

Boletín de la Sociedad Española de Matemática Aplicada SĒMA

Grupo Editor

J.J. Valdés García (U. de Oviedo) E. Fernández Cara (U. de Sevilla)
B. Dugnot Álvarez (U. de Oviedo) M. Mateos Alberdi (U. de Oviedo)
C.O. Menéndez Pérez (U. de Oviedo) P. Pérez Riera (U. de Oviedo)

Comité Científico

E. Fernández Cara (U. de Sevilla) A. Bermúdez de Castro (U. de Santiago)
E. Casas Rentería (U. de Cantabria) J.L. Cruz Soto (U. de Córdoba)
J.M. Mazón Ruiz (U. de Valencia) I. Peral Alonso (U. Aut. de Madrid)
J.J. Valdés García (U. de Oviedo) J.L. Vázquez Suárez (U. Aut. de Madrid)
L. Vega González (U. del País Vasco) E. Zuazua Iriondo (U. Comp. de Madrid)

Responsables de secciones

Artículos: E. Fernández Cara (U. de Sevilla)
Resúmenes de libros: F.J. Sayas González (U. de Zaragoza)
Noticias de SĒMA: R. Pardo San Gil (Secretaria de SĒMA)
Congresos y Seminarios: J. Mazón Ruiz (U. de Valencia)
Matemáticas e Industria: M. Lezaun Iturralde (U. del País Vasco)
Educación Matemática: R. Rodríguez del Río (U. Comp. de Madrid)

Página web de SĒMA

<http://www.uca.es/sema/>

Dirección Editorial: Boletín de SĒMA. Dpto. de Matemáticas. Universidad de Oviedo. Avda. de Calvo Sotelo, s/n. 33007-Oviedo. boletin_sema@orion.ciencias.uniovi.es

ISSN 1575-9822

Depósito Legal: AS-1442-2002

Imprime: Grupo Bitácora. C/ Instituto, 17, Entresuelo. 33201 Gijón (Asturias)

Diseño de portada: Ana Cernea

Estimados socios:

Con la publicación de este boletín se cumple nuestro primer año al frente de la edición del mismo. Después de estos primeros pasos, se pretende que esta publicación tenga un carácter periódico, por lo que a partir de ahora, como se acordó en la última Asamblea de SĒMA, intentaremos editar un boletín cada trimestre. La experiencia acumulada en este tiempo también nos aconseja elaborar unas normas de estilo más precisas, con el fin de conseguir una mayor uniformidad en la presentación de los trabajos.

En este número hemos de destacar dos hechos relevantes: por un lado, la inauguración de la sección sobre Matemáticas e Industria, y por otro, la publicación de dos artículos de los galardonados con los Premios SĒMA 2002 (Quinto Premio SĒMA al Joven Investigador y Tercer Premio SĒMA de Divulgación de la Matemática Aplicada).

El primer artículo de la sección de Matemáticas e Industria ha sido escrito por José Antonio Garrido, Vicepresidente de Iberdrola y Doctor Honoris Causa por la UPV-EHU. J.A. Garrido desarrolla una intensa actividad empresarial y ha recibido numerosos premios y distinciones, entre las que destacaremos la Medalla Puig Adam de la Fundación para el Fomento de la Innovación. Este artículo es un excelente trabajo sobre proyectos empresariales en la sociedad del siglo XXI y esperamos que sea el primero de una larga serie que sirva para estrechar lazos entre las Matemáticas y la Industria. Agradecemos sinceramente al Dr. Garrido su participación en la citada sección, y a la empresa que representa su colaboración como socio patrocinador de SĒMA.

El Premio SĒMA al Joven Investigador ha sido concedido al profesor Carlos Castro, de la Universidad Politécnica de Madrid. El trabajo que publicamos en este boletín, “Algunos problemas de control y diseño óptimo en sistemas gobernados por ecuaciones en derivadas parciales”, pone de manifiesto el excelente nivel investigador de este joven matemático, a quien le auguramos un brillante futuro que contribuirá a elevar el nivel investigador de nuestro país. Enhorabuena.

El Premio SĒMA de Divulgación de la Matemática Aplicada ha recaído en nuestro compañero Mikel Lezaun, profesor de la Universidad del País Vasco, por el trabajo “Predicciones del Tiempo y Matemáticas”, que reproducimos en este boletín. Desde estas líneas le deseamos transmitir nuestra doble enhorabuena: en primer lugar, por el premio conseguido, y en segundo lugar, como responsable de la sección Matemáticas e Industria, cuya andadura se inicia ahora y que, sin lugar a dudas, será muy fructífera.

Grupo Editor
boletin_sema@orion.ciencias.uniovi.es

Algunos problemas en sistemas dinámicos discretos

F. BALIBREA

Departamento de Matemáticas
Universidad de Murcia

balibrea@um.es

Resumen

Tomando como punto de partida la definición de Sistema Dinámico Discreto, introducimos algunos ejemplos que nos parecen relevantes desde el punto de vista histórico y que generalmente son de dimensiones uno y dos. De tales sistemas presentamos algunos resultados, señalamos problemas pendientes y realizamos comentarios sobre los mismos. Esta presentación permite obtener, aunque sea muy parcial, una visión sobre el trabajo que se realiza en esta parte de las Matemáticas.

Palabras clave: *Período, estructura periódica, orden de Sharkovsky, función de Newton, Dinámica Simbólica, puntos atractores, puntos repulsivos, variedad estable, ecuaciones en diferencias finitas con retraso.*

Clasificación por materias AMS: *37E05, 37E99*

1 Introducción

La formulación general de los problemas a los que nos referiremos en el presente artículo es la siguiente. Sea \mathbb{X} un espacio métrico y $f : \mathbb{X} \rightarrow \mathbb{X}$ una función continua. Denominamos *Sistema Dinámico Discreto* al par (\mathbb{X}, f) , donde es el *espacio de fases del sistema* y $\Phi(x, n) = f^n(x)$ el *flujo* del mismo que representa el estado, en el espacio de fases, que ocupa en el *instante* n el punto que en el instante inicial ocupa la posición x (esta terminología está tomada del lenguaje de la Mecánica). Con f^n representamos $f \circ f^{n-1}$ para $n \geq 1$ y f^0 representa la función *identidad* en \mathbb{X} .

En este marco general, el problema que tratamos de resolver es el de averiguar la evolución de cada uno de los puntos de \mathbb{X} por la acción del flujo del sistema. Tal problema es equivalente al de comprender cuál es su *dinámica*. El conocer la dinámica de un sistema consiste en entender el comportamiento de todas las sucesiones $(f^n(x))_{n=0}^{\infty}$ correspondientes a todos los puntos de \mathbb{X} . Para $x_0 \in \mathbb{X}$, tal sucesión se denomina la *órbita de* x_0 en el sistema y se representa por $Orb_f(x_0)$. Es evidente que tendremos tantas órbitas como puntos de \mathbb{X} . Denominamos *trayectoria* de una órbita al conjunto de puntos distintos del

espacio de fases que pertenecen a la misma. El comportamiento dinámico más simple es el de *órbita periódica*. Se dice que la órbita de $x \in \mathbb{X}$ es periódica si existe m tal que $f^m(x) = x$ y al menor de los números para los que tal propiedad se verifica se denomina el *período* de x . Cuando tal número es uno, se dice que el punto es *fijo*.

Aunque históricamente existen precedentes de la *Teoría de los Sistemas Dinámicos* (por ejemplo en Poincaré), podemos decir, sin embargo, que sus orígenes y desarrollo tienen que ver con los de la *Teoría Ergódica* en el primer tercio del siglo XX. En la mayor parte de los problemas la función f solía ser un homeomorfismo y los resultados que se obtenían dependían de la métrica y también de las medidas presentes en el espacio de fases. Cuando f es un homeomorfismo se puede usar la noción de *órbita total* que viene dada por la sucesión $(f^n(x))_{-\infty}^{\infty} = (f^n(x))_0^{\infty} \cup (f^n(x))_{-\infty}^0$, es decir, como la unión de una *semiórbita hacia adelante* y una *semiórbita hacia atrás*.

Sin considerar el punto de vista de la medida, el problema de estudiar sistemas dinámicos donde f sea un homeomorfismo, dio nacimiento a la *Dinámica Topológica*, es decir, cómo estudiar propiedades dinámicas de los sistemas, utilizando fundamentalmente las herramientas estructurales de la Topología.

Las dificultades encontradas eran bastante grandes y pronto se vio además que lo que era más interesante de cara a las aplicaciones, era el considerar endomorfismos continuos. En estas condiciones, los problemas son aún más complicados, por lo que se intentó tratar con endomorfismos pero usar espacios de fase más sencillos y manejables. En cierto modo complicamos la función asociada al sistema, pero a cambio hacemos más simple el espacio de fases.

En los años 60 hubo un renacimiento del interés por estudiar *ecuaciones en diferencias finitas no lineales*. Las ecuaciones en diferencias finitas aparecieron con anterioridad a las ecuaciones diferenciales, sin embargo fueron eclipsadas por éstas debido a que siempre puede pensarse en términos de que una ecuación en diferencias finitas no es más que la imagen que se obtiene de una ecuación diferencial cuando esta última se observa en saltos discretos de tiempo. Por otra parte es indudable la gran potencia de los métodos asociados a las ecuaciones diferenciales y su enorme versatilidad en el modelado de fenómenos.

Aunque puedan haber otros precedentes, se puede suponer que la primera ecuación en diferencias finitas fue planteada hace aproximadamente 800 años por Leonardo di Pisa, alias Fibonacci, en un libro que publicó con el nombre de *Liber Abaci*. Fibonacci planteaba la determinación de un modelo para la descripción de la evolución de una población de conejos. Los términos del problema eran los siguientes: cada pareja de conejos en los instantes $t = 1$ y $t = 2$ de su vida dan lugar a dos nuevos conejos cada instante, uno macho y otro hembra, siendo el instante $t = 0$ el del nacimiento de tal pareja. Si denotamos por x_n el número de parejas de conejos en el instante n , entonces se verifica la siguiente ecuación

$$x_n = x_{n-1} + x_{n-2}, \quad \text{con } n > 1 \text{ y donde } x_0 = x_1 = 1.$$

El número de nuevas parejas sigue la sucesión de números (1, 1, 2, 3, 5, 8, 13, ...) conocida desde entonces como la *sucesión de Fibonacci*. Procediendo por

inducción se ve que

$$x_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1} \right].$$

Es interesante que

$$\lim_{n \rightarrow \infty} \frac{x_{n+1}}{x_n} = \frac{1 + \sqrt{5}}{2},$$

es decir, para n grande, la sucesión se comporta aproximadamente como una sucesión geométrica cuya razón es el *número de oro*. Esto también quiere decir que la población de conejos vive en condiciones en las que no se restringe su reproducción (comida suficiente, espacio, no existen depredadores, etc). Es evidente que tales condiciones son ideales y no se pueden dar en la realidad.

La ecuación en diferencias de Fibonacci era una ecuación lineal con dos retardos que no tiene dificultades para su resolución. Sin embargo, los casos más interesantes son aquellos en los que la ecuación es no lineal. Algunas de tales ecuaciones son de la forma

$$x_{n+1} = f(x_n),$$

especialmente las que provienen de la *Dinámica de Poblaciones*. Tales ecuaciones se pueden estudiar como el sistema dinámico (\mathbb{I}, f) , donde $\mathbb{I} = [0, 1]$, ya que la variable de estado x suele ser la densidad de una sola especie que ocupa un determinado hábitat.

Cuando la ecuación no proviene de modelos de población, entonces puede estudiarse como el sistema dinámico (\mathbb{R}, f) .

El desarrollo de los ordenadores supuso un gran avance en el estudio de estos sistemas dinámicos por las facilidades de calcular las órbitas de los puntos. En los años 60 un momento relevante fue la aparición en 1964 en la revista *Ukrainian Mathematical Journal* de un artículo del matemático Alexander N. Sharkovsky sobre la estructura periódica de las funciones continuas de la recta real en ella misma. Este bello artículo atrajo la atención de muchos matemáticos y el interés de investigar en estos temas, lo que en los años setenta y ochenta llevó a un enorme desarrollo de esta parte de las Matemáticas.

En la demostración del *teorema de Sharkovsky* se usa casi exclusivamente, como herramienta topológica, el teorema de los valores intermedios de las funciones continuas reales de variable real y mucha teoría combinatoria. De hecho la extensión de las ideas de este teorema a espacios de fase fundamentalmente de dimensión uno más generales que el intervalo o la recta real, ha dado lugar a la aparición de una parte de la Dinámica que se denomina *Dinámica Combinatoria*, donde se manejan objetos de naturaleza combinatoria tales como permutaciones, ciclos, grafos, digrafos, etc.

Uno de tales espacios es la circunferencia \mathbb{S}^1 , sobre la que se han desarrollado muchos modelos que se ajustan a un sistema dinámico del tipo (\mathbb{S}^1, f) que curiosamente fueron tratados por primera vez por H. Poincaré a finales del siglo XIX a propósito del problema de los tres cuerpos.

2 El Teorema de Sharkovsky

En cualquier exposición que se precie sobre Sistemas Dinámicos Discretos, es de obligado cumplimiento hacer mención a este resultado.

En los años cincuenta, en plena progresión de la guerra fría y en la extinta Unión Soviética, se desarrolló un programa para lanzar vehículos al espacio y sobre todo para hacerlo antes que los americanos. A tal fin se construyeron y desarrollaron ordenadores rudimentarios que a pesar de todo, cumplían con su misión. Estos ordenadores fueron usados en los centros de alta tecnología y también en las Academias de Ciencias (las instituciones más emblemáticas de la ciencia soviética). Todo este proceso culminó en parte, como es bien conocido, con el lanzamiento el día 4 de Octubre de 1957 de un vehículo espacial con el nombre de *Sputnik 1* (que en ruso quiere decir satélite). En esta época, entró a formar parte de la Academia de Ciencias de Ucrania en Kiev, un matemático joven: *Alexander N. Sharkovsky*. Su tutor le encomendó comprobar con la ayuda de un ordenador los resultados que existían en la literatura sobre la iteración de funciones reales de variable real. En particular, que si una función continua real de variable real tiene un punto periódico de período dos, entonces tiene puntos fijos (resultado debido a Poincaré) y si la función no tiene ningún punto periódico de período dos, entonces la órbita de cada punto de \mathbb{R} es covergente (Coppel). Sharkovsky probó dichos resultados y también el hecho de que si la función tiene un punto periódico de período mayor que dos, entonces siempre existe un punto periódico de período dos. Finalmente en 1964 publicó su famoso resultado. Es interesante señalar que tal resultado fue presentado en la Academia de Ciencias de Ucrania como parte de su Tesina de Licenciatura.

Supongamos que establecemos una ordenación (que denotamos por $>_s$, respectivamente por $<_s$) de los elementos del conjunto $\mathbb{N}^* = \mathbb{N} \cup \{2^\infty\}$ que denominaremos *orden de Sharkovsky*, de la forma

$$3 >_s 5 >_s 7 >_s \dots >_s 3 \cdot 2 >_s 5 \cdot 2 >_s 7 \cdot 2 >_s \dots >_s 3 \cdot 2^2 >_s 5 \cdot 2^2 >_s 7 \cdot 2^2 >_s \dots \\ >_s 3 \cdot 2^n >_s 5 \cdot 2^n >_s 7 \cdot 2^n >_s \dots >_s \{2^\infty\} >_s \dots >_s 2^n >_s \dots >_s 2^2 >_s 2 >_s 1.$$

Denotamos por $S(m)$ el *segmento inicial* del orden de Sharkovsky que termina en $m \in \mathbb{N}^*$,

$$S(m) = \{n \in \mathbb{N}^* : n \leq_s m\},$$

e igualmente

$$S(2^\infty) = \{1, 2, 2^2, \dots, 2^n, \dots\}.$$

Si ahora denotamos por $Per(f)$ el conjunto de *períodos* que una función continua puede tener, tenemos

Teorema 1 *Sea $f \in C(\mathbb{R})$. Entonces $Per(f) = \emptyset$ o $Per(f) = S(m)$ para algún $m \in \mathbb{N}^*$. Si es $f \in C(I)$, entonces es $Per(f) = S(m)$ para algún $m \in \mathbb{N}^*$.*

Recíprocamente, sea $m \in \mathbb{N}^$, entonces existe $f \in C(I)$ tal que $S(m) = Per(f)$. Dado el conjunto vacío \emptyset , existe $f \in C(\mathbb{R})$ tal que $Per(f) = \emptyset$.*

La interpretación del resultado es clara. Si por ejemplo, una función continua tiene un punto periódico de período 15, entonces tiene puntos periódicos de períodos todos los que le siguen a 15 en el orden de Sharkovsky. El símbolo $\{2^\infty\}$ significa que la función tiene puntos periódicos de cualquier potencia de dos.

En la literatura ha habido comentarios curiosos acerca de este teorema. En [31], J. Smítal señala que “de acuerdo con la opinión de matemáticos sobresalientes, el teorema está entre lo mejor que se ha obtenido en matemáticas en los últimos veinte años” (frase de 1988). En [16] se dice: “este bonito e interesante resultado depende únicamente de la continuidad de la función en cuestión y debe servir para poner en cuarentena la opinión de aquellos que piensan que los teoremas realmente buenos que no sean de los del tipo de los que se puedan comparar con los cazabombarderos B-1 se probaron antes del final del siglo XIX”. En [29], D. Ruelle comenta que: “este teorema demuestra, entre otras cosas, que haríamos mal en menospreciar el contenido de las revistas matemáticas ucranianas”.

Un caso particular es el siguiente resultado, probado curiosamente sin tener conocimiento del resultado anterior por los matemáticos Li y Yorke en [20] en 1975.

Teorema 2 *Si $f \in C(I)$ tiene un punto periódico de período tres, entonces tiene puntos periódicos de todos los períodos.*

El artículo donde está contenido tal resultado, publicado en *American Mathematical Monthly*, ejerció una influencia enorme en las investigaciones sobre Sistemas Dinámicos Discretos de los años 70, sobre todo por el título: “*Period three implies chaos*” ya que era la primera vez que en un artículo de matemáticas aparecía la palabra *caos*, dando lugar además a una noción que desde entonces se conoce como *caos en el sentido de Li y Yorke*.

Aunque existen varias estrategias a seguir para la demostración del teorema de Sharkovsky, es especialmente relevante señalar la introducida por P. Straffin en [32] mediante el uso de *grafos* y *digrafos*, ya que es exportable a los casos de funciones continuas en otros espacios de dimensión uno tales como la *circunferencia*, *árboles*, *grafos*, *dendritas*, etc. En estos casos se obtienen resultados semejantes a los de Sharkovsky, pero aparecen nuevos órdenes entre los períodos de las funciones, por lo que la *estructura periódica* se complica.

Por estructura periódica entendemos una relación de forzamiento entre los períodos de una función que permita averiguar si la misma posee un punto periódico de un determinado período sabiendo que posee puntos periódicos de otros períodos. Este es el caso del orden de Sharkovsky, donde el orden entre los elementos de \mathbb{N}^* es total. En las referencias [25], [1], [2] y también en el artículo de tutoría [9] se hace referencia detallada a estos hechos. A modo de ejemplo, consideremos el caso de $f \in C(\mathbb{Y})$, donde $\mathbb{Y} = \{z \in \mathbb{C} : z^3 \in [0, 1]\}$ es el espacio conocido como *triodo*, donde se supone que 0 es un punto fijo de la función. En tales condiciones, además del citado orden de Sharkovsky, tenemos

que introducir otros dos órdenes, el *orden verde* en $\mathbb{N} \setminus \{2\}$:

$$5, 8, 4, 11, 14, 7, 17, 20, 10, 23, 26, 13, \dots, 3.3, 3.5, 3.7, \dots, 3.2.3, 3.2.5, 3.2.7, \dots, \\ \dots, 3.2^2.3, 3.2^2.5, 3.2^2.7, \dots, 3.2^3, 3.2^2, 3.1, 1.$$

El *orden rojo* consiste en ordenar el conjunto $\mathbb{N} \setminus \{2, 4\}$:

$$7, 10, 5, 13, 16, 8, 19, 22, 11, 25, 28, 14, \dots, 3.3, 3.5, 3.7, \dots, 3.2.3, 3.2.5, 3.2.7, \dots \\ \dots, 3.2^2.3, 3.2^2.5, 3.2^2.7, \dots, 3.2^3, 3.2^2, 3.2, 3.1, 1.$$

Si m_s, m_v, m_r representan un elemento de los órdenes de Sharkovsky, verde y rojo, tenemos:

Teorema 3 *Se satisfacen las siguientes condiciones:*

(a) *Sea $f \in C(\mathbb{Y})$, entonces $Per(f) = S(m_s) \cup S(m_v) \cup S(m_r)$.*

(b) *Supongamos dados m_s, m_v y m_r , entonces existe $f \in \mathbb{Y}$ tal que*

$$Per(f) = S(m_s) \cup S(m_v) \cup S(m_r).$$

La demostración y consecuencias de este resultado pueden verse en [3].

Usando los resultados sobre estructura periódica en espacios de dimensión uno, se pueden obtener resultados semejantes sobre ecuaciones en diferencias finitas con retardo del tipo

$$x_n = f(x_{n-k}), \quad \text{con } k \geq 2,$$

y donde la variable x pertenece a \mathbb{X} , que es cualquier espacio de dimensión uno de los estudiados o bien un espacio del que se conozca la estructura periódica. La formulación de esta estructura periódica se realizará fácilmente con ayuda de transformaciones n -dimensionales en la última sección de este trabajo.

3 El método de Newton-Raphson

Como una ilustración de los métodos iterativos usados en los Sistemas Dinámicos Discretos de la forma (I, f) , vamos a poner de manifiesto algunos hechos significativos en relación al método de Newton-Raphson para la búsqueda de soluciones de las ecuaciones $f(x) = 0$. Supongamos que f sea una función real de variable real diferenciable y que x_0 sea una aproximación razonable de una solución de la ecuación anterior. Se enseña en los cursos elementales de Cálculo que si es

$$N_f(x) = x - \frac{f(x)}{f'(x)},$$

entonces, en muchas ocasiones, encontramos una solución aproximada de la ecuación mediante la ecuación en diferencias finitas

$$x_{n+1} = N_f(x_n).$$

Pero es también conocido que esta ecuación, cuando empezamos en el punto x_0 , no genera necesariamente una sucesión convergente a algún valor y para el que se cumpla $f(y) = 0$. Es más, en muchos casos las iteradas sucesivas del punto x_0 se comportan de una manera errática y vagan por la recta real de una forma caótica.

El comportamiento de todas las órbitas $Orb_f(x_0) = (N_f^n(x_0))_{n=0}^{\infty}$ es conocido cuando la función f es una función polinómica sencilla tal como la cuadrática y la cúbica. Tal comportamiento se puede entender mediante el uso de la noción de *conjugación topológica*. Sean D, E dos espacios topológicos y $f : D \rightarrow D$, $g : E \rightarrow E$ funciones. Decimos que f y g son *topológicamente conjugadas* si existe un homeomorfismo $h : D \rightarrow E$ tal que se verifica $h \circ f = g \circ h$ en D . Esta definición puede tener carácter local si con la misma definición nos limitamos a dos entornos de dos puntos en los espacios D y E .

Supongamos que $p(x)$ sea un polinomio y que hemos realizado la cancelación (si es posible) de factores comunes en la expresión de $N_p(x)$ (la función de Newton asociada a $p(x)$). Entonces tenemos el siguiente resultado, donde decimos que un punto *periódico* x de período k es *atractor* (respectivamente *repulsor*) si se verifica que $|(p^k)'(x)| < 1$ (respectivamente $|(p^k)'(x)| > 1$). Para funciones diferenciables valen las mismas definiciones.

Teorema 4 *Un número es un punto fijo de $N_p(x)$ si, y solo si, es una raíz del polinomio y todos los puntos fijos de $N_p(x)$ son atractores.*

En el contexto de funciones diferenciables, una consecuencia de la demostración del anterior resultado es la siguiente.

Corolario 5 *Si f es diferenciable, entonces un punto fijo de $N_f(x)$ debe ser un punto atractor de la misma función y una raíz de la ecuación $f(x) = 0$. Por otra parte, si un punto es una raíz de f y $N_f(x)$ está definida en tal punto, entonces el punto es un punto fijo atractor de $N_f(x)$.*

Volviendo ahora a los polinomios de grado dos y tres, tenemos los siguientes resultados.

Proposición 6 *Sean $p(x) = ax^2 + bx + c$ y $q(x) = x^2 - A$, donde $A = b^2 - 4ac$. Entonces $h(x) = 2ax + b$ es una conjugación topológica entre $N_f(x)$ y $N_q(x)$.*

Como consecuencia de esta Proposición, podemos restringirnos en el análisis del método de Newton para polinomios cuadráticos al de polinomios de la forma $q(x) = x^2 - c$. Si $c > 0$, la función $N_q(x)$ tiene puntos fijos en $\pm\sqrt{c}$ y ambos son puntos atractores. Además el conjunto estable de \sqrt{c} es $(0, \infty)$ y el de $-\sqrt{c}$ es el intervalo $(-\infty, 0)$ (dado un punto periódico x de período k , la *variedad estable asociada a x* es el conjunto de todos los puntos y para los que se verifica que $\lim_{n \rightarrow \infty} f^{nk}(x) = y$).

Si $c = 0$ es inmediato que 0 es el único punto fijo de $N_q(x)$ y que \mathbb{R} es su conjunto estable.

Cuando $c < 0$, entonces tras un análisis no muy complicado obtenemos que la función $N_q(x)$ se comporta de una forma caótica en el conjunto $(-\infty, 0) \cup (0, +\infty)$, es decir las órbitas de todos los puntos de tal conjunto toman valores de una forma errática sin convergencia.

El caso de los polinomios cúbicos es mucho más interesante. Los resultados que se enuncian a continuación nos dan los comportamientos de las respectivas funciones de Newton.

Proposición 7 *Si $p(x) = ax^3 + bx^2 + cx + d$ y $q(x) = x^3 + Ax + B$ donde $A = 9ac - 3b^2$ y $B = 27a^2d + 2b^3 - 9abc$, entonces la función $h(x) = 3ax + b$ es una conjugación topológica de $N_p(x)$ en $N_q(x)$.*

Proposición 8 *Sean $p(x) = x^3 + ax + b^3$ con $b \neq 0$ y $q(x) = x^3 + cx + 1$ donde $c = \frac{a}{b^3}$. Entonces $N_p(x)$ y $N_q(x)$ son topológicamente conjugadas a través de $\tau = \frac{1}{b}x$.*

Estos dos resultados implican que la dinámica de la función de Newton para un polinomio cúbico es la misma que la dinámica de los polinomios de la familia parametrizada $f_c(x) = x^3 + cx + 1$ o los de la familia $g_a(x) = x^3 + ax$ (correspondiente al caso $b = 0$).

En este último caso, a su vez, la dinámica de las correspondientes funciones de Newton son topológicamente conjugadas a uno de los tres polinomios

$$p_+(x) = x^3 + x, \quad p_-(x) = x^3 - x \quad \text{o} \quad p_0(x) = x^3,$$

siendo la conjugación la dada por $\tau(x) = \frac{1}{a}x$.

Si ahora hacemos el análisis de lo que le ocurre a los polinomios p_+, p_- y p_0 , obtenemos el siguiente resultado,

Proposición 9 *Sean $p_+(x) = x^3 + x$ y $p_0(x) = x^3$. Entonces 0 es el único punto fijo de $N_{p_+}(x)$ y de $N_{p_0}(x)$ y el correspondiente conjunto estable para ambos casos es \mathbb{R} .*

El caso $p_-(x) = x^3 - x$ es mucho más complicado y hay que efectuar un detallado estudio del comportamiento de todos los puntos de \mathbb{R} , donde se encuentra por ejemplo que el conjunto estable del punto fijo 0 es $(-\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}})$.

Para la familia $f_c(x) = x^3 + cx + 1$, la dinámica de la función de Newton es sencilla en los casos $c > 0$ y $c = 0$. En efecto, en el primer caso la función de Newton tiene exactamente una raíz y su conjunto estable es \mathbb{R} . En el segundo, existe también una raíz real y su conjunto estable contiene a todos los números reales excepto al cero.

Cuando $c < 0$, la situación vuelve a ser complicada y el análisis detallado del comportamiento de la función de Newton hay que hacerlo en cada caso particular. Sin embargo, por analogía con la dinámica de la función logística $f_c(x) = cx(1 - x)$, en [16] se ha establecido la siguiente conjetura.

Conjetura *Sea $f(x) = x^3 - cx + 1$, entonces existe $c > 0$ y un conjunto de números reales en los que la función de Newton es caótica. Además este conjunto es atractor.*

De todo lo anterior se desprende la importancia que tiene una adecuada elección del punto inicial x_0 al aplicar el método de Newton. Puede ocurrir que con el punto elegido el método no converja, errando periódicamente o teniendo un comportamiento caótico. Afortunadamente existen tests para determinar si el punto elegido inicialmente converge o no a la raíz más cercana de la ecuación. De todas formas, uno se puede preguntar por qué si el método de Newton tiene tantos problemas, cómo se sigue usando. La contestación está en la velocidad de convergencia, que en este caso es cuadrática. Esto significa esencialmente que el número de dígitos correctos a la derecha de la coma donde comienza la parte decimal de los números de nuestra estimación, se dobla con cada iteración del método de Newton. Como consecuencia, si podemos elegir una aproximación razonablemente cerca de la raíz, entonces podríamos obtener una muy buena aproximación de la raíz con unos pocos pasos. Como contraste a lo dicho, en el método de *bisección*, que se aplica frecuentemente, la convergencia es geométrica. Es decir, nos lleva aproximadamente 3 iteraciones el poder añadir un dígito correcto a la parte decimal de la aproximación.

Como ilustración de lo anterior, el siguiente resultado que puede verse en [26] nos da con precisión la razón de convergencia en el método de Newton.

Teorema 10 *Sea $f(x)$ una función dos veces derivable y supongamos que el valor $f'(x_0) \neq 0$. Sean*

$$\varepsilon = \left| \frac{f(x_0)}{f'(x_0)} \right|, \quad x_1 = N_f(x_0), \quad e \quad I = [x_1 - \varepsilon, x_1 + \varepsilon].$$

Si $|f''(x)| < \frac{|f'(x_0)|^2}{2|f(x_0)|}$ para todo $x \in I$, entonces $N_f^n(x_0)$ converge a una raíz de f en I y

$$|N_f^{n+1}(x_0) - N_f^n(x_0)| \leq \frac{1}{\varepsilon} |N_f^n(x_0) - N_f^{n-1}(x_0)|^2.$$

El comportamiento de la función de Newton para un polinomio está lejos de haber sido bien entendido. Podemos presentar para terminar esta sección, el siguiente problema abierto,

Problema 1 *Sea $q(x) = ax^4 + bx^3 + cx^2 + dx + e$. Determinar un polinomio de cuarto grado con un número menor de parámetros para el que la función de Newton sea topológicamente conjugada a $N_q(x)$. ¿Cuál es el mínimo número de parámetros que se necesitan?*

Con relación al tema del conjunto de números reales donde no converge el método de Newton para polinomios, podemos recordar el resultado de B. Barna en [10]:

Teorema 11 *Sea $p(x) = \prod_{i=0}^n (x - r_i)$, donde $r_i \neq r_j$ si $i \neq j$, $r_i \in \mathbb{R}$ para $i = 0, 1, \dots, n - 1$ y $n > 4$. Si*

$$S = \{x \in \mathbb{R}, N_p^m(x) \rightarrow r_i \text{ cuando } m \rightarrow \infty \text{ para algún } i = 0, \dots, n - 1\},$$

entonces el conjunto $R \setminus S$ es homeomorfo a un conjunto de Cantor (excluyendo un conjunto numerable).

Ni qué decir tiene que dado un polinomio p de coeficientes reales, parte de sus raíces pueden ser complejas. El sistema dinámico $(\mathbb{C}, N_q(z))$ donde q es un polinomio con coeficientes complejos es muy interesante y ha sido objeto de consideración. A tal respecto señalaremos otro resultado clásico que tiene que ver con la condición de atractividad o no de las raíces de la ecuación $q(z) = 0$.

Teorema 12 *Para cada $d \geq 2$ existe un polinomio p con coeficientes reales tal que la función asociada de Newton tiene $2d - 2$ órbitas periódicas diferentes y atractoras en el plano complejo.*

De acuerdo con un teorema de Julia, éste es el número más grande de órbitas atractoras que cualquier función racional de grado d puede poseer.

4 Sistemas dinámicos discretos bidimensionales

En este apartado, nos centraremos en la consideración de sistemas dinámicos de la forma (\mathbb{X}^2, F) , donde el espacio de fases \mathbb{X} es de dimensión dos. La razón es que aparecen en muchas aplicaciones y en otros casos ayudan a comprender el comportamiento de otros sistemas más complicados.

Nos vamos a concentrar en cuatro tipos de aplicaciones que conducen a dichos sistemas aunque existen muchas otras:

- (a) las aplicaciones de Poincaré para el estudio de sistemas de tres ecuaciones diferenciales ordinarias,
- (b) ecuaciones en diferencias finitas no lineales con dos retardos y modelos de la Dinámica de Poblaciones,
- (c) transformaciones bidimensionales invertibles que poseen atractores extraños,
- (d) modelos de la Teoría Económica.

4.1 Aplicaciones de Poincaré

Es una de las herramientas más útiles desarrolladas por la *Teoría Cualitativa* de Ecuaciones Diferenciales Ordinarias. Sea $f \in C^1$ y supongamos que queremos entender el comportamiento de las órbitas del sistema cerca de un punto crítico del mismo o de una órbita periódica. Esto se puede hacer a través de la de la noción de *aplicación de Poincaré*. La idea es muy simple. Imaginemos un sistema autónomo de ecuaciones en \mathbb{R}^3 dado por

$$\dot{x} = f(x)$$

y supongamos que de una órbita periódica tenemos que decidir sobre su estabilidad orbital o bien ver sobre lo que puede ocurrir cerca de una *órbita*

homoclínica a un punto crítico. Por un punto de dicha órbita trazamos una *sección local* a la misma (esto se puede hacer de muchas formas), que en este caso es un trozo de plano S , y supongamos que somos capaces de determinar los puntos en los que la órbita que inicialmente pasa por x_0 vuelve a cortar a la sección transversal y así sucesivamente. De esta forma se puede construir una aplicación P de S en sí mismo, que resulta ser un homeomorfismo y que es la aplicación de Poincaré citada.

El resultado es el sistema dinámico bidimensional (S, P) . El conocimiento de la órbita de un punto de este sistema nos da una idea de la evolución de la órbita del sistema de ecuaciones tridimensional.

Esta forma de actuar se ha mostrado de gran utilidad en bastantes problemas. Uno de ellos, de interés histórico, es el siguiente.

El sistema no lineal de ecuaciones diferenciales

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y, \\ \dot{y} &= \rho x - y - xz, \\ \dot{z} &= xy - \beta z,\end{aligned}$$

fue introducido en 1963 por E. Lorenz como un modelo muy rudimentario para la predicción del tiempo atmosférico. De una forma bastante casual y haciendo análisis numérico, Lorenz descubrió un fenómeno que posteriormente se denominó *dependencia sensible a las condiciones iniciales* que equivalía a afirmar la *impredecibilidad* del desarrollo futuro de las órbitas del sistema de ecuaciones y en consecuencia la *impredecibilidad del tiempo atmosférico*.

Para los valores de los parámetros $\sigma = 10$, $\rho = 8/3$ y $\beta = 28$, Lorenz descubrió además que existía una región en \mathbb{R}^3 que atraía a casi todas las órbitas del sistema de ecuaciones y cuya estructura geométrica era bastante complicada. Posteriormente para designar a este tipo de conjuntos se acuñó el nombre de *atractores extraños* y se precisó su definición. En el caso del sistema de Lorenz, se denominó *el atractor de Lorenz*. El probar de forma rigurosa que efectivamente dicho conjunto era un atractor extraño ha sido un problema abierto durante más de 35 años. Finalmente en 1999, W. Tucker (ver [34],[35]) lo ha conseguido. Su demostración usa como hecho fundamental la construcción de una adecuada aplicación de Poincaré que verifica ciertas desigualdades que a su vez se han probado rigurosamente con ayuda del ordenador.

4.2 Ecuaciones en diferencias finitas con dos retardos y modelos de la dinámica de poblaciones.

Las ecuaciones en diferencias finitas con dos retardos no lineales pueden ser formuladas como

$$x_{n+1} = F(x_n, x_{n-1}),$$

haciendo ahora

$$\begin{aligned}x_{n-1} &= x, \\ x_n &= y.\end{aligned}$$

La ecuación anterior se puede ver como la transformación

$$T(x, y) = (y, F(x, y))$$

y tratarla como un sistema dinámico (\mathbb{R}^2, T) bidimensional.

Muchas de estas ecuaciones han sido propuestas como modelos de comportamiento de poblaciones de una sola especie donde la generación de nuevos individuos está influenciada por ciertos mecanismos de retraso. Un caso particular interesante nos lo proporciona el modelo de dinámica de evolución de una sola especie en un hábitat cuando la densidad de la misma en un instante determinado depende del nivel alimenticio disponible en dos instantes anteriores. En tal caso se obtiene la ecuación

$$x_{n+1} = ax_n(1 - x_{n-1}),$$

que se conoce como *ecuación logística bidimensional* y que se puede estudiar como el sistema dinámico (I^2, F) donde $I^2 = [0, 1] \times [0, 1]$ y

$$F(x, y) = (y, ay(1 - x)),$$

ya que x e y representan densidades de población. Este modelo fue propuesto por Maynard Smith en 1968 ([22]) y ha sido muy estudiado (véase por ejemplo las referencias [27], [4] y [28]), pero quedan muchas cosas de él que no han sido suficientemente entendidas. Vamos a hacer una descripción general y señalar algunos de los problemas pendientes.

En el cuadrado unidad, el sistema tiene un comportamiento dinámico que depende del valor del parámetro a y posee dos puntos fijos $P(0, 0)$ y $Q(\frac{a-1}{a}, \frac{a-1}{a})$ que son conectados por una *curva invariante atractora* para $a < 2$. Para el valor $a = 2$ se tiene una *bifurcación de Hopf* y el punto Q que era atractor pasa a ser repulsor y aparece una *curva cerrada invariante y atractora* que lo contiene. La curva va creciendo hasta ser tangente a los ejes coordenados. Al mismo tiempo, cerca de P van apareciendo múltiples lazos y bucles que producen una estructura geométrica cada vez más complicada. El correspondiente valor de $a = a^*$ para el que la curva es tangente a $y = 0$ resulta crítico, ya que a partir de él las iteradas de algunos puntos que están dentro del recinto de la curva comienzan a salir fuera de I^2 y a tender a infinito, que se pueden interpretar biológicamente como situaciones en las que la especie se extingue.

A partir de a^* el proceso se agudiza, los bucles se hacen cada vez más complicados y la mayor parte de los puntos de I^2 escapan al exterior en un número finito de iteraciones. A partir de un nuevo valor crítico $a = a^{**}$, en el cuadrado unidad conjeturamos que sólo permanece un conjunto *invariante de tipo Cantor repulsor*.

El comportamiento descrito se ha llevado a cabo usando aproximaciones numéricas y simulaciones del modelo y está pendiente el que se dé una justificación rigurosa del mismo a partir de $a = 2$.

Problemas abiertos.

En este modelo se pueden señalar los siguientes:

- (1) Precisar los períodos de las órbitas periódicas que aparecen en la transformación y probar si existe o no una estructura periódica de los mismos.
- (2) La curva invariante cerrada es derivable cuando a está próximo a 2, pero cuando a va aumentando, aparecen intervalos del parámetro donde la curva pierde la derivabilidad y se transforma en un conjunto con una estructura que recuerda a la del atractor de Hénon. Después de intervalos donde esto ocurre vuelven a aparecer otros intervalos donde la curva recupera su derivabilidad.
- (3) El conjunto interior a la curva invariante cerrada adquiere una estructura filamentososa donde se presenta el fenómeno de la *dependencia sensible a las condiciones iniciales*.
- (4) Aclarar la estructura del conjunto tipo Cantor.

Otros modelos aparecidos en la literatura que son interesantes de estudiar son los que siguen:

- La ecuación introducida por Lauwerier en 1986 ([19])

$$x_{n+2} = ax_{n+1}[1 - bx_n - (1 - b)x_{n+1}]$$

extiende la ecuación logística bidimensional anterior y se puede estudiar mediante un sistema dinámico con un amplio espectro de comportamientos que están lejos de estar bien entendidos.

- Otro ejemplo destacable en las aplicaciones es el modelo *depredador-presa discreto* introducido por Maynard Smith en 1968 ([23]),

$$\begin{aligned} x_{n+1} &= ax_n - bx_n^2 - cx_ny_n, & a > 1, b > 0, c > 0, \\ y_{n+1} &= dx_ny_n, & d > 0 \end{aligned}$$

que conduce al sistema dinámico,

$$F(x, y) = (ax - bx^2 - cxy, dxy)$$

- Como un ejemplo de modelo propuesto para describir las relaciones huéspedes-parásitos tenemos el de Crofton de 1971 ([13]),

$$\begin{aligned} x_{n+1} &= \lambda x_n [1 + y_n]^{-k}, \\ y_{n+1} &= x_n y_n [1 + y_n]^{-k-1}. \end{aligned}$$

4.3 Algunas transformaciones bidimensionales invertibles.

La posibilidad de poder obtener órbitas hacia adelante y hacia atrás en los sistemas dinámicos bidimensionales es una ventaja que tienen algunas de las transformaciones que asociadas a diferentes clases de modelos han aparecido desde los años 70, motivadas por diferentes problemas.

Consideremos sistemas de ecuaciones en diferencias finitas de la forma

$$\begin{aligned}x_{n+1} &= y_n, \\y_{n+1} &= f(a, y_n) + bx_n,\end{aligned}$$

donde $f(a, y)$ es una función no lineal dependiente del parámetro a y donde b es un parámetro de acoplamiento entre las variables x e y . La forma del correspondiente sistema dinámico es (\mathbb{R}^2, F) , donde

$$F(x, y) = (y, f(a, y) + bx).$$

Históricamente, una de las primeras transformaciones del tipo citado fue propuesta por M. Hénon en 1976 ([14]) y por Hénon y Pomeau en 1977 ([15]) y corresponde a la elección $f(a, y) = 1 - ay^2$. La motivación original para estudiarla fue “encontrar un modelo discreto que fuera muy sencillo y que tuviera esencialmente las mismas propiedades que el sistema de ecuaciones diferenciales de Lorenz”. En efecto, para $a = 1,8$ y $b = 0,4$, usando simulación numérica se conjeturó que el sistema posee un atractor extraño. Su demostración matemática es un problema que aún permanece abierto para el rango de parámetros citado. Para $a = 0,4$ y b próximo a cero, utilizando las propiedades dinámicas de las funciones $f(a, y) = 1 - ay^2$ fue demostrada la existencia de un atractor extraño ([11]).

En 1978, M. Lozi introdujo una transformación semejante a la de Hénon pero usando una función lineal a trozos, es decir, $f(a, y) = 1 - a|y|$. Dicha transformación tiene atractores extraños. Este hecho fue probado por Misiurewicz en 1980 ([24]) para un rango limitado de parámetros y después para una rango más grande, $0 < b < 2/3$ y $a + b < a < 2 - b/2$, por Liu, Xie, Zhu y Lu en 1992 ([21]).

En 1983, T. Tél ([33]) en conexión con la dimensión fractal de un atractor extraño, introdujo otra transformación con $f(a, y) = ay - \text{sgn}(y)$, donde

$$\text{sgn}(y) = \begin{cases} 1, & \text{si } y > 0, \\ -1, & \text{si } y < 0. \end{cases}$$

Los tres ejemplos citados tienen en común que son difeomorfismos para ciertos valores de los parámetros. La mayoría de los resultados sobre ellos de la literatura lo han sido por simulación numérica. Quedan muchos fenómenos sin ser explicados rigurosamente. Como ocurría con la aplicación logística bidimensional, dichos fenómenos dependen fuertemente de los valores de los parámetros.

Como una técnica sistemática para poder justificar la aparición de órbitas periódicas y órbitas caóticas, se ha introducido una *Dinámica Simbólica* adaptada a estos modelos bidimensionales. Esta técnica, que proporciona muy buenos resultados para los casos unidimensionales, tiene en el caso bidimensional serias limitaciones derivadas del hecho de que no se puede disponer de la relación de orden de los números reales. Sin embargo, se ha podido obtener resultados parciales, como por ejemplo la existencia en las tres transformaciones de puntos periódicos de períodos hasta siete.

Queda por ver si existe algún tipo de estructura periódica, aunque todo parece indicar que esto no es posible.

Finalmente, señalar una transformación de un anillo bidimensional en sí mismo dada por

$$\begin{aligned}r_{n+1} &= a + br_n + K \operatorname{sen} 2\pi\theta_n, \\ \theta_{n+1} &= \theta_n + r_{n+1} \pmod{1},\end{aligned}$$

donde r y θ representan las coordenadas polares. Esta transformación se denomina la *transformación disipativa estándar* y la dinámica simbólica que se puede aplicar con relación a la misma sirve para entender el comportamiento del sistema de dos ecuaciones diferenciales conocido como *Brusselator excitado periódicamente* que modela la cinética química tri-molecular. Se trata del sistema de ecuaciones

$$\begin{aligned}\dot{x} &= A - (B + 1)x + x^2y + \alpha \cos \omega t, \\ \dot{y} &= Bx - x^2y,\end{aligned}$$

donde x e y son las concentraciones de productos intermedios, A y B son concentraciones de productos químicos bajo control, y α y ω son respectivamente la amplitud del forzamiento y la frecuencia del mismo. La existencia de un atractor previsto numéricamente en este sistema de ecuaciones puede ser justificado parcialmente por el conocimiento de la dinámica de la anterior transformación.

4.4 Transformaciones bidimensionales que provienen de la teoría económica

En los últimos veinte años han proliferado en la Teoría Económica las transformaciones bidimensionales que modelan diferentes fenómenos del comportamiento económico. Estos modelos casi siempre incluyen parámetros que representan comportamientos exógenos de los mismos y que como consecuencia, sus valores no pueden ser fijados con precisión. La Teoría de las Bifurcaciones juega un papel determinante, es decir, estudia los cambios que se experimentan en la dinámica del sistema cuando los parámetros alcanzan unos determinados valores críticos.

En lo que sigue nos referiremos a dos problemas antiguos y que han tenido cierta relevancia.

En 1836, Cournot introdujo en [12] un modelo que modernamente se puede interpretar a la luz de la Teoría de los Juegos, en el que existen dos jugadores o empresas que están presentes en un cierto sector productivo y que su producción en cada instante de tiempo depende de lo que la otra empresa haya producido en el instante anterior. Si se denota por q_1 y q_2 la cantidad de producto de cada empresa, entonces resulta

$$\begin{aligned}q_2 &= f(q_1), \\ q_1 &= g(q_2),\end{aligned}$$

y en consecuencia

$$(q_{1(n+1)}, q_{2(n+1)}) = (g(q_{2n}), f(q_{1n})).$$

Esta transformación se conoce como el *duopolio de Cournot*.

Este modelo se puede ver como un sistema dinámico bidimensional de la forma (\mathbb{R}^2, F) donde

$$F(x, y) = (g(y), f(x)).$$

Este sistema dinámico se ha denominado *sistema dinámico antitriangular* y resulta ser un caso particular de una transformación n -dimensional que veremos en el siguiente epígrafe.

En 1940, N. Kaldor [18] propuso el siguiente modelo para estudiar las relaciones entre las inversiones y los niveles de stock de capital en un sector productivo:

$$\begin{aligned} Y_{n+1} &= \alpha(I(Y_n, K_n) - S(Y_n, K_n)) + Y_n, \\ K_{n+1} &= I(Y_n, K_n) + (1 - \delta)K_n, \end{aligned}$$

donde $I(Y_n, K_n)$ representa la función de inversión en un sector productivo de la economía que depende de los ingresos y del nivel de stock de capital; $S(Y_n, K_n)$ representa la función de ahorro del sector, α es un coeficiente de ajuste y δ representa una constante de razón de depreciación. El sistema anterior de dos ecuaciones en diferencias finitas se puede ver como un sistema dinámico bidimensional de la forma (\mathbb{R}^2, F) donde

$$F(x, y) = (f(x, y), g(x, y)),$$

que es muy difícil de estudiar. De hecho, solo son conocidos resultados parciales. Variando el parámetro α , cuando alcanza el valor

$$\alpha_0 = \frac{\delta - \frac{\partial I}{\partial K}}{\left(\frac{\partial I}{\partial Y} - \frac{\partial S}{\partial Y}\right)\left(\frac{\partial I}{\partial K} + 1 - \delta\right) - \frac{\partial I}{\partial Y}\left(\frac{\partial I}{\partial K} - \frac{\partial S}{\partial K}\right)}$$

se tiene una *bifurcación de Hopf*, en el que un punto fijo atractor explota a una curva cerrada invariante atractora. Este hecho tiene relevancia en la interpretación económica. Sin embargo, poco más se sabe del modelo.

