural Code* (MIT Press, Cambridge, Mass., (1997).
- [10] Toral, R.; Mirasso, C.R.; Gunton, J.D.; *System size coherence resonance in coupled FitzHugh-Nagumo models*. Europhysics Letters 61, (2003), (162-167).
- [11] Guantes, R.; de Polavieja, G.G.; *Variability in noise-driven integrator neurons*. Physical Review E 71, 011911 (2005).
- [12] Chialvo, D.R.; *Critical brain networks*. Physica A 340, (2004), (756-765).
- [13] McCulloch, W.S.; Pitts, W.; *A logical calculus of the ideas immanent in nervous activity*. Bulletin of Mathematical Biophysics 5, (1943), (115-133).
- [14] Hebb, D.O.; *The Organization of Behavior* (Wiley, Nueva York, 1949).
- [15] Hopfield, J.J.; *Neural Networks and Physical Systems with Emergent Collective Computational Abilities*. Proceedings of the National Academy 79, (1982), (2554-2558).
- [16] Rumelhart, D.E.; McClelland, J.L.; University of California San Diego. PDP Research Group. *Parallel Distributed Processing : Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, Mass., 1986).
- [17] Rowley, H.A.; Baluja, S.; Kanade, T.; *Neural network-based face detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, (1998), (23-38).
- [18] Omidvar, O.; Dayhoff, J.E.; *Neural Networks and Pattern Recognition* (Academic Press, San Diego, CA, 1998).
- [19] Hunt, K.J.; Sbarbaro, D.; Zbikowski, R.; Gawthrop, P.J.; *Neural Networks for Control-Systems - a Survey*. Automatica 28, (1992), (1083-1112).
- [20] Baxt, W.G.; *Use of an Artificial Neural Network for the Diagnosis of Myocardial-Infarction*. Annals of Internal Medicine 115, (1991), (843-848).
- [21] Rost, B.; Sander, C.; *Combining Evolutionary Information and Neural Networks to Predict Protein Secondary Structure*. Proteins-Structure Function and Genetics 19, (1994), (55-72).
- [22] Hsu, K.L.; Gupta, H.V.; Sorooshian, S.; *Artificial Neural-Network Modeling of the Rainfall-Runoff Process*. Water Resources Research 31, (1995), (2517-2530).
- [23] Shadbolt, J.; Taylor, J.G.; *Neural Networks and the Financial Markets : Predicting, Combining, and Portfolio Optimisation* (Springer, London, 2002).
- [24] Pérez, M.; *Artificial neural networks and bankruptcy forecasting: a state of the art*. Neural Computing and Applications 15, 154-163 (2006).
- [25] Hertz, J.; Palmer, R.G.; Krogh, A.S.; *Introduction to the Theory of Neural Computation* (Addison-Wesley Pub. Co., Redwood City, Calif., 1991).
- [26] Vapnik, V.N.; *The Nature of Statistical Learning Theory* (Springer, New York, 2000).
- [27] Parrondo, J.M.R.; Van den Broeck, C.; *Vapnik-Chervonenkis Bounds for Generalization*. Journal of Physics A 26, (1993), (2211-2223).
- [28] Damasio, A.R.; *The Feeling of what Happens : Body and Emotion in the Making of Consciousness* (Harcourt Brace, New York, 1999).
- [29] Grush, R.; *The emulation theory of representation: motor control, imagery, and perception*. En Behavioral and Brain Sciences (Cambridge University Press, 2003).



- [30] Freeman, W.J.; *Neurodynamics: an Exploration in Mesoscopic Brain Dynamics* (Springer, London; New York, 2000).
- [31] Eguiluz, V.M.; Chialvo, D.R.; Cecchi, G.A.; Baliki, M.; Apkarian, A.V.; *Scale-free brain functional networks*. Physical Review Letters 94, (2005), (18102).
- [32] Ward, L.M.; Neiman, A.; Moss, F.; *Stochastic resonance in psychophysics and in animal behavior*. Biological Cybernetics 87, (2002), (91-101).
- [33] Kitajo, K.; Nozaki, D.; Ward, L.M.; Yamamoto, Y.; *Behavioral stochastic resonance within the human brain*. Physical Review Letters 90, (2003), (218103).
- [34] Barabasi, A.L.; *The origin of bursts and heavy tails in human dynamics*. Nature 435, (2005), (207-211).
- [35] Paulus, M.P.; *Long-range interactions in sequences of human behavior*. Physical Review E 55, (1997), (3249- 3256).



Capítulo XVIII

SISTEMAS COMPLEJOS: LA CIENCIA DEL SIGLO XXI

POR ÁNGEL SÁNCHEZ SÁNCHEZ
UNIVERSIDAD CARLOS III DE MADRID
Grupo Interdisciplinar de Sistemas Complejos
Departamento de Matemáticas
28911 Leganés
Correo electrónico: anxo@math.uc3m.es

*"Creo que el próximo siglo será el siglo
de la Complejidad"*

Stephen Hawking, *San Jose Mercury News*,
23 de enero de 2000.

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

This paper is a brief introduction to the research on Complex Systems and its relationship with Mathematics. Complex Systems, systems with very many parts whose interaction gives rise to new emergent phenomena, are receiving very much attention both from public and private institutions, due to the great applicability of this paradigm. The thirty-year history of the field is outlined, and its roots on interdisciplinary approaches to science are discussed, with examples of Complex Systems in Physics, Economy, Sociology and Biology. Accordingly, the prospects for the research on Complex Systems are analyzed and its role in frontier research is identified, particularly as an anti-reductionist, bottom-up approach to understanding different problems.

Finally, a summary of the mathematical tools currently used to study Complex Systems is provided, along with a number of questions and avenues for further developments needed from Mathematics for this research.

¿Qué son los sistemas complejos?

La Ciencia de los Sistemas Complejos o de la Complejidad es un nuevo campo [1] que estudia *comportamientos emergentes* en sistemas con gran número de partes, componentes o agentes, es decir, comportamientos y fenómenos colectivos que aparecen debido a la interacción no predecibles ni entendibles a partir de los individuales. Los sistemas sociales, formados por personas, el cerebro, formado por neuronas, las proteínas, formadas por átomos, el sistema atmósfera-océanos, formado por flujos de aire y agua o los ecosistemas, formados por especies animales y vegetales, son buenos ejemplos de sistemas complejos. Esta disciplina, que tiene un carácter horizontal intrínseco que toca todas las ramas tradicionales de la Ciencia y la Ingeniería, se centra en el estudio de cómo las interacciones dan lugar a los patrones de comportamiento, el diseño de nuevos métodos de describir y analizar sistemas complejos, y el proceso de formación de tales sistemas. Este estudio, inicialmente teórico, se aplica hoy a problemas concretos de gran importancia, desde sanitarios a militares, pasando por la gestión de organizaciones [2].

Un campo de interés público y privado

Los sistemas complejos reciben cada vez más atención pública y privada. Así, en el contexto público europeo, el Programa NEST (*New and Emerging Science and Technology*, nuevo en el 6º Programa Marco y antecedente del flamante *European Research Council*) financia una gran cantidad de iniciativas en el campo, que se coordinan a través de la Acción GIACS [3] y la Red ONCE-CS [4]. Esta actividad ha dado lugar en 2004 a la creación de la *European Complex Systems Society*. Existe una iniciativa ERA-NET en marcha para estudiar un incremento de la

financiación por parte de las agencias de los estados miembros. Se han creado institutos de investigación específicos, como el *Max-Planck-Institut für Physik komplexer Systeme*, de Dresde, Alemania, o el Instituto de Biocomputación y Física de Sistemas Complejos de Zaragoza. En cuanto al sector privado, el interés es grande en Estados Unidos, donde hay instituciones como el *Santa Fe Institute for Complex Systems* [5], financiadas por empresas que van desde la informática (Intel, Sun) a la banca (Credit Suisse First Boston) pasando por los automóviles (Ford, Toyota) o la tecnología (Lockheed Martin). En España, la participación privada es todavía incipiente, pero hay ya entidades económicas y financieras relevantes con interés en el campo.

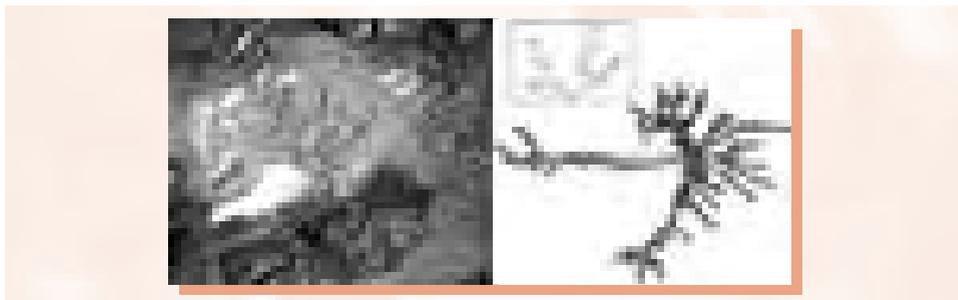
Treinta años de sistemas complejos

En 1972, el premio Nobel de Física Phil Anderson publicó el artículo *More is different* [6], cuyo título se refiere precisamente al concepto de emergencia, y que fue el primero que reconoció el campo de los sistemas complejos como tal. Sin embargo, habría que esperar a los años noventa para que la generalización del acceso a los ordenadores permitiera el desarrollo del estudio de los sistemas complejos, a la vez que herramientas básicas como la Mecánica Estadística y la Ciencia No Lineal alcanzaban su madurez. La *Mecánica Estadística* es originalmente la aplicación de la Estadística, con su bagaje matemático para tratar con poblaciones grandes, a la Mecánica, que se ocupa del movimiento de partículas u objetos sujetos a una fuerza [7]. Sus pioneros fueron J. Willard Gibbs y Ludwig Boltzmann, a finales del siglo XIX, proporcionando un marco para relacionar las propiedades microscópicas de, por ejemplo, átomos o moléculas individuales con las propiedades macroscópicas de los materiales (como por ejemplo la presión o el volumen de un gas). El salto cualitativo de los noventa fue darse cuenta de que el concepto es mucho más general, ya que no necesita aplicarse a átomos o moléculas sino que describe cualquier colectividad de objetos cuya dinámica individual e interacciones mutuas son conocidas, es decir, los sistemas complejos. El mismo espíritu generalista está en la base de la *Ciencia No Lineal*, que va más allá del significado puramente matemático de “no lineal” (véase la reciente y monumental enciclopedia compilada por Alwyn Scott [8]). Simplificando drásticamente y sin pretender ser exhaustivo, podríamos decir que la ciencia no lineal comprende cuatro paradigmas básicos: el caos determinista, las estructuras coherentes, la formación, competición y selección de patrones, y la dinámica adaptativa o evolutiva. El impacto de esos paradigmas se puede entender en términos de su relevancia interdisciplinar. El caos aparece, por ejemplo, en la actividad eléctrica de sistemas biológicos, en la transición de un fluido a la turbulencia, y en el movimiento de planetas. Las estructuras coherentes describen tanto la gran mancha roja de Júpiter como los tsunamis o los sistemas de comunicación óptica. Nos encontramos problemas relacionados con patrones en fenómenos tan distintos como la recuperación de petróleo, las interacciones láser-plasma o la morfogénesis. Y por último, la dinámica evolutiva nos habla de cosas como la evolución biológica, las redes neuronales o la vida artificial.



FIGURA 31

Ejemplos de sistemas complejos



Izquierda: representación esquemática de la Corriente del Golfo. Este sistema complejo natural consta de la combinación de numerosos flujos de agua, que unidos a los flujos atmosféricos dan lugar a la Corriente del Golfo, responsable central del clima de la Europa atlántica. Derecha: red de e-mail de la Universitat Rovira i Virgili. Los distintos niveles de gris representan los departamentos, y los enlaces unen personas que intercambian correo electrónico [tomado de R. Guimerà *et al.*, *Phys. Rev. E* 68, 65103 (2003)]. La figura representa entonces la *red social* del colectivo.

La interdisciplinariedad como fundamento

El carácter distintivo fundamental de la investigación en sistemas complejos es su *interdisciplinariedad*. Como dice Anderson [6], muchas veces los objetos elementales pertenecen a una ciencia (p.ej., la química) y los emergentes a otra (p.ej., la biología). Y lo mismo ocurre con los métodos de la complejidad: suelen caer en el espacio interdisciplinar: “Donde aparece la interdisciplinariedad es precisamente en el salto entre dos niveles de la estructura jerárquica de la Ciencia, y/o en el espacio no colonizado entre dos disciplinas bien establecidas, y/o en la creación de nuevos campos del saber. La investigación interdisciplinar no se refiere pues a un carácter pluri o multidisciplinar que implica una simple superposición de expertos o especialistas en campos diversos que ofrecen sus técnicas para la solución de un problema. Por el contrario, se refiere a la voluntad decidida de cruzar fronteras entre campos establecidos y de forzar la imaginación para transferir conceptos de unos campos a otros, partiendo de que la constatación de que el estudio a través de esas fronteras es fuente de grandes progresos del conocimiento [9].”

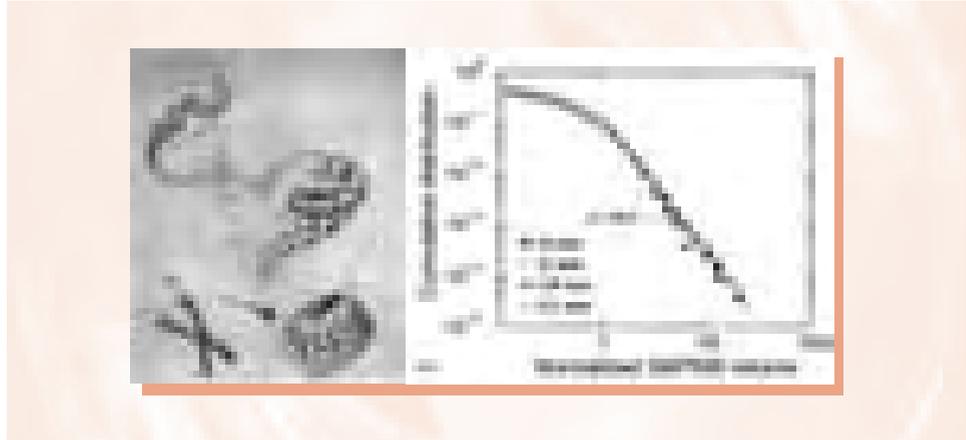
Como muestra de este carácter, se recogen a continuación aplicaciones de los sistemas complejos en distintas ciencias:

Física: Las ideas base de la Complejidad, que partieron de la Física Estadística y No Lineal, se aplican hoy en campos que van de la Nanotecnología a la Computación Cuántica pasando por la Astronomía. Quizá uno de los retos más importantes hoy en Física es el estudio de sistemas desordenados, que casi siempre son sistemas complejos formados por partes heterogéneas. Éste

es un problema en el que se necesitan nuevas matemáticas con urgencia, y avances significativos en esta cuestión serían de inmediato interés para los demás campos, donde generalmente la aproximación de suponer que las partes son iguales es muy pobre.

