

# 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](mailto: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 amigos,

Con gran satisfacción, encabezamos con estas breves palabras un nuevo número de nuestro Boletín. En él encontraréis una gran variedad de artículos (científicos, de carácter aplicado/industrial, educativos) que esperamos sean de vuestro agrado.

En esta ocasión, tenemos además la gran alegría de anunciar que nuestros compañeros Fuensanta Andreu-Vaillo, Vicent Caselles y José M. Mazón han sido galardonados con el Premio Ferrán Sunyer i Balagué 2003 por la monografía titulada *Parabolic quasilinear equations minimizing linear growth functionals*. De todo ello se ofrecen más detalles en el apartado referente a las noticias de SēMA.

Seguimos a vuestra disposición para cualquier sugerencia que permita mejorar nuestra labor.

Grupo Editor  
boletin\_sema@orion.ciencias.uniovi.es

Queridos amigos:

Tras el relevo de nuestros compañeros de la Universidad de Córdoba, el nuevo grupo editor ha cumplido ya el año de trabajo. En este tiempo, nuestros compañeros de la Universidad de Oviedo han dado un nuevo impulso al boletín. El boletín no solamente ha mejorado su imagen, sino que además se ha conseguido un objetivo que nos habíamos marcado: lograr una periodicidad en su publicación. Marzo, junio, septiembre y diciembre son los meses en los que nuestro boletín aparecerá. Os aseguro que no es tarea fácil cumplir con este compromiso, pero ellos, con la ayuda de Enrique Fernández Cara, están teniendo éxito en esta empresa. Nuestras felicitaciones y agradecimientos por el trabajo que están realizando.

Con el curso académico terminado, hago un breve balance de la actividad de SĒMA durante los últimos nueve meses. En este tiempo hemos firmado nuevos convenios de reciprocidad con la Società Italiana di Matematica Applicata e Industriale (SIMAI), la Societé de Mathématiques Appliquées et Industrielles de Francia (SMAI) y la Societat Catalana de Matemàtiques (SCM). Con todas ellas el acuerdo de reducción en la cuota es del 50%. Estos acuerdos se unen a los ya firmados anteriormente con la Real Sociedad Matemática Española (RSME) (60%), Gesellschaft für Angewandte Mathematik und Mechanik (GAMM) (33%) y la American Mathematical Society (AMS) (50%). Además en estos momentos estamos dando pasos para la firma de acuerdos con la Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC) y la Association des Mathématiques et Applications de Túnez.

Por otro lado, intentamos que estos acuerdos se consoliden mediante la organización de eventos conjuntos. Así el día 30 de mayo se desarrolló en la Universidad del País Vasco una jornada sobre Mecánica de Fluidos en colaboración entre la RSME y SĒMA y organizada por Mikel Lezaun, vicepresidente de nuestra sociedad. En estos momentos José M. Arrieta está ultimando la organización de una sesión en Brasil, en el XXVI Congreso Nacional de Matemática Aplicada e Computacional que organiza la SBMAC. Durante los días 22 y 25 de septiembre, tendrá lugar en Santiago de Compostela el *Encuentro de Sociedades Latinoamericanas de Matemáticas* en el que SĒMA participa. También participa SĒMA, junto a SCM, RSME y SEIO (Sociedad de Estadística e Investigación Operativa) en la organización del International Congress of Mathematicians (ICM2006) a celebrar en Madrid.

Seguramente me dejo en el tintero alguna otra actividad, pero creo que los ejemplos dados son suficientes para mostrar que SĒMA está viva y tiene una presencia importante en la matemática española. Por otro lado, de puertas adentro, también estamos trabajando y espero que próximamente podamos contar con una página *web* más dinámica, con foros de debate, que nos permitirá una mayor comunicación. Tengo que agradecer el trabajo que nuestro compañero de la Universidad de Cádiz, Rafael Rodríguez, está realizando para adaptar la página.

Por supuesto no me he olvidado del evento más importante que tenemos este año. Como todos habréis adivinado, me refiero al XVIII CEDYA/VIII CMA que nuestros compañeros de la Universidad Rovira i Virgili de Tarragona están organizando. La semana del 15 al 19 de septiembre espero encontraros allí, donde además celebraremos nuestra Asamblea General y entregaremos los premios SĒMA al *Joven Investigador* y a la *Divulgación de la Matemática Aplicada*. Os deseo un feliz verano.

*Eduardo Casas*

## Métodos iterativos para resolver sistemas lineales

R. BRU

Departament de Matemàtica Aplicada, Universitat Politècnica de València

rbru@mat.upv.es

### Resumen

En este trabajo se revisan diferentes métodos iterativos para la resolución de sistemas lineales  $Ax = b$ , mediante una pauta común basada en particiones de la matriz de coeficientes del sistema. En los diferentes métodos presentados se estudia la convergencia de cada uno de ellos, para diferentes tipos de matrices  $A$ . En particular, se trabaja con matrices monótonas,  $M$ -matrices,  $H$ -matrices y matrices simétricas definidas positivas. Algunas ideas a desarrollar se indican al final del trabajo.

**Palabras clave:** *Sistemas de ecuaciones lineales, métodos iterativos, particiones,  $M$ -matrices,  $H$ -matrices, matrices definidas positivas, métodos paralelos, método de Schwarz.*

**Clasificación por materias AMS:** *65F10, 15A06*

## 1 Introducción

En este trabajo vamos a estudiar diversos métodos iterativos que sirven para resolver sistemas lineales del tipo

$$Ax = b, \tag{1}$$

donde  $A \in \mathbb{C}^{n \times n}$ , es una matriz no singular,  $x \in \mathbb{C}^n$  y  $b \in \mathbb{C}^n$ , si bien algunos tipos de matrices consideradas se definen directamente en  $\mathbb{R}^{n \times n}$ . Estas ecuaciones lineales surgen, por ejemplo, cuando se discretizan ecuaciones en derivadas parciales mediante técnicas de diferencias finitas o de elementos finitos (ver [34] y [37]).

---

Fecha de recepción: 19 de Junio de 2002

Las diferentes discretizaciones dan lugar a diferentes propiedades de la matriz del sistema  $A$ . Por ello, es conveniente estudiar la resolución del sistema (1) bajo las diferentes situaciones que se pueden presentar.

Aunque sólo sea de forma rápida, indicaremos que los métodos iterativos tienen ciertas ventajas respecto a los métodos directos, cuyo paradigma es el método de Gauss. Cuando la matriz del sistema  $A$  está mal condicionada los métodos directos pueden dar soluciones con mucho error. Por otra parte, el coste computacional de los algoritmos directos que más se usan es del orden  $n^3$ ,  $O(n^3)$ , mientras que el coste de los iterativos es de  $O(n^2)$  por iteración. Además, si la matriz  $A$  es una matriz vacía o dispersa, con  $O(n)$  elementos no nulos, entonces el coste por iteración de los métodos iterativos es de  $O(n)$ , siendo en este caso la diferencia de costes muy significativa. También existen determinadas estructuras matriciales, como las matrices banda, que favorecen el uso de los métodos iterativos. Por último, los métodos iterativos son intrínsecamente paralelizables, mientras que los directos no lo son (ver [23] y [33]).

Iniciaremos el estudio con la puesta en común de los métodos clásicos mediante la técnica de las particiones. Con ella se pueden estudiar no solamente los métodos de Jacobi, Gauss-Seidel o SOR, sino también los métodos de dos etapas y los métodos de direcciones alternadas. Además, se pueden obtener algoritmos síncronos y asíncronos para la computación en paralelo, mediante el uso de multiparticiones. En un trabajo posterior queremos estudiar los métodos iterativos basados en los subespacios de Krylov, así como los preconditionadores de los mismos.

Estos dos grandes tipos de métodos iterativos matriciales para la resolución de sistemas lineales tienen diversos puntos en común. Por ejemplo, los métodos clásicos se pueden utilizar como preconditionadores de los métodos basados en los subespacios de Krylov, incluso algunos métodos pueden ser considerados por ambos tipos de esquemas iterativos (ver también [23] y [33]). Recordamos que consideraremos sólo sistemas lineales cuya matriz de coeficientes es invertible. En la última sección se indican unas notas sobre los sistemas singulares.

## 2 Resultados básicos

Para resolver el sistema (1) se puede realizar la siguiente *partición* de la matriz  $A$

$$A = M - N, \quad \text{donde } M \text{ es invertible y } N \neq O, \quad (2)$$

y por tanto escribir el sistema (1) como

$$Mx = Nx + b.$$

Esta ecuación inspira la construcción del método iterativo, del tipo de punto fijo, siguiente

$$Mx_l = Nx_{l-1} + b, \quad l = 1, 2, \dots \quad (3)$$

Este esquema da lugar al primer algoritmo que estudiaremos.

**Algoritmo 1. (Una partición).** Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,
  - 1.1. Resolver  $Mx_l = Nx_{l-1} + b$

Este esquema será el esquema iterativo básico que vamos a analizar. El análisis se centra en la convergencia del método iterativo cuyo esquema es (3) y, en caso de convergencia, en la rapidez de la misma.

Como por definición  $M$  es invertible, la ecuación (3) se puede escribir de la forma

$$x_l = M^{-1}Nx_{l-1} + M^{-1}b, \quad l = 1, 2, \dots \quad (4)$$

Por recursión, la ecuación (4) se puede escribir en función del vector inicial  $x_0$ , de donde se puede observar que la convergencia es función de la sucesión matricial  $\{(M^{-1}N)^l\}$ . Aquí el *radio espectral* de  $M^{-1}N$ , representado por  $\rho(M^{-1}N)$ , es decir, el máximo valor absoluto de los valores propios de la matriz, juega un papel esencial. De hecho se define la **velocidad media de convergencia** del esquema iterativo anterior, para  $l$  iteraciones, como

$$R_l(M^{-1}N) = -\ln \sqrt[l]{\|(M^{-1}N)^l\|},$$

donde  $\|\cdot\|$  representa una norma matricial inducida. La **velocidad asintótica de convergencia** se define como

$$R_\infty = \lim_{l \rightarrow \infty} R_l(M^{-1}N).$$

Puesto que

$$\lim_{l \rightarrow \infty} \sqrt[l]{\|(M^{-1}N)^l\|} = \rho((M^{-1}N)),$$

(ver [37]), se tiene que

$$R_\infty = -\ln \rho(M^{-1}N).$$

En consecuencia, el radio espectral de la *matriz de iteración*  $M^{-1}N$  mide la velocidad asintótica de convergencia. Por tanto, si tenemos dos particiones convergentes de una misma matriz  $A = M_1 - N_1 = M_2 - N_2$ , para comparar su convergencia tendremos que comparar sus radios espectrales, es decir, verificar si  $\rho(M_1^{-1}N_1) < \rho(M_2^{-1}N_2)$  o viceversa. Vale la pena hacer notar que puede darse el caso en que se de la desigualdad anterior entre radios espectrales pero que para un número pequeño de iteraciones la velocidad media de convergencia de la segunda partición sea mayor que la de la primera partición.

El estudio de la convergencia del Algoritmo 1 o del método correspondiente (3) se basa en el estudio del error

$$e_l = x_* - x_l, \quad l = 1, 2, \dots,$$

donde  $x_*$  es la solución exacta de la ecuación (1). Esta solución verifica la ecuación (3) y por tanto (4), es decir,

$$x_* = M^{-1}Nx_* + M^{-1}b. \quad (5)$$

Restando las ecuaciones (4) y (5), se obtiene

$$e_l = x_* - x_l = M^{-1}N e_{l-1}, \quad l = 1, 2, \dots,$$

o escrito en función del error inicial  $e_0$

$$e_l = (M^{-1}N)^l e_0, \quad l = 1, 2, \dots \quad (6)$$

En consecuencia, el método iterativo (3) converge a la solución del sistema (1) si y sólo si la sucesión del error  $\{e_l\}$  converge a cero, lo que es equivalente a que la sucesión de potencias de la matriz de iteración satisfice

$$\lim_{l \rightarrow \infty} (M^{-1}N)^l = O.$$

Para decidir cuándo detener la iteración es necesario estimar alguna norma del error  $\|x_* - x_l\|$ , lo que se suele hacer considerando  $\|x_l - x_{l-1}\|$ .

La forma canónica de Jordan de una matriz se utiliza para realizar el estudio de la convergencia de este método, obteniéndose el siguiente resultado básico (ver [37]).

**Teorema 1** *El esquema iterativo (3) —o equivalentemente, el Algoritmo 1— converge a la única solución del sistema (1), para cualquier vector inicial  $x_0$ , si y sólo si, el radio espectral de la matriz de iteración  $M^{-1}N$  es menor que 1.*

Puesto que el radio espectral de una matriz es el ínfimo del conjunto de todas las normas matriciales de dicha matriz, (ver [24]), intentamos buscar condiciones bajo las cuales el radio espectral es menor que 1 o que una determinada norma matricial sea inferior a 1. Para estudiar esta caracterización usaremos diferentes clases de matrices del sistema y diferentes particiones de la misma. En particular, investigaremos tres clases de matrices, las denominadas *matrices monótonas*, las *H-matrices*, y por último las *matrices simétricas definidas positivas*. La elección de estos tres grupos no es casual sino que obedece a causas científicas. En efecto, los tres tipos de matrices aparecen en la discretización de diferentes problemas de ecuaciones en derivadas parciales, como por ejemplo la ecuación del calor o la ecuación de difusión neutrónica [8]. Por ejemplo, si se discretiza con la técnica de diferencias finitas la ecuación del calor con condiciones de frontera constantes (condición de Dirichlet) en el cuadrado unidad, utilizando una discretización de cinco puntos, se obtiene la matriz

$$A = \begin{bmatrix} B & -I & & & \\ -I & B & -I & & \\ & -I & B & -I & \\ & & \ddots & \ddots & \ddots \\ & & & -I & B & -I \\ & & & & -I & B \end{bmatrix},$$



donde  $I$  denota la matriz identidad y  $B$  con el mismo tamaño que  $I$  es

$$B = \begin{bmatrix} 4 & -1 & & & & & \\ -1 & 4 & -1 & & & & \\ & -1 & 4 & -1 & & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 4 & -1 \\ & & & & & -1 & 4 \end{bmatrix}.$$

Es fácil ver que  $A$  es simétrica, también es fácil comprobar que su inversa es no negativa, luego es una matriz monótona. Además, se puede comprobar directamente que cumple la definición de  $M$ -matriz y por tanto de  $H$ -matriz (ver definiciones 1 y 2). En el clásico trabajo de Varga [38] o en el más reciente de Frommer y Szyld [18] se puede consultar una extensa bibliografía sobre aplicaciones de estas matrices y ejemplos de ellas. Sin embargo, hay otras matrices que no son ni  $M$ -matrices ni  $H$ -matrices y que aparecen en aplicaciones.

**Definición 1** *Se dice que una matriz no singular de tamaño  $n \times n$  es **monótona** si su inversa es una matriz no negativa.*

Una subclase muy interesante de matrices monótonas está formada por las llamadas  $M$ -matrices definidas a continuación. Así como una matriz simétrica definida positiva tiene muchas caracterizaciones, como por ejemplo, que todos sus valores propios son reales positivos, las  $M$ -matrices y  $H$ -matrices también tienen muchas caracterizaciones (ver [5]).

**Definición 2** *Sea  $A = [a_{ij}]$  una matriz cuadrada de orden  $n$ .*

- (i) *Si  $A$  es real, se dice que es una  **$M$ -matriz** si se puede escribir como  $A = sI - B$  donde  $s > \rho(A)$  y  $B$  es una matriz cuyos elementos son todos no negativos.*
- (i) *Si  $A$  es compleja, se dice que es una  **$H$ -matriz** si la matriz  $\langle A \rangle$  cuyos elementos fuera de la diagonal son  $-|a_{ij}|$ ,  $i \neq j$  y los de la diagonal  $|a_{ii}|$  es una  $M$ -matriz.*

Como se puede observar de la definición anterior, las propiedades de las  $M$ -matrices están ligadas a las de matrices no negativas. Toda matriz no negativa tiene la peculiar propiedad de que su radio espectral es valor propio. Además, asociado a dicho valor propio existe un vector propio cuyos elementos son no negativos. Estas propiedades se engloban en el conocido teorema de Perron-Frobenius (ver [37]). Por  $|A|$  denotamos a una matriz cuyos elementos son los de  $A$  en valor absoluto. De la misma forma,  $A \geq O$  significa que los elementos de  $A$ ,  $a_{ij}$ , son números reales no negativos.

Usaremos a continuación particiones de alguno de los tipos siguientes (ver [37], [18] y [34])

**Definición 3** *Sea  $A = M - N$  una partición de la matriz  $A$ .*

- (i) Se dice que dicha partición es **regular** si  $M^{-1} \geq O$  y  $N \geq O$ .
- (ii) Se dice que dicha partición es **débilmente regular** si  $M^{-1} \geq O$  y  $M^{-1}N \geq O$ .
- (iii) Se denomina **H-partición** si  $\langle M \rangle - |N|$  es una  $M$ -matriz.
- (iv) Se dice que es una **partición H-compatible** si  $\langle A \rangle = \langle M \rangle - |N|$ .
- (v) Si  $A$  es simétrica, se dice que es una **partición P-regular** si  $M^T + N$  es simétrica definida positiva.

De la no negatividad de las matrices implicadas en los dos primeros tipos de particiones se deduce que una partición regular es siempre débilmente regular. La partición débilmente regular definida aquí se suele decir que es del tipo 1. Permutando el papel de las matrices  $M$  y  $N$  de la partición se pueden obtener particiones del tipo 2 que también tienen propiedades interesantes. Un estudio detallado de estas particiones se puede seguir en [36].

Antes de finalizar esta sección, consideremos la iteración (4). Teniendo en cuenta que la partición (2) se puede escribir como

$$M^{-1}N = I - M^{-1}A \quad (7)$$

se tiene

$$x_l = x_{l-1} - M^{-1}(Ax_{l-1} - b), \quad l = 1, 2, \dots, \quad (8)$$

que utilizaremos en la Sección 6. Este esquema iterativo es el que se usa para presentar el método (estacionario) de Richardson, donde la matriz  $M^{-1}$  es un múltiplo de la matriz identidad (ver [22] y [23]).

### 3 Métodos síncronos

Si en la matriz  $A = [a_{ij}]$ , consideramos que los elementos de la diagonal son todos no nulos, es decir,  $a_{ii} \neq 0$ ,  $i = 1, 2, \dots, n$ , entonces al considerar la partición

$$A = D - L - U$$

donde  $D = [d_{ij}]$ , es una matriz diagonal con

$$d_{ij} = \begin{cases} a_{ij}, & \text{si } i = j, \\ 0, & \text{si } i \neq j, \end{cases}$$

y  $L = [l_{ij}]$  y  $U = [u_{ij}]$  son matrices estrictamente triangulares —inferior y superior, respectivamente— con

$$l_{ij} = \begin{cases} -a_{ij}, & \text{si } i > j, \\ 0, & \text{si } i \leq j, \end{cases} \quad u_{ij} = \begin{cases} -a_{ij}, & \text{si } i < j, \\ 0, & \text{si } i \geq j, \end{cases}$$

se pueden obtener los métodos clásicos de Jacobi ( $M = D$  y  $N = L + U$ ), Gauss-Seidel ( $M = D - L$  y  $N = U$ ) y SOR ( $M = \frac{1}{\omega}(D - \omega L)$  y  $N = \frac{1}{\omega}[(1 - \omega)D + \omega U]$ ), donde  $\omega \in \mathbb{R}$ .

El primer resultado interesante de convergencia que se obtiene es el siguiente.

**Teorema 2** Consideremos el sistema lineal (1). Entonces el método iterativo (3) —o equivalentemente, el Algoritmo 1— converge a la solución única del sistema (1), para cualquier vector inicial  $x_0$ , si se cumple alguna de las siguientes condiciones:

- (i)  $A$  es una matriz monótona y  $A = M - N$  es una partición débilmente regular de  $A$ .
- (ii)  $A$  y  $M$  son  $H$ -matrices y  $A = M - N$  es una  $H$ -partición de  $A$ .
- (iii)  $A$  es una  $H$ -matriz y  $A = M - N$  es  $H$ -compatible.
- (iv)  $A$  es una matriz simétrica definida positiva y  $A = M - N$  es una partición  $P$ -regular. Además, en este caso, el método (3) es convergente si y sólo si la matriz  $A$  es definida positiva.

El primer apartado de este teorema se demuestra en [37]. En realidad, esta parte se puede establecer con varias condiciones equivalentes así como bajo hipótesis más generales (ver [14]). La convergencia para  $H$ -matrices y  $H$ -particiones se puede ver en [18]. Además, en [15] pueden encontrarse resultados de convergencia para matrices no singulares cuya matriz de iteración es simétrica definida positiva. Cuando la matriz del sistema  $A$  es simétrica definida positiva y la partición (2) es  $P$ -regular, el resultado se encuentra en [34].

Como hemos mencionado en la sección anterior, dadas dos particiones convergentes de una misma matriz  $A = M_1 - N_1 = M_2 - N_2$ , resulta interesante comparar los radios espectrales de las respectivas matrices de iteración,  $\rho(M_1^{-1}N_1)$  y  $\rho(M_2^{-1}N_2)$ , para poder establecer cuál de ellas converge más rápidamente. Fueron Varga [37] y posteriormente Woźnicki [39] quienes introdujeron los primeros y más conocidos resultados de comparación que recogemos en el siguiente teorema.

**Teorema 3** Sea  $A$  una matriz monótona y  $A = M_1 - N_1 = M_2 - N_2$  dos particiones regulares. Si se cumple alguna de las siguientes condiciones:

- (i)  $N_1 \leq N_2$ , es decir,  $N_2 - N_1 \geq O$ ,
- (ii)  $M_1^{-1} \geq M_2^{-1}$ ,

entonces

$$\rho(M_1^{-1}N_1) \leq \rho(M_2^{-1}N_2) < 1.$$

Condiciones de comparación análogas a las del apartado (i) del Teorema 3 para matrices simétricas definidas positivas pueden encontrarse en [30]. Además, en [14] y [36] pueden encontrarse otras condiciones de comparación, así como las relaciones existentes entre ellas.

Inspirados posiblemente por el método de Jacobi por bloques, O'leary y White en 1985 [32] introdujeron el método iterativo basado en más de una partición, con el objeto de construir algoritmos para ordenadores paralelos, lo

que en la década de los 80 tuvo un gran impulso. Consideremos  $p$  particiones de la matriz del sistema (1)

$$A = M_k - N_k, \quad k = 1, 2, \dots, p,$$

y las matrices diagonales no negativas  $E_k$ ,  $k = 1, 2, \dots, p$ , tales que su suma es la matriz identidad. El conjunto de  $p$  particiones de  $A$  junto con las matrices de peso  $E_k$  se conoce con el nombre de *multipartición* de  $A$ . Entonces se puede construir el siguiente algoritmo.

**Algoritmo 2. (Multipartición).** Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,
  - 1.1. Para  $k = 1, 2, \dots, p$ 
    - 1.1.1. Resolver  $M_k y_{k,l} = N_k x_{l-1} + b$
  - 1.2.  $x_l = \sum_{k=1}^p E_k y_{k,l}$

Cada subsistema de este algoritmo se puede resolver en paralelo por cada uno de los  $p$  procesadores de un computador paralelo. Después de obtener cada procesador el correspondiente vector  $y_{k,l}$  en cada iteración  $l$ , se construye la solución global según la ecuación del paso 1.2 del Algoritmo 2, produciendo una *sincronización* de las soluciones parciales o de los correspondientes procesadores en cada iteración. En esta ecuación se observa que cada vector obtenido se multiplica por la matriz de peso  $E_k$  correspondiente, luego los elementos nulos de la diagonal de esta matriz anularán las correspondientes componentes de ese vector. En consecuencia, en ese vector el correspondiente procesador no necesita calcular dicha componente. De esta forma el método de multipartición distribuye el cálculo de las  $n$  componentes del vector en cada iteración entre los diferentes procesadores, de forma que el procesador  $k$  sólo calcula la componente  $i$  si el elemento  $i$  de la diagonal de la matriz  $E_k$  es diferente de 0. Además, si los elementos no nulos son distintos de 1, este método pondera la aportación de los diferentes procesadores.

Supongamos que  $A$  es una matriz *por bloques*  $A = [A_{ij}]$ ,  $i, j = 1, 2, \dots, p$ . Si se consideran las  $p$  particiones con  $M_k = \text{diag}(A_{11}, A_{22}, \dots, A_{pp})$  y las matrices peso  $E_k$  como las matrices diagonales que tienen 1 en los elementos diagonales correspondientes al bloque  $A_{kk}$  y nulos los demás elementos, se obtiene el método de Jacobi por bloques mencionado anteriormente.

Los resultados básicos de convergencia del método de multipartición son los siguientes.

**Teorema 4** *Considérese el sistema lineal (1) y una multipartición de la matriz del sistema  $A = M_k - N_k$ , con las matrices de peso  $E_k$ ,  $k = 1, 2, \dots, p$ . Entonces el esquema iterativo explicitado en el Algoritmo 2 converge a la solución del sistema (1), para todo vector inicial  $x_0$ , si se cumple alguna de las tres condiciones siguientes:*

- (i) la matriz del sistema es monótona y las  $p$  particiones de la multipartición débilmente regulares.
- (ii)  $A$  es simétrica definida positiva y las  $p$  particiones de una multipartición de  $A$  son  $P$ -regulares, con las matrices de peso  $E_k$ ,  $k = 1, 2, \dots, p$ , escalares (9). En este caso, la convergencia se da si y sólo si,  $A$  es simétrica definida positiva.
- (iii) la matriz del sistema lineal es  $H$ -matriz y las  $p$  particiones de la multipartición son particiones  $H$ -compatibles.

Las primeras dos condiciones se dan en [32]. Para matrices definidas positivas se obtiene el resultado con la particularidad de que las matrices peso  $E_k$  deben ser escalares, es decir,

$$E_k = \alpha_k I, \quad \text{con} \quad \sum_{k=1}^p \alpha_k = 1, \quad (9)$$

(ver [32]). El caso (iii) para  $H$ -matrices se estudia en [10].

Los dos algoritmos anteriores corresponden a métodos que se pueden denominar *estacionarios*, en el sentido de que en cada iteración principal los cálculos internos tienen los parámetros constantes. Por ejemplo, en el Algoritmo 2 se podría pensar que, en el bucle definido por el índice  $k$ , en alguna partición se resuelva el sistema con diferente número de iteraciones que en otras particiones. En este caso se tendría un método *no estacionario*. El método clásico no estacionario viene dado cuando el sistema del paso 2 del Algoritmo 1 o los subsistemas del paso 1.1.1 del Algoritmo 2 se resuelven de forma indirecta utilizando particiones de las matrices  $M_k$ .

Supongamos que la matriz  $M_k$  de la partición de la matriz  $A$  del sistema admite la partición *interna*  $M_k = B_k - C_k$ ; entonces se tiene el siguiente algoritmo no estacionario.

**Algoritmo 3.** (Síncrono de dos etapas con multipartición). Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,
  - 1.1. Para  $k = 1, 2, \dots, p$ 
    - 1.1.1.  $y_{k,0} = x_{l-1}$
    - 1.1.2. Para  $j = 1, 2, \dots, s(k, l)$ 
      - 1.1.2.1. Resolver  $B_k y_{k,j} = C_k y_{k,j-1} + N_k x_{l-1} + b$
  - 1.2.  $x_l = \sum_{k=1}^p E_k y_{k,s(k,l)}$

Otra vez, la solución global en cada iteración se obtiene por la ecuación del paso 1.2 del Algoritmo 3, produciendo una *sincronización* de todas las soluciones parciales o de todos los procesadores.

El modelo  $A$  presentado en [7] es un caso particular de este algoritmo cuando la partición externa se reduce a  $A = A - O$ . Por otra parte, el Algoritmo 3 se reduce al Algoritmo 2 cuando la partición *interna* de  $M_k$  es  $M_k = M_k - O$  y  $s(k, l) = s = 1$ , para  $k = 1, 2, \dots, p$  y  $l = 1, 2, \dots$ . Cuando se trabaja con una partición *externa*,  $k = 1$ , y otra interna, se obtiene el método de dos etapas, llamado a veces *iteraciones internas-externas*. Este método ha sido estudiado en [9], [17], [21], [25] y [31].

El estudio de la convergencia del Algoritmo 3 se basa, como en los otros algoritmos, en el estudio de la convergencia del error  $e_l = x_* - x_l$ , donde  $x_*$  es la solución exacta del sistema. Para ello, del algoritmo se desprende que

$$x_l = \sum_{k=1}^p E_k \left[ R_k^{s(k,l)} x_{l-1} + \sum_{j=0}^{s(k,l)-1} R_k^j B_k^{-1} (N_k x_{l-1} + b) \right],$$

donde,  $R_k = B_k^{-1} C_k$ . Luego

$$e_l = T_l e_{l-1} = T_l T_{l-1} \cdots T_1 e_0,$$

donde

$$T_l = \sum_{k=1}^p E_k \left[ R_k^{s(k,l)} + \sum_{j=0}^{s(k,l)-1} R_k^j B_k^{-1} N_k \right],$$

es la matriz de iteración en la iteración  $l$ -ésima.

Con estas consideraciones la convergencia del método no estacionario considerado se dará cuando  $\lim_{l \rightarrow \infty} \prod_{i=1}^l T_i = O$ . La demostración de este límite se basa en el siguiente resultado.

**Lema 5** *Sea  $T_1, T_2, \dots$ , una sucesión de matrices no negativas. Si existe un número real  $0 \leq \theta < 1$  y un vector  $v > 0$  de  $\mathbb{R}^n$  tal que*

$$T_l v \leq \theta v, \quad \text{para todo } l = 1, 2, \dots,$$

*entonces  $\rho(T(l)) \leq \theta^l$ , donde  $T(l) = T_l \cdots T_2 \cdot T_1$ , y por tanto  $\lim_{l \rightarrow \infty} T(l) = O$ .*

El teorema siguiente sólo aporta condiciones para que las matrices de iteración  $T_1, T_2, \dots$ , tengan las condiciones de este lema y así conseguir la convergencia. Esas condiciones son sobre la matriz de coeficientes  $A$  y para cada tipo de matriz considerada en los dos tipos de particiones, las externas e internas.

**Teorema 6** *Sea el sistema lineal (1). Entonces, el método no estacionario de multipartición en dos etapas, dado en el Algoritmo 3, es convergente a la solución exacta del sistema (1), para cualquier vector inicial  $x_0$  y cualquier número de iteraciones internas  $s(k, l) \geq 1$ ,  $k = 1, 2, \dots, p$ ,  $l = 1, 2, \dots$ , si se cumple alguna de las condiciones siguientes:*

- (i) *A es una matriz monótona, las particiones externas  $A = M_k - N_k$ ,  $k = 1, 2, \dots, p$ , son regulares y las particiones internas  $M_k = B_k - C_k$  son débilmente regulares.*
- (ii) *A es una H-matriz y las particiones externas  $A = M_k - N_k$ ,  $k = 1, 2, \dots, p$ , son H-compatibles, al igual que las particiones internas  $M_k = B_k - C_k$ .*
- (iii) *A hermítica definida positiva,  $A = M - N$ , con  $M = \text{diag}(M_1, M_2, \dots, M_p)$  hermítica y  $N$  semidefinida positiva,  $M = B - C$  partición P-regular y el número de iteraciones internas acotado.*

Las dos primeras condiciones se obtuvieron en [10], y la tercera en [12]. Obsérvese que, en este caso, la partición externa  $M$  es una matriz diagonal por bloques, que es un caso más general que el de la correspondiente matriz de Jacobi, permitiendo asignar un bloque a cada procesador y trabajar con bloques. El caso estacionario no exige la acotación de las iteraciones internas como se demuestra en [29]. Otros resultados relacionados se dan en [1].

## 4 Métodos asíncronos

Hemos comentado rápidamente que los dos métodos secuenciales y paralelos presentados en la Sección 3 son síncronos ya que el paso 1.2 del correspondiente algoritmo se realiza después de que todos los vectores  $y_{k,s(k,l)}$  que se necesitan han sido previamente calculados.

Supongamos que las matrices de peso  $E_k$  forman una partición de la matriz identidad, es decir,

$$E_k = \text{diag}[O \ I_k \ O],$$

con tamaños de bloques adecuados. Entonces en cada procesador se obtiene el valor de la parte del vector correspondiente  $E_k y_{k,s(k,l)}$ , que puede pasar a formar parte del vector de la iteración  $x_l$  sin tener que esperar a actualizar todas las componentes. De este modo, el vector actualizado en parte puede servir “más eficazmente” para los posteriores cálculos, dejando de ser válido el anterior vector  $x_{l-1}$  para los demás procesadores o cálculos. Así, el último vector actualizado puede tener grupos de componentes actualizados en diferentes momentos. Con ello, se evita la sincronización del cálculo y por tanto la sincronización de los procesadores, que en muchos casos puede ser importante en el coste computacional.

Normalmente, se considera una nueva iteración cuando algún grupo de componentes que corresponden a los elementos de un conjunto de índices,  $J_l$ , se actualiza. Para explicitar el algoritmo correspondiente interesa definir los siguientes parámetros: (i)  $k \in J_l$  es el  $k$ -ésimo bloque del vector  $x$  que se actualiza en la iteración  $l$ -ésima, (ii) el subíndice  $r(k, l)$  denota la iteración en la que el bloque  $k$ -ésimo del vector  $x$  está siendo usado para la actualización de cualquier bloque en la iteración  $l$ -ésima. Con estos parámetros podemos escribir el algoritmo del método asíncrono.

Algoritmo 4. (Asíncrono dos etapas con multipartición) . Dado un vector inicial  $x_0$ , donde sus componentes están agrupadas en  $p$  bloques,  $x_0^{(1)}, \dots, x_0^{(p)}$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,

1.1. Si  $k \in J_l$ , hacer

$$x_l^{(k)} = \sum_{k=1}^p E_k \left[ R_k^{s(k,l)} x_{r(k,l)}^{(k)} + \sum_{j=0}^{s(k,l)-1} R_k^j B_k^{-1} (N_k x_{r(k,l)}^{(k)} + b) \right],$$

en otro caso,  $x_l^{(k)} = x_{l-1}^{(k)}$ .

Para poder obtener convergencia en el método asíncrono hay que suponer las tres condiciones siguientes:

- $r(k, l) < l$  para todo  $k = 1, 2, \dots, p$ ,  $l = 1, 2, \dots$ ,
- $\lim_{l \rightarrow \infty} r(k, l) = \infty$  para todo  $k = 1, 2, \dots, p$ .
- El conjunto  $\{l \mid k \in J_l\}$  no está acotado para todo  $k = 1, 2, \dots, p$ .

La primera condición sólo asegura que grupos de componentes previamente calculadas son usados para los cálculos posteriores. La segunda condición indica que las componentes, que se utilizan para calcular otras, son actuales y no muy antiguas. Por su parte, la tercera condición indica que cada componente es actualizada un número infinito de veces. Esta condición es equivalente al concepto de sucesión regulada utilizada en [7].

Los siguientes resultados son válidos para este método asíncrono (ver [10]).

**Teorema 7** *El método del algoritmo asíncrono de dos etapas con multipartición converge a la solución exacta del sistema (1), para cualquier vector inicial  $x_0$  y cualquier sucesión de número de iteraciones internas  $s(k, l)$ ,  $l = 1, 2, \dots$ ,  $k = 1, 2, \dots, p$ , si se cumple alguna de las condiciones siguientes:*

- (i) *A es una matriz monótona, las particiones externas son regulares, las particiones internas son débilmente regulares, y los parámetros  $r(k, l)$  y los conjuntos  $J_l$ ,  $l = 1, 2, \dots$ ,  $k = 1, 2, \dots, p$ , satisfacen las tres condiciones mencionadas anteriormente.*
- (ii) *A es una H-matriz y las particiones externas e internas son particiones H-compatibles.*

El Algoritmo 4 (también los Algoritmos 1, 2 y 3) se pueden modificar ligeramente mediante la introducción de un parámetro de *relajación*  $\omega$  en la matriz de iteración. Con ello, se puede obtener mejor velocidad de convergencia. Los resultados de convergencia siguen siendo los mismos con la condición que el parámetro de relajación pertenezca a un cierto intervalo real, por ejemplo  $(0, 1)$  (ver [11] y [28]).



## 5 Método alternado

Entre el método de una partición (Algoritmo 1) y el de multipartición (Algoritmo 2) se encuentra el método alternado o de dos direcciones alternadas que fue introducido por Conrad y Wallach [13] y estudiado posteriormente por Marchuk [26]. Este método se basa en considerar dos particiones diferentes de la matriz del sistema de la forma

$$A = M - N = P - Q. \quad (10)$$

Alternando estas dos particiones se construye el siguiente algoritmo.

**Algoritmo 5.** (Direcciones alternadas). Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,
  - 1.1. Resolver  $Mx_{l+1/2} = Nx_{l-1} + b$
  - 1.2. Resolver  $Px_l = Qx_{l+1/2} + b$

La convergencia de este método se analiza mejor escribiendo las dos ecuaciones del algoritmo en una sola de la forma

$$x_l = P^{-1}QM^{-1}Nx_{l-1} + P^{-1}(QM^{-1} + I)b, \quad l = 1, 2, \dots \quad (11)$$

Con la matriz de iteración  $P^{-1}QM^{-1}N$ , Benzi y Szyld [4] obtienen los siguientes resultados.

**Teorema 8** *El método iterativo de direcciones alternadas —o equivalentemente el Algoritmo 5— converge a la solución exacta del sistema (1), para cualquier vector inicial  $x_0$ , si se cumple alguna de las condiciones siguientes:*

- (i) *la matriz  $A$  es monótona y las dos particiones (10) son particiones débilmente regulares.*
- (ii) *la matriz del sistema  $A$  es simétrica y definida positiva y las dos particiones (10) son  $P$ -regulares.*

Por otra parte, y también muy recientemente, Climent y Perea [16] introducen una generalización no estacionaria del Algoritmo 5 considerando una multipartición  $A = M_k - N_k$  para  $k = 1, 2, \dots, p$ , así como resultados de convergencia para dicho algoritmo análogos al Teorema 8.

Este método es una generalización del conocido método *ADI* que introdujeron Peaceman y Rachford [35] para resolver ecuaciones diferenciales parabólicas sobre un dominio  $\Omega$  discretizado con una malla de dos dimensiones simétrica.

Recientemente Bai, Golub y Ng [2] estudian la convergencia de sistemas lineales cuando la matriz del sistema es definida positiva pero no necesariamente

hermítica. Para ello, considerando la parte hermítica y anti-hermítica de la matriz de coeficientes

$$A = H + S = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*)$$

consideran las particiones

$$A = (\alpha I + H) - (\alpha I - S) = (\alpha I + S) - (\alpha I - H), \quad (12)$$

donde  $\alpha$  es una constante positiva. Con ello, construyen el esquema iterativo denominado *HSS*, cuyo algoritmo es el siguiente.

**Algoritmo 6.** (HSS). Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,
  - 1.1. Resolver  $(\alpha I + H)x_{l+1/2} = (\alpha I - S)x_{l-1} + b$
  - 1.2. Resolver  $(\alpha I + S)x_l = (\alpha I - H)x_{l+1/2} + b$

En este caso, la matriz de iteración es

$$T(\alpha) = (\alpha I + S)^{-1}(\alpha I - H)(\alpha I + H)^{-1}(\alpha I - S), \quad (13)$$

totalmente similar a la matriz de la ecuación (11). La convergencia de este método se consigue con las siguientes condiciones.