5 Transformaciones n -dimensionales

De cara a las aplicaciones, introducimos el siguiente sistema dinámico de dimensión n . Sea σ una permutación cíclica de los números $1, 2, \dots, n$ y \mathbb{X} un espacio métrico. La aplicación $F : \mathbb{X}^n \rightarrow \mathbb{X}^n$ se dice que es una σ -permutación si

$$F(x_1, x_2, \dots, x_n) = (f_{\sigma(1)}(x_{\sigma(1)}), f_{\sigma(2)}(x_{\sigma(2)}), \dots, f_{\sigma(n)}(x_{\sigma(n)})),$$

donde $f_i : \mathbb{X} \rightarrow \mathbb{X}$ es continua para $i = 1, 2, \dots, n$.

Estas aplicaciones aparecen en la Teoría Económica (duopolio de Cournot) y también en Dinámica de Poblaciones para modelar el comportamiento que se conoce como *cooperación cíclica* de varias especies animales en la sabana africana (ver [17]), donde cuatro especies de animales cooperan comiendo distintas partes de las plantas de unas determinadas extensiones y contribuyendo a la regeneración del espacio.

Cuando $\mathbb{X} = \mathbb{R}, \mathbb{I}$ o \mathbb{S}^1 y debido a que se conoce la estructura periódica de las funciones continuas sobre estos espacios, se ha podido caracterizar la estructura periódica de las aplicaciones σ -permutación (ver [6] y [7]). A modo de ejemplo, para el caso $\mathbb{X} = \mathbb{I}$ y $n \geq 2$ se puede dar el siguiente resultado, donde

$$S_n(m) = \{k \in \mathbb{N}; k \nmid n \text{ y } \frac{k}{\text{mcd}(k, n)} \in S(m) \cup \{1\}\}$$

y

$$f_i^{(j)} = f_{\sigma(i)} \circ f_{\sigma^2(i)} \circ \dots \circ f_{\sigma^j(i)}.$$

Teorema 13 *Sea F una σ -permutación en \mathbb{I}^n , $n \geq 2$. Entonces existe $m \in \mathbb{N}^*$ tal que*

$$\text{Per}(F) = S_n(m) \quad \text{o} \quad \text{Per}(F) = S_n(m) \cup \{p : p \mid n\}.$$

Tenemos además que $\text{Per}(f_i^{(n)}) = S(m)$ para cada $i \in \{1, \dots, n\}$. Más aún, $\text{Per}(F) \neq S_n(m)$ si, y solo si, $f_i^{(n)}$ tiene por lo menos dos puntos fijos.

Supongamos que $\mathcal{P} = S_n(m)$ o $\mathcal{P} = S_n(m) \cup \{p : p \mid n\}$ para algún $m \in \mathbb{N}^*$. Entonces existe una aplicación σ -permutación de \mathbb{I}^n en \mathbb{I}^n verificando $\text{Per}(F) = \mathcal{P}$.

Como ilustración de este resultado, consideremos el caso $n = 3$. La estructura periódica de la correspondiente aplicación σ -permutación puede obtenerse a través del siguiente diagrama, donde $a \Rightarrow b$ significa que la existencia de un período de período a implica la existencia de otro de período b . La notación $a \Leftrightarrow b$ significa que ambos períodos se fuerzan mutuamente.

$$\begin{aligned} 3.3 &\Rightarrow (3.5 \Leftrightarrow 5) \Rightarrow (3.7 \Leftrightarrow 7) \Rightarrow 3.9 \\ &\Rightarrow (3.11 \Leftrightarrow 11) \dots \Rightarrow 3.2.3 \Rightarrow \\ (3.2.5 \Leftrightarrow 2.5) &\Rightarrow (3.2.7 \Leftrightarrow 2.7) \Rightarrow 3.2.9 \Rightarrow \\ &\dots \\ &\Rightarrow 3.2^k.3 \Rightarrow (3.2^k.5 \Leftrightarrow 2^k.5) \Rightarrow \\ &\Rightarrow (3.2^k.7 \Leftrightarrow 2^k.7) \Rightarrow 3.2^k.9 \Rightarrow \\ &\dots \\ &\Rightarrow (3.2^m \Leftrightarrow 2^m) \Rightarrow \dots (3.2^2 \Leftrightarrow 2^2) \Rightarrow \\ &(3.2 \Leftrightarrow 2) \Rightarrow 1. \end{aligned}$$

A lo anterior hay que añadir que existirá un punto periódico de período 3 si, y solo si, las funciones $f_i^{(3)}$ para $i = 1, 2, \dots, n$ tienen por lo menos dos puntos fijos.

Otra aplicación interesante del resultado anterior se encuentra en la consideración de las ecuaciones en diferencias finitas con retardo

$$x_n = f(x_{n-k}), \quad \text{para } k \geq 2,$$

donde $f : \mathbb{X} \rightarrow \mathbb{X}$ es continua y $\mathbb{X} = \mathbb{R}, \mathbb{I}$ o \mathbb{S}^1 .

Dado el *vector inicial* de valores (x_1, x_2, \dots, x_k) , construimos la sucesión $(x_j)_{j=1}^{\infty}$ por

$$(x_1, x_2, \dots, x_k, f(x_1), f(x_2), \dots, f(x_k), f^2(x_1), f^2(x_2), \dots, f^2(x_k), \dots, f^j(x_1), f^j(x_2), \dots, f^j(x_k), \dots),$$

y decimos que esta sucesión es *periódica* de período p si $x_n = x_{n+p}$ para algún $p \geq 1$ y todo n . El menor $p \geq 1$ para el que lo anterior ocurre se denomina *período* de la sucesión. Entonces la sucesión es periódica de período p si, y solo si, la aplicación σ -*permutación* $F: \mathbb{R}^k \rightarrow \mathbb{R}^k$ dada por

$$F(z_1, z_2, \dots, z_k) = (z_2, z_3, \dots, z_k, f(z_1))$$

es periódica del mismo período. En este caso es $\sigma(1, 2, \dots, k) = (2, 3, \dots, k, 1)$, $f_{\sigma(1)} = \dots = f_{\sigma(k-1)} = \text{Identidad}$ y $f_{\sigma(k)} = f$.

Si ahora aplicamos el Teorema 13 obtenemos la estructura periódica asociada a la ecuación con retardo (para más detalles, ver [8]). Cuando $k = 3$ el diagrama de períodos que hemos escrito es también el que corresponde a tal ecuación.

Como comentario final, en todos los modelos presentados se pueden plantear otros problemas que pertenecen a la Dinámica Topológica, como por ejemplo, la estructura topológica de los conjuntos de puntos que verifican alguna propiedad global de retorno (puntos *casi periódicos*, *uniformemente recurrentes*, *recurrentes*, *conjuntos omega-límites*, *puntos "non-wandering"* y *puntos "chain-recurrent"*), existencia de atractores extraños en los sistemas, aparición o no de bifurcaciones locales (por ejemplo la citada *bifurcación de Hopf*) en modelos que dependen de parámetros. Pero el análisis de todos estos problemas excede de lo que nos hemos planteado en este artículo.

Referencias

- [1] Ll. Alsedà, J. Llibre y M. Misiurewicz, *Low Dimensional Combinatorial Dynamics*, Internat. J. Bifur. Chaos, **9**, 1687-1704 (1999).
- [2] Ll. Alsedà, J. Llibre y M. Misiurewicz, *Combinatorial Dynamics and Entropy in Dimension One*, Advanced Series in Nonlinear Dynamics, Vol **5**, (World Scientific, Singapore) (Second Edition, 2000).
- [3] Ll. Alsedà, J. Llibre, y M. Misiurewicz, *Periodic Orbits of Maps of Y*, Trans. Amer. Math. Soc., **313**, 475-583 (1989).
- [4] D.G. Aronson, M.A. Chory, G.R. Hall y R.P. McGehee, *Bifurcations from an Invariant Circle for Two-Parameter Families of Maps of the Plane: A Computer-Assisted Study*, Commun. Math.Phys. **83**, 303-354 (1982).
- [5] F. Balibrea y A. Linero, *On the Periodic Structure of the Antitriangular Maps on the Unit Square*, Annales Mathematicae Silesianae, **13**, 39-51 (1999).

- [6] F. Balibrea y A. Linero, *Periodic Structure of σ -Permutation Maps on I^n* , *Aequationes Mathematicae*, **62**, 265-279, (2001).
- [7] F. Balibrea y A. Linero, *Periodic Structure of σ -Permutation Maps II. The Case \mathbb{T}^n* , *Aequationes Mathematicae*, **64**, 1-19 (2002).
- [8] F. Balibrea y A. Linero, *On the Periodic Structure of Some Delayed Difference Equations*, aparecerá en *Journal of Difference Equations*.
- [9] F. Balibrea, L. Reich y J. Smital, *Iteration Theory: Dynamical Systems and Functional Equations*, aparecerá en *Internat. J. Bifur. Chaos*. Vol 13, No 7, July (2003).
- [10] B. Barna, *Über die Divergenzpunkte des Newtonschen Verfahrens zur Bestimmung von Wurzeln algebraischer Gleichungen I*, *Publ. Math. Debrecen* **3**, 109-118 (1953).
- [11] M. Benedicks y L. Carleson, *The Dynamics of the Hénon Map*, *Annals of Mathematics*, **133**, 73-169 (1991).
- [12] A.A. Cournot, *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, Librairie de L. Hachette, Paris (1838).
- [13] H.D. Crofton, *A Quantitative Approach to Parasitism*, *Parasitology*, **63**, 179-193 (1971).
- [14] M. Hénon, *A Two-dimensional Mapping with a Strange Attractor*, *Commun. Math. Phys.*, **50**, 67-77(1976).
- [15] M. Hénon e Y. Pomeau, *Two Strange Attractors with a Simple Structure*, *Lecture Notes in Mathematics*, **565**, Springer-Verlag 29 (1977).
- [16] R.A. Holmgren, *A First Course in Discrete Dynamical Systems*, Universitext, Springer-Verlag (1994).
- [17] G.E. Hutchinson, *Circular Causal Systems in Ecology*, *Annals N.Y. Acad. Sci.*, 221-246 (1949).
- [18] N. Kaldor, *A Model of the Trade Cycle*, *Economic Journal*, **50**, 78-92 (1940).
- [19] H.A. Lauwerier, *Two-dimensional Iterative Maps*, *Chaos* (ed. A.V.Holden), Princeton University Press, 58-95 (1986).
- [20] T.Y. Li y J.A. Yorke, *Period Three implies Chaos*, *Amer. Math. Monthly*, **82**, 985-992 (1975).
- [21] Z.R. Liu, H.M. Xie, Z.X. Zhu y Q.H. Lu, *The Strange Attractor of Lozi Mapping*, *Int. J. Bifur. Chaos*, **2**, 831 (1992).
- [22] J. Maynard Smith, *Mathematical Ideas in Biology*, Cambridge University Press (1968).

- [23] J. Maynard Smith, *Models in Ecology*, Cambridge University Press (1974)
- [24] M. Misiurewicz, *Strange Attractors for the Lozi Mappings*, Ann. N. Y. Acad. Sci., **375**, 348.
- [25] M. Misiurewicz, *Thirty Years after Sharkovsky's Theorem*, Internat. J. Bifur. Chaos, **5**, 1275-1281 (1995).
- [26] A.M. Ostrowskii, *Solutions of Equations and Systems of Equations*, Academic Press, second edition (1966).
- [27] J.R. Pounder y T.D. Rogers, *The Geometry of Chaos: Dynamics of a Nonlinear Second Order Difference Equation*, Bull. Math.Biol., **42**, 551-597 (1980).
- [28] T.D. Rogers y B.L. Clarke, *A Continuous Planar Map with Many Periodic Points*, Appl. Math. Computat., **8**, 17-33 (1981).
- [29] D. Ruelle, *Hasard et Chaos*, Odile Jacob (1991), versión española de Alianza Editorial (1993).
- [30] N.A. Sharkovsky, *Coexistence of Cycles of a Continuous Map of the Line into Itself*, Ukrain. Mat. Zh., **16**, 6-71 (1964, en ruso). Internat. J. Bifur. Chaos, **5** , 1263-1273 (1995, en inglés).
- [31] J. Smítal, *On Functions and Functional Equations*, Adam Hilger, Bristol and Philadelphia (1988).
- [32] Ph.D. Straffin, *Periodic Points of Continuous Functions*, Math. Mag. **51**, 99-105 (1978).
- [33] T. Tél, *Invariant Curves, Attractors, and Phase Diagram of a Piecewise Linear Map with Chaos*, J. Stat. Phys., **33**, 195 (1983).
- [34] W. Tucker, *The Lorenz Attractor exists*, C.R. Acad. Sci. Paris Sér. I Math., **328**, No 12, 1197-1202 (1999).
- [35] W. Tucker, *A rigorous ODE Solver and Smale's 14th Problem*, Found. Comput. Math., **2**, No 1, 53-117 (2002).

Sobre la regularidad, unicidad y justificación asintótica del problema de Ecuaciones Primitivas del Océano

F. GUILLÉN-GONZÁLEZ¹ M.A. RODRÍGUEZ-BELLIDO²

¹ Departamento de Ecuaciones Diferenciales y Análisis Numérico,
Universidad de Sevilla,

² Departamento de Matemática Aplicada I,
Universidad de Sevilla,

guillen@numer.us.es, angeles@us.es

Resumen

En este trabajo exponemos, y comentamos brevemente, los resultados obtenidos en la Tesis Doctoral de M. A. Rodríguez Bellido, dirigida por F. Guillén González, [15], sobre existencia, unicidad, regularidad y comportamiento asintótico en tiempo para las Ecuaciones Primitivas del Océano con condición de Dirichlet homogénea en el fondo, así como la obtención (a partir de las ecuaciones de Navier-Stokes), regularidad y unicidad de un nuevo modelo de Ecuaciones Primitivas con nuevas condiciones de contorno. Previamente se expone el marco físico y matemático en el que se encuadran dichas ecuaciones, así como los resultados que con anterioridad se conocían sobre este tema, prestando especial interés en la mejora que suponen los nuevos resultados de [15] sobre los ya existentes.

Palabras clave: *Ecuaciones Primitivas, existencia, unicidad, regularidad, comportamiento asintótico, condiciones de contorno de tipo Dirichlet, condiciones de contorno de tipo Navier.*

Clasificación por materias AMS: *35Q30, 35B40, 76D05*

1 Introducción

Si observamos nuestro planeta, veremos que al menos dos terceras partes están cubiertas por océanos, y todo él rodeado de la atmósfera. No es, por tanto, extraño que desde principios del siglo XIX algunos científicos, como Pierre

Fecha de recepción: 8 de Mayo de 2002

Simon de Laplace, pensasen que las leyes físicas que gobiernan la atmósfera y el océano podrían servir para predecir el tiempo y clima futuros. Ya en el siglo XX se empezó a considerar ese intento de predicción como la resolución de un problema de valores iniciales en física matemática.

Actualmente, según J. L. Lions, R. Temam y S. Wang [12], para entender el comportamiento turbulento de la atmósfera y el océano, y poder predecir el clima, necesitamos:

- (a) establecer las ecuaciones y modelos matemáticos que gobiernan el movimiento y estados de la atmósfera y el océano, y las interacciones que aparecen entre ellos;
- (b) establecer los fundamentos matemáticos de dichas ecuaciones y modelos;
- (c) establecer y resolver aproximaciones numéricas de dichas ecuaciones.

La atmósfera es un fluido compresible, descrito matemáticamente por las ecuaciones de la Hidrodinámica y la Termodinámica, en donde actúan también las fuerzas centrípetas y de Coriolis. Dichas ecuaciones describen los movimientos a gran escala, en donde las pequeñas escalas son consideradas como “ruidos” en los cálculos numéricos. Sin embargo, debido a que la escala vertical es mucho menor que la escala horizontal, podemos usar la aproximación hidrostática (que veremos a continuación) que permite obtener las Ecuaciones Primitivas de la atmósfera y del océano.

Una de las principales fuerzas que “mueven” el océano es la fuerza del viento, lo que pone de manifiesto que el comportamiento del océano depende de la atmósfera. Por otra parte, también es importante la influencia del océano en el comportamiento de la atmósfera y del clima. Todas estas observaciones justifican el estudio del mecanismo de acoplamiento entre la atmósfera y el océano.

En este trabajo nos centraremos en el tratamiento matemático de las Ecuaciones Primitivas del océano, tanto en su obtención mediante un argumento asintótico a partir de las ecuaciones de Navier-Stokes, como en los resultados de existencia, unicidad y regularidad de solución, comportamiento asintótico en tiempo, etc. Prestaremos especial atención a los resultados recogidos en la Tesis Doctoral de la autora, exponiendo los resultados de otros autores que con anterioridad a su defensa se conocían sobre este tema.

2 Obtención del modelo.

El océano es considerado como un fluido ligeramente compresible, con fuerzas de Coriolis y centrípetas. El conjunto de ecuaciones que rigen lo que se conoce como “large scale ocean model” son: **las ecuaciones de momentos, la ecuación de continuidad, la ecuación termodinámica o de la temperatura θ , la**

ecuación de difusión para la salinidad S y la ecuación de estado:

$$\left\{ \begin{array}{l} \rho \frac{d\mathbf{V}}{dt} + 2\rho \mathbf{W} \times \mathbf{V} + \rho \mathbf{W} \times (\mathbf{W} \times \mathbf{r}) + \nabla p + \rho \mathbf{g} = D \\ \frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{V} = 0 \\ \frac{d\theta}{dt} = Q_\theta \\ \frac{dS}{dt} = Q_S \\ \rho = f(\theta, S) \end{array} \right. \quad (1)$$

donde ρ es la densidad, \mathbf{V} la velocidad (3-dimensional) del flujo, p la presión, $\mathbf{g} = (0, 0, g)$ es la gravedad, $2\rho \mathbf{W} \times \mathbf{V}$ es el término de Coriolis y $\rho \mathbf{W} \times (\mathbf{W} \times \mathbf{r})$ las fuerzas centrípetas ($\mathbf{W} = f(0, \cos \lambda, \sin \lambda)$ es el vector de rotación de la Tierra, \mathbf{r} es el radio de la Tierra y $\lambda = \lambda(y)$ es la latitud), D es la disipación (molecular), Q_θ y Q_S las difusiones de temperatura y salinidad respectivamente.

Respecto a los operadores diferenciales utilizados, $\nabla = (\partial_x, \partial_y, \partial_z)$ es el gradiente tridimensional, con $\nabla \cdot$ el operador divergencia y $\frac{d}{dt}$ es la “derivada material o total”, es decir,

$$\frac{d}{dt} = \partial_t + \mathbf{V} \cdot \nabla$$

Si hacemos la aproximación llamada del β -plano, el dominio ocupado por el océano Ω , se describe en coordenadas cartesianas como:

$$\Omega = \{(x, y, z) = (\mathbf{x}, z) \in \mathbb{R}^3, \mathbf{x} \in S, -H(\mathbf{x}) < z < 0\}.$$

Su frontera $\partial\Omega = \overline{\Gamma}_b \cup \Gamma_l \cup \Gamma_s$ donde el fondo Γ_b , las paredes laterales Γ_l y la superficie Γ_s están definidas por:

$$\Gamma_b = \{(\mathbf{x}, z) \in \mathbb{R}^3 : \mathbf{x} \in S, z = -H(\mathbf{x})\},$$

$$\Gamma_l = \{(\mathbf{x}, z) \in \mathbb{R}^3 : \mathbf{x} \in \partial S, -H(\mathbf{x}) < z < 0\},$$

$$\Gamma_s = \{(\mathbf{x}, 0) : \mathbf{x} \in S\},$$

donde S (sección horizontal) es un abierto acotado en \mathbb{R}^2 y H (la profundidad) es una función continua no negativa sobre S . Las Figuras 1 y 2 ilustran las posibles configuraciones del dominio bidimensional (o las secciones verticales de un dominio tridimensional):

Las dificultades teóricas y computacionales que presenta el tratamiento del sistema (1) ha llevado a la consideración de dos simplificaciones importantes:

- a) **la aproximación de Boussinesq**, que consiste en despreciar las diferencias de densidad en todo el sistema salvo en el término de gravedad y la ecuación de estado. De ese modo, fijada una densidad media constante

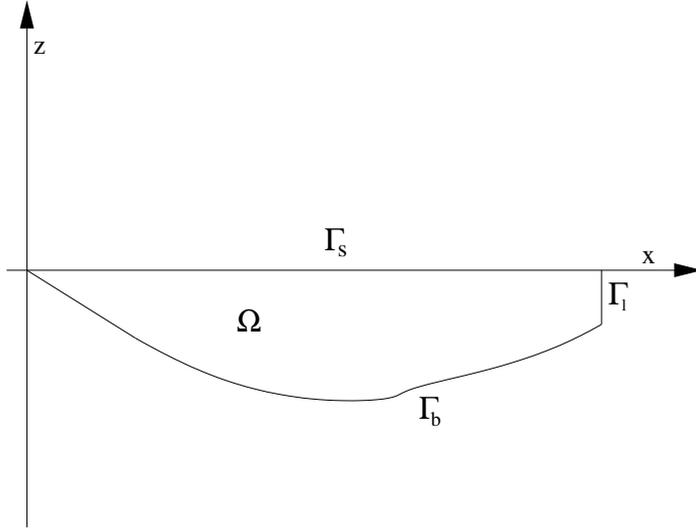


Figura 1: El dominio (2D) con paredes laterales.

ρ_0 , $\rho = \rho_0 + \rho'$ con $\rho' \ll \rho_0$. Dicha hipótesis hace que la ecuación de continuidad se transforme en la ecuación de incompresibilidad para la velocidad \mathbf{V} , que es una hipótesis físicamente válida debido a la lenta velocidad del agua respecto a la viscosidad en dicho medio. La inclusión de las fuerzas centrípetas en el gradiente de una función potencial P (junto con la presión), permiten reescribir (1) como:

$$(BEs) \left\{ \begin{array}{l} \frac{d}{dt} \mathbf{v} - \mu \Delta_{\mathbf{x}} \mathbf{v} - \nu \partial_{zz}^2 \mathbf{v} + 2f(\sin(\lambda) \mathbf{v}^\perp + \cos(\lambda) w \mathbf{e}_1) + \nabla_{\mathbf{x}} P = \mathbf{0} \\ \frac{d}{dt} w - \mu \Delta_{\mathbf{x}} w - \nu \partial_{zz}^2 w - 2f \cos(\lambda) v_1 + \partial_z P + \frac{\rho'}{\rho_0} g = 0 \\ \nabla_{\mathbf{x}} \cdot \mathbf{v} + \partial_z w = 0 \\ \partial_t \theta + \mathbf{v} \cdot \nabla_{\mathbf{x}} \theta + w \partial_z \theta - \mu_\theta \Delta_{\mathbf{x}} \theta - \nu_\theta \partial_{zz}^2 \theta = 0 \\ \partial_t S + \mathbf{v} \cdot \nabla_{\mathbf{x}} S + w \partial_z S - \mu_S \Delta_{\mathbf{x}} S - \nu_S \partial_{zz}^2 S = 0 \\ \rho = f(\theta, S), \end{array} \right.$$

donde $\mathbf{v} = (v_1, v_2)$ y w son las velocidades horizontales y vertical respectivamente, $\mathbf{v}^\perp = (-v_2, v_1)$, y μ , ν , μ_θ , ν_θ , μ_S , ν_S los coeficientes de difusión (turbulenta) anisótropos horizontal y vertical de (\mathbf{v}, w) , θ y S respectivamente. $\nabla_{\mathbf{x}}$, $\Delta_{\mathbf{x}}$ y $\nabla_{\mathbf{x}} \cdot$ son los operadores gradiente, laplaciano y divergencia en las variables horizontales. Dicho sistema recibe el nombre de Ecuaciones de Boussinesq del océano.

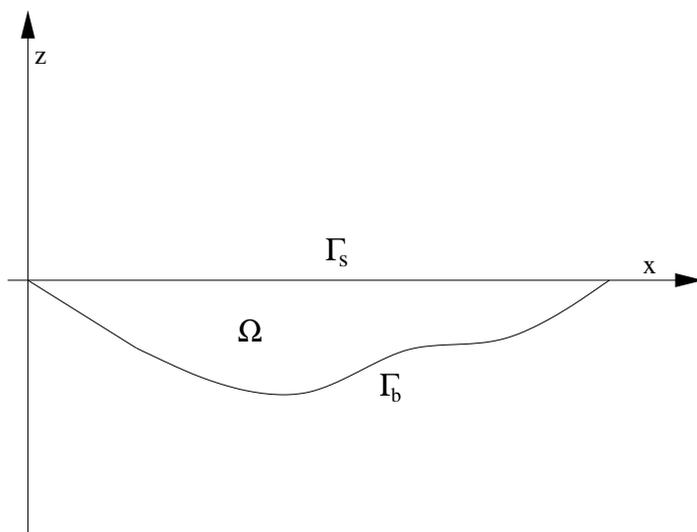


Figura 2: El dominio (2D) sin talud.

b) **la aproximación hidrostática.** Si queremos hacer un proceso de adimensionalización del sistema (*BES*) respecto a los valores de referencia de cada magnitud física, el análisis real de escalas nos dice que el cociente entre las longitudes vertical y horizontal es pequeño:

$$\delta = \frac{Z}{L} \approx 10^{-3},$$

siendo L el valor de referencia de las dimensiones horizontales, y Z de las verticales. Además, la velocidad vertical del agua es mucho menor que la velocidad horizontal, lo que se modela aproximando la tercera ecuación de momentos por la llamada **ecuación hidrostática**:

$$\frac{\partial p}{\partial z} = -\rho g,$$

que relaciona la presión y la densidad del océano con la gravedad, y que se ha convertido en una ecuación fundamental en Oceanografía. El análisis de escalas nos dice además que para que las viscosidades de las dos primeras componentes de la ecuación de momentos sean del mismo orden (respecto a δ), necesitamos suponer que:

$$\nu = \delta^2 \nu_v, \quad \mu = \nu_h, \quad \text{con } \nu_v = O(1) \text{ y } \nu_h = O(1). \quad (2)$$

Por simplicidad, sólo prestaremos atención al sistema de ecuaciones (no lineal) para la velocidad y presión, ya que el sistema acoplado con la temperatura y salinidad introduce ecuaciones de convección-difusión

(lineales si la velocidad es conocida) que no añaden dificultades esenciales (de tipo matemático). En consecuencia, las Ecuaciones Primitivas del océano vienen descritas por el siguiente sistema:

$$(EP) \begin{cases} \partial_t \mathbf{v} + \mathbf{v} \cdot \nabla_{\mathbf{x}} \mathbf{v} + w \partial_z \mathbf{v} - \nu_h \Delta_{\mathbf{x}} \mathbf{v} - \nu_v \partial_{zz}^2 \mathbf{v} + \alpha \mathbf{v}^\perp + \frac{1}{\rho_0} \nabla_{\mathbf{x}} p = \mathbf{0} \\ \partial_z p = -\rho g, \quad \nabla_{\mathbf{x}} \cdot \mathbf{v} + \partial_z w = 0, \end{cases}$$

donde $\alpha = 2f \operatorname{sen}(\lambda)$. La frontera superficie Γ_s no varía, siendo los nuevos Ω , Γ_l y Γ_b (donde $h = \frac{H}{Z}$):

$$\begin{aligned} \Omega &= \{(\mathbf{x}, z) \in \mathbb{R}^3, \mathbf{x} \in S, -h(\mathbf{x}) < z < 0\}, \\ \Gamma_l &= \{(\mathbf{x}, z) \in \mathbb{R}^3 : \mathbf{x} \in \partial S, -h(\mathbf{x}) < z < 0\}, \\ \Gamma_b &= \{(\mathbf{x}, z) \in \mathbb{R}^3 : \mathbf{x} \in S, z = -h(\mathbf{x})\}. \end{aligned} \quad (3)$$

Las Ecuaciones Primitivas del océano (*EP*) se obtienen a partir de la hipótesis de aproximación hidrostática en los trabajos de J. L. Lions, R. Temam y S. Wang, [10, 11]. Dicha hipótesis se puede justificar como límite de la formulación variacional de las ecuaciones de Navier-Stokes o (*BES*) cuando $\delta \rightarrow 0$ imponiendo (2), ver los resultados de O. Besson y M. R. Laydi, [2], para el caso estacionario y los de P. Azérad y F. Guillén-González, [1], para el caso de evolución.

2.1 Elección de condiciones de contorno.

La “fuente” de movimiento del océano es la atmósfera. No podemos olvidar pues la interacción atmósfera-océano a la hora de determinar cuáles son las condiciones de interfase, que se traducirán en condiciones de contorno sobre la superficie del océano. Una hipótesis simplificadora es considerar que dicha interfase está fija, hipótesis que se denomina de “techo rígido” (*rigid lid hypothesis*) y que está basada en dos hechos:

- (a) el agua es mucho más densa que el aire. De hecho, se observa que $\rho^a / \rho^s \approx 10^{-3}$, siendo ρ^a la densidad del aire y ρ^s la del agua oceánica, con lo que la interfase entre el aire y el agua es muy estable considerando grandes escalas espaciales, debido a la intensidad de la fuerza gravitacional.
- (b) a gran escala, el desplazamiento vertical de las ondas de mareas se puede despreciar, ya que aparecen como “ruidos en las altas frecuencias” no suelen ser tenidas en cuenta en la mayoría de modelos de Circulación Global.

Así pues, las condiciones de contorno que aparecen en la superficie son (denotando con el superíndice s las variables correspondientes al océano y con a las variables correspondientes a la atmósfera):

$$w|_{\Gamma_s} = 0, \quad \mathbf{v}^s|_{\Gamma_s} = \mathbf{v}^a|_{\Gamma_s}.$$

Sin embargo, debido a la diferencia de densidad entre ambos medios, aparece una capa límite fina en la atmósfera (de 1 km. de espesor) y muy fina en el océano (entre 10 y 100 m. de espesor) que no son tratables numéricamente, por el momento. Usando entonces unos fundamentos más prácticos, se considera que los esfuerzos de cizalla en la interfase son debidos a la fuerza horizontal del viento en superficie, lo que se traduce matemáticamente como:

$$-\rho_0^s \nu_v^s \partial_z \mathbf{v}^s = \rho^a C_D^a (\mathbf{v}^a - \mathbf{v}^s) |\mathbf{v}^a - \mathbf{v}^s|^\alpha \quad \text{sobre } \Gamma_s,$$

donde C_D^a es el coeficiente de transferencia de momento. Aquí vamos a considerar la siguiente simplificación:

$$w = 0, \quad \nu_v \partial_z \mathbf{v}^s = \tau \quad \text{sobre } \Gamma_s,$$

donde τ es la tensión del viento sobre la superficie del océano, que se toma como dato. Con respecto al fondo, imponemos condiciones de adherencia, permitiendo deslizamiento vertical en las paredes laterales, que vienen dadas por:

$$w = 0 \quad \text{sobre } \Gamma_b, \quad \mathbf{v}^s = \mathbf{0} \quad \text{sobre } \Gamma_b \cup \Gamma_l. \quad (4)$$

Fijaremos estas condiciones de contorno en los resultados que expondremos en la siguiente sección.

2.2 El modelo reducido.

Las variables del sistema de Ecuaciones Primitivas \mathbf{v} , w y p no son del mismo tipo: la velocidad horizontal \mathbf{v} verifica un problema de evolución y necesita por tanto datos iniciales para quedar determinado (*variable pronóstico*). La velocidad vertical w no verifica un problema de evolución pero se puede determinar a partir de la variable pronóstico \mathbf{v} (*variable diagnóstica*). Concretamente, integrando la ecuación de incompresibilidad en $(z, 0)$, como $w(t; \mathbf{x}, 0) = 0$, obtenemos:

$$w(t; \mathbf{x}, z) = \int_z^0 \nabla_{\mathbf{x}} \cdot \mathbf{v}(t; \mathbf{x}, s) ds. \quad (5)$$

Para la presión, integrando en $(z, 0)$ la ecuación hidrostática, obtenemos:

$$p(t; \mathbf{x}, z) = p_s(t; \mathbf{x}) + \int_z^0 (\rho g)(t; \mathbf{x}, s) ds = p_s(t; \mathbf{x}) + g \int_z^0 f(\theta, S)(t; \mathbf{x}, s) ds,$$

donde $p_s(t; \mathbf{x}) = p(t; \mathbf{x}, 0)$ es la presión en la superficie del océano. Se reescribe entonces el gradiente horizontal de presión como:

$$\frac{1}{\rho_0} \nabla_{\mathbf{x}} p = \frac{1}{\rho_0} \nabla_{\mathbf{x}} p_s + \mathbf{F}(\theta, S).$$

Imponiendo la condición de contorno (4) sobre w en la expresión de w (5), se llega a la restricción $\nabla_{\mathbf{x}} \cdot \langle \mathbf{v} \rangle = 0$ en $(0, T) \times S$, donde $\langle \mathbf{v} \rangle(t; \mathbf{x}) =$

Definición 1 (Solución débil) Sean $\mathbf{u}_0 \in H$, $\mathbf{F} \in L^2(0, T; H_{b,l}^{-1}(\Omega)^2)$ y $\tau \in L^2(0, T; H^{-1/2}(\Gamma_s)^2)$. Decimos que $\mathbf{u} : (0, T) \times \Omega \rightarrow \mathbb{R}^2$ es una **solución débil de (EP) en $(0, T)$** si

$$\mathbf{u} \in L^\infty(0, T; H) \cap L^2(0, T; V),$$

verifica la formulación variacional: $\forall \varphi \in C^1([0, T]; \mathcal{V})$ tal que $\varphi(T) = \mathbf{0}$,

$$\begin{aligned} & \int_0^T \int_\Omega (-\mathbf{u} \cdot (\partial_t \varphi + (\mathbf{u}_H \cdot \nabla_{\mathbf{x}}) \varphi + u_3 \partial_z \varphi) + \alpha \mathbf{u}^\perp \cdot \varphi) d\Omega dt \\ & + \int_0^T \int_\Omega (\nu_h \nabla_{\mathbf{x}} \mathbf{u} : \nabla_{\mathbf{x}} \varphi + \nu_v \partial_z \mathbf{u} \cdot \partial_z \varphi) d\Omega dt \\ & = \int_\Omega \mathbf{u}_0 \cdot \varphi(0) d\Omega + \int_0^T \langle \mathbf{F}, \varphi \rangle_\Omega dt + \int_0^T \langle \tau, \varphi \rangle_{\Gamma_s} dt, \end{aligned}$$

y, además, \mathbf{u} satisface la desigualdad de energía:

$$\begin{aligned} & \frac{1}{2} \|\mathbf{u}\|_{L^2(\Omega)}^2 + \int_0^t \left(\nu_h \|\nabla_{\mathbf{x}} \mathbf{u}\|_{L^2(\Omega)}^2 + \nu_v \|\partial_z \mathbf{u}\|_{L^2(\Omega)}^2 \right) ds \\ & \leq \frac{1}{2} \|\mathbf{u}_0\|_{L^2(\Omega)}^2 + \int_0^t \langle \mathbf{F}, \mathbf{u} \rangle_\Omega ds + \int_0^t \langle \tau, \mathbf{u} \rangle_{\Gamma_s} ds \quad \text{c.p.d. } t \in (0, T). \end{aligned} \tag{6}$$

En el caso $T = +\infty$, decimos que \mathbf{u} es una solución débil de (EP) en $(0, +\infty)$ si \mathbf{u} es una solución débil de (EP) en $(0, T)$, $\forall T < +\infty$.

Aquí, $\langle \cdot, \cdot \rangle_\Omega$ denota la dualidad entre $H_{b,l}^{-1}(\Omega)$ y $H_{b,l}^1(\Omega)$, mientras que $\langle \cdot, \cdot \rangle_{\Gamma_s}$ denota la dualidad entre $H^{-1/2}(\Gamma_s)$ y $H^{1/2}(\Gamma_s)$. Denotamos en esta sección por u_3 la velocidad vertical asociada a \mathbf{u} .

Finalmente, denotamos la norma en V por $\|\varphi\|_V^2 = \nu_h \|\nabla_{\mathbf{x}} \varphi\|_{L^2(\Omega)}^2 + \nu_v \|\partial_z \varphi\|_{L^2(\Omega)}^2$, y la norma en $H_{b,l}^1(\Omega)$ por $\|\varphi\|_{H^1(\Omega)}^2 = \|\nabla_{\mathbf{x}} \varphi\|_{L^2(\Omega)^N}^2 + \|\partial_z \varphi\|_{L^2(\Omega)}^2$.

Cuando aumentamos básicamente en un orden la regularidad de la solución, llegamos a la siguiente:

Definición 2 (Solución fuerte) Sean $\mathbf{u}_0 \in V$, $\mathbf{F} \in L^2(0, T; L^2(\Omega)^2)$, $\tau \in L^2(0, T; H^{1/2}(\Gamma_s)^2)$ y $\partial_t \tau \in L^2(0, T; H^{-1/2}(\Gamma_s)^2)$. Si \mathbf{u} una solución débil de (EP) en $(0, T)$, decimos que \mathbf{u} es una **solución fuerte** si verifica la siguiente regularidad adicional:

$$\mathbf{u} \in L^\infty(0, T; V) \cap L^2(0, T; H^2(\Omega)^2 \cap V), \quad \partial_t \mathbf{u} \in L^2(0, T; H).$$

La **existencia de solución débil de (EP)** es bien conocida, ver Lewandowski [9] y Lions-Teman-Wang [11], en dominios con profundidad acotada inferiormente (que también denominaremos **talud**, es decir, $h \geq h_{min} > 0$ en \bar{S}). En dichos trabajos, se usa un método de Galerkin para obtener la

velocidad \mathbf{u} en un espacio con la restricción $\nabla \cdot \langle \mathbf{u} \rangle = 0$. Más tarde, se recupera la presión a través de un lema de De Rham específico para los espacios de funciones definidas sobre la superficie. En dominios sin talud, la existencia de solución débil se obtiene como consecuencia de un proceso de paso al límite aplicado a las ecuaciones de Navier-Stokes con viscosidad anisótropa cuando el cociente entre la profundidad y el diámetro horizontal del dominio tiende a cero, ver Besson-Laydi [2] para el caso estacionario y Azerad-Guillén [1] para el caso de evolución. También pueden encontrarse otras demostraciones por argumentos de aproximación interna, ver [5] para el caso estacionario y [6] para el caso evolutivo.

Hasta donde sabemos, no había resultados de existencia de solución fuerte del problema (EP), excepto en el caso lineal estacionario [17]. Uno de los principales problemas para su estudio es el tratamiento de las condiciones de contorno; sobre la superficie tenemos una condición de Neumann no homogénea, mientras que en el fondo y paredes laterales tenemos condición de Dirichlet homogénea. Por otra parte, la unicidad de solución del problema (EP) es también un problema abierto, incluso suponiendo la existencia de soluciones fuertes.

3.1 Regularidad fuerte de (EP)

Como hemos dicho, los resultados de regularidad fuerte existentes para este tipo de problemas se conocían para el caso del sistema lineal estacionario (S_{st}), cuyo problema evolutivo correspondiente es:

$$(S) \left\{ \begin{array}{l} \partial_t \mathbf{v} - \nu_h \Delta_{\mathbf{x}} \mathbf{v} - \nu_v \partial_{zz}^2 \mathbf{v} + \nabla_{\mathbf{x}} q_s = \mathbf{g} \quad \text{en } (0, T) \times \Omega, \\ \nabla_{\mathbf{x}} \cdot \langle \mathbf{v} \rangle = 0 \quad \text{en } (0, T) \times S, \\ \mathbf{v}|_{t=0} = \mathbf{v}_0 \quad \text{en } \Omega, \\ \nu_v \partial_z \mathbf{v}|_{\Gamma_s} = \tau, \quad \mathbf{v}|_{\Gamma_b \cup \Gamma_l} = \mathbf{0} \quad \text{en } (0, T). \end{array} \right.$$

Nota 1 *Notemos que el término de Coriolis se ha omitido en los cálculos, ya que para la demostración de los resultados que presentaremos no supone ninguna dificultad adicional.*

Teorema 1 (Solución débil de (S_{st})) *Sea $S \subseteq \mathbb{R}^d$ ($d = 1$ o 2) y sea $\Omega \subseteq \mathbb{R}^{d+1}$, definido como (3), un dominio Lipschitz-continuo. Si $\mathbf{g} \in H_{b,l}^{-1}(\Omega)^d$ y $\tau \in H^{-1/2}(\Gamma_s)^d$, entonces el problema (S_{st}) tiene una única solución $\mathbf{v} \in H^1(\Omega)^d$. Además, existe una constante $C = C(\Omega) > 0$ tal que si $\nu = \min\{\nu_h, \nu_v\}$, obtenemos:*

$$\|\mathbf{v}\|_{H^1(\Omega)}^2 \leq \frac{C}{\nu^2} \left\{ \|\tau\|_{H^{-1/2}(\Gamma_s)}^2 + \|\mathbf{g}\|_{H_{b,l}^{-1}(\Omega)}^2 \right\}. \quad (7)$$

En [2], [5] y [9], hay diferentes demostraciones de este resultado.

Teorema 2 (Solución fuerte de (S_{st})) ([17]) Sea $S \subseteq \mathbb{R}^d$ ($d = 1$ o 2) un dominio de clase C^3 y $h \in C^3(\bar{S})$ con $h \geq h_{\min} > 0$ en \bar{S} . Si $\mathbf{g} \in L^2(\Omega)^d$ y $\tau \in H_0^{1/2+\varepsilon}(\Gamma_s)^d$ (para algún $\varepsilon > 0$), entonces existe una (única) solución fuerte \mathbf{v} de (S_{st}) (es decir, $\mathbf{v} \in H^2(\Omega)^d \cap V$). Además, existe una constante $C = C(\Omega) > 0$ tal que:

$$\|\mathbf{v}\|_{H^2(\Omega)}^2 \leq \frac{C}{\nu^2} \left\{ \|\mathbf{g}\|_{L^2(\Omega)}^2 + \|\tau\|_{H_0^{1/2+\varepsilon}(\Gamma_s)}^2 \right\}. \quad (8)$$

Aprovechando este resultado ([7]) (ver también [15]), extendemos el resultado de regularidad fuerte al caso del problema evolutivo lineal (S) . Para ello, definimos el operador $B : \mathbf{a} \in H^{-1/2}(\Gamma_s)^d \rightarrow \mathbf{u} = B\mathbf{a} \in V$, donde \mathbf{u} es la solución débil del problema de Stokes hidrostático (S_{st}) con $\mathbf{g} = \mathbf{0}$ y $\tau = \mathbf{a}$. Tomamos $\mathbf{e}(t) = B(\tau(t))$ que posee regularidad fuerte, y demostramos que $\partial_t \mathbf{e}(t)$ es igual a $B(\partial_t \tau(t))$ que posee regularidad débil, de donde deducimos que $\mathbf{e} \in C^0([0, T]; V)$. Estudiamos entonces el problema con condiciones de contorno homogéneas verificado por $\mathbf{y} = \mathbf{v} - \mathbf{e}$. Las estimaciones de energía para \mathbf{e} y $\partial_t \mathbf{e}$ que se deducen de los Teoremas 1 y 2 permitirán concluir el siguiente resultado:

Teorema 3 (Solución fuerte de (S)) Sea $S \subseteq \mathbb{R}^d$ ($d = 1$ o 2) un dominio C^3 y $h \in C^3(\bar{S})$ con $h \geq h_{\min} > 0$ en \bar{S} . Si $\mathbf{g} \in L^2((0, T) \times \Omega)^d$, $\mathbf{v}_0 \in V$, $\tau \in L^2(0, T; H_0^{1/2+\varepsilon}(\Gamma_s)^d)$, para algún $\varepsilon > 0$, con $\partial_t \tau \in L^2(0, T; H^{-1/2}(\Gamma_s)^d)$, entonces existe una única solución fuerte \mathbf{v} de (S) en $(0, T)$. Además, existe una constante $C > 0$ tal que:

$$\begin{aligned} \|\mathbf{v}\|_{L^\infty(V)}^2 + \|\mathbf{v}\|_{L^2(H^2(\Omega))}^2 + \|\partial_t \mathbf{v}\|_{L^2(H)}^2 &\leq C \left\{ \|\mathbf{v}_0\|_V^2 + \|\tau(0)\|_{H^{-1/2}(\Gamma_s)}^2 \right. \\ &\quad \left. + \|\mathbf{g}\|_{L^2(L^2(\Omega))}^2 + \|\tau\|_{L^2(H_0^{1/2+\varepsilon}(\Gamma_s))}^2 + \|\partial_t \tau\|_{L^2(H^{-1/2}(\Gamma_s))}^2 \right\} \end{aligned} \quad (9)$$

Pasamos entonces a obtener resultados de regularidad fuerte para el problema no lineal (EP) . Usamos el Teorema anterior para levantar las condiciones de contorno, y así estudiar el problema homogéneo que verifica (\mathbf{w}, π_s) , donde $\mathbf{w} = \mathbf{v} - \mathbf{e}$, $\pi_s = p_s - q_s$, para (\mathbf{e}, q_s) la solución de un problema de tipo (S) con \mathbf{F} y τ como datos y \mathbf{v}_0 como dato inicial:

$$(NL) \begin{cases} \partial_t \mathbf{w} - \nu_h \Delta_{\mathbf{x}} \mathbf{w} - \nu_v \partial_{zz}^2 \mathbf{w} + (\mathbf{w} + \mathbf{e}) \partial_x (\mathbf{w} + \mathbf{e}) \\ \quad + (w_3 + e_3) \partial_z (\mathbf{w} + \mathbf{e}) + \partial_x \pi_s = \mathbf{0} \text{ en } (0, T) \times \Omega, \\ \quad \nabla_{\mathbf{x}} \cdot \langle \mathbf{w} \rangle = 0 \text{ en } (0, T) \times S, \quad \mathbf{w}|_{t=0} = \mathbf{0} \text{ en } \Omega, \\ \quad \nu_v \partial_z \mathbf{w} = 0 \text{ sobre } (0, T) \times \Gamma_s, \quad \mathbf{w} = 0 \text{ sobre } (0, T) \times (\Gamma_b \cup \Gamma_l), \end{cases}$$

con $w_3 = \int_z^0 \nabla_{\mathbf{x}} \cdot \mathbf{w} ds$ y de forma similar para e_3 .

Aproximamos \mathbf{w} por \mathbf{w}_m las aproximaciones de Galerkin en los espacios m -dimensionales V_m formados por una base ortonormal (en V) de autovalores del operador hidrostático $A : V \rightarrow V'$ tal que:

$$\langle A\mathbf{u}, \mathbf{v} \rangle_{V', V} = \int_{\Omega} (\nu_h \nabla_{\mathbf{x}} \mathbf{u} : \nabla_{\mathbf{x}} \mathbf{v} + \nu_v \partial_z \mathbf{u} \cdot \partial_z \mathbf{v}) d\Omega \quad \forall \mathbf{u}, \mathbf{v} \in V, \quad (10)$$

asociado a condiciones de contorno homogéneas (Neumann sobre la superficie y Dirichlet sobre el fondo y las paredes laterales). Para obtener estimaciones en norma $H^2(\Omega)$, tomamos $A\mathbf{w}_m(t) \in V_m$ como funciones test, obteniendo:

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{w}_m\|_V^2 + \|A\mathbf{w}_m\|_{L^2(\Omega)}^2 = G(\mathbf{w}_m, \mathbf{e}, \mathbf{f}), \quad (11)$$

para una cierta función G . Intentamos acotar los términos de G con las estimaciones en norma fuerte de \mathbf{e} , las cotas de \mathbf{f} y controlando los términos en \mathbf{w}_m con las normas que aparecen a la izquierda de (11). La dificultad mayor la presentan los términos:

$$I_1 = - \int_{\Omega} (\mathbf{w}_m \cdot \nabla_{\mathbf{x}}) \mathbf{w}_m \cdot A\mathbf{w}_m d\Omega \quad I_2 = - \int_{\Omega} (w_m)_3 \partial_z \mathbf{w}_m \cdot A\mathbf{w}_m d\Omega$$

correspondientes al término no lineal de (EP). Observar que I_2 es menor regular que I_1 debido a la anisotropía en la regularidad de la velocidad vertical. Para la acotación de I_2 será fundamental el siguiente lema ([7]):

Lema 4 *Sea $\Omega \subseteq \mathbb{R}^N$ ($N = 2$ o 3) el dominio considerado anteriormente. Entonces, para toda función $\mathbf{v} \in W^{1,p}(\Omega)^{N-1}$ ($p > 1$), si se define v_3 como $v_3(\mathbf{x}, z) = - \int_{-h(\mathbf{x})}^z \nabla_{\mathbf{x}} \cdot \mathbf{v}(\mathbf{x}, s) ds$, tenemos:*

$$\|v_3\|_{L^p(\Omega)} \leq h_{max} \|\nabla_{\mathbf{x}} \cdot \mathbf{v}\|_{L^p(\Omega)}$$

En el caso 2D, usando además la desigualdad de Gagliardo-Nirenberg:

$$\begin{aligned} I_2 &\leq \|(w_m)_3\|_{L^4(\Omega)} \|\partial_z \mathbf{w}_m\|_{L^4(\Omega)} \|A\mathbf{w}_m\|_{L^2(\Omega)} \\ &\leq Ch_{\max} \|\partial_x \mathbf{w}_m\|_{L^4(\Omega)} \|A\mathbf{w}_m\|_{L^2(\Omega)}^{3/2} \|\partial_z \mathbf{w}_m\|_{L^2(\Omega)}^{1/2} \\ &\leq Ch_{\max} \|\mathbf{w}_m\|_{H^1(\Omega)} \|A\mathbf{w}_m\|_{L^2(\Omega)}^2 \\ &\leq \frac{C}{\nu^{1/2}} h_{\max} \|\mathbf{w}_m\|_V \|A\mathbf{w}_m\|_{L^2(\Omega)}^2 \end{aligned}$$

Acotando de forma similar el resto de los términos llegamos a la desigualdad:

$$\begin{aligned} \frac{d}{dt} \|w_m\|_V^2 + \|Aw_m\|_{L^2(\Omega)}^2 &\left(1 - C_1 h_{max} \|w_m\|_V\right) \\ &\leq C_2 \|w_m\|_V^4 + a(t) \|w_m\|_V^2 + b(t), \end{aligned} \quad (12)$$

donde $a = a(\tau, F)$, $b = b(\tau, F)$ son determinadas funciones de $L^1(0, T)$, lo que, bajo hipótesis de pequeñez sobre los datos, nos permite aplicar el Lema de Gronwall y obtener el siguiente resultado ([7]):

Teorema 5 (Solución global fuerte para datos pequeños en el caso 2D)
Sea $S \subseteq \mathbb{R}$ un intervalo y $h \in C^3(\bar{S})$ tal que $h \geq h_{min} > 0$ en \bar{S} . Supongamos que

$u_0 \in V$, $F \in L^2(0, T; L^2(\Omega))$ y $\tau \in L^2(0, T; H_0^{1/2+\varepsilon}(\Gamma_s))$, para algún $\varepsilon > 0$, con $\partial_t \tau \in L^2(0, T; H^{-1/2}(\Gamma_s))$. Si se verifica la siguiente **condición de pequeñez**: $\forall t \in [0, T]$,

$$(H)_{2D} \begin{cases} \exp\left(-\frac{1}{4K_2}t + \int_0^t a(s)ds\right) \left\{ 2\left(\|u_0\|_V^2 + K_1\|\tau(0)\|_{H^{-1/2}(\Gamma_s)}^2\right) \right. \\ \left. + \int_0^t \exp\left(\frac{1}{4K_2}s - \int_0^s a(\sigma)d\sigma\right) b(s)ds \right\} < M^2, \end{cases}$$

donde M es una constante positiva suficientemente pequeña, K_1 y K_2 son constantes, y a y b las funciones que aparecen en (12), entonces existe una única solución fuerte (u, p_s) de (EP) en $(0, T)$ (p_s es única salvo una función aditiva de t).

