

Boletín de la Sociedad Española de Matemática Aplicada SĒMA

Grupo Editor

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

Comité Científico

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

Responsables de secciones

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

Página web de SĒMA

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

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

ISSN 1575-9822

Depósito Legal: AS-1442-2002

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

Diseño de portada: Ana Cernea

Estimados amigos:

Tenéis en vuestras manos el número 30 del Boletín SEMA, que será el último que realice el actual Grupo Editor. Un grupo de compañeros de la Universidad de Salamanca han accedido con gran amabilidad a hacerse cargo de la edición en el futuro inmediato. Desde aquí, queremos expresar nuestro agradecimiento a todos ellos y desearles el mayor éxito posible. No nos cabe ninguna duda de que, dada la calidad y el empeño que siempre han estado presentes en su trabajo, llevarán a cabo la labor de edición a la perfección y corregirán los errores y las limitaciones que hayamos tenido.

Con éste han sido once boletines y un anuario, los que hemos editado desde abril de 2002. En el presente número, aparte de las contribuciones habituales, hemos conseguido incluir dos artículos de opinión, tres artículos de investigación, dos artículos correspondientes a los premiados con el *V Premio SĒMA a la Divulgación de la Matemática Aplicada* y el *VII Premio SĒMA al Joven Investigador* (P. Fernández y M. A. Fontelos, respectivamente) y un artículo adicional, dentro de la sección *Matemáticas e Industria*.

En esta última intervención, queremos agradecer a todos la colaboración que, con gran generosidad, habéis tenido con nosotros. En particular, queremos expresar nuestro agradecimiento a todos los responsables de secciones y, muy especialmente, a los distintos autores, por sus aportaciones y por la comprensión que han mostrado para facilitar nuestra labor.

Deseamos también pedir disculpas por los fallos que hayáis podido observar. Sin lugar a dudas, lo poco que hemos podido hacer ha sido gracias a todos vosotros. Para nosotros, ha sido un gran placer.

Queremos, desde estas líneas, felicitar al Presidente saliente, Eduardo Casas, por su gran labor y dedicación a nuestra Sociedad durante los últimos cuatro años. Al recién elegido Presidente de SĒMA, Juan Ignacio Montijano, y a los nuevos miembros del Consejo Ejecutivo les deseamos mucha suerte en esta nueva etapa en la vida de la Sociedad Española de Matemática Aplicada.

No queremos despedirnos sin recordar y agradecer la labor realizada por los Grupos Editores que nos precedieron, compañeros de Málaga, Zaragoza y Córdoba. Ellos han gestado y perfilado la estructura actual de nuestro Boletín, al que deseamos una buena evolución y larga vida.

Un cordial saludo,

Javier Valdés, Enrique Fdez-Cara, Benjamín Dugnol,
Mariano Mateos, Omar Menéndez y Pablo Pérez

Grupo Editor

Queridos compañeros, hace cuatro años me dirigía a vosotros a través de este boletín de SĚMA con el fin de agradecer la confianza que depositabais en mí para presidir la Sociedad. Hoy escribo estas líneas para agradecer vuestra colaboración durante este tiempo en el que fui presidente. Ha sido un honor para mí presidir la sociedad. Lamento si en algún momento no cumplí con las expectativas que pusisteis en mí. Os puedo asegurar que intenté hacerlo lo mejor que pude y el tiempo, que tuve que compaginar con muchas otras actividades, me permitió. Pero no debo ser yo quien haga balance de estos cuatro años, eso os corresponde a vosotros.

Quiero aprovechar esta circunstancia para agradecer la dedicación a SĚMA de todas aquellas personas que colaboraron estrechamente conmigo, muy especialmente a Rosa Pardo, secretaria de la Sociedad durante estos cuatro años y alguno más con nuestro anterior presidente, Enrique Fernández Cara. Ella fue una gran ayuda. También mi agradecimiento al tesorero, Luis Alberto Fernández, quien puso orden en las cuentas de SĚMA. Es justo reconocer el importante papel desempeñado por el Consejo Ejecutivo: gracias a todos sus sucesivos miembros. Tampoco puedo olvidar a los editores de nuestro boletín, al grupo de Córdoba que dirigió José Luis Cruz y después al de Oviedo que encabezó Javier Valdés. A nadie se le oculta que la edición del boletín es una de las actividades más importantes de SĚMA, por lo que todos nos sentimos en deuda con los editores, los señalados y los que anteriormente asumieron esta tarea.

Por último os pido vuestro apoyo para el nuevo presidente, Juan Ignacio Montijano. Estoy convencido que será un buen presidente y que dejará SĚMA en mejores condiciones que ahora ha encontrado, pero para ello precisará de la ayuda de todos nosotros. Monti, muchas gracias por asumir esta tarea y mucha suerte, que será la suerte de todos nosotros.

Eduardo Casas

Esta nueva edición del Boletín me permite llegar a todos los miembros de la Sociedad Española de Matemática Aplicada para expresar mi agradecimiento por la confianza que habéis depositado en mí al aceptarme como Presidente. En el escaso tiempo que llevo como tal he podido comprobar la buena disposición de los colegas y el sincero apoyo que me brindan.

Inicio esta etapa con enorme orgullo y satisfacción, contando con el bagaje y la experiencia de lo hecho en nuestros más de diez años de historia. SĒMA es hoy una sociedad matemática rica, obviamente no en términos económicos, sino por su capital humano, de indiscutible calidad científica, que gracias a la ilusión y al empuje de todos ha alcanzado un grado de madurez y un prestigio nacional e internacional considerable. Así, por ejemplo, hoy en día la participación de SĒMA en la organización de cualquier evento es un aval de calidad y seriedad científica.

Una parte importante de este éxito debemos achacarla a la labor de todos aquellos que de una manera directa se han implicado en las tareas de gestión de la Sociedad. En particular quiero mostrar mi agradecimiento a Eduardo Casas, nuestro Presidente saliente, a Rosa Pardo y José Luis Cruz, que han finalizado su labor como miembros del Consejo ejecutivo, y a Luis Alberto Fernández, que deja la tesorería de la sociedad. También terminan su etapa al frente de la edición del Boletín el grupo de la Universidad de Oviedo. Nuestra felicitación y agradecimiento por el excelente trabajo realizado. Para hacerse cargo de esta tarea se han comprometido Luis Ferragut Canals, María Isabel Asensio Sevilla, Antonio Martínez Fernández, María Teresa de Bustos Muñoz y Francisco Andrés Pérez, de la Universidad de Salamanca. Nuestra gratitud por su ofrecimiento y por el esfuerzo y el tiempo que sin duda van a dedicar a la sociedad.

Al mismo tiempo que yo, se han incorporado al Consejo ejecutivo José Antonio Carrillo (Universidad Autónoma de Barcelona) y Carlos Castro (Universidad Politécnica de Madrid), quien pasa a ser el nuevo Secretario. De la tesorería se hará cargo Pilar Laburta (Universidad de Zaragoza). No dudo que la aportación de todos ellos será muy beneficiosa para la Sociedad.

Nos une a todos el interés por las Matemáticas y su relación con las aplicaciones, y entre todos tenemos que hacer que SĒMA continúe su progresión ascendente. Desde el consejo ejecutivo nos proponemos seguir estrechando los lazos con otras sociedades científicas, tanto nacionales como internacionales, y potenciar nuestra presencia en los órganos de decisión y asesores, velando por los intereses de la Matemática aplicada, en especial por el avance en la investigación de calidad. Asimismo, queremos continuar promoviendo aquellas actividades que estimulen la comunicación científica y la formación continuada en el seno de la sociedad. Por ello animo a todos a colaborar en el Boletín, enviando artículos, reseñas opiniones, etc. También os animo a participar en los congresos, jornadas y demás eventos que organice o patrocine SĒMA. En especial espero vuestra presencia en “nuestro” congreso CEDYA-CMA que el

próximo septiembre se celebrará en la Universidad Carlos III de Madrid.