**Teorema 9** *Sea la matriz del sistema (1) una matriz compleja definida positiva. Entonces se cumple que el radio espectral de la matriz de iteración  $T(\alpha)$  está acotado por*

$$\sigma(\alpha) = \max \left\{ \left| \frac{\alpha - \lambda_i}{\alpha + \lambda_i} \right|; \lambda_i \in \lambda(H) \right\},$$

*y por consiguiente  $\rho(T(\alpha)) < 1$ , ya que  $\alpha > 0$ , y por tanto el algoritmo *HSS* converge a la solución exacta del sistema, para cualquier vector inicial  $x_0$ .*

El teorema anterior muestra que la velocidad asintótica de convergencia de este algoritmo está acotada por  $\sigma(\alpha)$  que sólo depende del espectro  $\lambda(H)$  de  $H$ , y no del espectro de la parte anti-hermítica  $S$ . Incluso se puede comprobar que

$$\|x_{l+1} - x_*\| \leq \sigma(\alpha)\|x_l - x_*\|,$$

indicando que la reducción del error, en una iteración, está acotada por el mismo factor que el radio espectral de la matriz de iteración. También se puede obtener que el valor óptimo del parámetro  $\alpha$  viene dado por el valor  $\sqrt{\lambda_{\max}\lambda_{\min}}$ .

## 6 Método de Schwarz

Supongamos que se quiere resolver una ecuación en derivadas parciales en un dominio  $\Omega$ . La técnica de *descomposición de dominios* consiste en considerar varios subdominios del dominio inicial y resolver dicha ecuación en cada subdominio. Estos subdominios pueden ser disjuntos o no y su elección puede venir dada no solamente por conseguir subproblemas de menor tamaño sino también porque el material de cada subdominio sea diferente o porque se quiere utilizar cálculo en paralelo.

Como estamos resolviendo la ecuación (1), consideraremos que dicha ecuación es el resultado de la discretización del problema inicial y no tendremos en cuenta la discretización misma y por tanto tampoco la malla considerada. Por ello, en esta sección consideraremos el punto de vista algebraico del método y no el variacional.

El método de Schwarz, desde el punto de vista algebraico, es un método de descomposición de dominios ya que descompone los vectores considerados del espacio vectorial  $\mathbb{E}$ , de dimensión  $n$ , en vectores de subespacios vectoriales  $\mathbb{E}_i$ , de dimensión  $n_i < n$ . Cuando

$$\sum_{i=1}^N n_i = n$$

los subdominios son disjuntos, mientras que si

$$\sum_{i=1}^N n_i > n$$

los subdominios se solapan, siendo  $N$  el número de subdominios considerados.

Para transformar vectores de  $\mathbb{E}_i$  en  $\mathbb{E}$  o viceversa, consideramos los operadores *prolongación*

$$p_i : \mathbb{E}_i \rightarrow \mathbb{E}$$

y los operadores *restricción*

$$r_i : \mathbb{E} \rightarrow \mathbb{E}_i,$$

que vienen representados en las correspondientes bases por las matrices  $P_i \in \mathbb{C}^{n \times n_i}$  y  $R_i \in \mathbb{C}^{n_i \times n}$ , respectivamente. En los dos casos de subdominios considerados, se supone que

$$\sum_{i=1}^N \mathbb{E}_i = \mathbb{E}.$$

Si en cada subdominio  $\Omega_i$  se consideran las submatrices

$$A_i = R_i A P_i,$$

el sistema inicial (1) puede ser descompuesto en varios subsistemas

$$A_i y_i = c_i.$$

Un punto importante a considerar es cómo realizar la construcción de la solución total a partir de las soluciones de los subsistemas. Cuando esta construcción se hace sin más que sumar las soluciones de cada subdominio se tiene el método de Schwarz *aditivo*, que se puede implementar con el siguiente algoritmo.

Algoritmo 7. (Schwarz aditivo). Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,

1.1. Hacer  $x_l = x_{l-1} + \theta \sum_{i=1}^N P_i A_i^{-1} R_i (Ax_{l-1} - b)$

donde se ha escrito el factor *damping*  $\theta > 0$ . En el paso 1.1 del algoritmo se observan las operaciones que se hacen en cada iteración: en cada subespacio, se restringe el vector residuo (cambiado de signo),  $r_i = Ax_{l-1} - b$ , mediante el operador  $R_i$ , se resuelve el problema local  $A_i e_i = r_i$ , después se prolonga el error local obtenido, se suman todas las soluciones locales y finalmente se añade esta corrección al vector anterior.

En este caso la matriz de iteración es

$$T = I - \theta \sum_{i=1}^N P_i A_i^{-1} R_i A.$$

De forma análoga, se puede construir la solución global a partir de las soluciones parciales de cada subespacio de forma multiplicativa, dando lugar al siguiente algoritmo.

Algoritmo 8. (Schwarz multiplicativo). Dado un vector inicial  $x_0$ ,

1. Para  $l = 1, 2, \dots$ , hasta convergencia,

1.1. Hacer  $x^{k+1} = T x^k + c$

siendo la matriz iteración

$$T = (I - P_p) \cdot (I - P_{p-1}) \cdot \dots \cdot (I - P_1) = \prod_{i=p}^1 (I - P_i)$$

donde

$$P_i = R_i^T A_i^{-1} R_i A.$$

Las operaciones que se realizan son las mismas que en el Algoritmo 7, excepto la última donde se construye la corrección componiendo las soluciones antes de añadir a la solución anterior. En este caso la matriz de iteración es

$$T = I - \prod_{i=1}^N P_i A_i^{-1} R_i A.$$

Una forma particular, aunque natural en muchos casos ([19]), de los operadores restricción y prolongación corresponde al caso en que cada columna de  $R_i^T$  está formada por los vectores canónicos sin orden establecido. Por ejemplo, si trabajamos en  $\mathbb{R}^6$  y la restricción la hacemos al subespacio  $\mathbb{R}^3$ , la matriz restricción sería

$$R_i^T = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \Pi_i \begin{bmatrix} I \\ O \end{bmatrix}, \quad (14)$$

donde  $\Pi_i$  representa una matriz permutación de orden 3.

Para mantener propiedades de simetría se suele trabajar con  $P_i = R_i^T$ . En estos casos las matrices de los subdominios  $A_i$  son submatrices principales de la matriz  $A$ .

Cuando no hay solapamiento (subdominios disjuntos) y además la matriz  $R_i^T$  se puede escribir como en (14) pero con  $\Pi_i = I$  y  $P_i = R_i^T$ , se tiene que el método aditivo de Schwarz es justamente el método de Jacobi descrito en la Sección 3, y el método multiplicativo de Schwarz es el método de Gauss-Seidel, (ver [23]).

Frommer y Szyld [19] dan un resultado de convergencia del método aditivo para  $M$ -matrices, cuando hay solapamientos entre subdominios. Para ello, hay que considerar el parámetro  $q$  que indica el número de colores necesarios para dibujar cada subdominio  $\Omega_i$ , de acuerdo con la siguiente regla: si  $\Omega_i \cap \Omega_j \neq \emptyset$ , entonces se dibujan con diferentes colores. Normalmente  $q < N$ .

**Teorema 10** *Sea la matriz del sistema (1) una  $M$ -matriz. Supongamos que las matrices  $R_i$  están definidas como en (14). Entonces, si  $\theta \leq 1/q$ , el método aditivo de Schwarz, o equivalentemente, el Algoritmo 7, converge a la solución exacta del sistema (1), para cualquier vector inicial  $x_0$ .*

Un resultado similar para el método multiplicativo de Schwarz ha sido dado por los mismos autores en [20], que expresamos a continuación. En este caso los autores prueban la convergencia para dos tipos de matrices:  $M$ -matrices y simétricas definidas positivas.

**Teorema 11** *Sea la matriz del sistema (1) una  $M$ -matriz (o simétrica definida positiva). Supongamos que las matrices  $R_i$  están definidas como en (14). Entonces, el método multiplicativo de Schwarz (Algoritmo 8) converge a la solución exacta del sistema (1), para cualquier vector inicial  $x_0$ .*

Otros resultados similares se encuentran en [3].

## 7 Consideraciones finales

Como se ha indicado en la introducción los métodos anteriores se utilizan sobre todo como preconditionadores de los métodos basados en subespacios de Krylov.

Cuando se preconditiona la matriz  $A$  mediante la matriz  $M$  de una partición de la forma

$$M^{-1}A = I - T,$$

se obtiene la matriz de iteración de alguno de los métodos estudiados. Se puede comprobar que preconditionar de esta forma equivale a resolver un sistema auxiliar donde la matriz de coeficientes es la matriz  $M$ .

Como se puede observar en los algoritmos anteriores, siempre hay un paso donde se debe resolver un sistema lineal con una matriz de coeficientes, en el caso más sencillo  $M$ . Este sistema se puede resolver invirtiendo  $M$  si es una matriz sencilla de invertir como en el caso del método de Jacobi. En otro caso, se puede utilizar un método directo si la estructura de la matriz lo aconseja, por ejemplo, cuando es triangular. También se puede resolver este segundo sistema por una nueva partición, como es el caso del método de dos etapas, o incluso por métodos de tipo subespacios de Krylov.

La teoría de los métodos iterativos basados en particiones tienen diferentes problemas a estudiar. Por ejemplo, en el método en dos etapas, las exigencias en las particiones internas y externas son diferentes. Se puede estudiar la convergencia cuando la partición externa es también débilmente regular.

El método de direcciones alternadas cuando se aproximan los dos sistemas del Algoritmo 6 por otras particiones no está estudiado. En el caso de los métodos de Schwarz, la convergencia está dada para  $M$ -matrices en los dos Algoritmos 7 y 8, y para matrices simétricas definidas positivas para el último algoritmo. Sin embargo, no hay resultados de convergencia para  $H$ -matrices.

No se ha comentado nada respecto a la convergencia de sistemas cuando la matriz del mismo no es invertible, ya sea cuadrada o rectangular. En este caso, también hay bastantes resultados análogos a los presentados para sistemas singulares (ver por ejemplo [6] y [27]), pero de cualquier modo falta mucho por obtener, especialmente en el caso asíncrono.

**Agradecimiento:** Quiero agradecer a los profesores y compañeros Joan Josep Climent, José Marín, Violeta Migallón, José Penadés y Carmen Perea por proporcionarme ideas para desarrollar este trabajo así como sugerencias que, sin duda, han mejorado la comprensión del mismo. En las tesis doctorales de todos ellos se encuentran otros resultados relacionados con el tema. También quiero agradecer a Enrique Fernandez-Cara y José Mas las sugerencias que han hecho sobre la versión preliminar. Finalmente, agradecer a la Dirección General de Investigación DGI y otros organismos oficiales que continuamente han subvencionado nuestra labor, actualmente con el proyecto BFM 2001-2641.

## Referencias

- [1] Z.-Z. Bai. A class of two-stage iterative methods for systems of weakly nonlinear equations. *Numerical Algorithms*, 14: 295–319, 1997.

- [2] Z.-Z. Bai, G.H. Golub y M.K. Ng. Hermitian and skew-hermitian splitting methods for non-hermitian positive definite linear systems. *SIAM J. on Matrix Analysis*, aparecerá, 2003.
- [3] M. Benzi, A. Frommer, R. Nabben y D.B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89: 605–639, 2001.
- [4] M. Benzi y D.B. Szyld. Existence and Uniqueness of splittings for stationary iterative methods with applications to alternating methods. *Numerische Mathematik*, 76: 309–321, 1997.
- [5] A. Berman y R.J. Plemmons. *Nonnegative matrices in the mathematical sciences*. SIAM, Philadelphia, Pensilvania, 1994. 77–106, 1998.
- [6] R. Bru, R. Cantó y J.J. Climent. On  $M$ -multisplittings of singular  $M$ -matrices with applications to Markov chains. *Numerical Linear Algebra with Applications*, 5: 299–311, 1998.
- [7] R. Bru, L. Elsner y M. Neumann. Models of parallel chaotic iteration methods. *Linear Algebra and its Applications*, 13:175–192, 1988.
- [8] R. Bru, D. Ginestar, T. Manteuffel, J. Marín, J. Mas y G. Verdú. Iterative schemes for the neutron diffusion equations. *Computers and Mathematics with Applications*, 44(10-11):1307–1323, 2002.
- [9] R. Bru, V. Migallón y J. Penadés. Chaotic methods for the parallel solution of linear systems. *Computing Systems in Engineering*, 6(4-5):385–390, 1995.
- [10] R. Bru, V. Migallón, J. Penadés y D.B. Szyld. Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Transactions on Numerical Analysis*, 3: 24-38, 1995.
- [11] M. Castel, V. Migallón y J. Penadés. Convergence of non-stationary parallel multisplitting methods for Hermitian positive definite matrices. *Mathematics of Computations*, 67(221): 209–220, 1998.
- [12] M. Castel, V. Migallón y J. Penadés. On parallel two-stage methods for Hermitian positive definite matrices with applications to preconditioning. *Electronic Transactions on Numerical Analysis*, 12:88–112, 2001.
- [13] V. Conrad e Y. Wallach. Alternating methods for sets of linear equations. *Numerische Mathematik*, 32:1105–108, 1979.
- [14] J.J. Climent y C. Perea. Some comparison theorems for weak nonnegative splittings of bounded operators. *Linear Algebra and its Applications*, 275-276:77–106, 1998.
- [15] J.J. Climent y C. Perea. Convergence and comparison theorems for multisplittings. *Numerical Linear Algebra with Applications*, 6: 93–107, 1999.

- [16] J.J. Climent y C. Perea. Convergence and comparison theorems for a generalized alternating iterative method. *Applied Mathematics and Computation*, 143:1–14, 2003.
- [17] A. Frommer y G. Mayer. Convergence of relaxed parallel multisplitting methods. *Linear Algebra and its Applications*, 119:141–152, 1989.
- [18] A. Frommer y D.B. Szyld.  $H$ -splittings and two-stage iterative methods. *Numerische Mathematik*, 63:345–356, 1992.
- [19] A. Frommer y D.B. Szyld. Weighted max norms, splittings and overlapping additive Schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.
- [20] A. Frommer y D.B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using matrix max norms. *SIAM J. Numerical Analysis*, 39(2):463–479, 2001.
- [21] G.H. Golub y M.L. Overton. Convergence of a two-stage Richardson iterative procedure for solving systems of linear equations. *En Numer. Anal., G.A. Watson, ed., Proc. of the ninth Biennial Conference, Dundee, Scotland, 1981, Lecture Notes in Mathematics 912*, Springer Verlag, New York, 128–139, 1982.
- [22] G.H. Golub y M.L. Overton. The convergence of inexact Chebyshev and Richardson iterative methods for solving linear systems. *Numerische Mathematik*, 53:571–593, 1988.
- [23] W.Hackbusch. *Iterative solutions of large sparse systems of equations*. Springer-Verlag, New York, New York, 1994.
- [24] P. Lancaster y L. Tismeneski. *The theory of matrices*. Academic press, New York, New York, 1985.
- [25] P.J. Lanzkron, D.J. Rose y D.B. Szyld. Convergence of nested classical iterative methods for linear systems. *Numerische Mathematik*, 58:685–602, 1991.
- [26] G.I. Marchuk. Splitting and alternating direction methods. *En Handbook of Numerical Analysis, P.G. Ciarlet y J.L. Lions, ed., volumen I*, North Holland, New York, 1990.
- [27] I. Marek y D.B. Szyld. Comparison theorems for the convergence factor of iterative methods for singular matrices. *Linear Algebra and its Applications*, 316(1-3): 67–87, 2000.
- [28] J. Mas, V. Migallón, J. Penadés y D.B. Szyld. Nonstationary parallel relaxed multisplitting methods. *Linear Algebra and its Applications*, 241-243:733–747, 1996.



- [29] V. Migallón y J. Penadés. Convergence of two-stage iterative methods for Hermitian positive definite matrices. *Applied Mathematics Letters*, 10(3):79–83, 1997.
- [30] R. Nabben. A note on comparison theorems for splittings and multisplittings of Hermitian positive definite matrices. *Linear Algebra and its Applications*, 233:67–80, 1996.
- [31] N.K. Nichols. On the convergence of two-stage iterative processes for solving linear equations. *SIAM J. Numerical Analysis*, 10:215–227, 1984.
- [32] D.P. O’Leary y R.E. White. Multi-splittings of matrices and parallel solution of linear systems. *SIAM J. on Algebraic and Discrete Methods*, 6:630–640, 1985.
- [33] J.M. Ortega. *Introduction to parallel and vector solution of linear systems*. Plenum Press, New York, New York, 1988.
- [34] J.M. Ortega. *Numerical Analysis: A second course*. SIAM, Philadelphia, Pensilvania, 1990.
- [35] D. Peaceman y H. Rachford. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3: 28–41, 1955.
- [36] C. Perea. *Teoremas de convergencia y comparación para particiones y multiparticiones*. Tesis doctoral, Univ. de Alicante, 1998.
- [37] R. Varga. *Matrix iterative analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1966. Segunda versión publicada por Springer-Verlag, 2000.
- [38] R. Varga. On recurring theorems on diagonal dominance. *Linear Algebra and its Applications*, 13:1–9, 1976.
- [39] Z.I. Woźnicki. Nonnegative splitting theory. *Japan Journal on Industrial and Applied Mathematics*, 11:289–342, 1994.



## Inclusiones diferenciales, Matemática Difusa y aplicaciones \*

E. FERNÁNDEZ-CARA<sup>1</sup>, M.A. ROJAS-MEDAR<sup>2</sup> Y  
A.J. VIERA BRANDÃO<sup>3</sup>

<sup>1</sup> Departamento de Ecuaciones Diferenciales y Análisis Numérico,  
Universidad de Sevilla

<sup>2</sup> Departamento de Matemática Aplicada,  
IMECC-UNICAMP, Campinas-SP, (Brasil)

<sup>3</sup> Departamento de Matemática  
ICEB-UFOP, Ouro Preto-MG (Brasil)

cara@numer.us.es, marko@ime.unicamp.br,  
adilson@iceb.ufop.br

### Resumen

En este trabajo diremos cómo se pueden modelar ciertos sistemas tomados de la realidad que usan conceptos propios de la Matemática Difusa (conjuntos, multifunciones e inclusiones diferenciales “fuzzy”). Consideraremos problemas de valor inicial para inclusiones diferenciales “fuzzy” y analizaremos la existencia de solución local. También nos referiremos a la estabilidad de los puntos de equilibrio de las inclusiones diferenciales “fuzzy”. Finalmente, mostraremos algunas aplicaciones de este desarrollo a problemas que aparecen en Biología.

**Palabras clave:** *Inclusiones diferenciales difusas, problemas de valor inicial, estabilidad de puntos críticos, aplicaciones a la dinámica de poblaciones.*

**Clasificación por materias AMS:** *46S40, 34K60, 34D20, 35Q80*

## 1 Introducción

Es bien conocida la amplia variedad de aplicaciones de las ecuaciones diferenciales y de la teoría de control de sistemas gobernados por éstas. Sin embargo, los sistemas deterministas parecen ser demasiado restrictivos para

---

\*El primer autor ha sido financiado por D.G.E.S.-España, Proyectos PB98-1134 y BFM2000-1317 y el segundo por CNPq-Brasil, Proyecto 300116/93-4 (RN) y FAPESP, Proyecto 01/07557-3.

Fecha de recepción: 2 de septiembre de 2002

describir la evolución de una gran clase de fenómenos que se dan en la realidad. Frecuentemente aparecen problemas con algunas de las siguientes características, conocidas con el término inglés “fuzzyness”:

- Desconocimiento del medio ambiente futuro del sistema.
- Falta de determinismo (lo que hace en algunos casos imposible describir de forma completa la dinámica).
- Desconocimiento exacto de la manera en la que el estado del sistema depende de los posibles parámetros de control.
- Múltiples dinámicas “admisibles”, etc.

En un primer análisis, es posible superar estas dificultades con lenguaje matemático apropiado haciendo uso de *inclusiones diferenciales*. En el caso más sencillo, éstas toman la forma

$$\dot{x} \in F(t, x), \quad (1)$$

donde  $x : [0, T] \mapsto \mathbb{R}$  es una función escalar que describe el estado del sistema y  $F : [0, T] \times \mathbb{R} \mapsto \mathcal{P}(\mathbb{R})$  es una multifunción. Aquí y en lo que sigue, la variable  $t$  tiene asignado el significado “tiempo”,  $\dot{x}$  denota la derivada de  $x$  respecto de  $t$  y  $\mathcal{P}(\mathbb{R})$  es la familia de todos los subconjuntos de  $\mathbb{R}$ .

Las inclusiones diferenciales (1) indican, *grosso modo*, la forma en la que en cada instante de tiempo  $\dot{x}(t)$  depende del estado del sistema  $x(t)$  y del valor de  $t$ . Generalmente, (1) es complementada con una condición inicial

$$x(0) = x_0 \quad (2)$$

con  $x_0 \in \mathbb{R}$  dado, que indica que conocemos el estado del sistema en un instante inicial.

Los problemas de valor inicial (1) – (2) han sido intensamente estudiados en los últimos años y se conocen diversas técnicas para su resolución, entre las cuales sobresalen los métodos de punto fijo y de selección continua (parte de este desarrollo ha sido resumido en la Sección 2).

Por otra parte, en una gran cantidad de ocasiones, los sistemas analizados permiten cierta capacidad de maniobra, es decir, pueden ser controlados al menos parcialmente por nosotros. Cuando la situación no es excesivamente compleja, un modelo aceptable es el siguiente:

$$\begin{cases} \dot{x} = f(t, x, v(t)), & v(t) \in U, \\ x(0) = x_0, \end{cases} \quad (3)$$

donde  $f : [0, T] \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$  y  $v : [0, T] \mapsto \mathbb{R}$  ( $v$  es el control, una función que podemos elegir). Bajo hipótesis razonables, puede probarse la equivalencia del problema de control (3) y el problema de valor inicial (1) – (2) con

$$F(t, x) = f(t, x, U) = \{ h \in \mathbb{R} ; \exists w \in U \text{ tal que } h = f(t, x, w) \}.$$

Aquí,  $U \subset \mathbb{R}$  juega el papel de una restricción del sistema (de origen físico, económico, etc.), en el sentido de que ha de respetarse la relación  $v(t) \in U$  para cada  $t \in [0, T]$ .

El sistema (3) es de carácter determinista: bajo condiciones aceptables de regularidad para  $f$ , fijados un dato inicial  $x_0$  y un control  $v$  con valores en  $U$  por ejemplo continuo, queda determinado el estado  $x = x(t)$ , i.e. la dinámica del sistema. No obstante, hay muchas situaciones en las que tan sólo es posible elegir el control de manera *difusa*. Esto ocurrirá, por ejemplo, cuando haya constancia de que nuestra elección está afectada por un “ruido” y no es captada nítidamente por el sistema. Entonces conviene modelar la situación recurriendo a una inclusión diferencial “fuzzy”. La descripción y análisis de estas inclusiones diferenciales es uno de los principales objetivos de este trabajo.

De momento, digamos tan sólo que una inclusión diferencial “fuzzy” es una relación

$$\dot{x} \in H(t, x), \quad (4)$$

donde  $H : [0, T] \times \mathbb{R} \mapsto \mathbb{E}(\mathbb{R})$  es una “multifunción fuzzy”. Aquí,

$$\mathbb{E}(\mathbb{R}) = \{ u : \mathbb{R} \mapsto [0, 1] ; L_a u \in \mathbb{K}(\mathbb{R}) \text{ para } 0 \leq a \leq 1 \},$$

$\mathbb{K}(\mathbb{R})$  es la familia de los compactos no vacíos de  $\mathbb{R}$ , para cada  $a > 0$   $L_a u = \{ x \in \mathbb{R} ; u(x) \geq a \}$  es el *conjunto de nivel a de u* y  $L_0 u$  es la adherencia del conjunto  $\{ x \in \mathbb{R} ; u(x) > 0 \}$ , i.e. el *soporte de u*.

Obviamente,  $\mathbb{E}(\mathbb{R})$  es una extensión de  $\mathbb{K}(\mathbb{R})$  vía la *inyección canónica*  $A \mapsto \chi_A$ , donde  $\chi_A$  es la función característica de  $A$ , véase (11).

Resolver (4) en el intervalo  $[0, T]$  consiste en encontrar una función  $x : [0, T] \mapsto \mathbb{R}$  absolutamente continua (al menos) tal que, para casi todo  $t$ , se tenga

$$\dot{x}(t) \in H(t, x(t)) \quad (5)$$

en un sentido difuso (véase más adelante).

También está claro que (4) es una extensión de (1) vía la inyección  $A \mapsto \chi_A$ . Conviene por tanto desarrollar una teoría “fuzzy” que permita abordar este problema y, para ello, hay que extender apropiadamente las ideas y conceptos habituales.

El resto de este trabajo está organizado como sigue. En las Secciones 2 y 3, recordaremos brevemente varios resultados relacionados con las inclusiones diferenciales “clásicas” (1). En la Sección 4, daremos un repaso a los conceptos básicos de la teoría “fuzzy”. Posteriormente, en la Sección 5, mostraremos algunos resultados para problemas de valor inicial para inclusiones diferenciales “fuzzy”. Finalmente, la Sección 6 está dedicada a mostrar en qué forma la resolución de inclusiones “fuzzy” análogas a (4) puede ser aplicada al análisis de la dinámica de poblaciones.

El primer autor desea agradecer las conversaciones sobre el tema tenidas con los Prof. Dr. Rodney C. Bassanezi de la UNICAMP (Brasil), Prof. Geraldo Nunes Silva de la UNESP, São José do Rio Preto (Brasil), Prof. Dr. Heriberto Román-Flores de la Universidad de Tarapacá, Arica (Chile) y, obviamente, con

sus alumnos de Doctorado en esta área: Lucelina Batista dos Santos y Yurilev Chalco Cano.

## 2 Multifunciones e inclusiones diferenciales “clásicas”

En esta sección presentaremos varias situaciones que justifican la conveniencia de introducir el concepto de inclusión diferencial. De manera no rigurosa y bastante simplificada, recordaremos algunas definiciones y cuestiones trascendentes y, también, motivaremos las técnicas utilizadas para responder a algunas (realmente muy pocas) de estas cuestiones.

En lo que sigue,  $\langle \cdot, \cdot \rangle$  y  $\|\cdot\|$  denotarán, para distintos valores de  $m$ , el producto escalar Euclídeo en  $\mathbb{R}^m$  y la norma asociada, respectivamente.

Comenzaremos recordando que el objeto de estudio de la teoría de control de sistemas gobernados por ecuaciones diferenciales es, en términos simples, *gobernar* el comportamiento de tales sistemas. Así, admitamos que estamos en presencia de un proceso sometido a control, modelado por la ecuación diferencial

$$\dot{x} = f(t, x, v(t)), \quad (6)$$

donde  $f : [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \mapsto \mathbb{R}^n$  es, por simplicidad, una función regular. La función  $v : [0, T] \mapsto \mathbb{R}^k$  es el *control* y debe satisfacer ciertas propiedades de regularidad (usualmente la exigencia más débil es la medibilidad) y una restricción geométrica:

Para todo  $t \in [0, T]$ , ha de tenerse  $v(t) \in U$ , donde  $U \subset \mathbb{R}^k$  es un conjunto dado.

Usualmente  $U$  está ligado a limitaciones físicas, económicas, etc. de los dispositivos de control. Un control  $v$  que verifica estas restricciones se denomina *control admisible*.

En (6), el control debe ser escogido de tal forma que la solución  $x$  satisfaga ciertas condiciones deseadas. Por ejemplo, que  $x$  minimice un cierto funcional (teoría de control óptimo), tome un valor dado para  $t = T$  (controlabilidad), satisfaga adecuadas restricciones geométricas (teoría de la viabilidad), etc.

Es importante comprender bien cómo depende la trayectoria  $x = x(t)$  del control  $v$ . Para ello, conviene razonar como sigue. Supongamos que el grafo de una solución  $x$  de (6) pasa por el punto  $(t_0, x_0) \in [0, T] \times \mathbb{R}^n$ . En este punto, la pendiente de la tangente a la trayectoria es  $f(t_0, x_0, w)$  para algún  $w \in U$ ; es decir, la trayectoria debe *elegir su velocidad* en el conjunto de velocidades permitidas

$$F(t_0, x_0) = \{ h \in \mathbb{R}^n ; \exists w \in U \text{ tal que } h = f(t_0, x_0, w) \}.$$

Por tanto, cualquier solución de (6) asociada a un control  $v$  satisface la inclusión diferencial

$$\dot{x} \in F(t, x).$$

Esto conduce de forma natural a la formulación del problema matemático siguiente:

**Cuestión 1:** Dada una multifunción  $F : [0, T) \times \mathbb{R}^n \mapsto \mathcal{P}(\mathbb{R}^n)$ , i.e. una relación que a cada punto  $(t, x) \in [0, T) \times \mathbb{R}^n$  asocia un conjunto no vacío  $F(t, x) \subset \mathbb{R}^n$ , fijado el punto  $x_0 \in \mathbb{R}^n$ , hallar una solución en  $[0, T)$  del problema de valor inicial para una *inclusión diferencial*

$$\begin{cases} \dot{x} \in F(t, x), \\ x(0) = x_0. \end{cases} \quad (7)$$

Aquí, por *solución en*  $[0, T)$  se entiende una función absolutamente continua  $x : [0, T) \mapsto \mathbb{R}^n$  que verifica (5) en casi todo  $t \in [0, T)$  y verifica también la *condición inicial*  $x(0) = x_0$ .

Además de los sistemas controlables de la forma (6), existe otra amplia clase de fenómenos que pueden ser descritos por inclusiones diferenciales. Esta clase incluye sistemas donde el control está sometido a restricciones que cambian con el tiempo y/o el estado

$$\dot{x} = f(t, x, u(t)), \quad u(t) \in U(t, x(t)),$$

sistemas de desigualdades diferenciales,

$$f_i(t, x, \dot{x}) \leq 0, \quad i \in I,$$

ecuaciones diferenciales implícitas, sistemas controlables con restricciones sobre el espacio de fases, etc.

Otro ejemplo está dado por las inecuaciones variacionales de evolución. En este caso, el problema se reduce a la búsqueda de una función absolutamente continua  $x : [0, T) \mapsto \mathbb{R}^n$  tal que

- i)  $x(t) \in K$  y
- ii)  $\sup_{y \in K} \langle \dot{x}(t) - f(t, x(t)), x(t) - y \rangle = 0$

para casi todo  $t \in [0, T)$ , donde  $K \subset \mathbb{R}^n$  es un cono convexo cerrado de  $\mathbb{R}^n$ . La multifunción que aparece ahora es  $F_K$ , con

$$F_K(t, x) = \{ h \in \mathbb{R}^n ; \sup_{y \in K} \langle h - f(t, x), x - y \rangle = 0 \}$$

para cada  $(t, x) \in [0, T) \times \mathbb{R}^n$ .

Las ecuaciones diferenciales  $\dot{x} = f(t, x)$  con segundo miembro  $f$  discontinuo en  $x$  también pueden ser tratadas sumergiendo  $f$  en una multifunción  $F$  con propiedades convenientes, de tal forma que la inclusión posea soluciones relacionadas con las soluciones de la ecuación original.

Pueden darse más ejemplos. La teoría de inclusiones diferenciales procura extraer lo que es común a todos ellos y proporciona resultados relativos a la existencia y otras propiedades de las soluciones.

Por supuesto, las inclusiones diferenciales (1) puede ser vistas como generalizaciones directas de ecuaciones de la forma

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

en el sentido de que éstas últimas son inclusiones diferenciales particulares correspondientes a multiaplicaciones univaluadas. En consecuencia, todas las cuestiones pertinentes a las ecuaciones diferenciales, como por ejemplo la existencia, prolongabilidad y limitación de las soluciones, la dependencia continua respecto de datos iniciales y parámetros, la bifurcación de soluciones, la existencia y estabilidad de puntos de equilibrio y órbitas especiales, la estabilidad global, el comportamiento asintótico, etc. también lo son para las inclusiones diferenciales.

**Nota 1** No obstante, para una inclusión diferencial dada, existe en general toda una familia de trayectorias que comienzan en cada punto inicial  $x_0$ . En otras palabras, con  $F$  multívoca, es previsible que el problema (7) posea más de una solución. Esta propiedad de no unicidad conduce a cuestiones que son específicas para sistemas diferenciales multivaluados, relativos a la convexidad de la familia de soluciones, la existencia de soluciones extremales, la selección de soluciones con propiedades dadas, etc.

## 2.1 Existencia de solución local del problema de valor inicial para una inclusión diferencial

Para motivar el aparato matemático que debe ser desarrollado, reflexionaremos a continuación sobre la existencia de solución de (7).

Recordemos que, en el caso de una ecuación diferencial ordinaria, es necesaria una cierta regularidad (por ejemplo la continuidad del segundo miembro) para garantizar la existencia de soluciones. Por otra parte, una multiaplicación  $F = F(t, x)$  puede ser pensada como una función usual  $F : [0, T] \times \mathbb{R}^n \mapsto \mathcal{P}(\mathbb{R}^n)$ , donde  $\mathcal{P}(\mathbb{R}^n)$  es el conjunto de las partes de  $\mathbb{R}^n$ . Así, para garantizar la existencia de solución de (7), parece adecuado extender las nociones de regularidad usuales (continuidad, medibilidad, la condición de Lipschitz, etc.) a las funciones con valores en  $\mathcal{P}(\mathbb{R}^n)$  (véase por ejemplo [4],[21]).

Ahora bien, incluso disponiendo de estas herramientas, ¿cómo podríamos proseguir y probar la existencia de solución de (7)?

Una posible vía es la siguiente. Supongamos asegurada la existencia de una función  $f : [0, T] \times \mathbb{R}^n \mapsto \mathbb{R}^n$  suficientemente regular (continua por ejemplo), con la propiedad de que

$$f(t, x) \in F(t, x) \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n.$$

Entonces cualquier solución del problema de Cauchy

$$\begin{cases} \dot{x} = f(t, x), \\ x(0) = x_0, \end{cases}$$



cuya existencia está garantizada por el Teorema de Carathéodory, será necesariamente solución de (7).

El argumento anterior conduce a la siguiente cuestión de carácter general:

**Cuestión 2:** Dada una multiaplicación

$$G : E_1 \rightarrow E_2,$$

donde  $E_1$  y  $E_2$  son espacios dotados de una estructura adecuada (espacios topológicos, espacios de medida, etc.), ¿cuándo puede afirmarse que existe una función univaluada  $g : E_1 \mapsto E_2$  tal que para cada  $e \in E_1$  se tiene que  $g(e) \in G(e)$  y, además,  $g$  verifica propiedades de regularidad deseadas (continuidad, medibilidad, etc.)?

Si existe, a una función  $g$  como la que precede se le llama una *selección* (continua, medible, etc.) de la multiaplicación  $G$ . Si supiéramos responder satisfactoriamente a la Cuestión 2 en el contexto de (7), estaríamos muy cerca de conseguir resultados de existencia para este problema. Obsérvese por otra parte que, si no imponemos nada a  $g$ , la cuestión precedente es trivial y dar una respuesta positiva a la misma equivale a aceptar el *axioma de elección*.

**Nota 2** Para responder a la Cuestión 2, estos últimos años se han seguido principalmente dos estrategias diferentes.

a) La primera de ellas consiste en buscar una *regla* que a cada  $e \in E_1$  asocia un punto de  $G(e)$  con una determinada propiedad y, a continuación, comprobar que, bajo ciertas condiciones, la función  $g$  así obtenida es regular. Este camino conduce a varios resultados importantes: el Teorema de la Selección Minimal, el Teorema de la Selección de Chevishev, el Teorema de la Selección Baricéntrica, etc., ver por ejemplo [3].

b) La segunda estrategia consiste en buscar selecciones de carácter local (es decir, definidas en una vecindad de cada punto) y construir una selección global (es decir, definida en todo  $E_1$ ) utilizando una partición de la unidad. Un Teorema bien conocido al que se llega de esta forma es el debido a Michael [24].

**Nota 3** Obsérvese que, cuando la inclusión diferencial (1) tiene su origen en el sistema de control (3), la existencia de selecciones y la existencia de controles admisibles son, hablando en términos generales, problemas equivalentes. Esto es esencialmente lo que afirma el llamado Lema de Filippov; para más detalles, véase [10],[20].

En lo que sigue, nos referiremos principalmente a multifunciones  $F$  que toman valores en  $\mathbb{K}(\mathbb{R}^n)$ . Por definición,  $\mathbb{K}(\mathbb{R}^n)$  es la familia

$$\mathbb{K}(\mathbb{R}^n) = \{ A \subset \mathbb{R}^n ; A \text{ es compacto y } A \neq \emptyset \}.$$

Hablando en términos generales, las multifunciones que admiten selecciones continuas son las denominadas *multifunciones semicontinuas inferiormente*. Se

dice que la multifunción  $F : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{K}(\mathbb{R}^n)$  es semicontinua inferiormente si, para cada  $(t_0, x_0) \in [0, T] \times \mathbb{R}^n$  y cada  $\varepsilon > 0$ , existe  $\delta > 0$  tal que, si  $(t, x) \in [0, T] \times \mathbb{R}^n$  y  $|t - t_0| + \|x - x_0\| \leq \delta$ , entonces

$$F(t, x) \cap F(t_0, x_0)_\varepsilon \neq \phi. \quad (8)$$

Aquí,  $F(t_0, x_0)_\varepsilon$  designa el entorno cerrado de radio  $\varepsilon$  de  $F(t_0, x_0)$  (la unión de las bolas cerradas de radio  $\varepsilon$  centradas en puntos de  $F(t_0, x_0)$ ).

Por otra parte, se dice que  $F : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{K}(\mathbb{R}^n)$  es semicontinua superiormente si para cada  $(t_0, x_0)$  y cada  $\varepsilon > 0$  se tiene la propiedad precedente, cambiando (8) por

$$F(t, x) \subset F(t_0, x_0)_\varepsilon.$$

Para las *multifunciones semicontinuas superiormente*, existe un Teorema de Selección Aproximada, debido a Cellina [9], que en muchas situaciones es suficiente para nuestros propósitos.

**Nota 4** Digamos también que, en el marco de la medibilidad, el resultado más conocido que garantiza la existencia de selecciones se debe a Kuratowski y Ryll-Nerzowski, cf. [20].

Combinando la existencia de selecciones continuas y el teorema de Carathéodory, se deduce el resultado de existencia siguiente:

**Teorema 1** *Supongamos que, en (7), la multifunción  $F : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{K}(\mathbb{R}^n)$  es semicontinua superiormente y  $x_0 \in \mathbb{R}^n$ . Entonces el problema (7) posee al menos una solución en  $[0, T^*]$  para algún  $T^*$  maximal que verifica  $0 < T^* \leq T$ .*

Hay otras formas de encarar la existencia de solución del problema de valor inicial para una inclusión diferencial que intentan simular las técnicas utilizadas en el caso de una ecuación diferencial.

Por ejemplo, consideremos el siguiente argumento de punto fijo. Dada la inclusión diferencial (1), introducimos la *multiaplicación integral*

$$Y : C^0([0, T]; \mathbb{R}^n) \rightarrow C^0([0, T]; \mathbb{R}^n),$$

dada como sigue:

Para cada  $z \in C^0([0, T]; \mathbb{R}^n)$ ,  $Y(z)$  es la familia de las funciones  $x : [0, T] \mapsto \mathbb{R}^n$  absolutamente continuas que verifican  $\dot{x}(t) \in F(t, z(t))$  para casi todo  $t \in [0, T]$  y  $x(0) = x_0$ .

Obviamente, una función  $x \in C^0([0, T]; \mathbb{R}^n)$  es solución de (1) en  $[0, T]$  si y sólo si

$$x \in Y(x),$$

esto es,  $x$  es un punto fijo de la multiaplicación  $Y$ . Es posible entonces demostrar un resultado análogo al conocido Teorema de Schauder que, aplicado en este contexto, conduce a la existencia de puntos fijos de  $Y$ . Como consecuencia del

mismo, se puede volver a deducir el Teorema 1. También es posible extender la Teoría del Grado Topológico de Brower y de Leray-Schauder al caso de una multiaplicación; para más detalles, véase [8],[19].