Además, en [7], se demuestra comportamiento asintótico en tiempo exponencialmente decreciente en norma $H^1(\Omega)$ cuando $t \uparrow +\infty$, si se impone $(H)_{2D} \forall t \in (0, +\infty)$ y una condición extra de pequeñez sobre los datos τ y F cuando $t \uparrow \infty$. Por último, mediante un argumento de punto fijo, se llega a que existe solución fuerte local en tiempo si h_{\max} es suficientemente pequeño.

En el caso 3D, aplicando desigualdades de interpolación, se obtiene ([7]):

$$I_2 \leq C \frac{h_{\min}}{\nu^{1/4}} \|A\mathbf{w}_m\|_{L^2(\Omega)}^{5/2} \|\mathbf{w}_m\|_V^{1/2},$$

lo que no permite aplicar el argumento del caso bidimensional. En la búsqueda de un remedio a dicho problema, en [8] nos centramos en la anisotropía de la velocidad vertical. Como dijimos, $\partial_z w_3 = -\nabla_{\mathbf{x}} \cdot \mathbf{w} \in L^2(\Omega)$, luego por una desigualdad de tipo Poincaré vertical $w_3 \in L^2(\Omega)$. Sin embargo, $\nabla_{\mathbf{x}} w_3 \notin L^2(\Omega)$ en general. Esto nos lleva a separar la regularidad en las variables \mathbf{x} y z , ajustando las estimaciones en cada dirección. La novedad está en considerar espacios y estimaciones anisótropas del tipo (ver [8] para las demostraciones):

Definición 3 *Dados $p, q \in [1, +\infty]$, diremos que una función \mathbf{u} pertenece a $L_z^q L_{\mathbf{x}}^p(\Omega)$ si:*

$$\mathbf{u}(\cdot, z) \in L^q(S_z) \quad \text{y} \quad \|\mathbf{u}(\cdot, z)\|_{L^q(S_z)} \in L^p(-h_{\max}, 0),$$

y su norma viene dada por la expresión:

$$\|\|\mathbf{u}(\cdot, z)\|_{L^q(S_z)}\|_{L^p(-h_{\max}, 0)}$$

Nota 2 *Las normas que usaremos más frecuentemente en el caso 3D serán:*

$$\begin{aligned} \|\mathbf{u}\|_{L_z^2 L_{\mathbf{x}}^4(\Omega)} &= \left(\int_{-h_{\max}}^0 \|\mathbf{u}(\cdot, z)\|_{L^4(S_z)}^2 dz \right)^{1/2} \\ \|\mathbf{u}\|_{L_z^\infty L_{\mathbf{x}}^4(\Omega)} &= \sup_{z \in (-h_{\max}, 0)} \|\mathbf{u}(\cdot, z)\|_{L^4(S_z)}, \end{aligned}$$

Lema 6 (Desigualdades de interpolación)

(a) Sea $v \in L^2(\Omega)$ una función tal que $\partial_z v \in L^2(\Omega)$ y $(vn_z)|_{\Gamma_b} = 0$. Entonces, $v \in L_z^\infty L_x^2(\Omega)$ y satisface la estimación:

$$\|v\|_{L_z^\infty L_x^2}^2 \leq 2 \|v\|_{L^2(\Omega)} \|\partial_z v\|_{L^2(\Omega)}. \quad (13)$$

Más generalmente, si $v \in H^1(\Omega)$ entonces $v \in L_z^\infty L_x^2(\Omega)$, y existe una constante $C = C(\Omega) > 0$ tal que:

$$\|v\|_{L_z^\infty L_x^2}^2 \leq C(\Omega) \|v\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)} \quad \forall v \in H^1(\Omega). \quad (14)$$

(b) Sea $v \in L^2(\Omega)$ una función tal que $\nabla_{\mathbf{x}} v \in L^2(\Omega)^2$ y $(vn_{x_i})|_{\Gamma_b \cup \Gamma_t} = 0$ ($i = 1, 2$). Entonces, $v \in L_z^2 L_x^4(\Omega)$ y verifica la estimación:

$$\|v_i\|_{L_z^2 L_x^4}^2 \leq 4 \|v_i\|_{L^2(\Omega)} \|\nabla_{\mathbf{x}} v_i\|_{L^2(\Omega)}. \quad (15)$$

Más generalmente, si $v \in H^1(\Omega)$ entonces $v \in L_z^2 L_x^4$, y existe una constante $C = C(\Omega) > 0$ tal que:

$$\|v\|_{L_z^2 L_x^4}^2 \leq C(\Omega) \|v\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)}. \quad (16)$$

Lema 7 (Nuevas estimaciones para v_3) Sea $\mathbf{v} \in L^2(\Omega)^2$ una función tal que $\nabla_{\mathbf{x}} \cdot \mathbf{v} \in H^1(\Omega)$. Entonces, si consideramos v_3 definida en función de $\nabla_{\mathbf{x}} \cdot \mathbf{v}$ como en el Lema 4, tenemos que $v_3 \in L_z^\infty L_x^4(\Omega)$ y verifica la estimación:

$$\|v_3\|_{L_z^\infty L_x^4} \leq C(\Omega) \|\nabla_{\mathbf{x}} \cdot \mathbf{v}\|_{L^2(\Omega)}^{1/2} \|\nabla_{\mathbf{x}} \cdot \mathbf{v}\|_{H^1(\Omega)}^{1/2}.$$

Usando dichas desigualdades, conseguimos acotar el término I_2 de la forma:

$$\begin{aligned} I_2 &\leq \|(w_3)_m\|_{L_z^\infty L_x^4} \|\partial_z \mathbf{w}_m\|_{L_z^2 L_x^4} \|\mathbf{A} \mathbf{w}_m\|_{L^2(\Omega)} \\ &\leq \frac{C}{\nu^{3/2}} \|\mathbf{A} \mathbf{w}_m\|_{L^2(\Omega)}^2 \|\mathbf{w}_m\|_V \end{aligned}$$

para $C = C(\Omega) > 0$ una constante. Siguiendo entonces un razonamiento similar al del Teorema 5, separando además por una parte la influencia de los datos de tipo $L^2(0, T)$ y $L^\infty(0, T)$, y por otra la dependencia explícita respecto a la viscosidad (con constantes que ahora sólo dependen del dominio), se llega al siguiente resultado [8]:

Teorema 8 (Solución global fuerte para datos pequeños en el caso 3D) Sea $S \subset \mathbb{R}^d$ ($d = 1$ o 2) un dominio C^3 y $h \in C^3(\bar{S})$ tal que $h \geq h_{\min} > 0$ en \bar{S} . Supongamos que $\mathbf{u}_0 \in V$, $\mathbf{F} = \mathbf{f}_1 + \mathbf{f}_2$ con $\mathbf{f}_1 \in L^2(0, T; L^2(\Omega)^d)$ y $\mathbf{f}_2 \in L^\infty(0, T; L^2(\Omega)^d)$, $\tau = \tau_1 + \tau_2$ con $\tau_1 \in L^2(0, T; H_0^{1/2+\varepsilon}(\Gamma_s)^d)$ y $\tau_2 \in L^\infty(0, T; H_0^{1/2+\varepsilon}(\Gamma_s)^d)$ para algún $\varepsilon > 0$, tal que $\partial_t \tau_1 \in L^2(0, T; H^{-1/2}(\Gamma_s)^d)$ y $\partial_t \tau_2 \in L^\infty(0, T; H^{-1/2}(\Gamma_s)^d)$. Si, además los datos verifican las siguientes “condiciones de pequeñez”:

$$(H)_{3D} \left\{ \begin{array}{l} \|\mathbf{f}_1\|_{L_T^2(L^2)} + \|\tau_1\|_{L_T^2(H_0^{1/2+\varepsilon})} < c\nu^{3/2}, & \|\partial_t \tau_1\|_{L_T^2(H^{-1/2})} < c\nu^{5/2}, \\ \|\mathbf{f}_2\|_{L_T^\infty(L^2)} + \|\tau_2\|_{L_T^\infty(H_0^{1/2+\varepsilon})} < c\nu^2, & \|\partial_t \tau_2\|_{L_T^\infty(H^{-1/2})} < c\nu^3, \\ \|\mathbf{u}_0\|_{H^1} < c\nu\sqrt{\frac{\nu}{\bar{\nu}}}, & \|\tau_1(0)\|_{H^{-1/2}} + \|\tau_2(0)\|_{H^{-1/2}} < c\nu^2\sqrt{\frac{\nu}{\bar{\nu}}}, \end{array} \right.$$

donde $\nu = \min\{\nu_h, \nu_v\}$, $\bar{\nu} = \max\{\nu_h, \nu_v\}$ y c es una constante suficientemente pequeña (que sólo depende de Ω), entonces existe una (única) solución fuerte (\mathbf{u}, p_s) de (EP) en $(0, T)$ (p_s es única salvo una función aditiva que sólo depende de t).

Nota 3 Hemos denotado $L_T^q(L^p) = L^q(0, T; L^p(\Omega))$, $H^{-1/2} = H^{-1/2}(\Gamma_s)$ y $H_0^{1/2+\varepsilon} = H_0^{1/2+\varepsilon}(\Gamma_s)$.

Tratamos ahora de evitar la hipótesis de pequeñez sobre los datos. Partimos entonces de la expresión relativa a (12) para el caso 3D:

$$\begin{aligned} \frac{d}{dt} \|\mathbf{w}\|_V^2 + \|A\mathbf{w}\|_{L^2(\Omega)}^2 &\leq \frac{C}{\nu^{3/2}} \|A\mathbf{w}\|_{L^2(\Omega)}^2 \|\mathbf{w}\|_V + \frac{C}{\nu^{11}} \|\mathbf{w}\|_V^{10} \\ &+ a(t) \|\mathbf{w}\|_V^2 + b(t), \end{aligned} \quad (17)$$

donde $a(t)$ y $b(t)$ son funciones de $L^1(0, T)$, dependientes de ν y de los datos. Al contrario del argumento (de punto fijo) hecho en [7] donde había que imponer que h_{\max} fuese suficientemente pequeño, en [8] hacemos un nuevo razonamiento que nos permite eliminar esta hipótesis. En síntesis, consiste en lo siguiente: Como $\mathbf{w}_m(0) = \mathbf{0}$ y \mathbf{w}_m es una función continua en tiempo con valores en $H^1(\Omega)$, podemos elegir un tiempo T_m^1 tal que:

$$\|\mathbf{w}_m(t)\|_V \leq \frac{\nu^{3/2}}{2C}, \quad \forall t \in [0, T_m^1].$$

Entonces, lo que tenemos que demostrar es que podemos elegir T_m^1 acotado inferiormente por un tiempo $T_* \geq 0$ e independientemente de m . Integrando la expresión (17) entre 0 y t , $t \in [0, T_m^1]$, y usando $(H)_{3D}$ se obtiene [8]:

$$\|\mathbf{w}_m(t)\|_V^2 + \int_0^t \|A\mathbf{w}_m(s)\|_{L^2(\Omega)}^2 ds \leq K(\nu, c) t + C\nu^2 \int_0^t \|\mathbf{e}(s)\|_{H^2(\Omega)}^2 ds.$$

Entonces, eligiendo T^2 tal que:

$$K(\nu, c) T^2 + C\nu^2 \|\mathbf{e}\|_{L_{T^2}^2(H^2(\Omega))}^2 < \frac{\nu^3}{4C^2},$$

no es difícil verificar que se puede elegir T_m^1 igual a T^2 para cada m . La prueba de la existencia de solución fuerte en $(0, T_2)$ a partir de aquí se puede concluir de forma estándar. ■

3.2 Comportamiento asintótico de las soluciones.

En [8] también se estudia el comportamiento en tiempo hacia un estado estacionario (generado por los datos segundo miembro \mathbf{f}_2 y condición de contorno Neumann τ_2 , supuestos independientes del tiempo). El objetivo es obtener un resultado de convergencia en norma V , lo que en principio nos obliga a conocer bajo qué condiciones se obtiene la regularidad fuerte del problema estacionario:

$$(EP)_{st} \begin{cases} -\nu_h \Delta_{\mathbf{x}} \mathbf{v} - \nu_v \partial_{zz}^2 \mathbf{v} + (\mathbf{v} \cdot \nabla_{\mathbf{x}}) \mathbf{v} + v_3 \partial_z \mathbf{v} + \nabla_{\mathbf{x}} p_s = \mathbf{f}_2 & \text{en } \Omega, \\ \nabla_{\mathbf{x}} \cdot \langle \mathbf{v} \rangle = 0 & \text{en } S, \\ \nu_v \partial_z \mathbf{v}|_{\Gamma_s} = \tau_2, \quad \mathbf{v}|_{\Gamma_b \cup \Gamma_l} = \mathbf{0}. \end{cases}$$

Centrándonos en el caso 3D, obtenemos el siguiente resultado [8]:

Teorema 9 *Si los datos \mathbf{f}_2 y τ_2 son suficientemente pequeños en las normas de $L^2(\Omega)^2$ y $H_0^{1/2+\varepsilon}(\Gamma_s)^2$ respectivamente, entonces existe una única solución (fuerte) \mathbf{v} para $(EP)_{st}$ y se satisfacen las siguientes estimaciones de regularidad débil y fuerte para las soluciones: existe $C = C(\Omega) > 0$ tal que*

$$\|\mathbf{v}\|_{H^1(\Omega)}^2 \leq \frac{C}{\nu^2} \left\{ \|\mathbf{f}_2\|_{H^{-1}(\Omega)}^2 + \|\tau_2\|_{H^{-1/2}(\Gamma_s)}^2 \right\}, \quad (18)$$

$$\|\mathbf{v}\|_{H^2(\Omega)}^2 \leq \frac{C}{\nu^2} \left\{ \|\mathbf{f}_2\|_{L^2(\Omega)}^2 + \|\tau_2\|_{H_0^{1/2+\varepsilon}(\Gamma_s)}^2 \right\}. \quad (19)$$

El resultado de comportamiento asintótico que obtenemos en [8] es:

Teorema 10 (Convergencia de la solución 3D de evolución hacia la solución 3D estacionaria) *Sea \mathbf{u} una solución fuerte de (EP) en $(0, +\infty)$ con segundo miembro $\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2$, donde $\mathbf{f}_1 \in L^2(0, +\infty; L^2(\Omega)^2)$ y $\mathbf{f}_2 \in L^2(\Omega)^2$ (independiente de t), y la condición de Newman $\tau = \tau_1 + \tau_2$, donde $\tau_1 \in L^2(0, +\infty; H_0^{1/2+\varepsilon}(\Gamma_s)^2)$ para algún $\varepsilon > 0$, tal que $\partial_t \tau_1 \in L^2(0, +\infty; H^{-1/2}(\Gamma_s)^2)$, y $\tau_2 \in H_0^{1/2+\varepsilon}(\Gamma_s)^2$ para algún $\varepsilon > 0$ (también independiente de t). Suponiendo condiciones de pequeñez $(H)_{3D}$ con $T = +\infty$, si \mathbf{v} es la solución fuerte estacionaria de $(EP)_{st}$ con segundo miembro \mathbf{f}_2 y condición de contorno Neumann τ_2 , entonces $\mathbf{u}(t) \rightarrow \mathbf{v}$ en la norma $H^1(\Omega)$ cuando $t \uparrow +\infty$.*

La demostración sigue los siguientes pasos:

- 1) Se considera el problema verificado por $\mathbf{w} = \mathbf{u} - \mathbf{v} - \mathbf{e}$ donde \mathbf{v} es la solución fuerte del problema estacionario (dada en el Teorema 9) y \mathbf{e} es la solución fuerte del problema lineal evolutivo (S) con datos $\mathbf{g} = \mathbf{f}_1$, $\mathbf{e}(0) = \mathbf{u}_0 - \mathbf{v}$ y $\tau = \tau_1$ (dada en el Teorema 3). Consideramos su formulación variacional y estimamos con funciones test del tipo $A\mathbf{w}$, obteniendo:

$$\begin{aligned} \frac{d}{dt} \|\mathbf{w}\|_V^2 + \|A\mathbf{w}\|_{L^2(\Omega)}^2 &\leq \frac{C}{\nu^{3/2}} \|A\mathbf{w}\|_{L^2(\Omega)}^2 \|\mathbf{w}\|_V + \frac{C}{\nu^{11}} \|\mathbf{w}\|_V^{10} \\ &+ (a_1(t) + a_2(t)) \|\mathbf{w}\|_V^2 + b(t), \end{aligned} \quad (20)$$

donde $a_1 \in L^\infty(0, \infty)$, a_2 y $b \in L^1(0, \infty)$, son funciones dependientes de los datos.

- 2) Gracias a la hipótesis de pequeñez sobre los datos, conseguimos que para un γ suficientemente pequeño:

$$\|\mathbf{w}(t)\|_V \leq \gamma \nu^{3/2}, \quad \forall t \in [0, +\infty).$$

Entonces, a partir de (20) se puede obtener:

$$\frac{d}{dt} \|\mathbf{w}\|_V^2 + \frac{\nu}{2C} \|\mathbf{w}\|_V^2 \leq C\nu^2 \|\mathbf{e}(t)\|_{H^2(\Omega)}^2.$$

Integrando la expresión anterior, obtenemos que:

$$\|\mathbf{w}(t)\|_V^2 \leq \int_0^t \exp\left(\frac{\nu}{2C}(s-t)\right) \|\mathbf{e}(s)\|_{H^2(\Omega)}^2 ds. \quad (21)$$

Como por hipótesis sobre los datos y la construcción de \mathbf{e} , $\|\mathbf{e}\|_{H^2(\Omega)}^2 \in L^1(0, +\infty)$, entonces para todo $\delta > 0$ existe un $T_* \in [0, +\infty)$ tal que

$$\int_{T_*}^{+\infty} \|\mathbf{e}(t)\|_{H^2(\Omega)}^2 dt < \delta.$$

Así pues, descomponiendo la cota de (21) como:

$$\begin{aligned} & e^{-\frac{\nu}{2C}t} \int_0^t e^{\frac{\nu}{2C}s} \|\mathbf{e}(s)\|_{H^2(\Omega)}^2 ds \\ & \leq e^{-\frac{\nu}{2C}t} e^{\frac{\nu}{2C}T_*} \int_0^{T_*} \|\mathbf{e}(s)\|_{H^2(\Omega)}^2 ds + \int_{T_*}^{+\infty} \|\mathbf{e}(s)\|_{H^2(\Omega)}^2 ds, \end{aligned}$$

podemos concluir que $\|\mathbf{w}(t)\|_V^2 = \|\mathbf{u}(t) - \mathbf{e}(t) - \mathbf{v}\|_V^2 \rightarrow 0$ cuando $t \uparrow +\infty$.

- 3) Razonando de manera análoga obtenemos que $\|\mathbf{e}(t)\|_V \rightarrow 0$ cuando $t \uparrow +\infty$, lo que nos permite concluir el resultado de convergencia $\|\mathbf{u}(t) - \mathbf{v}\|_{H^1(\Omega)}^2 \rightarrow 0$. ■

3.3 Unicidad de solución débil/fuerte.

La menor regularidad del término no lineal (de convección vertical) en el sistema de Ecuaciones Primitivas hace que se necesite mayor regularidad para demostrar la unicidad de solución que en el caso del sistema de Navier-Stokes. Para comprobarlo, basta con intentar reproducir la demostración de unicidad que aparece, por ejemplo, en el libro de P. L. Lions [13]. El principal inconveniente es la acotación del término:

$$\int_{\Omega} u_3 \partial_z \psi \cdot \mathbf{u} \, d\Omega$$

Usando las estimaciones isotropas, conseguimos una estimación para:

$$\begin{cases} \mathbf{u} \in L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega)), & u_3 \in L^2(0, T; L^2(\Omega)) \\ \partial_z \psi \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; L^\infty) \end{cases}$$

Sin embargo, el uso de las estimaciones anisótropas nos permite rebajar la regularidad a los siguientes espacios anisótropos:

$$\begin{cases} \mathbf{u} \in L^4(0, T; L_z^2 L_{\mathbf{x}}^4), & u_3 \in L^2(0, T; L_z^\infty L_{\mathbf{x}}^2), \\ \partial_z \psi \in L^4(0, T; L_z^2 L_{\mathbf{x}}^4), \end{cases}$$

de manera que podemos enunciar el siguiente resultado:

Teorema 11 (*Unicidad de solución débil/fuerte*) Sea \mathbf{u} solución débil de (EP) en $(0, T)$. Si existe $\underline{\mathbf{u}}$ una solución débil de (EP) en $(0, T)$ tal que:

$$\nabla_{\mathbf{x}} \underline{\mathbf{u}} \in L^2(0, T; L_z^\infty L_{\mathbf{x}}^2) \quad y \quad \partial_z \underline{\mathbf{u}} \in L^4(0, T; L_z^2 L_{\mathbf{x}}^4), \quad (22)$$

entonces ambas soluciones coinciden en $[0, T)$.

Nota 4 En 2D, la hipótesis (22) se rebaja sólo a $\partial_z \underline{\mathbf{u}} \in L^4(0, T; L^2(\Omega))$. En cualquier caso, dicha regularidad adicional no está asegurada en general para una solución débil.

4 El modelo de Ecuaciones Primitivas del océano con condiciones de contorno de tipo Navier.

Ya sabemos que las Ecuaciones Primitivas se obtienen a partir de las ecuaciones de Navier-Stokes anisótropas por un análisis asintótico cuando el cociente de aspecto δ tiende a cero, completando dicho modelo con determinadas condiciones de contorno (ver [1, 2] para condiciones Dirichlet en el fondo). Desde un punto de vista físico, la elección de condiciones de contorno de tipo Dirichlet homogéneas para la velocidad sobre el fondo sólo está justificada cuando la viscosidad molecular del fluido sea importante. Sin embargo, en muchos modelos geofísicos la viscosidad que se considera es turbulenta, siendo la viscosidad molecular despreciable. Por otra parte, se sabe que el uso de la condición de contorno de fricción de tipo Navier en las ecuaciones de Navier-Stokes previene la aparición de capas límites.

El objetivo de esta sección es doble. Primero, obtener nuevas condiciones de contorno para las Ecuaciones Primitivas, partiendo de condiciones de contorno de tipo Navier en Navier-Stokes con viscosidad anisótropa y haciendo tender el cociente de aspecto a cero ([3]). Segundo, analizar una cierta regularidad adicional global en tiempo para cualquier solución débil del modelo en un dominio 2D que implica, en particular, la unicidad ([4]).

4.1 Obtención del modelo con nuevas condiciones de contorno.

Consideramos un fluido gobernado por las ecuaciones de Navier-Stokes de evolución, con fuerza de Coriolis y viscosidad anisótropa (turbulenta). Supondremos que la densidad es constante (igual a 1). Suponemos que la viscosidad anisótropa ν viene dada por (ν_h, ν_h, ν_z) . Definimos el gradiente anisótropo como:

$$\nabla_\nu = (\nu_h \partial_{x_1}, \nu_h \partial_{x_2}, \nu_z \partial_z) = (\nu_h \nabla_{\mathbf{x}}, \nu_z \partial_z), \quad (23)$$

y el tensor de esfuerzos tangenciales debidos al campo de velocidades, viene dado por:

$$D_\nu(\mathbf{u}^\delta) = (\nabla_\nu \mathbf{u}^\delta + \nabla_\nu^t \mathbf{u}^\delta) / 2, \quad (24)$$

donde \mathbf{u}^δ es la velocidad del fluido, dada por $(\mathbf{v}^\delta, w^\delta)$, donde $\mathbf{v}^\delta = (v_1^\delta, v_2^\delta)$ es la componente horizontal y w^δ es la componente vertical. Supongamos el fluido en un dominio $\Omega_\delta = \{(\mathbf{x}, z) \in \mathbb{R}^3 / \mathbf{x} = (x_1, x_2) \in S, -\delta h(\mathbf{x}) < z < 0\}$, y que el flujo satisface una condición de tracción sobre la superficie Γ_s debida al viento, una condición de contorno Dirichlet homogénea sobre las paredes laterales Γ_l^δ (si existen) y una condición de contorno de tipo fricción sobre el fondo Γ_b^δ . Más precisamente, suponemos que la velocidad y el potencial $(\mathbf{u}^\delta, p^\delta)$ satisfacen el siguiente sistema:

$$\left\{ \begin{array}{ll} \partial_t \mathbf{u}^\delta + \mathbf{u}^\delta \cdot \nabla \mathbf{u}^\delta - 2 \nabla \cdot D_\nu(\mathbf{u}^\delta) + 2 \mathbf{W} \times \mathbf{u}^\delta + \nabla p^\delta = 0 & \text{en } (0, T) \times \Omega_\delta, \\ \nabla \cdot \mathbf{u}^\delta = 0 & \text{en } (0, T) \times \Omega_\delta, \\ \nu_z \partial_z \mathbf{v}^\delta = \alpha |\mathbf{v}_{air}| (\mathbf{v}_{air} - \mathbf{v}^\delta), \quad w^\delta = 0 & \text{sobre } (0, T) \times \Gamma_s^\delta, \\ (2 D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta + \gamma \mathbf{u}^\delta)_{tg} = 0 \quad \mathbf{u}^\delta \cdot \mathbf{n}_\delta = 0 & \text{sobre } (0, T) \times \Gamma_b^\delta, \\ \mathbf{u}^\delta = 0 & \text{sobre } (0, T) \times \Gamma_l^\delta, \\ \mathbf{u}^\delta|_{t=0} = \mathbf{u}_0^\delta & \text{en } \Omega_\delta, \end{array} \right. \quad (25)$$

con $D_\nu(\mathbf{u}^\delta)$ definido en (24). La fuerza de Coriolis viene dada por $2 \mathbf{W} \times \mathbf{u}^\delta$ con $\mathbf{W} = |\mathbf{W}|(0, \cos \theta, \sin \theta)$ el vector de rotación de la Tierra y $\theta = \theta(y)$ es la latitud. Denotamos por $\mathbf{v}_{air} : S \rightarrow \mathbb{R}^2$ la velocidad del aire sobre la superficie, y α y γ dos funciones de \mathbf{x} no negativas y acotadas. Recordemos que $D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta$ representa el producto matriz por vector definido por:

$$(D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta)_i = \sum_{j=1}^3 (D_\nu(\mathbf{u}^\delta))_{ij} n_\delta^j$$

y $(D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta)_{tg}$ corresponde a su componente tangencial, dada por:

$$(D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta)_{tg} = D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta - [(D_\nu(\mathbf{u}^\delta) \mathbf{n}_\delta) \cdot \mathbf{n}_\delta] \mathbf{n}_\delta,$$

donde \mathbf{n}_δ es el vector normal exterior a Γ_b^δ , dado por:

$$\mathbf{n}_\delta = (n_\delta^1, n_\delta^2, n_\delta^3) = \frac{1}{\sqrt{1 + |\delta \nabla_{\mathbf{x}} h|^2}} (-\delta \nabla_{\mathbf{x}} h, -1). \quad (26)$$

Nota 5 La condición $(25)_3$ representa el efecto (no lineal) que el aire ejerce sobre la velocidad horizontal del flujo de agua en el Océano. Por otra parte, en la condición $(25)_4$ el coeficiente $\gamma = \gamma(\mathbf{x})$ depende de la rugosidad del fondo. Así pues, la condición $(25)_4$ completa viene a expresar los efectos que la fricción con el fondo ejercen sobre la componente tangencial de $2D_\nu(\mathbf{u}^\delta)$.

Introducimos los siguientes espacios funcionales para el dominio de partida Ω_δ :

$$\begin{aligned} \mathcal{V}_\delta &= \{ \Phi = (\varphi, \psi) \in (C_l^\infty(\overline{\Omega}_\delta))^2 \times (C_l^\infty(\overline{\Omega}_\delta)) : \\ &\quad \nabla \cdot \Phi = 0 \text{ en } \Omega_\delta, \Phi \cdot \mathbf{n}_\delta = 0 \text{ sobre } \partial\Omega_\delta \} \end{aligned}$$

donde $C_l^\infty(\overline{\Omega}_\delta) = \{ \chi \in C^\infty(\overline{\Omega}_\delta) : \chi = \mathbf{0} \text{ en un entorno de } \Gamma_l^\delta \}$, y definimos H_δ y V_δ como la clausura de \mathcal{V}_δ en $(L^2(\Omega_\delta))^3$ y $(H^1(\Omega_\delta))^3$, respectivamente.

Notemos que el modelo que obtenemos es no estándar, ya que el término de segundo orden $-2\nabla \cdot (D_\nu(\mathbf{u}^\delta))$ no es necesariamente difusivo pues, en general, $-2\langle \nabla \cdot (D_\nu(\mathbf{u}^\delta)), \mathbf{u}^\delta \rangle \not\geq 0$.

Teorema 12 Supongamos $h \in W^{2,\infty}(S)$, $h > 0$ en S , $|\nabla_{\mathbf{x}} h| > c > 0$ sobre $\partial S \cap \{h = 0\}$, $\mathbf{u}_0^\delta \in H_\delta$, $\alpha^{1/3} \mathbf{v}_{air} \in L^3(0, T; L^3(\Omega)^2)$ y

$$\left\{ \begin{array}{l} C(\Omega) |\nu_h - \nu_z| \|\nabla_{\mathbf{x}} h\|_{W^{1/2,\infty}(S)} \leq \frac{\nu_h}{4}, \\ \delta^2 C(\Omega) |\nu_h - \nu_z| \left(1 + 2\|h\|_{W^{1,\infty}(S)}^2\right) \|\nabla_{\mathbf{x}} h\|_{W^{1/2,\infty}(S)} \leq \frac{\nu_z}{16}, \\ 2\delta |\nu_h - \nu_z| C(\Omega) \|\nabla_{\mathbf{x}} h\|_{W^{1/2,\infty}(S)} |h(\mathbf{x}) + \delta \nu_h| \partial_{ij}^2 h(\mathbf{x})| \\ \leq \frac{\gamma(\mathbf{x})}{2} \sqrt{1 + \delta^2 |\nabla_{\mathbf{x}} h(\mathbf{x})|^2}, \quad \forall i, j = 1, 2, \text{ c.p.d. } \mathbf{x} \in S, \end{array} \right. \quad (27)$$

donde $C(\Omega) = C(S)C'(\Omega)$ con $C'(\Omega)$ depende de las inyecciones de Sobolev de H^1 en $H^{1/2}$ -frontera. Entonces, existe una solución débil del problema tal que: $\forall \Phi = (\varphi, \psi) \in C^1([0, T]; \mathcal{V}_\delta)$ con $\Phi(T) = 0$,

$$\left\{ \begin{array}{l} - \int_0^T \int_{\Omega_\delta} \mathbf{u}^\delta \cdot (\partial_t \Phi + \mathbf{u}^\delta \cdot \nabla \Phi) + 2 \int_0^T \int_{\Omega_\delta} D_\nu(\mathbf{u}^\delta) \cdot \nabla \Phi \\ + \int_0^T \int_{\Gamma_s} \alpha |\mathbf{v}_{air}| (\mathbf{v}^\delta - \mathbf{v}_{air}) \cdot \varphi + 2 \int_0^T \int_{\Omega_\delta} (\mathbf{W} \times \mathbf{u}^\delta) \cdot \Phi \\ + \int_0^T \int_{\Gamma_b^\delta} \gamma (\mathbf{u}^\delta \times \mathbf{n}_\delta) \cdot (\Phi \times \mathbf{n}_\delta) = \int_{\Omega_\delta} \mathbf{u}_0^\delta \cdot \Phi(0). \end{array} \right. \quad (28)$$

Además, dicha solución satisface la siguiente desigualdad de energía, para todo

$t \geq 0$:

$$\begin{aligned} & \|\mathbf{u}^\delta(t)\|_{L^2(\Omega_\delta)}^2 + \nu_h \int_0^t \|\nabla_{\mathbf{x}} \mathbf{u}^\delta(s)\|_{L^2(\Omega_\delta)}^2 ds + \frac{\nu_z}{2} \int_0^t \|\partial_z \mathbf{u}^\delta(s)\|_{L^2(\Omega_\delta)}^2 ds \\ & + \int_0^t \int_{\Gamma_s} \alpha |\mathbf{v}_{air}| |\mathbf{v}^\delta|^2 + \int_0^t \int_{\Gamma_b^\delta} \gamma |\mathbf{u}^\delta|^2 \leq \|\mathbf{u}_0^\delta\|_{L^2(\Omega_\delta)}^2 + \int_0^t \int_S \alpha |\mathbf{v}_{air}|^3. \end{aligned} \quad (29)$$

Nota 6 Las hipótesis (27) aseguran la disipación del problema, necesaria para que el modelo sea físicamente admisible.

Esquema de la demostración: Se sigue un procedimiento de tipo Galerkin. Al acotar los términos que aparecen en la formulación variacional al tomar como funciones test las soluciones de Galerkin aproximadas, aparece un término conflictivo que proviene del gradiente traspuesto de la velocidad, que ahora no se anula porque tenemos un gradiente anisótropo. Dicho término se puede reescribir como un término sobre la frontera de la forma:

$$\begin{aligned} & \int_{\Omega_\delta} {}^t \nabla_{\nu} \mathbf{u}^\delta \cdot \nabla \mathbf{u}^\delta = \delta \nu_h \int_S \mathcal{H} \mathbf{v}^\delta|_{\Gamma_b^\delta} \cdot \mathbf{v}^\delta|_{\Gamma_b^\delta} d\mathbf{x} \\ & + \delta (\nu_h - \nu_z) \left\langle \nabla_{\mathbf{x}} \cdot \left(\mathbf{v}^\delta|_{\Gamma_b^\delta} \right), \mathbf{v}^\delta|_{\Gamma_b^\delta} \cdot \nabla_{\mathbf{x}} h \right\rangle_{H^{-1/2}(S), H^{1/2}(S)}, \end{aligned} \quad (30)$$

donde \mathcal{H} representa el hessiano de \mathbf{v}^δ , es decir, $\mathcal{H} = (\partial_{ij}^2 h)_{ij}$. Para acotar este término se usan desigualdades de interpolación de espacios de Sobolev H^s en sus correspondientes espacios de trazas, imponiendo las hipótesis (27) para llegar a controlar este término con las normas L^2 y H^1 que aparecen a la izquierda de la desigualdad de energía que se pretende conseguir (ver [3, 15] para los detalles). ■

Nota 7 [3]

- 1) Si $\gamma \neq 0$ y $|\Gamma_l^\delta| > 0$, se pueden reescribir las hipótesis (27) de manera más fácil ya que la seminorma del gradiente en $L^2(\Omega_\delta)$ es equivalente a la norma en $H^1(\Omega_\delta)$.
- 2) Si $\gamma \neq 0$ y $|\Gamma_l^\delta| = 0$, se puede obtener existencia de solución débil sin la hipótesis (27)₃, pero entonces no se garantiza la disipatividad del sistema.
- 3) Si $\gamma = 0$ y $|\Gamma_l^\delta| = 0$ (es decir, en el caso de deslizamiento sobre el fondo y sin paredes laterales), se puede garantizar la existencia de solución débil pero no la disipación del sistema.

Nos interesamos ahora por el comportamiento asintótico cuando $\delta \rightarrow 0$ de las soluciones débiles \mathbf{u}^δ (para cada dominio Ω_δ) obtenidas en el Teorema 12. Supondremos que $h \approx 1$ y:

$$\nu_h \approx 1, \quad \nu_z = \delta^2 \nu_v \quad (\nu_v \approx 1), \quad (31)$$

$$\alpha/\delta \rightarrow \bar{\alpha} \quad \text{y} \quad \gamma/\delta \rightarrow \bar{\gamma} \quad \text{en } L^\infty(S)\text{-débil*}, \quad (32)$$

e introducimos los siguientes cambios de variable:

$$Z = \frac{z}{\delta}, \quad \mathbf{V}^\delta(t; \mathbf{x}, Z) = \mathbf{v}^\delta(t; \mathbf{x}, z), \quad W^\delta(t; \mathbf{x}, Z) = \frac{w^\delta(t; \mathbf{x}, z)}{\delta}, \quad (33)$$

que nos llevan al dominio (adimensional)

$$\Omega = \{(\mathbf{x}, Z) \in \mathbb{R}^3 : \mathbf{x} \in S, -h(\mathbf{x}) < Z < 0\},$$

cuyas fronteras vienen dadas por $\Gamma_s \equiv \bar{S} \times \{0\}$, $\Gamma_l = \{(\mathbf{x}, Z) \in \mathbb{R}^3 : \mathbf{x} \in \partial S, -h(\mathbf{x}) < Z < 0\}$ y $\Gamma_b = \partial\Omega \setminus (\bar{\Gamma}_s \cup \bar{\Gamma}_l)$. Probamos entonces que cuando δ tiende a 0, entonces la solución $\mathbf{U}^\delta = (\mathbf{V}^\delta, W^\delta)$ reescalada según (33) a partir de una solución $(\mathbf{v}^\delta, w^\delta)$ de las ecuaciones de Navier-Stokes (25), posee una subsucesión que converge a una solución $\mathbf{U} = (\mathbf{V}, W)$ del modelo de Ecuaciones Primitivas:

$$\left\{ \begin{array}{ll} \partial_t \mathbf{V} + (\mathbf{U} \cdot \nabla) \mathbf{V} - \Delta_\nu \mathbf{V} + k \mathbf{V}^\perp + \nabla_{\mathbf{x}} P & = 0 & \text{en } (0, T) \times \Omega, \\ \partial_Z P = 0, \quad \nabla_{\mathbf{x}} \cdot \mathbf{V} + \partial_Z W & = 0 & \text{en } (0, T) \times \Omega, \\ \nu_v \partial_Z \mathbf{V} = \bar{\alpha} |\mathbf{v}_{air}| (\mathbf{v}_{air} - \mathbf{V}), \quad W & = 0 & \text{sobre } (0, T) \times \Gamma_s, \\ \nu_v (1 + \nu_h |\nabla_{\mathbf{x}} h|^2 / \nu_v) \partial_Z \mathbf{V} n^Z & = \mathcal{A} \mathbf{V}|_{\Gamma_b}, & \text{sobre } (0, T) \times \Gamma_b, \\ (\mathbf{V}, W) \cdot \mathbf{n} & = 0 & \text{sobre } (0, T) \times \Gamma_b, \\ \mathbf{V} & = 0 & \text{sobre } (0, T) \times \Gamma_l, \\ \mathbf{V}|_{t=0} & = \mathbf{V}_0 & \text{en } \Omega, \end{array} \right. \quad (34)$$

donde $\nabla_{\mathbf{x}} = (\partial_{x_1}, \partial_{x_2})$, $\Delta_\nu = \nu_h \Delta_{\mathbf{x}} + \nu_v \partial_Z^2$, $\Delta_{\mathbf{x}} = \partial_{x_1}^2 + \partial_{x_2}^2$, $\mathbf{V}^\perp = (-V_2, V_1)$, $k = |\mathbf{W}| \sin \theta$ y

$$\mathcal{A} \mathbf{V}|_{\Gamma_b} = \bar{\gamma} \mathbf{V}|_{\Gamma_b} n^Z + \nu_h \mathcal{H} \mathbf{V}|_{\Gamma_b} n^Z + \nu_h \nabla_{\mathbf{x}} \cdot (\mathbf{V}|_{\Gamma_b}) n^{\mathbf{x}} - \nu_h \nabla_{\mathbf{x}} (\mathbf{V}|_{\Gamma_b}) n^{\mathbf{x}}, \quad (35)$$

con $\mathbf{n} = (\mathbf{n}^{\mathbf{x}}, n^Z)$ el vector normal exterior a Γ_b . La condición (35) tendrá un sentido “dual” a través de la formulación débil (38) de (34).

Nota 8 En dimensión 2, la condición de contorno en el fondo se reduce a:

$$\nu_v \partial_Z \mathbf{V}|_{\Gamma_b} = \beta \mathbf{V}|_{\Gamma_b} \quad \text{con } \beta = \frac{(\bar{\gamma} + \nu_h h'')}{1 + \nu_h |h'|^2 / \nu_v}.$$

Considerando entonces los siguientes espacios funcionales en el límite:

$$\mathcal{V} = \{\tilde{\Phi} = (\tilde{\varphi}, \tilde{\psi}) \in (C_l^\infty(\bar{\Omega}))^3 : \nabla \cdot \tilde{\Phi} = 0 \text{ en } \Omega, \tilde{\Phi} \cdot \mathbf{n} = 0 \text{ sobre } \partial\Omega\},$$

donde $C_l^\infty(\bar{\Omega}) = \{\chi \in C^\infty(\bar{\Omega}) : \chi = \mathbf{0} \text{ en un entorno de } \Gamma_l\}$, y definiendo H^{prim} , V^{prim} y Y^{prim} como las clausuras de \mathcal{V} para las normas $\|\tilde{\Phi}\|_{H^{prim}} = \|\tilde{\varphi}\|_{(L^2(\Omega))^2}$, $\|\tilde{\Phi}\|_{V^{prim}} = \|\tilde{\varphi}\|_{(H^1(\Omega))^2}$ y $\|\tilde{\Phi}\|_{Y^{prim}} = \|\tilde{\varphi}\|_{(H^2(\Omega))^2}$ respectivamente, donde $\tilde{\varphi} = (\tilde{\Phi}_1, \tilde{\Phi}_2)$, obtenemos el siguiente resultado [3]:

Teorema 13 *En las hipótesis del Teorema 12, si suponemos además que:*

$$\begin{aligned} \|\mathbf{V}_0^\delta\|_{(L^2(\Omega))^2}^2 + \|\delta W_0^\delta\|_{L^2(\Omega)}^2 &\leq C_{inic}, \\ \mathbf{V}_0^\delta &\rightharpoonup \mathbf{V}_0 \quad \text{en } L^2(\Omega)^2\text{-débil*}, \end{aligned} \quad (36)$$

entonces, existe una subsucesión de $\mathbf{U}^\delta = (\mathbf{V}^\delta, W^\delta)$ (obtenida por el reescalamiento (33) de una solución \mathbf{u}^δ de las ecuaciones de Navier-Stokes) que converge a $\mathbf{U} = (\mathbf{V}, W)$ en el sentido siguiente:

$$\begin{aligned} (\mathbf{V}^\delta, W^\delta) &\rightarrow (\mathbf{V}, W) \quad \text{en } L^2(0, T; V^{Prim})\text{-débil} \\ \mathbf{V}^\delta &\rightarrow \mathbf{V} \quad \text{en } L^\infty(0, T; L^2(\Omega)^2)\text{-débil*}. \end{aligned} \quad (37)$$

Además, $\mathbf{U} = (\mathbf{V}, W)$ es una solución débil del sistema (34); es decir, $\forall \tilde{\Phi} = (\tilde{\varphi}, \tilde{\psi}) \in C^1([0, T]; \mathcal{V})$ con $\tilde{\Phi}(T) = 0$,

$$\left\{ \begin{aligned} & - \int_0^T \int_\Omega \mathbf{V} \cdot (\partial_t \tilde{\varphi} + \mathbf{U} \cdot \nabla \tilde{\varphi}) d\Omega dt + \int_0^T \int_\Omega k \mathbf{V}^\perp \cdot \tilde{\varphi} d\Omega dt \\ & + \int_0^T \int_\Omega (\nu_h \nabla_{\mathbf{x}} \mathbf{V} \cdot \nabla_{\mathbf{x}} \tilde{\varphi} + \nu_v \partial_Z \mathbf{V} \cdot \partial_Z \tilde{\varphi}) d\Omega dt \\ & + \int_0^T \nu_h \langle \nabla_{\mathbf{x}} \cdot (\mathbf{V}|_{\Gamma_b}), \tilde{\varphi}|_{\Gamma_b} \cdot \nabla_{\mathbf{x}} h \rangle_{H^{-1/2}(S), H^{1/2}(S)} + \nu_h \int_0^T \int_{\Gamma_b} \mathcal{H} \mathbf{V} \cdot \tilde{\varphi}(-n^Z) \\ & + \int_0^T \int_{\Gamma_s} \bar{\alpha} |\mathbf{v}_{air}| (\mathbf{V} - \mathbf{v}_{air}) \cdot \tilde{\varphi} + \int_0^T \int_{\Gamma_b} \bar{\gamma} \mathbf{V} \cdot \tilde{\varphi}(-n^Z) = \int_\Omega \mathbf{V}_0 \cdot \tilde{\varphi}(0) d\Omega \end{aligned} \right. \quad (38)$$

que verifica, además, la desigualdad de energía, para todo $t \geq 0$,

$$\begin{aligned} \|\mathbf{V}(t)\|_{L^2(\Omega)}^2 + \int_0^t \int_\Omega \left(\nu_h |\nabla_{\mathbf{x}} \mathbf{V}|^2 + \frac{\nu_v}{2} |\partial_Z \mathbf{V}|^2 \right) d\Omega ds \\ + \int_0^t \int_{\Gamma_s} \bar{\alpha} |\mathbf{v}_{air}| \|\mathbf{V}\|^2 + \int_0^t \int_{\Gamma_b} \bar{\gamma} |\mathbf{V}|^2 (-n^Z) \leq C_{inic} + \int_0^t \int_{\Gamma_s} \bar{\alpha} |\mathbf{v}_{air}|^3. \end{aligned} \quad (39)$$

Nota 9 *El límite cuando $\delta \rightarrow 0$ de las hipótesis (27) nos da las correspondientes hipótesis de disipación para el problema límite (34).*

Esquema de la demostración: Haciendo el cambio de variable (33) en la formulación variacional (28) y tomando funciones test adecuadas, podemos demostrar que la sucesión $(\mathbf{V}^\delta, W^\delta)$ está acotada en $L^2(0, T; V^{prim})$ y \mathbf{V}^δ en $L^\infty(0, T; L^2(\Omega)^3)$. Entonces existe una subsucesión (que denotaremos de la misma forma) convergente como en (37). Esto nos permite concluir la convergencia de los términos lineales de la formulación variacional.

También se puede demostrar que \mathbf{V}^δ converge fuerte a \mathbf{V} en $L^2(0, T; L^2(\Omega)^2)$, usando una variante dada en [1], de la compacidad en espacios L^p con valores

en un Banach [16], lo que permite la convergencia de los términos no lineales. Asimismo, la hipótesis (36) sobre los datos iniciales garantiza la convergencia de los términos dependientes de los datos iniciales.