Quiero dirigirme también a aquellos que no son socios de SēMA, pero que comparten buena parte de nuestras inquietudes y objetivos. Les animo a que se unan a nuestro colectivo, en la seguridad de que con su participación resultaremos todos fortalecidos.

Pienso poner de mi parte todo el esfuerzo de que sea capaz y espero no defraudar las expectativas de nadie. No dudéis en transmitirme cualquier queja o sugerencia que pueda servir para mejorar nuestra Sociedad. Las opiniones de todos y cada uno de nosotros son un activo muy valioso que no debemos desaprovechar. Espero poder contar con vosotros igual que vosotros podéis contar conmigo.

Un fuerte abrazo,

Juan Ignacio Montijano
monti@unizar.es

Matemática Aplicada en España: Presente y futuro

M. DE LEÓN Y E. ZUAZUA

Resumen

La intención de este artículo no es hacer un estudio exhaustivo de la situación del área de Matemática Aplicada en España, sino más bien la de iniciar un necesario debate que contribuya a una mejor definición de contenidos en el área y a una homologación internacional. Las circunstancias para ello son las mejores en muchos años; por una parte, la celebración en Madrid del *International Congress of Mathematicians* en agosto de 2006 debe suponer una reflexión profunda sobre el tema. Por otra, los nuevos grados y masters que deben ponerse en marcha dentro del Espacio Europeo de Educación Superior, junto con el desarrollo necesario del recién estrenado Programa Nacional de Matemáticas hace imprescindible tal reflexión, si queremos afrontar los desafíos de las matemáticas españolas en este comienzo de siglo, y muy en particular, en el ámbito de la investigación multidisciplinar y de su interacción con otras ciencias y en el contexto tecnológico.

1 Producción matemática en España

Las Matemáticas españolas han experimentado un crecimiento en los últimos 25 años que podríamos calificar de espectacular. Se ha pasado de una producción del 0,3% en artículos en 1980 hasta el casi 5% actual. Correspondiendo a esa importancia que la investigación matemática española había alcanzado en los últimos años, y con el objeto de propiciar un impulso y apostar por la calidad, el impacto y la visibilidad, la Dirección General de Investigación decidió que en 2004, el Plan Nacional de I+D+i incluyese por primera vez un Programa Nacional de Matemáticas. Podemos ver en la Tabla I la producción española por quinquenios, desde 1993, analizada en artículos en revistas ISI:

1993 – 97	1994 – 98	1995 – 99	1996 – 00	1997 – 01	1998 – 02	1999 – 03
3,46	3,66	3,88	4,18	4,42	4,53	4,65
-17	-14	-15	-16	-13	-13	-6

Tabla I: Producción española en Matemáticas y Factor de impacto medio

En la Tabla I se recoge además el factor de impacto. Sigamos de momento analizando los aspectos cuantitativos. Posteriormente analizaremos algunos

aspectos cualitativos. La Tabla II muestra la producción científica española en el quinquenio 1999-2003 en las diferentes disciplinas.

Field	Percentage of papers from Spain	Relative impact compared to world
Space Science	5,79	-5
Agricultural Sciences	5,30	+7
Mathematics	4,65	-6
Microbiology	4,41	-20
Chemistry	4,25	-1
Plant & Animal Sciences	3,88	-9
Ecology / Environmental	3,42	-15
Physics	3,14	+19
Spain's overall percent share, all fields: 3.02		
Materials Science	2,90	+1
Biology & Biochemistry	2,87	-29
Pharmacology	2,86	-21
Neurosciences	2,73	-18
Molecular Biology	2,66	-10
Economics & Business	2,64	-33
Immunology	2,57	-27
Clinical Medicine	2,54	-5
Geosciences	2,51	-17
Engineering	2,32	+5
Computer Science	2,26	-28
Psychology / Psychiatry	1,91	-40
Social Sciences	0,84	-12

Tabla II

Resulta también espectacular esta tabla: las matemáticas españolas se han convertido en la tercera ciencia española en términos relativos. Alguien podría pensar que al ser cifras relativas, y ser el número de publicaciones en matemáticas inferior a otras áreas, si lo viésemos en términos absolutos, las posiciones variarían notablemente. En la Tabla III, que recoge cantidades totales en el decenio 1994-2004, vemos que no es así. Se observa que las matemáticas se sitúan en octavo lugar en producción global.

Estas tablas muestran pues que las matemáticas españolas han crecido, y que además lo han hecho vertiginosamente. ¿Cuáles han sido las causas? En la década de los sesenta y los setenta el aumento del bienestar económico del desarrollismo franquista supuso que una gran cohorte de estudiantes accediera a las universidades españolas, con la consiguiente necesidad de profesorado para atenderlo. La investigación matemática es barata, lo que propició que la mayoría de estos jóvenes profesores realizaran sus tesis doctorales. A la vez, ya había comenzado una labor pionera de algunos matemáticos más veteranos, reiniciando la labor de investigación que había sido truncada con la Guerra Civil. Se refuerzan los contactos con la matemática extranjera, principalmente francesa y estadounidense, mediante estancias pre y postdoctorales. Todo este

proceso se acelera en la década de los 80, en la que comienza a subvencionarse de una manera más regular la investigación por el Ministerio de Educación y algunas comunidades autónomas.

	FIELD	PAPERS	CITATIONS	CITATIONS PER PAPER
1	Clinical Medicine	41,754	319,028	7,64
2	Chemistry	35,568	268,456	7,55
3	Physics	22,963	172,766	7,52
4	Plant & Animal Science	16,190	81,584	5,04
5	Biology & Biochemistry	13,696	130,640	9,54
6	Engineering	13,260	46,743	3,53
7	Materials Science	7,820	30,230	3,87
8	Mathematics	7,763	17,444	2,25
9	Agricultural Sciences	7,328	35,175	4,80
10	Neuroscience & Behavior	7,041	78,536	11,15
11	Environment / Ecology	6,145	37,374	6,08
12	Molecular Biology & Genetics	5,875	91,607	15,59
13	Microbiology	5,455	55,013	10,08
14	Space Science	5,005	52,288	10,45
15	Geosciences	4,644	26,078	5,62
16	Computer Science	4,431	7,017	1,58
17	Pharmacology & Toxicology	4,144	25,234	6,09
18	Psychiatry/Psychology	3,373	10,790	3,20
19	Immunology	2,682	32,058	11,95
20	Economics & Business	2,178	5,849	2,69
21	Social Sciences, General	1,866	4,752	2,55
22	Multidisciplinary	223	1,046	4,69

Tabla III

Una consecuencia del proceso de crecimiento descrito anteriormente es que la investigación se hace al margen de cualquier directriz u orientación: se trabaja en lo que buenamente se puede y se cree más interesante. Esa es una de las características de la investigación matemática española, y una de sus debilidades para el futuro. Y esto es particularmente relevante en el caso de la Matemática Aplicada. El Programa Nacional de Matemáticas puede constituir una herramienta útil para paliar en parte esta situación, tal y como veremos más abajo.

2 El área de Matemática Aplicada

El área de conocimiento de Matemática Aplicada incluyó en sus inicios, entre otros, a todos los matemáticos que impartían su docencia en las Escuelas de Ingeniería. En cuanto a la investigación, su característica más destacada era el entorno de las Ecuaciones Diferenciales Ordinarias, las Ecuaciones en Derivadas Parciales y el Análisis Numérico.