Otra técnica muy utilizada (y de hecho también de tipo punto fijo) se basa en el siguiente resultado, que es cierto para cualquier ecuación diferencial que satisfaga adecuadas restricciones de regularidad:

La función  $x : J \mapsto \mathbb{R}^n$  es solución en  $J$  del sistema  $\dot{x} = f(t, x)$  y verifica la condición inicial  $x(0) = x_0$  si y sólo si es continua en  $J$  y se tiene

$$x(t) = x_0 + \int_0^t f(s, x(s)) ds \quad \forall t \in J$$

( $J$  es un intervalo que contiene a 0).

En el caso multívoco, debe aclararse el sentido que tiene la integral de una multifunción. A este respecto, existe una noción debida a Aumann [5] y Debreu [17], con la cual se puede probar una equivalencia análoga a la precedente para las soluciones de (7) (véase la definición de integral de una multifunción en la Sección 3).

Existen otras técnicas (construcción de aproximaciones adecuadas y extracción de subsucesiones convergentes, métodos constructivos basados en poligonales generalizadas de Euler, etc.). Por razones de espacio, no es posible describirlas aquí con detalle. Los lectores interesados pueden consultar, por ejemplo, las referencias [2],[7],[8],[16] y las que en ellas se citan.

**Nota 5** Otras muchas ecuaciones funcionales tienen su correspondiente generalización en el contexto de las multifunciones. Así, para inclusiones integrales, existe una teoría desarrollada que conduce a resultados de existencia satisfactorios, véase [1],[16]. Las inclusiones funcionales con retardo han sido tratadas en [14],[15],[33] y sus referencias. Finalmente, para inclusiones diferenciales de la forma

$$\dot{x} \in Ax + F(t, x),$$

donde  $A : D(A) \subset H \mapsto H$  es (por ejemplo) un operador no acotado maximal monótono en el espacio de Hilbert  $H$ , véase [27].

## 2.2 Estabilidad según Lyapunov para Inclusiones Diferenciales

Otra cuestión que nos interesa analizar en relación con (7) es la estabilidad de los puntos de equilibrio. Por comodidad, nos limitaremos al caso en que la inclusión diferencial es autónoma, esto es

$$\dot{x} \in F(x), \tag{9}$$

con  $F : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$ . El problema de valor inicial asociado toma la forma

$$\begin{cases} \dot{x} \in F(x), \\ x(0) = x_0, \end{cases} \tag{10}$$

donde  $x_0 \in \mathbb{R}^n$ .

Recordemos en primer lugar el resultado fundamental de Lyapunov para ecuaciones diferenciales ordinarias:

**Teorema 2** *Sea  $f : \mathbb{R}^n \mapsto \mathbb{R}^n$  una función continuamente diferenciable en una vecindad del punto  $x = 0$ , con diferencial-Fréchet  $f'(0)$  en 0 y tal que  $f(0) = 0$ . Supongamos que todos los autovalores de  $f'(0)$  tienen parte real negativa, es decir, que el origen es asintóticamente estable para el sistema lineal*

$$\dot{y} = f'(0)y.$$

Entonces el origen es asintóticamente estable para el sistema no lineal

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

Para conseguir una extensión de este resultado en el contexto de (10), previamente necesitamos poder hablar de multiaplicaciones diferenciables.

Obsérvese que la definición de derivada de una multifunción no es trivial. En efecto, no es difícil probar que, si  $E_1$  y  $E_2$  son espacios vectoriales,  $G : E_1 \rightarrow E_2$  es una multiaplicación y se tiene

$$G(\alpha_1 e_1 + \alpha_2 e_2) = \alpha_1 G(e_1) + \alpha_2 G(e_2) \quad \forall \alpha_1, \alpha_2 \in \mathbb{R}, \quad \forall e_1, e_2 \in E,$$

entonces  $G$  es univaluada. Dicho de otro modo, la noción de función diferenciable no puede extenderse al caso de una multiaplicación palabra por palabra. No obstante, véase [8].

El conjunto  $\mathbb{K}(\mathbb{R}^n)$  se convierte en espacio métrico una vez dotado de la distancia de Hausdorff  $h(\cdot, \cdot)$ , donde

$$h(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\}.$$

Aquí,  $d(\cdot, \cdot)$  es la distancia euclídea habitual en  $\mathbb{R}^n$ .

El espacio métrico  $(\mathbb{K}(\mathbb{R}^n), h)$  es completo y separable. De hecho, se trata de un espacio *cuasi-normado*, con *cuasi-norma* dada por la aplicación

$$A \mapsto h(A, \{0\}).$$

La introducción de  $h(\cdot, \cdot)$  nos permitirá a continuación hablar de multifunciones *cuasi-lineales acotadas* y de multifunciones *Fréchet-diferenciables*.

**Definición 1** *Se dice que una multifunción  $Q : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  es cuasi-lineal si*

$$Q(x_1 + x_2) \subset Q(x_1) + Q(x_2) \quad \forall x_1, x_2 \in \mathbb{R}^n$$

y

$$Q(\lambda x) = \lambda Q(x) \quad \forall \lambda \in \mathbb{R}, \quad \forall x \in \mathbb{R}^n.$$

Se dice que  $Q$  es acotada si existe  $K > 0$  tal que

$$h(Q(x), \{0\}) \leq K \|x\| \quad \forall x \in \mathbb{R}^n.$$

**Definición 2** Se dice que la multifunción  $F : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  es Fréchet-diferenciable en  $x_0 \in \mathbb{R}^n$  si existe una multifunción cuasi-lineal acotada  $F'(x_0) : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  tal que

$$h(F(x), F(x_0) + F'(x_0)(x - x_0)) = o(\|x - x_0\|).$$

La multifunción  $F'(x_0)$  se denomina derivada de Fréchet de  $F$  en  $x_0$ .

Obviamente, si  $F : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  es cuasi-lineal y acotada, entonces es Fréchet-diferenciable en 0 y además  $F'(0) = F$ .

**Definición 3** Consideremos la inclusión diferencial (9). Supongamos que  $F(0) = \{0\}$ . Se dice entonces que  $x = 0$  es un punto de equilibrio de (9). Diremos que este punto de equilibrio es estable (en el sentido de Lyapunov) si se cumplen las siguientes condiciones:

1. Existe  $\delta_0 > 0$  tal que, si  $\|x_0\| \leq \delta_0$ , entonces existe una solución  $x = x(t)$  de (10) definida para todo  $t \geq 0$ .
2. Para todo  $\epsilon > 0$ , existe  $\delta_1 \in (0, \delta_0]$  tal que, si  $\|x_0\| < \delta_1$ , entonces  $\|x(t)\| < \epsilon$  para todo  $t \geq 0$ .

Finalmente, diremos que el punto de equilibrio Lyapunov-estable  $x = 0$  es asintóticamente estable si existe  $\delta_2 \in (0, \delta_0]$  tal que, si  $\|x_0\| \leq \delta_2$ , entonces  $\lim_{t \rightarrow +\infty} \|x(t)\| = 0$ .

Tenemos entonces el siguiente resultado relativo a la estabilidad de los puntos de equilibrio de (9):

**Teorema 3** Supongamos que la multifunción  $F : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  es Fréchet-diferenciable en 0 y que existe  $\delta_0 > 0$  tal que cualquier solución  $x = x(t)$  existe en el intervalo  $[0, +\infty)$  si  $\|x(0)\| \leq \delta_0$ . En estas condiciones, si el origen es un punto de equilibrio asintóticamente estable de la inclusión diferencial cuasi-lineal

$$\dot{y} \in F'(0)(y),$$

entonces también es un punto de equilibrio asintóticamente estable de (9). En tal caso, existen además  $\sigma > 0$ ,  $k > 0$  y  $\delta > 0$  tales que, si  $\|x_0\| \leq \delta$ , cualquier solución de (10) verifica

$$\|x(t)\| \leq k \|x(0)\| e^{-\sigma t} \quad \forall t \geq 0.$$

### 3 Otros resultados relacionados con las multifunciones

En esta sección, mostraremos brevemente otros aspectos del Análisis Matemático relacionados con las multifunciones que han tenido un desarrollo reciente. Comenzaremos hablando de multifunciones Lipschitz-continuas y de la generalización del Teorema de Banach del Punto Fijo.

**Definición 4** Se dice que la multiaplicación  $F : A \subset \mathbb{R}^m \mapsto \mathbb{K}(\mathbb{R}^n)$  es Lipschitz-continua en  $A$  si existe  $\gamma > 0$  tal que

$$h(F(x), F(y)) \leq \gamma|x - y| \quad \forall x, y \in A.$$

En particular, si se tiene esta propiedad con  $0 < \gamma < 1$ , se dice que  $F$  es una contracción multívoca o bien que  $F$  es contractiva.

El análogo del Teorema de Punto Fijo de Banach para una multifunción fue probado en [25] por Nadler en los años 60. En el caso de una multifunción definida en  $\mathbb{R}^n$ , dice lo siguiente:

**Teorema 4** Sea  $F : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  una contracción multívoca. Entonces  $F$  posee un punto fijo, es decir, existe  $x \in \mathbb{R}^n$  tal que

$$x \in F(x).$$

Para algunas mejoras y variantes del Teorema anterior, puede consultarse la referencia [19] (donde también se encontrarán los análogos multívocos del Teorema de Schauder).

Una herramienta esencial para el estudio de multifunciones con valores convexos compactos es el *funcional soporte* del Análisis Convexo, cuya definición es la siguiente:

**Definición 5** Sea  $A \subset \mathbb{R}^m$  un conjunto no vacío. El funcional soporte de  $A$  es la función  $\sigma_A : \mathbb{R}^m \mapsto \mathbb{R} \cup \{+\infty\}$ , dada por

$$\sigma_A(y) = \sup_{x \in A} \langle x, y \rangle \quad \forall y \in \mathbb{R}^m.$$

Algunas propiedades importantes de este funcional están recogidas en la Proposición siguiente:

**Proposición 5** Sean  $A \subset \mathbb{R}^m$  un conjunto no vacío,  $\sigma_A$  el correspondiente funcional soporte y  $D(\sigma_A) = \{y \in \mathbb{R}^m ; \sigma_A(y) < +\infty\}$  su dominio efectivo. Entonces se tiene:

1.  $D(\sigma_A)$  es un cono convexo con vértice en 0.
2.  $\sigma_A(\lambda y) = \lambda \sigma_A(y)$  para cada  $\lambda \geq 0$  y cada  $y \in \mathbb{R}^m$ , i.e.  $\sigma_A$  es homogéneamente positiva.
3.  $\sigma_A$  es una función convexa, semicontinua inferiormente y propia. En particular,  $\sigma_A \not\equiv +\infty$ .
4.  $\sigma_A$  coincide con el funcional soporte de  $\overline{\text{co}} A$ , la envolvente convexa cerrada de  $A$ .

La denominación “funcional soporte” está motivada por lo siguiente. Si  $A \subset \mathbb{R}^m$  es convexo, compacto y no vacío, para cada  $u \in \mathbb{R}^m$  con  $u \neq 0$ , existe  $x_u \in A$  tal que

$$\sigma_A(u) = \langle x_u, u \rangle$$

y el hiperplano  $\mathcal{H} = \{x \in \mathbb{R}^m ; \langle x, u \rangle = \sigma_A(u)\}$  soporta a  $A$  en  $x_u$ . Esto último quiere decir que  $A$  está contenido en  $\mathcal{H}^- = \{x \in \mathbb{R}^m ; \langle x, u \rangle \leq \sigma_A(u)\}$  y toca al hiperplano en  $x_u$ .

Otras propiedades interesantes del funcional soporte están recogidas en la siguiente

**Proposición 6** Sean  $A, B \subset \mathbb{R}^m$  conjuntos no vacíos. Entonces se tiene:

1.  $\sigma_A(y) \geq 0$  para todo  $y \in \mathbb{R}^m$  si y sólo si  $\overline{\text{co}} A \ni 0$ .
2.  $\sigma_{\lambda A + \mu B}(y) = \lambda \sigma_A(y) + \mu \sigma_B(y)$  para  $\lambda, \mu \geq 0$ .
3. Si  $A \subset B$ , entonces  $\sigma_A(y) \leq \sigma_B(y)$  para cada  $y \in \mathbb{R}^m$ . Recíprocamente, si para todo  $y \in \mathbb{R}^m$  se tiene que  $\sigma_A(y) \leq \sigma_B(y)$ , entonces  $\overline{\text{co}} A \subset \overline{\text{co}} B$ . En particular si  $A$  y  $B$  son convexos y cerrados, tenemos:

$$A \subset B \iff \sigma_A(y) \leq \sigma_B(y) \quad \forall y \in \mathbb{R}^m.$$

4. Si  $\{A_i ; i \in I\}$  es una familia de conjuntos no vacíos de  $\mathbb{R}^m$  y  $A = \bigcup_{i \in I} A_i$ , entonces

$$\sigma_A(y) = \sup_{i \in I} \sigma_{A_i}(y) \quad \forall y \in \mathbb{R}^m.$$

5.  $A$  es simétrico (es decir,  $A = -A$ ) si y sólo si su funcional soporte  $\sigma_A$  es par.
6. Si  $A$  es de la forma  $A = A_1 \times \dots \times A_q$ , entonces

$$\sigma_A(y) = \sum_{i=1}^q \sigma_{A_i}(y) \quad \forall y \in \mathbb{R}^m.$$

A continuación, vamos a indicar explícitamente qué es la *integral de una multifunción* en el sentido de Aumann.

Sea  $\{\Omega, \Sigma, \mu\}$  un espacio de medida completo. Dada una multifunción  $G : \Omega \mapsto \mathcal{P}(\mathbb{R}^n)$ , denotaremos  $S^1(G)$  el conjunto

$$S^1(G) = \{f \in L^1(\Omega; \mathbb{R}^n) ; f(w) \in G(w) \text{ } \mu\text{-c.t.p.}\}.$$

**Definición 6** La integral de  $G$  en  $\Omega$ , denotada

$$\int_{\Omega} G(w) d\mu,$$

es, por definición, el siguiente conjunto de  $\mathbb{R}^n$ :

$$\int_{\Omega} G(w) d\mu = \left\{ \int_{\Omega} f(w) d\mu ; f \in S^1(G) \right\}.$$

Se dice que la multifunción  $G$  es medible si, para cada conjunto medible  $B$  de  $\mathbb{R}^n$ , se tiene  $G^{-1}(B) \in \Sigma$ . Aquí,  $G^{-1}(B) = \{w \in \Omega ; G(w) \cap B \neq \emptyset\}$ .

Se dice que  $G$  está acotada integralmente si es medible y además la función  $\|G\|$ , dada por

$$\|G\|(w) = \sup\{\|f(w)\| ; f \in S^1(G)\} \quad \mu - c.t.p.,$$

es  $\mu$ -integrable.

Las principales propiedades de la integral multívoca aparecen en el resultado siguiente:

**Teorema 7** Con la notación precedente,  $\int_{\Omega} G(w) d\mu$  es un convexo. Por otra parte, si  $G$  está acotada integralmente, entonces  $S^1(G)$  es no vacío y, en consecuencia, también lo es  $\int_{\Omega} G(w) d\mu$ . Si los valores de  $G$  son cerrados, entonces  $\int_{\Omega} G(w) d\mu$  es un compacto y se da la igualdad

$$\int_{\Omega} G(w) d\mu = \int_{\Omega} (\overline{\text{co}} G)(w) d\mu.$$

Finalmente, si  $G$  está acotada integralmente y toma valores compactos y ponemos  $A = \int_{\Omega} G(w) d\mu$ , entonces

$$\sigma_A(y) = \int_{\Omega} \sigma_{G(w)}(y) d\mu \quad \forall y \in \mathbb{R}^n.$$

Con ayuda de este resultado, es posible calcular explícitamente algunas integrales. Para ello es suficiente construir la función soporte  $\sigma_{G(w)}$ , integrar la función  $\sigma_{G(w)}(y)$  respecto de  $w$  para cada  $y \in \mathbb{R}^n$  y, después, reconstruir el convexo compacto no vacío  $\int_{\Omega} G(w) d\mu$  a partir de la función soporte resultante. Veamos algunos ejemplos que permiten aclarar este procedimiento.

En primer lugar, sea  $A$  un convexo compacto no vacío de  $\mathbb{R}^n$ . Entonces, cualesquiera que sean  $t_1, t_2 \in \mathbb{R}$ , con  $t_1 \leq t_2$ , se tiene que

$$\int_{(t_1, t_2)} A dt = \int_{t_1}^{t_2} A dt = (t_2 - t_1)A.$$

En efecto, por el Teorema 7, tenemos

$$\begin{aligned} \sigma_{\int_{t_1}^{t_2} A dt}(y) &= \int_{t_1}^{t_2} \sigma_A(y) dt \\ &= \sigma_A(y) \int_{t_1}^{t_2} dt \\ &= (t_2 - t_1)\sigma_A(y). \end{aligned}$$



Luego

$$\begin{aligned} \int_{t_1}^{t_2} A dt &= \{ x \in \mathbb{R}^n ; (x, y) \leq (t_2 - t_1)\sigma_A(y) \quad \forall y \in \mathbb{R}^n \} \\ &= (t_2 - t_1)A. \end{aligned}$$

Si  $A$  no fuese convexo, entonces tendríamos

$$\int_{t_1}^{t_2} A dt = \int_{t_1}^{t_2} (\overline{\text{co}} A) dt = (t_2 - t_1) \overline{\text{co}} A.$$

Veamos un segundo ejemplo. Así, sea  $G : [-\pi, \pi] \rightarrow \mathbb{R}^2$  la multifunción definida por

$$G(t) = U(t)\{-v, v\} \quad \forall t \in [-\pi, \pi],$$

donde

$$U(t) = \begin{pmatrix} \text{sen } t & \text{tg } t \\ \text{cos } t & \text{cotg } t \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Entonces

$$\int_{-\pi}^{\pi} G(t) dt = \overline{B}(0; 4),$$

i.e. la bola cerrada de  $\mathbb{R}^2$  de centro el origen y radio 4. En efecto, tenemos que

$$\begin{aligned} \sigma_{\int_{-\pi}^{\pi} G(t) dt}(y) &= \int_{-\pi}^{\pi} \sigma_{G(t)}(y) dt \\ &= \int_{-\pi}^{\pi} |y_1 \text{sen } t + y_2 \text{cos } t| dt \\ &= 4\|y\|. \end{aligned}$$

Deducimos que

$$\begin{aligned} \int_{-\pi}^{\pi} G(t) dt &= \overline{\text{co}} \{ x \in \mathbb{R}^2 ; (x, y) \leq \sigma_{\int_{-\pi}^{\pi} G(t) dt}(y) \quad \forall y \in \mathbb{R}^2 \} \\ &= \overline{\text{co}} \{ x \in \mathbb{R}^2 ; (x, y) \leq 4\|y\| \quad \forall y \in \mathbb{R}^2 \} \\ &= \overline{B}(0; 4). \end{aligned}$$

Obsérvese que, en este ejemplo, no hubiera sido posible calcular la integral de  $G$  directamente usando la definición de integral multívoca.

## 4 Conceptos de la Teoría “Fuzzy”

Los conjuntos difusos (también denominados “fuzzy sets”) fueron introducidos por Zadeh [35] en 1965. Desde entonces, un gran número de publicaciones ha tenido como objetivo extender los conceptos clásicos al nuevo campo y resolver los problemas planteados.

En este trabajo, la motivación principal del concepto *conjunto difuso* es nuestro interés por matematizar aspectos *subjetivos* en fenómenos tomados de la realidad. Veremos que esto se consigue generalizando la idea de pertenencia clásica (de un elemento a un conjunto) e introduciendo la idea de *grado de pertenencia*.

Recordemos el argumento básico desarrollado por Zadeh. Sea  $X$  un conjunto. Los subconjuntos usuales  $A \subset X$  pueden ser identificados a partir de sus correspondientes funciones características  $\chi_A$ , con

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases} \quad (11)$$

Un punto  $x \in X$  pertenece a  $A$  (resp. al complementario de  $A$ ) si y sólo si  $\chi_A(x) = 1$  (resp.  $\chi_A(x) = 0$ ). Zadeh llamó *conjunto difuso* o *conjunto "fuzzy"* en  $X$  a toda función  $u : X \mapsto [0, 1]$ . De esta forma, dado un conjunto difuso  $u$ , el *grado de pertenencia* de un punto  $x \in X$  a este conjunto es  $u(x)$ , un número entre 0 y 1.

Obviamente, este concepto constituye una extensión de la idea habitual. Para hacerlo operativo, es preciso extender las definiciones y reglas usuales entre conjuntos, tales como determinar el complementario, la unión, la intersección, etc., véase [26].

En lo que sigue, consideraremos conjuntos difusos  $u : X \mapsto [0, 1]$ , donde  $X$  es un espacio topológico (usualmente tendremos  $X = \mathbb{R}^n$ ). Veamos en primer lugar que todo conjunto difuso puede identificarse con una familia (no numerable) de conjuntos "estándar" que cumplen ciertas propiedades.

Así, sea  $u : X \mapsto [0, 1]$  un conjunto difuso y denotemos  $L_a u$  el conjunto de nivel  $a$  de  $u$  para cada  $a \in (0, 1]$ , esto es,

$$L_a u = \{x \in X ; u(x) \geq a\}.$$

Por otra parte, pongamos

$$L_0 u = \text{sop}(u) = \overline{\{x \in X ; u(x) \neq 0\}}$$

(el soporte de  $u$ ). Entonces

$$L_0 u \supset L_a u \supset L_b u \quad \text{para } b \geq a \geq 0.$$

De este modo, al conjunto "fuzzy"  $u : X \mapsto [0, 1]$  le podemos asociar la familia  $\{L_a u ; a \in [0, 1]\}$ . Recíprocamente, tenemos el siguiente resultado debido a Negoita y Ralescu, cf. [26]:

**Teorema 8** *Sea  $X$  un espacio topológico y sea  $\{M_a ; a \in [0, 1]\}$  una familia de subconjuntos de  $X$  con las propiedades siguientes:*

1.  $M_0 \subset X$ ,
2. Si  $a \leq b$ , entonces  $M_b \subset M_a$ ,

3. Si  $0 \leq a_1 \leq a_2 \leq \dots \leq 1$  y  $\lim_{n \rightarrow \infty} a_n = a$ , entonces

$$M_a = \bigcap_{n=1}^{\infty} M_{a_n}.$$

Entonces existe un único conjunto difuso  $u : X \mapsto [0, 1]$  tal que  $M_a = L_a u$  para cada  $a \in [0, 1]$ .

Obsérvese que, en particular, si  $u : X \mapsto [0, 1]$  es un conjunto “estándar”, i.e. si  $u = \chi_A$  para algún  $A \subset X$ , entonces la familia  $\{L_a u ; a \in [0, 1]\}$  es la siguiente:

$$L_a u = \begin{cases} \bar{A} & \text{si } a = 0, \\ A & \text{si } 0 < a \leq 1. \end{cases}$$

Recordemos que  $\mathbb{K}(\mathbb{R}^n)$  denota la familia de los compactos no vacíos de  $\mathbb{R}^n$ . La extensión adecuada al contexto “fuzzy” de  $\mathbb{K}(\mathbb{R}^n)$  es la familia  $\mathbb{E}(\mathbb{R}^n)$ .

**Definición 7**  $\mathbb{E}(\mathbb{R}^n)$  es la familia de los conjuntos “fuzzy”  $u : \mathbb{R}^n \mapsto [0, 1]$  que cumplen las propiedades siguientes:

1.  $u$  es semicontinua superiormente,
2.  $L_0 u$  es compacto,
3. El conjunto  $L_1 u = \{x \in \mathbb{R}^n ; u(x) = 1\}$  es no vacío.

Dada  $u : \mathbb{R}^n \mapsto [0, 1]$ , no es difícil comprobar que

$$u \in \mathbb{E}(\mathbb{R}^n) \iff L_a u \in \mathbb{K}(\mathbb{R}^n) \quad \forall a \in [0, 1].$$

Esta caracterización es esencial para abordar las cuestiones mencionadas en la Sección 1.

Podemos definir una estructura métrica sobre  $\mathbb{E}(\mathbb{R}^n)$  poniendo

$$D(u_1, u_2) = \sup_{a > 0} h(L_a u_1, L_a u_2)$$

(recuérdese que  $h(\cdot, \cdot)$  es la distancia de Hausdorff habitual en  $\mathbb{K}(\mathbb{R}^n)$ ). Claramente,

$$(\mathbb{R}^n, d) \hookrightarrow (\mathbb{K}(\mathbb{R}^n), h) \hookrightarrow (\mathbb{E}(\mathbb{R}^n), D),$$

con inyecciones continuas (recuérdese que  $d(\cdot, \cdot)$  es la distancia euclídea en  $\mathbb{R}^n$ ).

En lo que resta de esta sección, mostraremos la extensión al contexto “fuzzy” de la mayor parte de conceptos y resultados presentados en la sección precedente.

Sea  $\mathbb{K}_c(\mathbb{R}^n)$  el subconjunto de  $\mathbb{K}(\mathbb{R}^n)$  formado por los compactos convexos no vacíos de  $\mathbb{R}^n$ . Indicaremos en primer lugar cuál es la generalización al contexto “fuzzy” de  $\mathbb{K}_c(\mathbb{R}^n)$ . Para ello, nos hace falta una definición:

**Definición 8** Se dice que el conjunto difuso  $u : \mathbb{R}^n \mapsto [0, 1]$  es *fuzzy-convexo* si, para cada  $x, y \in \mathbb{R}^n$  y cada  $\lambda \in [0, 1]$ , se verifica la desigualdad

$$u(\lambda x + (1 - \lambda)y) \geq \min[u(x), u(y)].$$

Esta definición es bastante natural si pensamos en  $u(x)$  como el *grado de pertenencia* de  $x$  al conjunto “fuzzy”  $u$ . De hecho,

$$u \text{ es fuzzy-convexo} \iff L_a u \text{ es convexo para cada } a \in [0, 1].$$

Pondremos ahora

$$\mathbb{E}_c(\mathbb{R}^n) = \{ u \in \mathbb{E}(\mathbb{R}^n) ; u \text{ es fuzzy-convexo} \}.$$

Claramente,  $\mathbb{K}_c(\mathbb{R}^n) \hookrightarrow \mathbb{E}_c(\mathbb{R}^n)$  vía la inclusión canónica  $A \mapsto \chi_A$ . Por otra parte,

$$u \in \mathbb{E}_c(\mathbb{R}^n) \iff L_a u \in \mathbb{K}_c(\mathbb{R}^n) \quad \forall a \in [0, 1].$$

El conjunto  $\mathbb{E}_c(\mathbb{R}^n)$  puede dotarse de una estructura de espacio vectorial, poniendo

- $(u_1 \oplus u_2)(x) = \sup_{y+z=x} \min[u_1(y), u_2(z)]$  para  $u_1, u_2 \in \mathbb{E}_c(\mathbb{R}^n)$ ,
- $(\lambda \odot u)(x) = u(x/\lambda)$  si  $\lambda \in \mathbb{R} \setminus \{0\}$  y  $(\lambda \odot u)(x) = \chi_{\{0\}}(x)$  si  $\lambda = 0$ , para cada  $u \in \mathbb{E}_c(\mathbb{R}^n)$ .

Aunque a primera vista estas operaciones parecen extrañas, son completamente naturales si pensamos en términos de conjuntos de nivel. En efecto, la definición precedente de suma permite escribir que

$$L_a(u_1 \oplus u_2) = L_a u_1 + L_a u_2 \quad \forall u_1, u_2 \in \mathbb{E}_c(\mathbb{R}^n)$$

para cada  $a \in [0, 1]$ . Por otra parte, con la definición del producto por un escalar, tenemos

$$L_a(\lambda \odot u) = \lambda L_a u \quad \forall \lambda \in \mathbb{R}, \quad \forall u_1, u_2 \in \mathbb{E}_c(\mathbb{R}^n).$$

Obsérvese por otra parte que una manera equivalente de definir la suma en  $\mathbb{E}_c(\mathbb{R}^n)$  es a partir de la *inf-convolución* del Análisis Convexo.

Se tiene el resultado siguiente (cf. [28] y [30] para la demostración):

**Teorema 9**  $(\mathbb{E}(\mathbb{R}^n), D)$  es un espacio métrico completo y no separable. Por otra parte,  $\mathbb{E}_c(\mathbb{R}^n)$  es cerrado en  $(\mathbb{E}(\mathbb{R}^n), D)$ .

Una propiedad importante de  $(\mathbb{K}_c(\mathbb{R}^n), h)$  es el Teorema de Inmersión de Minkowski. Una forma de enunciar este resultado es la siguiente:

**Teorema 10** La aplicación  $j_0 : \mathbb{K}_c(\mathbb{R}^n) \mapsto C^0(S^{n-1})$  definida por

$$j_0(A) = \sigma_A \quad \forall A \in \mathbb{K}_c(\mathbb{R}^n),$$

es aditiva, positivamente homogénea e isométrica, i.e.

$$\|j_0(A) - j_0(B)\|_{C^0(S^{n-1})} = h(A, B) \quad \forall A, B \in \mathbb{K}_c(\mathbb{R}^n).$$

Para poder generalizar este Teorema al contexto “fuzzy”, conviene que nos restrinjamos a un subconjunto adecuado de  $\mathbb{E}_c(\mathbb{R}^n)$ . Así, vamos a considerar el conjunto  $\mathbb{E}_{cc}(\mathbb{R}^n)$ , formado por las  $u \in \mathbb{E}_c(\mathbb{R}^n)$  tales que la función  $a \mapsto L_a u$  es continua de  $[0, 1]$  en el espacio métrico  $(\mathbb{K}_c(\mathbb{R}^n), h)$ .

No es difícil probar que  $(\mathbb{E}_{cc}(\mathbb{R}^n), D)$  es un subespacio métrico completo de  $(\mathbb{E}_c(\mathbb{R}^n), D)$ . Y la generalización del Teorema de Inmersión de Minkowski es como sigue:

**Teorema 11** *Existe una inyección  $j : \mathbb{E}_{cc}(\mathbb{R}^n) \mapsto C^0([0, 1] \times S^{n-1})$  que es aditiva, positivamente homogénea e isométrica, i.e. tal que*

$$\|j(u_1) - j(u_2)\|_{C^0([0,1] \times S^{n-1})} = D(u_1, u_2) \quad \forall u_1, u_2 \in \mathbb{E}_{cc}(\mathbb{R}^n).$$

Este resultado está probado en [29] (véase también [32]). Para la demostración, se usa el funcional soporte “fuzzy”  $S_u : [0, 1] \times S^{n-1} \mapsto \mathbb{R}$ , definido por

$$S_u(a, x) = \sigma_{L_a u}(x) \quad \forall (a, x) \in [0, 1] \times S^{n-1}.$$

Como consecuencia del resultado anterior, deducimos que  $(\mathbb{E}_{cc}(\mathbb{R}^n), D)$  es un subespacio métrico completo y separable de  $(\mathbb{E}_c(\mathbb{R}^n), D)$  y, además, que se trata del mayor subespacio que puede ser inmerso a través de  $j$  en  $C^0([0, 1] \times S^{n-1})$ .

**Definición 9** *Sea  $A \subset \mathbb{R}^m$  un conjunto no vacío. Se llama multifunción “fuzzy” a toda aplicación  $H : A \subset \mathbb{R}^m \mapsto \mathbb{E}(\mathbb{R}^n)$ .*

**Definición 10** *Diremos que  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  es una multifunción “fuzzy” Lipschitz-continua si existe  $\gamma > 0$  tal que*

$$D(H(x), H(y)) \leq \gamma \|x - y\| \quad \forall x, y \in \mathbb{R}^n.$$

*Diremos que  $H$  es contractiva si se tiene la propiedad anterior para algún  $\gamma \in (0, 1)$ . Finalmente, diremos que  $x$  es un punto fijo de  $H$  si*

$$H(x) \ni \chi_{\{x\}}.$$

El análogo del Teorema de Nadler para multifunciones “fuzzy” es el que sigue (véase [31] para la demostración del mismo):

**Teorema 12** *Sea  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  una multifunción “fuzzy” contractiva. Entonces  $H$  posee al menos un punto fijo.*

El problema de la estabilidad del punto fijo es resuelto por el Teorema 15, presentado más abajo. Este resultado es de gran importancia en el contexto de las inclusiones diferenciales “fuzzy” y también cuando hablamos de control óptimo “fuzzy”. Antes de enunciarlo, debemos definir el concepto de *conjunto difuso de puntos fijos* de una multifunción “fuzzy”. Para ello, usaremos las siguientes Proposiciones:

**Proposición 13** Sea  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  una multifunción “fuzzy” contractiva. Entonces, para cada  $a \in [0, 1]$ , la multifunción  $H_a : \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$ , dada por

$$H_a(x) = L_a[H(x)] \quad \forall x \in \mathbb{R}^n,$$

es contractiva en el sentido de la Definición 4.

**Proposición 14** En las condiciones de la Proposición precedente, sea  $S_a$  el conjunto de los puntos fijos de  $H_a$  para cada  $a \in [0, 1]$ . Entonces existe un único  $u \in \mathbb{E}(\mathbb{R}^n)$  con la propiedad siguiente:

$$L_a u = S_a \quad \forall a \in [0, 1].$$

Las demostraciones pueden encontrarse en [31]. En particular, para la prueba de la Proposición 14, se usa el Teorema de Representación de Nagoita y Ralescu (Teorema 8).

Podemos ahora dar la Definición anunciada:

**Definición 11** Sea  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  una multifunción “fuzzy” contractiva. El conjunto difuso  $u \in \mathbb{E}(\mathbb{R}^n)$  que proporciona la Proposición 14 es, por definición, el conjunto difuso de los puntos fijos de  $H$ .

Obsérvese que el conjunto difuso  $u$  que precede verifica lo siguiente:

$$u(x) = (H(x))(x) \quad \forall x \in \mathbb{R}^n.$$

Con estos ingredientes podemos ya enunciar el resultado de estabilidad mencionado más arriba:

**Teorema 15** Sean  $H_0$  y las  $H_i$  multifunciones “fuzzy” contractivas con la misma constante  $\gamma \in (0, 1)$ , tales que  $H_i \rightarrow H_0$  uniformemente en  $\mathbb{R}^n$ , i.e.

$$\sup_{x \in \mathbb{R}^n} D(H_i(x), H_0(x)) \rightarrow 0$$

cuando  $i \rightarrow +\infty$ . Sean  $u_0$  y los  $u_i$  los correspondientes conjuntos difusos de puntos fijos. Entonces  $D(u_i, u_0) \rightarrow 0$  cuando  $i \rightarrow +\infty$ .

Para finalizar esta sección, presentaremos una extensión de la integral de Aumann al caso multívoco “fuzzy” debida a Ralescu, cf. [28]. Previamente, definiremos la noción de medibilidad:

**Definición 12** Sea  $\{\Omega, \Sigma, \mu\}$  un espacio de medida completo y sea  $H : \Omega \mapsto \mathbb{E}(\mathbb{R}^n)$  una multifunción “fuzzy”. Diremos que  $H$  es medible si la multifunción  $H_a : \Omega \mapsto \mathbb{K}(\mathbb{R}^n)$ , definida por

$$H_a(w) = L_a[H(w)] \quad \forall w \in \Omega,$$

es medible para cada  $a \in [0, 1]$ . Diremos que  $H$  está acotada integralmente si cada  $H_a$  lo está.

**Teorema 16** Sean  $\{\Omega, \Sigma, \mu\}$  un espacio de medida completo y  $H : \Omega \mapsto \mathbb{E}(\mathbb{R}^n)$  una multifunción “fuzzy” medible y acotada integralmente. Entonces existe un único conjunto difuso  $u \in \mathbb{E}(\mathbb{R}^n)$  con la propiedad siguiente:

$$L_a u = \int_{\Omega} H_a d\mu \quad \forall a \in [0, 1].$$

Este  $u$  es, por definición, la integral de la multifunción “fuzzy”  $H$  y se denota  $\int_{\Omega} H d\mu$ .

Con esta definición de integral de una multifunción “fuzzy”, bajo condiciones razonables (por ejemplo, cuando  $\mu$  no es una medida atómica), se tiene que  $\int_{\Omega} H d\mu \in \mathbb{E}_c(\mathbb{R}^n)$ . En efecto, como consecuencia inmediata de las propiedades de la integral de Aumann, los conjuntos de nivel  $L_a[\int_{\Omega} H d\mu]$  son todos compactos y convexos.

Por otra parte, no es difícil comprobar que la integral “fuzzy” multívoca es lineal en el sentido siguiente:

$$\begin{aligned} L_a[\int_{\Omega} (H_1 \oplus H_2) d\mu] &= L_a \int_{\Omega} H_1 d\mu + L_a \int_{\Omega} H_2 d\mu, \\ L_a[\int_{\Omega} \lambda \odot H d\mu] &= \lambda L_a \int_{\Omega} H d\mu. \end{aligned}$$

Para más detalles sobre las propiedades de esta integral, se pueden consultar las referencias [12],[23],[28],[30].

## 5 Dinámica de poblaciones e inclusiones diferenciales “fuzzy”

Como hemos indicado ya en la Sección 1, las ecuaciones diferenciales “clásicas” no son capaces de describir todos los fenómenos observados en la realidad. Así, cuando nos interesamos por modelos de poblaciones, debemos afrontar con frecuencia hechos que no están completamente de acuerdo con la teoría clásica. En efecto,

- Los individuos pueden exhibir algunas estrategias o preferencias, tales como la elección de alimentos, la huida de depredadores, etc.
- Pueden existir restricciones adicionales que el sistema debe satisfacer (limitaciones de espacio, nutrientes o toxinas inducidas, etc.). Muchas veces éstas no pueden ser tenidas en cuenta por un sistema diferencial “estándar”.
- Puede haber indefiniciones en la dinámica debidas a ruidos producidos en el medio o en el crecimiento demográfico.

Consideremos por simplicidad un modelo determinista simple que, en ausencia de indefinición, está descrito por una ecuación diferencial ordinaria escalar

$$\dot{x} = f_0(t, x), \quad (12)$$

donde  $f_0 : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$  es una función regular dada. Podemos modelar la indefinición o ruido en la evolución de la población, introduciendo un “parámetro”  $v$ , por ejemplo una función o distribución que toma valores en  $U \subset \mathbb{R}$ . El sistema resultante es de la forma

$$\dot{x} = f(t, x, v(t)), \quad (13)$$

donde  $f : [0, T] \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$  es tal que  $f(t, x, 0) \equiv f_0(t, x)$ .

Al menos existen dos formas distintas de llegar a sistemas de este tipo:

a) Si aceptamos que la naturaleza de la indefinición es *aleatoria*, entonces lo más adecuado es utilizar una ecuación diferencial estocástica. En este caso, debido a dificultades técnicas, se suele suponer que la perturbación afecta linealmente a la dinámica. Más precisamente, se suele escribir

$$\dot{x} = f_0(t, x) + f_1(t, x)v,$$

donde  $v$  es un *ruido blanco*, esto es, la “derivada estocástica” de un movimiento Browniano. Para un primer análisis de estos modelos, véase por ejemplo [18].

b) Si, por el contrario, no parece adecuado suponer que el ruido posea estructura probabilística (debido a que las elecciones o preferencias de los individuos son subjetivas), parece más conveniente recurrir a la teoría “fuzzy”. Para más detalles, cf. [22].

Como hemos indicado en las Secciones anteriores, una manera adecuada de reescribir (13) es poner

$$\dot{x} \in F(t, x), \quad (14)$$

donde  $F$  es una multifunción determinada por  $f$  y el conjunto  $U$  donde debe tomar valores  $v$ . A la vista de las metodologías propuestas en trabajos previos, una generalización razonable de (12) para modelar sistemas dinámicos con indefiniciones se consigue admitiendo que, en (14),  $F$  es una multifunción “fuzzy” (es decir, que  $F(t, x)$  es un conjunto difuso para cada  $(t, x)$ ). Con este objetivo, Zhu y Rhao introdujeron en [34] las *inclusiones diferenciales difusas*.