Por último, para obtener la convergencia de los términos de frontera, usamos la convergencia fuerte de \mathbf{V}^δ y ciertas desigualdades de interpolación (ver [3] y [15] para los detalles). ■

Nota 10 *Aplicando el Lema de De Rham ([16]) en las formulaciones variacionales asociadas al problema aproximado (25) y al problema límite (34), podemos recuperar los potenciales respectivos P^δ y P , como distribuciones, tales que las ecuaciones de momentos relativas a (25) reescaladas con (33) y las ecuaciones de momentos relativas a (34) se satisfacen en $H^{-1}(0, T; H^{-1}(\Omega)^3)$ (espacio dual de $H_0^1(0, T; (H_0^1(\Omega))^3)$) y en $H^{-1}(0, T; (W^{-1, r'}(\Omega))^3)$ (espacio dual de $H_0^1(0, T; (W_0^{1, r}(\Omega))^3)$) respectivamente, siendo $r > 2$ y r' su exponente conjugado. Además, considerando la convergencia del resto de términos en las ecuaciones de momentos, podemos probar:*

$$\begin{aligned} \nabla_{\mathbf{x}} P^\delta &\rightarrow \nabla_{\mathbf{x}} P \quad \text{en } H^{-1}(0, T; (W^{-1, r'}(\Omega))^2)\text{-débil} \\ \|\partial_Z P^\delta\|_{H^{-1}(0, T; H^{-1}(\Omega))} &\leq C\delta. \end{aligned}$$

En particular, obtenemos convergencia fuerte de $\partial_Z P^\delta$ a $\partial_Z P (= 0)$ en $H^{-1}(0, T; H^{-1}(\Omega))$.

Llegados a este punto, sólo nos queda identificar las condiciones de contorno verificadas por la solución variacional límite (\mathbf{V}, W) obtenida en el Teorema anterior. Llamando T al tensor de esfuerzos:

$$T = \begin{pmatrix} \nu_h \partial_{x_1} V_1 - P & \nu_h \partial_{x_2} V_1 & \nu_v \partial_Z V_1 \\ \nu_h \partial_{x_1} V_2 & \nu_h \partial_{x_2} V_2 - P & \nu_v \partial_Z V_2 \\ 0 & 0 & -P \end{pmatrix},$$

obtenemos que $\nabla \cdot T \in H^{-1}(0, T; L^1(\Omega)^3)$. Por otra parte, también se tiene que $T \in H^{-1}(0, T; L^1(\Omega)^{3 \times 3})$. Estas regularidades nos permiten definir la traza normal de T en $(W^{1, \infty}(\partial\Omega)^3)'$, donde $W^{1, \infty}(\partial\Omega) = \{\phi|_{\partial\Omega} : \phi \in W^{1, \infty}(\Omega)\}$, dotado de la norma:

$$\begin{aligned} \|\chi\|_{W^{1, \infty}(\partial\Omega)} &= \inf_{\phi \in W^{1, \infty}(\Omega)} \|\phi\|_{W^{1, \infty}(\Omega)}, \\ \phi|_{\partial\Omega} &= \chi, \end{aligned}$$

como

$$\begin{aligned} \langle T\mathbf{n}; (\tilde{\varphi}, \tilde{\psi}) \rangle &= \int_{\Omega} T \cdot \nabla(\tilde{\varphi}, \tilde{\psi}) + \int_{\Omega} (\nabla \cdot T) \cdot (\tilde{\varphi}, \tilde{\psi}) \\ &= \int_{\Omega} (\nu_h \nabla_{\mathbf{x}} \mathbf{V} \cdot \nabla_{\mathbf{x}} \tilde{\varphi} + \nu_v \partial_Z \mathbf{V} \cdot \partial_Z \tilde{\varphi}) \\ &\quad - \int_{\Omega} P \nabla \cdot (\tilde{\varphi}, \tilde{\psi}) + \int_{\Omega} (\nabla \cdot T) \cdot (\tilde{\varphi}, \tilde{\psi}), \end{aligned} \tag{40}$$

para todo $(\tilde{\varphi}, \tilde{\psi}) \in (W^{1,\infty}(\Omega))^3$. Esta desigualdad se satisface en $\mathcal{D}'(0, T)$. Partiendo de la definición de traza normal para T en $(W^{1,\infty}(\partial\Omega)^3)'$, vamos a probar que la traza normal para el tensor

$$\mathcal{F} = \begin{pmatrix} \nu_h \partial_{x_1} V_1 & \nu_h \partial_{x_2} V_1 & \nu_v \partial_Z V_1 \\ \nu_h \partial_{x_1} V_2 & \nu_h \partial_{x_2} V_2 & \nu_v \partial_Z V_2 \end{pmatrix}$$

está definida en $(W^{1,\infty}(\partial\Omega)^2)'$. Con este objetivo, para cada $\tilde{\varphi} \in (W^{1,\infty}(\Omega))^2$ consideramos $\tilde{\psi}$ de tal forma que $(\tilde{\varphi}, \tilde{\psi}) \cdot \mathbf{n} = 0$ sobre $\partial\Omega$. Así, podemos definir $\langle \mathcal{F}\mathbf{n}; \tilde{\varphi} \rangle_{(W^{1,\infty}(\partial\Omega)^2)' \times (W^{1,\infty}(\partial\Omega))^2}$ por la expresión:

$$\langle \mathcal{F}\mathbf{n}; \tilde{\varphi} \rangle = \langle T\mathbf{n}; (\tilde{\varphi}, \tilde{\psi}) \rangle. \quad (41)$$

Las expresiones (40) y (41) nos permitirán encontrar las condiciones de contorno que aparecen en (34).

4.2 Resultados de regularidad y unicidad para el modelo 2D.

Nos centramos ahora en el modelo obtenido por paso al límite en la subsección anterior para el caso de un dominio 2D, definido como:

$$\Omega = \{(x, z) \in \mathbb{R}^2 / x \in S, -h(x) < z < 0\},$$

siendo S un intervalo abierto y $h : \bar{S} \rightarrow \mathbb{R}_+$ una función continua no negativa que se anula sobre ∂S . La frontera del dominio es $\partial\Omega = \bar{\Gamma}_b \cup \Gamma_s$ donde:

$$\Gamma_b = \{(x, z) \in \mathbb{R}^2 : x \in S, z = -h(x)\}, \quad \Gamma_s = \{(x, 0), x \in S\}.$$

El método que seguimos sigue siendo válido si consideramos paredes laterales.

Estudiamos entonces la velocidad del fluido (v, w) y la presión p que satisfacen el siguiente sistema:

$$(EP) \begin{cases} \partial_t v + v \partial_x v + w \partial_z v - \nu_h \partial_x^2 v - \nu_v \partial_z^2 v + \partial_x p = f & \text{en } (0, T) \times \Omega, \\ \partial_z p = 0 & \text{en } (0, T) \times \Omega, \quad \langle v \rangle = 0 & \text{en } (0, T) \times S, \\ \nu_v \partial_z v|_{\Gamma_s} = \alpha |v_{air}| (v_{air} - v), \quad \nu_v \partial_z v|_{\Gamma_b} = \beta(x)v & \text{en } (0, T), \\ v|_{t=0} = v_0 & \text{en } \Omega \end{cases}$$

donde

$$w(t, x, z) = \int_z^0 \partial_x v(t, x, \xi) d\xi \quad \text{y} \quad \langle v \rangle(t; x) = \int_{-h(x)}^0 v(t; x, z) dz.$$

Nota 11 Denotamos por $f : (0, T) \times \Omega \rightarrow \mathbb{R}$ una fuerza externa, $v_{air} : (0, T) \times S \rightarrow \mathbb{R}$ la velocidad horizontal del aire sobre la superficie y $v_0 : \Omega \rightarrow \mathbb{R}$ la velocidad horizontal inicial. Finalmente, $\alpha \in \mathbb{R}$ es una constante positiva y $\beta = \beta(x)$ una función positiva definida en S (que depende de la rugosidad del fondo). Lógicamente, hemos considerado una viscosidad anisótropa (ν_h, ν_v) .

Nota 12 *El modelo de Ecuaciones Primitivas con condiciones de Navier deducido anteriormente, estaba formado por $(EP)_1$, $\partial_z p = 0$ y $\partial_x v + \partial_z w = 0$ en $(0, T) \times \Omega$, y las condiciones de contorno $\nu_v \partial_z v = \alpha(v_{air} - v)$, $w = 0$ sobre $(0, T) \times \Gamma_s$, $\nu_v \partial_z v = \beta v$ y $(v, w) \cdot \mathbf{n} = 0$ sobre $(0, T) \times \Gamma_b$, y $v|_{t=0} = v_0$ en Ω . Así pues, la ecuación $\partial_x v + \partial_z w = 0$ junto con las condiciones de contorno para w permite deducir que $w(t; x, z) = \int_z^0 \partial_x v(t; x, s) ds$ y $\partial_x \langle v \rangle = 0$. Como $\langle v \rangle$ es una función de una variable, la hipótesis $\langle v \rangle = 0$ sobre $(0, T) \times \partial S$ implica que $\langle v \rangle = 0$ sobre $(0, T) \times S$.*

Introducimos los siguientes espacios funcionales:

$$\mathcal{V} = \{ \varphi \in C_s^\infty(\overline{\Omega}) : \langle \varphi \rangle = 0 \text{ en } S \},$$

donde $C_s^\infty(\overline{\Omega})$ es el espacio de funciones regulares sobre $\overline{\Omega}$ que se anulan en un entorno de ∂S ; H y V las clausuras de \mathcal{V} para las normas de $L^2(\Omega)$ y $H^1(\Omega)$, respectivamente.

En esta subsección usaremos la regularidad débil, que ya ha sido demostrada por un procedimiento asintótico (a partir de las ecuaciones de Navier-Stokes con condición de Navier) en la subsección anterior, y demostraremos una regularidad intermedia entre la débil y la fuerte que nos permitirá concluir en particular la unicidad de solución débil. Éste ha sido el trabajo hecho en [4] y desarrollado en un capítulo de [15]. Concretamente las definiciones son las siguientes:

Definición 4 (Solución débil) *Decimos que v es una solución débil de (EP) en $(0, T)$ si:*

$$v \in L^\infty(0, T; H) \cap L^2(0, T; V) \quad y \quad w(x, z) = \int_z^0 \partial_x v(x, s) ds$$

satisfacen la formulación variacional: $\forall \varphi \in C^1([0, T]; \mathcal{V})$ con $\varphi(T) = 0$,

$$\begin{aligned} & - \int_0^T \int_\Omega (\partial_t \varphi + v \partial_x \varphi + w \partial_z \varphi) v d\Omega dt + \int_0^T \int_\Omega (\nu_h \partial_x v \partial_x \varphi + \nu_v \partial_z v \partial_z \varphi) d\Omega dt \\ & + \int_0^T \int_S \beta(x) \left(1 + \frac{\nu_h}{\nu_v} |h'(x)|^2 \right) v |_{\Gamma_b} \varphi |_{\Gamma_b} dx dt \\ & + \int_0^T \int_S \alpha |v_{air}| (v |_{\Gamma_s} - v_{air}) \varphi |_{\Gamma_s} dx dt \\ & = \int_\Omega v_0 \varphi(0) d\Omega + \int_0^T \int_\Omega f v d\Omega dt + \nu_h \int_0^T \int_S v |_{\Gamma_b} \partial_x [\varphi |_{\Gamma_b} h'(x)] dx dt, \end{aligned} \tag{42}$$

y la desigualdad de energía:

$$\begin{aligned} & \frac{1}{2} \|v(t)\|_{L^2(\Omega)}^2 + \nu_h \int_0^t \|\partial_x v(s)\|_{L^2(\Omega)}^2 ds + \nu_v \int_0^t \|\partial_z v(s)\|_{L^2(\Omega)}^2 ds \\ & + \int_0^t \int_S \gamma(x) |v|_{\Gamma_b}|^2 dx ds + \frac{1}{2} \int_0^t \int_S \alpha |v_{air}| |v|_{\Gamma_s}|^2 dx ds \\ & \leq \frac{1}{2} \|v_0\|_{L^2(\Omega)}^2 + \int_0^t \int_\Omega f v d\Omega ds + \frac{1}{2} \int_0^t \int_S \alpha |v_{air}|^3 dx ds \end{aligned} \quad (43)$$

con $\gamma(x) = \left\{ \beta(x) \left(1 + \frac{\nu_h}{\nu_v} |h'(x)|^2 \right) - \frac{\nu_h}{2} h''(x) \right\}$. ■

Nota 13 Para asegurar que el sistema es disipativo, necesitaremos imponer que $\gamma(x) \geq 0$. Dicha hipótesis resulta ser la hipótesis límite en el caso 2D de (27) cuando $\delta \rightarrow 0$.

Definición 5 (Solución débil-vorticidad) Diremos que v es una solución débil-vorticidad de (EP) en $(0, T)$ si es una solución débil, que satisface además la regularidad adicional:

$$\partial_z v \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H^1(\Omega)).$$

Nota 14 Esta regularidad adicional para $\partial_z v$ implica que v verifica las condiciones de contorno en el fondo y en la superficie en el sentido de las trazas. Además, $\partial_z v$ se puede identificar con la vorticidad de las Ecuaciones Primitivas, ya que es el límite de la vorticidad del sistema de ecuaciones de Navier-Stokes 2D, de ahí su denominación.

Enunciamos entonces el resultado principal:

Teorema 14 Sea $h \in H^2(S)$ con $|h'| > 0$ sobre ∂S , $\beta \in H_0^1(S)$, $f \in L^2(0, T; L^2(\Omega))$, $\partial_z f \in L^2(0, T; H^{-1}(\Omega))$, $v_{air} \in L^\infty(0, T; H_0^1(S))$, $\partial_t v_{air} \in L^2(0, T; L^1(S))$, $v_0 \in H$, $\partial_z v_0 \in L^2(\Omega)$. Si además $\gamma(x) \geq 0$ en S , y la función profundidad h satisface $|h'(x)|/h(x) \leq c/\text{dist}(x, \partial S)$ en S , entonces existe una única solución débil para (EP) en todo $(0, T)$. Dicha solución es, además, una solución débil-vorticidad.

Esquema de la demostración: La búsqueda de soluciones débil-vorticidad responde al objetivo de obtener unicidad del modelo de (EP). En efecto, aplicando el argumento de unicidad hecho por ejemplo en P. L. Lions [13] al caso de (EP), observamos que una condición necesaria de unicidad es que una de las soluciones débiles verifique que $\partial_z v \in L^4(0, T; L^4(\Omega))$, ver Nota 4. Para conseguir dicho objetivo, buscamos de qué problema es solución $\partial_z v$. Derivando formalmente (EP)₁ respecto de z , obtenemos que $\partial_z v$ satisface en $\mathcal{D}'((0, T) \times \Omega)$:

$$\partial_t(\partial_z v) + v \partial_x(\partial_z v) + w \partial_z(\partial_z v) - \nu_h \partial_x^2(\partial_z v) - \nu_v \partial_z^2(\partial_z v) = \partial_z f.$$

Conocidas v y w , la ecuación anterior es lineal y parabólica para $\partial_z v$, y además la presión ha desaparecido al no depender de z . Por tanto, podemos esperar

regularidad débil para $\partial_z v$. Lo que ocurre es que como las condiciones de contorno para $\partial_z v$ sobre Γ_s y Γ_b dependen de v (y v y p están acopladas en el problema (EP)), en realidad el problema que verifica $\partial_z v$ también depende de la presión. Levantando las condiciones de contorno no homogéneas del sistema verificado por $\partial_z v$, con la función auxiliar:

$$\psi = \nu_v \partial_z v - \phi v - e$$

donde

$$\begin{cases} \phi(t; x, z) = -\alpha \left(1 + \frac{z}{h(x)}\right) |v_{air}(t; x)| - \frac{z}{h(x)} \beta(x) \\ e(t; x, z) = \alpha |v_{air}(t; x)| v_{air}(t; x) \left(1 + \frac{z}{h(x)}\right) \end{cases}$$

son funciones que garantizan que $\psi|_{\partial\Omega} = 0$, llegamos a que ψ verifica el problema:

$$(P) \begin{cases} \partial_t \psi + v \partial_x \psi + w \partial_z \psi - \nu_v \partial_{xx}^2 \psi - \nu_v \partial_{zz}^2 \psi = F & \text{en } (0, T) \times \Omega, \\ \psi = 0 & \text{sobre } (0, T) \times \partial\Omega, \\ \psi|_{t=0} = \nu_v \partial_z v_0 - \phi|_{t=0} v_0 - e|_{t=0} & \text{en } \Omega, \end{cases}$$

donde $F = G(\phi, v, w, e, f) + \phi \partial_x p$, para una cierta función G .

Los problemas a resolver llegados a este punto son dos: por una parte, necesitamos para la presión mayor regularidad de la habitual para obtener una solución débil de (P). En segundo lugar, y una vez obtenida ψ , necesitamos identificar $\psi + \phi v + e$ con $\nu_v \partial_z v$, donde la dificultad está en que $\nu_v \partial_z v$ en principio sólo tiene regularidad $L^2(0, T; L^2(\Omega))$.

El siguiente lema garantiza una cierta regularidad con peso para la presión que a posteriori será suficiente para obtener una solución débil del problema (P).

Lema 15 *En las hipótesis del Teorema 14, si (v, w) es una solución débil de (EP), entonces se verifica:*

$$\sqrt{h} \partial_x p \in L^2(0, T; H^{-1}(S)).$$

Nota 15 *Aquí estamos identificando la presión p como una distribución dependiente sólo de x , usando que $\partial_z p = 0$ (ver [4] y [15], para más detalles sobre esta identificación).*

Esquema de la demostración: En el marco de Navier-Stokes, la regularidad de la presión se obtiene a partir de la regularidad del resto de los términos de la ecuación de momentos, en el que el término de evolución $\partial_t v$ implica sólo una regularidad dual de tipo H^{-1} en tiempo. En este caso, aprovechando que $\langle v \rangle = 0$, tenemos que $\partial_t \langle v \rangle = 0$, lo que permite vislumbrar mejoras en la regularidad de la presión si integramos previamente la ecuación en altura.

Para considerar el término de presión, usaremos una formulación variacional mixta, para funciones test $\varphi \in C^1([0, T]; C_s^\infty(\bar{\Omega}))$ con $\varphi(T) = 0$, tales que existe una función ψ suficientemente regular verificando $(\varphi, \psi) \cdot \mathbf{n}|_{\partial\Omega} = 0$, de la forma:

$$\begin{aligned} & - \int_0^T \int_{\Omega} (\partial_t \varphi + v \partial_x \varphi + w \partial_z \varphi) v + \int_0^T \int_{\Omega} (\nu_h \partial_x v \partial_x \varphi + \nu_v \partial_z v \partial_z \varphi) \\ & + \int_0^T \int_S \delta(x) v |_{\Gamma_b} \varphi |_{\Gamma_b} + \int_0^T \int_S \alpha |v_{air}| (v |_{\Gamma_s} - v_{air}) \varphi |_{\Gamma_s} \\ & = \int_{\Omega} v_0 \varphi(0) + \int_0^T \int_{\Omega} f \varphi + \nu_h \int_0^T \int_S v |_{\Gamma_s} \partial_x (\varphi |_{\Gamma_b} h') + \int_0^T \int_{\Omega} p \nabla \cdot (\varphi, \psi). \end{aligned} \tag{44}$$

Centrándonos en la integral en presión, y como p es independiente de z , $\int_{\Omega} p \nabla \cdot (\varphi, \psi) d\Omega = \int_S p \partial_x \langle \varphi \rangle dx$. En consecuencia, al tomar funciones test φ independientes de z se genera un factor $h(x)$. En particular, eligiendo $\varphi = \zeta / \sqrt{h}$ con $\zeta \in C_0^1([0, T]; C_0^\infty(S))$ como funciones test (que son densas en $L^2(0, T; H_0^1(S))$), obtenemos para el término en presión:

$$\int_0^T \int_S p \partial_x (\sqrt{h} \zeta) dx dt.$$

Usando desigualdades de tipo Hardy (vertical) y las hipótesis sobre la función profundidad h , conseguimos (en [4]) acotar el resto de los términos de la formulación variacional mixta por $C \|\zeta\|_{L^2(0, T; H_0^1(S))}$, y concluimos la regularidad del enunciado. ■

Con la regularidad adicional para $\partial_x p$ del Lema anterior, no resulta difícil obtener la existencia (y unicidad, dado que es un problema lineal) de solución débil ψ de (P) .

Nos queda entonces identificar $\psi + \phi v + e$ con $\nu_v \partial_z v$. Como ya dijimos, la dificultad principal es que $\partial_z v$ sólo tiene la regularidad $L^2(0, T; L^2(\Omega))$; aunque se puede demostrar que $\nu_v \partial_z v$ verifica (P) pero en el sentido de solución por trasposición. Si intentamos demostrar unicidad entre una solución débil y otra por trasposición del problema (P) , tendríamos que demostrar mayor regularidad de dicho sistema (lo cual no parece fácil dada la regularidad de la velocidad de convección v). Lo que hacemos entonces es comparar v con una función adecuada \tilde{v} tal que $\nu_v \partial_z \tilde{v} = \psi + \phi v + e$. Para ello, llamamos $a = \psi + \phi v + e$ y definimos la función:

$$\tilde{v}(x, z) = -\frac{1}{\nu_v} \int_z^0 a(x, s) ds + \frac{1}{\nu_v} \frac{1}{h(x)} \left(\int_{-h(x)}^0 \left(\int_z^0 a(x, s) ds \right) dz \right),$$

que se puede ver que verifica $\tilde{v} \in L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega))$, $\nu_v \partial_z \tilde{v} = a$ en Ω , $\langle \tilde{v} \rangle = 0$ sobre S , y las mismas condiciones de contorno que v . Concretamente, \tilde{v} es solución de la e. d. p. :

$$\partial_t \tilde{v} + v \partial_x \tilde{v} + w \partial_z \tilde{v} - \nu_h \partial_{xx}^2 \tilde{v} - \nu_v \partial_{zz}^2 \tilde{v} + \partial_x \tilde{p}_s = G, \tag{45}$$

donde $G = v \partial_x \tilde{v} + \int_z^0 \partial_x (v \partial_z \tilde{v})(x, s) ds + f$. Es importante señalar que si $\tilde{v} = v$, entonces $G = f$. Aplicando ahora un razonamiento de unicidad y gracias a la regularidad adicional para $\partial_z \tilde{v}$ (por ser solución débil-vorticidad), se puede concluir (ver [4, 15]) que $v = \tilde{v}$. ■

Referencias

- [1] P. Azérad & F. Guillén, *Mathematical justification of the hydrostatic approximation in the Primitive Equations of Geophysical fluid dynamics*. Siam J. Math. Anal. ,Vol. 33, No. 4, 847-859.
- [2] O. Besson & M. R. Laydi, *Some Estimates for the Anisotropic Navier-Stokes Equations and for the Hydrostatic Approximation*, M²AN-Mod. Math. Ana.Nume., Vol. 7, 1992, pp. 855-865.
- [3] D. Bresch, F. Guillén-González, N. Masmoudi & M. A. Rodríguez-Bellido, *Asymptotic derivation of a Navier condition for the Primitive Equations*. Sometido a la revista *Asymptotic Anal.*
- [4] D. Bresch, F. Guillén-González, N. Masmoudi and M. A. Rodríguez-Bellido, *On the uniqueness of weak solutions of the two-dimensional Primitive Equations*. Aceptado en la revista *Diff. and Int. Eq.*
- [5] T. Chacón & F. Guillén, *An intrinsic analysis of existence of solutions for the hydrostatic approximation of the Navier-Stokes equations*, C. R. Acad. Sci. Paris, t. 330, Série I, pp. 841-846, 2000.
- [6] F. Guillén & M. V. Redondo, *Convergencia de algunos esquemas numéricos hacia el modelo evolutivo de Ecuaciones Primitivas*, Actas XVI CEDYA, VI CMA, University of Las Palmas de Gran Canaria 1999, pp. 1165-1172.
- [7] F. Guillén-González & M. A. Rodríguez-Bellido, *On the strong solutions of the 2D Primitive Equations problem*. Aparecerá en la revista *Journal of Nonlinear Analysis: Serie A Theory and Methods*, Vo. 50, 5 (2002), 621-646.
- [8] F. Guillén-González, N. Masmoudi & M. A. Rodríguez-Bellido, *Anisotropic estimates and strong solutions of the Primitive Equations*. *Diff. and Int. Eq.* , 14 (2001), no. 11, 1381-1408.
- [9] R. Lewandowski, *Analyse Mathématique et Océanographie*, Masson, 1997.
- [10] J. L. Lions, R. Teman, S. Wang, *New formulation of the primitive equations of the atmosphere and applications*. Nonlinearity, 5, 1992, pp. 237-288.
- [11] J. L. Lions, R. Teman, S. Wang, *On the equations of the large scale Ocean*. Nonlinearity, 5, 1992, pp. 1007-1053.
- [12] J.-L. Lions, R. Temam & S. Wang *Models of the Coupled Atmosphere and Ocean*, Computational Mechanics Advances, 1,(1993), pp. 5-54 and 55-119.

- [13] P.L. Lions, *Mathematical Topics in Fluids Mechanics*, Vol.1, Incompressible Models, Univ. Paris-Dauphine & École Polytechnique, Oxford Univ. Press. Inc., NY, 1996.
- [14] J. Pedlosky, *Geophysical fluid dynamics*, Springer-Verlag, (1987).
- [15] M. A. Rodríguez Bellido, *Análisis Matemático de Algunos Sistemas de Tipo Navier-Stokes: Fluidos Quasi-Newtonianos y Ecuaciones Primitivas del Océano*, Tesis Doctoral, Facultad de Matemáticas, Univ. Sevilla, Septiembre 2001.
- [16] J. Simon *Compact sets in the space $L^p(0, T; B)$* , Annali di Matematica Pure ed Applicata (IV), Vol. CXLVI (1987), pp. 65-96.
- [17] M. Ziane, *Regularity Results for Stokes Type Systems*. Applicable Analysis, Vol. 58, 1995, pp. 263-292.

Predicciones del Tiempo y Matemáticas

M. LEZAUN

Departamento de Matemática Aplicada, Estadística e I. O.
Universidad del País Vasco-Euskal Herriko Unibertsitatea

mpleitm@lg.ehu.es

Resumen

Desde la más remota antigüedad, el hombre ha intentado predecir el tiempo atmosférico, para tratar de librarse de sus efectos perniciosos o aprovecharse de sus beneficios.

En este artículo, dedicado a las predicciones numéricas del tiempo, se hace una introducción histórica de la meteorología, se definen las ecuaciones primitivas del movimiento de la atmósfera, se muestra la forma de resolver numéricamente esas ecuaciones, se presentan modelos de predicción numérica operativos y se concluye analizando la cuestión de ¿hasta qué plazo es posible predecir el tiempo?

1 Notas históricas. Introducción

Recopilar observaciones del tiempo, explicar el comportamiento de la atmósfera y pronosticar el viento y la lluvia son prácticas muy antiguas. Así, hasta el siglo XX, las personas interesadas de forma científica en el tiempo atmosférico realizaban tres actividades: una actividad empírica consistente en recopilar datos de observaciones y a partir de ellos intentar inferir algo, una actividad teórica dedicada a explicar los fenómenos atmosféricos basándose en leyes generales, y una actividad práctica de predicción del tiempo. Naturalmente, estas actividades siempre han estado relacionadas entre sí y el término meteorología se ha utilizado para las tres.

En el siglo XIX, al crecer el número de personas dedicadas a la meteorología, las actividades empíricas, teóricas y predictivas se fueron diferenciando. Muchas de las personas que trabajaban en la tradición empírica hicieron del tiempo promedio su principal preocupación y a mediados del siglo dieron origen, como una ciencia descriptiva, a la climatología. Muchos de los que trabajaban en la tradición teórica hicieron de las leyes de la física su punto de partida y dieron origen a la rama de la ciencia que se denominó meteorología dinámica. Por último, con el inicio en los años 1870 de las predicciones diarias por los servicios meteorológicos, la predicción del tiempo se convirtió en una profesión.

Estas tres tradiciones siguieron su camino diferenciado hasta mediados del siglo XX, época en la que la conexión entre ellas se hizo cada vez más estrecha y numerosa, y en la que los meteorólogos comenzaron a hablar de una unificación de la meteorología. Esta unificación, que culminó en los años 50 y 60 del siglo pasado, estuvo íntimamente ligada al desarrollo de las computadoras electrónicas y es a partir de aquí donde el concurso de las matemáticas es esencial.

Hagamos una breve exposición del desarrollo de cada una de estas tradiciones hasta comienzos del siglo XX, para así ver cual era el escenario previo a los comienzos de la predicción numérica del tiempo.

1.1 Una tradición empírica. Climatología



E. Torricelli

Aunque la moderna tradición en meteorología empírica se puede remontar hasta William Merle, rector de Driby, que anotó el tiempo diario en Oxford desde 1337 hasta 1344, las observaciones meteorológicas no se realizaron de forma sistemática hasta el siglo XVII con la invención del termómetro y del barómetro. En ese siglo tuvo lugar un cambio profundo en las descripciones del tiempo, que pasaron de tener un carácter meramente cualitativo a tener un carácter cuantitativo. Esto fue debido a que en el siglo XVII la temperatura, la humedad, la presión atmosférica, la cantidad de precipitaciones y la dirección y fuerza del viento se pudieron medir. Hay que recordar que se atribuye a Galileo Galilei (1564-1642) la construcción del primer termómetro en los últimos años del siglo XVI, que Evangelista Torricelli (1608-

1647) construyó el primer barómetro en 1643, que en esa centuria se inventaron aparatos para medir las precipitaciones, la dirección y la fuerza del viento, y que aunque los higrómetros ya estaban inventados raramente se habían utilizado antes del siglo XVII. Naturalmente, todos estos nuevos aparatos produjeron un considerable aumento de los datos meteorológicos, lo que a su vez propició el planteamiento de nuevas cuestiones teóricas.

Una vez que se fue capaz de obtener valores numéricos de las distintas variables atmosféricas, para su utilización universal hubo que establecer escalas, equivalencias, protocolos, ... Se inició así un largo proceso de estandarización que culminó con el establecimiento de acuerdos internacionales sobre los instrumentos a utilizar, sobre su calibración, los procedimientos de lectura y sobre la forma de registrar y transmitir los datos. Todo esto constituyó un aspecto importante de la segunda transformación de la ciencia, una transformación organizativa, que tuvo lugar en la segunda mitad del siglo

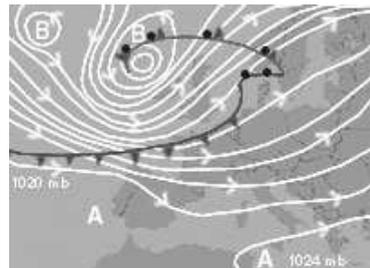
XIX con la creación de las distintas sociedades meteorológicas estatales europeas y la colaboración internacional entre meteorólogos.

A comienzos del siglo XIX mucha gente se interesó en la recopilación sistemática de todo tipo de información sobre un país y sus habitantes. Este interés está en el origen de la etimología del término estadística, que inicialmente designaba los estudios que trataban con los datos numéricos de un estado. Un tipo de información buscada fue los datos meteorológicos. Así, en el transcurso del siglo XIX se recogieron muchísimos datos meteorológicos por gente con distintos intereses, como por ejemplo la propagación de enfermedades, la mejora de la agricultura, el comercio marítimo, etc. Hay que indicar que a partir de la creación a finales del siglo XIX de los servicios meteorológicos estatales en casi todos los países desarrollados, estas instituciones fueron las encargadas de recopilar todos los datos sobre el tiempo atmosférico.

La gran cantidad de datos meteorológicos recopilados dio un continuo impulso a la climatología que, en la segunda mitad del siglo XIX, en gran parte devino una ciencia estadística. Como ciencia se establecieron numerosas leyes climáticas, muchas de ellas extraídas directamente del estudio de los datos. De hecho, se puede entender por clima la totalidad de los fenómenos meteorológicos que caracterizan el *estado medio* de la atmósfera y la probabilidad de la ocurrencia de sus valores extremos, en un lugar y estación anual determinada.

Como resultado de este crecimiento acelerado de los datos, algunos meteorólogos se preguntaron si éstos estaban siendo bien aprovechados e, incluso, si tal abundancia de datos no era un estorbo para la reflexión y el estudio teórico. Estas cuestiones impulsaron a los meteorólogos a idear métodos para descubrir o imponer orden. Para ello una primera opción fue tabular las observaciones del tiempo junto con la de otros fenómenos susceptibles de estar relacionados con él, y buscar patrones. En concreto se esperaba hallar relaciones con la astronomía, entre el movimiento celeste y el tiempo meteorológico, pero no se obtuvo ninguna significativa. También se vio que era muy difícil encontrar algún tipo de periodicidad o regularidad en las variaciones meteorológicas.

Otra forma de tratar los datos consistió en traducirlos a dibujos en los mapas. Así, en 1817 Alexander von Humboldt (1769-1859) introdujo una forma de representar la distribución de calor sobre la superficie terrestre: sobre un mapa de parte del hemisferio norte dibujó las líneas que unían puntos con la misma temperatura media. A estas líneas las denominó isotermas. Posteriormente se dibujaron mapas con líneas isotermas mensuales, con isobaras, con isolíneas de la precipitación anual, etc.



De todas estas representaciones en mapas se extrajeron muchos descubrimientos meteorológicos. Por ejemplo, de importancia capital fue el descubrimiento de que en las zonas templadas del hemisferio norte, las zonas de baja presión, que a menudo son regiones

de precipitaciones, se van desplazando regularmente del oeste hacia el este. De hecho, esta predictibilidad del movimiento de las borrascas, junto con la capacidad de comunicación rápida proporcionada por el telégrafo para construir mapas del tiempo diario, fueron los factores más importantes para la creación de los servicios meteorológicos nacionales en los años 1860 y 1870.

1.2 Una tradición teórica. Física de la atmósfera

La obra de Aristóteles (384-322 a.C.) *Meteorologica*, escrita alrededor de 340 a.C., estableció una tradición teórica en la meteorología y aseguró que ésta fuera estudiada como una parte de la filosofía natural. Este tratado fue la base de todos los estudios teóricos de meteorología hasta comienzos del siglo XVII. En ese siglo la meteorología teórica experimentó grandes cambios: la publicación en 1637 de la obra de René Descartes (1596-1650) *Discours de la méthode*, que contiene un apéndice titulado *Les météores*, estimuló la aparición de nuevas ideas sobre distintos fenómenos atmosféricos; las nuevas observaciones, muchas de ellas obtenidas gracias a la invención del barómetro y del termómetro, demandaron nuevas explicaciones; y el desarrollo de la ciencia de la mecánica propició la aparición de nuevas teorías sobre algunos fenómenos atmosféricos, como la de los vientos alisios de Edmond Halley (1656-1742) de 1686.

Hasta mediados del siglo XIX, a pesar de los datos proporcionados por el termómetro y el barómetro, y a pesar de la relevancia de las matemáticas en la mecánica a partir de la obra de Isaac Newton (1642-1727), la mayor parte de las teorías meteorológicas seguían siendo totalmente cualitativas. Sólo unas pocas cuestiones eran tratadas matemáticamente, por ejemplo la relación entre la altitud y la presión atmosférica. A finales de los años 1850, siguiendo el camino iniciado entre otros por William Ferrel (1817-1891), que dio una explicación teórica a que la dirección del viento generalmente es paralela a las isobaras locales, lo que se denominó como meteorología dinámica tuvo muchos cultivadores, entre los que destacaremos a Hermann von Helmholtz (1821-1894). Así, en la segunda mitad del siglo XIX se explicaron muchos fenómenos atmosféricos, lo cual a su vez produjo un gran desarrollo de esos estudios teóricos.

Los meteorólogos teóricos defendían que la meteorología debería ser física aplicada y que los datos observables deberían explicarse de forma deductiva. Ellos mismos se consideraban como pioneros de una nueva meteorología, en contraposición con los empíricos que afirmaban que la meteorología era una ciencia independiente cuyas leyes se tenían que inducir directamente de los datos.

Básicamente, los dinamicistas procedían de dos formas. Una práctica común consistía en tomar una teoría física y hacerla aplicable a las condiciones de la atmósfera terrestre. Así, se utilizaron las teorías sobre el comportamiento del aire saturado con vapor de agua para dar una teoría sobre la formación de nubes. Otra forma de proceder consistía en considerar un fenómeno meteorológico observado y ver cómo podía ser explicado de acuerdo a las leyes de la física. Como ejemplo diremos que Julius Hann (1839-1921) explicó en 1866 el viento

Föhn de los Alpes utilizando la termodinámica.

Hay que resaltar que las distintas actividades teóricas y empíricas no dieron como resultado dos comunidades de investigadores completamente separadas, pues los físicos aplicados no se oponían a descubrimientos de regularidades por métodos inductivos, ni los empíricos objetaban la aplicación de las leyes de la física a la atmósfera. Para muchos de ellos la diferencia sólo era una cuestión de conveniencia: ¿qué es más fructífero, trabajar a partir de la física o desde los datos? También hay que decir que en ningún momento hubo un total divorcio entre los resultados teóricos y los datos experimentales. Naturalmente, para mostrar que la teoría y las observaciones estaban relacionadas, que cuantitativamente eran iguales, había que hacer cálculos. Estos cálculos, entendidos en su sentido más amplio, adquirieron gran importancia ya que proporcionaron soporte a algunas teorías y, lo que fue más importante, sirvieron para refutar otras.

1.3 Una tradición práctica. Predicción del tiempo

¿Qué tiempo va a hacer? Ésta es una pregunta que siempre se ha hecho el hombre y que nunca pierde interés. Durante siglos los pronósticos se apoyaban en signos naturales, por ejemplo un pequeño halo alrededor del sol como presagio de lluvias, que en muchos casos se expresaban en forma de refranes. Citaremos como muestra

*Cielo empedrado, suelo mojado.
Cielo de lanas, si no llueve hoy, lloverá mañana.
Arco iris al mediodía, llueve todo el día.*

En todos los países perviven muchos refranes de este tipo y, aunque a veces sean contradictorios o tengan una interpretación ambigua, hay que reconocerles una cierta validez local para las previsiones a muy corto plazo. Hay que indicar que todavía hoy estos refranes constituyen verdaderas guías para los labradores y la población rural.

Se puede considerar a los astrólogos como los primeros hombres del tiempo profesionales. En la cultura occidental la astro-meteorología se remonta al menos hasta el *Tetrabiblos* de Claudio Ptolomeo (90-168), escrito hacia el año 160 de la era cristiana, que fue la principal autoridad para esta práctica durante toda la Edad Media. La invención de la imprenta en el siglo XV dio una gran difusión a la predicción astrológica del tiempo, debido principalmente a la inclusión de esas predicciones en almanaques. Así, hasta finales del siglo XVIII se publicaron muchos libros de astro-meteorología, siendo uno de los más populares *Della vera influenza degli astri sulle stagioni e mutazioni di tempo* de Giuseppe Toaldo (1719-1797), publicado en 1770. Esta práctica casi desapareció en el siglo XIX, a pesar de que en esa época las predicciones del tiempo eran más populares que nunca. De hecho era muy raro el diario escrito que no presentara en un lugar importante las previsiones del tiempo y escaseaban los lectores que no prestaran atención a esas predicciones. Esta gran popularidad fue generada por una nueva técnica de predicción del tiempo, que se denominó método sinóptico.

La hipótesis de partida de ese nuevo método consistía en admitir que

el conocimiento del tiempo actual en una amplia zona geográfica puede proporcionar una estimación del tiempo venidero en lugares de esa región. Para ponerlo en práctica había que elaborar mapas del tiempo diarios, y esto fue factible gracias a la invención del telégrafo. De hecho la capacidad proporcionada por el telégrafo para las comunicaciones a larga distancia hizo posible que en 1863, la Sociedad Meteorológica de Francia comenzara a confeccionar mapas del tiempo diarios, iniciativa que rápidamente fue imitada por los demás servicios meteorológicos nacionales. Hay que resaltar que el método sinóptico enseguida desplazó a los métodos locales y que hasta hace poco era la base de casi todos los pronósticos del tiempo.

Los mapas del tiempo pusieron en evidencia el papel de la presión atmosférica: parecía claro que las variaciones del tiempo dependían de las del barómetro. Así, se observó que en general el tiempo es especialmente frío en las regiones de alta presión y muy lluvioso en las de baja presión. En 1860, C.H.D. Buys-Ballot (1817-1890) descubrió la ley que permite determinar la velocidad y dirección del viento a partir del conocimiento de la distribución de la presión. Los dibujos en mapas de las trayectorias de las borrascas mostraron su carácter migratorio. En fin, los mapas del tiempo confirmaron que en la zona templada norte las condiciones atmosféricas en general se mueven del oeste hacia el este.

A finales del siglo XIX la predicción del tiempo se hacía como sigue. Cada día los pronosticadores construían mapas sinópticos con los valores de las distintas variables atmosféricas (presión, temperatura, precipitaciones, humedad, capa de nubes, etc.) recogidos a una misma hora en cien o más localidades y que les eran enviados por telégrafo. El más importante era un mapa en el que se habían dibujado las líneas isobaras y que mostraba la distribución de las presiones barométricas recogidas. Hay que resaltar que los pronosticadores prestaban una especial atención a los patrones de isobaras, pues determinados ordenamientos de las mismas se asociaban a tipos de tiempo particulares.

Una vez realizado el “retrato” del tiempo actual, la principal tarea de los pronosticadores consistía en hacer un mapa pronóstico que, la mayor parte de las veces, no era más que una estimación personal de lo que podía ser el mapa del tiempo del día siguiente. Para ello los hombres del tiempo seguían unas reglas obtenidas de su propia experiencia y de la de sus predecesores, las cuales les permitían estimar y pronosticar cómo se iba a desarrollar y mover cada estructura atmosférica. Luego, de ese mapa inferían las condiciones previstas del tiempo en los lugares considerados. Todo ese trabajo desembocaba en una predicción que era una simple descripción verbal del tiempo venidero, la mayor parte de las veces no más precisa que “lluvioso y ventoso” o “despejado y frío”. Hay que decir que las reglas de predicción son al sinóptico lo que las leyes físicas son al dinámico. Así, si se pudieran formular de manera que en cualquier situación imaginable dieran una única predicción, definirían un modelo matemático alternativo al obtenido a partir de las leyes generales de la física.

En todo ese proceso predictivo los considerables avances conseguidos por los meteorólogos teóricos no tenían prácticamente ninguna participación. El método sinóptico, con su exclusiva dependencia de informaciones expresadas en mapas, no requería de una comprensión teórica de la atmósfera. También hay

que decir que raramente se utilizaba de forma explícita la gran acumulación de datos de observaciones anteriores, que era la base de la climatología, pues aunque el método dependía de los datos, de hecho sólo se utilizaban los estrictamente necesarios para la confección de los últimos mapas del tiempo. Por todo ello se puede afirmar que en una visión panorámica las actividades de los pronosticadores, la de los teóricos y la de los climatólogos aparecerían bastante distanciadas entre sí.

Con la tradición predictiva del tiempo siempre ha convivido una tradición escéptica restrictiva. Así, a comienzos del siglo XIX el astrónomo francés François Arago (1786-1853) declaró que nadie que tuviera un prestigio científico que perder se debería arriesgar a profetizar el tiempo, y cuando a finales del siglo los estados comenzaron a subvencionar las predicciones los escépticos se hicieron más beligerantes. De hecho, la utilización de mapas para la predicción del tiempo se consideraba en muchos círculos científicos como algo sospechoso que, con ánimo de desprestigiarla, lo tildaban de “empírico”.

Tanto los defensores como los detractores de las predicciones intentaron ganar la batalla midiendo lo acertado de los pronósticos. La vaguedad de las previsiones, que salvo algunas excepciones hasta la primera guerra mundial era no cuantitativa, favorecía a los defensores. Como ejemplo de los detractores citaremos a Arnulph Mallock, que escribió un artículo en *Nature* en 1914 en el que comparaba la predicción diaria para Londres de todo el año 1913 con el tiempo real. Como conclusión afirmaba que alguien que siguiera una regla del tipo “el tiempo de mañana será como el de hoy” tendría el mismo nivel de aciertos que el del Servicio Meteorológico.

En cualquier caso al final vencieron los defensores de la predicción del tiempo, que siempre contaron con el abrumador apoyo del público general deseoso de ver en los periódicos alguna información sobre el tiempo más probable. Hay que indicar que a partir de la aparición de los métodos de previsión numérica a mediados del siglo XX, los escépticos fueron cada vez menos numerosos y menos ruidosos, lo cual permite afirmar que las predicciones del tiempo habían comenzado con 100 años de anticipación, en 1860 en vez de en 1960.

Conviene remarcar que las predicciones del tiempo atmosférico gozaron de poco prestigio entre los científicos. En parte eso era debido al modesto éxito de las previsiones, pero mucho más relevante era su percepción de que las predicciones no eran sistemáticas, de que su elaboración no estaba basada en conocimientos científicos, de que era más un arte que una ciencia. Hay que decir que esa percepción estaba bastante cercana de la realidad.

Para terminar diremos que en cada una de las tradiciones de la meteorología había que hacer cálculos. Ahora bien, aunque indudablemente los cálculos tenían su importancia, hay que reconocer que sólo jugaban un papel secundario. Sin embargo, se puede afirmar que fue el uso intensivo del cálculo el que condujo a mediados del siglo XX la unificación de las tres tradiciones.

1.4 La unificación de la meteorología

A finales del siglo XIX y comienzos del XX la meteorología había adquirido identidad propia. Su estatus como ciencia empírica se debía principalmente a los climatólogos, su estatus como ciencia teórica se había alcanzado aplicando la física a los fenómenos atmosféricos, y los gobiernos proporcionaban una amplia base institucional a la capacidad de los meteorólogos para predecir el tiempo.

En 1903, el noruego Vilhelm Bjerknes (1862-1951) propuso tratar la evolución de la atmósfera según las leyes de la termodinámica y de la mecánica de fluidos, y defendió que la predicción del tiempo fundamentalmente es un problema determinista de valores iniciales en el sentido matemático del término: *“Si, como piensa todo hombre que razona científicamente, los fenómenos atmosféricos se desarrollan a partir de los que le preceden siguiendo unas leyes físicas precisas, se puede afirmar que las condiciones necesarias y suficientes para una solución racional de la predicción en meteorología son:*

- *se debe conocer con una precisión suficiente el estado de la atmósfera en un instante dado;*
- *se deben conocer con una precisión suficiente las leyes según las cuales se desarrolla un estado de la atmósfera a partir del estado precedente.”*

El programa de Bjerknes atrajo la atención y el aplauso de los meteorólogos pero, como él mismo lo reconocía, la dificultad de este problema viene de la necesidad de tener que resolver un sistema de ecuaciones en derivadas parciales no lineales para las que no se dispone de soluciones analíticas.

La primera persona que llevó a cabo el programa de Bjerknes fue el matemático inglés Lewis Fry Richardson (1881-1953). En los años 1910, Richardson ideó un método aritmético (diferencias finitas) para resolver de forma aproximada ecuaciones en derivadas parciales, y se fijó en las ecuaciones de la previsión del tiempo propuestas por Bjerknes para aplicar su método numérico en un problema práctico importante. Para llevarlo a cabo Richardson se volcó en la meteorología, y aprovechó sus viajes por toda Francia como conductor de ambulancias durante la primera guerra mundial para recopilar un amplio



V. Bjerknes

conjunto de datos meteorológicos de un día concreto (20/05/1910). A partir de los datos de ese día necesitó seis semanas para realizar los miles de sumas, restas y multiplicaciones que hacían falta para hacer una predicción a seis horas vista en una pequeña región, obteniendo un resultado completamente insatisfactorio.

Sin desanimarse, Richardson buscó las razones de su fracaso. Sus trabajos se publicaron en 1922 en un libro que se hizo muy célebre, titulado *Weather Prediction by Numerical Process*, y que contiene una visión con carácter premonitorio. Remarcando que “harían falta 64.000 personas trabajando por turnos para prever el estado de la atmósfera con mayor rapidez que la de su evolución real”, Richardson imaginó una fábrica de predicciones meteorológicas formada por miles de calculadores humanos trabajando de forma sincronizada bajo la dirección de un responsable encargado de la buena marcha de las operaciones. Lo que no pudo prever fue el aparato que 25 años más tarde iba hacer el trabajo tan rápidamente como las 64.000 personas.



L.F. Richardson

Al margen de la meteorología, de capital importancia fue el hecho de que en 1928 los matemáticos alemanes Richard Courant (1888-1972), Kurt Friedrichs (1901-1982) y Hans Lewy (1904-1988) estudiaran de forma sistemática la manera de resolver las ecuaciones en derivadas parciales por “diferencias finitas”, y precisaran las condiciones que se deben respetar en la discretización (condiciones de estabilidad).

En el período de entre guerras sobresale la figura del meteorólogo dinamicista sueco Carl-Gustaff Rossby (1898-1957), que fue la persona que más influyó en los pronosticadores para que cambiaran su percepción de que la meteorología dinámica tenía muy poco interés para su trabajo. Una de sus contribuciones más importantes fue el convencimiento de que la clave para entender la atmósfera había que buscarla en el viento, en concreto en la componente vertical de la vorticidad, y no en la presión. En 1939 Rossby dedujo una ecuación cuya solución da la velocidad de propagación de ciertas ondas de longitud larga (denominadas hoy en día ondas de Rossby) presentes en las corrientes del oeste que circundan la tierra en altura en las latitudes medias. En 1940 Rossby propuso efectuar varias aproximaciones en las ecuaciones utilizadas por Richardson y obtuvo la ecuación de balance de la componente vertical de la vorticidad, que permite traducir el comportamiento de una atmósfera promedio.

La primera computadora electrónica, denominada ENIAC (Electronic Numerical Integrator Analyser and Computer), se construyó en la universidad de Pensilvania en 1946 gracias al determinante impulso del matemático americano de origen húngaro John von Neumann (1903-1957). Un objetivo importante para von Neumann consistía en demostrar, con un problema científico particular, el potencial revolucionario de la computadora. Aunque no era meteorólogo, von Neumann reconoció el problema de la predicción meteorológica como ideal para sus necesidades.