FIGURA 32

Ejemplos de sistemas complejos



Izquierda: niveles de plegamiento del ADN: la doble hélice se pliega en torno a unas proteínas llamadas histonas, y la estructura resultante sufre más plegamientos hasta formar los cromosomas, que empaquetan unos dos metros de molécula por célula en humanos. Derecha: fluctuaciones de las ganancias en el mercado de Nueva York, que muestra que no son gaussianas como predicen las teorías económicas usuales [tomado de P. Gopikrishnan *et al.*, *Physica A* 287, 362-373(2000)].

Economía: La afirmación de que la Economía es un sistema complejo (*muy complejo*) no resulta sorprendente [10]. Tampoco parece extraño que lo que gustaría a cualquier economista sea resolver el “problema micro-macro”: cómo se explica la Macroeconomía desde la Microeconomía, igual que la Termodinámica es explicada por la Mecánica Estadística. Aparte de esta pregunta fundamental, hay mucho interés hoy en día en modelar los mercados financieros como sistemas complejos, para intentar comprender su funcionamiento (y sacar partido de él, obviamente, o para detectar y evitar malas prácticas). Los problemas de gestión de las grandes empresas globalizadas son otra aplicación directa de los sistemas complejos.

Sociología: Cómo se forman las redes sociales, qué consecuencias conllevan, cómo dan lugar al establecimiento de normas e instituciones, cómo se ven afectadas por esas normas, qué tienen que ver con la cultura y su evolución son preguntas que están hoy en día en el ambiente de la sociología, y hay mucho trabajo por hacer. De hecho, la revista *Science* [11] incluyó el problema de la evolución de la cooperación humana entre los 25 más acuciantes para el siglo XXI. Otros problemas interesantes en el campo son la formación de opiniones y consenso, la dinámica electoral o la segregación de poblaciones (asunto en el que ya trabajó el premio Nobel de

Economía en 2005, Thomas Schelling, desde la perspectiva de los sistemas complejos pero anticipándose a ella [12]).

Biología: La Biología es el arquetipo de ciencia en la que la jerarquía de niveles de complejidad es muy intrincada. Del gran número de cuestiones a que eso da lugar, podemos destacar el funcionamiento del sistema inmunológico o el del cerebro; las redes de regulación genéticas o proteómicas o también ecológicas, o muy recientemente la cooperación en poblaciones bacterianas (*biofilms*). De particular relevancia es el estudio del plegamiento de proteínas: intentar predecir de que forma se va a estructurar en el espacio una secuencia lineal dada de aminoácidos, campo donde los sistemas complejos ya están haciendo aportaciones.

Hacia la frontera del futuro

Como hemos visto, muchos sistemas de distintos campos no pueden entenderse con la aproximación *top-down* o reduccionista habitual. Los sistemas complejos son a menudo impredecibles, pero cumplen su función incluso perturbados o sometidos a ruido, y aparecen no por un diseño ordenado, sino por un proceso de auto-organización, típicamente adaptativa. La Ciencia y la Ingeniería “tradicionales” carecen del marco conceptual para entender diseñar, o gestionar estos sistemas. La Complejidad viene a actuar de puente entre campos, no sólo entre la Ciencia y la Ingeniería, sino entre las Ciencias formales y naturales y las Ciencias sociales. Uno de sus propósitos fundamentales es, por tanto, el ayudarnos a prever las consecuencias de nuestras acciones cuando tratamos de diseñar o controlar los sistemas complejos que nos rodean. Así, nuevas oportunidades de negocio surgen de las redes P2P, mientras que Internet estimula maneras alternativas de relacionarse en sociedad. Sin embargo, no todo son buenas noticias, y esta interdependencia global, que abre oportunidades de innovación sin precedentes, trae consigo riesgos de enorme magnitud: somos incapaces de asegurar la eficiencia de internet (más del 68% del correo electrónico es “no deseado”), nos enfrentamos a epidemias globales (SIDA, SARS, gripe aviar) y luchamos para entender el medio ambiente y conseguir gestionarlo de manera sostenible a la vez que generamos cada vez mayores amenazas contra él. Afrontar estos problemas exige una nueva ciencia que, partiendo de una perspectiva *bottom-up*, sienta las bases de su solución a través del entendimiento de un sistema enormemente complejo del que nosotros mismos formamos parte. En el futuro, la innovación, la mayor fuente de crecimiento, sólo será posible si comprendemos cómo funcionan los “sistemas de sistemas”.

Estamos, pues, en el vértice de un cambio radical de paradigma en Ciencia que tendrá efectos profundos en todos los aspectos de la vida. La Ciencia de los sistemas complejos proporcionará nuevas maneras de coordinar sistemas humanos, de diseñar arquitecturas robustas de sistemas tecnológicos, y estimulará grandes novedades en las tecnologías de la información y las comunicaciones. Es decir, será la base de nuevas generaciones de tecnología en las próximas décadas, abriendo una oportunidad para la que Europa empieza a posicionarse y que España no puede dejar escapar.

Matemáticas y computación, de y en la frontera

En el corazón mismo de esta nueva Ciencia de los sistemas complejos están las Matemáticas y la Simulación, sin las cuales no será posible avanzar en ninguno de los tres aspectos de la I+D+i. De hecho, el paradigma tradicional del modelado en *Matemática Aplicada* [14], con sus tres fases: *formulación*, o traducción del problema científico a términos matemáticos; *análisis*, o solución matemática del modelo, e *interpretación y verificación* de la solución en términos del problema original, es precisamente el *modus operandi* de la investigación en sistemas complejos. En el proceso de modelado, la interacción con el ordenador abre un mundo de posibilidades, ya que “con el uso juicioso de los ordenadores podemos penetrar en nuevas áreas y descubrir conexiones entre campos diversos de las matemáticas que escaparon a nuestros predecesores [15]”. Sólo ese maridaje entre Matemáticas y Simulación permitirá avances sustanciales en el entendimiento de los sistemas complejos, un maridaje que, sin pretender ser exhaustivos y sin ningún orden particular, involucra:

- Un equivalente conceptual al Análisis, una manera de ver las consecuencias de la multitud de interacciones que definen un sistema complejo y que traslade nociones del continuo al discreto [16].
- Teoría de grafos, como base matemática para entender las redes de interacción subyacentes a todo sistema complejo [17, 18].
- Procesos estocásticos, cadenas de Markov, matrices aleatorias, y métodos novedosos de tratar con sistemas desordenados [19].
- Análisis funcional, para tratar problemas como las propiedades espectrales de operadores, cruciales en el estudio de transiciones de fase [20].
- Una formalización matemática de las nuevas técnicas de simulación numérica desarrolladas para los sistemas complejos, en particular las basadas en agentes, Monte Carlo y las simulaciones multiescala, incluyendo técnicas de homogeneización [21, 22].
- Teoría de control, como punto de partida para elaborar las matemáticas del diseño de sistemas con propiedades emergentes requeridas en aplicaciones concretas [23, 24].
- Sistemas complejos adaptativos, con el ingrediente añadido de la evolución para adaptarse, bien al ambiente, bien al comportamiento global del sistema, bien a ambos, cuyo estudio involucra muchas Matemáticas clásicas, desde los sistemas Lotka-Volterra a la teoría de juegos evolutiva pasando por los problemas de reacción difusión, y su relación con los algoritmos genéticos [25, 26].
- Estadística, incluyendo nuevos algoritmos de inferencia bayesiana en redes, y técnicas de tratamiento de cantidades masivas de datos (*data mining*) [27].
- Teoría de la señal, identificación de sistemas lineales y no lineales, problemas inversos, análi-



sis de series temporales (algoritmos y nuevas herramientas matemáticas más allá del análisis armónico, redes neuronales) [28, 29].

- Teoría de bifurcaciones, sistemas dinámicos, formas normales como modelos genéricos de clases de fenómenos, técnicas asintóticas de auto-semejanza (grupo de renormalización) [30, 31].
- Complejidad computacional, con especial énfasis en la complejidad algorítmica, base de la clasificación de algoritmos y por tanto de modelos para la simulación de sistemas complejos [32].
- Modelos formales de programación orientada a objetos basados en UML (*Unified Modelling Language*) como lenguaje alternativo para entender matemáticamente los sistemas complejos [33].

A modo de conclusión

La nueva ciencia de la Complejidad, desarrollada fundamentalmente en los años noventa para el estudio de sistemas compuestos por muchas unidades en interacción y los comportamientos emergentes resultantes, está aquí para quedarse, más aún, será una pieza clave de la I+D+i del siglo XXI. Matemáticas y Computación se entrelazan inextricablemente en la investigación en Complejidad, y sólo una nueva perspectiva que involucre ambas trabajando desde la frontera entre ellas y entre las ciencias establecidas permitirá aprovechar la ventana de oportunidad que tenemos en este momento para transferir ese esfuerzo investigador al desarrollo y a la innovación.

Bibliografía

- [1] Waldrop, M. A.; *Complexity: The Emerging Science at the Edge of Order and Chaos*. Simon & Schuster, (1992).
- [2] Bar-Yam, Y.; *Making Things Work: Solving Complex Problems in a Complex World*. Knowledge Press, (2005).
- [3] <http://www.giacs.org>.
- [4] <http://www.once-cs.net>
- [5] <http://www.santafe.edu>.
- [6] Anderson, P.W.; *More is different*. Science 177, (1972), (393-396).
- [7] Hay una infinidad de libros sobre Mecánica Estadística, pero uno reciente, relativamente sencillo y muy auto-contenido es el del físico matemático D. C. Mattis. *Statistical Mechanics Made Simple*. World Scientific, (2003).
- [8] Scott, A.; editor. *Encyclopedia of Nonlinear Science*. Taylor and Francis, (2005).
- [9] San Miguel, M.; Interdisciplinarietà: Comentarios desde la perspectiva de un físico. *El papel social de la ciencia en baleares: Un homenaje a Javier Benedí*. Eds. C. Duarte y F. Grases, Universitat Illes Balears, 2003, (235-250).
- [10] Anderson, P.W.; Arrow, K.J.; Pines, D.; editores. *The Economy as an Evolving Complex System*. Addison-Wesley, (1988).
- [11] Pennisi, E.; *How did cooperative behavior evolve?* Science 309, (2005), (93).

- [12] Schelling, T.C.; *Micromotives and Macrobehavior*. W. W. Norton and Co., (1978).
- [13] En http://gisc.uc3m.es/_anxo/math.html se puede consultar una lista de centros de Matemáticas donde se trabaja en estos campos.
- [14] Lin, C.C.; Segel, L.; *Mathematics Applied to Deterministic Problems in the Natural Sciences*. SIAM Classics in Applied Mathematics vol. 1, (1998).
- [15] Zabusky, N.; *Computational synergetics and mathematical innovation*. J. Comput. Phys. 43, (1981), (195).
- [16] Strogatz, S.; *Sync: The Emerging Science of Spontaneous Order*. Hyperion, (2003).
- [17] Newman, M.E.J.; *The structure and function of complex networks*. SIAM Review 45, (2003), (167-256).
- [18] Boccaletti, S.; Latora, V.; Moreno, Y.; Chavez, M.; Hwang, D.-U.; *Complex networks: Structure and Dynamics*. Physics Reports 424, (2006), (175-308).
- [19] García-Ojalvo, J.; Sancho, J.M.; *Noise in Spatially Extended Systems*. Springer, (1999).
- [20] Cuesta, J.A.; Sánchez, A.; General non-existence theorem for phase transitions in one-dimensional systems with short range interactions, and physical examples of such transitions. *J. Stat. Phys.* 115, (2004), (869-893).
- [21] Siettos, C.I.; Rico-Martinez, R.; Kevrekidis, I.G.; A systems-based approach to multiscale computation: Equation-free detection of coarse-grained bifurcations. <http://arxiv.org/abs/nlin.CG/0510007>.
- [22] Grimm, V.; *et al.* Pattern-oriented modeling of agent based complex systems: Lessons from ecology. *Science*, 310, (2005), (987-991).
- [23] Astrom, K.J.; Albertos, P.; Blanke, M.; *Control of Complex Systems*. Springer, (2000).
- [24] Capasso, V.; Periaux, J.; eds. *Multidisciplinary Methods for Analysis Optimization and Control of Complex Systems vol. 6*. Springer, (2005).
- [25] Levin, S.A.; Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bull. Am. Math. Soc.* 40, (2002), (3-19).
- [26] Hofbauer, J.; Sigmund, K.; *Evolutionary Games and Population Dynamics*. Cambridge University Press, (1998).
- [27] Hastie, T.; Tibshirani, R.; Friedman, J.H.; *The Elements of Statistical Learning*. Springer, (2003).
- [28] Kantz, H.; Schreiber, T.; *Nonlinear Time Series Analysis*. Cambridge University Press, (2003).
- [29] Neelakanta, P.S.; *Information Theoretic Aspects of Neural Networks*. CRC Press, (1999).
- [30] Kaneko, K.; Tsuda, I.; *Complex Systems: A Constructive Approach with Applications in Life Sciences*. Springer, (2000).
- [31] Goldenfeld, N.; *Lectures on Phase Transitions and the Renormalization Group*. Addison-Wesley, (1992).
- [32] Flake, G.W.; *The Computational Beauty of Nature*. M.I.T. Press, (1998).
- [33] Gnesi, S.; Formal specification and verification of complex systems. *Electronic Notes in Theoretical Computer Science* 80, (2003). <http://www.elsevier.nl/locate/entcs/volume80.html>.