El concepto es como sigue. Sean  $H : [0, T] \times \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  una multifunción “fuzzy”,  $\alpha : \mathbb{R}^n \mapsto [0, 1]$  una función y  $J \subset \mathbb{R}$  un intervalo y consideremos el problema siguiente: Hallar  $x : J \mapsto \mathbb{R}^n$  absolutamente continua y tal que

$$\dot{x}(t) \in L_{\alpha(x(t))} [H(t, x(t))], \quad t \in J \text{ c.t.p.} \quad (15)$$

Diremos que (15) es una *inclusión diferencial “fuzzy”*. Abreviadamente, escribiremos (15) en la forma

$$\dot{x} \in L_{\alpha(x)} [H(t, x)], \quad t \in J. \quad (16)$$



Vemos pues que una inclusión diferencial “fuzzy” está determinada por la multifunción “fuzzy”  $H$  y por la *función de nivel*  $\alpha$ . Si  $x = x(t)$  es una solución de (16) en el intervalo  $J$ , entonces

$$H(t, x(t))(\dot{x}(t)) \geq \alpha(x(t)) \tag{17}$$

para  $t \in J$  c.t.p.

Supongamos que

$$\alpha(x) > 0 \quad \forall x \in \mathbb{R}^n.$$

Si  $H$  toma valores en  $\mathbb{K}(\mathbb{R}^n)$ , i.e. si

$$H(t, x) = \chi_{F(t,x)} \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n,$$

donde  $F : [0, T] \times \mathbb{R}^n \mapsto \mathbb{K}(\mathbb{R}^n)$  es una multifunción “clásica”, entonces la inclusión diferencial “fuzzy” (16) es equivalente a la inclusión diferencial clásica

$$\dot{x} \in F(t, x) \quad t \in J \text{ c.t.p.}$$

Más adelante veremos que, para otras elecciones apropiadas de  $H$  y  $\alpha$ , (17) también equivale a una inclusión diferencial “clásica”.

A continuación, hablaremos de la existencia de solución del problema de valor inicial para (16) y, posteriormente, de la estabilidad de los puntos de equilibrio.

El problema de valor inicial para (16) consiste en hallar una solución  $x = x(t)$  en un intervalo de la forma  $[0, T^*)$  que, además, verifique  $x(0) = x_0$ , donde  $x_0 \in \mathbb{R}^n$  es un punto dado. Abreviadamente, pondremos

$$\begin{cases} \dot{x} \in L_{\alpha(x)} [H(t, x)], \\ x(0) = x_0. \end{cases} \tag{18}$$

Combinando el Teorema 1 con las definiciones y propiedades que preceden, es posible deducir el resultado de existencia siguiente:

**Teorema 17** *Supongamos que, en (18), la multifunción “fuzzy”  $H : [0, T] \times \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  y la función continua  $\alpha : \mathbb{R}^n \mapsto [0, 1]$  son tales que la multifunción*

$$(t, x) \mapsto L_{\alpha(x)} [H(t, x)]$$

*es semicontinua superiormente. Entonces, para cada  $x_0 \in \mathbb{R}^n$ , el problema (18) posee al menos una solución en  $[0, T^*)$  para algún  $T^*$  maximal que verifica  $0 < T^* \leq T$ .*

Para analizar las propiedades de estabilidad de los puntos de equilibrio de (16), como en la Sección 2, nos restringiremos al caso autónomo. Consideraremos por tanto la inclusión diferencial “fuzzy”

$$\dot{x} \in L_{\alpha(x)} [H(x)], \quad t \in J, \tag{19}$$

y el problema de valor inicial

$$\begin{cases} \dot{x} \in L_{\alpha(x)}[H(x)], \\ x(0) = x_0, \end{cases}$$

donde  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$ ,  $\alpha : \mathbb{R}^n \mapsto [0, 1]$  y  $x_0 \in \mathbb{R}^n$ .

Convendrá hablar previamente de multifunciones “fuzzy” *cuasi-lineales* y de la diferencial de una multifunción “fuzzy”. Recordaremos en consecuencia estos conceptos, que han sido introducidos en [11] y [12] y generalizan los que fueron presentados en la Sección 2.

**Definición 13** *Se dice que una multifunción “fuzzy”  $Q : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  es cuasi-lineal si*

$$Q(x_1 + x_2) \subset Q(x_1) \oplus Q(x_2) \quad \forall x_1, x_2 \in \mathbb{R}^n,$$

y

$$Q(\lambda x) = \lambda \odot Q(x) \quad \forall \lambda \in \mathbb{R}, \quad \forall x \in \mathbb{R}^n.$$

*Se dice que  $Q$  es acotada si existe  $K > 0$  tal que*

$$D(Q(x), \chi_{\{0\}}) \leq K\|x\| \quad \forall x \in \mathbb{R}^n.$$

Diremos que la inclusión (19) es cuasi-lineal si la correspondiente multifunción  $H$  lo es.

**Definición 14** *Se dice que la multifunción  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  es Fréchet-diferenciable en  $x_0 \in \mathbb{R}^n$  si existe una multifunción “fuzzy” cuasi-lineal acotada  $H'(x_0) : \mathbb{R}^n \mapsto \mathbb{E}_{cc}(\mathbb{R}^n)$  tal que*

$$D(H(x), H(x_0) + H'(x_0)(x - x_0)) = o(\|x - x_0\|).$$

*La multifunción “fuzzy”  $H'(x_0)$  se llama derivada de Fréchet de  $H$  en  $x_0$ .*

No es difícil probar que, si  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  es cuasi-lineal y acotada, entonces es Fréchet-diferenciable en 0 y además  $H'(0) = H$ .

**Definición 15** *Consideremos la inclusión diferencial fuzzy (19). Supongamos que  $H(0) \ni \chi_{\{0\}}$ . Se dice entonces que 0 es un punto de equilibrio de (19). Diremos que este punto de equilibrio es estable (en el sentido de Lyapunov) si se cumplen las siguientes condiciones:*

1. *Existe  $\delta_0 > 0$  tal que, si  $\|x_0\| \leq \delta_0$ , entonces existe una solución  $x = x(t)$  de (19) con la condición inicial  $x(0) = x_0$  definida para todo  $t \geq 0$ .*
2. *Para todo  $\epsilon > 0$ , existe  $\delta_1 \in (0, \delta_0]$  tal que, si  $\|x(0)\| < \delta_1$ , entonces  $\|x(t)\| < \epsilon$  para todo  $t \geq 0$ .*

*Finalmente, diremos que el punto de equilibrio Lyapunov-estable  $x = 0$  es asintóticamente estable si existe  $\delta_2 \in (0, \delta_0]$  tal que, si  $\|x(0)\| \leq \delta_2$ , entonces  $\lim_{t \rightarrow +\infty} \|x(t)\| = 0$ .*

El resultado principal relativo a la estabilidad de los puntos de equilibrio de (19) es el siguiente:

**Teorema 18** *Supongamos que la multifunción “fuzzy”  $H : \mathbb{R}^n \mapsto \mathbb{E}(\mathbb{R}^n)$  es Fréchet-diferenciable en 0 y que existe  $\delta_0 > 0$  tal que cualquier solución  $x = x(t)$  existe en el intervalo  $[0, +\infty)$  si  $\|x(0)\| \leq \delta_0$ . En estas condiciones, si el origen es un punto de equilibrio asintóticamente estable de la inclusión diferencial cuasi-lineal*

$$\dot{y} \in L_a[H'(0)(y)]$$

para algún  $a \in [0, 1]$ , entonces también es un punto de equilibrio asintóticamente estable de (19). En tal caso, existen además  $\sigma > 0$ ,  $k > 0$  y  $\delta > 0$  tales que cualquier solución de (19) para la que se tenga  $\|x(0)\| \leq \delta$  verifica

$$\|x(t)\| \leq k \|x(0)\| e^{-\sigma t} \quad \forall t \geq 0.$$

## 6 Algunos casos particulares (dinámica de poblaciones con ruido)

En esta Sección, consideraremos varias situaciones particulares que aparecen en Dinámica de poblaciones y pueden ser modeladas de forma satisfactoria con las herramientas precedentes. Consideramos por simplicidad sólo sistemas unidimensionales, correspondientes al comportamiento de una población aislada.

Así, sea  $x(t)$  la densidad de población en el tiempo  $t$  y consideremos los modelos de crecimiento exponencial y logístico, dados por la ecuación  $\dot{x} = f(x)$ , con

$$f(x) = rx \quad \text{y} \quad f(x) = rx \left(1 - \frac{x}{k}\right), \quad (20)$$

respectivamente.

En Biología existen dos interpretaciones posibles para las perturbaciones que pueden aparecer en modelos como éstos: el ruido *demográfico* (que se debe a aptitudes individuales distintas) y el ruido *ambiental* (que tiene su origen en variaciones imprevistas). Nisbet y Gurney propusieron la siguiente aproximación para el ruido demográfico:

$$g(x) = \sqrt{(b(x) + d(x))x}, \quad (21)$$

donde  $b(x)$  y  $d(x)$  son las razones de natalidad y mortalidad instantáneas, respectivamente. En particular, para el modelo exponencial, tenemos  $b(x) \equiv r$ ,  $d(x) \equiv 0$  y  $g(x) \equiv \sqrt{rx}$ .

Supongamos que, de algún modo, hemos tenido conocimiento de que el ruido está acotado por una constante positiva  $C$ . Entonces tiene perfecto sentido considerar la inclusión diferencial

$$\dot{x} \in f(x) + Cg(x)[-1, 1], \quad (22)$$

donde  $f$  es alguna de las funciones de (20) y  $g$  está dada por (21). Un estudio detallado de (22) se lleva a cabo en [22].

Sin embargo, en poblaciones con gran densidad, el ruido demográfico es insignificante frente al ruido ambiental. Parece por tanto mucho más natural aceptar que el ruido afecta sólo a los parámetros. Suponiendo además que solamente el parámetro  $r$  se ve afectado, tenemos los modelos perturbados

$$\dot{x} = (r + v)x$$

y

$$\dot{x} = (r + v)x \left(1 - \frac{x}{k}\right).$$

Basándonos en estos modelos, en los conceptos anteriores y en la indefinición del parámetro  $r$ , se ha introducido en [13] un nuevo modelo, donde hemos interpretado  $r$  como un conjunto “fuzzy”. La presencia de este conjunto “fuzzy” está justificada por la nebulosidad de los factores que influyen sobre el aumento de población y/o la expectativa de vida.

Más precisamente, consideraremos inclusiones diferenciales del tipo

$$\dot{x} \in L_{\alpha(x)}[(\lambda_1 + \lambda_2 u) \cdot x], \quad t \in [0, +\infty), \quad (23)$$

donde  $\alpha : \mathbb{R} \mapsto [0, 1]$  es una función continua dada,  $\lambda_1$  y  $\lambda_2$  son funciones de la variable  $t$  y  $u$  es un conjunto “fuzzy”.

Cuando  $\lambda_1$  y  $\lambda_2$  son independientes de  $t$ , vemos que (23) es una inclusión diferencial autónoma. La multifunción “fuzzy” asociada es cuasi-lineal y acotada y, por tanto, Fréchet-diferenciable en 0. También tenemos que  $x = 0$  es un punto de equilibrio de (23).

### 6.1 Ejemplo 1: Un modelo “fuzzy” con crecimiento exponencial

Supongamos que, en ausencia de ruido, la población considerada está descrita por el modelo de crecimiento exponencial

$$\dot{x} = rx, \quad (24)$$

donde  $r$  es la tasa de natalidad cuando la población posee todas las condiciones necesarias para subsistir, i.e. cuando hay espacio suficiente, nutrientes, etc. ( $r$  es una constante positiva).

Ahora supongamos que el parámetro  $r$  es perturbado por la saturación del espacio y que esto determina el grado de competición entre los componentes de la especie. Introduzcamos el conjunto “fuzzy”  $u$ , con

$$u(s) = \frac{k}{k + s} \quad \forall s \geq 0.$$

La variable  $s$  mide la *cantidad de espacio libre* y  $u$  modela la *competición entre especies* de acuerdo con la capacidad límite  $k$ , basándonos en el principio de que, cuanto mayor es el espacio, menor es el nivel de competición dentro de la población.

La perturbación de  $r$  basada en el conjunto “fuzzy”  $u$  conduce a la ley

$$\dot{x} \in L_{\alpha(x)}[rx - x \odot u], \quad t \in [0, +\infty),$$

que por analogía con las leyes “estándar” escribiremos

$$\dot{x} \in L_{\alpha(x)}[(r - u)x], \quad t \in [0, +\infty). \quad (25)$$

Naturalmente, la función  $\alpha = \alpha(x)$  debe ser dada.

Obsérvese que, si  $s \rightarrow +\infty$ , tenemos  $u(s) \rightarrow 0$ . Luego, en el límite  $s \rightarrow +\infty$ , (25) se reduce formalmente al modelo determinista (24).

Elijamos  $\alpha$  como sigue:

$$\alpha(x) = \frac{k^2}{k^2 + x^+} \quad \forall x \in \mathbb{R}.$$

Entonces

$$\begin{aligned} L_{\alpha(x)}[u] &= \left\{ s \in [0, +\infty) ; \frac{k}{k+s} \geq \frac{k^2}{k^2+x^+} \right\} \\ &= \left[ 0, \frac{x^+}{k} \right]. \end{aligned}$$

Por tanto, con esta elección de  $\alpha$ , (25) es equivalente a la inclusión diferencial ordinaria

$$\dot{x} \in \left( r - \frac{x}{k}[0, 1] \right) x. \quad (26)$$

En (25) tenemos que  $x = 0$  es un punto de equilibrio. Pero no se trata de un punto Lyapunov-estable, pues existen soluciones de (26) de la forma

$$x(t) = \frac{x_0 r k e^{rt}}{\theta x_0 e^{rt} + (rk - x_0 \theta)} \quad \text{con } \theta \in [0, 1].$$

Así, si  $0 \leq \theta < rk \text{ mín}(1, 1/x_0)$ , las soluciones verifican  $x(t) \rightarrow \infty$  cuando  $t \rightarrow +\infty$ .

## 6.2 Ejemplo 2: Un modelo para la expectativa de vida de una población.

Supongamos dada una población con  $x(t)$  individuos en cada instante de tiempo  $t \geq 0$ . Presentaremos a continuación un modelo para determinar la expectativa de vida en el que admitiremos que la pobreza es un factor que contribuye a la mortalidad de los individuos.

Para modelar la pobreza, podríamos utilizar varios indicadores distintos, como el consumo de vitaminas, el saneamiento básico, etc. Aquí, elegiremos como factor determinante la renta económica del grupo estudiado.

Razonando como en [6], vemos que el conjunto “fuzzy” de los pobres de una determinada localidad queda bien descrito poniendo

$$u(s) = \begin{cases} \left[ 1 - \left( \frac{s}{r_0} \right)^2 \right]^k & \text{si } 0 \leq s < r_0, \\ 0 & \text{si } s \geq r_0, \end{cases}$$

donde  $k$  es un parámetro positivo característico del grupo,  $s$  es una variable proporcional a la renta del individuo considerado y  $r_0$  es la renta mínima, a partir de la cual los individuos no tienen ya categoría de pobres y, en consecuencia, no influyen en la tasa de mortalidad.

El modelo viene dado por la inclusión diferencial “fuzzy”

$$\dot{x} \in L_{\alpha(x)}[-\lambda_1 x + \lambda_2 x \odot u], \quad t \in [0, +\infty),$$

que, de nuevo por analogía con las inclusiones diferenciales clásicas, escribiremos en la forma

$$\dot{x} \in L_{\alpha(x)}[-(\lambda_1 + \lambda_2 u)x], \quad t \in [0, +\infty). \quad (27)$$

Aquí,  $\lambda_1$  y  $\lambda_2$  son constantes positivas. Más precisamente,

- $\lambda_1$  es la tasa de mortalidad natural, obtenida a partir del análisis de un grupo que dispone de condiciones satisfactorias de supervivencia.
- $\lambda_2$  es un factor de proporcionalidad de la influencia de la pobreza en el aumento de la tasa de mortalidad del grupo.

Observamos que, si  $s \geq r_0$ , entonces  $u(s) = 0$  y (27) se reduce, al menos formalmente, al modelo determinista

$$\dot{x} = -\lambda_1 x.$$

Pongamos

$$\alpha(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ x^k & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Entonces, si  $x > 0$ , tenemos

$$\begin{aligned} L_{\alpha(x)}[u] &= \{s \geq 0; u(s) \geq \alpha(x)\} \\ &= r_0 [0, \sqrt{1 - \text{mín}(1, x)}]. \end{aligned}$$

La inclusión diferencial “fuzzy” (27) para  $0 < x \leq 1$  es, por tanto, equivalente a la inclusión diferencial

$$\dot{x} \in -\lambda_1 x - \lambda_2 r_0 x \sqrt{1 - \text{mín}(1, x)} [0, 1] \quad (28)$$

(justamente este hecho hace pensar que la idea de enfocar los problemas de Dinámica de poblaciones usando inclusiones diferenciales “fuzzy” es realmente útil).

Así, buscando selecciones apropiadas de la multifunción que aparece en (28), podemos determinar soluciones. Por ejemplo, las funciones  $f_1$ ,  $f_2$  y  $f_3$ , con

$$\begin{aligned} f_1(x) &= -\lambda_1 x - \lambda_2 r_0 x \sqrt{1 - \text{mín}(1, x^+)}, \\ f_2(x) &= -\lambda_1 x - \lambda_2 r_0 x (1 - \text{mín}(1, x^+)) \end{aligned}$$

y

$$f_3(x) = -\lambda_1 x,$$

son selecciones continuas. Para cada  $i$ , podemos fácilmente resolver la ecuación diferencial

$$\dot{x} = f_i(x),$$

completada con la condición inicial  $x(0) = x_0$ . Esto conduce a las soluciones deseadas.

En particular, tenemos las siguientes:

$$x_2(t) = \frac{(\lambda_1 + \lambda_2 r_0)x_0}{[(\lambda_1 + \lambda_2 r_0) - \lambda_2 r_0 x_0]e^{(\lambda_1 + \lambda_2 r_0)t} + \lambda_2 r_0 x_0};$$

$$x_3(t) = x_0 e^{-\lambda_1 t}.$$

Veamos a continuación que  $x = 0$  es un punto de equilibrio asintóticamente estable para (27).

En esta inclusión diferencial “fuzzy” tenemos la multifunción  $H(x) \equiv -(\lambda_1 + \lambda_2 u)x$ , que es cuasi-lineal acotada y, por tanto, Fréchet-diferenciable en 0, con

$$H'(0)x = -(\lambda_1 + \lambda_2 u)x \quad \forall x \in \mathbb{R}.$$

Probaremos a continuación que, para algún  $a \in [0, 1]$ , el origen es un punto de equilibrio asintóticamente estable de la inclusión diferencial cuasi-lineal

$$\dot{y} \in L_a[H'(0)(y)]. \quad (29)$$

En efecto, tomemos  $a = (\frac{1}{2})^k$ . Entonces (29) se reduce a

$$\dot{x} \in -\left(\lambda_1 + \lambda_2 r_0 [0, 1/\sqrt{2}]\right)x \quad (30)$$

Las soluciones de (30) son de la forma  $x(t) = x(0) \exp(-(\lambda_1 + \theta \lambda_2 r_0)t)$ , con  $\theta \in [0, \frac{1}{\sqrt{2}}]$  y existen para todo  $t \geq 0$ . Por otra parte, dado  $\epsilon > 0$ , existe  $\delta_1 = \epsilon$  tal que  $\|x(t)\| = \|x(0) \exp(-(\lambda_1 + a \lambda_2 r_0)t)\| \leq \|x(0)\|$  para todo  $t \geq 0$ . Además,  $\lim_{t \rightarrow \infty} \|x(t)\| = 0$ .

Luego el origen es un punto de equilibrio asintóticamente estable para la inclusión (29) con  $a = (\frac{1}{2})^k$ . Usando el Teorema 16, deducimos que  $x = 0$  es también un punto de equilibrio asintóticamente estable para la inclusión diferencial “fuzzy” (27). Además, existen  $\sigma > 0$ ,  $k > 0$  y  $\delta > 0$  tales que cualquier solución  $x = x(t)$  que cumpla que  $\|x(0)\| < \delta$  verifica

$$\|x(t)\| \leq k \|x(0)\| \exp(-\sigma t) \quad \forall t \geq 0.$$

## Referencias

- [1] T.S. Angell, *Existence of solutions of multivalued Uryshon integral equations*, J. Optim. Theory Appl. 46 (1985), no. 2, 129–151.

- [2] J.P. Aubin, *A survey of viability theory*, SIAM J. Control and Optimization, Vol. 28 (1990), p. 749–788.
- [3] J.P. Aubin, A. Cellina, *Differential Inclusions, Set-Valued Maps and Viability Theory*, Springer-Verlag, Berlin 1984.
- [4] J.P. Aubin, H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston 1990.
- [5] R. Aumann, *Integral of set-valued functions*, J. Math. Anal. Appl. 12 (1965), p. 1–12.
- [6] R.C. Bassanezi and L.C. Barros, *A simple model of life expectancy with subjective parameters*, Kibernetics: Int. J. of Systems and Cybernetics 24, Vol. 9, p. 91–98 (1995).
- [7] V.I. Blagodatskii y A.F. Filippov, *Differential inclusions and optimal control*, Proc. of the Steklov Inst. Math. 4 (1986), p. 199–259.
- [8] Y.G. Borisovich, B.D. Gelman, A.D. Myshkis, V.V. Obukhovskii, *Multivalued mappings*, J. Soviet Math. 49 (1990), no. 1, p. 800–855.
- [9] A. Cellina, *Approximation of set valued functions and fixed point theorems*, Ann. Mat. Pura Appl. (4) 82 (1969), p. 247–262.
- [10] L. Cesari, *Optimization - Theory and Applications*, Springer-Verlag, Berlin 1983.
- [11] Y. Chalco-Cano, M.A. Rojas-Medar and A.J.V.Brandão, *On the differentiability of fuzzy-valued mappings and the stability of a fuzzy differential inclusion*, submitted (2001).
- [12] Y. Chalco-Cano, M.A. Rojas-Medar and A.J.V. Brandão, *Fuzzy quasilinear spaces*, preprint (2001).
- [13] Y. Chalco-Cano, R.C. Bassanezi, M.A. Rojas-Medar and T. Mizukoshi, *Fuzzy differential inclusions and applications*, in preparation.
- [14] F.H. Clarke, G. Watkins, *Necessary conditions, controllability and the value function for differential-difference inclusions*, Nonlinear Anal., Vol. 10, no. 11, p. 1155–1179, 1986.
- [15] B. Cornet, G. Haddad, *Théorèmes de viabilité pour les inclusions différentielles du second ordre*, Israel J. Math., Vol. 57, no. 2, 1987.
- [16] Fan Van Cyong, *A smooth choice theorem and its application to multivalued integral equations*, Mat. USSR Sb. 105 (147) (1978), no. 4, p. 622–637, 640.
- [17] G. Debreu, *Integration of correspondences*, in “Proc. of the Fifth Berkeley Symp. Math Statistic. Probability”, p. 531–372, University of California Press, Berkeley 1966.



- [18] W.H. Fleming, R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Applications of Mathematics, No. 1. Springer-Verlag, Berlin-New York 1975.
- [19] V.I. Istrăţescu, *Fixed point theory. An introduction*, Reidel Publishing Co., Dordrecht-Boston 1981.
- [20] M. Kisielewicz, *Differential Inclusions and Optimal Control*, Kluwer, Warsaw 1991.
- [21] E. Klein, A. Thompson, *Theory of Correspondences*, John Wiley & Sons, New York 1984.
- [22] V. Krivan, G. Colombo, *A non-stochastic approach for modelling uncertainty in population dynamics*, Bull. Math. Biology 60, p. 721–751, 1998.
- [23] L. Lushu, *Random fuzzy sets and fuzzy martingales*, Fuzzy Sets and Systems, 69 (1995), p. 181–192.
- [24] E. Michael, *Continuous selections I*, Ann. of Math., (2) 63 (1956), p. 361–382.
- [25] S.B. Nadler, *Multi-valued contraction mappings*, Pacific J. Math. 30 (1969), p. 475–478.
- [26] M.L. Nagoita, D. Ralescu, *Applications of Fuzzy Sets to Systems Analysis*, Wiley, New York (1975).
- [27] N.S. Papageorgiou, *On multivalued evolution equations and differential equations in Banach spaces*, Comment. Math. Univ. St. Paul. 36 (1987), no. 1, p. 21–39.
- [28] M. Puri and D.A. Ralescu, *Fuzzy random variables*, J. Math. Anal. Appl. 114 (1986), p. 409–422.
- [29] M.A. Rojas-Medar, R.C. Bassanezi, H. Román-Flores, *A generalization of the Minkowski embedding theorem and applications*, Fuzzy Sets and Systems, 102 (1999), no. 2, p. 263–269.
- [30] H. Román-Flores, R.C. Bassanezi, M.A. Rojas-Medar, A. Flores-Franulic, *Tópicos en Análisis Fuzzy-Multívoco (minicurso)*, Actas XI Jornada de Matemáticas, Universidad de Tarapacá, Arica, Chile (1993).
- [31] H. Román-Flores, A. Flores-Franulič, M.A. Rojas-Medar, R.C. Bassanezi, *Stability of the fixed points set of fuzzy contractions*, Applied Math. Letters, Vol. 11, no. 4, p. 33–37, 1998.
- [32] H. Román-Flores, M.A. Rojas-Medar, *Embedding of level-continuous fuzzy sets on Banach spaces*, Inf. Sciences 144 (2002), p. 227–247.

- [33] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York 1972.
- [34] Yuanguo Zhu and Ling Rao, *Differential inclusions for fuzzy maps*, Fuzzy Sets and Systems, 112 (2000), p. 257–261.
- [35] L.A. Zadeh, *Fuzzy sets*, Information and Control 8 (1965), p. 338–353.

## Metasolutions

J. LÓPEZ-GÓMEZ

Departamento de Matemática Aplicada.  
Universidad Complutense de Madrid

Lopez\_Gomez@mat.ucm.es

### Abstract

Metasolutions are a class of very weak generalized solutions introduced to describe the dynamics of reaction diffusion equations with spatially heterogeneous coefficients. They have been introduced by the author and a number of coworkers in order to clarify a number of quantitative features specific of spatially heterogeneous media. This paper goes back up to the generation of the first metasolution fully describing the temporary evolution and main hits of the theory and eventually providing with some new very recent findings. The theory has a huge number of applications in wide areas of mathematics and in many applied sciences, e.g., ecology, chemistry, and biology. As most of the first part of the paper is of expository nature, an appreciable number of applied mathematicians might enjoy its reading.

**Key words:** *semilinear parabolic equations, classical solutions, large solutions, metasolutions, attracting properties.*

**AMS subject classifications:** *35K20, 35B40, 35B30.*

## 1 Going back to the first metasolution

Metasolutions are a class of very weak generalized solutions introduced to describe the dynamics of reaction diffusion equations with spatially heterogeneous coefficients. Although they are extensions by infinity of large solutions and large solutions have been extensively studied in the literature, the concept of metasolution has been introduced by R. Gómez-Reñasco and the author in [20] and [18]; [20] was primarily submitted to Professor H. Brezis on September 24th 1998 for publication in the *Arch. Rat. Mech. Anal.*, and, on March 4th 1999, it was formally rejected. Then, it was submitted to Professor

---

Fecha de recepción: 4 de septiembre de 2002

P. H. Rabinowitz for publication in *Nonl. Anal. TMA*, where it appeared after three years time. Fixing dates might be imperative to avoid some eventual priority dispute.

Although metasolutions arose from the mathematical analysis of the logistic equation with diffusion—a paradigmatic model in Ecology—, we believe that their role in understanding the dynamics of more realistic reaction diffusion equations will be shown to be crucial upon the next few years; as they will be necessary in order to understand the dynamics of all spatial heterogeneous models. Nevertheless, in this sort of introduction to the theory of metasolutions we are going to concentrate on the mathematical analysis of the model that originally motivated its discovery. So, attention will be focused into the problem of ascertaining the asymptotic behavior of the solutions of

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = \lambda u - a(x)u^p & \text{in } \Omega \times (0, \infty) \\ u(x, t) = 0 & \text{on } \partial\Omega \times (0, \infty) \\ u(\cdot, 0) = u_0 \geq 0 & \text{in } \Omega \end{cases} \quad (1)$$

where  $\Omega$  is a bounded domain of  $\mathbb{R}^N$ ,  $N \geq 1$ , with smooth boundary  $\partial\Omega$ , e.g., of class  $\mathcal{C}^3$ ,  $\lambda \in \mathbb{R}$ ,  $p \in (1, \infty)$ , and  $a \geq 0$ ,  $a \neq 0$ , is a smooth function, e.g., of class  $\mathcal{C}^1$ , such that

$$\Omega_+ := \{x \in \Omega : a(x) > 0\}$$

is a subdomain of  $\Omega$  with  $\bar{\Omega}_+ \subset \Omega$ , whose boundary,  $\partial\Omega_+$ , is of class  $\mathcal{C}^3$ , and the open set

$$\Omega_0 := \Omega \setminus \bar{\Omega}_+$$

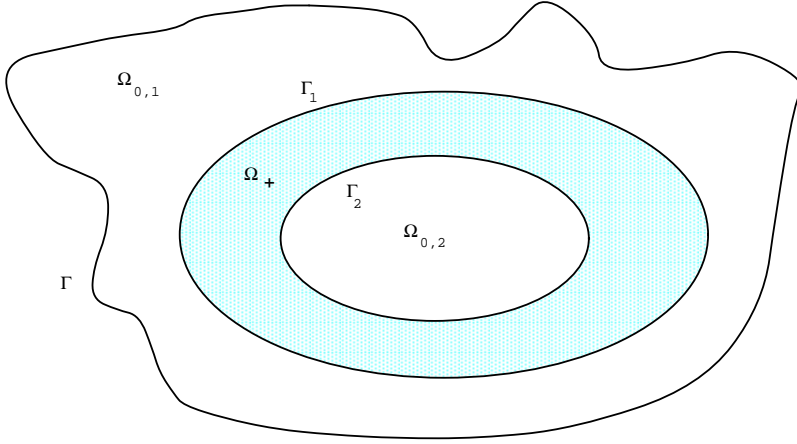
consists of two components,  $\Omega_{0,i}$ ,  $i \in \{1, 2\}$ , such that  $\bar{\Omega}_{0,1} \cap \bar{\Omega}_{0,2} = \emptyset$  and  $\bar{\Omega}_{0,2} \subset \Omega$ . Figure 1 represents typical situation where all the assumptions made on  $a(x)$  are satisfied. Note that

$$\Gamma = \partial\Omega, \quad \Gamma_1 = \partial\Omega_{0,1} \setminus \partial\Omega, \quad \Gamma_2 := \partial\Omega_{0,2}, \quad \partial\Omega_+ = \Gamma_1 \cup \Gamma_2.$$

Subsequently, given a regular subdomain  $D$  of  $\Omega$  and  $V \in C(\bar{D})$ , we denote by  $\sigma[-\Delta + V, D]$  the principal eigenvalue of  $-\Delta + V$  in  $D$  under homogeneous Dirichlet boundary conditions, and, using this notation, it will be throughout assumed that

$$\sigma_1 := \sigma[-\Delta, \Omega_{0,1}] < \sigma_2 := \sigma[-\Delta, \Omega_{0,2}]. \quad (2)$$

Thanks to Faber-Krahn inequality, (2) can be reached by taking  $\Omega_{0,2}$  with sufficiently small Lebesgue measure (e.g., [24, Section 5] and [7, Section 10]). Actually, one might think in (2) as a sort of hierarchical order size between the components  $\Omega_{0,1}$  and  $\Omega_{0,2}$  establishing that  $\Omega_{0,1}$  is larger than  $\Omega_{0,2}$ , although one should be careful since the principal eigenvalue  $\sigma[-\Delta; D]$  also depends on some hidden geometrical properties of  $D$ , and not merely on its size.



**Figure 1:** The nodal behavior of  $a(x)$ .

In Ecology, (1) provides us with the spatial-temporal distribution of a single species  $u(x, t)$  diffusing in the inhabiting area  $\Omega$  according to the classical Fick-Fourier law, where  $\lambda$  represents the neat birth —or dead— rate and  $a(x)$  measures the intra-specific competition in  $\Omega_+$ , while in  $\Omega_0$  the species is allowed to grow exponentially, according to Malthus's law. The inhabiting area  $\Omega$  is assumed to be fully surrounded by completely hostile regions, because of the homogeneous Dirichlet boundary conditions, and  $u_0 \in \mathcal{C}(\bar{\Omega})$  represents the initial population distribution.

Apart from the intrinsic interest of model (1) in the applied sciences and engineering, analyzing (1) is imperative because it provides us with a sort of cellular component of more sophisticated —and, so, more realistic— multi-species interacting models. As a consequence, the analysis carried out throughout the next pages should be of huge importance in wide areas of science and engineering.

Under the assumptions above, (1) possesses a unique smooth solution  $u_{[\lambda, a, \Omega]}(x, t; u_0)$  which is non-negative and globally defined in time, since  $au^p \geq 0$ . Subsequently, given any function  $v \in \mathcal{C}(\bar{\Omega})$  it is said that  $v > 0$  if  $v(x) \geq 0$  for each  $x \in \bar{\Omega}$  and  $v \neq 0$ , and, if  $v \in \mathcal{C}_0^1(\bar{\Omega})$ , then it is said that  $v \gg 0$  if  $v(x) > 0$  for each  $x \in \bar{\Omega}$  and  $\frac{\partial v}{\partial n}(x) < 0$  for each  $x \in \partial\Omega$ , where  $n$  stands for the outward unit normal to  $\Omega$  at  $x \in \partial\Omega$ ;  $\mathcal{C}_0^1(\bar{\Omega})$  is the space of functions of class  $\mathcal{C}^1$  in  $\bar{\Omega}$  vanishing on  $\partial\Omega$ . By the parabolic maximum principle, one has that  $u_{[\lambda, a, \Omega]} = 0$  if  $u_0 = 0$ , while  $u_{[\lambda, a, \Omega]}(\cdot, t; u_0) \gg 0$  for each  $t > 0$  if  $u_0 > 0$ . Our main goal here is ascertaining the behavior of the population distribution as time passes by, and, particularly, finding out the limiting behavior of the population, i.e., determining the limit

$$U_{[\lambda, a, \Omega]}(\cdot; u_0) := \lim_{t \nearrow \infty} u_{[\lambda, a, \Omega]}(\cdot, t; u_0) \quad (3)$$

—if such a limit exists!—, according to each of the admissible values of the several parameters used in the formulation of the problem.

Suppose  $D \subset \Omega$  is an open set such that

$$U_{[\lambda,a,\Omega]}(x; u_0) := \lim_{t \nearrow \infty} u_{[\lambda,a,\Omega]}(x, t; u_0) \in \mathbb{R} \quad \text{for each } x \in D.$$

Then, under some minimal regularity requirements, passing to the limit as  $t \uparrow \infty$  in the differential equation of (1), gives

$$-\Delta U_{[\lambda,a,\Omega]} = \lambda U_{[\lambda,a,\Omega]} - a U_{[\lambda,a,\Omega]}^p \quad \text{in } D$$

and, henceforth, it is rather natural postulating that all limiting profiles of the positive solutions of (1) will be given through the non-negative solutions of the elliptic boundary value problem

$$\begin{cases} -\Delta u = \lambda u - a(x)u^p & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \quad (4)$$

Actually, all classical known results concerning (1) corroborated this claim. By classical results it is meant all those available for (1) in the very special case when  $a$  is positive and bounded away from zero, e.g.,

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = \lambda u - (a + \mu)u^p & \text{in } \Omega \times (0, \infty) \\ u(x, t) = 0 & \text{on } \partial\Omega \times (0, \infty) \\ u(\cdot, 0) = u_0 \geq 0 & \text{in } \Omega \end{cases} \quad (5)$$

where  $\mu > 0$  is constant, whose associated steady-states are the non-negative solutions of

$$\begin{cases} -\Delta u = \lambda u - (a + \mu)u^p & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \quad (6)$$

Nevertheless, there is a crucial difference between (1) and (5) as a result of the fact that  $a$  vanishes somewhere. Indeed, although due to the continuous dependence of the solutions of (1) with respect to  $\mu$  for any  $T > 0$ —arbitrarily large—and  $\varepsilon > 0$ —arbitrarily small—there exists  $\mu_0 > 0$  such that

$$\|u_{[\lambda,a+\mu,\Omega]}(\cdot, \cdot; u_0) - u_{[\lambda,a,\Omega]}(\cdot, \cdot; u_0)\|_{L^\infty(\bar{\Omega} \times [0, T])} \leq \varepsilon \quad \text{for each } 0 \leq \mu \leq \mu_0,$$

it turns out that in case  $\lambda \geq \sigma_1$  the limiting profiles of the solutions of (1) differ considerably from those of the solutions of (5), independently of the size of  $\mu > 0$ . Namely, however  $U_{[\lambda,a+\mu,\Omega]}$  provides us with the unique positive solution of (6) for any  $\lambda \geq \sigma_1$  and  $\mu > 0$ ,  $U_{[\lambda,a,\Omega]} = \infty$  in  $\bar{\Omega}_{0,1} \setminus \partial\Omega$  for any  $\lambda \geq \sigma_1$ , and  $U_{[\lambda,a,\Omega]} = \infty$  in  $(\bar{\Omega}_{0,1} \setminus \partial\Omega) \cup \bar{\Omega}_{0,2}$  if  $\lambda \geq \sigma_2$ . Actually, explaining all those differences is the main goal of this work.

Throughout the rest of this paper we will use the notations

$$\sigma_0 := \sigma[-\Delta, \Omega], \quad \tilde{\Omega} := \Omega_+ \cup \bar{\Omega}_{0,2} = \Omega \setminus \bar{\Omega}_{0,1},$$

(see Figure 1). Then, by the monotonicity of the principal eigenvalue with respect to the underlying domain (e.g., [24, Proposition 3.2(iii)]), (2) gives

$$\sigma_0 < \sigma_1 < \sigma_2.$$

The following result collects some of the most important results for the perturbed problem (5); it is a classical result attributable to H. Amann [1], D. Cohen and T. Laetsch [8], and D. Sattinger [36], among others.