J. Charney

En 1950, el americano Jule Charney (1917-1981), el noruego Ragnar Fjørtoft y von Neumann realizaron la primera predicción numérica del tiempo. Para esta experiencia consideraron que lo mejor era utilizar un modelo simplificado que tuviera validez meteorológica, y eligieron el modelo barotrópico propuesto por Rossby. Los cálculos numéricos los realizaron en el ENIAC instalado en Aberdeen (Maryland, USA). Hay que resaltar que necesitaron 33 días con sus noches para programar y ejecutar tres predicciones para un plazo de 24 horas. Los resultados obtenidos para la previsión del movimiento medio de la troposfera fueron muy alentadores y esta experiencia histórica marca el punto de partida de la predicción numérica moderna.

En 1951 el meteorólogo Norman Phillips (1924-), que pertenecía al mismo equipo que los tres anteriores, intentó incluir la estructura vertical de la atmósfera e introdujo un modelo baroclínico formado por dos niveles barotrópicos, incompresibles, homogéneos de diferente densidad.

En los años 1950 y 1960 el ordenador se convirtió en una herramienta estándar en meteorología y se fueron abandonando la mayoría de los restantes medios de cálculo. A comienzos de los años 1960 se hizo evidente que las ecuaciones que habían hecho posible la predicción numérica con los primeros ordenadores no iban a dar predicciones de la calidad que algunos esperaban. Con la llegada de ordenadores más potentes y siguiendo el camino iniciado por Arnt Eliassen (1915-2000) en 1956, a mediados de los años 1960 se volvió a las denominadas ecuaciones primitivas de Richardson. Desde entonces los modelos de predicción numérica se han ido haciendo cada vez más complejos y para su resolución es necesario un considerable aparato matemático que ha sido desarrollado por matemáticos.

A comienzos de los años 1960 se suscitó la cuestión de si los modelos atmosféricos utilizados realmente servirían para la predicción del tiempo o sólo para la comprensión de la dinámica de la atmósfera. Así, a principios de 1961 Edward Lorenz (1917-) se dedicó a simular mediante ordenador el comportamiento de la atmósfera sobre largos períodos de tiempo, y en 1963 publicó su famoso artículo *Deterministic non periodic flows* en el cual introdujo el primer sistema dinámico caótico, que se denominó *sistema de Lorenz*.

2 La predicción numérica del tiempo

La predicción numérica del tiempo se lleva a cabo a partir de un modelo matemático formulado por ecuaciones en derivadas parciales, las cuales traducen

las leyes generales de la física que rigen la atmósfera terrestre.

La atmósfera es un fluido, por lo que las ecuaciones utilizadas serán las ecuaciones generales de la mecánica de fluidos, eso sí, para el caso de una capa aislada de aire (seco o conteniendo vapor de agua) cuyo movimiento se observa desde un sistema no inercial, como lo son todos los que giran con la Tierra. Un análisis del orden de magnitud de los distintos términos de esas primeras ecuaciones permite simplificarlas según sean las escalas de espacio y tiempo de los fenómenos meteorológicos que se consideren.

En lo que respecta a la predicción meteorológica clásica, las escalas espaciales horizontales van de 10.000 km a 10 km (escalas planetaria, sinóptica y de mesoescala) y las temporales de varios días a algunas horas. Las ecuaciones matemáticas obtenidas son no lineales y, en general, sus soluciones no se pueden obtener de forma analítica. Para resolverlas hay que apelar al cálculo numérico, que proporciona una solución aproximada. Como veremos más adelante, la “numerización” básicamente consiste en reemplazar las ecuaciones en variables continuas por ecuaciones en las que las variables son discretas, cuyas soluciones se obtienen mediante un algoritmo apropiado.

Existe una gran variedad de modelos, cada uno con sus ventajas e inconvenientes. Como norma general, cuanto más preciso sea el método más cálculos habrá que hacer y, por tanto, más tiempo se tardará en ejecutarlo. Evidentemente, para una precisión dada se elegirán métodos rápidos, ya que para que una predicción tenga sentido se debe poder hacerla en un plazo de tiempo relativamente corto. Hay que resaltar que la mayor parte de las veces una predicción meteorológica es una verdadera carrera contra el reloj, pues de nada sirve calcular con gran precisión el tiempo que hará mañana si para ese cálculo se necesitan más de veinticuatro horas. Así pues, siempre habrá que hacer un compromiso entre la precisión del cálculo y la duración de su ejecución.

Los resultados de una predicción numérica dependen de las simplificaciones que se hayan hecho para obtener el sistema de ecuaciones matemáticas del modelo y, también, de los efectos de la numerización adoptada. Por eso es muy importante estudiar analíticamente los distintos esquemas numéricos y conocer con precisión los errores que inevitablemente se cometen al utilizar uno u otro.

Quien dice cálculo numérico dice ordenador, ya que ésta es la herramienta necesaria para realizar los cálculos. Esto introduce una nueva degradación pues los cálculos no se hacen de forma exacta sino de forma aproximada, y al trabajar con un número limitado de decimales (8 a 16) se comete un error de redondeo. Sin embargo, la mayor parte de las veces ese error de redondeo introducido por el ordenador es mucho menor que el que se comete al resolver el problema



E. Lorenz

matemático con los métodos numéricos.

Una vez formuladas las ecuaciones matemáticas del modelo, para resolverlas hay que partir del estado inicial de la atmósfera en el instante $t = 0$. Esto significa que en ese instante se debe conocer el valor de todas las variables que caracterizan el estado de la atmósfera. Esto es posible gracias a la creación y al funcionamiento del servicio mundial de observación meteorológica que, con métodos convencionales tanto en superficie como en altitud y con la ayuda de satélites artificiales, obtiene una descripción aceptable de la atmósfera en un instante inicial dado. Ahora bien, los datos iniciales no se deben introducir brutalmente en el modelo (este fue el fallo de Richardson), ya que en un instante fijado las variables atmosféricas están ligadas entre sí pues, al menos teóricamente, también son un resultado de las ecuaciones de evolución. Por tanto, para evitar que se engendren fuertes oscilaciones debidas a la propagación de ondas gravitatorias de amplitud irrealista, se tienen que retocar los valores iniciales recogidos de las variables, de manera que la presión y viento verifiquen un determinado equilibrio denominado equilibrio masa-viento. Para ello se utiliza una técnica denominada *inicialización no lineal por modos normales*.

El diseño de esquemas numéricos más precisos y más rápidos y el aumento de la potencia de los ordenadores han hecho que los errores debidos a las técnicas numéricas y al cálculo electrónico tiendan a disminuir. Ahora bien, queda una fuente de error fundamental que resulta de que las medidas efectuadas a partir de la red de observación no permiten determinar el valor exacto de las variables que definen el estado inicial del modelo. Así, al definirlo siempre va a aparecer un cierto grado de inexactitud.

Debido a la no linealidad de las ecuaciones de evolución, la inevitable introducción de una ligera desviación en el estado inicial se amplifica a medida que evoluciona el modelo, para al cabo de un tiempo dar una solución que se aparta totalmente de la solución de referencia. Esto plantea la siguiente cuestión: ¿en qué momento deja de tener valor la solución obtenida?, o lo que es lo mismo, ¿cuál es el límite de la predictibilidad de la atmósfera? Se ha intentado determinar este límite de dos formas: mediante estudios de la turbulencia y realizando experimentos numéricos paralelos. Hoy en día la opinión unánime es que el límite de predictibilidad de la atmósfera para los problemas de valores iniciales está entre 10 y 15 días.

Para que el problema matemático esté bien propuesto, además de las ecuaciones de evolución y de las condiciones iniciales se necesitan las condiciones de contorno; es decir, es necesario conocer los valores de las variables atmosféricas sobre la frontera del dominio atmosférico D considerado. En el caso en que D sea toda la atmósfera terrestre sólo hay que tener en cuenta sus límites inferior y superior. En el caso en el que esté limitado por una frontera lateral, además habrá que hacer suposiciones más o menos justificadas sobre la evolución de las variables sobre esa frontera. Por ejemplo, en dominios grandes se podrá imponer las mismas condiciones en sus bordes laterales, lo cual equivale a una periodicidad espacial. Naturalmente, la calidad de la predicción en el interior del dominio dependerá de lo acertado de la estimación de la evolución de las variables sobre su frontera lateral. También, cuanto mayor sea el dominio más

tardará la predicción para su interior en ser afectada por los errores cometidos en la frontera.

Los progresos en la predicción numérica del tiempo se han visto muy favorecidos por el desarrollo de ordenadores electrónicos cada vez más potentes. Así, la potencia de los ordenadores utilizados en meteorología ha ido subiendo desde 3 mil operaciones/segundo (IBM 701 instalado en 1955) a 2,5 millones de operaciones/segundo (CDC 6600 del año 1966), a 7 mil millones de operaciones/segundo (Cray C98 del año 1991), hasta los 100 mil millones de operaciones/segundo (Fujitsu VPP 5000 instalado en Météo-France a finales de 1999). Esta creciente potencia de cálculo esencialmente se ha utilizado para aumentar la resolución horizontal y vertical de los modelos. Tampoco hay que olvidar que esos progresos también deben mucho a los esfuerzos de los matemáticos, que proponen esquemas cada vez más precisos y más rápidos, y a los de los físicos, que van mejorando los métodos para describir toda la complejidad de los procesos físicos atmosféricos.

El proceso de predicción del tiempo no acaba con el resultado numérico. ¡Los ordenadores no solucionan todo! En particular, para los plazos más breves (de algunas horas a uno o dos días) es indispensable la destreza del hombre del tiempo encargado de la predicción quien, como gran conocedor del clima regional y de los límites de los modelos, ajusta e incluso modifica los resultados de la simulación y los traduce a términos de tiempo observable: intensidad de las precipitaciones, temperaturas máxima y mínima del día, posible aparición de nieblas, de tormentas, de ráfagas de viento, etc.

Desde un punto de vista práctico la predicción mediante modelos matemáticos abarca un plazo que va desde las 8 ó 10 horas hasta, en el mejor de los casos, los nueve o diez días siguientes, y ello con serias limitaciones dependientes de la situación atmosférica concreta y de la época del año. No son muy raras las situaciones para las que es muy difícil ir más allá de las 60 ó 72 horas. Para períodos de tiempo inferiores a seis u ocho horas, hoy en día los modelos matemáticos no son los adecuados y es preciso utilizar otras técnicas denominadas de predicción inmediata o de predicción a muy corto plazo. Éstas se desarrollan a partir de extrapolaciones más o menos complejas de los datos meteorológicos obtenidos mediante estaciones automáticas de superficie y teledetección (radares, satélites, redes de detección de rayos, ...) con gran resolución espacial y temporal. Ello requiere una vigilancia continua de la evolución atmosférica y una rápida toma de decisiones sobre posibles envíos de avisos rectificando las predicciones.

A modo de conclusión veamos cómo ha mejorado la predicción meteorológica en los últimos años. Así, la calidad de las previsiones que se hacían en 1954 con el modelo barotrópico para un plazo de 24 horas es equivalente a las que se realizaban en 1995 en el Centro Europeo de Predicciones Meteorológicas a Medio Plazo para un plazo de seis días. Sin ir tan lejos en el pasado, las estadísticas de los últimos quince años sobre Europa publicadas por Météo-France muestran que la calidad de la predicción en 1995 para un plazo de 72 horas es equivalente a la que se tenía en 1980 para un plazo de 24 horas.

3 La parte dinámica de las ecuaciones de los modelos

Comenzaremos esta sección escribiendo las ecuaciones que traducen los principios generales de la física para el caso particular de la atmósfera terrestre. Esto nos dará un sistema de siete ecuaciones con siete incógnitas que, como hemos dicho antes, puede dar origen a modelos más sencillos, algunos de los cuales se describirán a continuación. Más adelante se describirá un método numérico de resolución de ese sistema de ecuaciones y luego se tratará brevemente la parte física del modelo.

3.1 Las ecuaciones del movimiento en coordenadas inerciales

En un sistema de referencia inercial la aplicación de la segunda ley de Newton (fuerza igual a masa por aceleración) a una “pequeña” parcela de fluido se escribe ([6]):

$$\frac{\partial v_i}{\partial t} + \sum_{j=1}^3 v_j \frac{\partial v_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} - g\delta_{i3} + f_{ri}, \quad i = 1, 2, 3, \quad (1)$$

donde $\vec{v} = (v_1, v_2, v_3)$ es la velocidad, ρ la densidad y p la presión. El primer miembro de esta ecuación es la derivada material de la velocidad y el segundo representa la suma de las fuerzas que actúan sobre la parcela de fluido. Así, el término $-\frac{1}{\rho} \frac{\partial p}{\partial x_i}$ representa la fuerza producida por las diferencias de presión, $-g\delta_{i3}$ proviene de la fuerza gravitatoria y $\vec{f}_r = (f_{r1}, f_{r2}, f_{r3})$ representa las fuerzas de rozamiento del resto del fluido sobre la superficie de la parcela seleccionada.

La ley de conservación de la masa afirma que la *masa de una parcela de fluido, cuando se sigue en su movimiento, se conserva*. La traducción a lenguaje matemático de esta ley de conservación es la denominada ecuación de continuidad ([6]):

$$\frac{\partial \rho}{\partial t} + \sum_{j=1}^3 v_j \frac{\partial \rho}{\partial x_j} = -\rho \operatorname{div}(\vec{v}). \quad (2)$$

De las leyes termodinámicas descubiertas por Robert Boyle (1627-1691), Jacques Charles (1746-1823) y Joseph Louis Gay-Lussac (1778-1850), se sigue que el aire seco considerado como gas perfecto verifica la siguiente relación entre la densidad (ρ), la presión (p) y la temperatura absoluta (T) ([6]):

$$p = \rho RT. \quad (3)$$

Aquí, R es la denominada constante de los gases para el aire seco, que se calcula mediante la ley de Dalton teniendo en cuenta los diferentes gases que constituyen el aire y sus respectivas proporciones ([11]). Esta relación recibe el nombre de *ecuación de estado de los gases perfectos* que, modificando R o T , puede ser refinada para así tener en cuenta la humedad del aire (*cf.* [11]).

Haciendo un balance de la energía (primer principio de la termodinámica) en una parcela de fluido atmosférico, y teniendo en cuenta la ley de estado anterior, se obtiene la ecuación en derivadas parciales ([12]):

$$C_p \left(\frac{\partial T}{\partial t} + \sum_{j=1}^3 v_j \frac{\partial T}{\partial x_j} \right) = \frac{RT}{p} \left(\frac{\partial p}{\partial t} + \sum_{j=1}^3 v_j \frac{\partial p}{\partial x_j} \right) + Q, \quad (4)$$

donde T es la temperatura absoluta, C_p el calor específico a presión constante del aire seco y Q el aporte de calor por unidad de masa.

Por último, la séptima ecuación es la correspondiente a la concentración de vapor de agua ([5]):

$$\frac{\partial q}{\partial t} + \sum_{j=1}^3 v_j \frac{\partial q}{\partial x_j} = Q', \quad (5)$$

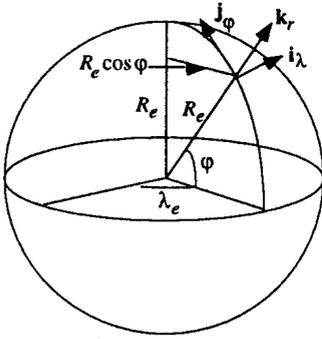
donde q es la humedad específica y Q' el aporte de vapor de agua por unidad de masa.

Los términos \vec{f}_r , Q y Q' representan las fuentes y los sumideros de cantidad de movimiento, de calor y de vapor de agua, respectivamente. Si no se tienen en cuenta estos términos se obtiene un sistema de ecuaciones que describe una atmósfera adiabática y sin rozamiento (hipótesis del sistema aislado), que constituye *la parte dinámica del modelo*. Por contra, *la parte física del modelo* consiste en la determinación precisa de esos términos.

Observemos que si nos restringimos a la parte dinámica del modelo, o lo que es equivalente, si suponemos que \vec{f}_r , Q y Q' son conocidos, el sistema (1)-(5) es completo en el sentido de que hay siete ecuaciones para siete funciones incógnita (la densidad, las tres componentes de la velocidad, la presión, la temperatura y la humedad específica). Como ya se ha dicho antes, junto con las condiciones iniciales y las condiciones de contorno conforman un problema tratable desde el punto de vista matemático.

3.2 Las ecuaciones meteorológicas del movimiento

Las ecuaciones del movimiento de los fluidos que acabamos de describir sólo son válidas para un sistema de coordenadas inercial. Ahora bien, en meteorología se utilizan como referencia sistemas ligados a la superficie de la Tierra. Esto supone dos alteraciones en las ecuaciones, las debidas a la rotación terrestre (fuerza de Coriolis) y las debidas a la forma esférica de la Tierra. Se tiene por tanto que modificar esas ecuaciones para escribirlas en las nuevas coordenadas, que pueden ser coordenadas cartesianas o coordenadas esféricas. En general, se usarán coordenadas cartesianas cuando se considere que la curvatura de la tierra es pequeña (aproximaciones f -plano y β -plano) y coordenadas esféricas cuando no se pueda despreciar esa curvatura. Aquí, para favorecer la claridad en la exposición, nos fijaremos sólo en las coordenadas cartesianas.



Sea (x, y, z) un sistema de coordenadas cartesianas que está fijado a la superficie terrestre (ver la figura). La coordenada x es la distancia sobre el paralelo terrestre (hacia el este x es positiva), la coordenada y es la distancia sobre el meridiano (hacia el norte y es positiva) y la coordenada z es la altura sobre la superficie terrestre.

Designaremos por u la componente del viento en la dirección de x , por v la componente en la dirección de y y por w la componente en la dirección de z .

Si no se tiene en cuenta la curvatura de la superficie terrestre, y si aprovechando que la capa atmosférica es “delgada” aproximamos la distancia de un punto de la atmósfera al centro de la Tierra por el radio medio de la Tierra, las ecuaciones (1) se escriben ([6]):

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - \frac{uv \operatorname{tg} \phi}{R_e} + \frac{uw}{R_e} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + 2\Omega v \operatorname{sen} \phi - 2\Omega w \operatorname{cos} \phi + f_{rx} \quad (6)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + \frac{u^2 \operatorname{tg} \phi}{R_e} + \frac{uw}{R_e} = -\frac{1}{\rho} \frac{\partial p}{\partial y} - 2\Omega u \operatorname{sen} \phi + f_{ry} \quad (7)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} - \frac{u^2 + v^2}{R_e} = -\frac{1}{\rho} \frac{\partial p}{\partial z} - g + 2\Omega u \operatorname{cos} \phi + f_{rz} \quad (8)$$

donde ϕ es la latitud, R_e el radio medio de la Tierra y Ω su velocidad de rotación. Los términos $f = 2\Omega \operatorname{sen} \phi$ y $f^* = 2\Omega \operatorname{cos} \phi$ provienen de la rotación con la Tierra del sistema de coordenadas y, en honor del matemático francés de la primera mitad del siglo XIX Gaspard Gustav de Coriolis (1792-1843), se denominan respectivamente parámetro de Coriolis y parámetro recíproco de Coriolis.

Éstas son las *ecuaciones básicas de la conservación del momento de la meteorología dinámica*. Junto con la ecuación de la conservación de la masa:

$$\frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + v \frac{\partial \rho}{\partial y} + w \frac{\partial \rho}{\partial z} = -\rho \operatorname{div}(\vec{v}), \quad (9)$$

la ecuación de estado:

$$p = \rho RT, \quad (10)$$

la primera ley de la termodinámica:

$$C_p \left(\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + w \frac{\partial T}{\partial z} \right) = \frac{RT}{p} \left(\frac{\partial p}{\partial t} + u \frac{\partial p}{\partial x} + v \frac{\partial p}{\partial y} + w \frac{\partial p}{\partial z} \right) + Q, \quad (11)$$

y la ecuación del vapor de agua:

$$\frac{\partial q}{\partial t} + u \frac{\partial q}{\partial x} + v \frac{\partial q}{\partial y} + w \frac{\partial q}{\partial z} = Q', \quad (12)$$

forman el modelo básico completo, o lo que es lo mismo, *las ecuaciones primitivas* del movimiento de la atmósfera.

3.3 La hipótesis hidrostática

Un análisis del orden de magnitud de los distintos términos de las ecuaciones (6)-(12) (latitud 45° N, velocidades horizontales del orden de m/seg, velocidades verticales del orden de cm/seg, ...) permite obviar los términos de menor influencia atendiendo a la escala de los fenómenos meteorológicos que se quieran tratar. Para las latitudes medias del hemisferio norte y escala sinóptica se obtiene así el sistema simplificado ([6]):

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + f v + f_{rx}, \quad (13)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial y} - f u + f_{ry}, \quad (14)$$

$$\frac{\partial p}{\partial z} = -g \rho, \quad (15)$$

$$\frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial x} + v \frac{\partial \rho}{\partial y} + w \frac{\partial \rho}{\partial z} = -\rho \operatorname{div}(\vec{v}), \quad (16)$$

$$p = \rho R T, \quad (17)$$

$$C_p \left(\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + w \frac{\partial T}{\partial z} \right) = \frac{RT}{p} \left(\frac{\partial p}{\partial t} + u \frac{\partial p}{\partial x} + v \frac{\partial p}{\partial y} + w \frac{\partial p}{\partial z} \right) + Q, \quad (18)$$

$$\frac{\partial q}{\partial t} + u \frac{\partial q}{\partial x} + v \frac{\partial q}{\partial y} + w \frac{\partial q}{\partial z} = Q'. \quad (19)$$

Notemos que la ecuación (15) es la ecuación hidrostática, que expresa la variación de la presión con la altura. Hay que resaltar que la simplificación hidrostática permite “filtrar” las ondas acústicas, que sin embargo son solución de las ecuaciones generales.

Básicamente, éste es el sistema propuesto por V. Bjerknes a principios del siglo XX y que intentó resolver L.F. Richardson a finales de los años 1910. Más adelante, en la sección 4, haremos nuevos cambios de coordenadas y sobre el sistema obtenido describiremos los principios básicos de la resolución numérica de los modelos atmosféricos.

Las ecuaciones en derivadas parciales anteriores son una variante de las denominadas ecuaciones de Navier-Stokes y como para éstas, sólo se han hallado soluciones analíticas (expresiones matemáticas explícitas) en unos pocos casos particulares. La dificultad esencial reside en las no linealidades, en concreto en la presencia de términos del tipo $u \frac{\partial}{\partial x}$, $v \frac{\partial}{\partial y}$ y $w \frac{\partial}{\partial z}$, denominados términos advectivos. Como ya se ha repetido varias veces, para resolverlas es necesario apelar a aproximaciones numéricas que puedan tratarse con ordenador. Respecto de las ecuaciones de Navier-Stokes es interesante resaltar que uno de los siete problemas que el Clay Mathematics Institute of Cambridge, Massachusetts, ha elegido como problemas del nuevo milenio es: *Existencia y regularidad de las ecuaciones de Navier-Stokes*.

Hay que indicar que la comprensión de este último problema puede ser un paso fundamental para la comprensión del gran problema de la turbulencia. Es casi una paradoja que a comienzos del siglo XXI se hayan podido comprender y modelar numerosos fenómenos de escala de las micropartículas (dominio de la mecánica cuántica) o de la escala del Universo (dominio de la mecánica relativista), mientras que flujos de los fluidos más corrientes, como por ejemplo del agua o del aire, que son fenómenos que están al alcance de nuestra vista y pertenecen al dominio de la mecánica clásica, todavía hoy casi son un misterio. Este gran desafío científico se puede resumir en una palabra: *turbulencia*.

3.4 Ecuaciones del viento geostrófico

Los términos principales de las ecuaciones (13) y (14) para movimientos atmosféricos de gran escala generalmente son el gradiente de presión y el término de Coriolis. Considerando sólo esos términos, o lo que es lo mismo, admitiendo que el movimiento es no acelerado y despreciando las fuerzas de rozamiento \vec{f}_r , se obtiene el sistema ([8]):

$$v = \frac{1}{\rho f} \frac{\partial p}{\partial x}, \quad u = -\frac{1}{\rho f} \frac{\partial p}{\partial y}. \quad (20)$$

Estas dos ecuaciones se denominan ecuaciones del *viento geostrófico*, ya que éste es el hipotético resultado de un perfecto balance entre la fuerza debida a la rotación terrestre (fuerza de Coriolis) y la fuerza del gradiente de presión. A menudo este balance es una muy buena primera aproximación de los movimientos atmosféricos de gran escala en altitud (por encima de los 3 km) y latitudes extratropicales. Sin embargo, en las capas bajas de la atmósfera no se pueden obviar las fuerzas de rozamiento y la aproximación geostrófica no es válida. Esta aproximación tampoco es válida en las regiones ecuatoriales, ya que el parámetro de Coriolis f del denominador se anula en el ecuador.

Un análisis directo de las ecuaciones (20) muestra que en el hemisferio norte, en una depresión, estos vientos fluyen paralelamente a las isobaras (si por ejemplo una isobara es paralela al eje x , $\frac{\partial p}{\partial x} = 0$ y la componente del viento v en la dirección del eje y es nula) y giran en sentido contrario al de las manecillas de un reloj. El viento real, sin embargo, no es estrictamente geostrófico y tiene una pequeña componente ageostrófica dirigida hacia las bajas presiones. Esto hace posible que se rellenen las depresiones y que desaparezcan los anticiclones.



En regiones intertropicales o para ciertos fenómenos atmosféricos de pequeña escala en otras latitudes, como por ejemplo los tornados, los términos dominantes en las ecuaciones (6) y (7) tienden a ser la aceleración centrípeta y el gradiente de presión. Esto conduce a un balance que se denomina balance *ciclotrófico* (ver [7]). Hay que indicar que este sistema no tiene la fuerte predilección por un sentido de rotación que tiene el sistema en el que la fuerza de Coriolis es la predominante.

3.5 Balance de la vorticidad vertical. Ondas de Rossby

Las ondas (planetarias) de Rossby son importantes para los modelos meteorológicos de escala sinóptica y planetaria. Estas ondas son debidas a la variación del parámetro de Coriolis $f = 2\Omega \sin\phi$ con la latitud. En esta sección introduciremos la ecuación de balance de la componente vertical de la vorticidad y de ella deduciremos las ondas de Rossby.

Por definición, la componente vertical de la vorticidad es $\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$. Derivando la ecuación (14) respecto de x , la ecuación (13) respecto de y y restando las dos expresiones se obtiene:

$$\frac{\partial \zeta}{\partial t} + u \frac{\partial \zeta}{\partial x} + v \frac{\partial \zeta}{\partial y} + v \frac{df}{dy} = -w \frac{\partial \zeta}{\partial z} - (\zeta + f) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + J \left(p, \frac{1}{\rho} \right) + \frac{\partial w}{\partial y} \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \frac{\partial v}{\partial z} + \frac{\partial f_{ry}}{\partial x} - \frac{\partial f_{rx}}{\partial y},$$

donde hemos puesto $J \left(p, \frac{1}{\rho} \right) = \frac{\partial p}{\partial x} \frac{\partial}{\partial y} \frac{1}{\rho} - \frac{\partial p}{\partial y} \frac{\partial}{\partial x} \frac{1}{\rho}$.

Si estimamos el tamaño de los términos de la ecuación anterior para los movimientos de gran escala en latitudes medias del hemisferio norte, se observa que los dominantes son los del primer miembro. Si retenemos solamente esos términos se obtiene la ecuación correspondiente a la aproximación β -plano ([12]):

$$\frac{\partial \zeta}{\partial t} + u \frac{\partial \zeta}{\partial x} + v \frac{\partial \zeta}{\partial y} + v \beta_0 = 0, \tag{21}$$

donde $\beta_0 = \left[\frac{df}{dy} \right]_{y=0}$. La aparición de este término se debe a que en el desarrollo en serie del parámetro de Coriolis se retienen solamente los términos lineales ([7]).

Ésta es una ecuación con dos variables espaciales que expresa el balance de la componente vertical de la vorticidad. Aunque ya era conocida antes, Rossby fue la primera persona que identificó esa ecuación en los años 1930 como una ecuación clave para describir los movimientos atmosféricos de gran escala.

La ecuación (21) todavía se puede simplificar más si sólo se permiten variaciones de u y v en la dirección de x , y se descompone u en una parte constante U y una pequeña perturbación u' de forma que $u = U + u'$. Suponiendo que $U \gg u'$ y despreciando los términos en los que aparece la perturbación u' , se obtiene la ecuación simplificada ([12]):

$$\frac{\partial^2 v}{\partial x \partial t} + U \frac{\partial^2 v}{\partial x^2} + \beta_0 v = 0.$$

Esta ecuación tiene la solución con forma de onda:

$$v(x, t) = A \cos \left[\frac{2\pi}{L}(x - ct) \right],$$

donde L es la longitud de onda y

$$c = U - \beta_0 \frac{L^2}{4\pi^2} \quad (22)$$

es la velocidad de onda. Así pues, en el hemisferio norte, estas ondas se propagan hacia el oeste con respecto al flujo zonal medio U .

Rossby conocía los primeros mapas de los movimientos del aire de las capas altas de la atmósfera hechos a partir de los datos atmosféricos obtenidos por sondas enviadas a gran altura. Por ello, Rossby sabía que la característica dominante del flujo de esas capas de aire en latitudes medias es una corriente cíclica en dirección este que tiene una forma ligeramente serpenteada. Este serpenteamiento incluye ondas de longitud de onda tan grande que sólo hay dos o tres ondas a lo largo de todo un paralelo. Estas ondas, denominadas ondas planetarias, están acompañadas por patrones sobreimpuestos compuestos por entre seis y ocho ondas alrededor del hemisferio. Estas últimas también son ondas de larga longitud (unos pocos miles de kilómetros) y se dice que son ondas de escala sinóptica. Ahora bien, a estas observaciones les faltaba una explicación teórica que las justificara y Rossby lo consiguió utilizando la fórmula (22), que muestra que la velocidad de movimiento de las ondas atmosféricas de gran escala sólo depende de la velocidad media U y de la longitud de onda L . A las ondas que verifican (22) se les denominó ondas de Rossby. Valores típicos de esas variables son $L \approx 3000$ km, $U \approx 10$ m/seg y $\beta \approx 10^{-11} \text{ m}^{-1} \text{ seg}^{-1}$, lo que implica que las ondas de Rossby de escala sinóptica se desplazan a unos 8 m/seg.

Teniendo en cuenta la severidad de las suposiciones realizadas, esta fórmula es sorprendentemente exitosa a la hora de explicar gran parte del

comportamiento de los patrones atmosféricos de gran escala. Hay que resaltar que desde una perspectiva histórica, el éxito de la teoría de Rossby proporcionó a los meteorólogos el coraje necesario para intentar la predicción numérica del tiempo. De hecho, para el primer pronóstico computacional realizado en 1950, Charney, Fjörtoft y von Neumann resolvieron por diferencias finitas una ecuación pronóstico similar a la del balance de la componente vertical de la vorticidad:

$$\frac{\partial \zeta}{\partial t} + u \frac{\partial \zeta}{\partial x} + v \frac{\partial \zeta}{\partial y} + v \beta = 0.$$

En un sentido estricto, la ecuación anterior sólo es válida para una atmósfera promedio en la que el viento no cambia con la altura (no aparece la variable z). En concreto esa ecuación proporcionó una buena primera aproximación al movimiento medio de la troposfera, o lo que es lo mismo, al movimiento a un nivel intermedio entre los 500 y los 600 mbar como si solamente fuese bidimensional. Este tipo de modelo se denomina *barotrópico*. En general un flujo se dice barotrópico si la presión es constante sobre las superficies de densidad constante. Por tanto, en esos casos, la temperatura no está conectada a la presión o a los cambios de densidad.

La principal ventaja de los modelos barotrópicos es su simplicidad. Sin embargo, como en esos modelos se excluyen los fenómenos térmicos, tienen severas limitaciones, pues por ejemplo no contemplan la formación de centros ciclónicos o anticiclónicos ni la influencia de la radiación.

Hay que resaltar que el modelo barotrópico inicial sufrió diversas modificaciones, algunas de las cuales resultaron interesantes. Así, el principio del balance de la vorticidad vertical absoluta fue reemplazado por el de la conservación de la vorticidad geostrófica absoluta, y a los modelos que utilizaron el viento geostrófico se les denominó *modelos geostróficos*. Posteriormente, a los modelos obtenidos aproximando el viento y la vorticidad (de forma selectiva, sólo en algunos términos) por unos valores geostróficos se les llamó *cuasigeostróficos* (cf. [7] y [8]).

3.6 Modelos baroclínicos

Como contraposición a los modelos barotrópicos están los modelos *baroclínicos*, en los que las superficies de igual densidad están inclinadas con respecto a las superficies de igual presión, o lo que es lo mismo, la presión no es constante sobre las superficies con densidad constante, sino que varía con las variaciones de la temperatura.

Mientras que los modelos barotrópicos predicen el movimiento de la troposfera media, los modelos baroclínicos también incluyen la estructura vertical de la atmósfera. Se tiene entonces modelos de 2, 3, 4, ... niveles, los cuales predicen el flujo atmosférico en esos niveles. Ni que decir tiene que existen varias versiones para cada una de esas situaciones. Hay que resaltar que estos modelos admiten intercambios caloríficos con fuentes calientes o frías, son capaces de describir procesos donde las energías potencial e interna se convierten

en cinética, dan cuenta de algunas formas de inestabilidad atmosférica, y a partir de tres niveles pueden incluso contener cambios en la estabilidad estática.

Veamos ahora cómo variaciones horizontales de la temperatura implican variaciones del viento con la altura y, por tanto, la necesidad de incluir la estructura vertical de la atmósfera. Combinemos por ejemplo la ecuación de estado $p = \rho RT$ con las ecuaciones (20) del viento geostrófico y la ecuación hidrostática (15). Procediendo de esta forma se obtiene ([12]):

$$\frac{g}{T} = -R \frac{\partial(\ln p)}{\partial z}, \quad \frac{u}{T} = -\frac{R}{f} \frac{\partial(\ln p)}{\partial y}, \quad \frac{v}{T} = \frac{R}{f} \frac{\partial(\ln p)}{\partial x}.$$

Derivando las dos últimas ecuaciones respecto de z y teniendo en cuenta la primera de las tres ecuaciones anteriores se tiene:

$$\frac{\partial}{\partial z} \left(\frac{u}{T} \right) = -\frac{g}{fT^2} \frac{\partial T}{\partial y}, \quad \frac{\partial}{\partial z} \left(\frac{v}{T} \right) = \frac{g}{fT^2} \frac{\partial T}{\partial x}.$$

Estas ecuaciones se denominan ecuaciones del *viento térmico* y reflejan el hecho esencial de que en una atmósfera baroclínica las variaciones horizontales de la temperatura implican variaciones verticales del viento.

Naturalmente, los cálculos para resolver los modelos baroclínicos son mucho más largos que los de los modelos barotrópicos. En la siguiente sección mostraremos un modelo numérico para el caso en que se admita la hipótesis hidrostática.

Comentario a las referencias. Los libros [6], [7], [8] y [12] son extensas guías introductorias a la dinámica atmosférica. En ellos se describe la estructura de la atmósfera, se introducen los principios físicos básicos necesarios para explicar su comportamiento, para luego obtener las ecuaciones primitivas de su movimiento. Después se estudian las implicaciones de las ecuaciones en una variedad de casos especiales. Así, se desarrollan distintos modelos utilizando suposiciones simplificadoras, que en muchos casos restringen su aplicación a fenómenos pertenecientes a un determinado rango de escalas espaciales y temporales. En [8] y [12] se presta atención a la resolución numérica de los modelos introducidos y en [8] también se estudian fenómenos de circulación oceánica.

4 La resolución numérica

Vamos a describir aquí un esquema numérico para resolver un modelo baroclínico en el que la coordenada vertical es la presión normalizada. En concreto se utilizará un método en diferencias finitas con integración temporal explícita. Para ello se dividirá la atmósfera en N capas de igual espesor en presión y se definirá un mallado regular sobre las variables horizontales. En realidad, este mallado horizontal a menudo suele ser una cuadrícula colocada sobre un mapa, es decir sobre una proyección plana (o esférica como en el modelo ARPEGE de Météo-France) de parte de la Tierra curva. En general los meteorólogos utilizan mapas que conservan la forma de cualquier pequeño elemento de área de la Tierra (mapas conformes, como por ejemplo una

proyección estereográfica polar), pero que pueden distorsionar el tamaño relativo del elemento. En este caso, si el mallado es regular sobre la proyección, no lo es sobre la superficie terrestre, por lo que en las ecuaciones del modelo escritas para la proyección habrá que introducir un *factor de mapa*. En lo que sigue y para no complicar la exposición, razonaremos sobre una malla horizontal fija y un factor de mapa igual a uno que, por tanto, no aparecerá en las ecuaciones.

Esta presentación nos permitirá mostrar de manera bastante detallada cómo funciona un modelo de predicción numérica del tiempo. Los modelos operativos modernos utilizan técnicas más complicadas (mallado variable en las variables horizontales y en la vertical, métodos espectrales, integración semi-lagrangiana y semi-implícita), pero una representación detallada de esos métodos nos conduciría a unos desarrollos que sobrepasan el marco de este trabajo. En cualquier caso, el modelo que presentamos es un modelo totalmente realista que, como hemos dicho, nos va a permitir comprender el funcionamiento general de un modelo baroclínico y las técnicas de numerización adoptadas, en particular el tratamiento de la variable vertical.

4.1 El modelo en coordenada vertical presión normalizada

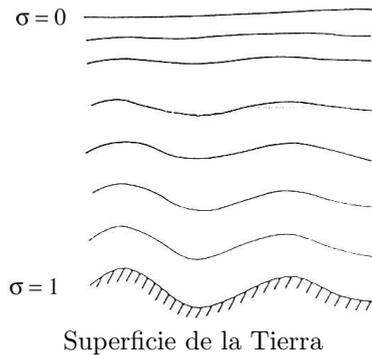
Volvamos al sistema (13)-(19). La ecuación de la hidrostática (15) permite sustituir la coordenada vertical z por la coordenada $\sigma = p/p_s$, donde p es la presión en el punto de altura z y p_s la presión en la proyección de ese punto sobre el suelo. Así pues, aquí la posición vertical viene dada por la presión normalizada.

En la cima de la atmósfera se tiene $p = 0$, luego $\sigma = 0$, y en la base de la atmósfera, en el nivel del relieve, se tiene $p = p_s$, luego $\sigma = 1$. Tenemos pues que esta nueva coordenada vertical σ varía desde 0 en la cima de la atmósfera a 1 en su base. De aquí le viene el nombre de coordenada *presión normalizada*.

La definición de esta nueva coordenada vertical necesita introducir una velocidad vertical generalizada $\dot{\sigma} = \frac{d\sigma}{dt}$.

Para que el problema propuesto se pueda resolver hay que fijar el valor de $\dot{\sigma}$ en la base y en la cima de la atmósfera. En concreto, en esas dos superficies frontera se tomará $\dot{\sigma} = 0$, lo cual asegura la nulidad de los flujos para las cantidades conservativas en la cima y en la base de la atmósfera, supuesta adiabática y sin rozamiento.

Esta coordenada σ facilita formular correctamente las condiciones de contorno en el límite inferior de la atmósfera en presencia de relieve. En contrapartida, su introducción complica ligeramente las ecuaciones, y obliga a realizar una interpolación de los valores de las variables medidos por el sistema



internacional de observación en los niveles de presión, a los niveles sigma del modelo.

En las nuevas coordenadas (x, y, σ) el sistema (13)- (19) se escribe:

$$\frac{\partial u}{\partial t} = -u \frac{\partial u}{\partial x} - v \frac{\partial u}{\partial y} - \dot{\sigma} \frac{\partial u}{\partial \sigma} + f v - \frac{\partial \Phi}{\partial x} - RT \frac{\partial \ln p_s}{\partial x} + f_{rx}, \quad (23)$$

$$\frac{\partial v}{\partial t} = -u \frac{\partial v}{\partial x} - v \frac{\partial v}{\partial y} - \dot{\sigma} \frac{\partial v}{\partial \sigma} - f u - \frac{\partial \Phi}{\partial y} - RT \frac{\partial \ln p_s}{\partial y} + f_{ry}, \quad (24)$$

$$\frac{\partial T}{\partial t} = -u \frac{\partial T}{\partial x} - v \frac{\partial T}{\partial y} - \dot{\sigma} \frac{\partial T}{\partial \sigma} - \frac{RT}{C_p} \frac{1}{\sigma p_s} \int_0^\sigma \left[\frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y} \right] d\sigma + \quad (25)$$

$$\frac{RT}{C_p} \left[u \frac{\partial \ln p_s}{\partial x} + v \frac{\partial \ln p_s}{\partial y} \right] + \frac{Q}{C_p},$$

$$\frac{\partial q}{\partial t} = -u \frac{\partial q}{\partial x} - v \frac{\partial q}{\partial y} - \dot{\sigma} \frac{\partial q}{\partial \sigma} + Q', \quad (26)$$

$$\frac{\partial p_s}{\partial t} = - \int_0^1 \left[\frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y} \right] d\sigma, \quad (27)$$

$$\Phi = \Phi_s + \int_1^\sigma RT d(\ln \sigma), \quad (28)$$

$$\dot{\sigma} = -\frac{1}{p_s} \int_0^\sigma \left[\frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y} \right] d\sigma + \frac{\sigma}{p_s} \int_0^1 \left[\frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y} \right] d\sigma, \quad (29)$$

donde Φ es el geopotencial ($\Phi = gh$ donde $z = h(x, y, p)$ define la superficie de presión constante igual a p) y Φ_s es el geopotencial al nivel del suelo. De la definición de geopotencial se sigue que $\Phi_s(x, t) = gh_s(x, y)$ donde $h_s(x, y)$ es la altitud del relieve del modelo. Notemos que, por ejemplo, en el capítulo 3 de [5] se obtienen estas ecuaciones en la variable σ y que en [6] se introducen las coordenadas isentrópicas en las que utiliza la temperatura potencial como medida de la posición vertical.

Determinada la parte física del modelo (es decir f_{rx}, f_{ry}, Q, Q'), este sistema consta de siete ecuaciones con siete incógnitas ($u, v, T, q, p_s, \Phi, \dot{\sigma}$).

Principio de resolución

Dividamos la atmósfera en N capas de igual espesor en presión. En el instante inicial conocemos los valores de u, v, T y q en los N niveles del modelo (superficies mitad de cada capa) y de p_s en la superficie del suelo ($\sigma = 1$). Mediante las ecuaciones diagnóstico (28) y (29) se calcula los valores de Φ y $\dot{\sigma}$ en las N superficies intercapa. Esta distinta localización de las variables en las

superficies mitad de capa y en las superficies intercapa se denomina mallado de Lorenz, debido a E. Lorenz.

Las ecuaciones pronóstico (23)-(27) permiten calcular las derivadas $\frac{\partial}{\partial t}$ de las variables evolutivas u, v, T, q y p_s en el instante inicial, y un esquema explícito en tiempo (en adelante α designará cualquiera de las variables u, v, T, q y p_s),

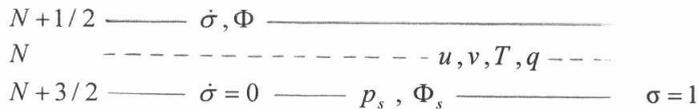
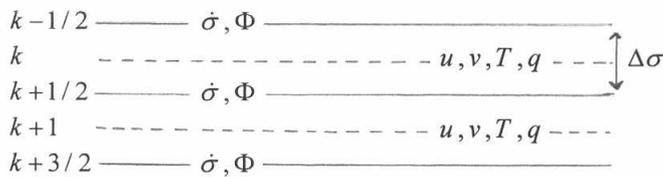
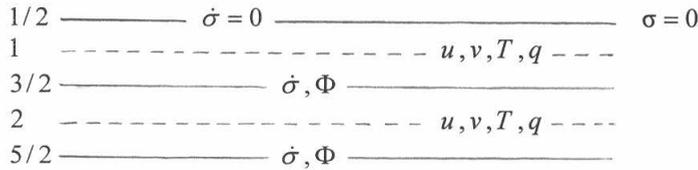
$$\alpha(\Delta t) = \alpha(0) + \Delta t \left(\frac{\partial \alpha}{\partial t}(0) \right),$$

nos da valores de u, v, T y q en los N niveles del modelo y de p_s en la superficie del suelo, en el instante Δt .

Hecho esto se vuelve a calcular las variables diagnóstico Φ y $\dot{\sigma}$ en el instante Δt , luego las tendencias de las variables evolutivas (derivadas temporales) y mediante un esquema explícito de la forma

$$\alpha(2\Delta t) = \alpha(0) + 2\Delta t \left(\frac{\partial \alpha}{\partial t}(\Delta t) \right),$$

obtenemos los valores de u, v, T, q y p_s en el instante $2\Delta t$.



Discretización vertical y emplazamiento de las variables.
Los límites de las capas están representados por líneas continuas y sus superficies mitad por discontinuas.

Se puede continuar este proceso utilizando un esquema *leap-frog*, de orden 2

de precisión,

$$\alpha(t + \Delta t) = \alpha(t - \Delta t) + 2\Delta t \left(\frac{\partial \alpha}{\partial t}(t) \right),$$

teniendo en cuenta que en el cálculo de las tendencias intervienen varias derivadas horizontales que se calculan con esquemas en diferencias centrados de orden 2 de precisión,

$$\left[\frac{\partial \alpha}{\partial x} \right]_i \cong \frac{\alpha(x + \Delta x) - \alpha(x - \Delta x)}{2\Delta x}.$$

Un mallado clásico para la discretización por diferencias finitas en las variables horizontales es el de Arakawa, debido a Akio Arakawa, actualmente profesor emérito en UCLA. Dada la complejidad de los distintos términos a discretizar no lo desarrollaremos en este trabajo. En cualquier caso, en el capítulo 7 de [8] se desarrollan métodos en diferencias finitas para hallar soluciones de las ecuaciones de la dinámica de la atmósfera. En concreto se estudia el mallado de Lorenz, el de Arakawa, el esquema *leap-frog*, el de Matsuno y combinaciones de los mismos.

Mostremos ahora los cálculos que hay que realizar en los términos con la variable vertical σ . Para más detalles ver [5], donde se describe el tratamiento de la discretización en esa variable con el mallado de Lorenz y se comentan distintas alternativas para mejorar los métodos presentados.

Cálculo del geopotencial

La expresión del geopotencial está dada por la ecuación (28). Se puede por tanto calcular $\Phi_{k-1/2}$ sobre las superficies intercapas evaluando la integral de (28) por un método de rectángulos:

$$\Phi_{k-1/2} = \Phi_s - \sum_{l=k}^N RT_l \Delta(\ln \sigma), \quad \text{para } k = 2, 3, \dots, N.$$

Ahora bien, para el cálculo de las tendencias de u y de v mediante las ecuaciones (23) y (24) se necesita conocer el valor de Φ sobre los niveles del modelo. Estos valores se calculan con la fórmula

$$\Phi_k = \frac{\Phi_{k-1/2} + \Phi_{k+1/2}}{2}, \quad \text{para } k = 2, 3, \dots, N.$$

Esta fórmula no es válida para $k = 1$, ya que en la superficie $\sigma = 0$ no está definido $\Phi_{1/2}$. Por ello, para determinar Φ_1 se discretiza la ecuación de la hidrostática $\frac{\partial \Phi}{\partial \ln \sigma} = -RT$ (que se obtiene derivando la expresión (28)) para la primera semicapa,

$$\Phi_1 = \Phi_{3/2} + RT_1 \left(\ln \Delta \sigma - \ln \frac{\Delta \sigma}{2} \right) = \Phi_{3/2} + RT_1 \ln 2.$$

Cálculo de la presión en el suelo

Se tiene que calcular

$$\frac{\partial p_s}{\partial t} = - \int_0^1 \left[\frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y} \right] d\sigma = - \int_0^1 D d\sigma,$$

donde hemos puesto $D = \frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y}$.

D está definido en los mismos emplazamientos que u y v y su integral se puede evaluar mediante un método de rectángulos:

$$\frac{\partial p_s}{\partial t} = - \sum_{k=1}^N D_k \Delta\sigma.$$

Cálculo de la velocidad vertical generalizada

Vayamos a la ecuación (29). Evaluando las integrales por un método de rectángulos se tiene

$$\dot{\sigma}_{k+1/2} = - \frac{1}{p_s} \sum_{l=1}^k D_l \Delta\sigma + \frac{\sigma}{p_s} \sum_{l=1}^N D_l \Delta\sigma,$$

que permite obtener la velocidad vertical generalizada en las superficies intercapas.

Cálculo de la integral de la divergencia

En la ecuación de la temperatura tenemos que evaluar

$$I_k = \int_0^\sigma \left[\frac{\partial(p_s u)}{\partial x} + \frac{\partial(p_s v)}{\partial y} \right] d\sigma = \int_0^\sigma D d\sigma$$

en los niveles del modelo (superficies mitad de capa). Teniendo en cuenta que al llegar al punto medio de la capa k -ésima hemos recorrido $k - 1$ capas y la mitad de esta capa k -ésima, utilizaremos la fórmula de integración numérica:

$$I_k = \sum_{l=1}^{k-1} D_l \Delta\sigma + D_k \frac{\Delta\sigma}{2}, \quad \text{para } k = 1, 2, \dots, N.$$

Las advecciones verticales

Calculemos ahora $\dot{\sigma} \frac{\partial\alpha}{\partial\sigma}$ en los niveles del modelo.