Capítulo XIX

CURVAS Y SUPERFICIES ALGEBRAICAS:
COMPUTACIÓN Y APLICACIONES

POR RAFAEL SENDRA PONS
UNIVERSIDAD DE ALCALÁ
Departamento de Matemáticas
Facultad de Ciencias
CP 28871
Correo electrónico: rafael.sendra@uah.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

The research activity in this area follows three different, but related, lines: curves and surfaces are the main objects to be studied, the computation and algorithmic treatment of them is the working instrumental tool, and their applications the final goal. Although, the most intuitive way of understanding the notion of algebraic curve and surface is, probably, by its graphic representation, the concept is usually introduced as the set of solutions, in the plane or in the space (in general, the space dimension can be higher than 3), of one or several polynomial equations (its implicit equations). Plane curves (curves in the plane) are defined by a single polynomial in 2 variables. Space curves (now the curves lies in the space) are defined by at least two polynomials in 3 variables. Surfaces (in the 3-dimensional space) are defined by a polynomial in 3 variables. Thus, an space curve can be seen as the intersection of several surfaces.

In the 20th century, with the appearing of computers, many branches of Mathematics moved part of their research goals to more computational statements, even appearing new lines of research. As a consequence, an interesting phenomenon has been generated: on one hand, the mathematicians study the resolution of computational problems from their field, and on the other, the technological development provides new problems being a real research challenge for mathematicians. Research on curves and surfaces has not been an exception to this process.

Previously, the computation with curves and surfaces was non-viable, with the exception of simple examples. However, with the development of algorithmic methods and with the help of computers, the situation has changed transforming curves and surfaces as a powerful tool in applications.

Historically, curves and surfaces have played an instrumental distinguished role in solving problems, not only in Mathematics, but also in other branches of science such as Physics, Biology, Chemistry or Engineering. However, the development of computers, and the expansion of hardware and software in the industrial word, have implied a considerable increment of the applicability of these type of geometric objects; mainly provided by the necessity of automatizing the industrial processes. A clear example of this assertion is the computer aided geometric design (CAGD) that, in some sense, appears as a response of the requirement of automatizing geometric models in automobile and aeronautic industry. One may also mention other applications of curves and surfaces as codification, cryptography, computer graphics, etc.

Future work in this field constitutes a real challenge for mathematicians. There exist yet many questions related to curves and surfaces, and to their applications, that require algorithmic answers. Also, some of the current algorithmic solutions involve a high computational complexity, and therefore the feasibility in practice of some methods is not so clear. Nevertheless, an special mention in this brief abstract should be done to the new evolution that this research field faces, and that is known as "approximated algorithms, and hybrid symbolic-numeric algorithms".

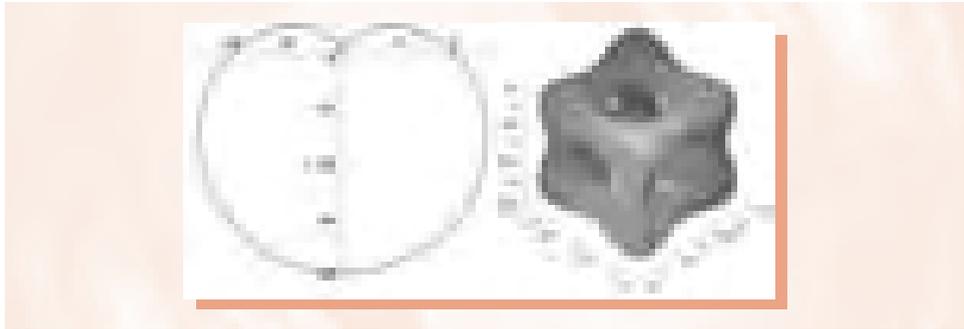
Introducción

Este trabajo pretende dar una visión panorámica de la investigación en el ámbito de las curvas y las superficies algebraicas, con especial énfasis en su manipulación algorítmica y aplicabilidad, y está dirigida a un espectro amplio de lectores, no necesariamente especialistas en Matemáticas.

Parece razonable comenzar con una breve descripción del marco temático en el que se ubica esta disciplina. Se trata de un ámbito de actuación con tres facetas interrelacionadas: las curvas y superficies son los objetos principales de estudio, la computación o algoritmia el instrumento de trabajo, y las aplicaciones el objetivo. A continuación se desarrollan brevemente cada uno de estos aspectos y se muestran algunos problemas concretos.

FIGURA 33

Izquierda: curva plana $(x^2 + 4y + y^2)^2 - 16x^2 - 16y^2 = 0$
 Derecha: superficie $(x^2 - 1)^2 + (y^2 - 1)^2 + (z^2 - 1)^2 - 3/2 = 0$



Contexto matemático: curvas y superficies algebraicas

El estudio de este tipo de objetos geométricos tiene una larga y notoria historia dentro de las Matemáticas y ocupa un lugar destacado dentro del ámbito mas general de la geometría algebraica, constituyendo un campo activo de investigación, tanto desde el punto de vista teórico como práctico. Aunque, posiblemente, la manera más intuitiva de visualizar una curva o superficie es mediante su representación gráfica (véase Figura 33), el concepto de curva o superficie algebraica se suele introducir como el conjunto de soluciones, en el plano o el espacio², de uno o varios polinomios (sus ecuaciones implícitas). Las curvas planas (curvas en el plano) están definidas por una ecuación polinomial en 2 variables. Las curvas espaciales (la curva está en el espacio) están definidas por al menos dos ecuaciones polinomiales en 3 variables. Las superficies (en el espacio tridimensional) están definidas por una ecuación polinomial en 3 variables; por tanto, una curva espacial se puede ver como intersección de varias superficies (véase Figura 34).

² En general, la dimensión del espacio puede ser superior a 3.



Como ejemplos ilustrativos de los conceptos anteriores obsérvese que el punto del plano $(1, 0)$ pertenece a la curva plana de ecuación $x^2 + y^2 = 1$ (circunferencia unidad) y que el punto del espacio $(1, 0, 0)$ está en la superficie $x^2 + y^2 + z^2 = 1$ (esfera unidad).

Otros ejemplos de curvas algebraicas son las rectas (polinomios de grado 1), circunferencias, elipses, parábolas, hipérbolas (polinomios de grado 2). Asimismo, en el caso de superficies, se pueden mencionar como ejemplos los planos (polinomios de grado 1), esferas, elipsoides, cilindros (polinomios de grado 2). En la Figura 33 aparecen representadas una curva y una superficie, ambas de grado 4; obsérvese que en ambos casos los polinomios son de grado 4.

FIGURA 34

Curva espacial definida por $\{x^2 + y^2 + z^2 = 1, x^3 + z = 1/2\}$



El papel de la computación y la algoritmia

Durante el siglo XVIII y principios del XIX, muchos matemáticos eran verdaderos expertos calculando, así por ejemplo Gauss en 1801 abandonó temporalmente sus investigaciones en aritmética y teoría de números para calcular la órbita del asteroide Ceres, recientemente descubierto. Durante el siglo XIX, la investigación en matemáticas cambió su enfoque, centrándose más en aspectos cualitativos. En el siglo XX, con la aparición de los ordenadores y el desarrollo del software matemático, muchas ramas de la matemática dirigen parte de su labor científica hacia planteamientos computacionales, apareciendo incluso nuevas líneas de investigación. Como consecuencia de este hecho se produce un fenómeno interesante: por una parte el matemático se plantea la resolución de problemas computacionales propios del área y, por otra, el avance tecnológico genera problemas nuevos que constituyen un reto científico. Como muestra, cabe indicar la criptografía y en particular el criptosistema RSA, desarrollado en 1976 por R. Rivest, A. Shamir y L. Adleman, y que se basa en la complejidad algorítmica del problema de factorizar números enteros. La investigación en curvas y superficies no es una excepción en este proceso y en la actualidad existe una actividad científica importante en esta dirección. Previamente, los cálculos con curvas y superficies eran inviables, salvo en

ejemplos sencillos. Sin embargo, con el desarrollo de los métodos algorítmicos y con la ayuda de los ordenadores, esta situación ha cambiado, dando mayor potencialidad a las aplicaciones. En particular, cabe señalar el desarrollo de algoritmos para la obtención de propiedades intrínsecas (invariantes), el estudio de singularidades (nudos, auto-intersecciones, etc), la determinación de intersecciones, la obtención de ecuaciones canónicas de representación, la automatización en cambios de representación (problemas de implícitación, parametrización, inversión, etc), el estudio de aspectos numéricos, el análisis algorítmico de formas (plotting en pantalla), etc. (véase [3], [5]).

Aplicaciones

Históricamente, las curvas y superficies han desempeñado un papel destacado como instrumento en la resolución de problemas, no sólo en la propia matemática, sino también en otras ramas de la ciencia como la física, la biología, la química o la ingeniería. No obstante, como consecuencia del desarrollo de los ordenadores y de la implantación de los medios informáticos en el mundo empresarial e industrial, se ha provocado un aumento considerable en el potencial aplicado de este tipo de objetos geométricos; en gran parte originado por la necesidad o voluntad de automatizar informáticamente los procesos industriales. Un ejemplo claro de esta afirmación es el diseño geométrico asistido por ordenador (CAGD) que, en cierta forma, nace como respuesta al requerimiento de automatización en el modelado de formas geométricas en la industria automovilística y aeronáutica (véase [1], [4]). Así, en 1962, P. Bézier y, en 1959, P. de Casteljaou, desarrollaron paralela e independientemente las curvas de Bézier, el primero para Renault y el segundo para Citroën. Dentro del CAGD cabe señalar aplicaciones concretas como la utilización de las curvas y superficies offsets en análisis de tolerancia, troquelados, etc, las curvas y superficies blending en la modelación de volúmenes, el seccionado de superficies, o el cálculo de errores de forma (véase [2], [3], [4]). Por último mencionar que, además de las aplicaciones en CAGD, existen otras aplicaciones importantes como son la codificación, la criptografía o la visualización gráfica.

Primer ejemplo concreto: procesos de offsetting

Le propongo al lector el siguiente experimento sencillo. Se coloca una moneda de radio d sobre una circunferencia de radio r , donde $d \neq r$, de forma que el centro de la moneda esté sobre la circunferencia. A continuación, se desplaza la moneda a lo largo de la circunferencia, siempre manteniendo su centro sobre la circunferencia, y en cada instante nos fijamos en los “polos norte y sur” de la moneda en la dirección de la normal a la circunferencia (es decir en la dirección de la recta perpendicular a la recta tangente a la circunferencia en el centro de la moneda).

Como resultado se tiene que todos los puntos obtenidos mediante este proceso forman otras dos circunferencias, a distancia d de la circunferencia inicial. Este proceso se puede repetir substituyendo la circunferencia por cualquier curva dando lugar al proceso de offsetting (véase Figura 35). La curva, o curvas resultantes, se denominan curvas offsets a distancia d de la curva motriz (o,



en la terminología clásica de Leibnitz, curvas paralelas). Quizas, la forma más intuitiva de visualizar una curva offset es pensar en una carretera de doble sentido, tomar como curva inicial la curva divisoria de carriles y como offset las curvas que separan la calzada del arcén. Ahora, el mismo proceso se puede repetir tomando una superficie y una bola cuyo centro se desliza por la superficie (véase Figura 35). Surgen así las superficies offset a distancia el radio de la bola.

FIGURA 35

Arriba: Offset de la parábola; Abajo: Offset del paraboloido



Las curvas y superficies offsets (cuyo estudio fue reactivado en los años 60, del siglo pasado, desde el ámbito aplicado) juega un papel fundamental en muchas aplicaciones en el contexto del CAGD. Así, por ejemplo, aparecen en: tolerancia geométrica (las piezas son aceptadas en función del número de mediciones situadas en la región offset del contorno), troquelado de formas (con un troquel circular o esférico que se desplaza por la curva/superficie patrón), robótica (diseño de trayectorias de máxima seguridad), modelización geométrica (el volumen se modela posicionando la esfera en el esqueleto de la estructura), etc.

FIGURA 36

Regla de curvas francesas



Aparentemente, puesto que la offset es una curva/superficie, y teniendo el cuenta el desarrollo algorítmico existente para la manipulación de curvas y superficies, no debería haber dificultades para trabajar con este tipo de objetos. No obstante el proceso de offsetting, aun partiendo de un objeto sencillo, genera curvas y superficies con ecuaciones extremadamente complicadas para las que resulta prácticamente inviable la aplicación de las herramientas estándar. Así por ejemplo, si se parte del Folium de Descartes, que es la curva definida por la ecuación $x^3 + y^3 - 3xy = 0$ (obsérvese que el polinomio sólo tiene grado 3 y 3 términos), la correspondiente offset tiene grado 14 y 114 términos. Para abordar esta dificultad, la investigación se dirige hacia la construcción de algoritmos que permitan deducir propiedades de las offset a partir de propiedades de la curva/superficie origen, sin necesidad de manejar o conocer la ecuación de la offset. En este sentido ya ha habido avances importantes en: la determinación de representaciones paramétricas sencillas, la obtención de auto-intersecciones, el análisis de formas paralelas, etc.

Segundo ejemplo concreto: procesos de blending

Imagínese que se quiere dibujar una silueta plana con formas curvas. Una solución “casera” podría ser utilizar una regla de curvas francesas (véase Figura 36), uniendo con cierta destreza cada arco de curva para evitar la aparición de “picos”. Ahora, imagínese que el trazado de la silueta debe ser automatizado y que, además, la silueta no es plana sino que corresponde al contorno de un volumen. Aparece así el problema del blending. Es decir, dados varios trozos tridimensionales en el espacio (conocidos a partir de sus ecuaciones), se trata de unirlos con suavidad usando nuevos trozos (que vendrán dados por nuevas ecuaciones).

En terminología matemática el problema es como sigue (véase Figura 37): dadas las ecuaciones de varias superficies (superficies primarias), dada una curva en cada una de las superficies primarias (curvas de unión), que viene expresada como intersección de las superficies primarias con otras nuevas superficies (superficies secundarias), se quiere determinar las ecuaciones de una nueva superficie (superficie blending) que se una con suavidad, en cada curva de unión, con cada superficie primaria.