Hoy en día, el área engloba matemáticos con muy distintos bagajes matemáticos: investigadores en ecuaciones diferenciales ordinarias y sistemas dinámicos, ecuaciones en derivadas parciales, aspectos numéricos en ambos casos, investigadores en análisis funcional, teoría de matrices, álgebra lineal, mecánica de fluidos, mecánica de sólidos, elasticidad, teoría de control, cálculo de variaciones y optimización, matemática discreta, finanzas, etc. En definitiva, un amplio panorama.

Estas circunstancias hacen que algunas personas opinen que hay muchos matemáticos puros que se han camuflado en el área. Una reflexión diametralmente opuesta, que suele acompañar a ésta, es que en Matemática Aplicada no se prueban teoremas, o que los que se prueban carecen de profundidad suficiente. Son reflexiones muy limitadas. Si nos fijamos en las actividades que realiza la *Society for Industrial and Applied Mathematics* (SIAM), nos damos cuenta que hoy en día no hay aspecto de la matemática llamada pura que no tenga un reflejo en las aplicaciones. De hecho, uno de los grandes logros de la Matemática Aplicada surgido en la segunda mitad del siglo XX como síntesis de las áreas de Ecuaciones Diferenciales y Análisis Numérico, es que ha servido de estímulo para todas las demás áreas de las matemáticas, que han sabido responder al reto de impulsar la actividad matemática hacia una investigación más pluridisciplinar y orientada a los ámbitos industrial, tecnológico y financiero. Basta echar un vistazo por las *webs* de los centros de investigación de referencia internacional para constatar hasta qué punto esto es así a día de hoy. Por otra parte es obvio que en este área se prueban teoremas, de muy diversa índole y en muchos casos, de gran calado. Pensemos por ejemplo en la Teoría de Códigos con profundos teoremas de Geometría Algebraica y Teoría de Números, o en las aplicaciones de la Mecánica Simpléctica a la Robótica y Planificación de Movimientos, y de la Geometría en el diseño de elementos finitos para sistemas complejos de EDPs. Cabe asimismo resaltar que el área de Matemática Aplicada recoge alguno de los retos más ambiciosos de la matemática actual. Así, el *Clay Mathematical Institute*¹ incluyó entre los problemas elegidos para sus jugosos premios la unicidad y regularidad de las soluciones de las ecuaciones de Navier-Stokes para un fluido incompresible tridimensional.

El Área de Matemática Aplicada incluye cerca de 1500 profesores de plantilla entre CU, TU, CE y TE, lo que supone el 50% de los efectivos de nuestras facultades y escuelas. Es cierto que es el área menos productiva por persona (veáse el informe coordinado por C. Andradás y E. Zuazua²), pero lo que se

¹<http://www.claymath.org/>

²*La investigación matemática en España en el periodo 1990-1999*, coordinado por Carlos

está reflejando no es más que la realidad de la propia matemática española (e internacional) donde solo una parte del profesorado universitario mantiene una actividad investigadora continuada y de cierto relieve, agravado este hecho porque la investigación matemática en nuestras escuelas de ingeniería es inferior a la de las facultades de matemáticas.

3 El papel de la investigación en Matemática Aplicada y el Plan Nacional

Nuestros equipos de investigación están recogiendo los frutos de un prolongado esfuerzo de establecer contactos con otras instituciones de investigación de orientación más tecnológica y con la industria. Todos conocemos en nuestro entorno grupos que trabajan en esa dirección y nos congratulamos de ese aún modesto pero creciente éxito.

Sin embargo, estas realizaciones están resultando más costosas de lo que sería deseable. No es en nuestra opinión fruto de la falta de interés o de formación de nuestra comunidad de investigadores en el área, sino más bien de un problema estructural del país en el que tradicionalmente la comunidad de científicos y la industria han vivido demasiado frecuentemente de espaldas el uno del otro. No es casual que aquéllos a quienes tan difícil resulta establecer contactos fructuosos con la industria en España vean que sus colegas en otros países como Francia o Alemania lo hagan con la máxima naturalidad. El Programa Nacional de Matemáticas se hace eco de esta realidad y propone algunas medidas que puedan contribuir a paliarlo, como son la dotación de recursos humanos y materiales adicionales para los equipos que se impliquen en estas tareas, y que éstas sean tenidas en cuenta en los procesos de evaluación.

Con respecto a esta ponencia de Matemáticas del Plan Nacional, creemos que sería justo empezar diciendo que el hecho de que las Matemáticas dispongan de una ponencia propia en el nuevo Plan es ya un éxito para toda nuestra comunidad matemática y que esto supone una herramienta importante que tenemos en gran medida a nuestra disposición para mejorar las condiciones de nuestra investigación y reorientar esfuerzos. La ponencia de Matemáticas del Plan Nacional fue concebida abierta, versátil, con clara vocación de incorporar no sólo a la buena investigación matemática que se realiza en las disciplinas más tradicionales que se ejercitan en el país, sino también a las nuevas áreas emergentes. En la ponencia³ se presenta una clasificación de líneas prioritarias de investigación⁴ que, obviamente, no pretende recoger todo el espectro; más bien, se trata de un primer paso en un proyecto de clasificación que intenta recoger las líneas cultivadas en España y que debe entenderse como una clasificación dinámica, que debe ir perfeccionándose en próximas convocatorias. En definitiva, debería convertirse en el mapa al que nuestra investigación

Andradas y E. Zuazua, CEAM2000.

³Publicada en el *Boletín SEMA*, no. 28 (2004), 192–209 y LA GACETA DE LA RSME Vol. 7.1 (2004), 5–26, y accesible en la web del Ministerio de Educación y Ciencia <http://www.mec.es>

⁴Que se recoge en la convocatoria anual de proyectos con el objetivo de clasificar las solicitudes por temas.

matemática debería converger. Y en el caso de la Matemática Aplicada, la tarea es quizás más urgente. En cualquier caso, la ponencia debería servir como fuente de debate para nuestra comunidad matemática. No olvidemos que de la experiencia que adquiramos del Plan Nacional 2004-2007 deberían salir nuevas ideas para el Plan Nacional 2008-2011.

La ponencia del Plan Nacional fue producto del esfuerzo de un grupo ampliamente representativo de una docena de personas que trabajaron con entusiasmo en una tarea que por primera vez se nos encomendaba, en el que participó un grupo mucho más extenso aún de investigadores en diversas rondas de consultas, habiéndose incorporado a la ponencia final las sugerencias recibidas. Éste es obviamente un proceso abierto y en el que, nuevamente, las sugerencias que sean trasladadas al Gestor de Matemáticas del Plan Nacional serán debidamente atendidas y trasladadas al Subdirector de Proyectos de la Dirección General de Investigación.

Por otra parte, el aumento de recursos para los proyectos de Matemáticas en los últimos años ha sido muy importante. Como botón de muestra cabe decir que si en el 2000 la ponencia de proyectos de Matemáticas dispuso de 2 millones de euros, en la convocatoria del 2004 ha dispuesto de casi 5 millones y medio. Simultáneamente, se ha pasado de un total de 133 proyectos solicitados a 185, mientras que el número de proyectos aprobados de 101 a 133 con un número de EJC's que supera los 850.

Desde el 2001, las Matemáticas se han incorporado a los modos de gestión que tradicionalmente eran propias a las áreas prioritarias del Plan Nacional. Así, tal y como se explica detalladamente en las convocatorias del Plan Nacional, los proyectos son primeramente evaluados por la ANEP en un proceso dirigido por el Coordinador de Matemáticas de la ANEP y sus colaboradores, y en el que participan unos 400 investigadores españoles y extranjeros. Después los proyectos son nuevamente considerados y priorizados por una comisión nombrada a tal efecto por la Subdirección General de Proyectos en el que cada año participan más de una docena de investigadores que anualmente cambian para ir cubriendo las diferentes áreas, grupos de investigación, universidades, etc. En este proceso de selección participan también el Coordinador de Matemáticas de la ANEP y sus adjuntos, en un proceso transparente y objetivo. Desde que se adoptó este modo de selección y priorización de proyectos, son más de cincuenta los investigadores españoles que han participado en estas comisiones finales de selección.