Como la variable $\dot{\sigma}$ está estimada en las superficies intercapas, empezaremos estimando $\dot{\sigma} \frac{\partial\alpha}{\partial\sigma}$ en esas superficies. Para ello utilizaremos la discretización

$$\left[\dot{\sigma} \frac{\partial\alpha}{\partial\sigma} \right]_{k+1/2} = \dot{\sigma}_{k+1/2} \frac{\alpha_{k+1} - \alpha_k}{\Delta\sigma},$$

$$\left[\dot{\sigma} \frac{\partial \alpha}{\partial \sigma} \right]_{k-1/2} = \dot{\sigma}_{k-1/2} \frac{\alpha_k - \alpha_{k-1}}{\Delta \sigma}.$$

Hecho esto estimaremos el valor de $\dot{\sigma} \frac{\partial \alpha}{\partial \sigma}$ en los niveles del modelo calculando la media entre los valores en la superficie y en la base de la capa,

$$\left[\dot{\sigma} \frac{\partial \alpha}{\partial \sigma} \right]_k = \frac{1}{2} \left(\dot{\sigma}_{k+1/2} \frac{\alpha_{k+1} - \alpha_k}{\Delta \sigma} + \dot{\sigma}_{k-1/2} \frac{\alpha_k - \alpha_{k-1}}{\Delta \sigma} \right), \text{ para } k = 1, 2, \dots, N.$$

Aquí es necesario tener en cuenta que $\dot{\sigma}_{1/2} = \dot{\sigma}_{N+1/2} = 0$, ya que hemos considerado que $\dot{\sigma}$ se anula en las superficies $\sigma = 0$ y $\sigma = 1$.

Recapitulación

Como acabamos de mostrar, resulta relativamente fácil integrar de forma explícita un modelo baroclínico en ecuaciones primitivas utilizando técnicas numéricas elementales. Como contrapartida a esta sencillez, en este caso es necesario tomar un paso de tiempo que satisfaga la condición de estabilidad de Courant, Friedrichs y Lewy (CFL):

$$c \frac{\Delta t}{\Delta x} \leq \frac{1}{\sqrt{2}}, \quad (30)$$

donde c es la velocidad de fase de las ondas gravitatorias más rápidas, que está próxima de la velocidad del sonido ($c \approx 300$ m/seg). El factor $\sqrt{2}$ aparece al calcular la condición de estabilidad para problemas bidimensionales sobre la horizontal.

En el caso de un modelo real los cálculos son un poco más complicados, ya que en las ecuaciones se debe introducir un factor de mapa. En lo que respecta a la horizontal, se puede elegir una discretización más sofisticada en la que las variables u , v y T están situadas sobre redes ligeramente desplazadas. Además, aunque ello implique complicaciones suplementarias, también se procura utilizar discretizaciones que tengan propiedades similares a las de las ecuaciones analíticas (conservación de la masa, de la energía, del momento angular, ...) y sean susceptibles de proporcionar representaciones más fieles de la atmósfera real (ver [8] y [12]).

La condición CFL (30) exige que el paso de tiempo Δt sea dos mil veces más pequeño que el paso de espacio Δx . Ahora bien, esto se puede mejorar con sólo tratar de forma implícita los términos responsables de la generación de las ondas gravitatorias. En este caso la condición de estabilidad CFL a respetar es

$$U \frac{\Delta t}{\Delta x} \leq \frac{1}{\sqrt{2}},$$

donde U es la velocidad de propagación de las ondas de Rossby que, como mucho, es la del viento sinóptico ($U \approx 50$ m/seg). Así pues, la utilización de un método semi-implícito permite tomar un paso de tiempo seis veces más grande que el de un explícito.

Los métodos semi-implícitos para las ecuaciones primitivas todavía se pueden mejorar más si se tratan de forma lagrangiana los términos advectivos. Así, combinando advección lagrangiana y tratamiento implícito de los términos generadores de las ondas gravitatorias, se obtiene un modelo *semi-lagrangiano semi-implícito* incondicionalmente estable. Estos modelos son más complejos y costosos de ejecutar que los explícitos pero como permiten un paso de tiempo mayor, al final son los que ganan la partida.

Básicamente, los métodos espectrales consisten en desarrollar el campo considerado en una serie de funciones conocidas y realizar los cálculos correspondientes en la serie truncada. En meteorología, si se usa, la técnica espectral solamente se utiliza en el cálculo de las derivadas horizontales. En lo que concierne a la vertical se mantienen las diferencias finitas en la forma que hemos descrito. En ese caso, para trabajar sobre la esfera se utilizan como funciones de base los armónicos esféricos de superficie, que están definidos por

$$Y_n^m(\lambda, \mu) = P_n^m(\mu) \exp(im\lambda),$$

donde λ representa la longitud terrestre, μ el seno de la latitud y $P_n^m(\mu)$ son las funciones (reales) asociadas de Legendre de primera especie definidas sobre el intervalo $[-1, 1]$.

Los métodos espectrales inicialmente sólo se utilizaron para los modelos globales. Hoy en día su campo de acción se ha extendido para tratar dominios geográficos limitados. En estos casos, considerando un dominio de trabajo lo suficientemente grande para que se pueda suponer que los campos tengan periodicidad lateral, se han probado con éxito métodos espectrales basados en una descomposición en funciones trigonométricas.

Otra alternativa para los modelos de dominio limitado es tomar como condiciones de contorno laterales los valores obtenidos en otro modelo de escala más grande. En esta situación es necesario introducir un término disipativo destinado a atenuar las perturbaciones engendradas por la artificiosidad de las condiciones de contorno laterales, perturbaciones que se propagan al interior del dominio. Se llega así a los modelos encajados. Este es el caso del modelo ALADIN, encajado en el ARPEGE, ambos desarrollados por Météo-France (ver [9]).

Paralelamente, la utilización de una transformación “conforme” de la esfera sobre sí misma ha desembocado en el concepto de malla variable que, aunque el modelo sea global, aumenta la resolución en la zona de mayor interés (modelo ARPEGE).

Otra cuestión a la que se ha dedicado mucha atención y que ha supuesto una notable mejoría de los modelos, es la determinación precisa de un estado inicial de la atmósfera a partir de las observaciones meteorológicas disponibles. Esta operación, denominada análisis objetivo, al principio se realizaba utilizando métodos de interpolación geométrica. A mediados de los años 1980 se propuso una formulación variacional (búsqueda del mínimo de un funcional) muy general, que puede resolverse utilizando métodos de control óptimo. En este caso se habla de asimilación variacional de los datos de observación.

Esta última aproximación permite tener en cuenta la información suministrada por una gran variedad de sistemas de observación, en particular de los datos teledetectados por el sistema de satélites, los cuales están ligados a las variables de los modelos por relaciones no lineales. Además, la minimización se puede extender fácilmente a datos repartidos en el espacio y tiempo: se habla entonces de asimilación variacional cuatridimensional (en forma abreviada 4D-VAR). En este caso el mismo modelo asegura que los campos verifican la restricción de coherencia interna y no es necesario recurrir a un proceso de inicialización. El desarrollo de estos métodos ha desembocado en la obtención de nuevas herramientas derivadas del modelo, como son el modelo lineal tangente y el modelo adjunto. Señalemos que en [10] se describe la asimilación 4D-VAR de los datos de observación y los problemas relativos a la asimilación de los datos de observación para la predicción numérica del tiempo. Para un estudio más exhaustivo ver su bibliografía.

5 La parte física de los modelos

La parte física de los modelos permite calcular los distintos intercambios energéticos entre la atmósfera y las fuentes externas. Considerada la atmósfera de forma global, las fuentes son la radiación solar, los océanos y en menor medida los continentes. Los procesos de intercambio de energía entre esas fuentes externas y la atmósfera son extremadamente variados y altamente interactivos, por lo tanto muy complejos. Como estos procesos no están tratados de forma explícita en las ecuaciones de la parte dinámica del modelo (sea por su naturaleza física, sea por las escalas consideradas), se deben parametrizar. Esto significa que se tiene que determinar e introducir el efecto medio de estos procesos en las variables dinámicas del modelo, o lo que es lo mismo, se tienen que calcular los términos \vec{f}_r , Q y Q' de las ecuaciones.

En esta sección nos limitaremos a describir brevemente los fenómenos físicos más relevantes que intervienen en estos procesos. Para un estudio más detallado ver por ejemplo [1], [5], [8] y [11]. Así, en [5] se describen con sencillez los principales fenómenos físicos y la manera de parametrizarlos para incluirlos en las ecuaciones de la dinámica atmosférica. [1] es un libro de texto a nivel de licenciatura que trata principalmente aspectos físicos de la atmósfera. [11] es una buena introducción a la meteorología y dedica especial atención a la formación de nubes, al estudio de las masas de aire y a la frontología. Como novedad respecto de los libros anteriores, en [8] se estudian procesos químicos atmosféricos.

Vapor de agua y aire húmedo

Toda porción de atmósfera contiene algo de vapor de agua. Ahora bien, como indica el hecho de que las nubes solo ocupan una pequeña parte de la atmósfera, se puede decir es muy raro que este aire esté saturado. La mezcla de aire seco con vapor de agua se denomina *aire húmedo*. Para tratar de forma rigurosa las transformaciones del aire húmedo, se tienen que modificar

ligeramente la ecuaciones del modelo. En concreto, basta con reemplazar los términos RT y C_p por $R_a T_v$ y C_{ph} , donde R_a es la constante del aire seco, T_v la temperatura virtual y C_{ph} el calor específico a presión constante para el aire húmedo. Notemos que estas dos últimas cantidades dependen de la humedad específica q .

Los efectos de la radiación

El sol es el gran motor de la máquina atmosférica, por tanto es muy importante conocer con precisión sus efectos directos e indirectos. Los efectos combinados de la radiación solar, de la radiación atmosférica y de la radiación terrestre van a contribuir a modificar la temperatura de las diferentes capas del modelo.

Dada la gran complejidad de los fenómenos que intervienen en la radiación, básicamente la absorción y la dispersión por las moléculas de aire, para calcular efectivamente todas las interacciones es necesario conocer la repartición de los distintos constituyentes atmosféricos.

El cálculo de los efectos de la radiación reposa sobre el cálculo de los flujos en la base y en la cima de cada capa considerada. Así, el aumento de temperatura debido a los efectos de la radiación vendrá dado por la divergencia de los flujos de radiación. Teóricamente se pueden calcular esos flujos si se conoce la distribución de los distintos constituyentes atmosféricos y su temperatura.

Los flujos de radiación en el suelo (radiación solar global y radiación atmosférica) son muy dependientes de la cobertura nubosa. Para determinar la evolución de la temperatura de la superficie del suelo (condición de contorno) es esencial conocer esos flujos radiativos, el flujo de radiación terrestre y los flujos turbulentos de calor sensible y de evaporación.

La interfase tierra-atmósfera

La capa límite atmosférica (CLA) es la parte de la atmósfera comprendida entre los 0 y 1500 metros de altitud, dependiendo de la topografía del terreno. En esta capa se tienen movimientos de tipo turbulento, los cuales contribuyen a que el suelo y la atmósfera intercambien cantidad de movimiento (rozamiento), calor (calor sensible) y vapor de agua (evaporación). La CLA se puede descomponer en dos partes: la capa límite superficial (CLS), en la cual se pueden despreciar los efectos de la fuerza de Coriolis, y la capa límite planetaria (CLP), en la que hay que introducir esos términos. La CLS es variable, pero más o menos llega hasta los 50 o los 100 metros de altitud. La CLP llega hasta los 1500 metros de altitud.

La parametrización de la capa límite consiste en determinar los valores de los flujos turbulentos en el suelo, en la CLS y en la CLP, en función de los valores de las variables de estado proporcionadas por la parte dinámica del modelo, de variables en el suelo proporcionadas por un modelo de superficie y de parámetros característicos del suelo. Se tiene pues que conocer los valores del viento en la CLS, de la energía estática seca ($s = C_p T + gz$) y de la humedad específica. En lo que respecta al viento se supone nulo en la superficie. En lo que respecta a la temperatura y humedad, el problema es más complejo y lleva

a introducir ecuaciones pronóstico suplementarias para describir esas variables de superficie, gobernadas por los flujos de radiación, el flujo de precipitación y los flujos de calor sensible y latente.

Las precipitaciones de gran escala

Se incluyen aquí las precipitaciones que no son de origen convectivo. El principio básico consiste en examinar si una capa está saturada o no. En el primer caso el exceso de vapor de agua se condensa y precipita en la capa inferior y, paralelamente, la capa en la que se produce la condensación se calienta debido a la liberación de calor latente.

Sobre este fenómeno habitualmente se hacen las siguiente hipótesis:

- el agua se condensa en cuanto se alcanza la saturación,
- en el transcurso de las transformaciones la energía se conserva,
- el agua condensada precipita inmediatamente en la capa inferior (no se tiene en cuenta el agua líquida de las nubes),
- parte del agua transmitida a una capa se evapora hasta que eventualmente se alcance la saturación, y parte se transmite a la capa inmediatamente inferior.

Las convección intensa

Recordemos que la hipótesis hidrostática se obtiene suponiendo que el orden de magnitud de las velocidades verticales es menor que el de las velocidades horizontales. Si bien esto es cierto a gran escala, no lo es a pequeña escala, por ejemplo en las tormentas. Así, en los modelos de previsión numérica basados en las ecuaciones primitivas (13)- (19) no se tienen directamente en cuenta los movimientos convectivos. Sin embargo, como sus efectos se dejan sentir en la escala sinóptica, resulta indispensable evaluar los efectos de la convección sobre las variables de gran escala.

La convección se desencadena si existen capas inestables debido a la estratificación térmica o, adicionalmente, si hay un aporte de vapor de agua debido a una fuerte evaporación o a una importante convergencia de humedad. En este caso el efecto de la convección es una homogeneización vertical de la energía estática seca, de la humedad específica y de la cantidad de movimiento. Naturalmente, la atmósfera convectiva es sede de movimientos ascensionales.

Efectos de las ondas gravitatorias orográficas

Las ondas gravitatorias orográficas pueden aparecer cuando un flujo atmosférico estable circula por encima de un relieve accidentado. En determinadas condiciones, la heterogeneidad del relieve induce un flujo de cantidad de movimiento que se puede propagar verticalmente y disiparse o ser absorbido en la parte alta de la atmósfera. Como su propagación de onda generalmente es débil con relación a la resolución horizontal del modelo, es necesario parametrizar el efecto de estos fenómenos sobre la circulación a gran escala.

Transportes turbulentos horizontales

Introducida teóricamente para parametrizar el efecto de los transportes turbulentos horizontales, la difusión horizontal también permite eliminar el ruido de pequeña escala que puede generarse en el modelo.

En general, a las ecuaciones de evolución se les añade un laplaciano iterado multiplicado por un coeficiente de difusión horizontal. Se introduce este término difusivo en las ecuaciones de evolución para el viento horizontal, la energía estática seca y la humedad específica. Es inmediato constatar que el efecto de estos términos difusivos es suavizar las heterogeneidades.

Validación de las parametrizaciones

Como se ha mencionado al comienzo de esta sección, las parametrizaciones físicas son muy interactivas y dependen de parámetros que deben ser ajustados. En general se intenta limitar el número de parámetros ajustables, ya que cada reglaje necesitará una serie de ensayos del modelo. Las pruebas de las distintas parametrizaciones se pueden realizar examinando la mejora de la predicción al modificar los parámetros.

6 Modelos de predicción operativos

La creciente complejidad de los modelos de predicción numérica y las dificultades para poner a punto programas realmente eficaces en superordenadores científicos han propiciado el paso de una forma de trabajo artesanal individual a un gran proyecto científico. De hecho, mientras que los primeros modelos de predicción numérica pudieron ser concebidos, desarrollados y probados por una sola persona, la puesta a punto de los modelos actuales implica la cooperación de numerosos equipos que sobrepasa las disponibilidades de un único servicio nacional. Así, sobre todo en Europa, se ha asistido a la realización de “modelos comunitarios” o de “modelos unificados”, destinados a varias categorías de utilizadores de distintos servicios meteorológicos.

Entre las realizaciones de modelos comunes a distintos servicios meteorológicos citaremos:

- el modelo HIRLAM (High Resolution Limited Area Model), resultado del trabajo común llevado a cabo desde 1985 por los países escandinavos, Irlanda, Holanda y España;
- el modelo ARPEGE-IFS (Integrated Forecast System), desarrollado por Météo-France y el ECMWF (European Center for Medium-Range Weather Forecast) desde 1987;
- el modelo ALADIN (Aire Limitée, Adaptation Dynamique, Développement InterNational), desarrollado a partir de 1992 por Météo-France en colaboración con investigadores de Europa central. Este modelo es explotado por Austria, Bélgica, Bulgaria, Croacia, Eslovaquia, Eslovenia, Francia, Hungría, República Checa y Rumania.

Dos centros meteorológicos se explotan en común en Europa:

- El ECMWF, creado en 1974 y localizado en Reading (Reino Unido). Agrupa a dieciocho países de Europa occidental y suministra diariamente previsiones a medio plazo (hasta diez días).
- El centro RCLACE (Regional Center for Limited Area Modelling in Central Europe), creado en 1994 en Praga. Agrupa a seis países de Europa central y oriental (Austria, Croacia, Hungría, República Checa, Eslovaquia y Eslovenia) y hace funcionar dos veces al día un modelo de dominio limitado de escala fina que proporciona predicciones para hasta dos días de plazo.

En el mundo sólo alrededor de una docena de Servicios meteorológicos nacionales hacen predicciones numéricas operativas para el conjunto del planeta. En Europa, además del ECMWF la hacen Météo-France en Toulouse, el Meteorological Office británico en Bracknell y el Deutscher Wetterdienst en Offenbach.

Los modelos ARPEGE y ALADIN

El 6 de diciembre de 1993, Météo-France puso en servicio operativo el modelo ARPEGE (acrónimo de Action de Recherche Petite Échelle Grande Échelle). Éste es un modelo que abarca todo el globo terrestre y su principal innovación es su resolución horizontal variable: las dimensiones de la malla son mínimas sobre Francia y crecen progresivamente hasta sus antípodas donde son máximas. De esta forma, con un único modelo se puede hacer predicciones a escala fina sobre el territorio francés y describir con una precisión suficiente los grandes fenómenos de escala planetaria. En [9] se describe de forma detallada los modelos ARPEGE y ALADIN. Los otros modelos operativos existentes son similares a estos.

El modelo ARPEGE operativo es una versión de un modelo espectral de resolución variable sobre la horizontal y en diferencias finitas sobre la vertical, con una coordenada vertical híbrida $p - \sigma$ (presión y presión normalizada). La resolución horizontal máxima corresponde a una malla de 19,1 km de lado y la mínima a una de 234 km de lado. La atmósfera está dividida en 31 capas, su nivel más alto está en los 5 mbar de presión y el más bajo 20,5 metros por encima del suelo. El esquema numérico es semi-lagrangiano de dos niveles temporales y el paso de tiempo es de 900 segundos. Se tiene pues que la relación entre el tamaño de malla y el paso de tiempo es 21 m/seg, que probablemente es la mejor de todos los modelos operativos actuales. Las ecuaciones pronóstico están escritas para la presión en el suelo, las componentes horizontales de la velocidad, la temperatura y la humedad específica, y son similares a las que hemos descrito en la sección 4. Por último, este modelo operativo contiene un juego completo de parametrizaciones físicas que introducen los efectos de los fenómenos descritos en la sección 5.

Las observaciones para los datos iniciales se realizan intermitentemente, con 4 análisis por día: 0, 6, 12 y 18 h UTC (Coordinated Universal Time);

y se obtienen de observatorios en superficie, aviones (temperatura y viento), radiosondas y satélites.

Desde su inicio, el modelo ARPEGE fue concebido como multiuso, lo mismo permite efectuar predicciones a corto plazo, de hasta tres días, que a medio plazo, del orden de la semana. Ahora bien, su función más importante es la predicción general para un plazo de hasta dos o tres días. Más allá de tres o cuatro días los pronosticadores europeos se basan casi exclusivamente en las previsiones numéricas proporcionadas por el ECMWF. Hay que indicar que una predicción para un plazo de 48 horas con el modelo ARPEGE en el superordenador Fujitsu VPP700 (20 mil millones de operaciones por segundo) que Météo-France tiene en Toulouse costaba en 1998 un poco menos de un cuarto de hora.

Para previsiones a corto plazo (uno o dos días) se utilizan modelos de dominio limitado en los que la malla es lo suficientemente fina para simular correctamente las escalas pequeñas del movimiento, en particular las inducidas por el relieve. Este es el caso del modelo ALADIN, que se planteó como un complemento al modelo planetario ARPEGE para realizar predicciones a corto plazo (típicamente hasta 48 horas) en un dominio limitado (*cf.* [9]). En este modelo el dominio es un cuadrado de 2740 km de lado, en proyección Lambert, centrado en el punto de resolución máxima de ARPEGE. Los niveles verticales son los mismos que los de ARPEGE. La resolución horizontal de 9,9 km es casi el doble que la resolución máxima de ARPEGE y el paso de tiempo es de 470 segundos. Las partes dinámica y física son totalmente idénticas a las del modelo ARPEGE.

ALADIN sobre todo permite refinar los detalles del escenario previsto por ARPEGE para el mismo día y el día siguiente, salvo en el caso en que el escenario de ARPEGE se juzgue incorrecto. De hecho, la información de base de ALADIN (condiciones iniciales y condiciones de contorno) viene enteramente de ARPEGE y no contiene, por ejemplo, información analizada a una escala más fina. El refinamiento de los detalles viene de la mejor resolución horizontal y, en particular, de la mejor adaptación de los campos al relieve del suelo y a otros parámetros de superficie (vegetación, albedo, etc.), que están dos veces mejor descritos. Es natural que una evaluación de ALADIN realizada por pronosticadores de distintas regiones contenga mejores informaciones que ARPEGE en las zonas montañosas y cerca de las costas, debido a que en ALADIN el relieve y el contraste tierra-mar están descritos con más detalle.

7 Sensibilidad respecto de los datos iniciales

A pesar de las continuas mejoras de las predicciones meteorológicas realizadas con los modelos numéricos, hay que rendirse a la evidencia de la imposibilidad de proveer predicciones precisas para más allá de un cierto límite. La existencia de un límite para la previsibilidad es debida a varias razones: las imperfecciones de los modelos numéricos que simulan la atmósfera, las incertidumbres inherentes a las medidas efectuadas para determinar el estado inicial y el carácter no lineal

del modelo.

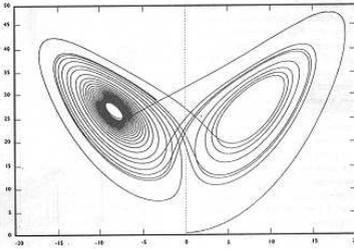
A comienzos de los años 1960 el meteorólogo americano Edward Lorenz se preguntaba: *¿Cómo es posible que conociendo las ecuaciones de la circulación atmosférica y las condiciones de partida, no se llegaba a predecir con un grado de fiabilidad aceptable el tiempo que haría tres días después?* Para tratar de responder a esta pregunta, Lorenz simplificó drásticamente las ecuaciones de la circulación atmosférica hasta llegar al sistema que lleva su nombre:

$$\frac{dx}{dt} + \sigma(x - y) = 0, \quad \frac{dy}{dt} + y - rx + xz = 0, \quad \frac{dz}{dt} + bz - xy = 0.$$

Para los parámetros Lorenz tomó los valores $\sigma = 10$, $b = 8/3$ y $r = 28$. Notemos que en el anexo D de [4] se detalla la obtención de estas ecuaciones de Lorenz a partir de las ecuaciones inerciales de la dinámica de la atmósfera, en concreto a partir del sistema de Bénard.

Es fácil adivinar que simplificado hasta este punto, este sistema no tenga ninguna utilidad para previsiones meteorológicas reales. Sin embargo posee los ingredientes necesarios para ser representativo de los movimientos atmosféricos y constituye el modelo teórico de caos determinista más célebre y más estudiado.

Lo mismo que los modelos meteorológicos completos, y a pesar de su aspecto tan sencillo, estas tres ecuaciones no tienen solución analítica. Por tanto hay que resolverlas por métodos numéricos utilizando un ordenador. Así lo hizo Lorenz y se topó con numerosas sorpresas.



Atractor de Lorenz

La primera fue que la evolución de cada una de las componentes de la solución tiene un comportamiento que parece que sólo obedece a la fantasía del azar.

El segundo descubrimiento fue que si dibujamos la sucesión de valores que en el transcurso del tiempo toman las soluciones, se obtiene una trayectoria que se enrolla sobre un curioso objeto de dos lóbulos. Se descubría así el primer “*atractor extraño*”.

La tercera sorpresa se produjo cuando comprobó que la introducción de un pequeñísimo error inicial crecía exponencialmente con el cálculo, de forma que se obtenía un resultado radicalmente diferente. Esto ponía en evidencia la sensibilidad del sistema de Lorenz respecto de las condiciones iniciales.

Los “verdaderos” modelos atmosféricos presentan la misma extrema sensibilidad respecto de los datos iniciales. Así, dos datos iniciales muy cercanos y cuyas diferencias son del orden de las incertidumbres inevitables derivadas de las imperfecciones de la red de observación, para plazos de tiempo superiores a un cierto límite pueden conducir a soluciones divergentes. Lorenz dio de este efecto una imagen muy impactante, que bautizó “efecto mariposa”: *una perturbación tan débil como el aleteo de las alas de una mariposa puede producir en la otra punta del Globo, un mes más tarde, un efecto considerable, como por*

ejemplo el desencadenamiento de un ciclón o lo contrario, el término de una tempestad.

Hay que resaltar que de forma general, el límite de previsibilidad está más lejos cuanto mayor es la escala de los fenómenos meteorológicos considerados. Así, para los fenómenos de escala sinóptica, es decir de una dimensión característica de 1000 a 2500 km y una duración de vida de uno a tres días, el límite para los modelos actuales es de tres o cuatro días.

Conocido el estado inicial y sus incertidumbres, para predicciones superiores a cuatro o cinco días actualmente se utiliza lo que se denomina previsiones de conjunto. Con estos métodos se pretende suministrar a los pronosticadores un número tan grande como sea posible de evoluciones diferentes de la atmósfera que se puedan derivar del estado inicial de referencia. Se tiene así la predicción probabilística, consistente en prever para un plazo dado la distribución de los valores previstos y las probabilidades asociadas. Para ello hay que calcular una función de probabilidad en cada uno de los puntos y para cada una de las magnitudes atmosféricas.

Una alternativa para acceder a la misma información es la utilización de un método Monte-Carlo, consistente en efectuar un cierto número de predicciones equiprobables y, a partir del conjunto de previsiones obtenidas, calcular los valores medios y los momentos estadísticos de orden superior. Se ha aplicado esta idea efectuando varias integraciones de un mismo modelo con distintos estados iniciales obtenidos haciendo en el estado inicial de referencia pequeñas modificaciones aleatoriamente repartidas y compatibles con los errores del análisis.

El problema fundamental de la predicción de conjunto consiste en efectuar una elección razonable de las variantes de la situación inicial de referencia, de manera que utilizando un mínimo de estados iniciales se pueda obtener un máximo de soluciones relativamente alejadas las unas de las otras. El empleo del modelo adjunto permite identificar las modificaciones del estado inicial que sufren las mayores amplificaciones para un plazo dado (en el sentido de una norma convenientemente elegida). Esta técnica es la utilizada por el ECMWF para efectuar una cincuenta de integraciones que conducen a un conjunto de previsiones para el cual es posible calcular la media y la varianza de las distintas magnitudes meteorológicas. Señalemos que en [2] se describe los métodos de interpretación de las predicciones numéricas a medio plazo adoptadas por los servicios de previsión de Météo-France, que son similares a las de otros servicios operativos. Para un estudio más exhaustivo ver su bibliografía.

8 Conclusión

En sus inicios, los hombres del tiempo pudieron considerar los modelos numéricos como algo anecdótico capaz de reproducir con mayor o menor fidelidad las características de la circulación atmosférica. Sin embargo, gracias a su constante mejora, actualmente son una herramienta imprescindible para la predicción. Hoy en día nadie duda de que los modelos numéricos realizan

mucho mejor que el hombre la síntesis de todos los procesos de interacción física, que antes sólo eran parcialmente comprendidos. Ahora bien, aunque estos modelos suelen dar previsiones realistas, no siempre dan una predicción exacta y a veces aparecen importantes errores, incluso para predicciones a muy corto plazo. Esto significa que los resultados de los modelos numéricos deben ser continuamente confrontados con la realidad: a posteriori para intentar corregir sus defectos sistemáticos, y en tiempo real para detectar a tiempo divergencias en las simulaciones que sean susceptibles de conducir a una estimación errónea de los fenómenos meteorológicos implicados.

Para terminar, las palabras de Miguel Azevedo (1916-1965) en su discurso de ingreso en la Real Academia de Ciencias y Artes de Barcelona ([3]): “*La gente que, a pesar de todo, está dispuesta siempre a dejarse engañar muy a gusto, como dice un proverbio latino, seguirá creyendo en los intuitivos de esta ciencia, como también ocurre en Biología o en Medicina, y en las predicciones realizadas con recursos extremadamente simples. Forma parte de la naturaleza humana el querer saber, de modo inmediato, a qué atenerse en todos los aspectos que afectan a su ámbito vital, antes que el comprender las dificultades de cualquier problema que requiere una formación especializada; y adopta esta actitud incluso en cuestiones de mayor categoría cultural o humana, sin que nadie se asombre por ello.*”

Bibliografía

Artículos históricos

Bjerkness V., 1904: Das problem von der Wettervorhersage, betrachtet vom Standpunkt der Mechanik und der Physik. *Meteor. Zeitschrift*, 21, pp. 1- 7.

Charney J.G., Fjörtoft R. and von Neumann J., 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, vol. 2, pp. 237-254.

Courant R., Friedrichs K.O., Lewy H., 1928: Ueber die partiellen Differenzengleichungen der mathematischen Physik. *Math. Annalen*, 100, pp. 32-74.

Eliassen A., 1956: A procedure for numerical integration of the primitive equations of the two-parameter model of the atmosphere. *Scientific Report n° 4*, Department of Meteorology, UCLA.

Lorenz E., 1963: Deterministic non periodic flow. *Journal of Atmospheric Sciences*, vol. 20, n° 2, pp. 130-141.

Phillips N.A., 1951: A simple three-dimensional model for the study of large-scale extratropical flow patterns. *J. Meteorol. Soc.*, vol 8, pp. 381- 394.

Richardson L.F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, reprinted Dover, 1965.

Rossby C.G., 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *Journal of Marine Research* (Sears Foundation), pp. 38-55.

Generalidades. Historia de la Meteorología

Coiffier J., 2000: Un demi-siècle de prévision numérique du temps. *La météorologie*, 8^e série, n^o 30, juin 2000, (*Spécial Prévision numérique du temps*), pp. 11-31.

Dahan-Dalmedico A., 2001: History and Epistemology of Models: Meteorology (1946- 1963) as a Case Study. *Arch. Hist. Exact Sci.*, vol 55, pp. 395-422.

Hunt J.C.R., 1998: Lewis Fry Richardson and his contributions to Mathematics, Meteorology and Models of Conflict. *Annu. Rev. Fluid Mech.*, vol 30, xiii-xxxvi.

Lorenz E., 1993: *The Essence of Chaos*, University of Washington Press, Seattle. (Traducción en español *La esencia del caos*, Editorial Debate, Madrid, 2000).

Nebeker F., 1995: *Calculating the Weather. Meteorology in the 20th Century*. Academic Press, San Diego.

Referencias

- [1] Aguirre de Cárcer I. y Jaque F., 2001: *Introducción a la Meteorología Ambiental*. Ediciones de la Universidad Autónoma de Madrid, Cantoblanco, Madrid.
- [2] Atger F., 2000: La prévision du temps à moyenne échéance en France. *La météorologie*, 8^e série, n^o 30, juin 2000, (*Spécial Prévision numérique du temps*), pp. 61-86.
- [3] Azpiroz M., 1965: Predicciones del tiempo. *Memorias de la Real Academia de Ciencias y Artes de Barcelona*, vol. XXXVII, núm. 1.
- [4] Bergé P., Pomeau Y. et Vidal C., 1984: *L'ordre dans le chaos*. Hermann, Paris.
- [5] Coiffier J., 1996: *Eléments de prévision numérique du temps*. École Nationale de la Météorologie, Météo-France, Toulouse.
- [6] Dutton J.A., 1995: *Dynamics of Atmospheric Motion*. Dover Publications, New-York.
- [7] Holton J.R., 1979: *An Introduction to Dynamic Meteorology*. Academic Press, New-York. (Traducción en español *Introducción a la Meteorología Dinámica*. Instituto Nacional de Meteorología, Madrid, 1990).
- [8] Jacobson M.Z., 1999: *Fundamentals of Atmospheric Modeling*. Cambridge University Press, Cambridge.
- [9] Pailleaux J., Geleyn J.F. et Legrand E., 2000: La prévision numérique du temps avec les modèles Arpège et Aladin. Bilan et perspectives. *La*

météorologie, 8^e série, n^o 30, juin 2000, (*Spécial Prévision numérique du temps*), pp. 32-60.

- [10] Rabier F., Mahfouf J.F. et Klinker E., 2000: Une nouvelle technique d'assimilation des données d'observation au CEPMMT: L'Assimilation variationnelle quadridimensionnelle. *La météorologie*, 8^e série, n^o 30, juin 2000, (*Spécial Prévision numérique du temps*), pp. 87-101.
- [11] Triplet J.P. et Roche G., 1971: *Météorologie Générale*. Troisième édition 1986, École Nationale de la Météorologie, Toulouse. Réimpression 1996, Météo-France, Paris.
- [12] Washington W.M. and Parkinson C.L., 1986: *An Introduction to Three-Dimensional Climate Modeling*. University Science Books, Mill Valley, California.

Algunos problemas de control y diseño óptimo en sistemas gobernados por ecuaciones diferenciales parciales

C. CASTRO

Departamento de Matemática e Informática Aplicadas a la
Ingeniería Civil
ETSI Caminos, Canales y Puertos
Universidad Politécnica de Madrid

ccastro@caminos.upm.es

Resumen

Este artículo contiene una breve descripción de cuatro problemas diferentes relacionadas con el control y el Diseño Óptimo en sistemas gobernados por Ecuaciones Diferenciales Parciales. El primer problema que se presenta consiste en el análisis y control de sistemas híbridos compuestos por cuerdas y vigas acopladas con masas de Dirac. El segundo problema se refiere al estudio de la propagación de ondas en medios muy heterogéneos. El tercer problema es una adaptación del método de la homogeneización al diseño óptimo de un conocido problema en ingeniería nuclear. Por último, describimos una forma óptima de estabilizar las vibraciones de una cuerda vibrante.

Palabras clave: *Control, Sistemas Híbridos, Homogeneización, Diseño Óptimo*

Clasificación por materias AMS: *35B37, 35B27, 35L05, 35L30, 35P15*

1 Introducción

Este artículo presenta una breve descripción de los principales problemas que han constituido la actividad investigadora que he desarrollado durante los últimos años.

La motivación principal de nuestro trabajo es *el estudio de propiedades de control y optimización en sistemas mecánicos gobernados por ecuaciones diferenciales parciales*.

Fecha de recepción: 8 de Mayo de 2002

Concretamente hemos trabajado en cuatro tipos de problemas relacionados con este tema: análisis y control de sistemas híbridos, análisis y control de sistemas sobre medios muy heterogéneos, diseño óptimo mediante homogeneización y disipación óptima de una cuerda vibrante.

A continuación presentamos una breve descripción de los cuatro problemas en los que hemos trabajado y los resultados obtenidos. La última sección contiene un resumen de las principales conclusiones.

2 Análisis y control de sistemas híbridos

Se llaman sistemas híbridos a aquellos que acoplan estructuras flexibles de diferente naturaleza.

En los últimos años el estudio de este tipo de sistemas ha tenido un gran auge justificado por los numerosos problemas de ingeniería en los que suelen aparecer. Un ejemplo de estos problemas, que además dio lugar a los primeros modelos matemáticos de sistemas híbridos, es el estudio de las vibraciones de un satélite con una antena adosada. El objetivo a la hora de estudiar este sistema era en primer lugar analizar la interacción entre las dos estructuras (la antena y el satélite) para posteriormente actuar mediante un control, situado en el satélite, que permitiese anular las vibraciones no deseadas en ambas estructuras.

El modelo matemático más simple utilizado para estudiar este sistema consistió en una cuerda vibrante con una masa puntual en un extremo (ver [37] y [36]). Este modelo constituye una buena aproximación siempre que la longitud de la antena sea significativamente grande con respecto al diámetro del satélite. Se trata de un modelo lineal en el que se acoplan la ecuación de ondas para describir las vibraciones de la cuerda con la ley de Newton para las vibraciones de la masa.



Figura 1: Sistema cuerpo-cuerda elástica.

Un modelo más complejo, pero a la vez más realista, del mismo problema consiste en describir las vibraciones de la antena mediante la ecuación de vigas. Este modelo, que también conduce a un sistema lineal de ecuaciones, fué estudiado en [34] y [35].

En todos estos trabajos se abordó el problema de la controlabilidad del sistema cuando se actúa sobre la masa puntual.

Otro problema interesante en el que aparece un sistema híbrido es el estudio de la propagación y control de ruido en cavidades con paredes elásticas. Es el caso, por ejemplo, de la cabina de un avión que constituye un medio con un alto nivel de ruido rodeado de una estructura elástica. El modelo matemático más sencillo para estudiar este tipo de problemas fué introducido por H. Banks, W. Fang, R.J. Wilcox and R.C. Smith [7] y consiste en un fluido contenido en un cuadrado en el que uno de sus lados es elástico. Más tarde, S. Micu y

E. Zuazua estudiaron rigurosamente la controlabilidad y estabilización de este sistema cuando se actuaba en diferentes partes del dominio (ver [38] y [39]) y dieron una descripción completa del espectro del problema a altas frecuencias.

La dificultad principal a la hora de estudiar este tipo de sistemas híbridos es la interacción entre las distintas estructuras que lo forman. Esta interacción constituye un problema difícil y que aún no se comprende bien.

Nuestro estudio se ha centrado fundamentalmente en modelos matemáticos unidimensionales que nos han permitido analizar con detalle la interacción entre ciertos tipos de estructuras muy particulares y deducir propiedades de control. A pesar de que la particularidad de los problemas considerados hace que las técnicas utilizadas sean a veces difícilmente transportables a marcos más generales, creemos que algunos de los comportamientos detectados son significativos y de carácter genérico, es decir, seguirán presentes en otros modelos más complejos en este campo.

El punto de partida fué un trabajo realizado por S. Hansen y E. Zuazua sobre un modelo híbrido compuesto por una cuerda vibrante con una masa puntual en su interior (ver [28]). En dicho trabajo se muestra que la interacción entre la masa y la cuerda produce el siguiente fenómeno sorprendente: *El sistema está bien puesto en espacios de diferente regularidad a ambos lados de la masa sin que exista una energía natural asociada.* Más concretamente, si consideramos un dato inicial con un orden de regularidad mayor a uno de los lados de la masa entonces la solución mantiene esa propiedad a lo largo del tiempo.



Figura 2: Sistema cuerda-cuerpo-cuerda. Hemos estudiado la influencia del cuerpo en las vibraciones del sistema

La interpretación física de este hecho es la siguiente: cuando una vibración llega a la masa se descompone en dos nuevas vibraciones: una de ellas se refleja manteniendo la regularidad inicial mientras que otra parte se transmite al otro lado ganando un orden de regularidad. Este efecto regularizante es debido a la presencia de la masa ya que, como es bien conocido, en una cuerda las vibraciones se propagan a través de las características sin ningún efecto regularizante.

La demostración de este resultado está basada en una construcción explícita de las soluciones mediante la fórmula de D'Alambert que se verifica a ambos lados de la masa. Esta fórmula tan sólo sirve para construir soluciones de la ecuación de ondas y esto hace que no se pueda generalizar el resultado a sistemas gobernados por otras ecuaciones diferenciales.

El modelo introducido en [28] para el sistema híbrido compuesto de dos

cuerdas unidas por una masa es el siguiente:

$$\begin{cases} u_{tt} = u_{xx}, & \text{en } -1 < x < 0, t > 0 \\ u_{tt} = u_{xx}, & \text{en } 0 < x < 1, t > 0 \\ [u(0, t)]_{x=0} = 0 & \text{en } t > 0 \\ [u_x(0, t)]_{x=0} = u_{tt}(0, t) & \text{en } t > 0 \\ u(-1, t) = u(1, t) = 0, & \text{en } t > 0, \end{cases} \quad (1)$$

donde $[u(x, t)]_{x=0}$ denota el salto de la función en $x = 0$, es decir, $[u(x, t)]_{x=0} = u(0^+, t) - u(0^-, t)$. En (1) la función $u(x, t)$ mide los desplazamientos verticales de las cuerdas situadas en los intervalos $(-1, 0)$ y $(0, 1)$, y $u(0, t)$ mide el desplazamiento vertical de la masa, situada en $x = 0$. Observemos que las dos primeras ecuaciones modelizan el movimiento de las cuerdas mientras que la tercera y cuarta ecuaciones representan las condiciones de transmisión de energía entre las cuerdas y la masa.

El estudio de este modelo tenía dos objetivos concretos. En primer lugar pretendíamos justificarlo desde el punto de vista matemático. Los modelos sobre sistema híbridos aparecidos hasta el momento se deducían a partir de consideraciones físicas o argumentos matemáticos formales. Para cumplir este objetivo, estudiamos el sistema híbrido como límite de un conjunto de sistemas S_ϵ formados por tres cuerdas vibrantes conectadas entre sí, en los que la cuerda central es de longitud ϵ y densidad $1/\epsilon$. Los sistemas aproximados que consideramos son:

$$\begin{cases} u_{tt} = u_{xx}, & \text{en } -1 < x < -\epsilon, t > 0 \\ \frac{1}{2\epsilon}u_{tt} = u_{xx}, & \text{en } -\epsilon < x < \epsilon, t > 0 \\ u_{tt} = u_{xx}, & \text{en } \epsilon < x < 1, t > 0 \\ u(\pm\epsilon^-, t) = u(\pm\epsilon^+, t), & \text{en } t > 0 \\ u_x(\pm\epsilon^-, t) = u_x(\pm\epsilon^+, t), & \text{en } t > 0 \\ u(-1, t) = u(1, t) = 0, & \text{en } t > 0. \end{cases} \quad (2)$$

De esta forma, en el límite cuando $\epsilon \rightarrow 0$ obtenemos una densidad con una delta de Dirac, es decir, una masa puntual.

Para estudiar la convergencia de los problemas aproximados al problema límite utilizamos dos métodos diferentes: El primero de los métodos consiste en definir las soluciones introduciendo las formulaciones débiles (soluciones en sentido de las distribuciones) y después pasar al límite en dichas formulaciones. La convergencia de las soluciones resulta un problema delicado debido a que la densidad de la cuerda central en los problemas aproximados no está acotada uniformemente cuando $\epsilon \rightarrow 0$.

El segundo método consiste en definir las soluciones mediante su desarrollo en series de Fourier y después pasar al límite en las series. Este último método requiere un detallado conocimiento del espectro de los problemas aproximados y su dependencia con respecto al parámetro $\epsilon \rightarrow 0$. En este sentido, los métodos clásicos de perturbación de operadores lineales permiten encontrar la tasa de convergencia de un número finito fijo de autovalores y autofunciones, lo que resulta insuficiente para probar la convergencia de las series de Fourier. Un

cálculo explícito del espectro del problema proporciona mejores resultados y se puede establecer la convergencia uniforme de los autovalores λ_k^ϵ y autofunciones ϕ_k^ϵ para un número finito $k \leq C\epsilon^{-1/6}$ pero que crece cuando $\epsilon \rightarrow 0$. Esto permite obtener la convergencia de las series de Fourier. Al mismo tiempo se pone de manifiesto que el sistema límite sólo refleja el comportamiento de la parte de las soluciones con frecuencias por debajo de la tasa $\epsilon^{-1/6}$. El comportamiento de la parte de las soluciones asociada a frecuencias por encima de este rango no se conoce bien. Como veremos en la próxima sección, este tipo de soluciones puede dar lugar a patologías muy sorprendentes.

Este análisis permite convencerse de que el sistema híbrido (1) refleja fielmente el comportamiento límite de los sistemas aproximados, al menos en lo que respecta a las "bajas" frecuencias (ver [10]).

El segundo objetivo que nos planteamos consistió en dar una demostración del fenómeno descrito por Hansen y Zuazua en [28] basada en propiedades espectrales y que no utilizase la fórmula de D'Alambert. Con ello pretendíamos deducir el mismo fenómeno en otros modelos gobernados por ecuaciones diferenciales diferentes de la ecuación de ondas, como por ejemplo las ecuaciones de vigas.

El análisis realizado demostró que las propiedades espectrales que caracterizan el fenómeno descrito son por un lado la falta de separación entre las raíces cuadradas de dos autovalores consecutivos, es decir,

$$\inf_k (\sqrt{\lambda_{k+1}} - \sqrt{\lambda_k}) = 0, \quad (3)$$

y por otro lado un comportamiento singular de las autofunciones que permite establecer combinaciones lineales de las mismas con una apariencia asimétrica (ver Figura 3).

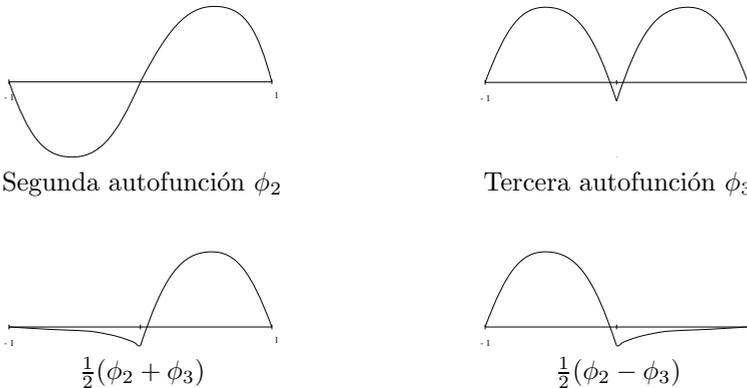


Figura 3: Se observa el comportamiento singular de la tercera autofunción que permite encontrar combinaciones lineales sencillas con un comportamiento asimétrico respecto de la masa.

La dificultad principal radica en la caracterización de los espacios con diferente regularidad a ambos lados de la masa usando sólo la propiedad

espectral descrita. Este problema fué resuelto mediante una técnica basada en la construcción de una base de Riesz para estos espacios asimétricos generada por una perturbación de ciertas combinaciones lineales de autofunciones como las presentadas en la Figura 3.

Posteriormente ensayamos con éxito este método en el problema de la viga con una masa en el interior. Para modelizar las vibraciones de la viga hemos considerado dos modelos lineales diferentes: el modelo de Timoshenko y el modelo de Euler-Bernouilli.

El modelo de Timoshenko incorpora en la ecuación un término de inercia de rotación que hace que el sistema sea hiperbólico:

$$\left\{ \begin{array}{ll} \gamma u_{xxtt}(x, t) - u_{tt}(x, t) - u_{xxxx}(x, t) = 0 & x \in (-1, 0) \cup (0, 1), t > 0 \\ [u(x, t)]_{x=0} = [u_x(x, t)]_{x=0} = 0 & t > 0 \\ u_{tt}(0, t) + [u_{xxx}(x, t)]_{x=0} = 0 & t > 0 \\ u_{ttx}(0, t) - \gamma^{-1}[u_{xx}(x, t)]_{x=0} = 0 & t > 0 \\ u(-1, t) = u(1, t) = 0 & t > 0 \\ u_{xx}(-1, t) = u_{xx}(1, t) = 0 & t > 0. \end{array} \right. \quad (4)$$

La constante $\gamma > 0$ es la constante de inercia de rotación. Cuando consideramos este modelo, la situación es análoga a la que se produce en el caso de la cuerda con una masa en el interior, es decir, es posible probar a partir de propiedades espectrales que el sistema está bien puesto en espacios con un orden de regularidad mayor a uno de los lados de la masa (ver [16]).

La situación es muy diferente cuando consideramos el modelo (más clásico) de Euler-Bernouilli que corresponde al caso $\gamma = 0$ en (4). Entonces, el sistema deja de ser hiperbólico y las vibraciones se propagan con velocidad infinita. La singularidad que produce la presencia de la masa no es suficiente para reducir la separación espectral asintótica de manera que se verifique la propiedad espectral (3). Este hecho permite deducir, en particular, que no existe ningún efecto regularizante de las vibraciones al atravesar la masa (ver [17]).