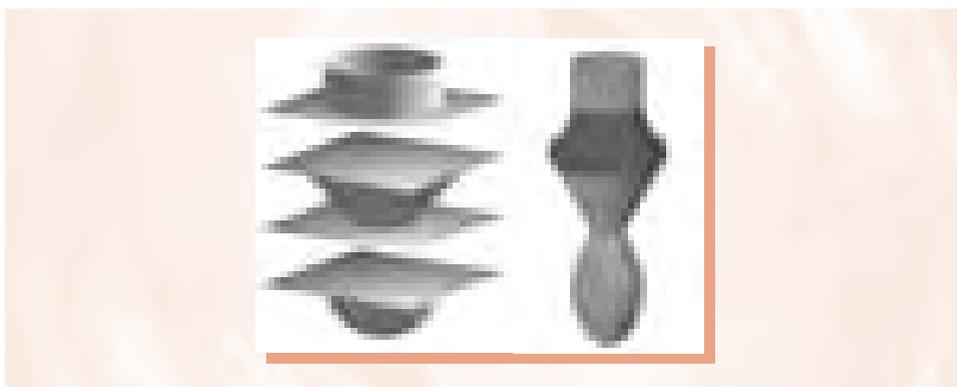
Este tipo de problemas aparece de forma natural en la fabricación de piezas y especialmente en la industria automovilística y aeronáutica. Inicialmente, el problema había sido abordado utilizando curvas (para el caso plano) y superficies (para el caso espacial) sencillas. Sin embargo, con el desarrollo de la algoritmia y la utilización de los ordenadores, se dispone de mayor versatilidad en la gama de curvas y superficies a utilizar en el modelado de los trozos primarios, consiguiéndose así resultados mas satisfactorios. El avance en este contexto ha permitido demostrar que todas las posibles soluciones al problema (es decir todas las posibles superficies blending correspondientes a unos datos fijados) están fuertemente relacionadas y estructuradas (si los datos de partida vienen dados por ecuaciones implícitas, las soluciones forman un ideal del que se conoce una base de Gröbner y si los datos vienen dados por ecuaciones paramétricas, las soluciones forman un modulo libre de rango 3 del que se conoce una base). Como consecuencia de esta estructura



del conjunto de soluciones, los algoritmos existentes permiten dar “todas” las soluciones del problema en función de parámetros; parámetros que posteriormente son utilizados por el diseñador para seleccionar la solución mas idónea al problema concreto con el que esté trabajando.

FIGURA 37

Izquierda: superficies primarias y secundarias en un proceso de blending
Derecha: Resultado del proceso de blending



Perspectivas y líneas futuras

Ciertamente las perspectivas científicas en este ámbito son amplias y prometedoras. Existen todavía muchas cuestiones propias de las curvas y superficies, y de sus aplicaciones, que requieren respuesta algorítmica³. Asimismo, algunas de las soluciones existentes tienen gran complejidad de cómputo haciéndolas inviables en muchas aplicaciones⁴. En el ámbito de la manipulación de

³ Por ejemplo, en la práctica se suele trabajar con una doble representación de la curva o superficie: representación implícita (véase Sección 2) y paramétrica (las variables vienen despejadas en función de parámetros). La obtención de dichas representaciones ha sido resuelta algorítmicamente en los últimos años. Sin embargo en los procesos aplicados, al trabajar con dos representaciones distintas (por ejemplo al calcular intersecciones) se generan nuevas parametrizaciones, cuyos coeficientes ya no son números reales sino complejos. Esta nueva representación, aunque paramétrica, es poco útil a nivel práctico. Se trata por tanto de reparametrizar el dato obtenido para conseguir una representación con coeficientes adecuados.

⁴ Para resolver algunos problemas computacionales se utilizan técnicas globales como las bases de Gröbner. Se trataría de desarrollar herramientas dirigidas a situaciones especiales, derivadas del contexto aplicado que, aunque menos potentes, actúen de forma mas eficiente.

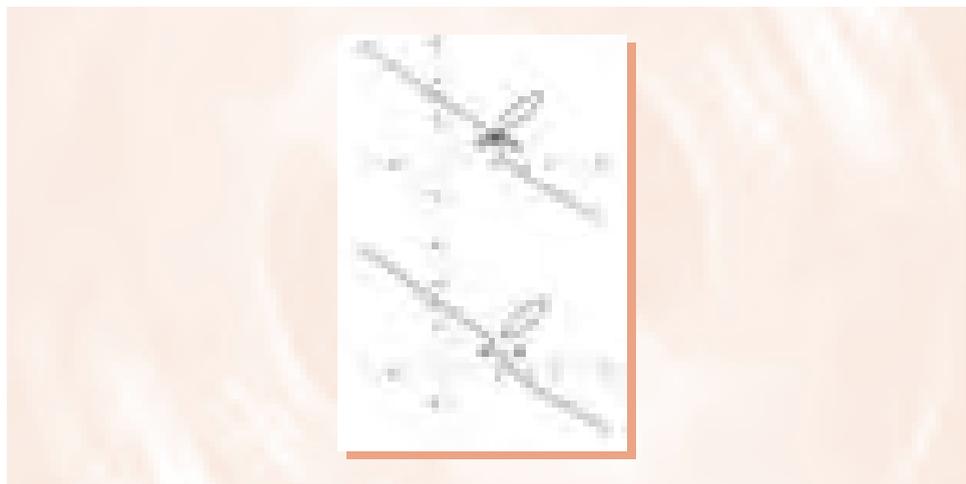
offsets se puede mencionar, entre otros, la determinación de intervalos de seguridad para la distancia d de offsetting; es decir, intervalos en los que, si d toma valores, la forma de la curva/superficie inicial coincide (duplicada) con la de su offset. En relación a procesos de blending, se puede señalar el diseño de estrategias para la elección adecuada de las curvas de unión.

No obstante, quizás, merezca una mención especial, en este breve resumen, la nueva evolución a la que se enfrenta el área y que se ha dado en llamar “algoritmos aproximados o algoritmos híbridos simbólico-numéricos”.

En muchas aplicaciones prácticas resulta insuficiente la utilización de los métodos y algoritmos derivados de un planteamiento teórico que asume la exactitud de los datos, ya que con frecuencia la información de la que se dispone ha sufrido alguna perturbación inicial. Aparece así la necesidad de desarrollar algoritmos aproximados.

FIGURA 38

Arriba: curva C
Abajo: curva C'



Para entender mejor la situación, les propongo un ejemplo. Supóngase que, debido a la naturaleza del problema aplicado se sabe que el resultado buscado se corresponde con una curva C como la que aparece en la Figura 38 (arriba), es decir con un triple lazo⁵. No obstante, como resultado de los procesos previos, o como consecuencia de mediciones, se recibe la curva C' que aparece en la Figure 6 (abajo). La cuestión es como recuperar la curva C a partir de la curva C' . Dar una respuesta exacta a este problema resulta imposible pues en la práctica los datos son

⁵En este caso, la importancia no reside en el hecho de tener un triple lazo, sino en ser una curva parametrizable; propiedad que se deduce de la primera afirmación.



insuficientes. Para abordar la cuestión, la estrategia consiste en deducir una curva C'' próxima a C' y tenga el mismo aspecto que C aunque no sea exactamente C .

Evidentemente, este ejemplo de dificultad media se puede (y de hecho esa es la situación real) complicar tanto como se quiera. En términos generales, la descripción del problema es como sigue. Por la entidad del problema aplicado, o del experimento que se está realizando, se sabe que cierto objeto matemático O (en el ejemplo anterior, $O = C$) verifica cierta propiedad P (en el ejemplo anterior, P es el triple lazo). Sin embargo, debido a errores del proceso, el dato de entrada no es O , sino una pequeña perturbación O' de O (en el ejemplo anterior, $O' = C'$). La dificultad aparece al observar que, en la mayoría de las situaciones, la propiedad P ya no se verifica en O' . Entonces, el problema consiste en, dada una tolerancia de trabajo $\varepsilon > 0$ (error permitido), determinar un nuevo objeto O'' próximo a O' que verifique la propiedad P . La noción de proximidad depende del problema concreto que se esté analizando. Sin embargo, en el caso de curvas y superficies algebraicas la proximidad se introduce exigiendo que O'' se encuentre en la región offset de O' (véase Sección sobre procesos de offsetting), a distancia ε , y viceversa.

Bibliografía

- [1] Bajaj, C.; (ed.). *Algebraic Geometry and its Applications*. Springer-Verlag, Berlin Heidelberg New York. (1994), (154).
- [2] González-Vega, L.; Necula, I.; Pérez-Díaz, S.; Sendra, J.; Sendra, J.R.; *Algebraic Methods in CAGD: Theoretical and Practical Applications*. *Geometric Computation*. Lecture Notes on Computing, Vol. 11. Chap. 1, World Scientific Publishing Co., Singapore. (2004), (1-33).
- [3] Hoffmann, C.M.; Sendra J.R.; Winkler, F.; (eds.) *Special Issue on Parametric Algebraic Curves and Applications*. *Journal of Symbolic Computation*, vol. 23. (1997).
- [4] Hoschek, J.; Lasser, D.; *Fundamentals of Computer Aided Geometric Design*. A.K. Peters Wellesley MA. Ltd. (1993).
- [5] Sendra, J.R.; *Rational Curves and Surfaces: Algorithms and Some Applications*. *Geometric Computation*. Lecture Notes on Computing, Vol. 11. Chapter 3, World Scientific Publishing Co., Singapore. (2004), (63-125).

Capítulo XX

SISTEMAS DINÁMICOS

POR CARLES SIMÓ TORRES
UNIVERSITAT DE BARCELONA
Departament de Matemàtica Aplicada i Anàlisi,
Facultat de Matemàtiques
Gran Via de les Corts Catalanes, 585, 08007 Barcelona
Correo electrónico: carles@maia.ub.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

The goal of Dynamical Systems, in a wide sense, is to study “everything which moves”, this is, all the phenomena in which there is some magnitude which evolves in time.

In this note we review some historical facts, the concepts of phase and parameter spaces and the multiple relations with other areas of sciences. Furthermore we emphasize that the milestones in the study of (any family of) Dynamical Systems are the objects which are invariant under the action of the dynamics, how these objects are connected and how they evolve when parameters change. The role of the computational aspects is shortly described. The note closes with a perspective on the possibilities and the problems to be faced in the near future.

Introducción

En un sentido amplio el objetivo de los Sistemas Dinámicos (S.D.) es estudiar “todo lo que se mueve”, es decir, todos los fenómenos en los que hay alguna magnitud que evoluciona con el tiempo.

Contienen sistemas cuya evolución viene regida por ecuaciones diferenciales ordinarias (EDO), ecuaciones en derivadas parciales (EDP), ecuaciones con retardo o memoria (EDR), sistemas discretos, etc. Asimismo puede considerarse el efecto de términos estocásticos. Sin embargo la teoría de los S.D. contiene importantes diferencias de enfoque respecto a lo que clásicamente se engloba en la teoría de ecuaciones diferenciales.

Poincaré fue el iniciador de los S.D. Dado que es “imposible” obtener fórmulas explícitas para las soluciones de esas ecuaciones, propuso dar la vuelta al problema y considerar no sólo algunas soluciones sino la totalidad de las mismas. Pero intentando describir propiedades cualitativas en lugar de cuantitativas. Inicialmente entroncaba con la Mecánica Celeste y la Mecánica Analítica, pero hoy en día su metodología influye en muchos más dominios, como veremos.

Un objeto fundamental en todo S.D. es el “espacio de estados” o espacio de fases del sistema, E . En un instante dado la posición en este espacio caracteriza el estado del sistema. En ejemplos elementales ese espacio puede ser \mathbb{R}^n o una variedad finito-dimensional. Pero en el caso de EDP, EDR, con componentes estocásticas y otros, E es un espacio funcional, o un espacio de medidas, etc.

En la actualidad, y dada la importancia de las aplicaciones, se intenta de nuevo retomar los aspectos cuantitativos en la teoría de los S.D. Mi punto de vista es que no se pueden dissociar unos de otros y la combinación de ambos es muy fructífera.

Por otra parte un sistema puede depender de parámetros, pertenecientes a un espacio P . Pueden ser constantes físicas, de control, etc. Cómo varían las propiedades es importante, tanto para entender cambios en la fenomenología como para diseñar estrategias de control. En realidad es

siempre conveniente considerar el espacio producto $E \times P$ como el más adecuado para comprender el sistema.

Veremos someramente la relación con otras áreas de la Matemática, algunos de los problemas y métodos para resolverlos. Haremos también énfasis en aspectos computacionales y en perspectivas de futuro.

En aras de la brevedad sólo se da una referencia bibliográfica, que contiene una exposición técnica sobre varios aspectos de los S.D. y algunas referencias.

Relación con otras áreas

Pocas áreas de conocimiento se relacionan con tantas otras como los S.D. Dentro de la Matemática no existe prácticamente ninguna área ajena. No sólo eso: ha dado lugar a áreas que hoy en día se desarrollan independientemente.

De manera natural se relaciona con todos los temas de ecuaciones diferenciales y en general del Análisis (teoría de funciones, una o varias variables complejas, análisis armónico, análisis funcional, teoría de la medida, problemas espectrales, problemas inversos, etc). Pero también con las estructuras algebraicas (álgebra multilineal, grupos de transformaciones, teoría de cuerpos,...) y con la teoría de números.

Por otra parte los objetos de interés en S.D. son objetos geométricos en $E \times P$, por lo que existen fuertes conexiones con la geometría diferencial, geometría algebraica y analítica y con múltiples aspectos de la topología. Es especialmente relevante la relación con la teoría de singularidades que se enriquece notablemente con los problemas de S.D. que consideran no sólo los aspectos “estáticos” de las singularidades sino los “dinámicos”.

Aún en sistemas deterministas la descripción de la dinámica debe hacerse en muchos casos mediante existencia de medidas invariantes. Enlaza así con las probabilidades. Especialmente interesante es la teoría ergódica, a caballo entre ambas disciplinas. Conviene señalar también el papel de la estadística en fases del proceso de modelización que da lugar a parte de las ecuaciones estudiadas en S.D.