**Theorem 1** *Suppose  $\mu > 0$  and  $u_0 > 0$ . Then,*

- (a) *The problem (6) possesses a positive solution if, and only if,  $\lambda > \sigma_0$ . Moreover, it is unique if it exists and if we denote it by  $\theta_{[\lambda, a+\mu, \Omega]}$ , then*

$$\lim_{t \nearrow \infty} u_{[\lambda, a+\mu, \Omega]}(\cdot, t; u_0) = \theta_{[\lambda, a+\mu, \Omega]} \quad \text{uniformly in } \bar{\Omega}.$$

- (b) *In case  $\lambda \leq \sigma_0$ , we have that*

$$\lim_{t \nearrow \infty} u_{[\lambda, a+\mu, \Omega]}(\cdot, t; u_0) = 0 \quad \text{uniformly in } \bar{\Omega}.$$

*In particular, for each  $\lambda \in \mathbb{R}$  the maximal non-negative solution of (6) is a global attractor for the positive solutions of (5).*

**Sketch of the Proof:** It is based upon a rather simple application of the method of sub and supersolutions for both models (5) a (6) (cf. [2] and [36]). Eventually, we will use the characterization of the strong maximum principle found in [31]; the details of the proof should be consulted in [24, Section 2]. We first focus our attention on (6). For each  $\lambda \in \mathbb{R}$  large positive constants are supersolutions. Let  $\varphi > 0$  denote the principal eigenfunction associated with  $\sigma_0$ . Then, for each  $\lambda > \sigma_0$ ,  $\varepsilon\varphi$  is a positive subsolution of (6) if  $\varepsilon > 0$  is sufficiently small. Thus, for each  $\lambda > \sigma_0$  the problem (5) possesses an ordered sub and supersolution pair and, therefore, it admits a minimal and a maximal solution in the corresponding order interval, necessarily positive. Now, suppose (6) possesses a positive solution  $u > 0$ . Then,

$$[-\Delta + (a + \mu)u^{p-1}]u = \lambda u \quad \text{in } \Omega \quad \text{and} \quad u|_{\partial\Omega} = 0.$$

Thus,  $\lambda$  is an eigenvalue of  $-\Delta + (a + \mu)u^{p-1}$  associated with it there is a positive eigenfunction  $u > 0$ . Henceforth, by Krein-Rutman theorem and the uniqueness of the principal eigenvalue we obtain that  $u \gg 0$  and

$$\lambda = \sigma[-\Delta + (a + \mu)u^{p-1}, \Omega].$$

Thus, by the monotonicity of the principal eigenvalue with respect to the potential, one has  $\lambda > \sigma_0$  and, therefore, (6) cannot admit a positive solution if  $\lambda \leq \sigma_0$ . Actually, this entails that (6) cannot admit a positive subsolution if  $\lambda \leq \sigma_0$ , since large positive constants provide us with arbitrarily large supersolutions.

To show the uniqueness of the positive solution suppose  $\lambda > \sigma_0$  exists for which (6) possesses two positive solutions  $u \neq v$ . We already know that  $u \gg 0$  and  $v \gg 0$ , and that  $u$  and  $v$  are the principal eigenfunctions of  $\sigma[-\Delta + (a + \mu)u^{p-1}, \Omega]$  and  $\sigma[-\Delta + (a + \mu)v^{p-1}, \Omega]$ , respectively. Moreover,

$$\lambda = \sigma[-\Delta + (a + \mu)u^{p-1}, \Omega] = \sigma[-\Delta + (a + \mu)v^{p-1}, \Omega]. \quad (7)$$

On the other hand, since

$$u^p - v^p = \int_0^1 \frac{d}{dt} [tu + (1-t)v]^p dt = p \int_0^1 [tu + (1-t)v]^{p-1} dt (u - v),$$

it is easily seen that

$$\left[ -\Delta + (a + \mu)p \int_0^1 [tu + (1-t)v]^{p-1} dt \right] (u - v) = \lambda(u - v).$$

Thus,  $w := u - v \neq 0$  provides us with an eigenfunction of  $-\Delta + (a + \mu)p \int_0^1 [tu + (1-t)v]^{p-1} dt$  associated with the eigenvalue  $\lambda$ . Hence, since the principal eigenvalue is the lowest one, we find that

$$\lambda \geq \sigma[-\Delta + (a + \mu)p \int_0^1 [tu + (1-t)v]^{p-1} dt, \Omega]. \quad (8)$$

On the other hand,  $(1-t)v > 0$  implies that

$$p \int_0^1 [tu + (1-t)v]^{p-1} dt > p \int_0^1 t^{p-1} u^{p-1} dt = u^{p-1}$$

and, hence, taking into account (7),

$$\sigma[-\Delta + (a + \mu)p \int_0^1 [tu + (1-t)v]^{p-1} dt, \Omega] > \sigma[-\Delta + (a + \mu)u^{p-1}, \Omega] = \lambda,$$

which contradicts (8). This contradiction shows the uniqueness of the positive solution for each  $\lambda > \sigma_0$ .

The global attractive character of the maximal non-negative solution of (6) is a consequence from the general theory of [36]. Indeed, suppose  $\lambda > \sigma_0$  and consider  $u_1 := u_{[\lambda, a + \mu, \Omega]}(\cdot, 1; u_0)$ . By the parabolic maximum principle  $u_1 \gg 0$ . Thus, if  $\varepsilon > 0$  is sufficiently small and  $M > 0$  is sufficiently large,  $\varepsilon\varphi < u_1 < M$  and, hence, for each  $t > 0$ ,

$$u_{[\lambda, a + \mu, \Omega]}(\cdot, t; \varepsilon\varphi) \leq u_{[\lambda, a + \mu, \Omega]}(\cdot, t; u_1) = u_{[\lambda, a + \mu, \Omega]}(\cdot, t+1; u_0) \leq u_{[\lambda, a + \mu, \Omega]}(\cdot, t; M) \quad (9)$$

Since  $\varepsilon\varphi$  is a subsolution of (6),  $u_{[\lambda, a + \mu, \Omega]}(\cdot, t; \varepsilon\varphi)$  grows with  $t$  and it approximates the minimal positive solution of (6) in the order interval  $[\varepsilon\varphi, M]$ ,  $\theta_{[\lambda, a + \mu, \Omega]}$ . Similarly, since  $M$  is a supersolution of (6),  $u_{[\lambda, a + \mu, \Omega]}(\cdot, t; M)$  decreases with  $t$  and it approximates the maximal positive solution of (6) in



the order interval  $[\varepsilon\varphi, M]$ ,  $\theta_{[\lambda, a+\mu, \Omega]}$ . Therefore, passing to the limit as  $t \uparrow \infty$  in (9) gives the desired result. When  $\lambda \leq \sigma_0$ , the function  $u_{[\lambda, a+\mu, \Omega]}(\cdot, t; M)$  must converge to zero, since otherwise (6) would have a positive solution and we already know that this is not the case. This concludes the proof.  $\square$

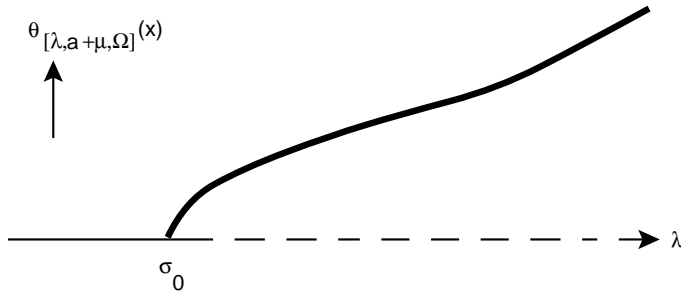
By using some further continuation techniques —implicit function theorem— it is easy to see that the mapping

$$\begin{aligned} (\sigma_0, \infty) &\longrightarrow C_0^2(\Omega) \\ \lambda &\mapsto \theta_{[\lambda, a+\mu, \Omega]} \end{aligned}$$

is differentiable and point-wise increasing. Moreover,

$$\lim_{\lambda \searrow \sigma_0} \theta_{[\lambda, a+\mu, \Omega]} = 0 \quad \text{uniformly in } \bar{\Omega}$$

i.e.,  $\theta_{[\lambda, a+\mu, \Omega]}$  bifurcates from  $u = 0$  at the value of the parameter  $\lambda = \sigma_0$ . In Figure 2 we have represented the positive solution  $\theta_{[\lambda, a+\mu, \Omega]}$  versus the parameter  $\lambda$ ; most precisely, we have taken  $x \in \Omega$  and represented  $\theta_{[\lambda, a+\mu, \Omega]}(x)$  versus  $\lambda$ . Stable solutions are indicated by solid lines, and unstable solutions by dashed lines. The state  $u = 0$  loses stability as  $\lambda$  crosses  $\sigma_0$ , such stability being gained by  $\theta_{[\lambda, a+\mu, \Omega]}$ . It turns out that local stability actually entails global attractiveness.



**Figure 2:** The steady-states of (5).

Nevertheless, when dealing with (4) —(6) for the special choice  $\mu = 0$ — live is severely much harder. The first difficulty coming from the fact that (4) cannot admit a positive solution if  $\lambda \geq \sigma_1$ . Indeed, if  $u$  is a positive solution of (4), then, arguing as in the proof of Theorem 1, we find that  $u \gg 0$  and  $\lambda = \sigma[-\Delta + au^{p-1}, \Omega]$ . Hence, by the monotonicity of the principal eigenvalue with respect to the domain,

$$\lambda = \sigma[-\Delta + au^{p-1}, \Omega] < \sigma[-\Delta + au^{p-1}, \Omega_{0,1}] = \sigma[-\Delta, \Omega_{0,1}] = \sigma_1$$

because  $a = 0$  in  $\Omega_{0,1}$ . Moreover, by the monotonicity with respect to the potential,

$$\lambda = \sigma[-\Delta + au^{p-1}, \Omega] > \sigma[-\Delta, \Omega] = \sigma_0$$

and, therefore, condition

$$\sigma_0 < \lambda < \sigma_1 \quad (10)$$

is necessary for the existence of a positive solution. As for any  $\lambda > \sigma_0$  the function  $\varepsilon\varphi$  constructed in the proof of Theorem 1 still provides us with a positive subsolution of (4) for any  $\varepsilon > 0$  small enough, the method of sub and supersolutions shows that (4) cannot admit a positive supersolution above  $\varepsilon\varphi$  if  $\lambda \geq \sigma_1$ , in strong contrast with the situation described by Theorem 1, where large positive constants are supersolutions of (6) above  $\varepsilon\varphi$ .

Nevertheless, for  $\lambda$  within the range (10) the problem (4) possesses arbitrarily large supersolutions above  $\varepsilon\varphi$  and, therefore, the proof of Theorem 1 can be easily adapted to get the following result.

**Theorem 2** *Suppose  $u_0 > 0$ . Then,*

- (a) (4) has a positive solution if, and only if,  $\sigma_0 < \lambda < \sigma_1$ . Moreover, it is unique if it exists and if we denote it by  $\theta_{[\lambda, a, \Omega]}$ , then  $\lim_{t \nearrow \infty} u_{[\lambda, a, \Omega]}(\cdot, t; u_0) = \theta_{[\lambda, a, \Omega]}$  uniformly in  $\bar{\Omega}$ .
- (b) In case  $\lambda \leq \sigma_0$ ,  $\lim_{t \nearrow \infty} u_{[\lambda, a, \Omega]}(\cdot, t; u_0) = 0$  uniformly in  $\bar{\Omega}$ .
- (c) In case  $\lambda \geq \sigma_1$ ,  $\lim_{t \nearrow \infty} u_{[\lambda, a, \Omega]}(\cdot, t; u_0) = \infty$  uniformly in compact subsets of  $\Omega_{0,1} \setminus \partial\Omega$ .

Part (a) is attributable to [14], although some related results had already been published at that time (e.g., T. Ouyang [35] and M. A. Del Pino [9]), none of them with the generality, amplitude and simplicity of techniques exhibited by [14]. Part (c) is attributable to [32, Theorem 4.3]. The proof given here is based upon [20], [25] and [32]; its key point being the construction of an adequate supersolution within the range (10) —those supersolutions go back to [22]—.

**Sketch of the proof:** It follows the general scheme of the proof of Theorem 1, except for the construction of the supersolution and the proof of Part (c). So, suppose (10) is satisfied and, for each  $\delta > 0$  sufficiently small, consider the open  $\delta$ -neighborhoods

$$\Omega_{\delta, j} := \Omega_{0, j} \cup \{x \in \Omega : \text{dist}(x, \Omega_{0, j}) < \delta\}, \quad j \in \{1, 2\}.$$

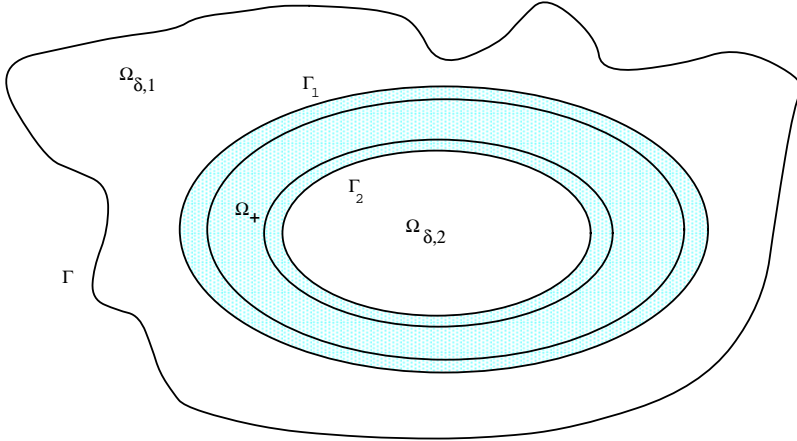
We have represented those neighborhoods in Figure 3. Note that  $\Omega_{\delta, 1}$  is formed by  $\Omega_{0, 1}$ ,  $\Gamma_1$  and the set of  $x \in \Omega_+$  such that  $\text{dist}(x, \Gamma_1) < \delta$ . Similarly,  $\Omega_{\delta, 2}$  is formed by  $\bar{\Omega}_{0, 2}$  and the set of  $x \in \Omega_+$  such that  $\text{dist}(x, \Gamma_2) < \delta$ . By the continuous dependence of the principal eigenvalues with respect to the domains (e.g., [24, Theorem 4.2]),

$$\lim_{\delta \searrow 0} \sigma[-\Delta, \Omega_{\delta, j}] = \sigma[-\Delta, \Omega_{0, j}] = \sigma_j, \quad j \in \{1, 2\}.$$

Thus, using the monotonicity of the principal eigenvalues with respect to the domains, it is apparent that

$$\sigma_0 < \lambda < \sigma[-\Delta, \Omega_{\delta, 1}] < \sigma_1 < \sigma[-\Delta, \Omega_{\delta, 2}] < \sigma_2 \quad (11)$$

if  $\delta > 0$  is sufficiently small. Pick up one of those  $\delta$ 's.



**Figure 3:** The  $\delta$ -neighborhoods  $\Omega_{\delta,1}$  and  $\Omega_{\delta,2}$ .

For each  $j \in \{1, 2\}$ , let  $\varphi_{\delta,j} \gg 0$  denote a principal eigenfunction associated with  $\sigma[-\Delta, \Omega_{\delta,j}]$ —unique up to a multiplicative constant—and consider the function  $\Phi$  defined through

$$\Phi := \begin{cases} \varphi_{\delta,1} & \text{in } \bar{\Omega}_{\frac{\delta}{2},1} \\ \varphi_{\delta,2} & \text{in } \bar{\Omega}_{\frac{\delta}{2},2} \\ \varphi_+ & \text{in } \Omega \setminus \left( \bar{\Omega}_{\frac{\delta}{2},1} \cup \bar{\Omega}_{\frac{\delta}{2},2} \right) \end{cases} \quad (12)$$

where  $\varphi_+$  is any smooth extension of  $\varphi_{\delta,1} \vee \varphi_{\delta,2}$  to

$$\Omega \setminus \left( \bar{\Omega}_{\frac{\delta}{2},1} \cup \bar{\Omega}_{\frac{\delta}{2},2} \right) = \left\{ x \in \Omega_+ : \text{dist}(x, \partial\Omega_+) > \frac{\delta}{2} \right\}$$

positive and bounded away from zero. Note that  $\varphi_+$  exists, since  $\varphi_{\delta,j}$  is positive and bounded away from zero on  $\Omega_+ \cap \partial\Omega_{\frac{\delta}{2},j}$ ,  $j \in \{1, 2\}$ . We claim that  $\kappa\Phi$  is a supersolution of (4) if  $\kappa$  is sufficiently large. Indeed,  $\kappa\Phi = 0$  on  $\partial\Omega$ . Moreover, for each  $j \in \{1, 2\}$ , we have that

$$-\Delta(\kappa\Phi) \geq \lambda\kappa\Phi - a(\kappa\Phi)^p \quad \text{in } \Omega_{\frac{\delta}{2},j}$$

if, and only if,

$$\kappa\sigma[-\Delta, \Omega_{\delta,j}]\varphi_{\delta,j} \geq \lambda\kappa\varphi_{\delta,j} - a(\kappa\varphi_{\delta,j})^p \quad \text{in } \Omega_{\frac{\delta}{2},j}$$

or, equivalently,

$$a(\kappa\varphi_{\delta,j})^{p-1} \geq \lambda - \sigma[-\Delta, \Omega_{\delta,j}] \quad \text{in } \Omega_{\frac{\delta}{2},j}$$

which is satisfied for every  $\kappa > 0$ , since, due to (11),

$$a(\kappa\varphi_{\delta,j})^{p-1} \geq 0 > \lambda - \sigma[-\Delta, \Omega_{\delta,j}] \quad \text{in } \Omega_{\delta,j}.$$

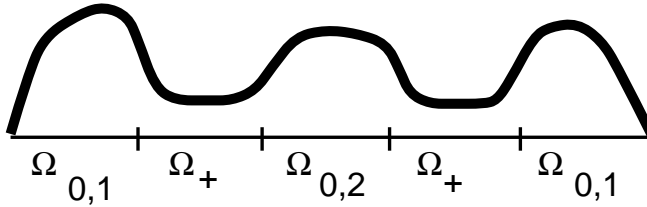
Further,

$$-\Delta(\kappa\Phi) \geq \lambda\kappa\Phi - a(\kappa\Phi)^p \quad \text{in } \Omega \setminus \left(\bar{\Omega}_{\frac{\delta}{2},1} \cup \bar{\Omega}_{\frac{\delta}{2},2}\right)$$

if, and only if,

$$\frac{-\Delta\varphi_+}{\varphi_+} \geq \lambda - a(\kappa\varphi_+)^{p-1} \quad \text{in } \Omega \setminus \left(\bar{\Omega}_{\frac{\delta}{2},1} \cup \bar{\Omega}_{\frac{\delta}{2},2}\right)$$

which is satisfied for all  $\kappa > 0$  sufficiently large, since  $a\varphi_+^{p-1}$  is positive and bounded away from zero in  $\Omega \setminus \left(\bar{\Omega}_{\frac{\delta}{2},1} \cup \bar{\Omega}_{\frac{\delta}{2},2}\right) \subset\subset \Omega_+$ . Figure 4 shows the sliced profile of  $\Phi$ . Note that  $\varepsilon\varphi \leq \kappa\Phi$  if  $\kappa$  is sufficiently large.



**Figure 4:** The sliced profile of the *supersolution element*  $\Phi$ .

Now, the details of the proofs of Parts (a), (b) can be easily completed by adapting the proof of Theorem 1. Our proof of Part (c) is based upon the following fundamental property

$$\lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda,a,\Omega]} = \infty \quad \text{uniformly in compact subsets of } \bar{\Omega}_{0,1} \setminus \partial\Omega. \quad (13)$$

To prove (13) one can argue as follows. Pick up  $\lambda_1 \in (\sigma_0, \sigma_1)$  and consider  $\eta > 0$  such that

$$\theta_{[\lambda_1,a,\Omega]} > \eta\varphi_{0,1} \quad \text{in } \Omega_{0,1}$$

where  $\varphi_{0,1}$  stands for the principal eigenfunction associated to  $\sigma_1 = \sigma[-\Delta, \Omega_{0,1}]$ . By differentiating (4) with respect to  $\lambda$ , particularizing the result at  $\theta_{[\lambda,a,\Omega]}$ , rearranging terms and setting

$$\Theta_\lambda := \frac{d}{d\lambda}\theta_{[\lambda,a,\Omega]}$$

gives

$$\begin{cases} \left(-\Delta + pa\theta_{[\lambda,a,\Omega]}^{p-1} - \lambda\right)\Theta_\lambda = \theta_{[\lambda,a,\Omega]} & \text{in } \Omega \\ \Theta_\lambda = 0 & \text{on } \partial\Omega \end{cases} \quad (14)$$

Since

$$\sigma[-\Delta + pa\theta_{[\lambda,a,\Omega]}^{p-1} - \lambda, \Omega] > \sigma[-\Delta + a\theta_{[\lambda,a,\Omega]}^{p-1} - \lambda, \Omega] = 0$$

applying the maximum principle to (14) gives  $\Theta_\lambda \gg 0$  and, hence, the mapping  $\lambda \mapsto \theta_{[\lambda,a,\Omega]}$  is increasing. In particular, for each  $\lambda \in [\lambda_1, \sigma_1]$ , we have that

$$\theta_{[\lambda,a,\Omega]} > \eta \varphi_{0,1} \quad \text{in } \Omega_{0,1}$$

and, so, (14) gives

$$\begin{cases} (-\Delta - \lambda) \Theta_\lambda > \eta \varphi_{0,1} & \text{in } \Omega_{0,1} \\ \Theta_\lambda > 0 & \text{on } \partial\Omega_{0,1} \end{cases} \quad (15)$$

because  $\Theta_\lambda|_{\Gamma_1} > 0$ . Moreover,

$$\sigma[-\Delta - \lambda, \Omega_{0,1}] = \sigma_1 - \lambda > 0$$

and, therefore, thanks to the maximum principle, we find from (15) that

$$\Theta_\lambda > \Psi_\lambda \quad \text{in } \Omega_{0,1} \quad (16)$$

where  $\Psi_\lambda$  is the unique solution of

$$\begin{cases} (-\Delta - \lambda) \Psi_\lambda = \eta \varphi_{0,1} & \text{in } \Omega_{0,1} \\ \Psi_\lambda = 0 & \text{on } \partial\Omega_{0,1} \end{cases} \quad (17)$$

A direct calculation shows that  $\Psi_\lambda = \frac{\eta}{\sigma_1 - \lambda} \varphi_{0,1}$  and, hence, (16) gives  $\lim_{\lambda \nearrow \sigma_1} \Theta_\lambda = \infty$ . In particular,  $\lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda,a,\Omega]} = \infty$  uniformly in compact subsets of  $\Omega_{0,1}$ . The completion of the proof of (13) can be done by using a device coming from [32], where we send for further details. To end the proof of Part (c) one can argue as follows. By the parabolic maximum principle, for each  $\lambda \geq \sigma_1$  and  $\varepsilon > 0$ , we have that

$$u_{[\lambda,a,\Omega]}(\cdot, t; u_0) > u_{[\sigma_1 - \varepsilon, a, \Omega]}(\cdot, t; u_0) \quad \text{in } \Omega.$$

Moreover, for any  $\varepsilon > 0$  sufficiently small,

$$\lim_{t \nearrow \infty} u_{[\sigma_1 - \varepsilon, a, \Omega]}(\cdot, t; u_0) = \theta_{[\sigma_1 - \varepsilon, a, \Omega]}$$

uniformly in  $\bar{\Omega}$ . Thus,

$$\liminf_{t \nearrow \infty} u_{[\lambda,a,\Omega]}(\cdot, t; u_0) \geq \theta_{[\sigma_1 - \varepsilon, a, \Omega]}$$

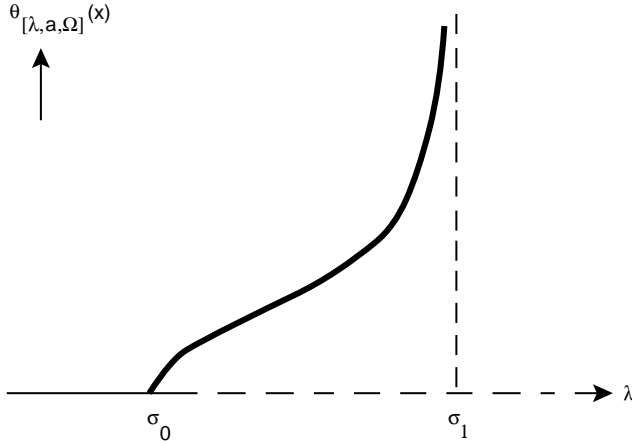
for each  $\varepsilon > 0$  sufficiently small. Passing to the limit as  $\varepsilon \downarrow 0$  in the previous inequality, (13) concludes the proof of Part (c).  $\square$

By using some further continuation techniques it is easy to see that the mapping  $\lambda \mapsto \theta_{[\lambda,a,\Omega]}$  is differentiable and increasing. Moreover,  $\theta_{[\lambda,a,\Omega]}$  bifurcates from

$u = 0$  at  $\lambda = \sigma_0$ . In Figure 5 we have taken  $x \in \Omega_{0,1}$  and represented  $\theta_{[\lambda,a,\Omega]}(x)$  versus  $\lambda \in (\sigma_0, \sigma_1)$ . Thanks to (13),

$$\lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda,a,\Omega]}(x) = \infty$$

and, hence, the diagram exhibits a bifurcation from infinity at  $\lambda = \sigma_1$ . In Figure 5, stable solutions are indicated by solid lines, and unstable solutions by dashed lines. The state  $u = 0$  loses stability as  $\lambda$  crosses  $\sigma_0$ , such stability being gained by  $\theta_{[\lambda,a,\Omega]}$ ; local stability actually entails global attractiveness.



**Figure 5:** The steady states of (1).

Subsequently, we consider the compact set  $K$  defined by

$$K := \{ x \in \Omega : \text{dist}(x, \partial\Omega) \geq \text{dist}(\partial\Omega, \Gamma_1)/2 \} .$$

Thanks to Harnack inequality, for each  $\lambda \in (\sigma_0, \sigma_1)$  a constant  $C_\lambda > 0$  exists such that

$$\max_K \theta_{[\lambda,a,\Omega]} \leq C_\lambda \min_K \theta_{[\lambda,a,\Omega]} . \quad (18)$$

Thus, if for some  $\varepsilon > 0$  and  $C > 0$  an estimate of the form

$$\sup_{\sigma_1 - \varepsilon < \lambda < \sigma_1} C_\lambda \leq C \quad (19)$$

would be satisfied, then, necessarily, (13) and (18) would imply that

$$\lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda,a,\Omega]} = \infty \quad \text{uniformly in } K \quad (20)$$

and, hence,  $\theta_{[\lambda,a,\Omega]}$ ,  $\sigma_0 < \lambda < \sigma_1$ , would exhibit entire blow up in  $K$  as  $\lambda \uparrow \sigma_1$ . Therefore, under condition (19) the diagram of positive solutions of (4) looks like shows Figure 5 for every  $x \in K$ , and, henceforth, for all  $x \in \Omega$ , because of (13). Since the length of the interval  $(\sigma_1 - \varepsilon, \sigma_1)$  is  $\varepsilon$  —arbitrarily small—, the

temptation of believing in the validity of (19) is extremely strong to be resisted. As a result, several P.D.E's experts asked about by the author during the period 1995–96 tried to convince him of the complete blow up of  $\theta_{[\lambda,a,\Omega]}$  as  $\lambda \uparrow \sigma_1$ . Actually, after spending one week in the Canary Islands during September 1996 in order to take part in the Master Degree Committee of Professor J. C. García-Melián —an student of Professor J. C. Sabina at that time—, the author received from professor J. C. Sabina a fax containing the proof of (20) in the radially symmetric case. Experts in history of mathematics might wish to know that the fax was sent to the Dean Office and that it was changing of hands during one week until it reached the author. As a result, when the author finally got the faxed manuscript several distinguished members of the Faculty already new Sabina's theorem. Unfortunately, it was wrong. The reader should not jock too much, as Professor J. C. Sabina is one of the very best experts in P.D.E's that we ever had in Spain, and... at the end of the day no body is free from making mistakes; the only way of avoiding mistakes is doing nothing. Anyway... after receiving Sabina's fax, the author was able to prove that the unique positive solution of the one dimensional problem

$$\begin{cases} -u'' = \lambda u - au^p & \text{in } J := (-R, R) \\ u(-R) = u(R) = M \end{cases}$$

where  $a > 0$  is fixed and  $M > 0$  is sufficiently large, subsequently denoted by  $\theta_{[\lambda,a,J,M]}$ , approximates to a fixed profile in  $J$ , say  $L_{[\lambda,a,J]}$ , as  $M \uparrow \infty$ . In the specialized literature, such a solution  $L_{[\lambda,a,J]}$  is referred to as a *large solution* of  $-u'' = \lambda u - au^p$  in  $J$ . Although, some time later we learned that a huge industry on large solutions was available in the literature (e.g., C. Bandle & M. Marcus [4] and M. Marcus & L Véron [33]), it should be remarked that the a priori bounds of J. B. Keller [21] and R. Osserman [34] do not apply straight away to (4), and that, occasionally, they have been used in the wrong way (cf. [25] for further details). Undoubtedly, a genuine mathematical “Bisociation” in the sense of F. Browder [6]. Bisociation... a word coined by the political and scientific writer A. Koestler for designating unexpected —sharply hidden— connections between *a priori* unrelated separated fields; the motor of Science.

Following with the previous historical remarks, when the right result together with a right proof were electronically sent to the Canary Islands, Professor J. C. Sabina and his student J. García-Melián were able to generalize the author's result to the radially symmetric case and could produce a first draft of the pioneering paper [16]. The final draft was prepared in Madrid, where, among other things, the 1-dimensional numerical experiments of R. Gómez-Reñasco were included; these computations confirmed and complemented the analytical results.

Summarizing, it turns out that (20) fails to be true and, therefore, (19) is false. Actually, the following result is satisfied. Subsequently, we set

$$\tilde{\Omega} := \Omega_+ \cup \bar{\Omega}_{0,2}. \quad (21)$$

**Theorem 3** *The point-wise limit*

$$L_{[\sigma_1, a, \tilde{\Omega}]}(x) := \lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda, a, \Omega]}(x) \quad (22)$$

is well defined for each  $x \in \tilde{\Omega}$ , and it provides us with the minimal positive solution of

$$\begin{cases} -\Delta u = \sigma_1 u - a(x)u^p & \text{in } \tilde{\Omega} \\ u = \infty & \text{on } \partial\tilde{\Omega} = \Gamma_1 \end{cases} \quad (23)$$

subsequently denoted by  $\mathfrak{L}_{[\sigma_1, a, \tilde{\Omega}]}^{\min}$ .

Note that Theorem 3 was not covered in [16], where  $\Omega_0$  was assumed to be connected. Although not extremely involved, the changes in the proof are far from straightforward, because of the presence of a second vanishing component of  $a$ ,  $\Omega_{0,2}$ . Those degenerated cases were covered by R. Gómez-Reñasco in [18]—under author’s supervision— and in [20] and [25].

To illustrate how behave  $\theta_{[\lambda, a, \Omega]}$  as  $\lambda$  grows approaching  $\sigma_1$ , we shall consider a model coming from [20]. Suppose  $N = 2$ ,  $p = 5$ ,

$$\Omega = B_{0.5}, \quad \Omega_{0,1} = A_{(0.3, 0.5)}, \quad \Omega_+ = A_{(0.1, 0.3)}, \quad \Omega_{0,2} = B_{0.1},$$

where, for any  $0 < \alpha < \beta$ ,  $A_{(\alpha, \beta)}$  and  $B_\alpha$  stand for

$$A_{(\alpha, \beta)} := \{x \in \mathbb{R}^2 : \alpha < |x| < \beta\}, \quad B_\alpha := \{x \in \mathbb{R}^2 : |x| < \alpha\},$$

respectively, and

$$a(x) := \begin{cases} -\sin(5\pi(|x| + 0.5)), & x \in \Omega_+, \\ 0, & x \in \Omega \setminus \Omega_+. \end{cases}$$

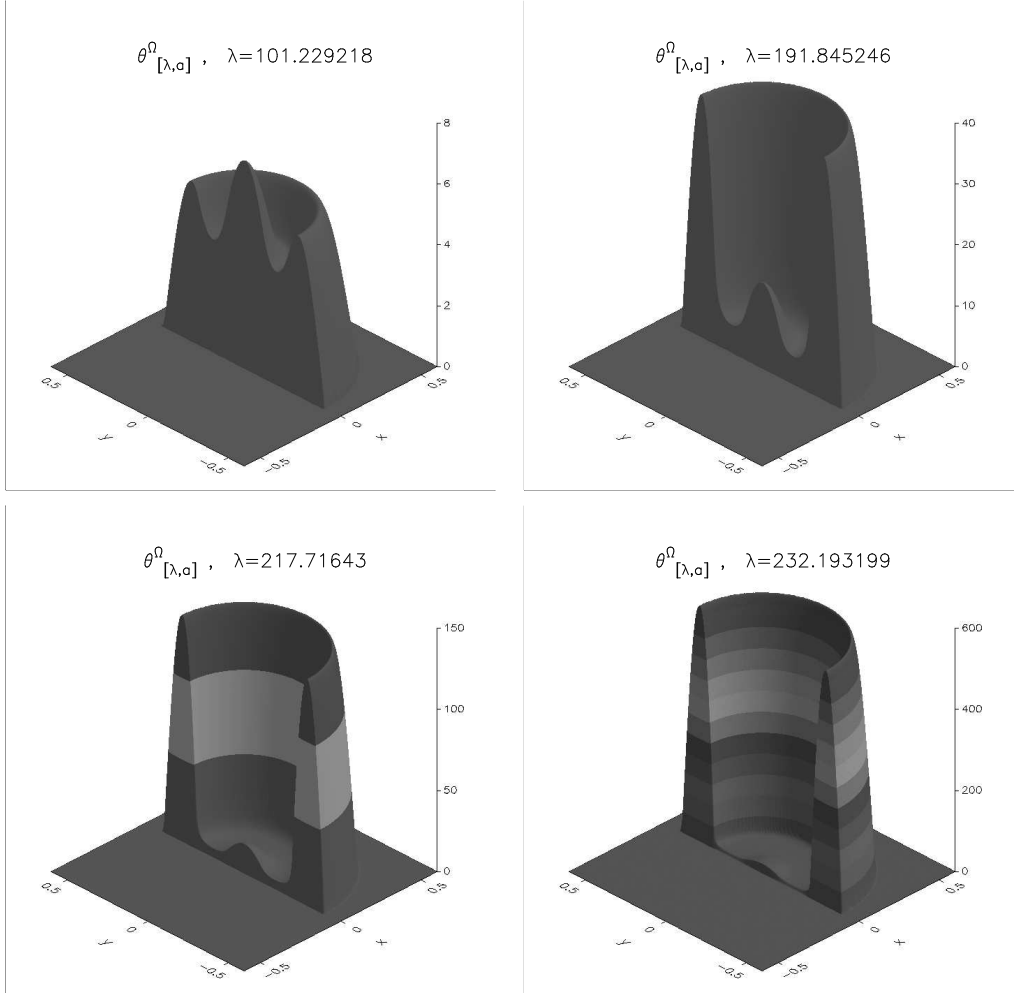
For this choice, one has that

$$\sigma_0 \sim 23.13 < \sigma_1 \sim 245.138590 < \sigma_2 \sim 578.306304$$

(cf. [20, Section 1]) and, hence, all our conditions here are satisfied. Figure 6 shows the profiles of some of the positive solutions  $\theta_{[\lambda, a]}^\Omega := \theta_{[\lambda, a, \Omega]}$  computed in [20] coupling a pseudo-spectral-collocation method with path-following. Note how these solutions grow in  $\Omega_{0,1} \setminus \partial\Omega$ , while they stabilize in  $\Omega \setminus \bar{\Omega}_{0,1}$  as  $\lambda$  grows



approaching  $\sigma_1$ .



**Figure 6:** Profiles of some steady states of (1).

**Sketch of the proof of Theorem 3:** For each  $\delta > 0$  sufficiently small consider the open set

$$\tilde{\Omega}_\delta := \{x \in \tilde{\Omega} : \text{dist}(x, \partial\tilde{\Omega}) > \delta\}.$$

If  $\delta < \text{dist}(\Gamma_1, \Gamma_2)$ , then  $\Gamma_{1,\delta} := \partial\tilde{\Omega}_\delta$  is a compact subset of  $\Omega_+$  and, hence, it can be covered by a finite number of balls contained in  $\Omega_+$ . Moreover, in each of these balls  $a$  is bounded away from zero by a positive constant. Combining this features with [16, Lemma 6.2] it is easily seen that there exists a constant  $M_\delta > 0$  such that

$$\theta_{[\lambda, a, \Omega]}(x) \leq M_\delta \quad \text{for each } (\lambda, x) \in (\sigma_0, \sigma_1) \times \Gamma_{1,\delta}. \quad (24)$$

By the reader's convenience, we recall [16, Lemma 6.2]; it reads as follows:

**Lemma 4** *Let  $B_R(x_0) \subset \mathbb{R}^N$  denote the ball of radius  $R$  centered at  $x_0 \in \mathbb{R}^N$  and consider the boundary value problem*

$$\begin{cases} -\Delta u = \lambda u - Au^p & \text{in } B_R(x_0) \\ u = M & \text{on } \partial B_R(x_0) \end{cases} \quad (25)$$

where  $\lambda > 0$ ,  $A > 0$  and  $M > \left(\frac{\lambda}{A}\right)^{\frac{1}{p-1}}$  are constants. Then, (25) possesses a unique positive solution  $u_M \in C^\infty(\bar{B}_R(x_0))$ , the mapping  $M \mapsto u_M$  is pointwise increasing in  $\bar{B}_R(x_0)$  and

$$u_\infty(x) := \lim_{M \nearrow \infty} u_M(x) < \infty \quad \text{for each } x \in B_R(x_0).$$

Moreover,  $u_\infty$  provides us with the minimal classical solution —large solution— of

$$\begin{cases} -\Delta u = \lambda u - Au^p & \text{in } B_R(x_0) \\ u = \infty & \text{on } \partial B_R(x_0) \end{cases} \quad (26)$$

Now, for each  $\delta > 0$  sufficiently small and  $M \geq M_\delta$ , we consider the problem

$$\begin{cases} -\Delta u = \lambda u - au^p & \text{in } \tilde{\Omega}_\delta \\ u = M & \text{on } \Gamma_{1,\delta} = \partial\tilde{\Omega}_\delta \end{cases} \quad (27)$$

Adapting the proof of Theorem 2 it is easy to see that (27) has a positive solution if, and only if,  $\lambda < \sigma_2$ , and that it is unique if it exists. Moreover, if we denote it by  $\theta_{[\lambda,a,\tilde{\Omega}_\delta,M]}$ , then the mapping  $\lambda \mapsto \theta_{[\lambda,a,\tilde{\Omega}_\delta,M]}$  is increasing. Indeed, if we suppose (27) to have a positive solution, say  $\theta$ , then  $\theta$  provides us with a positive strict supersolution of  $-\Delta + a\theta^{p-1} - \lambda$  (under Dirichlet boundary conditions) in  $\tilde{\Omega}_\delta$ . Hence, thanks to [24, Theorem 2.5],

$$\sigma[-\Delta + a\theta^{p-1} - \lambda, \tilde{\Omega}_\delta] > 0.$$

Thus,

$$\lambda < \sigma[-\Delta + a\theta^{p-1}, \tilde{\Omega}_\delta] < \sigma[-\Delta, \Omega_{0,2}] = \sigma_2,$$

since  $\Omega_{0,2} \subset \tilde{\Omega}_\delta$ , and, hence,  $\lambda < \sigma_2$  is necessary for the existence of a positive solution. To prove the converse, fix  $\lambda < \sigma_2$  and consider, for any  $\eta > 0$  sufficiently small,

$$\Phi := \begin{cases} \varphi_{\eta,2} & \text{in } \bar{\Omega}_{\frac{\eta}{2},2} \\ \varphi_+ & \text{in } \tilde{\Omega}_\delta \setminus \Omega_{\frac{\eta}{2},2} \end{cases}$$

where the notations introduced in the proof of Theorem 2 are kept and  $\varphi_+$  is any smooth extension of  $\varphi_{\eta,2}$  to  $\tilde{\Omega}_\delta \setminus \Omega_{\frac{\eta}{2},2}$  bounded away from zero and such

that  $\varphi_+ \geq M$  on  $\Gamma_{1,\delta} = \partial\tilde{\Omega}_\delta$ . As in the proof of Theorem 2,  $\bar{u} := \kappa\Phi$  is a supersolution of (27) if  $\kappa$  is sufficiently large. Moreover,  $\underline{u} := 0$  provides us with a subsolution of (27) such that  $\underline{u} \leq \bar{u}$ . Therefore, the existence of a positive solution follows from H. Amann [1], [2]. The remaining assertions follow from the strong maximum principle arguing as in the proof of Theorem 2. Actually, the proof of (13) can be easily adapted to show that

$$\lim_{\lambda \nearrow \sigma_2} \theta_{[\lambda, a, \tilde{\Omega}_\delta, M]} = \infty \quad \text{uniformly in compact subsets of } \bar{\tilde{\Omega}}_{0,2}. \quad (28)$$

Relation (28) explains why those positive solutions do not exist for  $\lambda \geq \sigma_2$ . Note that (27) admits arbitrarily large supersolutions everywhere in  $\tilde{\Omega}_\delta$  if  $\lambda < \sigma_2$ , since  $\kappa$  can be taken arbitrarily large. Moreover, thanks to (24), for each  $\lambda \in (\sigma_0, \sigma_1)$  and  $M \geq M_\delta$ ,  $\theta_{[\lambda, a, \Omega]}$  provides us with a positive subsolution of (27). Thus, for each  $\lambda \in (\sigma_0, \sigma_1)$ , we have that

$$\theta_{[\lambda, a, \Omega]} \leq \theta_{[\lambda, a, \tilde{\Omega}_\delta, M_\delta]} \leq \theta_{[\frac{\sigma_1 + \sigma_2}{2}, a, \tilde{\Omega}_\delta, M_\delta]} \quad \text{in } \tilde{\Omega}_\delta \quad (29)$$

and, therefore, for each  $x \in \tilde{\Omega}_\delta$ , the point-wise limit (22) is well defined. As this is true for any  $\delta > 0$  sufficiently small, (22) is well defined in  $\tilde{\Omega}$ . The fact that it is a solution of (23) follows by a rather standard argument; so, we will omit the details. The minimality is an easy consequence from (22).  $\square$

Summarizing, thanks to Theorems 2, 3, it is apparent that

$$M_{[\sigma_1, a, \Omega]} := \lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda, a, \Omega]} = \begin{cases} \infty & \text{in } \bar{\tilde{\Omega}}_{0,1} \setminus \partial\tilde{\Omega} \\ \mathfrak{L}_{[\sigma_1, a, \tilde{\Omega}]}^{\min} & \text{in } \tilde{\Omega} \end{cases}$$

Now, one is naturally tempted to propose as the limiting profile of  $u_{[\sigma_1, a, \Omega]}(\cdot, t; u_0)$  as  $t \uparrow \infty$  the *meta-generalized solution* —metasolution— of (4) given by the limiting profile  $M_{[\sigma_1, a, \Omega]}$ , at least for  $u_0$  small. As it can be easily realized from the proof of Theorem 3, these *metasolutions* actually exist for each  $\lambda < \sigma_2$ . Actually, they provide us with the limiting behavior of all solutions of (1) for each  $\lambda \in [\sigma_1, \sigma_2)$  and  $u_0$  sufficiently small.