Por último, hemos aplicado los resultados anteriores al estudio de la controlabilidad de los sistemas descritos cuando actuamos sobre uno de los extremos. Brevemente, la controlabilidad de un sistema consiste en buscar una fuerza externa, denominada control, que actúe sobre una parte del sistema y que permita conducirlo a un estado deseado como puede ser el equilibrio. Por ejemplo, consideremos el sistema (4) con un control $q(t)$ que actúa en el extremo $x = 1$:

$$\left\{ \begin{array}{ll} \gamma u_{xxtt}(x, t) - u_{tt}(x, t) - u_{xxxx}(x, t) = 0 & x \in (-1, 0) \cup (0, 1), t > 0 \\ [u(x, t)]_{x=0} = [u_x(x, t)]_{x=0} = 0 & t > 0 \\ u_{tt}(0, t) + [u_{xxx}(x, t)]_{x=0} = 0 & t > 0 \\ u_{ttx}(0, t) - \gamma^{-1}[u_{xx}(x, t)]_{x=0} = 0 & t > 0 \\ u(-1, t) = u(1, t) = u_{xx}(-1, t) = 0 & t > 0 \\ u_{xx}(1, t) = q(t) & t > 0. \end{array} \right. \quad (5)$$

El problema de control al equilibrio puede formularse como sigue: *Dado un tiempo $T > 0$ y unos datos iniciales u_0, u_1 encontrar un control $q(t)$ tal que la*

solución del sistema (5) que satisface las condiciones iniciales

$$u(x, 0) = u_0, \quad u_t(x, 0) = u_1$$

verifique

$$u(x, T) = u_t(x, T) = 0.$$

Naturalmente la respuesta al problema de control planteado dependerá en gran medida de los espacios donde se tomen tanto el control q como los datos iniciales. Para fijar ideas supondremos que tomamos los controles en $L^2(0, T)$ y los datos iniciales en el mayor espacio para el cual existe un control $q \in L^2(0, T)$. Entonces, el problema de control se reduce a calcular este espacio de *datos iniciales controlables*.

Para estudiar la controlabilidad hemos aplicado el conocido método HUM (Hilbert Uniqueness Method) introducido por J.-L. Lions en los años 80 (ver [33]). HUM es un método de dualidad que reduce la controlabilidad de un sistema al estudio de ciertas desigualdades de observabilidad para el sistema adjunto sin control.

En general, la dificultad principal del método es probar las desigualdades de observabilidad. En el caso que estamos considerando, la controlabilidad del sistema (5) se reduce a la siguiente desigualdad para las soluciones del problema adjunto sin control (4):

$$\|(u_0, u_1)\|_F^2 \leq Cte \int_0^T |u_x(1, t)|^2 \quad (6)$$

donde F (que se define como el espacio más grande para el que se verifica (6)) es el espacio de *datos iniciales observables*.

Observamos que una desigualdad de observabilidad es una desigualdad que estima la energía total del sistema, medida en alguna norma, con la información que se puede obtener de una parte del dominio (zona de observación). En la desigualdad (6) la zona de observación es el extremo $x = 1$.

Debido a la existencia de espacios asimétricos en los que el sistema está bien puesto, el espacio F de datos observables para el sistema (4) es un espacio formado por funciones que tienen un orden más de regularidad a la izquierda de la masa. Aplicando la dualidad del método HUM se deduce inmediatamente que el espacio de datos controlables para el sistema (5) está formado por funciones que tienen un orden de regularidad menor a la izquierda de la masa. Naturalmente este efecto es debido a la presencia de la masa. En ausencia de la masa, los espacios de datos iniciales observables y controlables son completamente simétricos.

La prueba de estos resultados se realiza con una técnica basada en el desarrollo en serie de Fourier de las soluciones. Esta técnica precisa un detallado análisis de las autofunciones y algunos resultados sobre series de Fourier no armónicas (ver [14], [15], [16] y [17]). El principal inconveniente es que, por lo general, requiere que el espectro del problema esté compuesto por autovalores simples, y esto hace que se aplique casi exclusivamente en problemas unidimensionales.

3 Propagación de energía en medios muy heterogéneos

Entendemos por medios heterogéneos aquéllos que están formados por diferentes tipos de materiales, o por un mismo material pero con diferentes densidades. Un especial interés tienen aquellos medios heterogéneos que presentan una estructura periódica con un periodo ϵ muy pequeño en comparación con el tamaño del dominio que se considera. De esta forma se crean modelos de materiales con microestructura que tienen una apariencia homogénea y en los que la heterogeneidad y la estructura periódica del medio sólo son observables a un nivel microscópico.

Analizar el comportamiento de las soluciones teniendo en cuenta las leyes de comportamiento de cada uno de los materiales que originan la estructura sería un proceso muy laborioso y muchas veces impracticable. Por ello, desde el punto de vista matemático se ha desarrollado toda una teoría en torno a este tipo de estructuras conocida por "homogeneización". Con ella se pretende describir las leyes que rigen el comportamiento macroscópico de estos sistemas.

En los últimos años estos modelos heterogéneos y la teoría de la homogeneización para describirlos han sido objeto de estudios exhaustivos. Véase por ejemplo las monografías de Bensoussan - Lions - Papanicolau [8] y Oleinik - Shamaev - Yosifian [41].

Aunque la teoría de la homogeneización nos ha permitido avanzar mucho en la comprensión de la propagación de energía en este tipo de medios existen aún muchos problemas interesantes por resolver. Voy a destacar dos preguntas interesantes en este sentido:

1. ¿Cómo afecta la estructura microscópica de un material compuesto en su comportamiento macroscópico?, o más concretamente, ¿cuándo son válidas las leyes macroscópicas obtenidas por la homogeneización?
2. La homogeneización permite aproximar las soluciones de una ecuación sobre un medio muy heterogéneo. ¿Es esto suficiente para describir la distribución de energía de una solución en cualquier instante de tiempo? Si no es así, ¿Cuál es la ley de propagación de energía en estructuras muy heterogéneas compuestas de diferentes materiales?

Ambas cuestiones se presentan en muchos diseños de construcciones complejas actuales en los que un modelo macroscópico que no tenga en cuenta la interacción entre los diferentes materiales puede producir resultados erróneos.

La respuesta a las cuestiones planteadas no se conoce bien y tan sólo se han hecho estudios en casos muy particulares que además han arrojado resultados sorprendentes. En nuestro análisis hemos intentado profundizar lo más posible en este tipo de problemas manejando los modelos más sencillos.

A continuación desarrollo brevemente nuestro trabajo en este campo:

1. Acerca de la influencia de la estructura microscópica de un material compuesto en su comportamiento macroscópico

Uno de los aspectos que mejor describen la propagación de energía y que está estrechamente relacionado con la controlabilidad es la observabilidad, y en

pocas palabras establece la relación entre la energía que posee un sistema y aquélla que uno es capaz de medir desde una reducida zona del mismo durante un periodo de tiempo $T > 0$ (ver Figura 4). Concretamente, si consideramos la ecuación de ondas sobre un dominio $\Omega \subset \mathbb{R}^n$

$$\begin{aligned} \rho(x)u_{tt} - \operatorname{div}(A(x)\nabla u) &= 0, & x \in \Omega, & 0 < t < T, \\ u(x, t) &= 0, & x \in \partial\Omega, & 0 < t < T, \end{aligned} \quad (7)$$

y $\omega \subset \Omega$ una región de observación, la observabilidad consiste en encontrar desigualdades del tipo

$$\int_{\Omega} [\rho(x)|u_t(x, 0)|^2 + |\nabla u(x, 0)|^2] dx \leq Cte \int_0^T \int_{\omega} [A(x)|\nabla u(x, t)|^2] dx dt, \quad (8)$$

para toda solución u de (7). Observamos que a la izquierda de esta desigualdad aparece la energía de las soluciones de (7) y a la derecha, cierta información obtenida sólo de la zona de observación ω durante un periodo de tiempo $T > 0$.

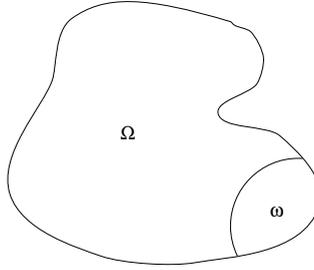


Figura 4: La observabilidad consiste en estimar la energía de un sistema mediante la que se puede observar en una pequeña región ω a lo largo de un cierto tiempo $T > 0$.

A la hora de estudiar este tipo de propiedades, las leyes macroscópicas que proporciona la homogeneización no son suficientes. Por ejemplo, se sabe que la interacción entre las ondas de alta frecuencia y la microestructura puede producir soluciones que concentran toda su energía en una parte de la frontera (ver [6] y los trabajos de G. Allaire y C. Conca [1]), es decir, se produce un cierto fenómeno de resonancia de las ondas con la microestructura.

Para ilustrar este hecho consideremos el modelo más simple formado por la ecuación de ondas unidimensional sobre un medio de densidad muy heterogénea con estructura periódica. Las ecuaciones que describen el sistema son:

$$\begin{cases} \rho^\epsilon(x) \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} & x \in (0, 1), \quad t > 0 \\ u(x, 0) = u^0(x), \quad \frac{\partial u}{\partial t}(x, 0) = u^1(x) & t > 0, \\ u(0, t) = u(1, t) = 0 & t > 0, \end{cases} \quad (9)$$

donde $\rho^\epsilon(x) = \rho(x/\epsilon)$ con $\rho(x)$ una función periódica tal que $0 < \rho_m \leq \rho(x) \leq \rho_M < \infty$, $\epsilon \rightarrow 0$ un parámetro pequeño que mide el periodo de la microestructura.

El problema de autovalores asociado a (9) posee autofunciones que se concentran en los extremos. En la Figura 5 se muestra el perfil de una de estas autofunciones ϕ que se concentra exponencialmente en $x = 0$. Estas autofunciones permiten construir soluciones de (9) en variables separadas del tipo

$$u(x, t) = \cos(\sqrt{\lambda t})\phi(x),$$

que mantienen el perfil de ϕ a lo largo del tiempo. De esta forma se construyen soluciones de la ecuación de ondas (9) que están muy concentradas en el extremo $x = 0$ a lo largo del tiempo. Por supuesto, estas soluciones son muy difíciles de *observar* desde el extremo $x = 1$ y, en realidad, debido a ellas la mejor constante en la desigualdad de observabilidad (8) para el problema (9) es del orden $\exp(1/\varepsilon)$, que explota cuando $\varepsilon \rightarrow 0$.

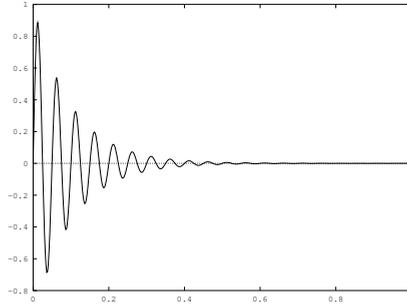


Figura 5: Comportamiento singular de algunas autofunciones de la ecuación de ondas unidimensional en un medio periódico. Se observa concentración de la energía en un extremo.

Observemos que este comportamiento singular de algunas soluciones de (9) se pierde en el problema límite. De hecho, es fácil deducir que las soluciones de (9) convergen a las del problema límite

$$\begin{cases} \bar{\rho} \frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u(x,t)}{\partial x^2} & x \in (0, 1), \quad t > 0 \\ u(x, 0) = u^0(x), \quad \frac{\partial u}{\partial t}(x, 0) = u^1(x) & \\ u(0, t) = u(1, t) = 0 & t > 0, \end{cases} \quad (10)$$

en donde la densidad oscilante se sustituye por su media $\bar{\rho} = \int_0^p \rho(x) dx$, con p el periodo de ρ . En este sistema las singularidades se propagan a través de las características con velocidad finita y no pueden existir soluciones que concentren su energía en el extremo $x = 0$ a lo largo del tiempo. La desigualdad de observabilidad (8) se cumple para el problema límite (10).

Vemos, por tanto, como el estudio del problema límite o problema homogeneizado se muestra insuficiente para detectar ciertos fenómenos como la localización de energía de las soluciones.

El análisis de este tipo de situaciones singulares constituyen la motivación principal de nuestro trabajo.

Un cuidadoso análisis espectral del sistema (9) permite acotar el rango de frecuencias que producen un efecto de resonancia con el medio. Se trata de aquellas frecuencias cuya longitud de onda coincide con la periodicidad del medio (ver Figura 6).

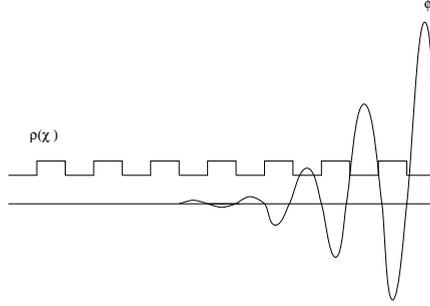


Figura 6: Sólo las autofunciones ϕ que oscilan con el mismo orden que la densidad ρ presentan fenómenos de resonancia con el medio.

Concretamente, las frecuencias mayores o menores que este rango crítico no muestran estos fenómenos de resonancia.

Para explicar este resultado introducimos el problema de autovalores asociado a (9):

$$\begin{cases} \varphi''(x) + \lambda \rho^\epsilon \varphi(x) = 0, \\ \varphi(0) = \varphi(1) = 0. \end{cases} \quad (11)$$

Para cada $\epsilon > 0$, existen un conjunto de autovalores

$$0 < \lambda_1^\epsilon < \lambda_2^\epsilon < \dots < \lambda_k^\epsilon < \dots \rightarrow \infty \text{ cuando } k \rightarrow \infty$$

y un conjunto de autofunciones asociadas $\{\phi_k^\epsilon\}$.

El sistema límite de los sistemas (11) cuando $\epsilon \rightarrow 0$ es el siguiente:

$$\begin{cases} \varphi''(x) + \lambda \bar{\rho} \varphi(x) = 0, \\ \varphi(0) = \varphi(1) = 0, \end{cases} \quad (12)$$

en el que se sustituye la densidad oscilante ρ^ϵ por su media $\bar{\rho}$. Los autovalores $(\lambda_k)_{k \in \mathbb{N}}$ y autofunciones asociadas $(\phi_k)_{k \in \mathbb{N}}$ pueden calcularse explícitamente

$$\begin{cases} \lambda_k = \frac{k\pi}{\sqrt{\bar{\rho}}}, \\ \phi_k(x) = \text{sen}(k\pi x). \end{cases}$$

Observemos que las autofunciones del problema límite no exhiben ninguna concentración en ninguna parte del dominio.

Se sabe que los autovalores y autofunciones de los problemas (11) convergen a los del problema límite (12) cuando $\epsilon \rightarrow 0$. Concretamente, se tiene el siguiente resultado (ver [41])

$$\begin{cases} |\lambda_k^\epsilon - \lambda_k| \leq c_1(k)\epsilon \\ \|\phi_k^\epsilon - \phi_k\|_{H_0^1(0,1)} \leq c_2(k)\epsilon. \end{cases} \quad (13)$$

Puesto que las autofunciones del problema límite no presentan ninguna concentración, está claro que aquellas autofunciones que se encuentren suficientemente cerca de las del problema límite no presentarán tampoco ninguna concentración. Resulta por tanto esencial conocer bien las constantes $c_1(k)$ y $c_2(k)$ en (13).

Para estudiar esta dependencia realizamos un desarrollo asintótico de los autovalores y autofunciones en función de k y ϵ con el conocido método WKB. Esto permite analizar la validez del sistema límite como aproximación de los sistemas (11). Concretamente, se obtiene

$$\begin{aligned} |\lambda_k^\epsilon - \lambda_k| &\leq c_1 k^{3/2} \epsilon \\ \|\phi_k^\epsilon - \phi_k\|_{H_0^1(0,1)} &\leq c_2 k \epsilon. \end{aligned} \quad (14)$$

Observamos que los autovalores del problema límite constituyen una buena aproximación de los autovalores λ_k^ϵ si $k \leq c\epsilon^{-2/3}$. Por encima de este rango, es necesario introducir correctores a los autovalores límite para obtener una aproximación uniforme. La necesidad de introducir correctores para encontrar mejores aproximaciones de los autovalores ya se conocía desde hace tiempo (ver [30]) pero un cálculo detallado del primer corrector tan sólo se produjo recientemente para un problema similar al que estamos tratando pero con el coeficiente oscilante en la parte principal del operador (ver [42]). El método WKB proporciona una fórmula explícita para un conjunto infinito de correctores (ver [21]). Por supuesto, cada corrector mejora la convergencia, sin embargo, aún con un número infinito de correctores, sólo es posible aproximar los autovalores λ_k^ϵ por debajo del rango crítico $k \leq c\epsilon^{-1}$ con c suficientemente pequeña.

En cuanto a las autofunciones, vemos que las correspondientes al problema límite constituyen una buena aproximación a las de los problemas aproximados cuando estamos por debajo del rango crítico $k \leq c\epsilon^{-1}$ con c suficientemente pequeña.

Esto permite deducir que por debajo del rango crítico no hay autofunciones que concentren energía en un extremo del dominio.

En el caso de las bajas frecuencias G. Lebeau ha extendido recientemente el resultado de la densidad oscilante al caso de varias dimensiones usando técnicas de análisis microlocal (ver [32]).

En cuanto a las altas frecuencias, es decir aquéllas que están por encima del rango crítico, también realizamos un análisis de su comportamiento (ver [22]). En este caso, se comprueba que si $\rho \in C^2$ es periódica, los autovalores y autofunciones $(\lambda_k^\epsilon, \phi_k^\epsilon)$ del problema (11) con $k \geq c\epsilon^{-2}$ y c suficientemente grande, pueden aproximarse por los del siguiente problema:

$$\begin{cases} \phi'' + \lambda \rho_\epsilon^* \phi = 0, \\ \phi(0) = \phi(1) = 0, \end{cases} \quad (15)$$

en el que la densidad oscilante se sustituye por la constante $\rho_\epsilon^* = \left(\int_0^1 \sqrt{\rho(x/\epsilon)} dx\right)^2$.

Los autovalores y autofunciones (λ_k^*, ϕ_k^*) de (15) también pueden calcularse explícitamente:

$$\begin{cases} \sqrt{\lambda_k^*} = \frac{k\pi}{\sqrt{\rho_\epsilon^*}}, \\ \phi_k^*(x) = \text{sen}\left(k\pi \frac{\int_0^x \sqrt{\rho(x/\epsilon)} dx}{\sqrt{\rho_\epsilon^*}}\right). \end{cases}$$

Podemos observar que las autofunciones del problema (15) no exhiben ninguna concentración de energía en los extremos del dominio y por tanto tampoco van a exhibirla las autofunciones ϕ_k^ϵ de (11) si $k \geq c\epsilon^{-2}$ con c suficientemente grande.

A medida que consideramos autofunciones ϕ_k^ϵ con k más cercano al rango crítico $k \sim \epsilon^{-1}$ se hace necesario introducir correctores para obtener aproximaciones uniformes. Cada corrector requiere una derivada más de la densidad ρ de manera que si $\rho \in C^{N+1}$, podemos encontrar una fórmula asintótica para los autovalores y las autofunciones de (11) por encima del rango $k \geq \epsilon^{-1-1/N}$. En particular, si $\rho \in C^\infty$ podemos deducir que las autofunciones de (11) no presentan ninguna concentración de energía si $k \geq c\epsilon^{-\alpha}$ con c suficientemente grande y $\alpha > 1$ cualquiera.

Nuevamente, este resultado está basado en un desarrollo asintótico de los autovalores y autofunciones mediante el método WKB similar al realizado para las bajas frecuencias.

Cabe destacar que el espectro de (11) tiene un comportamiento muy diferente para las frecuencias altas y bajas. Concretamente, mientras las bajas frecuencias se comportan como las de una ecuación con densidad constante $\bar{\rho}$, las altas frecuencias se comportan como las de una densidad $\rho_\epsilon^* \leq \bar{\rho}$. Además la igualdad sólo se da si ρ es constante. En consecuencia, las velocidades de propagación de las altas y bajas frecuencias no son iguales.

En el caso de las altas frecuencias, la complejidad del problema en varias dimensiones hace que existan muy pocos resultados en este contexto por el momento.

El caso crítico $k \sim \epsilon^{-1}$ es el más difícil de tratar y hasta el momento tan sólo hemos podido dar una descripción detallada cuando $\epsilon = 1/n$ con $n \in N$ (ver [12]). En él se mezclan autofunciones que se concentran en los extremos, autofunciones que se concentran en el interior y autofunciones que no presentan ningún tipo de concentración. Este caso también ha sido recientemente estudiado por G. Allaire y C. Conca [1], donde se propone un test para determinar las autofunciones que se concentran en una parte del dominio basado en ondas de Bloch. El test resulta válido sólo para $\epsilon = 1/n$ (ver [19]).

Finalmente, hemos aplicado los resultados espectrales al estudio de la controlabilidad del sistema (9) cuando actuamos en uno de los extremos (ver [20] que contiene resultados preliminares para las soluciones que contienen bajas frecuencias y [11] para un análisis global). Hemos probado que, si $\rho \in C^N$, las proyecciones de las soluciones sobre los subespacios generados por las autofunciones ϕ_k^ϵ con $k \leq c\epsilon^{-1}$ y $k \geq C\epsilon^{-1-1/N}$ verifican las desigualdades de observabilidad con constantes uniformes en ϵ . El método utilizado para la demostración, al igual que ocurre con el usado en los problemas híbridos

presentados en la sección anterior, está basado en el método HUM y dos propiedades espectrales:

- Separación uniforme de dos autovalores consecutivos, es decir

$$\lambda_{k+1}^\epsilon - \lambda_k^\epsilon \geq \alpha > 0.$$

- Las autofunciones no se concentran en un extremo.

Ambas propiedades se verifican para las altas y las bajas frecuencias pero dejan de verificarse en el rango crítico $k \sim \epsilon^{-1}$.

2. Acerca de la propagación de ondas en medios compuestos por materiales de diferente naturaleza

Consideramos ahora el problema de la observabilidad (y por tanto la controlabilidad) de sistemas gobernados por ecuaciones de ondas cuando los coeficientes son $C^{0,\alpha}$ con $0 < \alpha < 1$. Observemos que un caso particular de estos coeficientes, con regularidad incluso menor, lo constituyen las funciones constantes a trozos que modelizan estructuras compuestas de diferentes materiales.

En este caso la propagación de energía aún no se entiende bien y ni siquiera es posible establecer una definición de rayos como medio de propagación. Cabe mencionar que la regularidad mínima de los coeficientes para poder definir rayos que se propagen según las leyes de la óptica geométrica es C^2 (ver [9]).

Como muestra de las posibles patologías que pueden producirse se sabe que es posible construir coeficientes, con la regularidad arriba indicada ($C^{0,\alpha}$), para los cuales existen soluciones que se concentran en un entorno de un punto a lo largo del tiempo. Por tanto, dichas soluciones no exhiben la propagación finita de energía característica de las ecuaciones de ondas. En otras palabras, es posible construir rayos que permanecen atrapados en torno a un punto y que no pueden escapar de un entorno del mismo (ver [23]).

Existen muchas implicaciones interesantes de este resultado que puede extenderse al caso multidimensional o la ecuación de Schrödinger. Por ejemplo, es claro que cualquier ley de propagación que se enuncie para coeficientes poco regulares debe contener soluciones que representen este tipo de rayos inmóviles.

Otra implicación interesante se refiere a las conocidas propiedades de dispersión que se ponen de manifiesto con las desigualdades de Strichartz y que resultan ser falsas en nuestros ejemplos.

4 Optimización mediante homogeneización

En esta sección presentamos un problema de diseño óptimo mediante el método de la homogeneización.

La descripción general de este tipo de problemas es la siguiente: Se considera un dominio y un conjunto de materiales diferentes con una proporción establecida a priori. *El objetivo es encontrar una distribución adecuada de los*

materiales sobre el dominio (diseño) que optimice alguna propiedad física o mecánica.

Existen una gran variedad de problemas que se adaptan al esquema anterior. Entre ellos podemos mencionar el diseño de estructuras en los que se trata de encontrar aquélla que, con una cantidad de material fija, alcance la máxima resistencia. Se trata de un caso particular en el que se considera el esquema anterior con un sólo material que se combina con el vacío.

Es claro que el diseño de estructuras es un campo interesante con muchas aplicaciones industriales. El proceso habitual que se ha seguido hasta el momento en este campo es el ensayo sucesivo de prototipos basados en la intuición de los ingenieros. Sin embargo, cada vez más, esta técnica manual se está sustituyendo por técnicas de modelización matemática y optimización que permiten un análisis mucho más eficiente.

Los esquemas numéricos que se han ensayado están basados en una representación de un diseño inicial por un conjunto de parámetros descriptivos (generalmente controles en la frontera). A continuación, se realiza un análisis basado en variaciones de estos parámetros que permita mejorar la forma inicial.

Los principales problemas a la hora de implementar estos métodos son por un lado el costo excesivo y, lo que es más grave, una fuerte dependencia tanto del diseño inicial elegido como del conjunto de parámetros elegidos para hacer variaciones.

En los últimos años se ha desarrollado un nuevo método para abordar los problemas de diseño estructural basado en la homogeneización. El origen de este método se encuentra en los trabajos de F. Murat y L. Tartar de los años 70 (ver [40]) en los que introdujeron un método para resolver problemas de optimización mediante homogeneización. La adaptación definitiva del método al diseño de estructuras es más reciente (ver, por ejemplo, [2]). A continuación describimos brevemente los puntos más importantes de este método.

Como hemos dicho, el problema típico consiste en calcular la estructura que con una cierta cantidad de materia fija aguante la mayor fuerza externa en una o varias direcciones. Desgraciadamente, este problema no suele tener solución. Esto se debe al hecho de que mecánicamente resulta más efectivo hacer varios agujeros pequeños que uno grande. De esta forma, la configuración óptima resulta ser un material con infinitos agujeros que forman una microestructura. Desde el punto de vista numérico esta situación produce una fuerte dependencia de la solución con respecto a la fineza del mallado.

La forma de solucionar este problema consiste en introducir una extensión de la noción de estructura que incluya los materiales con microestructura. Esto se consigue a través de la homogeneización que permite caracterizar las estructuras a través de sus propiedades efectivas.

La nueva noción de estructura ya no queda determinada por una función característica que distingue los puntos en los que hay materia o agujero. En su lugar aparecen dos nuevos parámetros que son la densidad de materia en cada punto y la geometría de la microestructura.

La estructura óptimal estará compuesta principalmente por material compuesto, es decir con microestructura y se hace necesario buscar una forma

de encontrar la estructura más cercana que contenga sólo materia y agujeros, es decir sin microestructuras. Una aproximación adecuada no es difícil conseguir mediante un proceso de penalización.

Nuestro trabajo ha consistido en desarrollar una generalización del método expuesto para problemas de diseño óptimo más generales en los que, en particular, aparezcan estructuras compuestas por más de un material y sistemas de ecuaciones.

La motivación física parte de un problema ya clásico en el ámbito de la energía nuclear. Se trata del problema de la repartición de combustible en el núcleo de un reactor nuclear que describimos a continuación.

El núcleo de un reactor es un medio compuesto por 157 ensamblados diferentes pero con una estructura común (ver Figura 7). Cada ensamblado contiene básicamente un conjunto de 17×17 barras de combustible rodeadas por agua. Se trata por tanto de un medio tremendamente heterogéneo.

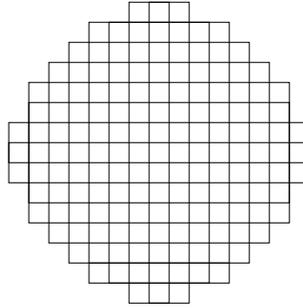


Figura 7: Estructura de la sección transversal del núcleo de un reactor con 157 ensamblados.

Para simplificar supondremos que los ensamblados son completamente homogéneos. Sin embargo, tomaremos las características físicas de cada ensamblado diferentes debido fundamentalmente al mayor desgaste del combustible en unos ensamblados que en otros.

Cada cierto tiempo se realiza un ajuste en la distribución de ensamblados en el que se remplazan aquellos que contienen el combustible más usado por otros nuevos. Desde el punto de vista técnico, es posible combinar este proceso de recambio con un proceso de permutación de ensamblados. El problema consiste en encontrar la distribución óptima de los ensamblados para que el flujo sea lo mayor y más uniforme posible. Esto hará que el rendimiento del reactor sea máximo sin peligro de que existan zonas localizadas donde el flujo se concentre creando una reacción incontrolada.

Lo ideal sería realizar el cálculo con todas las posibles configuraciones del núcleo que son $157!$ En la actualidad es imposible llevar a cabo un cálculo de esta naturaleza y por ello se han desarrollado procesos de optimización similares a los descritos para los problemas de diseño estructural en los que se hacen pequeñas

variaciones de una configuración inicial elegida. Estos métodos están aún lejos de conducir a la configuración óptima principalmente por su enorme dependencia de la configuración inicial y la dificultad para encontrar una configuración inicial lo bastante buena.

El método que hemos implementado está basado en el método descrito para el diseño de estructuras y tiene la particularidad que permite encontrar diseños óptimos que combinan más de dos materiales diferentes, a diferencia del diseño estructural en el que se combinan tan sólo materia y vacío.

La formulación matemática del problema es como sigue: Sea Ω un abierto de \mathbb{R}^2 y consideremos el sistema de ecuaciones:

$$\begin{cases} -\operatorname{div}(D(x)\nabla\phi) - \Sigma(x)\phi = \lambda\sigma(x)\phi & x \in \Omega, \\ \phi(x) = 0, & x \in \partial\Omega. \end{cases} \quad (16)$$

En (16) ϕ representa el flujo neutrónico, D es el coeficiente de difusión, Σ es coeficiente de absorción y σ el coeficiente de fisión. El proceso es el siguiente: Para conseguir calor se produce una reacción nuclear en la que el flujo neutrónico sufre cambios debidos a la fisión de los núcleos, la difusión del flujo y la absorción de neutrones por otros núcleos. Las ecuaciones (16) representan un modelo simplificado de este proceso en estado estacionario. El coeficiente λ es el valor mínimo para el cual existe solución de (16) y mide el equilibrio entre el aumento del flujo debido a la fisión y su disminución debida a la absorción y la difusión. Si $\lambda > 1$ el flujo tiende a decrecer y la reacción se extingue, mientras que si $\lambda < 1$ la reacción crece y puede hacerse incontrolada.

En la práctica, existen medios de reducir la reacción mediante mecanismos eficientes de control. Por ello, nuestro objetivo es maximizar el flujo, es decir, minimizar el autovalor λ . Por otro lado, la concentración de flujo puede ser peligrosa en este tipo de reacciones por lo que impondremos una restricción al flujo para que sea lo más uniforme posible. Por último, el conjunto sobre el que minimizaremos será el conjunto de coeficientes (D, Σ, σ) que representen una configuración admisible del núcleo (ver Figura 8).

Nuestro problema de diseño óptimo se reduce a minimizar el siguiente funcional:

$$\min_{(D, \Sigma, \sigma)} J(D, \Sigma, \sigma) = \lambda_1 + l \|\phi_1\| - \int_{\Omega} \phi_1 \|_{L^2(\Omega)} \quad (17)$$

donde λ_1 es el primer autovalor en (16) y ϕ_1 su autofunción asociada. Observamos que la condición sobre la uniformidad del flujo está incorporada en el funcional J mediante el multiplicador de Lagrange $l \in \mathbb{R}$ que constituye un parámetro desconocido.

El conjunto de coeficientes (D, Σ, σ) sobre el que se minimiza es el conjunto de todos aquellos coeficientes que representan una posible configuración del núcleo. Para simplificar vamos a suponer que sólo hay dos tipos de ensamblados que se encuentran caracterizados por las funciones $(d_i, \Sigma_i, \sigma_i)$ con $i = 1, 2$. De esta forma,

$$\begin{cases} D(x) = d^1\chi(x) + d^2(1 - \chi(x)), \\ \Sigma(x) = \Sigma^1\chi(x) + \Sigma^2(1 - \chi(x)), \\ \sigma(x) = \sigma^1\chi(x) + \sigma^2(1 - \chi(x)), \end{cases} \quad (18)$$

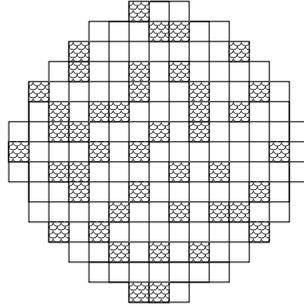


Figura 8: Posible configuración de ensamblados en un reactor.

donde $\chi(x)$ es la función característica que vale 1 en los puntos donde esta el ensamblado de tipo 1 y 0 en donde se encuentra el ensamblado de tipo 2.

Supondremos también que cada ensamblado homogéneo, es decir los coeficientes sobre cada ensamblado $(d_i, \Sigma_i, \sigma_i)$ son constantes. Por último, para que el problema no sea trivial, introducimos una restricción de volumen sobre los ensamblados, es decir $\int_{\Omega} \chi = \gamma_1$ constante.

Con estas consideraciones el problema de optimización que resolvemos es el siguiente

$$\begin{aligned} & \text{mín} && J(\chi) \\ & (\chi \text{ t.q. } \int_{\Omega} \chi = \gamma_1) \end{aligned} \tag{19}$$

Para resolver este problema, en una primera etapa introducimos una extensión del conjunto de posibles coeficientes a un conjunto continuo en el que no imponemos ninguna condición geométrica a los ensamblados. Matemáticamente, consideramos en (18) cualquier función característica χ con $\int_{\Omega} \chi(x) dx = \gamma_1$ dado, que mide la proporción de ensamblados del tipo 1 (ver Figura 9). Observemos que esta extensión transforma el espacio de configuraciones admisibles de un conjunto finito a un conjunto infinito.

Esta primera extensión no es suficiente porque nos conduce a un problema sobre un espacio no convexo, con multitud de mínimos locales, en el que es muy difícil implementar un método numérico para calcular el mínimo absoluto. Por ello, es necesario introducir una nueva extensión que contenga lo que en homogeneización se denomina medios compuestos con microestructura.

En la segunda etapa extendemos el funcional J a este nuevo conjunto de coeficientes valiéndonos de la teoría matemática de la homogeneización. Este es el proceso más delicado. La dificultad está en caracterizar todas las posibles microestructuras que se pueden obtener por combinación de varios materiales. En el caso de dos materiales este conjunto está bien caracterizado desde los años 60 (ver los trabajos de Hashin-Shtrikman [29]). En el caso de 3 o más materiales, este problema ha sido muy estudiado pero aún permanece abierto (ver [31] y [5]). Afortunadamente, no es necesaria una caracterización completa de estos materiales para resolver el problema de optimización y es posible

probar que trabajando con un conjunto más pequeño (materiales construidos por laminación) se puede alcanzar el óptimo.

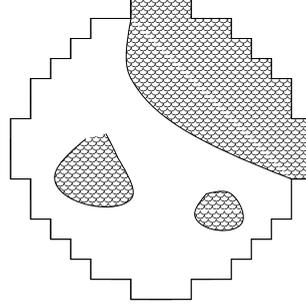


Figura 9: Posible configuración de los coeficientes tras la primera extensión que no respeta la geometría de los ensamblados.

En la tercera etapa efectuamos el proceso de minimización. El espacio sobre el que se minimiza es convexo y resulta relativamente sencillo encontrar condiciones de optimalidad para los mínimos.

Estas condiciones se implementan numéricamente sin dificultad. Para ello hemos utilizado el paquete informático PDETOOLBOX del programa matemático MATLAB que permite resolver una gran variedad de ecuaciones diferenciales parciales mediante el método de elementos finitos.

Una observación importante es que el mínimo no se corresponde, en general, con una configuración real del núcleo puesto que contiene zonas con microestructura. Por ello, en la cuarta etapa intentamos reconstruir una configuración real que se encuentre próxima del mínimo. Este proceso se realiza introduciendo una penalización en el funcional de manera que aquellas partes que contengan más densidad se saturen y las que tengan menos densidad se aproximen al vacío.

Con este método hemos resuelto el problema obteniendo una configuración óptima por laminación con cuatro tipos de ensamblados (ver [3]).

En general, no sabemos si el mínimo encontrado es o no único aunque todos los ensayos que hemos realizado nos proporcionan la misma configuración final.

Seguidamente hemos aplicado nuestro método al siguiente sistema más complejo:

$$\begin{cases} -\operatorname{div}(D_1(x)\nabla\phi) + \Sigma(x)\phi = \lambda(\sigma_1(x)\phi + \sigma_2(x)\psi) & x \in \Omega, \\ -\operatorname{div}(D_2(x)\nabla\psi) + \Sigma_2(x)\psi = \sigma_r(x)\phi & x \in \Omega, \\ \phi(x) = \psi(x) = 0, & x \in \partial\Omega. \end{cases} \quad (20)$$

En (20) aparecen dos tipos de flujos neutrónicos, ϕ y ψ , correspondientes a neutrones de velocidad alta y térmicos respectivamente. El modelo (20) se conoce como modelo de dos grupos y es más realista que el modelo (16).

En este caso, nuestro método tropieza con una dificultad matemática seria ya en la primera etapa. Los resultados conocidos sobre homogeneización no

permiten extender el funcional a todo el conjunto de configuraciones con microestructura porque este conjunto no está aún bien caracterizado al haber dos matrices en juego. Existen cotas optimales en ciertos casos (ver [24]) pero el problema general parece difícil de solucionar. Hemos salvado esta dificultad introduciendo una nueva hipótesis en el problema: hemos supuesto que las propiedades físicas de cada ensamblado son similares. Esta hipótesis es realista desde el punto de vista físico puesto que los ensamblados se rempazan antes de que el combustible se agote. Bajo esta nueva hipótesis es posible introducir un funcional aproximado que podemos extender a las configuraciones con microestructura (ver [4]).

5 Disipación óptima en una cuerda vibrante

En esta sección planteamos el siguiente problema: *encontrar la forma más eficaz de amortiguar las vibraciones de una cuerda vibrante sujeta en los extremos*. Para ello consideramos mecanismos pasivos, es decir mecanismos de disipación distribuidos sobre la longitud de la cuerda, como puede ser el rozamiento con un medio externo.

Sea $u(x, t)$ el desplazamiento de una cuerda vibrante de longitud 1, fija en sus extremos y en presencia de un amortiguamiento distribuido $2a(x)$. Se satisfacen las siguientes ecuaciones

$$\begin{aligned} u_{tt}(x, t) - u_{xx}(x, t) + 2a(x)u_t(x, t) &= 0, & 0 < x < 1, & 0 < t, \\ u(0, t) = u(1, t) &= 0, & 0 < t, \end{aligned} \quad (21)$$

junto con las ecuaciones del estado inicial

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = v_0(x). \quad (22)$$

Es bien conocido que dado un coeficiente positivo $a \in L^\infty(0, 1)$ la energía viene dada por

$$E(t) \equiv \int_0^1 u_x^2(x, t) + u_t^2(x, t) dx$$

y verifica $E(t) \leq CE(0)e^{2\omega t}$ con $C > 0$ y $\omega < 0$, independiente del dato inicial. Al menor de tales ω ,

$$\omega(a) \equiv \inf \left\{ \omega : \exists C(\omega) > 0 \text{ s.t. } E(t) \leq CE(0)e^{2\omega t}, \right. \\ \left. \text{para cada solución de energía finita de (1.1)} \right\}$$

se le conoce como **tasa de decaimiento** asociada a a .

Buscamos el coeficiente a que produzca la mínima tasa de decaimiento.

De forma intuitiva parece que la disipación constante $a = a_0$ es la mejor forma de amortiguar. Sin embargo, se observa que al elegir la disipación constante no parece cierto el hecho de que 'más es mejor'. Es decir, una mayor disipación en la cuerda no produce, en general, una mayor tasa de decaimiento. Concretamente, cuando $a = a_0$ constante, al aumentar a_0 hasta el valor π la

tasa de decaimiento decrece. Sin embargo, a partir de este valor crítico la tasa de decaimiento comienza a crecer rápidamente (ver por ejemplo [26]).

Este hecho sorprendente es conocido como *overdamping*. En la Figura 10 se ilustra este fenómeno tomando una cuerda vibrante sometida a un rozamiento constante en toda su extensión. Se puede apreciar como el fuerte rozamiento hace que el primer modo de vibración apenas se mueva y tarde un tiempo muy largo en amortiguarse.

A pesar de todo, durante mucho tiempo se pensó que la distribución constante de rozamiento a lo largo de la cuerda era la mejor forma de amortiguar. De esta forma, el fenómeno del *overdamping* proporcionaba un límite a la cantidad de rozamiento que debemos utilizar para obtener una tasa de decaimiento óptima. Apoyando esta conjetura S. Cox y M. Overton [25] mostraron además que la constante óptima es un mínimo local entre todas las posibles distribuciones de rozamiento. Sin embargo, P. Freitas ([27]) daba posteriormente algunos ejemplos numéricos que mostraban evidencias de disipaciones no constantes que mejoraban la disipación constante óptima.

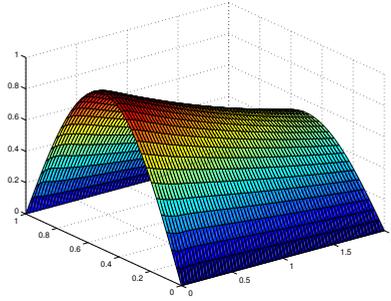


Figura 10: Desplazamiento $u(x, t)$ de una cuerda vibrante que ocupa el intervalo $(0, 1)$ a lo largo del tiempo $0 < t < 2$ con un fuerte rozamiento uniformemente distribuido.

Nuestro estudio resolvió de forma negativa esta conjetura (ver [13]). Existen formas mejores de amortiguamiento que la distribución constante de rozamiento. Concretamente se tiene el siguiente resultado:

Si $a_n(x) = 1/(x + 1/n)$ entonces $\omega(a_n) \rightarrow -\infty$ cuando $n \rightarrow \infty$.

Todavía más sorprendente es el hecho de que cuando $n \rightarrow \infty$, las disipaciones $a_n(x) \rightarrow 1/x \notin BV(0, 1)$ y en este caso las soluciones se extinguen en tiempo finito (ver Figura 11). Concretamente, si $a(x) = 1/x$, dado cualquier dato inicial $(u_0, u_1) \in H_0^1(0, 1) \times L^2(0, 1)$ la solución $u(x, t)$ de (21)-(22) satisface

$$u(x, t) = u_t(x, t) = 0, \quad \forall t > 2. \quad (23)$$

La demostración de estos resultados está basada en dos hechos fundamentales: Por un lado, la tasa de decaimiento puede caracterizarse mediante la abscisa espectral del problema de autovalores asociado a la ecuación

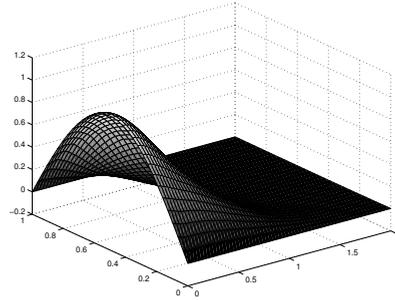


Figura 11: Amortiguamiento de una cuerda vibrante con un rozamiento localizado en el extremo izquierdo del tipo $a(x) = 1/x$.

amortiguada. En segundo lugar, un estudio cuidadoso del espectro muestra que la abscisa espectral se hace cada vez más negativa a medida que tomamos $a(x) = 1/(x+c)$ con $c > 0$ más pequeño. En el caso en que $c = 0$ un argumento directo basado en la transformada de Laplace muestra que no hay autovalores del problema asociado.

6 Conclusiones

Sistemas híbridos: Hemos visto que se pueden caracterizar en términos de series de Fourier espacios asimétricos en los que los modelos híbridos están bien puestos o, en su caso, ver que estos espacios no existen. Para ello es necesario un análisis riguroso del comportamiento asintótico del espectro a altas frecuencias y en particular, un análisis de la distancia espectral asintótica.

Además, hemos deducido algunas consecuencias correspondientes en lo que respecta a la propagación, observación y control de ondas en estos modelos.

El análisis de estos fenómenos en varias dimensiones espaciales o en modelos no lineales son temas mucho más complejos y sobre los que apenas hay resultados. Es de esperar que los fenómenos descritos tengan su análogo en varias dimensiones.

Propagación de ondas en medios heterogéneos: Hemos planteado el problema de la propagación de ondas en un medio muy heterogéneo con estructura periódica de periodo $\epsilon \rightarrow 0$ en una dimensión de espacio.

El análisis realizado proporciona una descripción detallada de algunas situaciones singulares que se producen. Hemos visto que existe un valor crítico ϵ , que coincide con el periodo de la estructura, de forma que las soluciones que contienen frecuencias con longitud de onda mayor o menor que ϵ no presentan ningún fenómeno de localización de energía. En cambio, aquellas soluciones que contienen frecuencias del orden de ϵ pueden producir resonancia con la estructura creando localización de energía en uno de los extremos del dominio.

Estos resultados se pueden aplicar para obtener condiciones observabilidad

y control, uniformes en ϵ , de las soluciones con frecuencias de longitud de onda mayor y menor que ϵ , es decir, aquellas que no presentan resonancia.

Por último hemos expuesto el tipo de patologías que se pueden esperar en el comportamiento de estructuras muy heterogéneas compuestas por diferentes materiales. En particular, hemos visto que pueden existir ondas que no se propaguen en ninguna dirección. Cabe destacar que un resultado similar en la ecuación del calor constituye un problema de gran interés que no se conoce.

Diseño óptimo mediante homogeneización: Hemos planteado un problema de diseño óptimo motivado por su aplicación dentro del campo de la energía nuclear.

El método que hemos utilizado para resolverlo está basado en un planteamiento similar en el diseño de estructuras. Este método se basa en el uso de la homogeneización para encontrar una adecuada relajación del problema de optimización que generalmente está mal puesto.

La aportación principal consiste en una generalización del método de la homogeneización a sistemas de ecuaciones no necesariamente autoadjuntos, y que permite combinar más de un material bajo la hipótesis adicional de que todos los materiales sean similares.

Disipación óptima en una cuerda vibrante: Hemos visto que existen formas de repartir la disipación en una cuerda vibrante que producen una tasa de decaimiento arbitraria. Incluso existe una disipación, que no está acotada, y que produce extinción en tiempo finito de las soluciones. Este hecho responde negativamente a la conjetura que establecía que la constante es la mejor forma de disipación en una cuerda vibrante.

El resultado está basado en un cálculo explícito. Sería interesante probar un resultado similar en dimensión mayor que uno y en otros sistemas similares como en la ecuación de vigas.

Agradecimientos

La mayor parte del trabajo presentado en este artículo es fruto de sucesivas colaboraciones con E. Zuazua (Director de Tesis), G. Allaire (Director de trabajo durante una estancia postdoctoral en el centro de investigación 'Commissariat à l'Energie Atomique' (CEA) en Francia) y S. Cox. Desearía agradecerles su acogida y disposición.

Referencias

- [1] G. Allaire y C. Conca, Analyse Asymptotique Spectrale de l'Equation des ondes. Homogénéisation par ondes de Bloch, C.R. Acad. Sci. Paris, **321**, Série I (1991), 293-298 y 557-562.
- [2] G. Allaire, E. Bonnetier, G. Franfort y F. Jouve, Shape Optimization by the Homogenization Method, Numerische Mathematik **76** (1997), 27-68.

- [3] G. Allaire y C. Castro, A new approach for the optimal distribution of assemblies in a nuclear reactor, *Numerische mathematik*, **89** (1) (2001), 1-29.
- [4] G. Allaire y C. Castro, Optimization of nuclear fuel reloading by the homogenization method, *Structural and multidisciplinary optimization*, por aparecer.
- [5] G. Allaire y V. Kohn, Optimal Design for Minimum Weigh and Compliance in Plane Stress using Extremal Microstructures, *Eurp. J. Mech. A/Solids*, **12**, 6 (1993), 839-878.
- [6] M. Avellaneda, C. Bardos y J. Rauch, Contrôllabilité Exacte, Homogénéisation et Localization d'ondes dans un milieu non-homogène, *Asymptotic Analysis*, **5** (1992), 481-494.
- [7] H.T. Banks, W. Fang, R.J. Silcox y R.C. Smith, Approximation Methods for Control of Acoustic/Structure Models with Piezoceramic Actuators, *Journal of Intelligent Material Systems and Structures*, **4** (1993), 98-116.
- [8] A. Bensoussan, J.-L. Lions y G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, Studies in Mathematics and its Applications, North Holland, Amsterdam, 1978.
- [9] N. Burq, Contrôllabilité exacte des ondes dans des ouverts peu réguliers, *Asymptot. Anal.* **14** (2) (1997), 157-191.
- [10] C. Castro, Asymptotic analysis and control of a hybrid system composed by two vibrating strings connected by a point mass, *Control, Optimisation and Calculus of Variations (ESAIM COCV)*, <http://www.emath.fr/cocv/>, Vol 2 (1997), 231-280.
- [11] C. Castro, Boundary controllability of the one-dimensional wave equation with rapidly oscillating density, *Asymptotic Analysis* **20** (1999), 317-350.
- [12] C. Castro, Some controllability results in the one-dimensional wave equation with oscillating density, en preparación.
- [13] C. Castro y S. J. Cox, Achieving arbitrarily large decay in the damped wave equation, *SIAM J. Control Optim.* **39** (6), 1748-1755, 2001.
- [14] C. Castro y E. Zuazua, Analyse spectrale et contrôle d'un système hybride composé de deux poutres connectées par une masse ponctuelle, *C. R. Acad. Sci. de Paris*, t. 322, Série I, 351-356, 1996.
- [15] C. Castro y E. Zuazua, Une remarque sur les séries de Fourier non-harmoniques et son application à la contrôllabilité des cordes avec densité singulière, *C. R. Acad. Sci. de Paris*, t. 323, Série I, 365-370, 1996.