En todo este proceso, desde la evaluación a la selección, no existe ninguna preferencia o penalización de ninguna línea particular de investigación, como no sean la calidad y la oportunidad de los proyectos. La evaluación de la ANEP se encarga a dos, tres o a veces cuatro expertos en el tema específico del proyecto. Estos informes son sintetizados por el Coordinador y sus adjuntos y enviados al Gestor. La reunión con el Gestor y su equipo tiene como finalidad la selección y priorización de los proyectos y minimiza los errores.

La filosofía con la que se realiza esta selección ha sido difundida en varias ocasiones y en particular a través de cartas enviadas a todos los IP's de proyectos de los últimos años. Esencialmente, en el Plan Nacional se busca financiar buenos

proyectos y no equipos, aunque en la práctica, obviamente, ambas cosas van juntas, y se busca primar la calidad, la originalidad y el impacto. Los proyectos mejor evaluados reciben ahora un nivel de financiación bastante satisfactorio y, tal y como reflejábamos más arriba en las cifras citadas, hoy en día son muchos más los proyectos financiados, lo cual es fiel reflejo del crecimiento de nuestra comunidad matemática y de su productividad.

Como se comentaba antes, los proyectos de corte más aplicado reciben una atención especial, de acuerdo con las directrices recogidas en la ponencia del Plan Nacional puesto que parece lógico que en esos casos, más aún que en todos los demás, el número e impacto de las publicaciones no puede ser el criterio más relevante a la hora de priorizar el proyecto. La posibilidad de solicitar financiación en el apartado “Personal” con el objeto de contratar Personal Técnico de apoyo a la investigación, y de solicitar becas para Técnicos en Formación, refuerzan esta intención.

Es cierto también que algunos matemáticos aplicados deciden enviar sus proyectos, que muy bien podrían caer en la ponencia de Matemáticas, a otras ponencias como Medio Ambiente, Materiales, Diseño y Producción Industrial, TIC, ... Pero la motivación para ello es positiva, considerando la aprobación de su proyecto en estas ponencias como un reto adicional y que con ello liberen recursos para los proyectos que sólo pueden competir en Matemáticas. Esa situación es de hecho natural y habitual en todas las áreas multidisciplinares del Plan. En cualquier caso, los matemáticos aplicados españoles han de saber que en la ponencia de Matemáticas del Plan Nacional tienen un lugar claramente identificado y que encontrarán una evaluación realizada por expertos competentes en el área en un proceso que arroja muchas garantías por su transparencia y la multiplicidad de agentes que participan en la misma. Eso no quita, obviamente, que estos investigadores tengan perfecto derecho a someter sus proyectos a otras ponencias afines, lo cual, como decíamos, es algo perfectamente previsto en el Plan Nacional. A este respecto cabe añadir que en el Plan Nacional también está previsto que un proyecto sometido a una ponencia pueda ser reconducido a otra, si con ello se considera que aumentan sus posibilidades de éxito.

Como se ha puesto de manifiesto, y como no podía ser de otra manera, el MEC intenta financiar proyectos competitivos. Pero desde el Plan Nacional de Matemáticas y su gestión y evaluación desde el Ministerio y la ANEP, esto se ha hecho de la manera más positiva, de acuerdo a una filosofía que, entre otros, inspiró Jacques Louis Lions, uno de los matemáticos aplicados más influyentes de la segunda mitad del siglo XX: “Apoyemos de manera incondicional a los grupos de excelencia, apoyemos a los buenos equipos para que alcancen el grado de excelencia de los anteriores, apoyemos por último a todos los demás para que mejoren y alcancen cuotas de calidad competitivas”.

Otra de las iniciativas que se recogen en el Programa Nacional de Matemáticas es la creación de un Centro Nacional de Matemáticas (CNMat). Se trata de una iniciativa aún en fase de debate⁵. Pero sin duda alguna, uno de los grandes

⁵Recientemente se ha celebrado en Santiago de Compostela las Segundas Jornadas Abiertas

objetivos del CNMat debería ser la de promover una investigación matemática cada vez más multidisciplinar y con mayor interacción con los demás ámbitos del sistema de I+D+i.

4 El impacto de la investigación en Matemática Aplicada

En la Sección 1, se informaba de la producción española en el contexto internacional. Vayamos ahora a un análisis de la calidad. Se suele argumentar que el factor impacto de la investigación matemática española está por debajo de la media internacional. Y así es en efecto. Pero esto no debe ser tomado al pie de la letra. El factor de impacto del quinquenio 1999-2003 es -6. Digamos en primer lugar que esto significa que las citas de un artículo español están en media un 6% por debajo de la media de citas de un artículo en el mundo. Si observamos la Tabla I, veremos sin embargo que la tendencia en la serie es claramente creciente, con lo que, sin temor a sufrir un error grave, es de esperar que en uno o dos años, ese indicador sea positivo. Ese mismo fenómeno ha ocurrido en algunos países como Francia, que exhibía en el quinquenio 1995-99 un factor de impacto medio de -3, a pesar de ser las matemáticas la primera ciencia del país, y cuenta ahora en el quinquenio 1999-03 con un favorable +9. Esa ha sido también la evolución de Italia, como se puede comprobar visitando *Web of Knowledge*⁶. En general, la ciencia española presenta un retraso que va reduciéndose año a año.

Es ilustrativo reproducir la siguiente tabla

- Agricultural Sciences 26
- Biology & Biochemistry 31
- Chemistry 22
- Clinical Medicine 52
- Computer Science 39
- Ecology / Environment 36
- Economics & Business 33
- Engineering 14
- Geosciences 49
- Immunology 40
- Material Sciences 21
- Mathematics 27
- Microbiology 37
- Molecular Biology & Genetics 40
- Multidisciplinary 45

en torno a este tema <http://www.usc.es/imat/cnm2/>

⁶<http://go5.isiknowledge.com/portal.cgi>

- Neuroscience & Behavior 29
- Pharmacology & Toxicology 32
- Physics 22
- Plant & Animal Science 32
- Psychology / Psychiatry 61
- Social Sciences, general 40
- Space Science 19

que muestra la posición relativa de España en diversas áreas entre todos los países del mundo en cuanto a citaciones por artículo. Se observa que Matemáticas ocupa en esta lista de 22 disciplinas el lugar séptimo. Creemos que este indicador refleja también que las matemáticas están en muy buen camino.

En cualquier caso, tras unas décadas en las que la necesidad de publicar en revistas internacionales ha sido asumida con creces por nuestra comunidad matemática, el mensaje de que es importante apuntar a unas mejoras de calidad ha empezado a calar y es de esperar que en breve veamos los frutos en este cambio de tendencia. Por tanto, esta situación nada parece tener de alarmante sino que más bien se enmarca en el proceso de normalización de las publicaciones españolas en matemáticas, proceso que se desarrolla a buen ritmo.

En 2000, con ocasión del Año Mundial de las Matemáticas y mediante una Acción Especial del MECD, cuyo Investigador Responsable fue uno de los firmantes del presente artículo (M. de León) se elaboró un estudio sobre la investigación matemática en la década de los 90. Este estudio fue diseñado desde el Comité Español para el Año Mundial de las Matemáticas (CEAMM2000) que elaboró los parámetros del mismo. La coordinación y elaboración del informe fue encargada a Carlos Andradas y Enrique Zuazua, siempre en colaboración con el propio CEAMM2000. El resultado fue la primera aportación seria y objetiva para conocer a fondo la investigación española, convirtiéndose en una referencia obligada para cualquier estudio bibliométrico sobre las matemáticas.