In this section we have shortly described the genesis of the first metasolution found in the literature. In the next section we shall fix some crucial related concepts for describing the dynamics of (1).

## 2 Large solutions and metasolutions

In this section we collect the precise concepts of *large solution* and *metasolution* and give their main existence results. Most of the results go back to [18], [20] and [25], where we sent the interested reader for all necessary details, included the proofs of the theorems, which might be easily completed from the basic ingredients introduced in Section 1. Throughout the remaining of this paper,

the dependence on  $a(x)$  of the several solutions, generalized solutions and meta-generalized solutions will be dropped, unless emphasizing it be strictly necessary. In particular, we will set

$$\theta_{[\lambda, \Omega]} := \theta_{[\lambda, a, \Omega]}.$$

## 2.1 Large solutions

Given a subdomain  $D \subset \Omega$ , the classical positive solutions of the singular problem

$$\begin{cases} -\Delta u = \lambda u - au^p & \text{in } D \\ u = \infty & \text{on } \partial D \end{cases} \quad (1)$$

will be referred to as *large solutions* in  $D$  of

$$-\Delta u = \lambda u - au^p. \quad (2)$$

A smooth function  $u$  is said to be a solution of (1) if it solves (2) in  $D$  and

$$\lim_{\substack{x \in D \\ \text{dist}(x, \partial D) \searrow 0}} u(x) = \infty.$$

The following result, characterizes the existence of large solutions of (2) in  $D \in \{\tilde{\Omega}, \Omega_+\}$ .

**Theorem 5** *The following assertions are true:*

- ( $\tilde{\Omega}$ ) *Equation (2) has a large solution in  $\tilde{\Omega}$  if, and only if,  $\lambda < \sigma_2$ . Moreover, in such case, there is a minimal, and a maximal, large solution. The minimal (resp. maximal) large solution will be throughout denoted by  $\mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$  (resp.  $\mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\max}$ ).*
- ( $\Omega_+$ ) *For each  $\lambda \in \mathbb{R}$  the equation (2) has a large solution in  $\Omega_+$ . Actually, it has a minimal, and a maximal, large solution; throughout denoted by  $\mathfrak{L}_{[\lambda, \Omega_+]}^{\min}$  and  $\mathfrak{L}_{[\lambda, \Omega_+]}^{\max}$ , respectively.*

Subsequently, for any constant  $M > 0$ , any subdomain  $D \subset \Omega$  of class  $\mathcal{C}^3$ , and  $u_0 \in \mathcal{C}(\bar{D})$ , we denote by  $u_{[\lambda, D, M]}(x, t; u_0)$  the unique solution of the evolutionary problem

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = \lambda u - au^p & \text{in } D \times (0, \infty) \\ u = M & \text{on } \partial D \times (0, \infty) \\ u(\cdot, 0) = u_0 \geq 0 & \text{in } D \end{cases} \quad (3)$$

Since  $a \geq 0$ ,  $u_{[\lambda, D, M]}(x, t; u_0)$  is globally defined in time. Moreover, by parabolic regularity is smooth. Also, we shall consider the associated elliptic problem

$$\begin{cases} -\Delta u = \lambda u - au^p & \text{in } D \\ u = M & \text{on } \partial D \end{cases} \quad (4)$$

The following result is the key to prove Theorem 5.

**Theorem 6** *The following assertions are true:*

( $\tilde{\Omega}$ ) *Suppose  $D = \tilde{\Omega}$ . Then, (4) possesses a positive solution if, and only if,  $\lambda < \sigma_2$ , and it is unique if it exists. It will be throughout denoted by  $\theta_{[\lambda, \tilde{\Omega}, M]}$ . Moreover, for each  $\lambda < \sigma_2$ ,*

(a)  $\mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min} := \lim_{M \nearrow \infty} \theta_{[\lambda, \tilde{\Omega}, M]}$ , *where  $\mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$  is the large solution given by Theorem 5.*

(b) *For each  $u_0 > 0$ ,  $\lim_{t \nearrow \infty} u_{[\lambda, \tilde{\Omega}, M]}(\cdot, t; u_0) = \theta_{[\lambda, \tilde{\Omega}, M]}$  uniformly in  $\tilde{\Omega}$ .*

( $\Omega_+$ ) *Suppose  $D = \Omega_+$ . Then, for each  $\lambda \in \mathbb{R}$ , (4) possesses a unique positive solution; subsequently denoted by  $\theta_{[\lambda, \Omega_+, M]}$ . Moreover, the following properties are satisfied:*

(a)  $\mathfrak{L}_{[\lambda, \Omega_+]}^{\min} := \lim_{M \nearrow \infty} \theta_{[\lambda, \Omega_+, M]}$ , *where  $\mathfrak{L}_{[\lambda, \Omega_+]}^{\min}$  is the solution given by Theorem 5.*

(b) *For each  $u_0 > 0$ ,  $\lim_{t \nearrow \infty} u_{[\lambda, \Omega_+, M]}(\cdot, t; u_0) = \theta_{[\lambda, \Omega_+, M]}$  uniformly in  $\Omega_+$ .*

*Moreover, for each  $D \in \{\tilde{\Omega}, \Omega_+\}$ , the mapping  $M \mapsto \theta_{[\lambda, D, M]}$  is increasing, and, if  $\underline{u}$  (resp.  $\bar{u}$ ) is a subsolution (resp. supersolution) of (4), then  $\underline{u} \leq \theta_{[\lambda, D, M]}$  (resp.  $\bar{u} \geq \theta_{[\lambda, D, M]}$ ).*

**Remark 1** Theorem 6 is also true if  $\tilde{\Omega}$  and  $\Omega_+$  are we inter-exchanged by their respective  $\delta$ -neighborhoods

$$\tilde{\Omega}_\delta := \{x \in \tilde{\Omega} : \text{dist}(x, \partial\tilde{\Omega}) > \delta\} \quad \text{and} \quad \Omega_+^\delta := \{x \in \Omega_+ : \text{dist}(x, \partial\Omega_+) > \delta\}$$

provided  $\delta > 0$  is sufficiently small.

## 2.2 Metasolutions

Given  $D \in \{\tilde{\Omega}, \Omega_+\}$ , a function  $\mathfrak{M}_{[\lambda, D]} : \Omega \rightarrow [0, \infty]$  is said to be a *metasolution* of (2) —or (4)— *supported in  $D$*  if there exists a large solution of (2) in  $D$ , say  $\mathfrak{L}_{[\lambda, D]}$ , for which

$$\mathfrak{M}_{[\lambda, D]} = \begin{cases} \infty & \text{in } \Omega \setminus \bar{D} \\ \mathfrak{L}_{[\lambda, D]} & \text{in } D \end{cases}$$

In other words, metasolutions are extensions by  $\infty$  to the whole of  $\Omega$  of large solutions in  $D$ . As a consequence from the results of large solutions, the next result is obtained.

**Theorem 7** *The following assertions are true:*

( $\tilde{\Omega}$ ) *The equation (2) possesses a metasolution supported in  $\tilde{\Omega}$  if, and only if,  $\lambda < \sigma_2$ . Moreover, if this is the case, then there are a minimal and a maximal metasolution supported in  $\tilde{\Omega}$ ; throughout denoted by  $\mathfrak{M}_{[\lambda, \tilde{\Omega}]}^{\min}$  and  $\mathfrak{M}_{[\lambda, \tilde{\Omega}]}^{\max}$ , respectively.*

( $\Omega_+$ ) *For each  $\lambda \in \mathbb{R}$ , the equation (2) possesses a metasolution supported in  $\Omega_+$ . Actually, there are a minimal and a maximal metasolution supported in  $\Omega_+$ ; throughout denoted by  $\mathfrak{M}_{[\lambda, \Omega_+]}^{\min}$  and  $\mathfrak{M}_{[\lambda, \Omega_+]}^{\max}$ , respectively.*

Furthermore, the following fundamental relations are satisfied

$$\lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda, \Omega]} = \mathfrak{M}_{[\sigma_1, \tilde{\Omega}]}^{\min}, \quad \lim_{\lambda \nearrow \sigma_2} \mathfrak{M}_{[\lambda, \tilde{\Omega}]}^{\min} = \mathfrak{M}_{[\sigma_2, \Omega_+]}^{\min}. \quad (5)$$

The first limit is uniform in compact subsets of  $\bar{\Omega}_{0,1} \setminus \partial\Omega$  and  $\tilde{\Omega}$ . The second limit is uniform in compact subsets of  $\Omega \setminus \Omega_+$  and  $\Omega_+$ .

Figure 7 shows the plots of some of the minimal large solutions of (2) in  $\Omega_1 := \tilde{\Omega} = B_{0,3}$  for the special choice for which the positive solutions of Figure 6 were computed. In our setting,  $\Xi_{[\lambda, a]}^{\Omega_1} := \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$ . In [20, Section 6] we explained how to compute all those large solutions. It should be noted how  $\mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$  stabilizes in  $\Omega_+$  as  $\lambda \uparrow \sigma_2$ , whereas it blows up in  $\bar{\Omega}_{0,2} = B_{0,1}$ .

Finally, note that, for each  $D \in \{\tilde{\Omega}, \Omega_+\}$ , we have

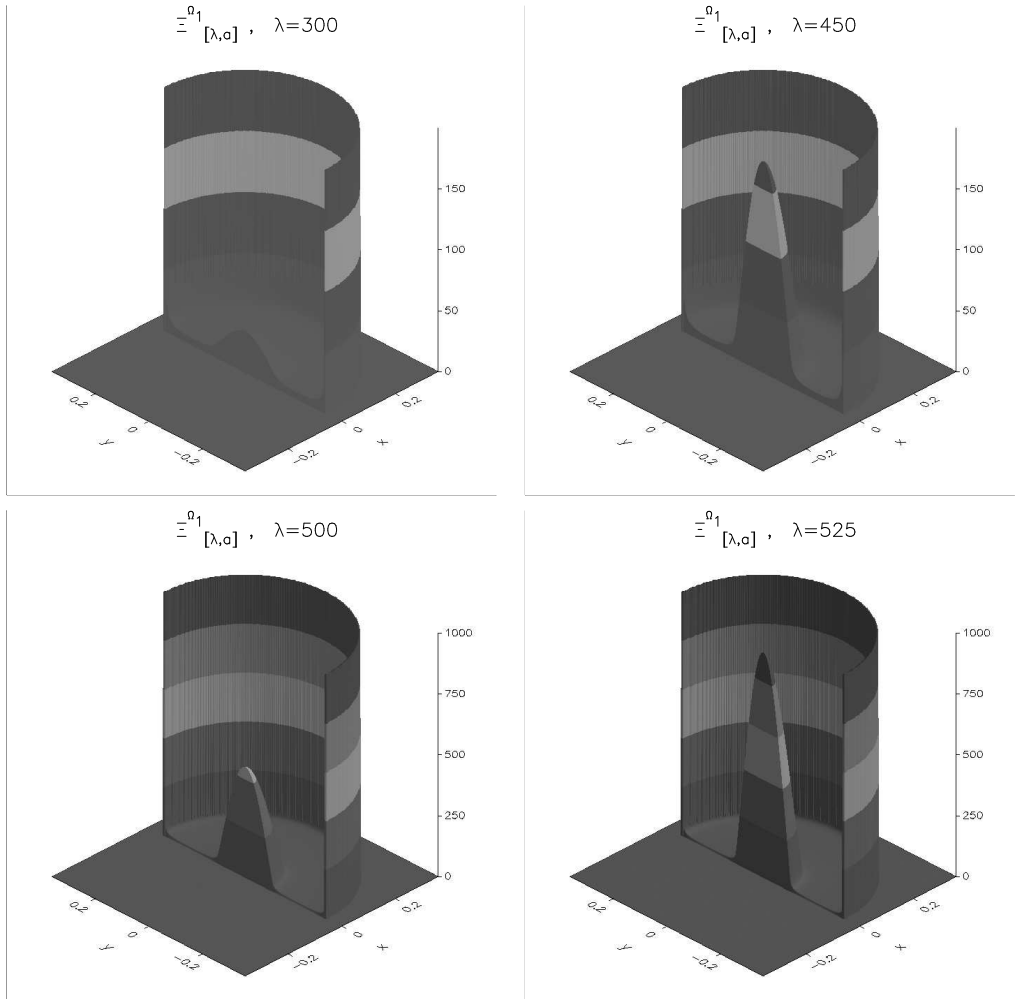
$$\mathfrak{M}_{[\lambda, D]}^{\max(\min)} = \begin{cases} \infty & \text{in } \Omega \setminus \bar{D} \\ \mathfrak{L}_{[\lambda, D]}^{\max(\min)} & \text{in } D \end{cases} \quad (6)$$

and, henceforth, the first relationship of (5) reads as follows

$$\lim_{\lambda \nearrow \sigma_1} \theta_{[\lambda, \Omega]} = \begin{cases} \infty & \text{in } \bar{\Omega}_{0,1} \setminus \partial\Omega \\ \mathfrak{L}_{[\sigma_1, \tilde{\Omega}]}^{\min} & \text{in } \tilde{\Omega} = \Omega \setminus \bar{\Omega}_{0,1} \end{cases} \quad (7)$$

whereas the second one can be stated under the form

$$\lim_{\lambda \nearrow \sigma_2} \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min} = \begin{cases} \infty & \text{in } \bar{\Omega}_{0,2} \\ \mathfrak{L}_{[\sigma_2, \Omega_+]}^{\min} & \text{in } \Omega_+ \end{cases} \quad (8)$$



**Figure 7:** Profiles of some large solutions supported in  $\tilde{\Omega}$ .

### 3 Characterizing the dynamics of (1)

Subsequently, the notations introduced in Section 2 will be kept. The following result completely describes the dynamics of (1) for  $\lambda \geq \sigma_1$ .

**Theorem 8** *Suppose  $u \in \mathcal{C}(\bar{\Omega})$ ,  $u_0 > 0$ . Then:*

1.– *In case  $\sigma_1 \leq \lambda < \sigma_2$ , the following assertions are true:*

- (a)  $\lim_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \infty$  uniformly on compact subsets of  $\bar{\Omega}_{0,1} \setminus \partial\Omega$ .

$$(b) \quad \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min} \leq \liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\max} \quad \text{in } \tilde{\Omega}.$$

(c) When  $u_0$  is a subsolution of (4), then  $\lim_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$  uniformly on compact subsets of  $\tilde{\Omega}$ .

2.– In case  $\lambda \geq \sigma_2$ , the following assertions are true:

$$(a) \quad \lim_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \infty \quad \text{uniformly on compact subsets of } \Omega \setminus \Omega_+.$$

$$(b) \quad \mathfrak{L}_{[\lambda, \Omega_+]}^{\min} \leq \liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \mathfrak{L}_{[\lambda, \Omega_+]}^{\max} \quad \text{in } \Omega_+.$$

(c) When  $u_0$  is a subsolution of (4), then  $\lim_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \mathfrak{L}_{[\lambda, \Omega_+]}^{\min}$  uniformly on compact subsets of  $\Omega_+$ .

Theorem 8 complements to Theorem 2 by establishing that, if  $\sigma_1 \leq \lambda < \sigma_2$ , then the limiting behavior of the solutions of (1) is governed by the metasolutions of (4) supported in  $\tilde{\Omega}$ , whereas it is governed by the metasolutions of (4) supported in  $\Omega_+$  if  $\lambda \geq \sigma_2$ .

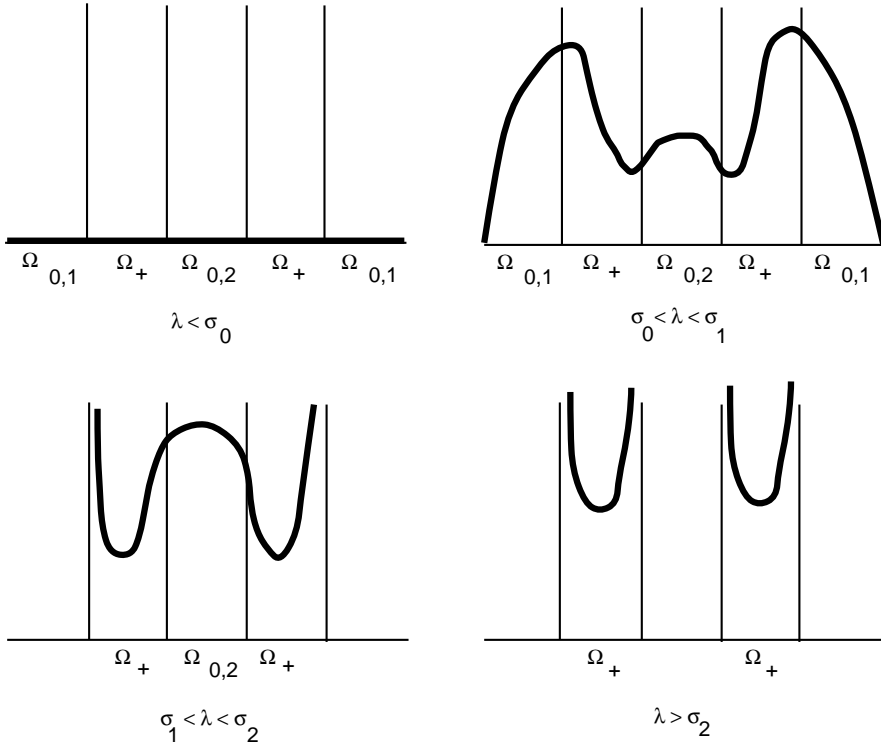
Parts 1(a), 2(a) of Theorem 8 have been already shown in [25], and Parts 1(b), 2(b) were stated in [25, Theorem 6.2], although the proof of [25, Theorem 6.2] contains a gap, since the comparison carried out in between formulas (6.4) and (6.5) is not necessarily true! Parts 1(c), 2(c) of Theorem 8 are completely new, even in the simplest case when  $\Omega \setminus \tilde{\Omega}_+$  is connected (cf. [13, Theorem 1.3]).

In general, large solutions, and, so, metasolutions, supported in  $\tilde{\Omega}$  and  $\Omega_+$  are not necessarily unique. As a consequence, the results of Parts 1(c), 2(c) are not necessarily true when  $u_0$  fails to be a subsolution of (4). Some uniqueness results when  $\Omega \setminus \tilde{\Omega}_+$  is connected and

$$a(x) = \beta [\text{dist}(x, \partial\Omega)]^\gamma [1 + \rho \text{dist}(x, \partial\Omega) + o(\text{dist}(x, \partial\Omega))] \quad \text{as } \text{dist}(x, \partial\Omega) \searrow 0$$

for some constants  $\beta > 0$ ,  $\gamma \geq 0$ , and  $\rho \in \mathbb{R}$ , were given by Y. Du and Q. Huang [13] and J. García-Melián et al. [17]. These results were substantially extended in [30], where  $\beta > 0$ ,  $\gamma \geq 0$  and  $\rho$  were allowed to be arbitrary continuous functions.





**Figure 8:** The limiting profiles of the solutions of (1).

In Figure 2 we have represented all possible limiting profiles according to each of the possible ranges of variation of  $\lambda$ . In all cases we have represented a one-dimensional slice of the limiting profile. To discuss the diagram we will assume that either the corresponding metasolution is unique or else  $u_0$  is a subsolution. In the last case, the next discussion should refer to minimal metasolutions, instead of to metasolutions. When  $\lambda \leq \sigma_0$ , all solutions approach zero. When  $\sigma_0 < \lambda < \sigma_1$  all solutions approach the unique positive steady state. As  $\lambda \uparrow \sigma_1$ , the steady states grow to infinity in  $\Omega_{0,1}$  while they stabilize in  $\tilde{\Omega} = \Omega_+ \cup \Omega_{0,2}$ ; the limiting steady state being the minimal metasolution of (4) supported in  $\tilde{\Omega}$  for  $\lambda = \sigma_1$ . Quite naturally, this metasolution provides us with the limiting profile as  $t \uparrow \infty$  of the solutions of (1) at  $\lambda = \sigma_1$ . Actually, the metasolutions of (4) supported in  $\tilde{\Omega}$  provide us with the limiting profiles of the solutions of (1) for every  $\sigma_1 \leq \lambda < \sigma_2$ . As  $\lambda \uparrow \sigma_2$  the metasolutions of (4) supported in  $\tilde{\Omega}$  grow to infinity in  $\Omega_{0,2}$  while they stabilize to a bounded profile in  $\Omega_+$ ; the limiting generalized solution being the metasolution supported in  $\Omega_+$  for  $\lambda = \sigma_2$ . Such metasolution is the limiting profile of all positive solutions of (1) as  $t \uparrow \infty$ . Actually, the metasolutions supported in  $\Omega_+$  provide us with the asymptotic behaviors of all positive solutions of (1) for any  $\lambda \geq \sigma_2$ .

**Proof of Theorem 8:** Let  $u_0 \in \mathcal{C}(\bar{\Omega})$  with  $u_0 > 0$ , and suppose

$$\sigma_1 \leq \lambda < \sigma_2. \quad (1)$$

Thanks to the parabolic maximum principle, for each  $\varepsilon > 0$  and  $t \geq 0$ ,  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq u_{[\sigma_1 - \varepsilon, \Omega]}(\cdot, t; u_0)$ , since  $\lambda > \sigma_1 - \varepsilon$ . Thus, thanks to Theorem 2(a),

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \lim_{t \nearrow \infty} u_{[\sigma_1 - \varepsilon, \Omega]}(\cdot, t; u_0) = \theta_{[\sigma_1 - \varepsilon, \Omega]}. \quad (2)$$

As (2) is valid for each  $\varepsilon > 0$  sufficiently small, using (5) it is apparent that

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \lim_{\varepsilon \searrow 0} \theta_{[\sigma_1 - \varepsilon, \Omega]} = \mathfrak{M}_{[\sigma_1, \bar{\Omega}]}^{\min}. \quad (3)$$

In particular,

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \infty \quad \text{in } \bar{\Omega}_{0,1} \setminus \partial\Omega = \Omega \setminus \tilde{\Omega} \quad (4)$$

uniformly on compact subsets, and the limit is uniform on the component  $\Gamma_1 (= \partial\tilde{\Omega})$  of  $\partial\Omega_+$ . So, for each  $M > 0$  a constant  $T_M > 0$  exists such that  $u_{[\lambda, \Omega]}(x, t; u_0) \geq M$  for each  $(x, t) \in \partial\tilde{\Omega} \times [T_M, \infty)$  and, hence,  $u_{[\lambda, \Omega]}(\cdot, t; u_0)$  provides us with a supersolution of the problem

$$\begin{cases} \frac{\partial v}{\partial t} - \Delta v = \lambda v - av^p & \text{in } \tilde{\Omega} \times (T_M, \infty) \\ v = M & \text{on } \partial\tilde{\Omega} \times (T_M, \infty) \\ v(\cdot, 0) = u_{[\lambda, \Omega]}(\cdot, T_M; u_0) & \text{in } \tilde{\Omega} \end{cases} \quad (5)$$

Thus, by the parabolic maximum principle,

$$u_{[\lambda, \Omega]}(x, t; u_0) \geq u_{[\lambda, \bar{\Omega}, M]}(x, t - T_M; u_{[\lambda, \Omega]}(\cdot, T_M; u_0))$$

for each  $(x, t) \in \tilde{\Omega} \times (T_M, \infty)$  (cf. (3)). So, thanks to Theorem 6( $\tilde{\Omega}$ )(b), we have that

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \lim_{t \nearrow \infty} u_{[\lambda, \bar{\Omega}, M]}(\cdot, t - T_M; u_{[\lambda, \Omega]}(\cdot, T_M; u_0)) = \theta_{[\lambda, \bar{\Omega}, M]}$$

in  $\tilde{\Omega}$ . Therefore, passing to the limit as  $M \uparrow \infty$ , Theorem 6( $\tilde{\Omega}$ )(a) gives

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \mathfrak{L}_{[\lambda, \bar{\Omega}]}^{\min} \quad \text{in } \tilde{\Omega}.$$

Now, combining this estimate with (4), it is easily realized that

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \mathfrak{M}_{[\lambda, \bar{\Omega}]}^{\min} \quad \text{in } \Omega. \quad (6)$$

It should be noted that (6) is independent of  $u_0$ , and that it provides us with Theorem 8 1(a) and the lower estimate of Theorem 8 1(b).

Suppose, in addition, that  $u_0$  is a subsolution of (4). Then, for each  $t > 0$ , the frozen function  $u_t(x) := u_{[\lambda, \Omega]}(x, t; u_0)$ ,  $x \in \tilde{\Omega}$ , is a subsolution of (4) (cf. [36]), because the mapping  $t \mapsto u_{[\lambda, \Omega]}(x, t; u_0)$  is non-decreasing. Now, fix  $t > 0$  and set

$$M_t := \max_{\tilde{\Omega}} u_{[\lambda, \Omega]}(\cdot, t; u_0).$$

Then, due to last assertion of Theorem 6,  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \theta_{[\lambda, \tilde{\Omega}, M]}$  in  $\tilde{\Omega}$  for each  $M \geq M_t$  and, hence, Theorem 6( $\tilde{\Omega}$ )(a) implies  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \lim_{M \nearrow \infty} \theta_{[\lambda, \tilde{\Omega}, M]} = \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$  in  $\tilde{\Omega}$ . Therefore,  $\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \mathfrak{L}_{[\lambda, \tilde{\Omega}]}^{\min}$  in  $\tilde{\Omega}$  and it is apparent from (6) that

$$\lim_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \mathfrak{M}_{[\lambda, \Omega]}^{\min} \quad \text{in } \Omega. \quad (7)$$

This concludes the proof of Theorem 81(c).

Now, suppose  $u_0 > 0$  is arbitrary —not necessarily a subsolution of (4)— and

$$\lambda \geq \sigma_2. \quad (8)$$

Then, thanks to the parabolic maximum principle,  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq u_{[\sigma_2 - \varepsilon, \Omega]}(\cdot, t; u_0)$  in  $\Omega$  for each  $\varepsilon > 0$  and  $t > 0$ , and, hence, thanks to (6), for each  $\varepsilon > 0$  sufficiently small,

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \liminf_{t \nearrow \infty} u_{[\sigma_2 - \varepsilon, \Omega]}(\cdot, t; u_0) \geq \mathfrak{M}_{[\sigma_2 - \varepsilon, \tilde{\Omega}]}^{\min} \quad \text{in } \Omega \quad (9)$$

since  $\sigma_1 < \sigma_2 - \varepsilon < \sigma_2$ ; it should be noted that we cannot apply (7), because  $u_0$  is not necessarily a subsolution of (4). As (9) holds true for any  $\varepsilon > 0$ ,  $\varepsilon \sim 0$ , and, thanks to the second relationship of (5),  $\lim_{\varepsilon \searrow 0} \mathfrak{M}_{[\sigma_2 - \varepsilon, \tilde{\Omega}]}^{\min} = \mathfrak{M}_{[\sigma_2, \Omega_+]}^{\min}$ , we obtain from (9) that

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \mathfrak{M}_{[\sigma_2, \Omega_+]}^{\min} \quad \text{in } \Omega. \quad (10)$$

In particular,

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \infty \quad \text{in } \Omega \setminus \Omega_+, \quad (11)$$

which concludes the proof of Theorem 8 2(a). Moreover, (11) is satisfied uniform on  $\partial\Omega_+$ . Thus, for each  $M > 0$  a constant  $T_M > 0$  exists such that  $u_{[\lambda, \Omega]}(x, t; u_0) \geq M$  for all  $(x, t) \in \partial\Omega_+ \times [T_M, \infty)$ . Hence,  $u_{[\lambda, \Omega]}(\cdot, t; u_0)$  provides us with a supersolution of

$$\begin{cases} \frac{\partial v}{\partial t} - \Delta v = \lambda v - av^p & \text{in } \Omega_+ \times (T_M, \infty) \\ v = M & \text{on } \partial\Omega_+ \times (T_M, \infty) \\ v(\cdot, 0) = u_{[\lambda, \Omega]}(\cdot, T_M; u_0) & \text{in } \Omega_+ \end{cases} \quad (12)$$

and, so,  $u_{[\lambda, \Omega]}(x, t; u_0) \geq u_{[\lambda, \Omega_+, M]}(x, t - T_M; u_{[\lambda, \Omega]}(\cdot, T_M; u_0))$  if  $(x, t) \in \Omega_+ \times (T_M, \infty)$ . So, thanks to Theorem 6( $\Omega_+$ )(b), in  $\Omega_+$  we have that

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \lim_{t \nearrow \infty} u_{[\lambda, \Omega_+, M]}(\cdot, t - T_M; u_{[\lambda, \Omega]}(\cdot, T_M; u_0)) = \theta_{[\lambda, \Omega_+, M]}$$

and, henceforth, passing to the limit as  $M \uparrow \infty$  in this estimate, Theorem 6( $\Omega_+$ )(a) gives  $\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \mathfrak{L}_{[\lambda, \Omega_+]}^{\min}$  in  $\Omega_+$ . Therefore, taking into account (11), we find that

$$\liminf_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \geq \mathfrak{M}_{[\lambda, \Omega_+]}^{\min} \quad \text{in } \Omega. \quad (13)$$

Relation (13) provides us with the lower estimate of Theorem 8 2(b).

Now, besides (8), suppose  $u_0$  is a subsolution of (4). Then, for each  $t > 0$ , the function  $u_{[\lambda, \Omega]}(\cdot, t; u_0)$  is a subsolution of (4). Thus, fixing  $t > 0$  and setting

$$M_t := \max_{\Omega_+} u_{[\lambda, \Omega]}(\cdot, t; u_0),$$

it is apparent from Theorem 6 that  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \theta_{[\lambda, \Omega_+, M]}$  in  $\Omega_+$  for each  $M \geq M_t$  and, hence, Theorem 6( $\Omega_+$ )(a) implies  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \lim_{M \nearrow \infty} \theta_{[\lambda, \Omega_+, M]} = \mathfrak{L}_{[\lambda, \Omega_+]}^{\min}$  in  $\Omega_+$ . Therefore,  $\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq \mathfrak{L}_{[\lambda, \Omega_+]}^{\min}$  in  $\Omega_+$  and, due to (13), we find that

$$\lim_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t; u_0) = \mathfrak{M}_{[\lambda, \Omega_+]}^{\min} \quad \text{in } \Omega., \quad (14)$$

which concludes the proof of Theorem 8 2(c).

To complete the proof of the theorem it remains to obtain the upper estimates when  $u_0 > 0$  is an arbitrary positive initial data. In order to get these estimates, we will obtain upper bounds in  $\Omega_+$  for any positive solution of (1). Those bounds can be obtained arguing as follows. Fix  $\lambda \geq \sigma_1$  and froze the solution of (1) at time unity by considering the function  $u_{[\lambda, \Omega]}(\cdot, 1; u_0)$ . By the parabolic maximum principle,  $u_{[\lambda, \Omega]}(\cdot, 1; u_0)$  lies in the interior of the cone of positive functions of  $\mathcal{C}_0^1(\bar{\Omega})$ . Thus, a constant  $\kappa = \kappa(u_0) > 0$  exists for which

$$u_{[\lambda, \Omega]}(\cdot, 1; u_0) < \kappa \varphi \quad (15)$$

where  $\varphi$  is the principal eigenfunction associated with  $\sigma_0$ . We claim that  $\Lambda > \max\{\lambda, \sigma_2\}$  exists for which  $\kappa \varphi$  is a subsolution of

$$\begin{cases} -\Delta u = \Lambda u - a u^p & \text{in } \Omega \\ u = 0 & \text{on } \partial \Omega \end{cases} \quad (16)$$

Indeed, since  $\kappa \varphi = 0$  on  $\partial \Omega$ ,  $\kappa \varphi$  is a subsolution of (16) if, and only if,

$$-\Delta(\kappa \varphi) \leq \Lambda \kappa \varphi - a(\kappa \varphi)^p \quad \text{in } \Omega$$

or, equivalently,  $a(\kappa \varphi)^{p-1} \leq \Lambda - \sigma_0$  in  $\Omega$ , which is satisfied for any  $\Lambda$  sufficiently large. Now, thanks to the parabolic maximum principle, (15) implies that

$$u_{[\lambda, \Omega]}(\cdot, t + 1; u_0) = u_{[\lambda, \Omega]}(\cdot, t; u_{[\lambda, \Omega]}(\cdot, 1; u_0)) \leq u_{[\lambda, \Omega]}(\cdot, t; \kappa \varphi)$$

for any  $(x, t) \in \Omega \times (0, \infty)$ . Similarly,  $\Lambda > \lambda$  implies  $u_{[\lambda, \Omega]}(\cdot, t; \kappa\varphi) \leq u_{[\Lambda, \Omega]}(\cdot, t; \kappa\varphi)$  and, hence, for each  $t > 0$  we have that

$$u_{[\lambda, \Omega]}(\cdot, t + 1; u_0) \leq u_{[\Lambda, \Omega]}(\cdot, t; \kappa\varphi) \quad \text{in } \Omega. \quad (17)$$

Moreover, since  $\kappa\varphi$  is a subsolution of (16), it follows from Theorem 8 2(c) that

$$\lim_{t \nearrow \infty} u_{[\Lambda, \Omega]}(\cdot, t; \kappa\varphi) = \mathfrak{M}_{[\Lambda, \Omega]}^{\min} \quad \text{in } \Omega.$$

Thus, thanks to (17),

$$\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(\cdot, t + 1; u_0) \leq \mathfrak{M}_{[\Lambda, \Omega]}^{\min} \quad \text{in } \Omega \quad (18)$$

and, therefore,  $u_{[\lambda, \Omega]}(\cdot, t; u_0)$ ,  $t > 0$ , is uniformly bounded above in any compact subset of  $\Omega_+$ . This provides us with the necessary a priori bounds to conclude the proof of the theorem.

Suppose (1), and, for each  $\delta > 0$  sufficiently small, consider

$$\tilde{\Omega}_\delta := \{x \in \tilde{\Omega} : \text{dist}(x, \partial\tilde{\Omega}) > \delta\}.$$

Since  $\partial\tilde{\Omega}_\delta \subset \Omega_+$ , thanks to the existence of the a priori bounds, a constant  $M_0 > 0$  exists such that, for each  $M \geq M_0$  and  $t \geq 0$ ,

$$u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq M \quad \text{on } \partial\tilde{\Omega}_\delta.$$

Thus, the parabolic maximum principle shows that

$$u_{[\lambda, \Omega]}(x, t; u_0) \leq u_{[\lambda, \tilde{\Omega}_\delta, M]}(x, t; u_0) \quad \text{for each } (x, t) \in \tilde{\Omega}_\delta \times [0, \infty)$$

where  $u_{[\lambda, \tilde{\Omega}_\delta, M]}(x, t; u_0)$  stands for the solution of (3) with  $D = \tilde{\Omega}_\delta$ . As the vanishing region of  $a(x)$  in  $\tilde{\Omega}_\delta$  equals  $\Omega_{0,2}$  and we are assuming  $\lambda < \sigma_2$ , Theorems 2, 6 also apply in  $\tilde{\Omega}_\delta$ , besides in  $\Omega$ , and maintaining the notations introduced above, we have that, in  $\tilde{\Omega}_\delta$ ,

$$\limsup_{t \nearrow \infty} u_{[\lambda, \tilde{\Omega}_\delta, M]}(x, t; u_0) = \theta_{[\lambda, \tilde{\Omega}_\delta, M]}$$

and, hence, for each  $\delta > 0$  sufficiently small,  $\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(x, t; u_0) \leq \theta_{[\lambda, \tilde{\Omega}_\delta, M]}$  in  $\tilde{\Omega}_\delta$ . As this estimate holds true for all  $M \geq M_0$ , passing to the limit as  $M \uparrow \infty$  gives

$$\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(x, t; u_0) \leq \mathfrak{L}_{[\lambda, \tilde{\Omega}_\delta]}^{\min} \quad \text{in } \tilde{\Omega}_\delta$$

and, therefore, passing to the limit as  $\delta \downarrow 0$ , it is easy to see that

$$\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(x, t; u_0) \leq \mathfrak{L}_{[\lambda, \Omega]}^{\max} \quad \text{in } \tilde{\Omega}.$$

It should be noted that such a large solution is well defined because  $\lambda < \sigma_2$ . This estimate concludes the proof of Theorem 8 1. Finally, suppose  $\lambda \geq \sigma_2$  and, for each  $\delta > 0$  sufficiently small, consider the open  $\delta$ -neighborhood  $\Omega_+^\delta := \{x \in \Omega_+ : \text{dist}(x, \partial\Omega_+) > \delta\}$ . Arguing as above, a constant  $M_0 > 0$  exists such that  $u_{[\lambda, \Omega]}(\cdot, t; u_0) \leq M$  on  $\partial\Omega_+^\delta \subset \Omega_+$  for each  $M \geq M_0$  and  $t \geq 0$ . Hence,

$$u_{[\lambda, \Omega]}(x, t; u_0) \leq u_{[\lambda, \Omega_+^\delta, M]}(x, t; u_0) \quad \text{for each } (x, t) \in \Omega_+^\delta \times [0, \infty).$$

Moreover,  $\limsup_{t \nearrow \infty} u_{[\lambda, \Omega_+^\delta, M]}(x, t; u_0) = \theta_{[\lambda, \Omega_+^\delta, M]}$  in  $\Omega_+^\delta$ , and, henceforth, for each  $\delta > 0$  sufficiently small, we have that  $\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(x, t; u_0) \leq \theta_{[\lambda, \Omega_+^\delta, M]}$  in  $\Omega_+^\delta$ . As this estimate holds true for all  $M \geq M_0$ , passing to the limit as  $M \uparrow \infty$  gives  $\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(x, t; u_0) \leq \mathfrak{L}_{[\lambda, \Omega_+^\delta]}^{\min}$  in  $\Omega_+^\delta$ . Therefore, passing to the limit as  $\delta \downarrow 0$ , we find that

$$\limsup_{t \nearrow \infty} u_{[\lambda, \Omega]}(x, t; u_0) \leq \mathfrak{L}_{[\lambda, \Omega_+]}^{\max} \quad \text{in } \Omega_+.$$

This concludes the proof of the theorem.  $\square$

## 4 Final comments and further reading

Although Theorems 1, 8 show that the dynamics of (1) and (5) are different, the main result of [28] shows that the limiting profiles of the positive solutions of (5) approach the limiting profiles of the solutions of (1) as  $\mu \downarrow 0$ . Therefore, enlarging the spaces of smooth functions to include metasolutions gives the structural stability of the dynamics of (5) as  $\mu \downarrow 0$  when metasolutions are unique.