Mucha información relevante en S.D. no se puede obtener de forma precisa con las actuales herramientas teóricas. Ello obliga a usar métodos numéricos. Los S.D. constituyen un motor para el Análisis Numérico, no sólo en los aspectos que son ahora más clásicos (resolución de ecuaciones diferenciales) sino también en aspectos novedosos como cálculo efectivo de bifurcaciones, de toros invariantes y de todo tipo de variedades invariantes. Esas técnicas deben combinarse con métodos de cálculo simbólico, y no hay que olvidar los temas de la teoría de la complejidad y de la visualización gráfica, que plantean problemas formidables en dimensión elevada.



En cuanto a las aplicaciones los S.D. surgen de manera natural como un nuevo enfoque de la Mecánica clásica. Pero actualmente se aplican a todas las ramas de la Física, desde la cosmología a la cuántica o desde los fluidos al mundo nanométrico. Y, por supuesto, eso repercute en su aplicación a la industria y a la descripción del entorno: meteorología, oceanografía, climatología, dispersión de contaminantes, etc. Más recientemente han entrado en aspectos de la cinética química y en el diseño de moléculas. Se empiezan a aplicar en Biología, Medicina, Economía, si bien en esas áreas muchos modelos son aún poco fiables. Pero en S.D. se dispone de técnicas de análisis de sistemas que no precisan conocer el modelo matemático del fenómeno estudiado y extraen información relevante directamente de las medidas experimentales, si éstas son suficientemente abundantes.

Problemas y métodos

Un problema básico es calcular lo que puede llamarse “el esqueleto del S.D.”, es decir, los objetos geométricos en E que “guían” la dinámica. Són objetos invariantes (O.I.) bajo la acción del sistema. Puede ser que para verlos como invariantes se tengan que usar sistemas de referencia móviles. Los O.I. más simples son los puntos fijos (o, en lenguaje de EDP, estados estacionarios). Ya su cálculo puede presentar enormes dificultades, como son el poder demostrar que se han calculado todos, problema que, en su versión más simple, enlaza con la geometría computacional.

Los siguientes O.I. son las órbitas periódicas y su generalización, los toros invariantes, con 2 o más frecuencias independientes sobre \mathbb{Q} . Para un sistema dado el paso siguiente es estudiar la estabilidad de los O.I. Muy relevantes son los de tipo “hiperbólico” (quizás en algún sentido débil), esto es, con direcciones en las que hay soluciones que se acercan al O.I. y otras en las que se alejan de él. El conjunto de soluciones tendiendo al (alejándose del) O.I. forma la llamada variedad invariante estable, W^s (inestable, W^u).

Las variedades invariantes de los distintos O.I. pueden cortarse. Por supuesto, el corte sólo puede ocurrir entre una W^s y una W^u . Las soluciones en las que se cortan forman las llamadas conexiones homoclínicas (si se cortan las W^s y W^u de un mismo O.I.) o heteroclínicas (si son de O.I. distintos). Esos fenómenos “clínicos” son los responsables, si los cortes son transversales, de la existencia de dinámica impredecible, popularmente conocida como “caos”. El entramado de conexiones actúa como guía de lo que pueden hacer las soluciones del problema. Si dicho entramado es complicado pueden aparecer objetos invariantes que no son variedades, como los llamados atractores extraños en sistemas disipativos.

Además de la estabilidad de un sistema concreto, interesa también estudiar la “robustez” del sistema frente a cambios de los parámetros o, en general, frente a pequeños cambios del S.D. Ello da lugar a la estabilidad estructural y la teoría de bifurcaciones. La teoría de los S.D. permite, en ciertos casos de naturaleza local, alrededor de un O.I. simple, describir cuáles son todos los posibles cambios en la dinámica que pueden aparecer al perturbar un sistema dado. Aparecen así los llamados “despliegamientos universales”.

Mientras que problemas de existencia de ciertas soluciones pueden abordarse por métodos topológicos o geométricos que, sin embargo, dan poca información sobre las características de las mismas, los métodos analíticos son muy útiles en problemas perturbativos, cuando el sistema es cercano a otro que sea más simple y cuyas soluciones sean bien conocidas. Entre ambos enfoques hay una amplísima “tierra de nadie” en la que es indispensable contar, también, con métodos numéricos rigurosos.

El cálculo de O.I. presenta importantes retos incluso para los métodos numéricos más refinados.

Aspectos Computacionales

No sólo es posible usar métodos numéricos, simbólicos y su combinación para estudiar casos concretos sino también en el análisis de problemas teóricos. Pero veamos primero qué es posible calcular. Un procesador actual puede hacer 10^9 operaciones básicas (+, -, *) por segundo. Consideremos S.D. deterministas y, en el caso de sistemas continuos (EDO, EDP que al discretizar se convierten en EDO en dimensión elevada), la llamada aplicación de Poincaré (por ejemplo, una de las variables del S.D. vuelve a un valor prefijado conveniente). Así todo S.D., discreto o continuo, puede estudiarse como discreto. Para fijar ideas vamos a llamar “imagen” al cálculo de una aplicación de Poincaré. Son relevantes para obtener soluciones periódicas, toros invariantes, variedades invariantes, conexiones, etc. Costes “típicos” si el S.D. original es discreto, EDO o EDP pueden ser 10^3 , 10^6 y 10^{12} operaciones básicas, respectivamente.

Por ello, para EDO en dimensiones pequeñas (digamos, hasta 7 u 8, incluyendo parámetros) es posible hacer estudios razonablemente completos con una red de 100 procesadores, a un ritmo cercano a 10^{10} imágenes por día. Eso se reduce de forma brutal en el caso de EDP. La evolución temporal de una sola solución hasta tener una idea de cómo se comporta puede requerir 1 día completo en un tal cluster. Un estudio paramétrico, análisis de bifurcaciones, etc, puede requerir meses, incluso con esos recursos.

Es bien sabido que uno de los puntos débiles del país en matemáticas es la escasez de publicaciones en métodos y análisis numérico. Ello es debido a dos razones. La primera es el poco énfasis en ese área dentro de las licenciaturas, especialmente a nivel de cómputo efectivo. Ello no se ve en modo alguno favorecido con las actuales directrices sobre enseñanza de la Matemática. La segunda es que, ciertamente, hay diversos grupos de investigación que son expertos en aspectos computacionales. Pero para ellos esos aspectos son sólo una herramienta. No van a publicar en revistas especializadas en temas numéricos porque no son esos sus objetivos básicos.

Por otra parte NO es razonable suponer que, dotando de clusters a los grupos de investigación que deberían ser sus usuarios, se resuelve el problema. Es necesario, primero, formar a dichos investigadores en la resolución de pequeños problemas que son asequibles con los recursos disponibles en la actualidad. Mientras las simulaciones se hagan sólo con paquetes comerciales no se estará en condiciones de llegar a la frontera del conocimiento actual.



Aún sólo de paso conviene mencionar el uso de los llamados “indicadores de la dinámica” (exponentes de Lyapunov, entropía, análisis de frecuencias, velocidades de difusión, dimensiones de Hausdorff, etc.) que permiten una idea rápida de algunas características de los S.D. Y también el amplio campo de las “computer assisted proofs” (CAP) en las que la combinación de métodos topológicos, geométricos y analíticos con resultados numéricos rigurosos (uso de multiprecisión, análisis intervalar y otros) permiten demostrar (en el sentido matemático) resultados “in the middle of nowhere”.

Perspectivas

El estado actual en S.D. es interesante y prometedor. Pero creo que hay que desarrollar herramientas y métodos nuevos. Podemos esperar:

- a) La aplicabilidad de los S.D. se incrementará, incluyendo biología, medicina, etc. En esas áreas conviene hacer un enorme esfuerzo en conocimiento básico y modelización. Los aspectos cualitativos y cuantitativos se fusionarán.
- b) Métodos típicos de dimensión baja (EDO) se extenderán a EDP, EDR y ecuaciones con parte estocástica. Por razones de dificultad obvias, esos dominios están lejos del “state of the art” en EDO.
- c) Los resultados teóricos tenderán hacia métodos constructivos y eficientes. Se sistematizarán las CAP a muchos dominios. Se reducirá el “gap” entre el dominio de validez de un resultado y el dominio en el que hay evidencia física o numérica de que es aún cierto.
- f) Se hará un gran esfuerzo en la comprensión de los regímenes transitorios. Recordemos que nuestro mundo está permanentemente en estado transitorio. Según las escalas de tiempo, la influencia de cambios temporales en el modelo no se puede despreciar.
- g) En el aspecto teórico la Matemática tiene herramientas con amplio dominio de aplicación, pero que dan información muy parcial, y otras que dan información precisa pero muy local. Es deseable hallar nuevos campos matemáticos que den información precisa y global.
- h) Dado que muchos S.D. sólo admiten una descripción correcta usando medidas probabilísticas y su evolución, no parece que los ordenadores digitales sean la mejor opción. Quizás los basados en la cuántica sean la solución. Pero no hay que esperarla en un futuro inmediato.

Bibliografía

- [1] Simó, C.; Dynamical systems, numerical experiments and super-computing. *Memòries de la Reial Acadèmia de Ciències i Arts de Barcelona*, Núm. 987, Vol. LXI, (2003), (1–36).

Capítulo XXI

ANÁLISIS ARMÓNICO:
PERSPECTIVAS Y APLICACIONES

POR ANA VARGAS REY
UNIVERSIDAD AUTÓNOMA DE MADRID
Departamento de Matemáticas,
Facultad de Ciencias
CP 28049
Correo electrónico: ana.vargas@uam.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Harmonic Analysis was born in the nineteenth century as a tool to treat some Partial Differential Equations, which are models for natural, social and economical phenomena. The basic idea is the representation of a function in terms of some elementary building blocks. That may be useful to understand the action of operators (transformations) on functions. Some canonical operators studied with these methods are singular integrals (which originally appeared in connection with elliptic equations) and oscillatory integrals (fundamental to understand wave propagation).

Among the applications of Harmonic Analysis we may highlight two:

- a) Some inverse problems that are models for medical non invasive scan methods, such as ultrasound tomography and the X-ray.
- b) Image processing and compression using wavelets, cosine basis and other recently developed Analysis tools.

Introducción

En 1807 Fourier presentó una memoria en el Instituto de Francia, en la que afirmaba que toda función periódica (de período 1) se puede representar como una serie (suma de infinitos términos) de senos y cosenos.

$$(20) \quad f(x) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(2\pi nx) + b_n \operatorname{sen}(2\pi nx)] = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i nx}$$

Este trabajo estaba motivado por el estudio de la difusión del calor, que, como muchos otros fenómenos naturales (y económicos o sociales) se modeliza con una ecuación diferencial. El Análisis Armónico (también llamado Análisis de Fourier) ha sido desde entonces una herramienta básica para el estudio de estas ecuaciones.

Las funciones se pueden utilizar para representar multitud de fenómenos: una función $f(t)$ del parámetro tiempo, puede representar la presión del aire en el tímpano (que nuestro sistema nervioso transformará en un sonido). Una imagen plana, en blanco y negro se representa con una función $f(x, y)$ que nos dice el nivel de gris en cada punto; una imagen en color es la superposición de tres funciones que representan los niveles de rojo, verde y azul (los tres canales de color). La temperatura en una barra de hierro es una función $F(x, t)$ del tiempo y de la posición.

La serie (20) recibe el nombre de Serie de Fourier de la función f . Toda función de cuadrado integrable se representa de esta forma y además la representación es única, es decir, la función determina los coeficientes $\{a_n, b_n\}$ de forma unívoca. Para funciones no periódicas existe una

expresión análoga, en la que la suma se sustituye por una integral y la sucesión de coeficientes por otra función $\widehat{f}(\omega)$ a la que se conoce como Transformada de Fourier de f . La relación entre una función y su transformada de Fourier es integral:

$$(21) \quad \widehat{f}(\omega) = \int e^{-2\pi i x \cdot \omega} f(x) dx, \quad f(x) = \int e^{2\pi i x \cdot \omega} \widehat{f}(\omega) d\omega.$$

Existen versiones de las series de Fourier y la Transformada de Fourier para funciones que dependen de más de una variable. Una expresión como (20) nos permite sustituir una función por una sucesión de coeficientes $\{a_n, b_n\}$ a partir de los cuales podemos reconstruirla. Esto es una forma de digitalizar la función. La sucesión de senos y cosenos que aparece en (20) se puede sustituir por otra “base” adecuada al problema que estemos tratando y a la familia de funciones con que trabajemos. Un problema fundamental es entender en qué sentido y para qué tipo de funciones, se tiene la igualdad (20), o la análoga para la base con que trabajemos. Dicho de otro modo, si almacenamos sólo un número finito (suficientemente grande) de los coeficientes y con ellos “construimos” una función, esperamos que ésta sea una buena aproximación de la función de partida. Querríamos además saber cuán buena es esta aproximación (y esto dependerá del número de coeficientes almacenados y del tipo de función).

En la siguiente sección hablaremos del tipo de problemas que se abordan en Análisis de Fourier. Para no extendernos demasiado nos restringiremos al caso euclídeo, dejando de lado un amplio campo que es el Análisis de Fourier en grupos. En las secciones 3 y 4 hablaremos de algunas aplicaciones prácticas.

Operadores [5]

En general, el Análisis de Fourier no se ocupa de estudiar ejemplos concretos de funciones, sino de familias de funciones genéricas que cumplen ciertas propiedades de regularidad o de tamaño. El tamaño y la regularidad se pueden medir de varias formas distintas que se denominan normas. Por ejemplo nos puede interesar el valor máximo de la función (por ejemplo, la intensidad máxima en una imagen) o su tamaño medio, o su tamaño cuadrático medio. Respecto a la regularidad, podemos trabajar con funciones más o menos suaves (diferenciables, Lipschitz, Hölder continuas, ...). Todo esto se puede expresar con una norma adecuada. En lo que sigue denotaremos $\|f\|$ una de estas normas medida en la función f .