En este informe se observaba una producción por profesor en Matemática Aplicada inferior a la media de otras áreas, lo cuál no es más que el reflejo de ser un área que supone la mitad de las matemáticas españolas y que encierra una enorme diversidad de circunstancias, tal y como señalábamos anteriormente.

En agosto del 2006 se celebrará en Madrid el ICM06, uno de los más claros exponentes del éxito de nuestra matemática. Este congreso, organizado por las cuatro sociedades matemáticas RSME, SCM, SEIO y SEMA, constituye una excelente ocasión para publicar un nuevo informe de esta índole. Por el momento, está a punto de finalizarse un estudio de los grupos de matemáticos españoles, basado en la coautoría, realizado en colaboración con el CINDOC, y cuyos responsables por la parte matemática son Manuel de León y David Martín de Diego. Esperamos que este informe sea el inicio de una actualización del referido a los 90.

Por otra parte, hoy en día disponemos de muchos otros parámetros para evaluar la salud de nuestra investigación matemática. Las Matemáticas

presentan signos de excelencia e impacto rotundos. Así, en la lista de los investigadores españoles que hasta el momento han recibido el calificativo de “Muy citados” (*Highly Cited Researchers*) por parte de ISI tenemos un total de 14 investigadores que desarrollan su labor en España (véase Tabla V).

Name	Institution	Category
Aguilar-Benitez De Lugo, M.	Ciemat	Physics
Barbacid, Mariano	Centro Nacional de Investigaciones Oncologicas	Molecular Biology & Genetics
Gonzalez, Antonio G.	Inst. de Productos Naturales Orgánicos, CSIC	Agricultural Sciences
Herrera, Carlos M.	Estación Biológica de Doñana, CSIC	Ecology / Environment
Nualart, David	University of Barcelona	Mathematics
Oro, Luis A.	Universidad de Zaragoza / CSIC	Chemistry
Palacios, José Maria	Almirall Prodesfarma Research Center	Neuroscience Pharmacology
Rodés, Juan	University of Barcelona	Clinical Medicine
Rodriguez, Benjamin	CSIC	Agricultural Sciences
Rohrer, Heinrich	CSIC	Computer Science
Sanchez-Madrid, Francisco	Universidad Autónoma de Madrid	Immunology
Sanz-Serna, J. M.	Universidad de Valladolid	Mathematics
Vazquez, Juan Luis	Universidad Autónoma de Madrid	Mathematics
Zuazua, Enrique	Universidad Autónoma de Madrid	Mathematics

Tabla V

Cuatro de ellos son matemáticos, y tres son del área de Matemática Aplicada. Estas cifras contrastan con las de otras disciplinas supuestamente con más impacto. Añadamos que de las 29 revistas españolas incluidas en ISI, 3 son de matemáticas. Con cierto conocimiento de causa, por razón de nuestros cargos, de las debilidades que aún muestran nuestras matemáticas, creemos que debemos sentirnos orgullosos de estar en donde estamos, habiendo partido de una investigación matemática prácticamente testimonial a principios de los años 80.

Hay otros indicios de la salud de nuestra investigación en Matemática Aplicada. Hoy en día los investigadores españoles de matemática aplicada están presentes en los Comités Editoriales en algunas de las revistas más prestigiosas del área como *European J. Applied Mathematics*, *SIAM J. Math. Anal.*, *SIAM J. Control Optim.*, *M3AS (Mathematical Models and Methods in Applied Sciences)*, y en muchas otras. Incluso, en algún caso, como *ESAIM:COCV*, el Redactor en Jefe es un investigador español.

Finalmente, en el estudio en marcha citado sobre grupos españoles en matemáticas, se desprende que una parte importante de los grupos más productivos y de mayor impacto están encuadrados en el área de Matemática Aplicada.

Nuevamente pues se puede decir que la Matemática Aplicada española va por buen camino y que en breve tendremos indicadores adicionales de este hecho tanto en lo que se refiere a a producción de corte más básico como en lo que se refiere a las aplicaciones.

5 La formación de nuestros licenciados en Matemática Aplicada

Debemos hacer alusión también a la formación en Matemática Aplicada de los licenciados en matemáticas de nuestras universidades. En realidad se trata de una preocupación compartida por todos los que estamos implicados en las licenciaturas de Matemáticas y ha sido recogida ya en el Libro Blanco recientemente editado por encargo de la ANECA. Por otra parte, está aún abierto el debate de cómo conseguir que la formación de nuestros jóvenes licenciados sea más completa en relación a lo que debería ser una visión más multidisciplinar de la Ciencia y la Tecnología. De hecho, muy recientemente se han celebrado unas Jornadas de Modelización en la Facultad de Matemáticas de la UCM organizadas por la Comisión de Educación de la RSME para analizar esta cuestión. Se trata pues de un tema de indudable interés en el que se trabaja en la actualidad pero en el que las soluciones no son del todo obvias, especialmente en el nuevo escenario de “Bolonia” en el que nuestras licenciaturas no sólo están destinadas a los estudiantes españoles sino a todos los europeos. Es en este contexto del Espacio Único de Educación Superior donde el debate está todavía abierto y esperando aportaciones de la comunidad matemática española.

Evidentemente, el debate debería centrarse en lo que se entiende por “formación en Matemática Aplicada” de nuestros matemáticos: cálculo numérico, optimización, EDPs, modelización, elementos de Mecánica, . . . Todo queda abierto ante los acuerdos de Bolonia, y hacerlo bien va a suponer un salto cualitativo también en lo que respecta a nuestra investigación. Podríamos decir que con Bolonia: “Cada uno se hace matemático a su manera”. Y hay bastante de verdad es esta frase.

Pero no sólo se ha de trabajar en lo que concierne a la formación en Matemática Aplicada de los licenciados en Matemáticas, sino también en el caso de las escuelas de ingeniería. En efecto, las escuelas de ingeniería española han de recorrer un importante camino para que podamos disponer en nuestro país de Escuelas Politécnicas de referencia internacional donde los alumnos tengan la posibilidad de complementar su formación en ingeniería con una importante base científica y en particular matemática como ocurre en centros tales como la “Ecole Polytechnique de Paris”, el “Politécnico de Milano” o la “Ecole Polytechnique Federale de Lausanne” en Suiza.

6 Conclusiones

De este breve estudio, se pueden extraer unas cuantas conclusiones:

- El área de Matemática Aplicada espera una mejor definición de contenidos, desde unas condiciones iniciales provenientes de la LRU, que la homologue a la situación internacional.
- Es un área muy grande, en la que por tanto se refleja la situación general

española. En cualquier caso, anteriormente hemos dado datos objetivos de que el área cuenta con un alto número de investigadores y docentes dedicados y capaces.

- El área debe aprovechar la ocasión que proporcionan los acuerdos de Bolonia para contribuir a que los nuevos grados y *masters* sirvan para definir una mayor interdisciplinariedad.
- El diseño de nuevas herramientas de investigación dentro del marco del Programa Nacional de Matemáticas (como el posible Centro Nacional de Matemáticas o acciones estratégicas orientadas a las aplicaciones) deberían tener una enorme influencia en el área de Matemática Aplicada, y propiciar la transferencia del conocimiento a los sectores productivos, tecnológicos y financieros.

En cualquier caso, conseguir un área de Matemática Aplicada acorde con los desafíos científicos y tecnológicos del siglo XXI es una tarea de todos, y exige generosidad y dedicación, y sobre todo, olvidar hábitos que se han quedado ya obsoletos: no se trata de luchar por quién se reparte el pasado sino de cómo construiremos el futuro.