- [16] C. Castro y E. Zuazua, A hybrid system consisting of two flexible beams connected by a point mass: spectral analysis and well-posedness in asymmetric spaces, en *Elasticité , Viscoelasticité et Contrôle Optimal. Aspects Theoretiques et Numeriques*, J. Blum and A. Raoult eds., ESAIM Proceedings, **2** (1997), 17-53, <http://www.emath.fr>.
- [17] C. Castro y E. Zuazua, Boundary controllability of a hybrid system consisting in two flexible beams connected by a point mass, *SIAM J. Control and Optim.* **36**, 5 (1998), 1576-1595.
- [18] C. Castro y E. Zuazua, Exact boundary controllability of two Euler-Bernoulli beams connected by a point mass, *Mathematical & Computer Modelling.* **32** (2000), 955-969.
- [19] C. Castro y E. Zuazua, Une remarque sur l'analyse asymptotique spectrale en homogénéisation, *C. R. Acad. Sci. de Paris*, t. 322, Série I, 1043-1047, 1996.
- [20] C. Castro y E. Zuazua, Contrôlabilité de l'équation des ondes à densité rapidement oscillante à une dimension d'espace, *C. R. Acad. Sci. Paris*, t. 324, Série I, 1237-1242, 1997.
- [21] C. Castro y E. Zuazua, Low frequency asymptotic analysis of a string with rapidly oscillating density, *SIAM J. Appl. Math.*, **60** (2000), 1205-1233.
- [22] C. Castro y E. Zuazua, High frequency asymptotic analysis of a string with rapidly oscillating density, *European J. Appl. Math.*, **11** (2000), 592-622.
- [23] C. Castro y E. Zuazua, Concentration and lack of observability of waves in highly heterogeneous media, *Arch. Rat. Mech. Anal.*, por aparecer.
- [24] A.V. Cherkaev y L.V. Gibiansky, The exact coupled bounds for effective tensors of electrical and magnetic properties of two-components two-dimensional composites, *Proc. Royal Soc. Edinburgh*, **122 A** (1992), 93-125.
- [25] S. J. Cox y M. L. Overton, Perturbing the Critically Damped Wave Equation, *SIAM J. Appl. Math.* 56(5) (1996), pp. 1353-1362.
- [26] S. J. Cox y E. Zuazua, The Rate at Which Energy Decays in a Damped String, *Commun. in Partial Differential Equations*, Vol. 19, Nos. 1& 2 (1994), pp. 213-243.
- [27] P. Freitas, Optimizing the rate of decay of solutions of the wave equation using genetic algorithms: a counterexample to the constant damping conjecture, *SIAM J. Control & Optim.* 37(2) (1999), pp. 376-387.
- [28] S. Hansen y E. Zuazua, Exact Controllability and Stabilization of a Vibrating String with an interior Point Mass, *SIAM J. Control Optim.*, **33**, 5 (1995), 1357-1391.

- [29] Z. Hashin y S. Shtrikman, A Variational Approach to the Theory of Effective Magnetic Permeability of Multiphase Materials, *J. Appl. Phys.* **33** (1962), 3125-3131.
- [30] S. Kesavan, Homogenization of Elliptic Eigenvalue Problems, *Appl. Math. Optim.*, **5** (1979), 153-167 y 197-216.
- [31] V. Kohn, The Two-well Problem, por aparecer en *Proc. Royal Soc. Edinburgh*.
- [32] G. Lebeau, The wave equation with oscillating density: observability at low frequency, *ESAIM Control, Optimization and Calculus of variations*, **5** (2000), pp. 219.
- [33] J.-L. Lions, *Contrôlabilité Exacte, Stabilization et Perturbations de Systèmes Distribués*, Tomos 1 y 2, Masson RMA8, Paris, 1988.
- [34] W. Littman y L. Marcus, Stabilization of a Hybrid System of Elasticity by Feedback Boundary Damping, *Ann. Mat. Pura Appl.*, **CLII**, 4 (1988), 281-330.
- [35] W. Littman y L. Marcus, Exact Boundary Controllability of a Hybrid System of Elasticity, *Archive Rat. Mech.*, **103**, 3 (1998), 193-236.
- [36] W. Littman y L. Marcus, *Some Recent Results on Control and Stabilization of Flexible Structures*, Univ. Minn., Mathematics Report, 1988.
- [37] E.B. Lee and Y.C. You, Stabilization of a hybrid (string/point mass) System, *Proc. Fifth Int. Conf. on Systems Engineering*, Dayton, Ohio (Sep. 1987).
- [38] S. Micu y E. Zuazua, Asymptotics for the Spectrum of a Fluid/Structure Hybrid System Arising in the Control of Noise. *SIAM J. Math. Anal.* **29**, 4 (1997), 967-1001.
- [39] S. Micu y E. Zuazua, Boundary Controllability of a Linear Hybrid System Arising in the Control of Noise, *SIAM J. Control Optim.*, **35**, 5 (1997), 1614-1637.
- [40] F. Murat y L. Tartar, Calcul des Variations et Homogénéisation, en *Les Méthodes de l'Homogénéisation: Théorie et Applications en Physique*, Eyrolles, 319-369 (1985).
- [41] O.A. Oleinik, A.S. Shamaev y B.A. Yosifian, *Mathematical Problems in Elasticity and Homogenization*, North Holland, 1992.
- [42] F. Santosa y M. Vogelius, First-order Corrections to the Homogenized Eigenvalues of a Periodic Composite Medium, *SIAM J. Appl. Math.*, **53** (1993), 1636-1668.

Hacia un marco común para los títulos de Matemáticas en Europa

THE MATHEMATICS TUNING GROUP¹

Presentación

En mayo de 2001 se puso en marcha, bajo los auspicios de la Comisión Europea, el programa piloto Tuning educational structures in Europe para facilitar e impulsar la construcción del espacio europeo de educación superior previsto en los acuerdos de Bolonia y Praga. En dicho programa se seleccionaron cinco titulaciones (Administración de Empresas, Geología, Historia, Matemáticas y Ciencias de la Educación) y para cada una se constituyó una red con universidades de los distintos países de la UE. El proyecto finalizó el 31 de mayo de 2002, y en el documento que publicamos se encuentran las conclusiones de la red de matemáticas, en la que han participado la Universidad Autónoma de Madrid y la Universidad de Cantabria. Estas conclusiones se van a publicar en los Boletines de varias sociedades matemáticas europeas.

Aprovechamos esta ocasión para informar de que la Conferencia de Rectores de Universidades Españolas (CRUE) ha decidido utilizar la experiencia de Tuning para desarrollar a su vez un proyecto piloto. Al grupo de matemáticas, coordinado por las dos Universidades ya mencionadas, se han incorporado la Universidad Autónoma de Barcelona, la Universidad de Santiago de Compostela y la Universidad de Sevilla. Tan pronto como el grupo termine sus trabajos, los documentos que los recojan se pondrán en conocimiento de la comunidad matemática española para su discusión.

*José Manuel Bayod (Universidad de Cantabria)
bayodjm@unican.es*

*Adolfo Quirós (Universidad Autónoma de Madrid)
adolfo.quirós@uam.es*

Tras la firma de la *Declaración de Bolonia* [1] en 1999 por Ministros responsables de la Educación Superior de 29 países europeos, y de su continuación, el *Comunicado de Praga* [2], un grupo de universidades puso en marcha el proyecto “*Tuning educational structures in Europe*” [4, 5]. Lo han coordinado las universidades de Deusto y Groningen y ha obtenido financiación

de la Unión Europea. Como su nombre indica, el objetivo principal del proyecto fue estudiar la forma de “afinar” (como los distintos instrumentos de una orquesta, *no* uniformizar) las estructuras educativas europeas, y colaborar así en la construcción del Espacio Europeo de la Educación Superior. Esto debería a su vez contribuir a la movilidad y mejorar las posibilidades laborales de los titulados europeos.

Uno de los campos incluidos en el proyecto *Tuning* fue el de las matemáticas, y este documento refleja el consenso unánime del grupo de matemáticas del proyecto. Pero dado que el grupo no pretende tener ningún papel representativo, consideramos necesario someterlo a discusión entre la comunidad de matemáticos europeos. Estamos convencidos de que cualquier clase de acción en las direcciones que aquí señalamos solamente será posible y fructífera cuando se haya alcanzado un amplio acuerdo. Por supuesto, todos los matemáticos pertenecientes al grupo recibirán gustosos cualquier comentario sobre este documento. Sus direcciones electrónicas aparecen al final.

El Grupo *Tuning* de Matemáticas quiere mostrar su agradecimiento a los coordinadores del proyecto *Tuning*, Julia González (Universidad de Deusto) y Robert Wagenaar (Rijksuniversiteit Groningen), y a la Comisión Europea por crear las condiciones que permitieron una agradable y provechosa comunicación entre sus miembros.

Resumen

- Este documento se refiere únicamente a las universidades (incluyendo las politécnicas), y ninguna de nuestras propuestas se aplica a otros tipos de instituciones de educación superior.
- La finalidad de disponer de un “marco común para los títulos de matemáticas en Europa” es la de facilitar un reconocimiento automático, que contribuya a la movilidad.
- La idea de un marco común debe ir ligada a la de un sistema de acreditación.
- Las dos componentes de un marco común son unas estructuras similares (aunque no necesariamente idénticas) y una parte troncal, básica y común, en los contenidos de los dos o tres primeros años del plan de estudios (permitiendo cierto grado de flexibilidad local).
- Más allá de la parte básica y troncal del plan de estudios, y sin duda en todo el segundo ciclo, los planes podrían diverger de modo significativo. Puesto que hay muchas áreas en matemáticas, y están enlazadas con otros campos del conocimiento, la flexibilidad es de la máxima importancia.
- La base común de todos los planes de estudios incluirá el cálculo en una y varias variables reales y el álgebra lineal.

- Proponemos una amplia lista de otras materias que nuestros graduados deberían conocer para ser inmediatamente reconocidos como matemáticos. No se propone que todos los planes incluyan asignaturas específicas que se dediquen a cada uno de estos temas.
- No presentamos una lista obligatoria de temas que haya que estudiar, pero sí que mencionamos tres destrezas que cualquier graduado en matemáticas debería poseer:
 - (a) la capacidad de idear demostraciones,
 - (b) la capacidad de modelizar matemáticamente una situación,
 - (c) la capacidad de resolver problemas con técnicas matemáticas.
- El primer ciclo normalmente debería incluir el aprendizaje de algo de computación y la adquisición del conocimiento de al menos uno de los más importantes campos de aplicación de las matemáticas.
- Se debería procurar que los segundos ciclos de matemáticas fueran de muy diversa índole. Su característica común debería ser que todos los estudiantes lleven a cabo una apreciable cantidad de trabajo individual. Para conseguir esto, parece necesario un mínimo de 90 créditos ECTS² para obtener un título de *Master*.
- Puede ser aceptable que coexistan titulaciones con diversos diseños, pero en el caso de que se dieran desviaciones significativas del estándar (en lo relativo a los contenidos mínimos o a la estructura cíclica), éstas han de estar fundamentadas en unos requisitos de ingreso adecuados o en otros factores específicos del plan que puedan ser juzgados en la acreditación externa. De otro modo, tales títulos corren el riesgo de no beneficiarse del reconocimiento automático europeo que dará el marco común, aunque puedan constituir títulos válidos de educación superior.

1. Un marco común: lo que significa y lo que no significa

1.1 El único objetivo posible de acordar un “marco común europeo” debería ser el de facilitar un reconocimiento automático de los títulos de matemáticas en Europa para contribuir a la movilidad. Esto significaría que cuando una persona con un título en matemáticas obtenido en un país A se traslada a un país B:

- a) Se le reconocerá oficialmente el título, y para ello las autoridades del país B no le exigirán ninguna otra prueba de su capacidad.

²ECTS son las siglas de *European Credit Transfer System*. Los créditos ECTS se utilizan para medir el aprendizaje de los alumnos. Por definición, los resultados del aprendizaje que se espera que un alumno medio a tiempo completo pueda obtener en un año académico, valen 60 créditos ECTS. En consecuencia, la carga de trabajo que se precisa para obtener 60 créditos ECTS debería corresponder a lo que se espera que un estudiante medio a tiempo completo realice durante un año académico.

- b) Quienquiera que vaya a contratarle en el país B podrá suponer que el poseedor del título tiene los conocimientos generales que se esperan de alguien con un título en matemáticas.

Naturalmente, ninguna de estas facilidades garantizará la obtención de un empleo: el titulado en matemáticas tendrá que pasar por cualesquiera procedimientos (oposiciones, entrevistas, análisis de su currículum vitae, valoración por parte del empresario de la universidad en la que obtuvo el título,...) que se utilicen en el país B para obtener un empleo, ya sea público o privado.

1.2 Una componente importante del marco común de los títulos europeos de matemáticas es que todos los planes tengan estructuras similares, aunque no necesariamente idénticas. Otra componente es un acuerdo sobre una parte troncal, básica y común del contenido del plan que permita cierto grado de flexibilidad local.

1.3 Queremos insistir en que de ningún modo pensamos que un acuerdo sobre un marco común pueda usarse como un instrumento para los traslados automáticos entre universidades. Los traslados deberán considerarse caso a caso, puesto que diferentes planes de estudios pueden llevar a los estudiantes hasta los mismos niveles de formas diferentes pero todas ellas coherentes, mientras que una mezcla inadecuada de varios planes puede no servir para el mismo fin.

1.4 En muchos países europeos existen instituciones de educación superior que difieren de las universidades tanto en el nivel que exigen a sus estudiantes como en su enfoque general de la enseñanza y el aprendizaje. Para no excluir de la enseñanza superior a un número importante de estudiantes, en la práctica es esencial mantener estas diferencias. Queremos declarar expresamente que **este documento se refiere solamente a las universidades (incluyendo las politécnicas)**, y que cualquier propuesta de un marco común diseñado para las universidades no sería automáticamente aplicable a instituciones de otro tipo.

2. Hacia una troncalidad común

2.1 Consideraciones generales

A primera vista, las matemáticas parecen idóneas para la definición de unos contenidos comunes, por ejemplo, para los dos o tres primeros años. Por la naturaleza misma de las matemáticas, y por su estructura lógica, habrá una parte común a todos los planes de estudios de matemáticas, que constará de las nociones fundamentales. Pero por otra parte, existen muchas áreas de las matemáticas, y muchas de ellas están relacionadas con otros campos del conocimiento (informática, física, ingeniería, economía, etc.). La flexibilidad es

de la máxima importancia para preservar esta variedad y las interrelaciones que enriquecen nuestra ciencia.

Podría alcanzarse un acuerdo sobre una lista de materias que con toda seguridad deben estar incluidas (álgebra lineal, cálculo/análisis) o que debieran estar incluidas (probabilidad/estadística, cierta familiaridad con la utilización matemática de un ordenador) en cualquier título de matemáticas. En el caso de algunos temas especializados, como física matemática, sin duda habrá variaciones entre países e incluso entre universidades del mismo país, sin que deba deducirse ninguna diferencia de calidad entre los distintos planes de estudios.

Por otra parte, actualmente existen en Europa planes de estudios de matemáticas muy variados, con diferentes requisitos de acceso y con distintas duraciones de las enseñanzas y distintos niveles de exigencia sobre los estudiantes. Es enormemente importante que se mantenga esta variedad, tanto para la eficiencia del sistema educativo como desde el punto de vista social, con objeto de conseguir atender a las demandas del mayor número posible de alumnos potenciales. La fijación de una única definición de los contenidos, las destrezas y los niveles para la totalidad de la educación superior europea excluiría del sistema a muchos estudiantes y, en conjunto, resultaría contraproducente.

De hecho en el grupo hay un acuerdo total acerca de que los planes puedan diverger de modo significativo en lo que sea adicional a la parte troncal básica (por ejemplo en la dirección de la matemática “pura”; o de la probabilidad-estadística aplicada a la economía o a las finanzas; o de la física matemática; o de la enseñanza de las matemáticas en la educación secundaria). Lo que hará que esos planes sean reconocibles como planes válidos de matemáticas será su forma de presentación y su nivel de rigor, admitiendo que hay y debe seguir habiendo variantes en la importancia que se dé a cada tema y, hasta cierto punto, en el contenido, incluso dentro de los dos o tres primeros años.

En cuanto al segundo ciclo, no sólo pensamos que los distintos planes pueden diferir, sino que estamos convencidos de que, para reflejar la diversidad de las matemáticas y de sus relaciones con otros campos, se deberían desarrollar en las diferentes instituciones todo tipo de segundos ciclos diferentes en matemáticas, aprovechando en particular los aspectos en los que más destaque cada institución.

2.2 La necesidad de la acreditación

La idea de una troncalidad básica debe combinarse con un sistema de acreditación. Con el objetivo de reconocer que un programa dado cumple con los requisitos de la troncalidad, hay que comprobar tres aspectos:

- una lista de contenidos
- una lista de destrezas o competencias
- el nivel del dominio de los conceptos.

No es posible reducir estos aspectos a una simple escala.

Para conceder la acreditación a un plan de matemáticas es imprescindible un análisis por parte de un grupo de evaluadores académicos, de los cuales la mayor parte serán matemáticos. Los aspectos clave a ser evaluados deberían ser:

- (a) el plan de estudios en su conjunto
- (b) las unidades o asignaturas (tanto en contenido como en nivel)
- (c) los requisitos de acceso al plan
- (d) los objetivos del aprendizaje (las destrezas y el nivel alcanzado)
- (e) una evaluación cualitativa tanto por los graduados como por quienes les contratan.

El grupo no cree que se necesite un (elaborado) sistema de acreditación europeo, sino que las universidades, buscando el reconocimiento, actuarán a nivel nacional. Para que este reconocimiento tenga valor internacional, parece necesario que entre los evaluadores se incluyan matemáticos de otros países.

3. Algunos principios para la troncalidad común del primer título (*Bachelor*) en matemáticas

No creemos que sea necesario, ni siquiera oportuno, fijar una lista detallada de los temas a cubrir. Sin embargo, creemos que es posible dar algunas directrices sobre el contenido común de un “primer título europeo en matemáticas”, y, lo que es más importante, sobre las destrezas que todos los titulados deberían poseer.

3.1 Contenido

3.1.1 Todos los titulados en matemáticas conocerán y entenderán, y serán capaces de usar, los métodos y las técnicas apropiados a su plan de estudios. La parte común de todos los planes incluirá

- cálculo en una y varias variables reales
- álgebra lineal.

3.1.2 Los titulados en matemáticas han de conocer las áreas básicas de las matemáticas, no solo las que históricamente han guiado la actividad matemática, sino también otras de origen más moderno. En consecuencia los titulados normalmente habrán de conocer la mayoría de las siguientes materias, y preferiblemente todas:

- ecuaciones diferenciales a nivel básico

- funciones de variable compleja a nivel básico
- algo de probabilidad
- algo de estadística
- algo de métodos numéricos
- geometría de curvas y superficies a nivel básico
- algunas estructuras algebraicas
- algo de matemáticas discretas.

No es necesario que estos temas se aprendan en asignaturas o módulos individuales que cubran en profundidad y desde un punto de vista abstracto cada materia. Por ejemplo, un estudiante podría aprender sobre los grupos en un curso de teoría de grupos (abstracta) o en el marco de un curso sobre criptografía. Las ideas geométricas podrían aparecer en varias asignaturas, dado su papel central.

3.1.3 De acuerdo con el carácter y las exigencias del plan de estudios, se desarrollarán otros métodos y otras técnicas, cuyos niveles serán definidos por el propio plan. En cualquier caso, todos los planes incluirán un número importante de asignaturas con contenido matemático.

3.1.4 En la práctica y hablando en términos algo imprecisos, hay dos tipos de estudios de matemáticas que coexisten actualmente en Europa, y ambos tipos de estudios son útiles. Podemos llamarlos, siguiendo [3]³, “basados en la teoría” y “basados en la práctica”. La incidencia de cada uno de estos dos tipos de enseñanzas varía ampliamente según el país, y podría ser interesante averiguar si la mayor parte de los estudios universitarios europeos de matemáticas son “basados en la teoría” o no.

Los graduados en planes de estudios basados en la teoría tendrán conocimiento y comprensión de los resultados de varios de los campos más importantes de las matemáticas. Ejemplos de tales campos son el álgebra, el análisis, la geometría, la teoría de números, las ecuaciones diferenciales, la mecánica, la teoría de la probabilidad y la estadística, pero hay otros muchos. Sobre este conocimiento y esta comprensión se apoyarán el conocimiento y la comprensión de los métodos y técnicas matemáticos, otorgándoles un contexto matemático bien fundamentado.

Los graduados en planes de estudios basados en la práctica también tendrán conocimiento de los resultados de varios campos matemáticos, pero este conocimiento normalmente estará diseñado para apoyar la comprensión de modelos y de cómo pueden aplicarse. Además de los mencionados más

³El grupo consideró enormemente útil este documento, y mostró un acuerdo unánime con su contenido. Incluso se han utilizado al pie de la letra algunas de sus frases.

arriba, estos campos incluyen el análisis numérico, la teoría de control, la investigación operativa, las matemáticas discretas, la teoría de juegos y muchos otros. (Naturalmente estos campos también pueden estudiarse en las enseñanzas más teóricas.)

3.1.5 Es necesario que todos los titulados conozcan al menos una de las más importantes áreas de aplicación de las matemáticas, en la que el uso de las matemáticas sea esencial para entender verdaderamente la materia. La naturaleza y la forma en que se estudia esta área de aplicación puede variar en función de si el plan de estudios está basado en la teoría o en la práctica. Algunas de las posibles áreas de aplicación pueden ser la física, la astronomía, la química, la biología, la ingeniería, la computación, la tecnología de la información y las comunicaciones, la economía, la contabilidad, las ciencias actuariales, las finanzas y muchas otras.

3.2 Destrezas

3.2.1 Para un concepto como la integración en una variable, el mismo “contenido” podría significar:

- calcular integrales sencillas
- comprender la definición de la integral de Riemann
- conocer las demostraciones de la existencia y de las propiedades de la integral de Riemann para ciertas clases de funciones
- usar las integrales para modelizar y resolver problemas en diversas ciencias.

Concluimos que por una parte el contenido ha de ser detallado claramente, y que por otra mediante el estudio de una misma materia se desarrollan varias destrezas.

3.2.2 Los estudiantes que se gradúan en matemáticas disponen de una amplia variedad de posibilidades de empleo. Los empresarios valoran en alto grado la capacidad y el rigor intelectual, y las habilidades de razonamiento que estos estudiantes han adquirido, así como sus demostradas capacidades numéricas y el enfoque analítico a la solución de problemas que constituyen sus cualidades más distintivas.

Por tanto, las tres destrezas clave que consideramos que cualquier titulado en matemáticas debería adquirir son:

- (a) la capacidad para idear demostraciones
- (b) la capacidad para modelizar matemáticamente una situación
- (c) la capacidad para resolver problemas con técnicas matemáticas.

Hoy en día está claro que resolver un problema debe incluir su resolución numérica y computacional. Para esto se requiere un firme conocimiento de algoritmos y de programación, así como del uso del software actualmente existente.

3.2.3 Conviene resaltar también que estas destrezas y el nivel de las mismas se desarrollan de forma progresiva a través de la práctica de varias materias. No se empiezan los estudios de matemáticas con una asignatura llamada “cómo hacer una demostración” y con otra llamada “cómo modelizar una situación” con la idea de que estas destrezas se adquieran inmediatamente, sino que se desarrollan practicándolas en todas las asignaturas.

3.3 Nivel

Todos los graduados habrán desarrollado el conocimiento y la comprensión a un alto nivel en algún área en particular. El nombre de los estudios o del título reflejará su contenido de materias a alto nivel. Por ejemplo, los poseedores de títulos que incluyan “estadística” tendrán un conocimiento y una comprensión sustanciales de la teoría central de la inferencia estadística y de muchas aplicaciones de la estadística. Quienes posean un título en “matemáticas” pueden tener conocimientos de muy distintas partes de las matemáticas, pero en todo caso habrán tratado en profundidad algunos temas.

4. El segundo título (*Master*) en matemáticas

Ya hemos dejado claro nuestro convencimiento de que sería un error establecer cualquier clase de currículum troncal para los estudios de segundo ciclo. Dada la diversidad de las matemáticas, los diferentes planes deberían dirigirse a una amplia gama de estudiantes, incluyendo muchos cuyo primer título no sea en matemáticas sino en otros campos más o menos relacionados (informática, física, ingeniería, economía, etc.). En consecuencia se debería procurar que los segundos ciclos de matemáticas fueran de muy diversa índole.

Pensamos que el denominador común de todos los segundos ciclos debería residir, más que en el contenido, en el nivel que se espera que los alumnos alcancen. Una característica unificadora podría ser el requisito de que todos los estudiantes de segundo ciclo lleven a cabo una apreciable cantidad de trabajo individual, lo que se podría plasmar en la presentación de un proyecto individual de cierta consideración.

Creemos que, en orden a lograr el nivel necesario para realizar un verdadero trabajo individual en matemáticas, el tiempo requerido para obtener un título de *Master* debería ser al menos el equivalente de 90 créditos ECTS. Por tanto el número de créditos ECTS de un *Master* estará comprendido normalmente entre 90 y 120, dependiendo de cuál sea la duración de cada uno de los dos ciclos en los distintos países.

5. Un marco europeo y el acuerdo de Bolonia

5.1 La forma en que los diferentes países implementen el acuerdo de Bolonia tendrá trascendencia sobre la troncalidad común. En particular, 3+2 puede no ser equivalente a 5, porque en una estructura de 3+2 años los 3 primeros años podrían conducir a un título profesional, lo que significaría que se invierte menos tiempo en las nociones fundamentales, o podrían conducir a los 2 años siguientes, en cuyo caso el espíritu del plan de estudios de los 3 años sería diferente.

5.2 Si es mejor que los estudios de matemáticas estén formados por un *Bachelor* de 180 créditos ECTS seguidos por un *Master* de 120 créditos ECTS (es decir, una estructura 3+2, en términos de años académicos), o si por el contrario es preferible una estructura 240+90 (es decir, 4+1+proyecto), dependerá de varias circunstancias. Por ejemplo, una estructura 3+2 seguramente facilitará la movilidad entre materias para estudiantes que decidan seguir un *Master* en un área distinta de aquella en la obtuvieron su *Bachelor*.

Un aspecto que no se puede ignorar, al menos en matemáticas, es la formación de los profesores de enseñanza secundaria. En caso de que la cualificación pedagógica haya de obtenerse durante los estudios de primer ciclo, éstos probablemente deberían durar 4 años. Pero si el ser profesor de enseñanza secundaria exige un *Master* (o algún otro tipo de cualificación postgraduada), entonces un *Bachelor* de 3 años puede ser adecuado, y en este caso la formación pedagógica sería una de las posibles opciones de postgrado (a nivel de *Master* o a otro nivel).

5.3 El grupo no ha intentado resolver las contradicciones que podrían aparecer en el caso de que haya diferentes implementaciones del acuerdo de Bolonia (es decir, si coexisten planes universitarios de tres años con otros de cinco años; o si se establecen diferentes estructuras cíclicas, ya que se han propuesto todos estos esquemas: 3+1, 3+2, 4+1, 4+1+proyecto, 4+2). Como se ha dicho más arriba, podría ser aceptable que coexistan diversos sistemas, pero creemos que si hay grandes alejamientos del estándar (como la estructura 3+1, o el incumplimiento de los principios enunciados en la sección 3) éstos tienen que estar fundamentados en unos requisitos adecuados sobre los niveles de acceso o en otros factores particulares del plan de estudios, que puedan ser juzgados en la acreditación externa. De otro modo, tales títulos corren el riesgo de no beneficiarse del reconocimiento automático europeo que dará el marco común, aunque puedan constituir títulos válidos de educación superior.

Referencias

[1] http://www.bologna-berlin2003.de/pdf/bologna_declaration.pdf

[2] http://www.bologna-berlin2003.de/pdf/Prague_communicuTheta.pdf

- [3] Documento para la evaluación comparada de los títulos de Matemáticas, Estadística e Investigación Operativa, de la *Quality Assurance Agency for Higher Education* del Reino Unido.
<http://www.qaa.ac.uk/crntwork/benchmark/phase2/mathematics.pdf>.
- [4] Los sitios oficiales del proyecto *Tuning educational structures in Europe*:
<http://www.relint.deusto.es/TuningProject/index.htm>,
<http://www.let.rug.nl/TuningProject/index.htm>
- [5] Información sobre el proyecto *Tuning educational structures in Europe* en el sitio de la Comisión Europea:
<http://europa.eu.int/comm/education/tuning.html>

Miembros de “The Mathematics Tuning Group”:

Stephen Adam, *University of Westminster* (Experto en educación superior)

José Manuel Bayod, *Universidad de Cantabria* (bayodjm@unican.es)

Martine Bellec, *Université Paris IX Dauphine* (martine.bellec@dauphine.fr)

Marc Diener, *Université de Nice* (diener@math.unice.fr)

Alan Hegarty, *University of Limerick* (Alan.Hegarty@ul.ie)

Poul Hjorth, *Danmarks Tekniske Universitet* (P.G.Hjorth@mat.dtu.dk)

Anne Mette Holt, *Danmarks Tekniske Universitet* (Experta en relaciones internacionales)

Günter Kern, *Technische Universität Graz* (Kern@opt.math.tu-graz.ac.at)

Frans J. Keune, *Katholieke Universiteit Nijmegen* (keune@sci.kun.nl)

Luc Lemaire, *Université Libre de Bruxelles* (llemaire@ulb.ac.be)

Andrea Milani, *Università degli Studi di Pisa* (milani@dm.unipi.it)

Julian Padget, *University of Bath* (jap@maths.bath.ac.uk)

María do Rosário Pinto, *Universidade de Porto* (mspinto@fc.up.pt)

Adolfo Quiros, *Universidad Autónoma de Madrid* (adolfo.quiros@uam.es)

Wolfgang Sander, *Technische Universität Braunschweig* (w.sander@tu-bs.de)

Hans-Olav Tylli, *University of Helsinki* (hojtylli@cc.helsinki.fi)

Proyectos empresariales en la sociedad del siglo XXI

J.A. GARRIDO¹

Iberdrola
Gardoqui 8, 48008 Bilbao

Hemos iniciado el siglo XXI sumidos en una gran confusión, sin duda acrecentada por los escándalos económicos de los últimos meses, cuando en realidad entre el siglo XX y el XXI, no ha habido ninguna frontera más que la artificial y convenida del calendario.

La mejor manera de comprender el siglo XXI es entender los cambios producidos a lo largo del siglo XX. Cambios que han suscitado riesgos e incertidumbres, que han sustituido capital por conocimiento, cantidad por calidad y todo ello en un entorno de abundancia, hasta de despilfarro podríamos decir, de información, tecnología y capital, frente a escasez de tiempo, atención y justicia social. Cambios que nos han conducido a una globalización donde las ideas universales han sido desplazadas por las interpretaciones locales, donde se ha producido la desaparición de lo extraño y donde el espacio geográfico ha sido sustituido por el espacio social. Cambios que nos permiten abordar el futuro utilizando, en vez de dejarnos llevar por, la tecnología; futuro que lo definimos como el destino al que queremos llegar.

En definitiva, salimos del siglo XX y entramos en el XXI con las mismas variables clave que nos deben de servir para la reflexión y la definición de un proyecto de futuro en el cual la complejidad, y no la confusión, deberá ser la base de todas nuestras acciones. Complejidad entendida como análisis de muchas variables; hoy en día cualquier situación que estudiemos a nivel familiar, empresarial, institucional, ... es compleja y, por lo tanto, debe de ser abordada conociendo las variables que la definen.

Una situación que ha contribuido a crear confusión ha sido la superposición en el tiempo de dos efectos muy distintos. Uno fue la explosión de algo que nunca existió: las empresas de la llamada nueva economía, cuya analogía con el traje invisible del señor duque fue notable. Otro, la capacidad real de transformación en la forma de desarrollar las actividades empresariales, con el soporte de las tecnologías de la información y comunicaciones.

La utilización de las capacidades de computación para potenciar muchas de las actividades “convencionales” y la capacidad de integrar el conocimiento especializado disperso mediante el uso de Internet, son realidades que no pueden ser ignoradas y que previsiblemente darán lugar a cambios importantes en los

¹Vicepresidente de Iberdrola. Doctor Honoris Causa por la UPV-EHU

esquemas organizativos de las empresas. El proceso está en marcha, el desarrollo tecnológico y organizativo continúa y las transformaciones se prolongarán aún bastante tiempo.

“En el siglo XXI, hacer empresa es hacer Sociedad, capitalizando a la persona”.

La Empresa es el lugar de concurrencia de personas creativas que se reúnen para aportar un valor que es apreciado por los grupos de interés que participan en ella, clientes, trabajadores, accionistas, proveedores y la Sociedad. La Empresa es un ente social en que cada persona puede y debe aportar su libertad creativa, a partir de los valores de la solidaridad, la responsabilidad, la tolerancia y la profesionalidad. Valores que deben articularse a través de los principios culturales sobre los que se basan las relaciones con los grupos de interés antes citados.

La Empresa, así concebida, se caracteriza por “crear riqueza”, con lo que: satisface a sus clientes, satisface a sus trabajadores, retribuye adecuadamente a sus accionistas, contribuye al desarrollo y crecimiento de sus proveedores y contribuye al perfeccionamiento sostenido del entorno económico y social.

Empresa y sociedad, al menos en el mundo occidental, son dos realidades estrechamente vinculadas, casi incluso dos aspectos de una misma realidad. Los cambios que experimente la sociedad influirán, están influyendo más bien, en los modelos y en la concepción de la empresa. La sociedad lleva camino de sufrir importantes transformaciones que tendrán su reflejo en la forma de concebir la empresa, en su estructura, en la forma de gestionarla y en su función dentro de la sociedad civil, la cual parece que toma cada vez más cuerpo en las sociedades más cultas y más avanzadas.

Por eso la gran necesidad de “civilizar” a la Sociedad, lo que significa dar mayor protagonismo a la “Sociedad civil” frente a la “Sociedad política”. Muchas veces se transmite la sensación de que los empresarios son subsidiarios de los políticos cuando en realidad están llamados a ser líderes sociales. Líderes sociales comprometidos con la creación de riqueza y no con otros objetivos cortoplacistas cuyas consecuencias negativas afectan a toda la Sociedad.

La “creación de riqueza”, y su justo reparto, es un compromiso de largo plazo con el que la empresa consigue legitimidad para poder seguir existiendo en un determinado contexto económico, social y político. Uno de los aspectos de la crisis actual del capitalismo es precisamente el enfrentamiento y la elección en la empresa entre valores sociales frente a los valores operativos tradicionales.

La empresa se caracteriza por su naturaleza simbólica y moral, de tal forma que se puede definir como el lugar de integración de conocimiento especializado, alrededor de las ideas. De esta forma la empresa del siglo XXI logra la transformación de Ideas en Beneficios, en contraste con la empresa del siglo XX que transformaba los Recursos en Beneficios.

De nuevo aparece la persona y la organización como referente empresarial, tal vez lo más importante hasta ahora de la Sociedad del Conocimiento es el descubrimiento, o la percepción, de que el conocimiento constituye un activo fundamental en los procesos de creación de valor añadido. El principal problema pendiente es que aún no se han encontrado soluciones efectivas para

optimizar la utilización del conocimiento disponible o para mejorar los procesos de producción de nuevos conocimientos. Europa se resiente de forma especial del primero de estos males y no consigue encontrar el remedio.

En estos momentos el desarrollo de las neurociencias cognitivas, con el apoyo de la tecnología con aportaciones como, por ejemplo, la posibilidad de obtención de neuroimágenes funcionales, es algo, que si bien tiene un origen reciente y un desarrollo todavía corto, su incidencia a lo largo del siglo puede ser más trascendente de lo que en estos momentos se puede estar pensando.

La Sociedad del Conocimiento podría contar, para su desarrollo, con unos instrumentos que hasta ahora ni siquiera había imaginado.

La transformación de Ideas en Beneficios da lugar a unas nuevas empresas que fundamentan su actividad en la capacidad de integrar el talento y la creatividad individual en una dimensión colectiva. La nueva estructura de estas empresas, basada en las potencialidades individuales y sociales de las personas, impone una nueva forma de entender la propia empresa y de gobernarla. Hablar de visión, misión y valores es apoyarse en la ilusión, el conocimiento y el comportamiento. Para definir la visión, objetivo a largo plazo ilusionante y alcanzable, es necesario tener bien clara la misión, conocimiento de lo que sabemos hacer (del negocio) y ambas deben estar soportadas por unos sólidos valores.

El empresario tiene que ser capaz de crearse su propia ilusión para que pueda ser compartida por sus colaboradores y por los distintos grupos de interés que participan de una forma u otra en la vida de la empresa. El líder empresarial tiene que ser un líder referente que colabore a la formación y despliegue de otros líderes dentro de su propio negocio: liderazgo compartido.

Una de las características relevantes de este nuevo líder, es su capacidad para seleccionar personas, quienes bajo la influencia de la cultura del liderazgo compartido se transformarán en nuevo líderes. Ser líder seleccionador significa buscar personas que compartan Proyectos, huir de la mediocridad, marcar un estilo acorde con el Proyecto, seleccionar a los “seniors” que quieran ayudar desde la experiencia y el conocimiento, no tener miedo a equivocarse...

La principal aportación de un líder empresarial es la de crear y potenciar la cultura en la que la competencia (conocimiento) y los comportamientos (valores) sean las bases del desarrollo del proyecto empresarial.

Es preciso ser suficientemente competente en los conocimientos básicos y en los fundamentos de cada negocio. Conocer, manejar y controlar las variables clave de cada negocio y en cada ocasión. La competencia tecnológica es un factor fundamental para un líder empresarial y vemos cómo, paulatinamente, va perdiendo su “glamour” social e intelectual.

El líder empresarial tiene un poder ético que debe desarrollarlo de manera positiva para su empresa, para sus colaboradores, para quienes se relacionan con su negocio y para la Sociedad. Los valores de los líderes empresariales están vinculados a los asumidos por la Sociedad en la que viven.

Los valores del líder son el marco de referencia para él y para su equipo (liderazgo compartido). Este líder tiene que elegir y llevar a su comportamiento una serie de principios culturales que dejen su impronta allí por donde pase.

Cumplir con los comportamientos exigidos por las leyes y los reglamentos no es ningún mérito especial, sino una condición previa que se debe suponer en cualquier empresario o directivo.

El nuevo capitalismo tiene que dar lugar a una nueva sociedad. A su vez, la nueva sociedad que se está configurando exige un nuevo paradigma del sistema económico. Algunos de los aspectos determinantes del cambio social son el incremento de la esperanza de vida, la transformación de los esquemas de valores, el desarrollo de Tecnologías de la Información y de las Comunicaciones (TIC), el progresivo desarrollo de la Sociedad del Conocimiento y desviación del esfuerzo investigador del campo de las TIC al de las Ciencias de la Vida.

Necesitamos definir la Nueva Corporación que ha de generar el nuevo capitalismo de siglo XXI.

La empresa entendida como corporación es uno de los logros y símbolos del desarrollo de la sociedad industrial y, como tal, ha sufrido su misma evolución y recibido sus mismos impactos. La “calidad de vida” fue definida como el espíritu de la sociedad industrial. Más tarde, este concepto de “calidad” invadiría la vida empresarial más allá de su actividad productiva, incorporando nuevos compromisos que hoy una empresa no puede eludir: Calidad en los productos, Calidad de las condiciones de Trabajo, Calidad del Medio Ambiente... Hasta el punto que estos factores impulsados por la presión de la opinión pública han llegado a transformar los valores que han soportado durante el siglo XX la actividad empresarial, sustituyendo:

- El Crecimiento por el Crecimiento equilibrado.
- La Tecnología por el uso tecnológico seguro, respetuoso con el medio ambiente y la estabilidad social.
- El Beneficio por la Responsabilidad Social de las Corporaciones.
- La Gestión autoritaria por la gestión participativa.
- La eficiencia económica por la eficiencia en la economía, en el factor humano, en la cultura empresarial.
- La supuesta lealtad a la empresa por la adhesión a proyectos compartidos.

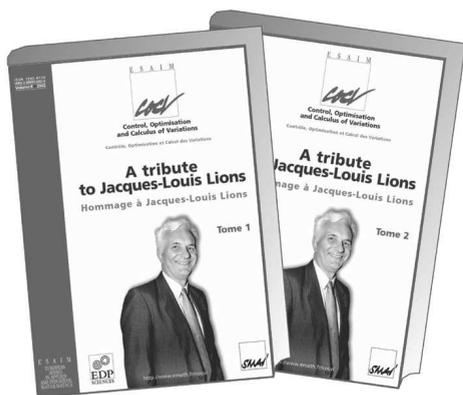
Hasta el punto que hoy una empresa no puede desarrollar su actividad permaneciendo ajena a las expectativas de la Sociedad.

En definitiva, se trataría de producir un cambio radical de tendencia en “nuestros Proyectos Empresariales” de manera individual y colectiva acertando a saber definir correctamente (no dejar que lo hagan otros), la visión, la misión y los valores; adquiriendo, manteniendo y mejorando el conocimiento de los aspectos clave de cada negocio (variables claves), lo que también es aplicable a nivel familiar y personal.

Varios autores*ESAIM:COCV. A tribute to Jacques-Louis Lions*

EDP Sciences, Les Ulis Cedex A, France, 2002.

ISBN: 2-86883-602-X (Tomo 1) y 2-86883-609-7 (Tomo 2)



Estos dos Tomos de ESAIM:COCV están dedicados a honrar la memoria de Jacques-Louis Lions.

La serie de revistas ESAIM (European Series in Applied and Industrial Mathematics) fue creada en septiembre de 1995 por la SMAI, Sociedad Francesa de Matemática Aplicada e Industrial, con el apoyo del Ministerio de Educación y Ciencia francés. Esta iniciativa fue posible gracias al impulso del entonces presidente de la SMAI Jean-Pierre Puel.

La serie comprende cuatro revistas, ESAIM: Modélisation Mathématique et Analyse Numérique (ESAIM: M2AN), ESAIM: Control, Optimisation and Calculus of Variations (ESAIM: COCV), ESAIM: Probability and Statistics (ESAIM: PS) y ESAIM: Proceedings.

Jean Michel Coron, ha sido el editor en jefe de ESAIM: COCV desde su creación.

Como su propio nombre indica, los temas en los que el interés de esta publicación se centra son la Teoría del Control, el Cálculo de Variaciones y la Optimización. Si bien en sus orígenes se trató de una publicación meramente electrónica, en la actualidad los artículos publicados cada año son recogidos en un volumen impreso que es enviado a las Instituciones que lo suscriben, cuyos miembros pueden a su vez consultar en tiempo real los contenidos de la revista de manera electrónica.

La editorial que en la actualidad se encarga de su edición y difusión es EDP Sciences (EDP = "Editions de Physique" en este caso y no Ecuaciones en Derivadas Parciales), cuya dirección es <http://www.edpsciences.org/cocv/>.

Desde su creación, ESAIM: COCV ha publicado más de 160 artículos y más de 5000 páginas. En mi opinión, las contribuciones publicadas hacen de esta revista una de las de mayor calidad en su campo.

Este año, además del volumen anual habitual, ESAIM: COCV ha publicado este volumen especial en dos tomos dedicados a la memoria del reciente y prematuramente desaparecido Jacques-Louis Lions, que ha sido un pionero de la Matemática Aplicada contemporánea (véanse las reseñas al respecto en el anterior número del Boletín) y que ha influenciado de manera decisiva los campos abarcados por esta revista.

J.L. Lions fue también el autor del primer artículo de la revista y ofreció su apoyo entusiasta en todo momento, lo cual contribuyó de manera decisiva al éxito de la empresa.

Mediante estos dos tomos, la revista ESAIM: COCV, su Comité Editorial y los numerosos autores que han contribuido a este volumen, que representan a su vez a un colectivo mucho más amplio de investigadores del campo, desean recordar y rendir homenaje a J.L. Lions, por sus contribuciones científicas y la labor de maestro realizada, que marcarán el futuro de este campo ya de manera permanente.

E. Zuazua
enrique.zuazua@uam.es

Título:	ANÁLISIS DE UN MODELO BICAPA DE AGUAS POCO PROFUNDAS.
Doctorando:	María de la Luz Muñoz Ruiz.
Director/es:	Carlos Parés Madroñal y Pierre Orenge.
Defensa:	28 de junio de 2002, Universidad de Málaga.
Calificación:	Sobresaliente cum Laude.

Resumen:

El objetivo principal de esta memoria es el análisis teórico de un sistema de ecuaciones en derivadas parciales acopladas, que corresponde a un modelo bicapa de aguas poco profundas, aplicable al estudio de la dinámica que tiene lugar en la zona del Mar de Alborán y el Estrecho de Gibraltar.

En primer lugar se aborda la construcción del modelo de aguas poco profundas bicapa en formulación velocidad-espesor, y se recuerdan aquellos resultados relativos al análisis de un modelo de aguas poco profundas de una sola capa que se extenderán al caso de dos capas.

A continuación se aborda el análisis del problema bicapa con condiciones de contorno homogéneas, ofreciendo un teorema de existencia de solución para datos controlados, y algunos resultados de regularidad que permiten probar un teorema de unicidad de solución. La principal dificultad es la aparición de términos de acoplamiento entre las capas. Esto plantea la necesidad de obtener estimaciones a priori en el espacio de las funciones de cuadrado sumable con el fin de obtener la existencia de soluciones. En el caso de una sola capa estas estimaciones eran obtenidas una vez probada la existencia de solución.

Después se estudia el problema bicapa con condiciones de contorno no homogéneas, para el que se ofrece un teorema de existencia de solución. A las diferencias propias del modelo bicapa se unen ahora las debidas a la aparición de términos de borde que es necesario estimar.

Por último se realiza el análisis teórico y la aproximación numérica de un problema bicapa unidimensional que modela un flujo bicapa en un canal con sección rectangular variable. Desde el punto de vista teórico se presentan resultados de existencia, regularidad y unicidad de solución. Desde el punto de vista numérico se propone un esquema de tipo volúmenes finitos para su resolución: el Q -esquema de Van Leer con descentrado de los términos fuente y se presentan algunos resultados numéricos obtenidos con este esquema.

Título:	ALGUNOS RESULTADOS DE EXISTENCIA, UNICIDAD Y ESTABILIDAD PARA EDP FUNCIONALES ESTOCÁSTICAS NO LINEALES.
Doctorando:	María José Garrido Atienza.
Director/es:	Tomás Caraballo Garrido y José Real Anguas.
Defensa:	29 de mayo de 2002, Universidad de Sevilla.
Calificación:	Sobresaliente cum Laude (por unanimidad).

Resumen:

En la Memoria se analiza la existencia y unicidad de solución para ecuaciones en derivadas parciales estocásticas con retardos en un contexto variacional, así como el comportamiento asintótico de las soluciones de sistemas diferenciales estocásticos, tanto en el caso en el que no haya características hereditarias como en el que la ecuación contenga retardos. La principal dificultad que surge al estudiar la existencia de solución de los sistemas planteados es que no se puede aplicar ni un método de Galerkin adecuado, puesto que las ecuaciones consideradas son retardadas, ni un esquema de Picard conveniente, debido a que los operadores que intervienen en los sistemas son muy generales. En concreto, en el Capítulo 3 se presentan resultados relativos a una ecuación en derivadas parciales estocástica de primer orden en tiempo, bastante general, y con características hereditarias, donde los retardos son de tipo variable. Motivados por los resultados obtenidos para un sistema diferencial estocástico de primer orden en tiempo y con retardos, en el Capítulo 4 se ha tenido como objetivo establecer resultados similares en el caso de que la ecuación sea retardada y de segundo orden en tiempo. Se han obtenido varios resultados para un problema de este tipo donde las familias de operadores que intervienen son no lineales y muy generales, constituyendo estos una de las aportaciones más significativas de esta Tesis, debido principalmente a que no conocemos resultados de este tipo para estos problemas. El Capítulo 5 se dedica a un análisis del comportamiento asintótico de sistemas diferenciales. Fundamentalmente se ha centrado en la ecuación de primer orden en tiempo, para la que se obtienen resultados de estabilidad trayectorial de una manera directa, es decir, sin necesidad de establecer resultados de estabilidad en media cuadrática. Además, el decaimiento que se analiza no es, en general, de tipo exponencial, sino que se trabaja con funciones de decaimiento más generales. Por otro lado, los resultados obtenidos relativos a estabilidad casi segura de las soluciones de ecuaciones diferenciales estocásticas de primer orden en tiempo se aplican a la estabilización de sistemas deterministas y estocásticos. Más concretamente, si el ruido no causa estabilidad exponencial, o no se sabe determinar si el sistema perturbado estocásticamente es o no exponencialmente estable, se investiga si el ruido produce algún tipo de estabilidad más débil, o incluso más fuerte. También se dan algunos criterios para la construcción de estabilizadores de tipo feedback. En este último Capítulo se establecen además algunos resultados relativos a la estabilidad de las soluciones de ecuaciones diferenciales estocásticas de segundo orden en tiempo. En concreto, se analiza la estabilidad exponencial en media cuadrática y, a partir de ahí, la estabilidad exponencial trayectorial del sistema.