The theory developed in this paper is valid in the general setting of [25], where we send the reader for further details. In particular, the proof of Theorem 8 can be adapted to cover the general framework of [25]—the technical details will appear elsewhere—; in this slightly expository paper we have preferred to keep controlled any irrelevant difficulty coming from the use of sophisticated assumptions and/or notations.

The most pioneering result concerning the positive solutions of (4) seems to be the one obtained by H. Brezis and L. Oswald in [5, Theorem 1], where it was shown that (4) possesses at most one positive solution and that a positive solution exists if and only if

$$\sigma[-\Delta - a_0(x), \Omega] < 0 < \sigma[-\Delta - a_\infty(x), \Omega], \quad (1)$$

where

$$a_0(x) := \lim_{u \searrow 0} (\lambda - a(x)u^{p-1}) = \lambda$$

and

$$a_\infty(x) := \lim_{u \nearrow \infty} (\lambda - a(x)u^{p-1}) = \begin{cases} -\infty & \text{if } x \in \Omega_+ \\ \lambda & \text{if } x \in \Omega \setminus \Omega_+ \end{cases}$$

Although it is rather obvious that the first inequality of (1) becomes  $\lambda > \sigma_0$  in our setting, it is far from easy to realize why the second inequality becomes  $\lambda < \sigma_1 = \sigma[-\Delta, \Omega_{0,1}]$ ; actually, this is an extremely sharp result going back to [22] and establishing that

$$\lim_{\xi \nearrow \infty} \sigma[-\Delta + a\xi^{p-1} - \lambda, \Omega] = \sigma[-\Delta - \lambda, \Omega_{0,1}] \quad (2)$$

and having a number of applications in Mathematical Biology (e.g., [23]) — certainly, another bisociation. As  $\sigma[-\Delta - \lambda, \Omega_{0,1}] = \sigma_1 - \lambda$ , it is apparent from (2) that (1) becomes into  $\sigma_0 < \lambda < \sigma_1$ , which is the characterization of our Theorem 2. Another significant pioneering work concerning with the existence of smooth positive solutions of (4) was done by T. Ouyang in [35] as a part of his Ph. D. Dissertation under the supervision of Professor W. M. Ni on Yamabe's problem [37] about compact Riemannian manifolds — a further bisociation.

Metasolutions are not only relevant for analyzing the dynamics of (1), as since their discovery, they have revealed to be crucial for understanding the dynamics of larger classes of reaction diffusion equations, e.g., [10] and [11], and systems, e.g., [29] and Y. Du [12]. Actually, they might play a crucial role in the problem of understanding the dynamics of superlinear indefinite problems (cf. the final discussion of [19]).

**Acknowledgements.** This work has been supported by the Ministry of Science and Technology of Spain under grant BFM2000-0797.

## References

- [1] H. Amann, On the existence of positive solutions of nonlinear elliptic boundary value problems. *Ind. Univ. Math. J.*, 21:125–146, 1971.
- [2] H. Amann, Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces. *SIAM Rev.*, 18:620–709, 1976.
- [3] H. Amann and J. López-Gómez, A priori bounds and multiple solutions for superlinear indefinite elliptic problems. *J. Diff. Eqns.*, 146:336–374, 1998.
- [4] C. Bandle and M. Marcus, Large solutions of semilinear elliptic equations: Existence, uniqueness and asymptotic behavior. *J. D'Analysis Math.*, 58:9–24, 1991.
- [5] H. Brezis and L. Oswald, Remarks on sublinear elliptic equations. *Nonl. Anal. TMA*, 10:55–64, 1986.
- [6] F. Browder, Reflection on the future of mathematics. *Notices of the AMS*, 49:658–662, 2002.

- [7] S. Cano-Casanova and J. López-Gómez, Properties of the principal eigenvalues of a general class of non-classical mixed boundary Value problems. *J. Diff. Eqns.*, 178:123–211, 2002.
- [8] D. Cohen and T. Laetch, Nonlinear boundary value problems suggested by chemical reactor theory. *J. Diff. Eqns.*, 7:217–226, 1970.
- [9] M. A. Del Pino, Positive solutions of a semilinear elliptic equation on a compact manifold. *Nonl. Anal. TMA*, 22:1423–1430, 1994.
- [10] M. Delgado, J. López-Gómez and A. Suárez, Nonlinear versus linear diffusion. From classical solutions to metasolutions. *Adv. Diff. Eqns.*, 7:1101–1124, 2002.
- [11] M. Delgado, J. López-Gómez and A. Suárez, Characterizing the existence of large solutions for a class of sublinear problems with nonlinear diffusion. *Adv. Diff. Eqns.*, 7:1235–1256, 2002.
- [12] Y. Du, Effects of a degeneracy in the competition model. Classical and generalized steady-state solutions. *J. Diff. Eqns.*, 181:92–132, 2002.
- [13] Y. Du and Q. Huang, Blow-up solutions for a class of semilinear elliptic and parabolic equations. *SIAM J. Math. Anal.*, 31:1–18, 1999.
- [14] J. M. Fraile, P. Koch-Medina, J. López-Gómez and S. Merino, Elliptic eigenvalue problems and unbounded continua of positive solutions of a semilinear equation. *J. Diff. Eqns.*, 127:295–319, 1996.
- [15] J. E. Furter and J. López-Gómez, Diffusion mediated permanence problem for an heterogeneous Lotka-Volterra competition model. *Proc. Roy. Soc. Edinburgh*, 127-A:281–336, 1997.
- [16] J. García-Melián, R. Gómez-Reñasco, J. López-Gómez and J. C. Sabina de Lis, Point-wise growth and uniqueness of positive solutions for a class of sublinear elliptic problems where bifurcation from infinity occurs. *Arch. Rat. Mech. Anal.*, 145:261–289, 1998.
- [17] J. García-Melián, R. Letelier-Albornoz and J. C. Sabina de Lis, Uniqueness and asymptotic behaviour for solutions of semilinear problems with boundary blow-up. *Proc. Amer. Math. Soc.*, 129: 3593–3602, 2001.
- [18] R. Gómez-Reñasco, *The effect of Varying Coefficients in Semilinear Elliptic Boundary Value Problems. From Classical Solutions to Metasolutions*. Ph. D. Dissertation, Universidad de La Laguna, Tenerife, March 1999.
- [19] R. Gómez-Reñasco and J. López-Gómez, The effect of varying coefficients on the dynamics of a class of superlinear indefinite reaction diffusion equations. *J. Diff. Eqns.*, 167:36–72, 2000.



- [20] R. Gómez-Reñasco and J. López-Gómez, On the existence and numerical computation of classical and non-classical solutions for a family of elliptic boundary value problems. *Nonl. Anal. Th. Meth. Appl.*, 48:567–605, 2002.
- [21] J. B. Keller, On solutions of  $\Delta u = f(u)$ . *Comm. Pure and Appl. Maths.*, X:503–510, 1957.
- [22] J. López-Gómez, On linear weighted boundary value problems, in *Partial Differential Equations; Models in Physics and Biology*, Akademie Verlag, Berlin 1994, 188–203.
- [23] J. López-Gómez, Permanence under strong competition, Dynamical Systems and Applications. *WSSAA*, 4:473–488, 1995.
- [24] J. López-Gómez, The maximum principle and the existence of principal eigenvalues for some linear weighted boundary value problems. *J. Diff. Eqns.*, 127:263–294, 1996.
- [25] J. López-Gómez, Large solutions, metasolutions, and asymptotic behaviour of the regular positive solutions of sublinear parabolic problems. *El. J. Diff. Eqns.*, Conf.05:135–171, 2000.
- [26] J. López-Gómez, Varying bifurcation diagrams of positive solutions for a class of indefinite superlinear boundary value problems. *Trans. Amer. Math. Soc.*, 352:1825–1858, 2000.
- [27] J. López-Gómez, *Spectral Theory and Nonlinear Functional Analysis*. Chapman & Hall/CRC Research Notes in Mathematics 426, Boca Ratón, 2001.
- [28] J. López-Gómez, Approaching metasolutions by classical solutions. *Diff. Int. Eqns.*, 14:739–750, 2001.
- [29] J. López-Gómez, Coexistence and metacoexistence states in competing species models. *Houston J. Maths.*, 29:485–538, 2003.
- [30] J. López-Gómez, The boundary blow-up rate of large solutions. *J. Diff. Eqns.* In press.
- [31] J. López-Gómez and M. Molina-Meyer, The maximum principle for cooperative weakly coupled elliptic systems and some applications. *Diff. Int. Eqns.*, 7:383–398, 1994.
- [32] J. López-Gómez and J. C. Sabina de Lis, First variations of principal eigenvalues with respect to the domain and point-wise growth of positive solutions for problems where bifurcation from infinity occurs. *J. Diff. Eqns.*, 148:47–64, 1998.
- [33] M. Marcus and L. Véron, Uniqueness and asymptotic behavior of solutions with boundary blow-up for a class of nonlinear elliptic equations. *Ann. Inst. Henri Poincaré*, 14:237–274, 1997.

- [34] R. Osserman, On the inequality  $\Delta u \geq f(u)$ . *Pacific J. of Maths.*, 7:1641–1647, 1957.
- [35] T. Ouyang, On the positive solutions of semilinear equations  $\Delta u + \lambda u - hu^p = 0$  on the compact manifolds. *Trans. Amer. Math. Soc.*, 331:503–527, 1992.
- [36] D. Sattinger, *Topics in Stability and Bifurcation Theory*. Lectures Notes in Mathematics 309, Springer, Berlin/New York, 1973.
- [37] H. Yamabe, On a deformation of Riemaniann structures on compact manifolds. *Osaka Math. J.*, 12:21–37, 1960.

## **Herramientas de modelado de superficies en un sistema CAD/CAM para diseño de buques**

F. ALONSO<sup>1</sup>

SENER, Ingeniería y Sistemas S.A. Las Arenas (Vizcaya)

`fernando.alonso@sener.es`

SENER, Ingeniería y Sistemas S. A. es una empresa privada española de ingeniería, fundada en Bilbao en 1956. Aunque SENER originalmente surgió como una oficina de ingeniería naval, evolucionó de forma muy rápida hasta convertirse en una empresa multidisciplinar trabajando en una gran variedad de campos de la ingeniería.

SENER nació con una gran vocación de innovación y desarrollo tecnológico, lo que le ha hecho participar en todo tipo de proyectos avanzados de ingeniería, tanto nacionales como internacionales, utilizando y desarrollando las tecnologías más avanzadas en cada campo de la ingeniería. En la actualidad SENER puede considerarse una de las ingenierías españolas de referencia en los campos aeronáutico, comunicaciones, espacial, energía, industrial, civil y naval.

El Sistema FORAN es un Sistema CAD/CAM específicamente desarrollado para construcción naval, utilizado en más de 120 astilleros en todo el mundo. El Sistema FORAN comenzó a desarrollarse inicialmente como una herramienta interna de SENER para facilitar y automatizar los procesos de diseño naval.

El primer desarrollo que se abordó fue el de una herramienta informática para la definición rápida de las formas de un buque a partir de unos pocos parámetros básicos definiendo las características principales del mismo (eslora, manga, calado, etc.).

Otros requisitos de partida de este desarrollo fueron el que las superficies generadas estuvieran intrínsecamente alisadas desde el punto de vista del proyectista naval (requisito difícil de expresar matemáticamente) y la disponibilidad de herramientas para una modificación rápida y eficaz de las superficies generadas.

El resultado de este desarrollo fue el primer módulo FORAN, de Generación de formas de buques, que también dio nombre al Sistema FORAN (acrónimo de FORMas ANALíticas).

---

<sup>1</sup>Jefe de Sistemas Navales de SENER, Ingeniería y Sistemas S.A.

Posteriormente el Sistema FORAN se fue ampliando a otras áreas del diseño del buque (arquitectura naval, definición de la estructura, información para fabricación, equipos, tuberías, diseño eléctrico, ...) hasta llegar a la situación actual del Sistema en la que prácticamente se cubre la mayor parte de las tareas de diseño y fabricación de un buque, a lo largo de sus distintas etapas de proyecto (diseño inicial, diseño de clasificación y diseño de detalle).

A finales de los años 60, a la vista del éxito interno de las herramientas informáticas desarrolladas y del interés detectado en distintos astilleros clientes de SENER, se decidió comercializar el Sistema FORAN, tanto en España como en el extranjero. La primera versión del Sistema funcionaba en ordenadores IBM 1130 con 8 K de memoria, por lo que los algoritmos desarrollados (por supuesto en Fortran) tenían que estar absolutamente optimizados, tanto en lo referente a memoria utilizada como a tiempo de proceso.

Una de las principales características de FORAN es la creación de un modelo de producto del buque en una base de datos relacional conteniendo toda la información del buque. La utilización de una base de datos relacional permite trabajar en un entorno concurrente multiusuario.

El modelo de producto es una representación completa del buque, conteniendo datos geométricos, información de la estructura de producto, atributos tecnológicos de los elementos, así como toda la información necesaria para el proceso de fabricación y montaje. Parte importante de esta información son las relaciones topológicas o asociativas entre los distintos elementos, que permiten gestionar de forma eficiente cambios o modificaciones en el diseño.

En la Figura 1 puede verse una representación tridimensional del modelo geométrico construido por el Sistema correspondiente a un buque de suministros *offshore*.

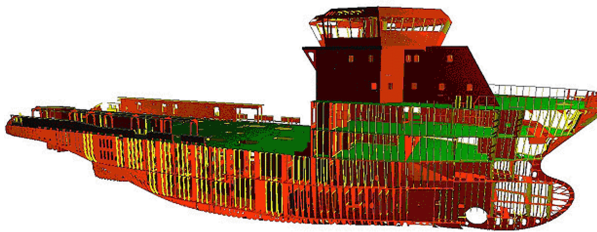


Figura 1: Modelo 3D del buque. Buque de suministros *offshore*

En la Figura 2 puede verse con más detalle una zona del mismo buque. Para poder visualizar el interior del modelo, en ambos casos se han eliminado las planchas del casco del buque.

Un aspecto de primordial importancia en el desarrollo del Sistema FORAN y en la evolución del mismo a lo largo de los años ha sido la utilización de las tecnologías básicas de modelado y visualización más avanzadas disponibles en cada momento. Especialmente importantes son las herramientas y los algoritmos

matemáticos básicos utilizados para la definición del modelo de superficies y sólidos manejado por el Sistema.

Estas herramientas y algoritmos básicos han sido desarrollados, bien internamente, como fue el caso de la Generación de formas que constituyó el origen del Sistema FORAN, o bien en colaboración con prestigiosos departamentos de algunas universidades españolas.

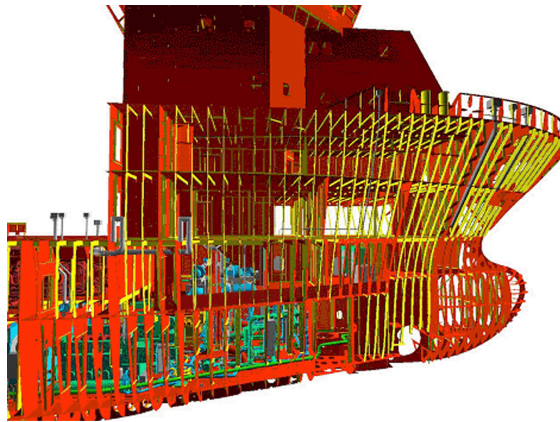


Figura 2: Zona de proa con equipos y tuberías. Buque de suministros *offshore*

Como ejemplo de colaboración con una universidad, se incluye a continuación una versión en español del trabajo *Mejoras en el modelado de la superficie del casco*, realizado por A. Rodríguez (SENER Ingeniería y Sistemas, S.A.), M. Vivo (Universidad Politécnica de Cataluña) y A. Vinacua (Universidad Politécnica de Cataluña), que fue presentado en COMPIT'2000 (1st International Euroconference on Computer Applications and Information Technology in the Maritime Industries).

## Mejoras en el modelado de la superficie del casco

A. Rodríguez, M. Vivo y A. Vinacua

### Resumen

Este trabajo describe una serie de herramientas desarrolladas para acelerar los procesos de definición de la superficie del casco utilizando algoritmos especializados. La primera etapa consiste en mejorar el proceso de ajuste cuando se parte de una cartilla de trazado. La segunda etapa consiste en automatizar al máximo posible el proceso de alisado una vez creado el modelo de superficies. En este contexto, el modelo de superficies se define mediante un conjunto de parches NURBS recortados. Los procesos de aproximación se implementan mediante algoritmos de minimización mientras que el alisado se realiza mediante mecanismos discretos más cercanos a la manera tradicional de realizar dicho proceso que un planteamiento basado en algoritmos variacionales.

## 1 Introducción

Las tareas de ajuste y alisado del plano de formas suelen ser el comienzo de un proceso completo de ingeniería naval. Estas tareas pueden consumir un tiempo considerable, mientras que el resultado de las mismas es una información esencial para los procesos subsidiarios.

Existen varias estrategias para enfrentarse a este cuello de botella, siendo la más efectiva la que presentan los Sistemas integrados que utilizan de forma extensiva una definición topológica del producto. En estos Sistemas, una vez que se tienen unas formas preliminares ajustadas dentro de un nivel de tolerancia determinado, pero no necesariamente alisadas, se puede comenzar a trabajar en los procesos subsidiarios. Mientras tanto se sigue progresando en las tareas de ajuste y alisado hasta alcanzar la calidad requerida por los procesos productivos.

Una vez completado el modelo del casco, éste se incorpora a los procesos subsidiarios de forma automática, ya que la definición de elementos realizada sobre las formas preliminares sigue siendo válida. Esta es una propiedad de la definición topológica, que es invariante ante pequeños cambios, como son los que se producen en el proceso de alisado.

Este esquema permite establecer una serie de requisitos que debe cumplir un sistema de modelado de superficies para construcción naval:

- Facilidad para un ajuste rápido dentro de un nivel de tolerancias establecido.
- Métodos de alisado rápidos e intuitivos, basados en algoritmos complejos pero evitando que el usuario tenga que involucrarse en procesos matemáticos complicados.
- Basado en una definición topológica, de manera que las cosas sólo se definen una vez.

Un Sistema que dispone de una definición topológica debe progresar por la vía de mejorar las herramientas de modelado de las superficies del buque, siendo el objetivo de este trabajo el de presentar dichas mejoras.

## 2 Entrada de datos

Una de las características más importantes de un Sistema de CAD es su capacidad de intercambiar información con otros Sistemas a través de un amplio espectro de formatos. La información de partida de un proceso de ajuste suele venir dada a través de una cartilla de trazado, por lo que es necesario poder leer ficheros con información numérica (coordenadas de puntos) en formato textual.

También es muy habitual recibir información de trazado en formato DXF (formato propio de Autodesk pero considerado estándar *de facto*) por lo que debe incluirse en los requisitos del Sistema. Para gestionar los datos importados de esta manera, el Sistema deberá disponer de entidades adecuadas que reflejen

esta información. Básicamente se consigue con puntos y polilíneas (por supuesto, en 3D).

Este tipo de información requiere ser procesada previamente a su utilización en el modelado. El tipo de operaciones requeridas es:

- Selección de puntos para crear polilíneas (por ejemplo, en un cierto plano).
- Construcción de polilíneas a partir de puntos.
- Unión y división de polilíneas.
- Recorte de polilíneas.
- Aplicación de transformaciones geométricas sobre puntos y polilíneas.
- Combinaciones lineales de dos polilíneas.
- Segmentación de polilíneas con un ángulo umbral.

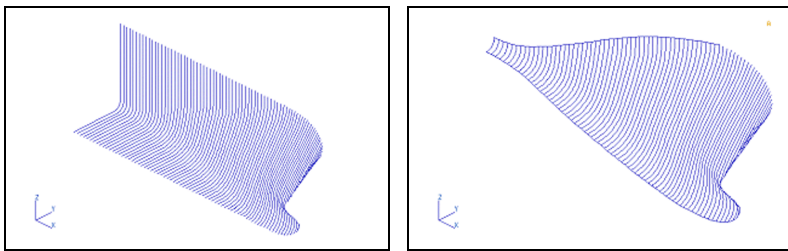


Figura 3: Antes y después de procesar la información proveniente de formato DXF

De esta manera se prepara la información para que pueda ser utilizada en las tareas siguientes. Mientras se realiza este trabajo, ya se está considerando cómo se va a subdividir la superficie en distintos parches, de acuerdo a las características geométricas del casco (codillos, zonas muy curvadas, zonas planas, etc.).

Otro formato que no se puede olvidar cuando se trabaja con NURBS es el formato IGES. En dicho formato existen entidades específicas para la definición de NURBS como es la entidad 128. A través de este formato se pueden intercambiar directamente y sin pérdida de información modelos geométricos basados en NURBS. Es también el formato habitual para intercambiar información con los programas de CFD, por lo que su importancia queda de manifiesto.

### 3 Definición de Curvas

El siguiente paso en la construcción del modelo es la creación de curvas que contienen la información geométrica de las superficies que se pretenden obtener.

Estas curvas se construirán a partir de los puntos y polilíneas obtenidos en la etapa anterior.

El usuario deberá disponer de herramientas sencillas pero efectivas, para realizar las siguientes tareas:

- Creación de curvas interpolantes a partir de un conjunto de puntos o de una polilínea. Opcionalmente deberán poder definirse condiciones de tangencia en los extremos.
- Creación de curvas por aproximación de un conjunto de puntos o de una polilínea. Este tipo de algoritmos requiere una tolerancia especificada por el usuario. Es normal trabajar habitualmente con una tolerancia típica (en construcción naval será de algunos milímetros) que sólo se cambiará cuando sea necesario. Para realizar esta aproximación, se ha seleccionado un algoritmo estándar basado en ajuste de *splines* suavizados. Se ha preferido este algoritmo frente a un ajuste típico de mínimos cuadrados en distancias porque produce mejores resultados en la distribución de curvaturas, aunque introduce más nudos en la definición del *spline*.
- Creación automática de secciones cónicas a partir de condiciones geométricas: círculo por tres puntos, círculo mediante tangencias y radio, elipses, parábolas, etc.

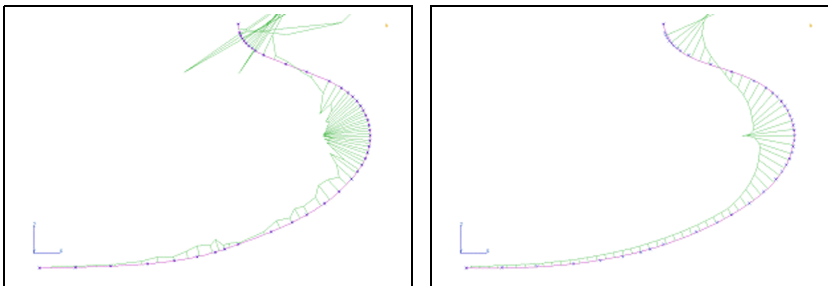


Figura 4: Curva generada por interpolación (izquierda) y por aproximación. Se muestran las curvaturas para verificar las diferencias

A veces puede ser útil que el usuario pueda definir curvas NURBS a partir de los elementos básicos de esta formulación. Definiendo el polígono de control, el grado y eventualmente los pesos de los puntos, se puede construir una curva de una manera bastante primitiva. No obstante, un usuario familiarizado con esta definición puede utilizarla de forma creativa, para rellenar zonas de las que se carece de información, por ejemplo.

Con este conjunto de herramientas es posible convertir de una manera rápida y efectiva un conjunto de puntos y polilíneas en una serie de curvas NURBS que contienen toda la información geométrica relevante para construir



la superficie. La calidad de las curvas, desde el punto de vista del alisado, se puede comprobar visualizando las curvaturas de las mismas. En cuanto al ajuste, ha sido controlado a lo largo del proceso mediante las tolerancias establecidas.

## 4 Manipulación de curvas

Una vez definido un modelo preliminar de curvas es necesario algún trabajo adicional para adecuarlas a la definición de superficies. Por un lado, es necesario alisar dichas curvas para mejorar la calidad de los parches que se construyan con ellas. También deberán ser recortadas, copiadas, transformadas, etc. para ajustarse a las fronteras de las superficies que se pretende construir.

Para realizar estas tareas el usuario necesita las siguientes funciones:

- Edición de curvas. Se puede hacer de varias maneras:
  - La más común, tratándose de NURBS, es mediante la manipulación de los puntos de control. La relación entre la disposición de los puntos de control y la curva en sí es algo abstracta. Esto condiciona que aunque se pueda usar esta herramienta para ajuste, es más adecuada para alisado, ya que la disposición de los puntos de control define las derivadas de la curva. Por tanto, una distribución regular de puntos de control genera una distribución de curvaturas más homogénea, que es un criterio de alisado.
  - Cuando se pretende utilizar la edición para ajuste es más conveniente poder manipular directamente puntos de la curva. Esto requiere un algoritmo algo más sofisticado, que permita mover una serie de puntos de la curva a sus nuevas posiciones. Dependiendo del número de puntos que se pretendan mover simultáneamente, el problema puede ser superdeterminado. Si este es el caso, se deberá conseguir la solución más cercana al objetivo.
- Transformación de curvas. Una ventaja de esta formulación (NURBS) es que es invariante para transformaciones afines (traslación, rotación, simetría, escalado respecto a un punto, recta o plano, ...) por lo que implementarlas es realmente sencillo.
- Alisado de curvas. Esta tarea requiere la utilización de un algoritmo complicado que calcule una mejor redistribución de las curvaturas, manteniendo las modificaciones dentro de cierta tolerancia. A veces es necesario un alisado previo “a mano”, antes de que estos algoritmos puedan ser efectivos. Como ya se mencionó antes, se ha seleccionado un algoritmo estándar basado en ajuste de *splines* suavizados para realizar esta tarea.
- Otras tareas necesarias, como son las operaciones del estilo de “cortar y pegar” con curvas y la redefinición de una curva con puntos igualmente espaciados sobre la misma.

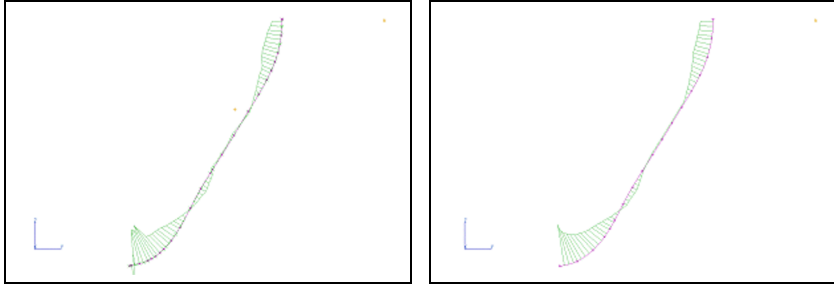


Figura 5: Curva generada por aproximación (izquierda) y la misma curva después de alisarse un poco. Se muestran las curvaturas para comprobar las diferencias

Finalmente el usuario obtiene un conjunto de curvas adecuadas para poder definir los parches. La calidad de los parches que se construyan dependerá de la calidad de las curvas que se estén utilizando, por lo que es conveniente dedicar esfuerzos en esta etapa a construir un modelo de curvas lo más correcto posible. De esta manera, la complejidad del proceso se descompone en etapas fácilmente abordables por un usuario normal.

## 5 Definición de superficies

Definir superficies a partir de curvas no debería ser muy complicado. Las curvas que se utilizan pueden contener información sobre secciones internas y/o bordes de las superficies. Las utilidades que necesita un usuario son:

- Generación de un parche recubriendo un conjunto de secciones genéricas (no necesariamente planas). Este proceso se denomina *skinning*.
- Generación de un parche rellenando una superficie a partir de los bordes. Los parches de Coon se pueden considerar un caso particular de este método. El parche generado se ajustará a los bordes, generando el interior de la superficie con algún método heurístico, como el de mínima energía o similar.

La eficacia de estos algoritmos es el factor clave para garantizar el éxito del proceso global.

## 6 Manipulación de superficies

En este momento se dispone ya de un modelo de superficies que representa la geometría del buque con un nivel de precisión determinado sobre la información inicial con la que ha comenzado el proceso. No obstante, este modelo puede no

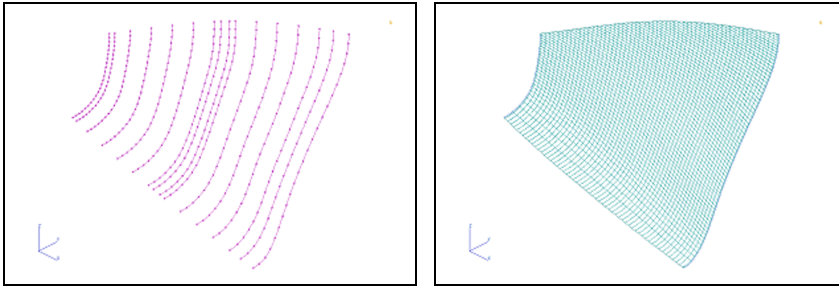


Figura 6: Un conjunto de curvas preparadas para ser recubiertas (izquierda) y el parche resultante

ser todavía lo suficientemente bueno, y no sólo con respecto al grado de alisado conseguido.

Puede haber desajustes entre los bordes de los parches y en las condiciones de tangencia de los mismos. Además, los algoritmos de generación de parches pueden producir resultados inadecuados en zonas difíciles, especialmente donde existen fuertes curvaturas.

En cualquier caso, el usuario necesita disponer de una serie de herramientas para realizar el ajuste final y conseguir un acabado lo más perfecto posible.

Las tareas que se han identificado como necesarias son:

- Edición de superficies. Estos comandos son similares a sus homólogos de curvas. Análogamente, se puede realizar de dos maneras:
  - Mediante la manipulación de la malla de puntos de control de la superficie. Una distribución regular de los puntos de la malla garantiza un buen grado de alisado a través de las derivadas de la superficie en las dos direcciones. En general, este proceso de alisado puede destruir el nivel de precisión obtenido en el ajuste. La buena noticia es que la mayor parte de las veces sólo se van a requerir pequeñas modificaciones locales en los parches, ya que una parte importante del alisado se ha realizado a través de las curvas.
  - La otra manera es moviendo directamente un conjunto de puntos sobre la superficie. De nuevo puede aparecer un problema de superdeterminación, en cuyo caso se deberá obtener la solución más cercana.

Combinando estos dos comandos, se puede realizar una corrección de los defectos locales originados durante el proceso previo, sin degradar el nivel de ajuste conseguido.

- Transformación de superficies. También en el caso de superficies la formulación de los NURBS es invariante ante transformaciones afines

(traslación, rotación, simetría, escalado respecto a un punto, recta o plano, ...) por lo que implementar dichas transformaciones es realmente sencillo.

- Recorte de superficies. Se trata de una herramienta básica para conseguir un buen modelado. La topología rectangular de los parches puede ser muy restrictiva para adaptarse a la geometría de un buque. Existe la posibilidad de colapsar todos los puntos de control de un borde del parche en un sólo punto para conseguir una forma triangular. No es recomendable abusar de este procedimiento ya que produce un punto singular donde el vector normal a la superficie no está definido. Esto puede originar inestabilidades numéricas en las computaciones que tengan lugar sobre dichos parches. Por tanto, es más recomendable utilizar el recorte para conseguir parches de topología no rectangular.
- Cosido de bordes. El objetivo es el de eliminar los huecos entre parches. Muchos programas evitan este problema definiendo la superficie del buque en un único parche. Esto es muy drástico ya que produce superficies con excesivos puntos en zonas donde no se necesitan. Por ejemplo, de esta manera se tiene que utilizar el mismo número de puntos para definir un pantoque circular (que en NURBS sólo necesita tres puntos) y el bulbo (que necesitará muchos más) Otro inconveniente es que el trabajo no puede ser repartido a distintos usuarios por zonas.

El cosido de parches requiere algoritmos sofisticados, pero las ventajas son muchas. De nuevo se requiere de una tolerancia para ajustar los bordes de los parches. Además este proceso introducirá sólo pequeñas modificaciones debido al ajuste previo.

- Edición de las condiciones de tangencia de los parches. De manera similar a como se maneja la continuidad en posición, se debe poder controlar la continuidad de las derivadas en bordes adyacentes.

A partir de este punto se dispone ya de un modelo completo, ajustado y alisado con lo que se puede considerar terminado el ciclo de vida del proceso. La practica más habitual, sin embargo, determina que este ciclo sea una espiral convergente y se pueda refinar el proceso en sucesivas etapas. Esto es realmente interesante cuando se dispone de un Sistema integrado basado en una definición topológica de los elementos, como ya se comentó al principio del artículo.

## 7 Conclusiones

Hay en el mercado bastantes herramientas para modelar cascos de buques utilizando la formulación de los NURBS. Sin embargo, la difusión de estas herramientas no ha tenido la amplitud que cabría esperar. Las causas de que esto ocurra podrían ser las siguientes:

- Estas herramientas se basan en el trabajo directo con una (o varias) superficies NURBS.

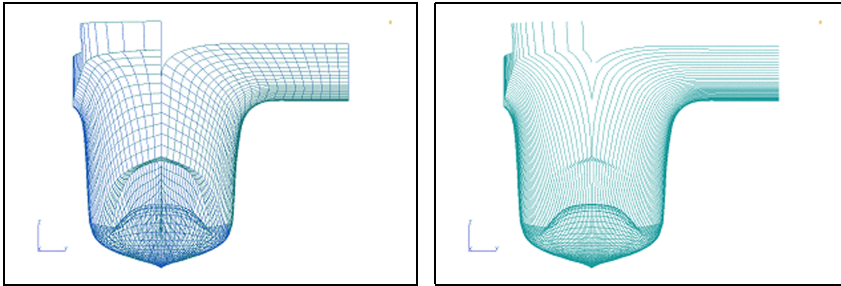


Figura 7: Un modelo de superficies de un catamarán terminado. Se visualizan curvas isoparamétricas de los parches (izquierda) y cuadernas (derecha)

- Tanto para conseguir el ajuste como el alisado, el usuario debe posicionar adecuadamente los puntos de control de las mallas. El ajuste se realiza “a mano”.
- El alisado se consigue cuando la malla tiene un “buen aspecto”.

Lo habitual es que el usuario deshaga un proceso cuando consigue el otro (ajuste y alisado). Deberá decidir sobre la marcha cuándo y dónde insertar más filas o columnas de puntos de control. El proceso de conseguir simultáneamente un buen grado de ajuste y alisado resulta sumamente laborioso y este tipo de herramientas se utilizan sobre todo en diseño libre.

Una herramienta de NURBS para producción debería contemplar otros aspectos:

- El ser humano se siente más cómodo trabajando con curvas. El trabajo con superficies directamente es inherentemente complejo y debería reducirse lo más posible.
- No obstante, construir superficies a partir de curvas que definan su geometría si resulta más intuitivo y abordable.
- Si la calidad de las curvas es buena y los algoritmos de construcción de parches funcionan razonablemente bien, el parche producido deberá estar ya bastante cerca de la solución final.
- De esta manera, aunque haya que manipular superficies, se trata de correcciones locales y terminaciones. Además, se disponen de algoritmos especiales que facilitan las tareas más engorrosas.
- Se debe disponer del recorte de parches para modelar trozos no rectangulares.
- Se debe disponer de algoritmos de alisado automático que solucionen muchas de las situaciones que se presentan.

El estado del arte en Geometría Computacional permite abordar la mayoría de estos problemas, por lo que los autores consideran que aparecerá pronto una nueva generación de herramientas de modelado de superficies con NURBS, que contemplen todas estas funcionalidades.

## 8 Algunas referencias FORAN

- [1] F. Alonso, C. Andújar, P. Brunet, L. García, I. Navazo y A. Vinacua. Virtual Reality Tools In Shipbuilding Design. In TeamCAD'97, pp.39-43, 1996.
- [2] F. Alonso, A. Cebollero, A. Gómez y A. Rodríguez. Collaborative Engineering in Shipbuilding, ICCAS'2002.
- [3] J. Durán, M.J. Puzas, F. Alonso y L. García. The use of a 3D product model orientated CAD/CAM System in a small/medium size shipyard. In CADAP'95, 1995.
- [4] L. García, V. Fernández y J. Torroja. The Role of CAD/CAE/CAM in engineering for production. In ICCAS'94, 1994.
- [5] A. Rodríguez, C. González, I. Gurrea y L. Solano. Kernel Architecture for the Development of CAD/CAM Applications in Shipbuilding Environments. In 2nd International Conference on Computer Applications and Information Technology in the Maritime Industries (COMPIT'2003), 2003.
- [6] A. Rodríguez, M. Vivo y A. Vinacua. New Tools for Hull Surface Modelling. In 1st International Euroconference on Computer Applications and Information Technology in the Maritime Industries (COMPIT'2000), pp. 400-411, 2000.
- [7] L. Solano, I. Gurrea y P. Brunet. Topological Constraints in Ship Design. Proceedings 4th IFIP WG 5.2 Workshop on Knowledge Intensive CAD KIC-4, U.Cugini and M.Wozny eds, Parma, 2000.

## Los nuevos currícula de Matemáticas en la Enseñanza Secundaria

R. RODRÍGUEZ DEL RÍO<sup>1</sup> E. ZUAZUA<sup>2</sup>

<sup>1</sup>Departamento de Matemática Aplicada  
Universidad Complutense de Madrid,

<sup>2</sup>Departamento de Matemáticas  
Universidad Autónoma de Madrid

rr\_delrio@mat.ucm.es, enrique.zuazua@uam.es

### Resumen

En este documento se recogen algunas de las ponencias presentadas por los participantes en las mesas redondas del Curso de Formación del Profesorado de Matemáticas en Secundaria “De la Aritmética al Análisis: los nuevos currícula de Matemáticas en la Enseñanza Secundaria” celebrado del 8 al 14 de julio de 2002 en el marco de los cursos de verano que la Universidad Complutense de Madrid organiza en San Lorenzo de El Escorial. Los participantes en estas mesas redondas y autores de las contribuciones que reproducimos fueron M. de León, F. Marcellán, T. Recio, S. Rodríguez y A. Ruíz.

### 1 Introducción

En este documento se recogen algunas de las ponencias presentadas por los participantes en las mesas redondas del Curso de Formación del Profesorado de Matemáticas en Secundaria “De la Aritmética al Análisis: los nuevos currícula de Matemáticas en la Enseñanza Secundaria” celebrado del 8 al 14 de julio de 2002 en el marco de los cursos de verano que la Universidad Complutense de Madrid y el Instituto Superior de Formación del Profesorado del MECD organizan en San Lorenzo de El Escorial.