Madrid, 29 de Noviembre de 2004

Manuel de León
Profesor de Investigación del CSIC
Coordinador de Matemáticas de la ANEP

Enrique Zuazua
Catedrático de Matemática Aplicada, UAM
Gestor del Programa de Matemáticas del Plan Nacional I+D+i

Matrices

J.M. GRACIA

Departamento de Matemática Aplicada y Estadística e I.O.,
Universidad del País Vasco

mepgrmej@vc.ehu.es

Intentaré probar que, al igual que Teruel o Soria, la teoría de matrices también existe. El primer Programa Nacional de Investigación en Matemáticas, que ha empezado a ejecutarse, parcela la investigación activa de Matemáticas en nuestro país en doce campos. En dicho programa el álgebra lineal y la teoría de matrices no aparecen. En los impresos de solicitud para la convocatoria última de ayudas a proyectos de investigación del M.E.C., a estas materias les cupo el dudoso honor de estar dentro de un campo que lleva por título “Otros campos”. Sin ninguna duda, esta ubicación evidencia la opinión que tienen sobre la investigación en estas matemáticas los redactores de la convocatoria, que supongo hereda el esquema del programa. Aunque en el prólogo del programa se piden disculpas a los marginados, es claro que no se trata de un olvido, sino de una decisión en curso.

En primer lugar, daré algunas pruebas de la actividad nacional e internacional de la investigación en álgebra lineal. En Vitoria–Gasteiz tuvo lugar en 1983 el primer congreso sobre este tema, fruto de una estrecha colaboración entre las universidades del País Vasco y Coimbra, y en el que Miguel de Guzmán impartió la conferencia inaugural. Siguió a éste, congresos habidos cada uno, dos o tres años. Particularmente importante, por la concurrencia internacional que concitó, fue el celebrado en Valencia 1987; poco después, surgiría la Sociedad Internacional de Álgebra Lineal, conocida por ILAS (su sigla en inglés), que organiza congresos anuales en Norteamérica, Europa e Israel (Haifa), habiéndose celebrado asimismo uno en China. Igualmente la SIAM organiza bianualmente un congreso sobre álgebra lineal aplicada. A los que nos educamos en matemáticas bajo la férula bourbakista, nos pareció que la teoría de matrices era un tema cultivado preferentemente por los anglosajones y que era debido a que éstos eran más pragmáticos y les gustaba más lo concreto que lo teórico; no obstante, debemos reconocer que el libro sobre matrices de Gantmacher nos llegó en traducción francesa del original ruso, aparecida en dos volúmenes en 1966. Muchos artículos sobre matrices citan este libro, que se ha convertido en una referencia canónica.

Para acabar este primer punto, permítaseme recordar que al álgebra lineal y multilineal y a la teoría de matrices está destinado el epígrafe 15 de la clasificación de la AMS utilizada por las revistas *Mathematical Reviews* y *Zentralblatt für Mathematik*; así como que en la clasificación de la UNESCO el álgebra lineal y la teoría de matrices tienen asignados los códigos 1201.10 y 1201.11, respectivamente. Amén de esto, hay tres revistas dedicadas a la publicación de artículos sobre estos temas y sus aplicaciones, que son *Linear Algebra and Its Applications*, *Linear and Multilinear Algebra*, y *SIAM Journal on Matrix Analysis and Applications*. Siendo la teoría de matrices una encrucijada de matemáticas, también aparecen artículos en muchas otras revistas. Creo pues, que tanto la actividad española como la internacional en álgebra lineal, merecían algún sitio en el programa de matemáticas.

En segundo lugar, osaré dar algunas razones de índole matemática para esta reivindicación. Dicen algunos que la investigación en álgebra lineal y teoría de matrices está acabada o es irrelevante. Para unos, las matrices se reducen a un baile de índices y prestidigitaciones con sus líneas; para otros, incluido Dieudonné, son la simple representación de una aplicación lineal. Pero las matrices no son sólo una notación utilísima; para algunos disidentes matemáticos, entre los que me incluyo, son objetos ricos en propiedades estructurales y combinatorias. Riqueza de la que dan cuenta la teoría combinatoria de matrices y la descripción de las formas canónicas de matrices genéricas mediante digrafos; simbiosis de combinatoria, matrices y programación lineal. También es conocido el estrecho maridaje que existe entre matrices y polinomios, y que fecunda ambos campos.

Dado que una de las señas de identidad de las matemáticas es la unidad entre todas sus ramas, la teoría de matrices es fagocitada con facilidad por otras disciplinas. Ejemplos abundan:

1. La teoría de matrices no negativas sería una rama del análisis funcional (operadores que dejan invariante un cono en espacios de dimensión infinita). Incluso más, cuando el teorema de Perron-Frobenius se utiliza en sistemas dinámicos o en las cadenas de Markov, puede pertenecer a las geometrías diferencial e integral o a la teoría de probabilidades. Bien es sabido que la teoría de Perron-Frobenius es de muy largo alcance, encontrando aplicaciones insospechadas en el estudio de funciones algebraicas al tratar sistemas implícitos de ecuaciones polinómicas con coeficientes no negativos. También se aplica en la resolución numérica de ecuaciones en derivadas parciales, y en el tratamiento de datos a lo Benzecri.
2. El análisis matricial sería una versión finito dimensional del análisis funcional.
3. La teoría de formas canónicas estaría contenida en la teoría de los *quivers* y de las representaciones, así como en la teoría de invariantes.
4. El estudio de las matrices cuyos elementos son funciones quedaría subsumido en el estudio de anillos.

Con estos ejemplos no es mi intención poner en entredicho estas ubicaciones, pues cada uno ve todo desde sus ojos. Sólo pido un poco de empatía hacia un grupo numeroso de matemáticos en activo.

En tercer lugar, la enseñanza del álgebra lineal se resiente por esta situación. Nuestras facultades de Matemáticas la han orientado como una introducción a las geometrías euclídea, afín, proyectiva, algebraica y diferencial. Esta elección, si bien impecable desde el punto de vista histórico y estructural, no hace justicia al tema y aleja al estudiante de nuevos, y no tan nuevos, desarrollos en álgebra lineal. Durante demasiado tiempo se han ignorado:

- Los valores singulares de matrices y su relación con el cálculo del rango.
- La localización de valores propios (y el teorema de Kharitonov).
- El campo de valores o rango numérico de matrices.
- Las desigualdades de entrelazamiento de Cauchy entre los valores propios de una matriz hermitica y los de sus submatrices principales.
- Las desigualdades de entrelazamiento de los valores singulares, y de los factores invariantes, que constituyen uno de los hitos de la teoría.
- La descripción de los pseudoespectros de matrices mediante los valores singulares.
- Las desigualdades de mayoración en el espacio de n -tuplas reales en el sentido de Hardy, Littlewood y Pólya, y su carácter ubicuo en las matemáticas; en particular,
 - implican que el poliedro convexo de las matrices doblemente estocásticas tiene por vértices a las matrices de permutación (teorema de Birkhoff, publicado originalmente en castellano);
 - dan cuenta de los cambios acaecidos en los invariantes de una matriz al variar algunos de sus elementos.
- La teoría de matrices no negativas y el teorema de Perron-Frobenius.
- El análisis espectral de una matriz mediante su digrafo.
- La forma canónica de Kronecker de haces de matrices, que sería de gran utilidad para:
 - los investigadores dedicados a las ecuaciones algebraico-diferenciales;
 - los ingenieros y físicos que estudian sistemas lineales con control (forma de Brunovsky, subespacios invariantes controlados).
- La estructura de matrices polinómicas y de fracciones racionales (forma de Jordan, matrices bipropias, índices de Forney, subespacios invariantes).
- El espacio métrico de los subespacios de un espacio vectorial de dimensión finita con la distancia de la abertura (*gap*).

- La estabilidad de subespacios invariantes y de formas canónicas de matrices frente a ligeras perturbaciones de las mismas.
- Los valores propios y la forma canónica de Jordan de relaciones lineales o aplicaciones lineales multívocas.
- etc.