El Análisis de Fourier estudia operadores que actúan sobre las funciones. Un operador es una transformación que se aplica a las funciones, como por ejemplo cambiar el contraste en una (función) imagen. Para estudiar el efecto de un operador sobre la función f , podemos ver cómo actúa sobre cada función elemental $\cos(2\pi n x)$, $\sin(2\pi n x)$, y (si es lineal) superponer esos efectos. En particular, si estudiamos fenómenos oscilatorios, su efecto sobre esas funciones elementales es en muchos casos sencillo. Un ejemplo típico son los operadores que consisten en multi-



plicar cada coeficiente a_n por una cantidad $m(n)$; a éstos se les llama operadores multiplicadores de Fourier o filtros de frecuencia. Todos los operadores “invariantes por traslaciones” (en el tiempo, en el espacio, . . .) responden a este modelo. Un operador multiplicador de Fourier, a pesar de su sencillez, puede distorsionar de forma radical la norma de las funciones de una familia. Diremos que un operador T es acotado si existe una desigualdad de la forma

$$\|Tf\| \leq C \|f\|,$$

que nos permite controlar el tamaño de la función que resulta al aplicar el operador T a una función f , en términos de la norma de f y una constante universal c (la misma para todas las funciones f). Un problema típico en Análisis de Fourier es estudiar si un determinado multiplicador de Fourier es acotado con una norma que nos interese. El problema de estudiar la convergencia de la serie (20) se traduce en el problema de acotación del operador de suma parcial, o, si nos interesa la convergencia puntual, del llamado operador de Carleson. Este último es un problema extremadamente difícil, que de hecho está sin resolver en el caso multidimensional. En esta sección presentaremos algunos de los operadores más estudiados en Análisis Armónico.

Los operadores multiplicadores se engloban dentro de una familia más general de operadores que se describen a través de una integral

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y) dy$$

donde la función K recibe el nombre del núcleo y es una función con singularidades de tamaño (es decir puede tomar valor infinito) en el conjunto $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : x = y\}$. Este tipo de operador recibe el nombre de integral singular. A. Calderón y A. Zygmund trabajaron con estos operadores en los años 50, para estudiar la regularidad de las soluciones de la ecuación de Laplace, que es la versión estacionaria de la ecuación del calor y también es importante en la modelización de problemas de electrostática clásica, como por ejemplo, la descripción del potencial generado por una distribución de cargas. Encontraron condiciones bastante generales sobre el tamaño del núcleo y su gradiente, bajo las que los operadores están acotados con ciertas normas naturales. En esta teoría una pieza fundamental es el operador maximal de Hardy–Littlewood, que se define como un supremo de los promedios de una función sobre bolas:

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r} f(y) dy$$

Hay una familia de operadores, llamados de tipo Kakeya (o Nikodym, según la nomenclatura de otros autores) que se definen de manera análoga al de Hardy–Littlewood, pero sustituyendo bolas por rectángulos o tubos en direcciones predeterminadas o con una excentricidad fija (la excentricidad es la razón entre el radio de la sección del tubo y su longitud). Están asociados a las integrales oscilatorias, que definiremos más adelante y a algunos problemas inversos, de los que hablaremos en la siguiente sección.

En los años posteriores al trabajo de Calderón y Zygmund, se estudiaron integrales singulares más generales, en los que el espacio \mathbb{R}^n se sustituye por otros no isotrópicos, es decir su estructura y propiedades varían dependiendo de la posición (imaginemos cualquier medio natural realista). Se demostraron resultados análogos a los de Calderón y Zygmund para los llamados espacios homogéneos, en los que las condiciones no varían excesivamente si nos movemos de una región a otra cercana. Estos resultados son fundamentales a la hora de trabajar con ecuaciones diferenciales elípticas fuertemente no lineales, como la ecuación de Monge–Ampère. Uno de los momentos cruciales del Análisis de Fourier de finales de los 90 fue la generalización de esos resultados a espacios no homogéneos. Aún se trabaja en el desarrollo completo de este tema. Este problema está íntimamente ligado a otras áreas como la Teoría Geométrica de la Medida o la Teoría del Potencial.

Si consideramos ecuaciones como las de ondas o de Schrödinger, relacionadas con fenómenos ondulatorios, aparecen otro tipo de operadores, las integrales oscilatorias. En una primera aproximación podemos considerar ecuaciones lineales y en ese caso, la solución de los problemas tiene una representación sencilla utilizando la Transformada de Fourier. Por ejemplo, si consideramos la ecuación libre de Schrödinger $i\partial_t u + \Delta u = 0$, su solución tiene la forma

$$(22) \quad u(x, t) = \int \exp [i(2\pi x \cdot \xi - 4\pi^2 t |\xi|^2)] \hat{u}_0(\xi) d\xi$$

donde u_0 es el dato del problema a tiempo inicial $t = 0$. Aquí aparece un operador de tipo integral

$$Tf(x) = \int e^{i\phi(x, \xi)} f(\xi) d\xi$$

La ecuación de ondas está asociada a otra integral oscilatoria. A diferencia de las integrales singulares descritas arriba, en este caso la dificultad para estudiar el operador no es el tamaño del núcleo, sino el hecho de que puede oscilar de forma poco controlada. Las acotaciones para integrales oscilatorias permiten estudiar la existencia de soluciones de ecuaciones diferenciales no lineales, su tamaño y regularidad y si hay explosiones en esas soluciones, es decir, si pueden tomar valor infinito en algún punto, instante de tiempo, o en alguna norma.

Las integrales oscilatorias están asociadas a los operadores maximales de tipo Keakeya. Si en la expresión (22) “cortamos” la transformada de Fourier \hat{u}_0 de forma suave, para restringirla a una pequeña bola (dicho de otro modo, nos quedamos solo con frecuencias en una región acotada) el efecto sobre la solución $u(x, t)$ es que la sustituimos por promedios sobre un cilindro en espacio–tiempo. O si consideramos un dato inicial con transformada de Fourier suave localizada en una pequeña bola, su solución, vista en espacio–tiempo, está localizada en un tubo cuya excentricidad depende del tamaño de la bola y su dirección depende de las frecuencias del dato inicial; esto es lo que se conoce como un paquete de ondas. Entender la interacción de los paquetes de ondas (y por tanto la solución general de las ecuaciones diferenciales asociadas) pasa por el estudio de las propiedades de acotación de los operadores de Keakeya.



Problemas inversos [4]

La Transformada de Fourier no es sólo un artefacto matemático, sino que aparece en la naturaleza: en óptica, si estudiamos la difracción producida por una pequeña abertura, aparece (en una primera aproximación) la transformada de Fourier de una función asociada a esa abertura; si se observa la difracción de rayos X por un cristal, lo que vemos es la transformada de Fourier de la densidad electrónica del cristal.

En líneas generales, la teoría de scattering (difracción) se preocupa del efecto de un medio no homogéneo sobre una onda incidente. El problema directo de scattering consiste en determinar la onda resultante a partir de la onda incidente. El problema inverso consiste en determinar la naturaleza de las singularidades del medio a partir del conocimiento de ciertas mediciones asintóticas de las ondas resultantes. De este modo es como obtenemos gran parte de la información sobre el mundo que nos rodea. Un ejemplo de la vida real es la visión humana, que reconstruye el mundo que nos rodea a partir de medidas de ondas de luz que alcanzan nuestras retinas; los murciélagos hacen lo mismo con ondas sonoras; muchos de los métodos modernos de exploración médica no invasivos se basan en scattering inverso de ondas de ultrasonidos y electromagnéticas; algunos sistemas de prospección petrolífera utilizan ondas sísmicas.

En el caso de tomografía de ultrasonidos el modelo matemático es el siguiente: la solución del problema directo (tras una transformación de Fourier para eliminar la dependencia del tiempo) correspondiente a una onda incidente plana, $u_i(x, t) = e^{i(k\theta \cdot x - kt)}$ ($k \in \mathbb{R}_+$, $\theta \in S^{n-1}$, es un vector de módulo 1 que indica la dirección de propagación), satisface una ecuación

$$(23) \quad \Delta u + k^2 n(x)u = 0$$

y debe tener la forma $u(x) = e^{ik\theta \cdot x} + u_s(x)$, donde u_s cumple una cierta condición (condición de radiación de Sommerfeld) que garantiza que representa una onda saliente. En esta ecuación, $n(x)$ es el índice de refracción del medio. Para simplificar, supongamos que el medio ocupa una región acotada del espacio (y fuera de esa región el índice de refracción es 1). Se prueba que asintóticamente (es decir, cuando $x \rightarrow \infty$), $u_s(x) = \frac{c}{|x|} e^{ik|x|} A\left(k, \frac{x}{|x|}, \theta\right)$, donde la función A se conoce como amplitud de scattering. Además,

$$(24) \quad A\left(k, \frac{x}{|x|}, \theta\right) = -k^2 \int \exp\left(-ik \frac{x}{|x|} \cdot y\right) (n(x) - 1) [e^{ik\theta \cdot y} + u_s(y)] dy$$

El problema inverso consiste en obtener información sobre el índice de refracción, $n(x)$, a partir de mediciones de la amplitud de scattering. Éste es un problema altamente no lineal; una forma de dar soluciones aproximadas es considerar el problema lineal que se obtiene eliminando “términos de orden superior”, es decir, suponemos que u_s es muy pequeña y que en una primera aproximación, podemos ignorar su efecto en (24). Ésta es la llamada aproxima-

ción de Born. Se obtiene un “valor aproximado”, n_B de $n(x)$ a partir de de la amplitud de scattering:

$$A(k, \frac{x}{|x|}, \theta) = -k^2(\widehat{n_B - 1})(k(\frac{x}{|x|} - \theta)).$$

Esta fórmula involucra mediciones de la amplitud de scattering en todas las direcciones y con todas las posibles ondas incidentes. Este es un problema sobredeterminado y tanto en teoría como en la práctica se trabaja con menos datos: en el problema de ángulo fijo, se considera la dirección de la onda incidente θ , fija; en el “backscattering” sólo se utilizan los datos de scattering en la dirección opuesta a la de las ondas incidentes, $A(k, -\theta, \theta)$. Un problema matemático subyacente en cada caso es determinar hasta qué punto ese supuesto “valor aproximado” de $n(x)$ realmente lo es, es decir, si la diferencia $n_B(x) - n(x)$ es pequeña (en alguna norma adecuada). Para establecer ese tipo de resultados es necesario probar que el operador que define la diferencia $n_B(x) - n(x)$ es un operador acotado. Éste es un problema abierto, ya que ese operador no se ajusta a ningún modelo clásico; se conocen algunos resultados para modelos más sencillos, como

$$\Delta u + k^2 u = V(x)u.$$

Obsérvese que en la ecuación (23), $V(x) = k^2(1 - n(x))$ depende del número de onda, k , lo que hace el problema mucho más complejo.

Otra clase de operador muy importante en las aplicaciones es el grupo de las transformaciones de tipo Radon. El ejemplo básico es el siguiente: dada una función f definida en \mathbb{R}^2 o \mathbb{R}^3 calculamos sus integrales sobre sobre todas las posibles rectas, π (transformada de Rayos X),

$$Xf(\pi) = \int_{\pi} f.$$

Este tipo de transformación modeliza, por ejemplo, las exploraciones médicas CT con rayos X (transmission computer tomography). f es el coeficiente de atenuación del objeto observado. Lo que el aparato nos ofrece es una colección de valores, $Xf(\pi)$ de esas integrales. A partir de ellas queremos reconstruir la función f de partida, es decir, hay que invertir la transformada de Radón (con un método que sea computacionalmente viable). La inversa de la transformada de Radón, X^{-1} , involucra una de las integrales singulares de Calderón-Zygmund, la llamada transformada de Hilbert.

Posiblemente, en este tipo de problemas sería más razonable sustituir la integral sobre rectas por una integral sobre tubos sólidos (resultado de engrosar esas rectas). Por tanto, los operadores de tipo Kakeya tendrían algo que decir en este problema.

En todos los problemas inversos es importante también estudiar la estabilidad, es decir, si pequeños errores en las mediciones inducen pequeños errores en las reconstrucciones. Esto se traduce en que queremos probar que el operador que asocia los errores en las medidas con los errores en las reconstrucciones es un operador acotado.



Wavelets [3], [2]

En muchas aplicaciones, por ejemplo en la codificación y compresión de imágenes, la transformada de Fourier resulta inadecuada, porque tiene una pobre localización tiempo–frecuencia. Por dar un ejemplo extremo, vemos en la expresión (21) que para calcular la transformada de Fourier necesitamos conocer toda la función; si pensamos en almacenar la información contenida en una pieza musical, esto nos está diciendo que necesitamos oír toda la pieza antes de comenzar a hacerlo. La transformada de Fourier no es adecuada para captar aspectos locales de la señal, como cambios súbitos, saltos, autosimilaridades locales. Notemos que estos detalles contienen gran parte de la información en una señal; pensemos en los contornos en una imagen, que nos hacen distinguir unos objetos de otros y son cambios bruscos de intensidad de luz.

En los años 50 Gabor y su escuela intentaron resolver este problema, introduciendo modificaciones de la transformada de Fourier, con las llamadas transformadas de Fourier con ventanas. Encontraron (es el teorema de Balian–Low) un obstáculo: no era posible discretizar una señal de esa manera de modo no redundante y con buena localización a la vez en tiempo y frecuencia. En los años 70, con los trabajos de Morlet y Grossman, comenzaron a utilizarse en análisis de señales las llamadas “wavelets” (“ondelettes” en francés, “ondículas” en castellano). La idea fundamental es procesar las funciones a todas las escalas (en el análisis con ventanas, la escala era única, el tamaño de la ventana). En lugar de la serie de senos y cosenos de la expresión (20) se utiliza una familia de funciones $\{\psi_{j,k} = 2^{j/2}\psi(2^jx - k)\}$ (una base de wavelets) dependiente de dos parámetros. La función f se representa como

$$(25) \quad f(x) = \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x)$$

donde

$$(26) \quad \langle f, \psi_{j,k} \rangle = \int f(y) \psi_{j,k}(y) dy$$

Meyer construyó la primera wavelet suave con buena localización en tiempo y frecuencia, mostrando que no hay análogo al teorema de Balian–Low en este contexto. Mallat y Meyer, en los 80, dieron estructura a este campo con la definición del Análisis Multirresolución. Daubechies construyó la primera wavelet suave con soporte acotado. En problemas bidimensionales (como es una imagen), se construyen bases de wavelets a partir de productos tensoriales de wavelets unidimensionales. En la actualidad se trabaja también con las bases biortogonales de wavelets, en las que la wavelet que se utiliza para codificar la función (26) es distinta a la que se utiliza para reconstruirla (25). Otros autores están explorando otros tipos de bases, como las “curvelets” de Donoho o las “contourlets” de Minh Do [2].