La mayor parte del evento giró en torno a los cursos y conferencias que reseñamos seguidamente:

- *El Cálculo Diferencial y la Modelización Matemática*. José M. Arrieta (Universidad Complutense de Madrid)

- *El análisis estadístico de grandes masas de datos: algunas tendencias recientes.* Antonio Cuevas (Universidad Autónoma de Madrid)
- *Publicación de contenidos matemáticos en la web.* Agustín Muñoz (Centro Nacional de Información y Comunicación Educativa/MECD)
- *Una introducción al Álgebra y la Geometría Computacional.* Tomás Recio (Universidad de Cantabria)
- *Matemáticas, Ciencia y Tecnología: una relación profunda y duradera.* Juan Luís Vázquez (Universidad Autónoma de Madrid)
- *Las Matemáticas del Control.* Enrique Zuazua (Universidad Autónoma de Madrid)

El objetivo del curso fue el de proporcionar a los Profesores de Matemáticas de Enseñanza Secundaria una visión divulgativa de algunos de los problemas más relevantes de la matemática actual y las técnicas que más comunmente se emplean, incluidas las de carácter computacional, a un nivel que, resultando accesible a los Profesores de Secundaria, pudiese resultarles de utilidad en su tarea docente.

Los contenidos de estos cursos y conferencias serán publicados en breve en el libro *De la Aritmética al Análisis: Historia y desarrollos recientes en Matemáticas*, en la colección *Aulas de Verano* del Instituto Superior de Formación del Profesorado del MECD, editado por los abajo firmantes, organizadores del curso. De este modo, con este nuevo volumen hemos querido continuar con la labor que se emprendió con el primer volumen *Temas relevantes de la Matemática actual: el reto de la enseñanza secundaria*, editado por M. V. Pinillos y E. Zuazua, publicado por el MECD en el año 2000 y correspondiente al curso del mismo nombre impartido en la Universidad Internacional Menéndez Pelayo en septiembre de 1999, en el marco de los mismos cursos de Formación del Profesorado de Enseñanza Secundaria.

En esta ocasión, el volumen no recoge las contribuciones realizadas por los ponentes en las mesas redondas. Sin embargo, creemos que las reflexiones que en ellas se hicieron son sumamente interesantes y recogen en buena medida la complejidad y diversos puntos de vista sobre la encrucijada actual de la enseñanza de las Matemáticas en la Educación Secundaria y en general.

Hemos creído pues que merecía la pena recoger estas contribuciones y publicarlas en este Boletín.

En el curso se celebraron dos mesas redondas con los títulos respectivos *Educación Secundaria en el nuevo marco legal: realidad actual y perspectivas y Metodologías y contenidos para las Matemáticas de la Educación Secundaria*. Sin embargo, hemos creído más conveniente no distinguirlas en esta publicación y reproducir literalmente los textos aportados por los intervinientes en las mesas tras realizar los retoques que nos solicitaron y con su visto bueno.



Aprovechamos estas líneas para agradecer nuevamente a los ponentes de estas mesas por haber hecho posible, junto con alumnos y conferenciantes, el éxito del curso, y a los editores del Boletín por hacer posible esta publicación.

*Roberto Rodríguez del Río y Enrique Zuazua*

<http://www.mat.ucm.es/~rrdelrio> <http://www.uam.es/enrique.zuazua>

## 2 Ponencia presentada por Manuel de León (Consejo Superior de Investigaciones Científicas)

En los últimos meses el MECD está ultimando una reforma a fondo del sistema educativo (la denominada **Ley de Calidad**) que sustituirá a la LOGSE.

Hay un consenso bastante generalizado en la necesidad de una tal reforma de la LOGSE, fundamentalmente entre los propios profesores de secundaria.

La LOGSE consiguió un modelo de enseñanza obligatoria hasta los 16 años, una gran conquista social, cuyo objetivo era conseguir que la totalidad de la población española poseyera unos estudios mínimos, un escalón más arriba de lo que supuso en su momento superar el analfabetismo de una parte importante de la población. En otras palabras, se trataba de lograr una auténtica instrucción pública.

La universalización de la enseñanza secundaria llevó desgraciadamente a una bajón en el nivel de contenidos así como en su exigencia, al no ser capaces de superar el compromiso necesario entre una educación para todos y una educación que cubriera las expectativas de los más capaces. Las causas de este proceso son difíciles de discernir, y una de ellas probablemente haya sido la siempre escasa dedicación presupuestaria de las administraciones por la educación. Otras causas están a mi entender en el cambio sociológico experimentado en nuestro país en los últimos 20 años. En cualquier caso, como consecuencia, aspectos tan importantes como el esfuerzo personal y la creación de hábitos de trabajo fueron abandonados paulatinamente. Así, el nivel de formación de nuestros estudiantes ha ido descendiendo año tras año hasta alcanzar mínimos insostenibles. Existe un consenso bastante amplio sobre este diagnóstico.

En el caso particular de las ciencias, disciplinas que exigen un trabajo continuado a lo largo de todos los años de escolarización, la situación se agrava notablemente. Es paradigmático el caso de las Matemáticas, edificio que se va construyendo desde las primeras etapas de primaria (incluso ya desde la educación preescolar), y en cuya enseñanza el trabajo personal del alumno es imprescindible (por ejemplo, en la resolución de problemas).

Por una parte, los avances científicos conseguidos a lo largo del último siglo (teoría de la relatividad, mecánica cuántica, genoma humano, etc.) deben formar parte de una cultura elemental del ciudadano, sin la cual no entenderá su lugar en el universo, ni su papel como habitante de un planeta cuyos recursos naturales estamos agotando, ni entenderá su contribución indispensable a un desarrollo sostenible.

Por otra parte, la creciente presencia de las nuevas tecnologías en la vida del ciudadano corriente (especialmente las asociadas a la llamada sociedad de

la información) hace imprescindible una buena preparación científica de los escolares, que los haga capaces de desenvolverse conscientemente en la sociedad tecnológica del siglo XXI y no limitarse a ser meros consumidores de mágicos e incomprensibles procesos.

Además, si queremos que España ocupe un lugar preeminente en la nueva Europa, necesitamos una élite científica y tecnológica que desarrolle tecnología propia y que nos libre de una vez por todas de ser meros importadores.

Esta preocupación por la deficiente enseñanza científica en nuestras aulas y sus nefastas consecuencias a medio y largo plazo, llevó a las Reales Sociedades de Física, Matemáticas y Química a proponer la creación de una ponencia en el Senado cuyo objetivo fuera poner de manifiesto esta problemática, y extraer conclusiones con la elaboración de un libro blanco.

Al mismo tiempo, existía la intención de conseguir un acercamiento de las sociedades científicas con las instituciones políticas, puesto que entendíamos que aquellas son asesores naturales de las mismas, tal y como ocurre en los países avanzados.

La experiencia ha resultado enriquecedora e instructiva para ambas partes y recientemente la Comisión de Educación del Senado aprobó las conclusiones provisionales de la ponencia que han sido elevadas al MECD para su posible inclusión en la futura Ley de Calidad [1]. Confiamos en haber abierto un camino para que las sociedades científicas sean un contrapunto equilibrado en la lógica confrontación de ideas propia del debate político.

Quisiera a continuación apuntar algunas de las sugerencias que han surgido de la primera parte de esta ponencia aderezadas con opiniones personales pues, como todos sabemos, la educación secundaria es un tema que levanta polémicas y pasiones y en el que cada profesional de la enseñanza mantiene opiniones en general muy definidas. Permítaseme añadir que cuando el tema se analiza de cerca y con profundidad, las soluciones ya no se ven con tanta claridad. Se trata, no cabe duda, de un tema de gran dificultad y en el que resulta difícil (que no imposible) llegar a acuerdos y compromisos.

#### *Consideraciones generales*

- El preámbulo de la Ley de Calidad es demasiado utilitarista. La intención primaria es conseguir ciudadanos bien formados, y será esa buena instrucción la que les permitirá ser competitivos en la Europa comunitaria. En el caso de las matemáticas, esa concepción utilitarista puede llegar a poner en peligro las propias razones para su enseñanza, en una peligrosa confusión entre utilidad y utilitarismo, como se señala en el informe Kahane [2].
- Asignaturas como Física y Química deben ser divididas, son muy diferentes. Las horas de laboratorio son en estas asignaturas de una gran importancia para un buen entendimiento por parte del alumno. En el caso de las Matemáticas (como en el de la Lengua) debería impartirse una clase diaria, y elevar el nivel de contenidos.

- Deberían eliminarse una gran cantidad de asignaturas optativas. Las tecnologías, por ejemplo, podrían perfectamente constituir, en gran medida, partes de las asignaturas científicas.
- Asignaturas como la Religión deberían ser impartidas en el ámbito privado, en las instalaciones habituales del colegio, pero no como parte del currículum.
- Sería deseable una buena coordinación entre la informática y las Matemáticas. No se trata de adquirir meras destrezas informáticas (manejo de internet, procesado de textos), tareas que pueden ser desarrolladas perfectamente en las propias clases de ciencias.

#### *Alumnos y padres*

- Es preciso inculcar el hábito de trabajo y el valor del esfuerzo individual, bases del aprendizaje.
- La promoción automática debe ser suprimida, necesitamos que los alumnos valoren lo que aprenden, y que sepan demostrarlo en las diferentes evaluaciones.
- Son precisos grupos de refuerzo, pero no basados en la segregación de los alumnos según sus aptitudes, sino como clases adicionales.
- Es preciso abordar de una vez por todas, de manera decidida, la realidad de la inmigración. Deben formarse grupos de refuerzo para incorporar a estos alumnos (con grandes dificultades de idioma muchas veces) lo más rápidamente posible. Las matemáticas, con su universalidad como parte de cualquier cultura, pueden jugar un papel en esta integración.
- Se precisa una atención especial a los alumnos superdotados, un problema que va creciendo año tras año (un país no puede renunciar a formar sus élites, sería un desperdicio).
- No parece aconsejable la existencia de itinerarios antes de los 14 años.
- Las familias deben involucrarse mucho más en la educación de los hijos y considerar los centros educativos como auténticas prolongaciones del hogar familiar.
- Se precisa definir una Formación Profesional de gran calidad, que sea apetecible a muchos alumnos, de manera que a los 14 años pudieran comenzar en ella pero que posea también contenidos de cultura general, a la vez que se permitan pasarelas con el Bachillerato.

#### *Profesores*

- Se constata una deficiente formación en materias científicas (especialmente en Matemáticas) en los profesores de primaria. Debe pues aumentarse el

contenido científico, o, mejor todavía, lograr una mayor especialización entre este colectivo. Los profesores de primaria ponen las primeras piedras del edificio educativo y son por lo tanto una pieza esencial del sistema.

- Es deseable que los mismos profesores impartan la docencia en los diferentes ciclos educativos, consiguiendo así una confianza por parte de los alumnos y un mejor conocimiento de las cualidades de estos por los docentes.
- Se precisa también una mejor definición de la carrera docente, que estimule a los profesores.
- Se necesita acabar con la situación de interinidad que viven los profesores que han logrado sus oposiciones y que se ven obligados a cambiar año tras año de instituto hasta lograr un destino definitivo. Es una situación totalmente incomprensible.
- La formación continuada de los profesores es una tarea clave, tanto en el caso de primaria como en el de secundaria. Sería deseable una mejor estructuración de los cursos de formación, y una mayor interacción entre los centros de las diferentes Comunidades Autónomas. La asistencia a congresos especializados tendría que ser estimulada, actualmente, los profesores de secundaria deben dedicar sus fines de semana a estas actividades.
- Es preciso reforzar la autoridad del profesor, dotándole de mecanismos que hagan posible su trabajo.

### *Finalmente*

- Es imprescindible que la reforma vaya acompañada de un incremento del gasto educativo, estamos a la cola de los países europeos en tanto por ciento del PIB y hay margen para este aumento. Existe una incongruencia en el mensaje que llega de nuestros gobiernos sobre la importancia de la educación para el futuro de España y el capítulo correspondiente al gasto educativo en los presupuestos generales.

### **Referencias**

[1] Información sobre la ponencia en la página web de la Real Sociedad Matemática Española: <http://www.rsme.es> (En esta web se pueden encontrar numerosos documentos elaborados por diversas sociedades matemáticas y la Real Academia de Ciencias relativos a la Ley de Calidad).

[2] L'Enseignement des Science Mathématiques, Rapport au ministre de l'Éducation nationale. Sous la direction de Jean-Pierre Kahane. Centre National de Documentation Pédagogique et Editions Odile Jacob, 2002.

### 3 Ponencia presentada por Francisco Marcellán (Universidad Carlos III de Madrid)

A la hora de enfocar el tema de la enseñanza de las Matemáticas en Educación Secundaria considero necesario incorporar el análisis de varios parámetros que no se limitan solo a metodología y contenidos. Básicamente son los siguientes:

1. ¿Qué objetivos se plantean acerca del papel y contribución de las Matemáticas en una enseñanza obligatoria que refleja tres principios básicos como son la comprensividad (currículum básico y común, retraso en la selección y especialización, ampliación del límite de edad escolar), la atención a la diversidad (medidas ordinarias y extraordinarias de atención y respuesta a las diferentes capacidades intelectuales e intereses de los alumnos) y la equidad (compensación de desigualdades y promoción de la igualdad de oportunidades educativas independientemente de los recursos socioeconómicos del alumno)?.
2. De acuerdo con los objetivos anteriores y los principios sustentadores del sistema de enseñanza, ¿cuáles deben ser los contenidos que se deben contemplar en la elaboración de un currículum que contemple tanto mínimos como desarrollos acordes a las diversidades de los alumnos?
3. ¿Qué metodología se debe seguir y cómo evaluar el proceso de aprendizaje de los alumnos? En una dirección complementaria, parece natural preguntarse acerca de qué características deben reunir los docentes desde el punto de vista formativo inicial como qué estímulos e incentivos se aplican para la mejora de su aportación al proceso de aprendizaje así como la evaluación de su actividad.

#### **En torno a objetivos.**

En documentos oficiales ([1]) relativos a la enseñanza de las Matemáticas en la ESO se manifiesta que “la finalidad fundamental de la enseñanza de las Matemáticas es el desarrollo de la facultad de razonamiento y de abstracción.... así como su carácter instrumental”. A continuación se describen ocho grandes objetivos que a efectos de discusión cito literalmente:

1. Utilizar las formas de pensamiento lógico en los distintos ámbitos de la actividad humana.
2. Aplicar con soltura y adecuadamente las herramientas matemáticas adquiridas a situaciones de la vida diaria.
3. Usar correctamente el lenguaje matemático con el fin de comunicarse de manera clara, concisa, precisa y rigurosa.
4. Utilizar con soltura y sentido crítico los diversos recursos tecnológicos (calculadoras, programas informáticos) de forma que supongan una ayuda en el aprendizaje y en las aplicaciones instrumentales de las Matemáticas.

5. Resolver problemas matemáticos utilizando diferentes estrategias, procedimientos y recursos, desde la intuición hasta los algoritmos.
6. Aplicar los conocimientos geométricos para comprender y analizar el mundo físico que nos rodea.
7. Emplear los métodos y procedimientos estadísticos y probabilísticos para obtener conclusiones a partir de los datos recogidos en el mundo de la información.
8. Integrar los conocimientos matemáticos en el conjunto de saberes que el alumno debe adquirir a lo largo de la ESO.

Después de este compendio, me permito plantear unas preguntas elementales: ¿Son conocidos y asumidos en la práctica estos objetivos por profesores y alumnos? ¿Representan la realidad del aprendizaje o no dejan de ser deseos de difícil cumplimiento teniendo en cuenta las motivaciones de los alumnos y la influencia de los valores dominantes socialmente?

#### **A la búsqueda de los contenidos.**

Si están correctamente definidos los objetivos, los contenidos deberían ser una consecuencia lógica de los mismos. Curiosamente aparecen cuatro grandes ejes naturales como son

1. Aritmética y Álgebra.
2. Geometría.
3. Tablas, funciones y gráficas.
4. Estadística y probabilidad.

No obstante, la jerarquización de mínimos no está definida y se denominan criterios de evaluación una serie de items que más bien son líneas de concreción curricular.

Mi cuestión básica es la acomodación de dichos contenidos al origen y a la evolución de los problemas, la dificultad de los mismos y los medios para abordarlos. Asimismo, la búsqueda del utilitarismo frente al valor del aprendizaje en sí, condiciona de una manera clara la formalización y la integración de los contenidos en el curriculum global. ¿En qué medida los contenidos antes expuestos permiten sacar a las Matemáticas de la torre de marfil que históricamente se le ha asignado en la enseñanza secundaria? ¿Qué debe aprenderse por sí o para sí en relación con otras asignaturas no solo científicas (con la aparente paradoja del menor reconocimiento y valoración por parte de una sociedad basada en el uso intensivo de la Ciencia y la Técnica) sino humanísticas y sociales? Finalmente, ¿en qué medida los contenidos se acomodan a actitudes sociales y permiten incrementar las aptitudes instrumentales de los alumnos ante el mundo que les rodea?

### **El inexistente discurso del método.**

Mi gran sorpresa ante las directrices oficiales anteriormente indicadas radica en que la discusión sobre las cuestiones metodológicas se limita a “la adaptabilidad a cada grupo de alumnos y situación, rentabilizando al máximo los recursos disponibles. Como criterio general, parecen aconsejables las actuaciones que potencien el aprendizaje inductivo, sobre todo durante los primeros años de la etapa, a través de observación y manipulación y refuercen, al mismo tiempo, la adquisición de destrezas básicas, esquemas y estrategias personales a la hora de enfrentarse ante una situación problemática cercana al alumno, sin perder de vista la relación con otras áreas del currículo”. La metodología debe significar estimular la experimentación y la curiosidad, dar a los docentes medios para ello, validar a través del contraste lo que se ha desarrollado. El papel del profesorado ante ello es central, debiéndose perfilar claramente su formación de partida (prerrequisitos científicos en el más amplio sentido), las estrategias para el aprendizaje permanente así como la adaptación al intercambio de experiencias. Todos estos elementos constituyen condiciones necesarias para la potenciación de los aspectos metodológicos que no pueden limitarse al voluntarismo individual sino que deben potenciar la relación entre los diferentes niveles educativos, las disponibilidades materiales en los centros y la orientación de los centros de formación de los titulados que imparten la docencia en Matemáticas. Mi criterio básico es que para enseñar Matemáticas hay que saber vivir las Matemáticas y conocer sus interioridades conceptuales con el fin de poderlas transmitir adecuadamente a personas que deberían incorporarlas de manera natural y consecuente a su formación integral como ciudadanos libres. En esta dirección la formación de los profesores de Matemáticas debería constituir uno de los ejes prioritarios de actuación no solo de las administraciones públicas sino de los colectivos profesionales (sindicatos, sociedades matemáticas, etc) que deben adquirir un papel central en el proceso. Como ejemplo de dicha preocupación me remito al documento [2] elaborado por la Comisión de Educación de la Real Sociedad Matemática Española.

Finalmente, me gustaría resaltar que el aprendizaje por parte de los alumnos no debe limitarse a las horas de aula dirigidas por el profesor (en clara competencia con otras materias del currículo escolar) sino que se debe fomentar el trabajo en equipo para un mayor contraste y autoevaluación del propio aprendizaje utilizando tecnologías de información y comunicaciones que permiten el trabajo en red. El nivel de esfuerzo en el aprendizaje de las Matemáticas exige la capacidad de contrastar los avances y dificultades de manera adecuada y proporcionar al alumno y al profesor los instrumentos para la evaluación de los mismos. La pasión por descubrir y entender la realidad no debe estar alejada de toda esta concepción de las Matemáticas más allá del utilitarismo o de la autocontemplación.

### **Del Bachillerato a la Universidad.**

Una última consideración, basada en mi experiencia cotidiana como profesor universitario, me lleva a formular la necesidad antes apuntada de una visión global del proceso educativo como condición sine qua non de nuestro trabajo

profesional junto a una permanente relación con profesores de los diferentes niveles educativos.

En primer lugar, constato que, en su mayor parte, los profesores de Universidad que imparten las asignaturas de Matemáticas en el primer año de las diferentes titulaciones, desconocen el nivel previo de formación de los alumnos y utilizan metodologías que no reflejan la evolución real del aprendizaje. Ello se traduce en lenguajes inasequibles para una mayoría de los estudiantes, una huída hacia adelante ante el fracaso académico (que se traduce en un elevado número de suspensos y abandonos, enseñanza paralela extrauniversitaria), desmotivación y ausencia de incentivos para la renovación metodológica y pedagógica por parte del profesorado universitario, escasa utilización de tecnologías de información y comunicaciones (pizarra y tiza o, en su versión “moderna”, rotulador, transparencias, etc., escasamente interactivos).

En segundo lugar, existe una escasa implicación en la formación del profesorado de Secundaria y en su actualización tanto en contenidos como en metodologías de aprendizaje. Los modelos “académicos” basados en actividades de postgrado no resultan especialmente atractivos y habría que plantear modelos alternativos que reflejaran análisis de realidades y respuestas a retos evidentes e inmediatos.

Finalmente, la pérdida de la inocencia de los estudiantes ante la validación social de la formación científica y la incorporación de los valores del mercado de trabajo por encima de otras consideraciones está ocasionando una notable disminución de las vocaciones en las áreas científicas y, en particular, en la titulación de Matemáticas, que pueden condicionar gravemente el futuro de la docencia de dicha disciplina en los niveles obligatorios.

### Referencias

- [1] Matemáticas de la Educación Secundaria (ESO).  
[http://www.mec.es/gabipren/documentos/realdecreto/eso\\_matemat.pdf](http://www.mec.es/gabipren/documentos/realdecreto/eso_matemat.pdf)
- [2] La situación de la enseñanza de las Matemáticas: Un documento inicial. Real Sociedad Matemática Española. 2002.

## 4 Ponencia presentada por Tomás Recio (Universidad de Cantabria)

1. *¿Cómo va a afectar la futura Ley de Calidad a la enseñanza en general?*

En general, cuando las Administraciones realizan un esfuerzo especial en un sentido, como, por ejemplo, con motivo de una nueva Ley, el resultado suele ser una mejora de muchos aspectos. No necesariamente por el contenido concreto del articulado, sino por el especial esfuerzo humano y material que se pone en tales circunstancias, por la renovación de las ilusiones de los agentes involucrados en su puesta en marcha y por las expectativas que se crean en el conjunto de la sociedad .

Desde esta visión (que puede calificarse de optimista o de pesimista, según se mire) de la eficacia de los actos administrativos, pienso que



la nueva Ley puede contribuir a una mejora de las enseñanzas. Otra cosa es que se piense que algunos problemas no se van a resolver o no se van a resolver bien, o que se podrían resolver mejor o que se van a crear nuevos problema... Pero también hay que considerar que una Ley tan amplia tendrá, necesariamente, luces y sombras. Lo verdaderamente importante son los Reglamentos, es decir, el desarrollo normativo y la práctica implantación de las medidas que establece la Ley. Es, por tanto, pronto para juicios tajantes al respecto. Por ejemplo, la disposición de este proyecto de Ley que señala la obligación que contrae la Administración para ofertar un número suficiente de plazas gratuitas en Educación Infantil, puede desarrollarse en el sentido de dotar a la red de titularidad pública de una cantidad abundante de tales plazas; o en el sentido de concertar con centros de titularidad privada todas aquellas plazas que ahora mismo no pueda ofrecer la Administración en los centros de su titularidad... Y no es lo mismo.

En el caso de las Matemáticas la ley contiene elementos de progreso tales como:

- la renovación del caduco modelo de formación inicial del profesorado
- la consideración de los alumnos con facultades extraordinarias

puntos todos ellos que fueron en su momento planteados en el documento de la RSME sobre la situación de la enseñanza de las matemáticas (La Gaceta de la RSME, Vol. 4, no. 3, 2001). Otros aspectos considerados en ese documento no han sido, de momento, contemplados en la Ley (lo que no significa que no puedan ser tenidos en cuenta en el desarrollo futuro de la misma).

2. *Los itinerarios (separar a los alumnos según su rendimiento) en la ESO: ¿convenientes o contraproducentes?*

En el documento de la RSME antes citado se hace referencia explícita a la necesidad de atender a la diversidad de intereses, capacidades y situaciones “mediante actuaciones regladas y con dotación de medios suficientes, especialmente en aquellos centros donde, por razones socio-económicas o de otro tipo, las situaciones problemáticas tengan una especial incidencia”. De nuevo, la pregunta es incontestable hasta que no se constate en el desarrollo normativo y práctico si, efectivamente, el desarrollo de itinerarios va a significar una real atención a la diversidad, es decir, enseñar mejor a todos (y no simplemente enseñar a los mejores).

En todo caso, como se señalaba también en el documento precitado, “no basta con la segregación: sería preciso en todo caso diseñar un currículo adecuado... dotar a sus centros de medios idóneos para ese currículo, formar a sus profesores en los retos específicos de su enseñanza, promover socialmente la imagen y la función de los itinerarios alternativos”.

Pero no parece estar en las prioridades ministeriales la urgencia de una modificación de contenidos (argumentando que se ha aprobado

recientemente una modificación del currículo, dic. 2001) ni de la formación inicial del profesorado (se está contemplando, en el momento de redactar estas líneas, la (nueva) prórroga del CAP hasta el 2005).

### 3. *Formación inicial y continua del Profesorado de Matemáticas.*

Desde luego los mecanismos que preconiza la Ley de Calidad para la formación inicial son sensiblemente mejores que los actuales, al menos sobre el papel, puesto que pretenden dotar al futuro profesor de una formación más completa en aquellos aspectos que, fuera del ámbito estricto de su disciplina, son cada vez más necesarios en el complejo mundo de la educación.

Sin embargo el anteproyecto de Ley no hace referencia, por ejemplo, a una modificación substancial de la formación inicial de los maestros, que tiene un peso más importante, si cabe, que la de los futuros profesores de secundaria, si pensamos en la formación matemática de los españoles. Tampoco hace referencia al modo de encardinar la formación inicial de los profesores de matemáticas en los planes de estudio de las Licenciaturas de Matemáticas, justamente cuando estas están abocadas a determinadas modificaciones requeridas por la convergencia europea en educación superior (Proyectos Tuning o CRUE). De nuevo habrá que esperar al desarrollo normativo correspondiente, esperando que se atiendan a estas y otras cuestiones no dilucidadas en el texto que hoy conocemos.

El modelo de formación continua no está apenas desarrollado, pero parece existir un excesivo intervencionismo de la Administración (tal vez como contrapartida a unos años de autogestión en este ámbito). La Universidad, en mi opinión, permanece alejada (por voluntad propia y por olvido ajeno) de este tema.

### 4. *Selección del profesorado: el sistema de oposición actual, (re)creación del Cuerpo de Catedráticos*

Dada la situación de la pirámide poblacional española este parece ser un tema de menor trascendencia. En todo caso, y como previsión de futuro, parece razonable insistir en que el mejor sistema de selección es aquel que atrae a los mejores, para lo cual es muy importante dignificar y valorar económica y socialmente la profesión de profesor. La Ley no avanza significativamente en este tema (o tal vez lo confunde con, o lo reduce a, un problema de disciplina escolar). La carrera docente, en la Ley, tiende excesivamente a considerar como meta suprema la Inspección Educativa y no debería ser así, sino que tendría que considerar la posibilidad de un progreso profesional “in situ”, obtenido por mérito y rendimiento propios. En lo que respecta al cuerpo de catedráticos, me temo que las disposiciones normativas, que establezcan las condiciones para la incorporación como catedráticos de una parte de los actuales profesores, se enmarañarán en una tupida red de derechos adquiridos, situaciones especiales, etc... que contribuirán poco al prestigio que se pretende tenga ese nuevo cuerpo.

Es una medida que llega, posiblemente, tarde. Desde otra perspectiva se me ocurre plantear si ¿se encargarán los catedráticos más prestigiosos de los itinerarios curriculares más complicados (al igual que los médicos más valorados son los que tratan los casos más difíciles)? Debería haber una auténtica carrera docente que no consista en ir eludiendo el aula... y en ella convendría enmarcar esta recreación del cuerpo de catedráticos.

## 5 Ponencia presentada por Soledad Rodríguez (Universidad Complutense de Madrid)

### 1. Metodologías.

- ¿Qué es una metodología para las matemáticas en Secundaria? ¿Cuántas hay? ¿Son distintas de las metodologías para las matemáticas en la Primaria y en la Universidad?
- Martin Gardner, en un artículo publicado en el New York Times Reviews el 24 de septiembre de 1998, titulado “The New New Math”, habla de una nueva metodología para la enseñanza de las matemáticas, que en nuestro país se ha implantado con el nombre de “método de resolución de problemas”. ¿Son tan grandes los riesgos de este método como denuncia Gardner? ¿En la fecha de hoy, podemos hablar, no ya de riesgos, sino de desastres? ¿Hasta qué punto está influyendo esta metodología en la formación académica de los estudiantes?
- Un poco de historia: ¿Cómo se enseñaban las matemáticas en la educación secundaria de nuestro país en los años 60, 70, 80 y 90? Aquella historia de “los conjuntos”, con esas generaciones que aprendieron de memoria lo que era una aplicación inyectiva y un anillo de polinomios, ¿para qué valió? ¿Murió de muerte natural, por lento contagio de lo que pasaba en otros países, o influyeron los matemáticos españoles en alguna dirección, bien alargando la vida de “los conjuntos” en la escuela, bien clavándoles el estoque? ¿Podemos hablar de alguna analogía con la situación actual? ¿Existirá ahora algún Morris Klein que publique algún libro similar a su tan leído y traducido “¿Por qué Juanito no sabe sumar? El fracaso de la matemática moderna” (Alianza Editorial) ¿Debemos involucrarnos los profesores de matemáticas de cualquier nivel en esto? Y los no matemáticos, científicos en general, técnicos, economistas, etc., ¿no tendrán también algo que decir? O, ¿debemos delegar todos en los autodenominados “investigadores en educación matemática”?

### 2. Contenidos.

- ¿Resulta conveniente que estén claramente especificados los contenidos mínimos?
- En la enseñanza universitaria, ¿están especificados?

- ¿Quién se atreve a contestar a las siguientes preguntas:
  - i) ¿qué hay que consolidar en las matemáticas básicas?
  - ii) ¿qué conocimientos debe tener un alumno al acabar la E.S.O.?
  - iii) ¿y al acabar el bachillerato?

## 6 Ponencia presentada por Andrés Ruíz (I.E.S. San Isidro (Madrid))

### Lo que hay

Las condiciones sociales han creado un ambiente que no favorece el esfuerzo y dedicación de los alumnos. La LOGSE, no sólo no ha ido contra esta nefasta corriente sino que la ha fomentado. La cultura del “pelotazo” arraigada en España a partir del 85, se ha visto reflejada en la LOGSE: se pueden conseguir buenos resultados sin trabajar.

La LOGSE mantiene la ilusión de que los objetivos de la ESO, en los casos de malos estudiantes que a los 16 años se encuentran todavía en 3º, “se alcanzarán con una metodología específica, a través de contenidos e incluso de áreas diferentes”. No hace falta ninguna experiencia para saber que se trata de una falsedad. Los pocos alumnos que han obtenido el título de Graduado en Educación Secundaria con este método, lo han hecho a base de una notable rebaja de contenidos.

La atención a la diversidad comporta la idea de adaptación curricular personalizada. Estas adaptaciones pueden ser “significativas”, es decir, con una rebaja de contenidos lo suficientemente amplia como para hacerlos alcanzables para los alumnos. El desarrollo de la LOGSE llega en Real Decreto 1345/91 a la locura de exigir adaptaciones de los mismísimos programas de diversificación para cada uno de los alumnos (Art. 17.3).

Con estas adaptaciones personalizadas, cada aula requiere el mismo trabajo que las llamadas antiguamente unitarias, en las que un maestro rural se las tenía que ver durante todo el día con alumnos desde los 6 hasta los catorce años.

La idea central de la LOGSE es que los conocimientos son privatizables por los alumnos a base de las adaptaciones curriculares y, en definitiva, se pueden “personalizar”. Y este es precisamente el gran error porque, **los contenidos de enseñanza, como tales, no pueden adaptarse a las idiosincrasias o condiciones personales de los alumnos, sino que deben ser estos los que tienen que adaptarse a la impersonal condición de los conocimientos.**

Lo que sí cabe en la escuela pública, como se ha hecho siempre por quien ha podido, es instaurar de modo sistemático los apoyos, eufemismo con que se denominan las clases particulares, a los alumnos que, aún esforzándose, no consiguen alcanzar los objetivos.

Por lo que respecta a la disciplina, esta idea, así como la de sanción, han desaparecido de los textos legales que se utilizan para poner orden en los centros. Ahora se habla de conductas contrarias a las normas de convivencia y de correcciones, en lugar de faltas de disciplina y sanciones. Si un alumno tiene

una conducta que impide el desarrollo de una clase se le llama hiperactivo. Para sancionar una falta grave hay que montar un aparato análogo al que se exige para juzgar delitos mayores y poner de acuerdo al jurado formado por los miembros del Consejo Escolar.

### **Lo que viene**

La ley de Calidad abandona la idea de adaptación curricular. En su Art. 41, en la Sección sobre alumnos con necesidades educativas especiales, habla de “determinados apoyos y atenciones educativas específicas” y nada de adaptaciones curriculares personalizadas. En el Art. 39 se ocupa de los alumnos extranjeros y propone programas específicos para su incorporación al sistema.

La organización en itinerarios parece la única diversificación legítima con objeto de que cada alumno elija lo que más le conviene.

Respecto del Bachillerato, exige, además de superar todas las asignaturas, una Prueba General. La LOGSE, en Art.29 decía que “para obtener este título será necesaria la evaluación positiva en todas las materias”. Parece que ahora la condición sigue siendo necesaria pero no suficiente. Esto los matemáticos lo entendemos bien. La LOGSE exige, además, superar una prueba de acceso para entrar en la universidad. Este es uno de los temas de debate actual.

Respecto de la vida en los centros llama la atención el hecho de que los Consejos Escolares no serán los encargados de imponer correcciones por conductas gravemente contrarias a las normas de convivencia, sino los directores. Esto podría ser el principio de una reforma sensata, pero queda por ver si continúa el actual sistema de expedientes disciplinarios.

### **Los contenidos**

Parece un problema insuperable el hecho de que un gran porcentaje de alumnos, mayor entre los que no obtienen el título de Graduado en Educación Secundaria, pero relevante entre los que siguen el Bachillerato y continúan en la Universidad no adquieran jamás lo que podríamos llamar saberes o conocimientos primordiales o básicos, a saber:

- Leer, escribir y hablar correcta y fluidamente el español.
- Realizar cálculos aritméticos, manejar la proporcionalidad (regla de tres) y conocer los cuerpos geométricos, aprendiendo a “cuadrar” y “cubicar” superficies y volúmenes.
- Situarse en el espacio y en el tiempo (rudimentos de geografía e historia), sobre todo en su entorno próximo.
- Conocer algunos aspectos de funcionamiento de la naturaleza y del propio cuerpo.

Es esta una lista que habría que consensuar, pero que una vez aceptada debería servir para evaluar la labor de la Escuela y para permitir que un alumno obtuviera el título de Graduado en Educación Secundaria, sin tener que pasar

forzosamente por todo lo que se le propone como obligatorio (¿para qué forzar de manera universal la enseñanza de los rudimentos del álgebra?).

El Título de Graduado debería estar al alcance de todos y debería incluir con algunos retoques lo que enseñaban los maestros en las escuelas, antes de la implantación de la EGB (Ley de 1970). ¿Qué sentido tienen esos cursos con diez o doce asignaturas, cuando no se aprende lo esencial? El problema de los contenidos de la ESO sigue en pie. Los actuales no harán sino agravar el fracaso escolar de los que quieren sólo obtener el Graduado y, al mismo tiempo, no preparar adecuadamente a los que quieren continuar con el bachillerato.

## Referencias

- [1] <http://www.mat.ucm.es/~rrdelrio/sanlorenzo2002/escorial2002.html> *En esta página web se encuentran todos los materiales relativos al curso.*

<b>Título:</b>	PROBLEMAS DE CONTORNO NO LINEALES DE TIPO ELÍPTICO.
<b>Doctorando:</b>	David Ruiz Aguilar.
<b>Director/es:</b>	Antonio Cañada Villar.
<b>Defensa:</b>	19 de Septiembre de 2002, Universidad de Granada.
<b>Calificación:</b>	Sobresaliente cum laude por unanimidad.

**Resumen:** Motivados por algunas aplicaciones a Ingeniería, Mecánica, Astrofísica y Geometría Diferencial, en esta tesis se estudian diversos problemas de contorno resonantes de tipo elíptico.

El capítulo primero se dedica al estudio de un problema resonante en el valor propio principal, con condiciones de contorno tipo Dirichlet y términos no lineales de tipo casiperiódico. Se aportan resultados novedosos, tanto desde el punto de vista cualitativo como cuantitativo, sobre la descripción del rango de los operadores no lineales que surgen en este tipo de problemas. El segundo capítulo trata sobre sistemas resonantes de ecuaciones con términos no lineales de tipo oscilatorio. En este caso la principal dificultad es el carácter multidimensional del núcleo del operador lineal asociado. El capítulo tercero describe nuevos resultados para problemas no lineales asociados al operador  $p$ -Laplaciano y condiciones de contorno tipo Neumann. El último capítulo contiene problemas variacionales asociados a problemas elípticos no lineales con exponente crítico y potencial singular tipo Hardy. Se presentan también algunos resultados para el caso de dominios no acotados y dominios invariantes por la acción de ciertos subgrupos de movimientos rígidos del espacio euclídeo  $n$  dimensional.

Las demostraciones usan diferentes métodos del Análisis no lineal, especialmente el método alternativa (o reducción de Liapunov-Schmidt), grado topológico, teoría de puntos críticos, desigualdades tipo Hardy y Liapunov, teoría de la medida, etc.

<b>Título:</b>	NONLINEAR BOUNDARY VALUE PROBLEMS.
<b>Doctorando:</b>	Antonio Jesús Ureña Alcázar.
<b>Director/es:</b>	Antonio Cañada Villar.
<b>Defensa:</b>	5 de Noviembre de 2002, Universidad de Granada.
<b>Calificación:</b>	Sobresaliente cum laude por unanimidad.

**Resumen:** En esta tesis doctoral se tratan problemas de contorno resonantes no lineales para Ecuaciones Diferenciales Ordinarias y Ecuaciones en Derivadas Parciales.

El primer capítulo está dedicado al estudio de problemas de contorno tipo péndulo no autoadjuntos, para condiciones tipo Dirichlet y donde la principal novedad estriba en la consideración de términos lineales que incluyen la derivada de la función incógnita (fricción lineal). Usando el método alternativa (o reducción de Liapunov-Schmidt), así como lemas generalizados del tipo Riemann-Lebesgue, el autor consigue hacer un estudio preciso de las propiedades fundamentales de la ecuación de bifurcación, poniendo de manifiesto las similitudes y diferencias básicas con el correspondiente problema de contorno periódico. En el segundo capítulo se abordan problemas similares para Ecuaciones en Derivadas Parciales. Usando métodos topológicos y variacionales se pone de manifiesto cómo la dimensión del problema puede influir claramente en el tipo de resultados obtenidos. En el tercer capítulo se aportan resultados novedosos sobre multiplicidad de soluciones para el problema de contorno periódico en ecuaciones tipo péndulo. Dentro de la filosofía de la teoría de Bifurcación, aquí se combinan diversas herramientas como la mencionada reducción de Liapunov-Schmidt, grado topológico de Leray Schauder y método de soluciones superiores e inferiores. El último capítulo se dedica al estudio de sistemas de ecuaciones diferenciales ordinarias con operadores del tipo  $p$ -Laplaciano y con condiciones de tipo Hartman-Nagumo sobre los términos no lineales.