Con todo soy consciente del malditismo que arrostra la teoría de matrices en muchos países. H. Schneider dijo que cuando llegó a los EE.UU. en los años 1960, la teoría de matrices era un tema “horrible”. Todavía en 1988 Olga Taussky nos sugería que cuando la gente desprecie las matrices, les recordemos que grandes matemáticos tales como Frobenius, Schur, C.L. Siegel, Ostrowski, Motzkin, Kac, etc. hicieron contribuciones importantes al tema. A esta lista podríamos añadir Leibniz, Jacobi, Hamilton, Cayley, Sylvester, Grassmann, Weiertrass, Kronecker, Hermite, Poincaré, Perron, Minkowski, Lyapunov, Courant, Hilbert, Kaplansky, Weyl, Jacobson, Wielandt, M.G. Krein, V. Pták, I. Gohberg, V.I. Arnold, P. Lax y W. Wasow. En la década pasada han aparecido dos libros rusos sobre álgebra lineal escritos por V.V. Prasolov y S.K. Godunov tras la estela del Gantmacher. En años recientes han hecho aportaciones notables D. Zeilberger, S. Friedland, A. Knutson, T. Tao, C. Woodward, A. Klyachko y W. Fulton entre otros.

Finite Dimensional Behavior in Semilinear Wave Equation

R. WILLIE

Department of Mathematics, University of Zimbabwe

rwillie@maths.uz.ac.zw

Abstract

The paper is devoted to the study of a semilinear wave equation with a linear damping term. For simplicity, the situation at the boundary is considered homogeneous of Robin type conditions. The asymptotic behavior in the limit of large diffusion in the equation is discussed in detail. In an associated linear wave evolution equation with no damping term, we show in a given sense that the limiting problem is a second order linear ordinary differential equation. In the semilinear case, under suitable dissipative conditions existence in the form of a global attractor is proved. Finally we demonstrate in the limit of large diffusion convergence of attractors of the semilinear wave equation to an attractor of a second order nonlinear finite dimensional equation.

Key words: *Semilinear Wave Equations, Global Attractor, Limit of large Diffusion, Upper semicontinuity of Attractors.*

AMS subject classifications: *47D07, 47D06, 34G10*

1 Introduction

In this paper, we study the dynamics generated and the asymptotic limit of large diffusion of a semilinear wave equation

$$u_{tt} + \beta u_t - \operatorname{Div}(d^\varepsilon(x)\nabla u) + \lambda u = f(u) \quad \text{in } Q_T \quad (1)$$

where $Q_T = \Omega \times (0, T)$ denote a cylinder of inferior base the set $\Omega \subset \mathbb{R}^N$, $N \geq 1$ an open bounded convex subset with smooth boundary $\partial\Omega = \Gamma$ and $\varepsilon > 0$ is a given parameter. Let $\Sigma_T = \Gamma \times (0, T)$ denote the lateral boundary surface

Fecha de recepción: 19/04/04

of the cylinder Q_T . The model equation in (1) is to be complemented by the conditions

$$\partial_{\vec{n}_\varepsilon} u + b^\varepsilon(x)u = 0 \quad \text{on } \Sigma_T, \quad u_t(0) = u_1^\varepsilon, \quad u(0) = u_0^\varepsilon \quad \text{in } \Omega, \quad (2)$$

where the first are of the equation in (1) on the lateral boundary surface, while the last two correspond to those of the initial data of the system at the initial instant zero of time.

Let us briefly describe the physical meaning of the terms arising in the model equations (1)-(2). Starting in the cylinder equation (1), the first two terms are of time displacements of second and first order respectively, of a wave function $u(x, t)$ which depend on a spatial variable x at each instant of time t . The physical notions of these respectively are of acceleration and of damped velocity by a parameter $\beta > 0$. The addendum term to these terms physically describe a spatial acceleration of the wave in divergence form. The acting operator defining this speed is called a diffusion operator. Here a non constant continuous coefficient $d^\varepsilon : \Omega \mapsto \mathbb{R}^+ \setminus \{0\}$ will be assumed and an adding $\lambda > 0$ to denote a constant potential function.

To make precise on the lateral boundary surface of the cylinder Q_T in the model equation (2), the notation

$$\partial_{\vec{n}_\varepsilon} \cdot = \langle d^\varepsilon(x) \nabla \cdot, \vec{n} \rangle \quad \text{with } \vec{n} \text{ being an exterior unit vector on } \Gamma$$

denote the co-normal derivative related to the diffusion operator in (1). The entire first conditions in (2) receive the name of homogeneous Robin type lateral boundary surface conditions and they will be assumed of a non continuous potential, such that if the Sobolev inclusions

$$H^1(\Omega) \subset L^r(\Gamma) \quad \text{are satisfied, then} \quad b^\varepsilon \in L^{q_0}(\Gamma) \quad \text{with} \quad \frac{1}{q_0} + \frac{2}{r} \leq 1. \quad (3)$$

Returning into the cylinder to balance the equation from the left hand side of (1), we shall consider a smooth function $f : \mathbb{R} \mapsto \mathbb{R}$ of class C^1 to describe the density of an externally applied force depending on u . This is to satisfy either no growth behavior, or conditions of the type

$$|f(u) - f(v)| \leq \mu(u, v)|u - v| \quad (4)$$

where the first Lipschitz factor is exponential of the form

$$\mu(u, v) = C_\eta (e^{\eta|u|^2} + e^{\eta|v|^2} + 1) \quad \forall \eta > 0, \quad \text{if } N = 2, \quad (5)$$

with $C_\eta > 0$, while will be assumed polynomial of the form

$$\mu(u, v) = C(|u|^{p-1} + |v|^{p-1} + 1), \quad \text{if } N \geq 3 \quad (6)$$

with $C > 0$ and order power $p > 1$ verifying that

$$p \leq 1 + \frac{4}{N-2}, \quad \text{or } p-1 = 2 \text{ if } N = 2. \quad (7)$$

First, let us comment that the above wave equation (1) is usually known as a nonlinear telegraph equation. In general, [17, 24, 31] the wave equation is an important partial differential equation of hyperbolic type. This is so, since [31] every second order hyperbolic equation may be reduced by a transformation of coordinates to a canonic form which at least at any given point has principal part the same as that of the wave equation. Therefore, to have a good understanding of the theory of hyperbolic equations a careful study of the wave equation is essential. For the physical scientist, engineer or applied mathematician the reason for studying the wave equations are obvious as these equations govern many wave propagation and vibration phenomena.

In the case of no damping term, the homogeneous equation considered in any space dimensions has simple solutions physically representing traveling waves with plane or spherical profiles. Still in this situation, in the case of initial and boundary value problems, the equations exhibits an interesting property that of the conservation of the total energy of the solution. In three space dimensions it known that Kirchhoff's formula gives the solution. Usual in this case the descent method is used to obtain the solution in two and one space dimensions. It is known that in three space dimensions the so called Huygen's principle holds, a consequence of which is the sharp propagation of signals. In two space dimensions, Huygen's principle does not hold and as a result propagating signals always have a decaying trailing edge. This phenomenon is known as a diffusion of waves. For further readings and detailed illustrative examples on this material on the wave equation we recommend the interested reader to consult [31], while for advanced treatments to see the literature [17, 24]. Finally, we note that the paper we are presenting is an improvements on the exposition of our results in [30]. We would like to thank the referees, for indicating to us that we supplement and further details on the following background notes of the problem.