Las wavelets tienen en la actualidad múltiples aplicaciones debido a dos propiedades: los coeficientes tienen un significado local (y están jerarquizados por escalas) y además se puede apro-

ximar eficientemente clases amplias de señales utilizando pocos coeficientes significativos (si se elige una wavelet adecuada al problema entre la biblioteca de wavelets disponible). Esto es lo que se hace cuando se “comprime” una imagen en jpeg2000; a grandes rasgos, lo que se hace es eliminar todos los coeficientes correspondientes a frecuencias altas (detalles). En otras versiones de jpeg se utilizan bases de cosenos discretas en lugar de wavelets. También se utilizan las wavelets para eliminar ruidos. En este caso, suponiendo que se trata de ruido blanco, se establece un umbral (que depende de los datos del problema concreto) y se eliminan los coeficientes menores que él (en otras versiones, se resta a todos los coeficientes una cantidad). Las wavelets son utilizadas por el FBI en su sistema de almacenamiento de huellas digitales WSQ (wavelet transform/scalar quantization image coding).

Bibliografía

- [1] Bruna, J.; *Representació i mostratge de funcions: de les sèries de Fourier a la teoria d'ondetes*. Butlletí de la Societat Catalana de Matemàtiques, Vol. 14, N. 1, (1999), (31-61).
- [2] Do, M.N.; Vetterli, M.; *Contourlets: a directional multiresolution image representation, I* Proceedings of the 2002 International Conference on Image Processing, Vol. 1, (2002), (357- 360).
- [3] Mallat, S.; *A wavelet tour of signal processing*. Academic Press, New York, (1998).
- [4] *Mathematics and Physics of Emerging Biomedical Imaging*, National Research Council, Institute of Medicine. National Academy Press, Washington D. C., (1996).
- [5] Stein, E.M.; *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Mathematical Series, 43. Princeton University Press, Princeton, NJ, (1993).



Capítulo XXII

MATEMÁTICAS Y SOCIEDAD
DE LA INFORMACIÓN

POR ORLANDO VILLAMAYOR URIBURU
UNIVERSIDAD AUTÓNOMA DE MADRID
Departamento de Matemáticas
Facultad de Ciencias
CP 28049
Correo electrónico: villamayor@uam.es

mat



MATEMÁTICAS EN LA FRONTERA
NUEVAS INFRAESTRUCTURAS MATEMÁTICAS
EN LA COMUNIDAD DE MADRID
COMPUTACIÓN E INTERACCIÓN I+D+i

Abstract

Mathematics has been playing a key role in the search of new forms of access to information for the general public. We illustrate this fact with two examples: the development of GPS (Global Positioning System); and with Google, as a searching tool in Internet.

A GPS, which requires interaction with a number of satellites in orbit, is now available for cars and bicycles.

As for the role of Google: it is clear that the vast amount of data available in Internet can only be of practical use with a clever form of search and selection of this information. When we search in Internet using Google the outcome is a long but ordered list of sites; and normally we obtain a significant answer to our requirement among the first sites in the list.

The objective of these notes is to indicate the role of mathematics within the development of these two areas. Both areas are subjects in progress, and hence of motivation of current mathematical research.

A good part of this presentation is based on the articles in the list of references, which we strongly suggest for further reading on these topics.

Introducción

Entre los aportes que la ciencia y la tecnología han proporcionado a la sociedad, en los últimos años, destaca las posibilidades que han dado al individuo para obtener y disponer de información, que antes le estaba vedada o le era de difícil acceso.

El papel de las matemáticas ha sido determinante en este proceso de socialización de la información. En efecto, este desarrollo se ha plasmado gracias a las respuestas que las matemáticas han dado a problemas planteados por la transmisión y por la acumulación de la información.

Aquí ilustraremos este hecho a través de dos ejemplos. Discutiremos por un lado el desarrollo de los sistemas GPS, relacionado a información vía satélite, y que en la actualidad es un instrumento asequible para cualquiera. Por otro lado discutiremos su papel en la selección de información disponible en Internet. Concretamente sobre los "buscadores", como es el ejemplo de Google, verdadero logro de ingeniería matemática.

Aprovechamos para indicar algunos retos que conlleva, para las matemáticas, el desarrollo futuro de estas técnicas; y que son de interés para los investigadores de la Comunidad de Madrid.

1) El buscador Google se ha diseñado para ordenar información disponible en Internet. Hay cientos de millones de páginas web, y es claro que la cantidad ingente de información disponi-

ble haría imposible su utilidad sin un criterio de selección apropiado. Si se hace una consulta concreta la base de datos puede contener miles de páginas con información. El objetivo del buscador es ordenar las páginas seleccionadas, de tal modo que entre las diez o veinte primeras haya información que sea útil y relevante para quien realiza la consulta.

El buscador va a ordenar las páginas web siguiendo algún criterio. Frente a una consulta concreta hará la búsqueda en una impresionante base de datos, y hará uso de esta ordenación de páginas al presentar la información recabada.

No vamos a discutir aquí sobre el problema que conlleva la búsqueda en una base de datos tan grande, aunque también en ello intervienen las matemáticas. Nos centraremos en el problema de ordenación de páginas.

Partimos de información sobre la red, en cuanto a sitios, contenidos, enlaces entre páginas, etc. Con estos datos el buscador hará su ordenación. Naturalmente esta se actualiza cada tantas semanas dada la constante aparición de nuevos sitios en la red.

En un primer paso para la modelización del problema vamos a etiquetar cada página, digamos P_1, P_2, \dots, P_N , y vamos a considerar los enlaces entre ellas. De este modo podemos interpretar al modelo como un **grafo** dirigido: el conjunto de puntos P_1, P_2, \dots, P_N , y una flecha de P_i en P_j si hay un enlace de la primera en la segunda. Podemos codificar este grafo dirigido con una matriz \mathbf{M} de dimensión $N \times N$, y con entrada m_{ij} , que definimos como 1 si hay enlace desde la página P_j a la P_i , y m_{ij} igual a cero en caso contrario.

De este modo obtenemos una matriz con ceros y unos. Nótese que los unos que aparecen en la fila i corresponden a los enlaces a la página P_i , y que los unos que aparecen en la columna j se corresponden con los enlaces *desde* la página P_j .

Recordemos que nuestro objetivo es establecer un orden de *importancia* de las P_i . Ahora bien, la suma de las entradas de \mathbf{M} en la fila i , es el número total de páginas P_j que enlazan con P_i .

En una primera aproximación al problema, ordenemos las P_i por la suma de las entradas en la fila i (por el número total de los P_j que enlazan con P_i). Un simple análisis nos va a indicar, sin embargo, que este criterio de ordenación no es el más adecuado. Fijemos, para ejemplificar, dos páginas P_1 y P_2 . Supongamos ahora que hay exactamente tres páginas, digamos $P_{i1}, P_{i2},$ y P_{i3} , que enlazan con P_i ; y que hay exactamente cuatro, digamos $P_{j1}, P_{j2}, P_{j3},$ y P_{j4} , que enlazan con P_2 . Si solamente tomamos en cuenta el número de enlaces, entonces P_1 tiene "importancia" tres, y P_2 "importancia" cuatro.

Sin embargo, con este criterio no tenemos en cuenta las "importancias" que hemos asignados a las $P_{i1}, P_{i2},$ y P_{i3} , ni las asignadas a $P_{j1}, P_{j2}, P_{j3},$ y P_{j4} . Supongamos, por ejemplo, que cada una de las $P_{i1}, P_{i2},$ y P_{i3} tiene, con este criterio, importancia 100; y que cada uno de los $P_{j1},$



P_{j_2} , P_{j_3} , y P_{j_4} , tiene importancia 2. Es claro que un criterio razonable debe asignar mayor importancia a P_1 que a P_2 .

Haremos una segunda aproximación al problema de asignación de importancias de un modo diferente. Digamos que a cada P_i le queremos asignar un número real positivo x_i (valor de "importancia"), y busquemos una asignación en la que el valor x_i tenga en cuenta la suma de las importancias de las páginas que enlazan con P_i .

Veamos si podemos hacerlo de tal modo que x_i sea proporcional a la suma de las importancias de las páginas que enlazan con P_i . Siguiendo nuestro ejemplo anterior, nos estamos preguntando si podemos encontrar reales positivos x_i , de tal modo que

$$x_1 = K(x_{i_1} + x_{i_2} + x_{i_3})$$

y

$$x_2 = K(x_{j_1} + x_{j_2} + x_{j_3} + x_{j_4})$$

Para cierta constante K real y positiva.

Sea $F_1 = (m_{1,1}, m_{1,2}, \dots, m_{1,N})$ la primera fila de la matriz \mathbf{M} , y $F_2 = (m_{2,1}, m_{2,2}, \dots, m_{2,N})$ la segunda fila. En el caso de nuestro ejemplo las entradas $m_{1,k}$ son nulas salvo para los caso $k = i_1, i_2, i_3$; y las $m_{2,k}$ son nulas salvo para los caso $k = j_1, j_2, j_3$ y j_4 . Observemos que entonces la primera fila del producto $\mathbf{M}\mathbf{x}$ es $x_{i_1} + x_{i_2} + x_{i_3}$ y la segunda $x_{j_1} + x_{j_2} + x_{j_3} + x_{j_4}$.

En el caso general, nos estamos preguntando si dada nuestra matriz \mathbf{M} , podemos encontrar valores reales positivos K , y x_1, x_2, \dots, x_N tales que

$$\mathbf{M}\mathbf{x} = K^{-1}\mathbf{x},$$

donde \mathbf{x} denota el vector columna con coordenadas x_1, x_2, \dots, x_N .

Desde un punto de vista matemático nos estamos preguntando por un *autovalor* positivo de la matriz, y por un *autovector* \mathbf{x} positivo (con todas sus entradas positivas). En efecto, este último daría lugar a "valores de importancia" para cada P_i , $i = 1, 2, \dots, N$.

La discusión anterior ilustra cómo un autovector positivo de \mathbf{M} proporcionará valores de importancia de las páginas según nuestro modelo. Pero hallar un autovector positivo de \mathbf{M} , y teniendo en cuenta que N es un número del orden de miles de millones, no es viable desde un punto de vista computacional.

Antes de presentar la versión definitiva del modelo usado por Google para valorar la importancia de las páginas, empecemos por indicar cómo hallar un autovector positivo, de manera computacionalmente efectiva, al menos para una clase muy particular de matrices cuadradas.

Nos vamos a centrar en las llamadas *matrices de transición*.

Se dice que una matriz T , de $N \times N$, es de transición, cuando todas las entradas son reales no negativos y tiene la propiedad adicional de que la suma de las entradas de cada columna es 1.

Si T es de transición, también lo son sus potencias: T^2, T^3, \dots . Estas matrices han sido bien estudiadas debido a sus aplicaciones al estudio de procesos *probabilísticos*, en las llamadas *cadena de Markov*.

Se dice que una matriz de transición T es *regular*, si para algún entero positivo k , todas las entradas de la matriz T^k son estrictamente positivos.

Sea \mathbf{x} un vector columna con N entradas reales no negativas, y con la propiedad adicional de que estas entradas sumen 1.

No es difícil comprobar que si \mathbf{x} tiene esa propiedad, y si T es de transición, también los vectores $T(\mathbf{x}), T^2(\mathbf{x}), T^3(\mathbf{x})$, etc. tienen esta propiedad.

En el caso particular en que T es regular se demuestra que existe un vector columna \mathbf{c} , con entradas *estrictamente* positivas que suman 1, y tal que:

1. Los vectores $T(\mathbf{x}), T^2(\mathbf{x}), T^3(\mathbf{x}) \dots$ convergen (¡rápidamente!) al vector \mathbf{c} .
2. $T\mathbf{c} = \mathbf{c}$.

Más aún, la convergencia al vector \mathbf{c} , indicada en (1), es independiente del vector \mathbf{x} que elijamos en las condiciones anteriores.

Notemos que (2) indica que \mathbf{c} es un autovector de la matriz T , con entradas estrictamente positivas, correspondiente al autovalor 1.

Como veremos, éste es el camino que va a seguir Google para hallar un vector de “importancias” de las páginas P_1, P_2, \dots, P_N . Se le asocia a la matriz \mathbf{M} una matriz de transición, digamos \mathbf{M}' , de manera natural como sigue: Si N_j denota la suma de las entradas en la columna j de \mathbf{M} , definimos

$$m'_{i,j} = \frac{m_{i,j}}{N_j}$$

Con esto garantizamos que la suma de las entradas de cada columna es 1, y por tanto \mathbf{M}' es una matriz de transición.

A continuación realizaremos la aproximación definitiva al vector de importancias, que será concretamente un autovector positivo de \mathbf{M}' .



Para comparar, supongamos que P_3 es citada sólo en la P_7 , entonces el vector de importancias que se obtenía de la matriz \mathbf{M} da lugar a una relación:

$$x_7 = Kx_3,$$

sin tener en cuenta el número de páginas que son citadas desde P_3 . Mientras que el autovector (vector de importancias) que resulta de \mathbf{M}' da lugar a una relación:

$$y_7 = K'y_3: \frac{1}{N_3}$$

donde N_3 denota el número total de páginas citadas desde P_3 . Este último parece un criterio más adecuado de valoración, y es el criterio de valoración que utiliza Google.

Por otro lado, desde el punto de vista estocástico la matriz \mathbf{M}' es muy natural: la entrada $m'_{i,j}$ denota la probabilidad de que un visitante de la página P_j enlace con la P_i . De esta forma la parte de la importancia de P_i debida a la página P_j queda multiplicada por la probabilidad de pasar de P_j a P_i .

Si bien \mathbf{M}' es de transición, no se puede garantizar que sea regular. Pero las matrices de transición regulares son suficientemente generales entre las matrices de transición; en el sentido que una pequeña deformación de \mathbf{M}' será regular. Google deforma la matriz \mathbf{M}' como sigue: se fija un vector $Q = (r_1, r_2, \dots, r_N)$, de números reales positivos que suman 1; y se define \mathbf{H}_Q la matriz de $N \times N$ con sus N columnas iguales a Q . Finalmente se define

$$\mathbf{M}'' = c\mathbf{M}' + (1 - c)\mathbf{H}_Q,$$

donde c un parámetro entre cero y uno. Nótese que \mathbf{M}'' es de transición, y la afirmación es que también es regular.

El autovector positivo de \mathbf{M}'' correspondiente al autovalor 1, será el vector de importancia buscado, con el que ordenará Google a las páginas en red.

Nótese que este orden que establece Google depende de la elección tanto del vector Q como de la constante c . El candidato más natural para elegir como vector Q sería aquel con $r_1 = r_2 = \dots = r_N = 1/N$. Sin embargo se utilizan otros vectores a efectos de trastocar el orden de las páginas, ¡y a efectos de censurar algunas!

Este sistema de ordenación de páginas web ("PageRank") conlleva el cálculo con matrices cuadradas de orden extremadamente grande. Un problema de investigación vigente consiste en mejorar los métodos conocidos.

Como hemos indicado anteriormente, la matriz queda determinada por las páginas web junto con la estructura de grafo dirigido. Estructura que queda definida por las referencias de una página en otra.

El conocimiento de esta estructura de grafo (Bow-tie Theory) permite expresar la matriz en forma “triangular superior en bloques”, con ciertas matrices cuadradas en la diagonal. De este modo el tratamiento con la matriz original se hace viable, por reducción al tratamiento de los bloques diagonales ([2], [3], [4]).

2) El GPS (Global Positioning System) es, hoy en día, un sistema muy divulgado, tanto para coches, barcos, como también para bicicletas e incluso caminantes.

Para fijar ideas, consideremos el caso de un barco, con un receptor GPS, y con localización desconocida. A través de este receptor, y con la información que éste reciba de cuatro satélites, se obtendrá la localización del barco de un modo muy preciso.

La modelización matemática es muy elemental, pero no por ello menos significativa.

Para entender el modelo matemático imaginemos el espacio con coordenadas xyz , con centro en el eje de la tierra, y con eje z atravesando los polos. La posición desconocida de nuestro barco se podrá así expresar en coordenadas (x, y, z) , que después se traducirán en paralelos y meridianos. Agregaremos una nueva variable t , que medirá el tiempo, y que, como veremos luego, juega un papel fundamental para nuestro propósito.

Cada satélite conocerá su posición en el espacio (su coordenada (x, y, z)), y un reloj le provee la hora (su coordenada t del tiempo).

Para simplificar, podemos marcar los tres ejes con una unidad igual al radio de la tierra. Por tanto, cualquier objeto en la superficie de la tierra tiene coordenadas (x, y, z) que verifican la ecuación $x^2 + y^2 + z^2 = 1$. Veremos ahora cómo nuestro proceso matemático nos permitirá obtener, no solamente la ubicación de nuestro barco en un instante dado, sino también la hora en dicho instante. Y todo esto a partir de los datos que recibe de los cuatro satélites.

Un dato fundamental para el sistema GPS es conocer las distancias del barco a cada satélite. Esto se hará teniendo en cuenta el tiempo que demora en llegar la señal de radio (velocidad de la luz). Esta velocidad es del orden de 0,047, en términos de unidades de radio de la tierra por milisegundos. En lo que sigue la unidad de tiempo será un milisegundo.

Digamos, para ejemplificar (ver [7]), que de cada unos de los satélites S_1, S_2, S_3 , y S_4 , el barco han recibido de manera simultánea los siguientes datos:





SATÉLITE	POSICIÓN	TIEMPO
S_1	(1, 2, 0)	19,9
S_2	(2, 0, 2)	2,4
S_3	(1, 1, 1)	32,6
S_4	(2, 1, 0)	19,9

Y el objetivo es hallar la posición (x, y, z) de nuestra nave, y el tiempo t en el momento de recibir esos datos.

Usando los datos del primer satélite S_1 , podremos hallar su distancia al barco. En efecto, la señal fue enviada a la hora 19,9 y llega a la hora (desconocida) t ; y dado que viaja a velocidad 0,047, resulta que la distancia es:

$$d = 0,047(t - 19,9)$$

Observemos que esta misma distancia se puede expresar en términos de (x, y, z) y de la posición de S_1 :

$$d = \sqrt{(x - 1)^2 + (y - 2)^2 + (z - 0)^2}$$

De donde resulta que:

$$(27) \quad (x - 1)^2 + (y - 2)^2 + z^2 = (0,047)^2(t - 19,9)^2$$

Finalmente, desarrollando cuadrados, obtenemos siempre a partir de los datos recibidos de satélite S_1 , la igualdad:

$$2x + 4y - 2(0,047)^2(19,9)t = 1^2 + 2^2 - (0,047)^2(19,9)^2 + x^2 + y^2 + z^2 - (0,047)^2t^2$$

Obtenemos, de manera análoga, igualdades a partir de los datos de los otros satélites. Escribamos las cuatro ecuaciones como sigue:

$$2x + 4y - 2(0,047)^2(19,9)t = 1^2 + 2^2 - (0,047)^2(19,9)^2 + x^2 + y^2 + z^2 - (0,047)^2t^2$$

$$4x + 0y + 4z - 2(0,047)^2(2,4)t = 2^2 + 0^2 + 2^2 - (0,047)^2(2,4)^2 + x^2 + y^2 + z^2 - (0,047)^2t^2$$

$$2x + 2y + 2z - 2(0,047)^2(32,6)t = 1^2 + 1^2 + 1^2 - (0,047)^2(32,6)^2 + x^2 + y^2 + z^2 - (0,047)^2t^2$$

$$4x + 2y - 2(0,047)^2(19,9)t = 2^2 + 1^2 - (0,047)^2(19,9)^2 + x^2 + y^2 + z^2 - (0,047)^2t^2$$



Observamos que las cuatro ecuaciones anteriores tienen los mismos términos cuadráticos. Por tanto estos se eliminan si restamos la primera igualdad a las tres igualdades siguientes. De este modo obtenemos un sistema de tres ecuaciones lineales en las cuatro variables x, y, z, t :

$$2x - 4y + 4z + 2(0,047)^2(17,5)t = 8 - 5 + (0,047)^2((19,9)^2 - (2,4)^2)$$

$$0x - 2y + 2z - 2(0,047)^2(12,7)t = 3 - 5 + (0,047)^2((19,9)^2 - (32,6)^2)$$

$$2x - 2y + 0z + 2(0,047)^2(0)t = 5 - 5 + (0,047)^2((19,9)^2 - (19,9)^2).$$

Dado que estas tres ecuaciones se obtuvieron de las cuatro primeras, resulta que si (x, y, z) denota la posición del barco en tiempo t , entonces (x, y, z, t) tiene que ser también solución del sistema lineal de ecuaciones anterior. Este tiene matriz asociada:

$$\begin{pmatrix} 2 & -4 & 4 & 0,077 & 3,86 \\ 0 & -2 & 2 & -0,056 & -3,47 \\ 2 & -2 & 0 & 0 & 0 \end{pmatrix}$$

Un sistema de tres ecuaciones lineales en cuatro variables tiene una recta de soluciones, entre las cuales figuran las de nuestro barco. Este sistema lineal es equivalente, por medio de operaciones elementales, al correspondiente a la matriz:

$$\begin{pmatrix} 1 & 0 & 0 & 0,095 & 5,41 \\ 0 & 1 & 0 & 0,095 & 5,41 \\ 0 & 0 & 1 & 0,067 & 3,67 \end{pmatrix}$$

que permite expresar las soluciones en la forma:

$$x = 5,41 - 0,095t$$

$$y = 5,41 - 0,095t$$

$$z = 3,67 - 0,067t$$

donde t es un parámetro libre.

Recordemos que las coordenadas buscadas son también solución de la ecuación (27). Reemplazando estas soluciones del sistema lineal en (27) obtenemos el polinomio de grado dos:

$$0,02t^2 - 1,88t + 43,56 = 0.$$

Éste tiene dos raíces reales: 43,1 y 50,0. Si escogiéramos $t = 43,1$, obtendríamos (del sistema lineal), la solución $(x, y, z) = ((1,317), (1,317), (0,79))$. Pero ésta tiene longitud aproximadamente 2 veces la unidad (duplica el radio de la tierra), y por tanto no representa un punto en la superficie.

En cambio para $t = 50$ obtenemos $(x, y, z) = ((0,667), (0,667), (0,332))$, con longitud 0,99997. Ésta será la respuesta del sistema GPS.



Este ejemplo supone la exactitud de los datos, mientras que GPS tiene que tratar con datos reales, susceptibles de errores. Estos errores provienen de los relojes internos, tanto de satélites como de receptor, y también se originan cuando la señal atraviesa la ionosfera.

Ya con dos receptores, uno de ellos fijos (GPS diferencial), se controlan estos errores y se mejora la fidelidad. Uno de los retos matemáticos aparece al considerar toda una red de receptores en contacto con satélites.

El sistema GPS mide distancias contando longitudes de ondas entre el satélite y el receptor. Esta cantidad tiene una parte entera más una fracción. El dato de interés es la parte entera. Si se trata de una red con 100 receptores se obtiene un vector con 100 coordenadas cuyas partes enteras se quiere determinar de manera simultánea. Se trata de hallar la distancia de un vector a una red de puntos. Es un problema de mínimos cuadrados: de minimizar

$$(\mathbf{x} - \mathbf{x}_0)^T A (\mathbf{x} - \mathbf{x}_0) \quad \mathbf{x} \in \mathbb{Z}^n.$$

Esto es simple cuando A es diagonal, porque no se mezclan las variables; y el reto está en el tratamiento del caso general.

Las técnicas con las que se aborda el problema es a través de cambios de la forma $Z^T A Z$, con Z matriz invertible en los enteros, y el estudio de los pivotes que aparecen por este tipo de cambio.

Bibliografía

- [1] Fernández, P.; *El secreto de Google y el Álgebra Lineal*, Bol. Soc. Esp. Mat. Apl. 30, (2004), (115-141).
- [2] Arasu, A.; Novak, J.; Tomkins, A.; Tomlin, J.; *Graph Structure in the Web: Experiments and Algorithms*, Technical Reports, IBM Almaden Research Center, (2001).
- [3] Broder, A. (1); Kumar, R. (2); Maghoul, F. (1); Raghavan, P. (2); Rajagopalan, S. (2); Stata, R. (3); Tomkins, A. (2); Wiener, J. (3); 1: AltaVista Company, San Mateo, CA. 2: IBM Almaden Research Center, San Jose, CA. 3: Compaq Systems Research Center, Palo Alto, CA.
Disponible en: www9.org/w9cdrom/160/160.html
- [4] Moler, C.; *The World's Largest Matrix Computation Google's PageRank is an eigenvector of a matrix of order 2.7 billion*, MATLAB News Notes - October 2002 ; www.mathworks.com
- [5] Wilf, H.S.; *Searching the web with eigenvectors*. www.cis.upenn.edu
- [6] Kalman, D.; *An underdetermined Linear System for GPS*. The College Mathematics Journal, Vol 33. No 5, (2002).
- [7] Strang, G.; *The Mathematics of GPS*, SIAM News 30: 5, (1997).

TÍTULOS PUBLICADOS POR LA DIRECCIÓN GENERAL DE UNIVERSIDADES E INVESTIGACIÓN

1. La innovación tecnológica en trece sectores de la Comunidad de Madrid
2. Cooperación tecnológica entre centros públicos de investigación y empresa
3. Investigación y desarrollo en la Comunidad de Madrid
4. Madrid, Centro de Investigación e Innovación
5. Generación de conocimiento e innovación empresarial
6. La prosperidad por medio de la investigación
7. I+D+I en pequeñas y medianas empresas de la Comunidad de Madrid
8. Los Parques Científicos y Tecnológicos en España: retos y oportunidades
9. La Innovación: un factor clave para la competitividad de las empresas
10. Creación de empresas de base tecnológica: la experiencia internacional
11. Madrid, nodo de comunicaciones por satélite
12. Capital intelectual y producción científica
13. El sistema regional de I+D+I de la Comunidad de Madrid
14. Guía de creación de bioempresas
15. Inteligencia económica y tecnología. Guía para principiantes y profesionales
16. Gestión del conocimiento en Universidades y Organismo Públicos de Investigación
17. Análisis de los incentivos fiscales a la Innovación
18. VI Programa Marco para Pymes
19. Indicadores de Producción Científica y Tecnológica de la Comunidad de Madrid (PIPCT) 1997-2001
20. GEM. Global Entrepreneurship Monitor. Informe ejecutivo 2004. Comunidad de Madrid
21. NANO. Nanotecnología en España
22. ISCI. Informe Spring sobre Capital Intelectual en la Comunidad de Madrid
23. AGE-CM. Análisis de la inversión en Ciencia y Tecnología, de la Administración General del Estado, en la Comunidad de Madrid
24. PRO-IN. La propiedad de la sociedad del conocimiento
25. ICCM. Indicadores Científicos de Madrid (ISI, Web of Science, 1990-2003)
26. OSLO. Manual de Oslo. Directrices para la recogida e interpretación de información relativa a Innovación
27. SEU-1. La sanidad en Europa. Fase 1
28. SEU-2. La sanidad en Europa. Fase 2

Colección dirigida por
Alfonso González Hermoso de Mendoza

Publicación especial
PRICIT: III y IV Plan Regional de Investigación Científica e Innovación Tecnológica 2005-2008

Disponibles en Internet
<http://www.madrimasd.org>

mat

MATEMÁTICAS EN LA FRONTERA

NUEVAS INFRAESTRUCTURAS MATEMÁTICAS

EN LA COMUNIDAD DE MADRID

COMPUTACIÓN E INTERACCIÓN I+D+i