To do this, we shall start by noting that, large diffusion occurs almost natural every time we sit down to solve a parabolic equation with generalized elliptic differential operator for which either the simple coercive inequality in Lax-Milgram's Theorem [5, 21, 25] or Gårding's inequality holds. The most typical cases are when the principal parts in these operators are linearly pertubated by non constant functions or potentials. The asymptotic limit case of large diffusion has been a theme of study in the last two and half decades [1, 4, 7, 6, 13, 15, 26, 28]. There has been two inverse approaches to the problem. In the first case, a formal limiting ordinary differential equation is postulated and assumed to have an attractor. It is then shown that if this attractor is considered as a subset of constant functions in the phase space and if the diffusion is sufficiently large then the orginal partial differential equation also has an attractor. Moreover, the attractors are upper semicontinuous in the limit case of large diffusion. Here the guess on the formal explicit form of the limiting problem, usually is sound founded from other natural sciences.

The second approach [4, 7, 27, 28, 29], shows under suitable boundedness conditions of the data, that the limit case of large diffusion has the effect of stablizing all spatial inhomogeneities in the equation, in such a way that there

is homogenization to a constant function in the domain.

It turns out therefore, that the asymptotic limiting problem is finite dimensional and describe limit processes that are defined in the spatial average solutions of the equations. The dynamics generated by the equations can then be related, if we prove convergence of their dynamical structures depicting the geometry of the solutions. Most recently, for example in [1, 28, 29] research problems have covered nonlinear boundary conditions, localized limit case of large diffusion and the asymptotic limit of large diffusion of the inertial manifold. More precise, in [28] we have studied the asymptotic limiting problem of large diffusion in a system of nonlinear reaction and diffusion (R-D) equations with nonlinear boundary conditions. Initially, we had proved results of this problem in [27] in some extended scales of Hilbert spaces. However, the asymptotic limit problem of large diffusion of the attractors generated by the nonlinear semigroups associated with the equations was obtained in sets topology of continuous functions in an open bounded smooth domain of restricted space dimensions. Following a careful and detailed study of the regularity and boundedness of the solutions in the spaces of Hölder continuous functions given in [28] we were able to prove convergence of attractors of the system of reaction and diffusion equations to one associated with a system of ordinary differential equations, in sets topology of uniformly continuous functions on an open bounded smooth domain in any space dimensions, when the limit of large diffusion in all parts of this spatial domain of the system of R-D equations was considered. On the other hand, very little is known [6, 22] of the long time dynamics of large diffusion in semilinear wave equations with nonlinear boundary conditions.

Thus, the aim of the present paper is to provide an initial step and a revisit of the study of the semilinear wave equation. The approach it considers in the investigation of the problem is the second one of the mentioned in the above paragraphs. Let us now outline its structure, this is organized as follows. In Section 2, we shall provide the most necessary properties of the second order elliptic differential operator in the equations (1)-(2) and formulate our problem in an evolution sense of the extended scales of Hilbert spaces. In Section 3, we shall study the asymptotic limit problem of the limit of large diffusion in an associated linear wave equation to the semilinear equation (1)-(2) with zero damping term. In Section 4, we therefore shall furnish a complete study of the local dynamical system generated by the equations (1)-(2) in the sense of existence and uniqueness of solutions in extended scales of Hilbert spaces. In addition, if we suppose a dissipative condition on the nonlinear term, in Section 5 we prove existence of solutions in the form of a global attractor.

In Section 6, we shall study the asymptotic limit of the global attractor in the notion of upper semicontinuity in the limit case of large diffusion. This convergence of attractors is obtained in the norm of the natural energy space of the existence and uniqueness of the solutions to the wave equation (1)-(2). Moreover, the results thus we obtain conclude a finite dimensional asymptotic behaviour of the solutions and relate the geometric structures of the solutions sets in the long time dynamics of the problems.

In Section 7, we give brief conclusive remarks of the results. Here in Subsection 7.1 we review natural extensions to the cases of nonlinear lateral boundary surface conditions in (2) of a generalized wave equation to (1).

2 Preliminaries

Throughout the paper, we shall assume familiarity with the notions [5, 21] of Sobolev spaces $H^s, s \in \mathbb{R}$.

Let us now consider, the stationary problem in time to the wave equation (1)-(2) and define the bilinear form $a^\varepsilon : H^1(\Omega) \times H^1(\Omega) \mapsto \mathbb{R}$ by

$$a^\varepsilon(u, \varphi) = \int_{\Omega} d^\varepsilon(x) \nabla u \nabla \varphi + \lambda \int_{\Omega} u \varphi + \int_{\Gamma} b^\varepsilon(x) u \varphi. \quad (8)$$

It is easy to see that (8) is symmetric and continuous. Now if we let $\lambda_0 : \mathbb{R}^+ \mapsto \mathbb{R}^+$ satisfying $\lambda_0(0) = 0$ and choose

$$\lambda \geq \lambda_0(\|b_-^\varepsilon\|_{L^{q_0}(\Gamma)}) \quad \text{where } b_-^\varepsilon(x) = \max\{-b^\varepsilon(x), 0\}.$$

Then, if the diffusion in the second order spatial differential operator of the wave equation (1)-(2) is adequately large, we have that

$$a^\varepsilon(u, u) \geq \beta_1^\varepsilon \int_{\Omega} |\nabla u|^2 + \beta_0 \int_{\Omega} |u|^2 \geq C \|u\|_{H^1(\Omega)}^2 \quad (9)$$

where $\beta_1^\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$, and $\beta_0, C > 0$ are independent of $\varepsilon > 0$. In other words, the bilinear form (8) is coercive in $H^1(\Omega)$ - norm.

Consequently, we have using Lax-Milgram's Theorem [5, 21, 25] that (8) define an isomorphism $L^\varepsilon u \in H^{-1}(\Omega)$ such that

$$\text{for each } u \in H^1(\Omega) \Rightarrow a^\varepsilon(u, \varphi) = \langle L^\varepsilon u, \varphi \rangle, \quad \forall \varphi \in H^1(\Omega),$$

where the notation $\langle \cdot, \cdot \rangle$ denote the inner product of the space $L^2(\Omega)$. Further, the restriction $L^\varepsilon u \in L^2(\Omega)$ define the elliptic operator $A_\varepsilon : D(A_\varepsilon) \subset L^2(\Omega) \mapsto L^2(\Omega)$ with realization

$$\begin{aligned} D(A_\varepsilon) &= \{u \in H^1(\Omega) : A_\varepsilon u = -\text{Div}(d(x)\nabla u) + \lambda u \in L^2(\Omega) \\ &\quad \text{and } \partial_{\vec{n}^\varepsilon} u + b^\varepsilon(x)u = 0 \text{ on } \Gamma\}, \end{aligned} \quad (10)$$

such that $L^\varepsilon u = A_\varepsilon u$ for all $u \in D(A_\varepsilon)$. It is well known [3, 16, 21] that A_ε is a sectorial operator and has a compact resolvent set in $L^2(\Omega)$. Hence from [3, 4, 16], we have the extended scales of Hilbert spaces $D(A_\varepsilon^\alpha) = X_\varepsilon^\alpha, \alpha \in \mathbb{R}$ are well defined and endowed with the graph norm

$$\|A_\varepsilon^\alpha u\|_0 = \|u\|_\alpha, \quad \forall u \in X_\varepsilon^\alpha, \alpha \in \mathbb{R}.$$

In particular, these scales of spaces satisfy

$$X_\varepsilon^1 \cong D(A_\varepsilon), \quad X_\varepsilon^{1/2} \cong H^1(\Omega), \quad X_\varepsilon^0 \cong L^2(\Omega), \quad \text{and } X_\varepsilon^{-1/2} \cong H^{-1}(\Omega).$$

Now we can introduce some abbreviations of the norm notations. To begin let us recall that the norm of the space $X_\varepsilon^{1/2}$, is equivalent to one obtained from the bilinear form (8) i.e

$$\|u\|_{1/2,\varepsilon}^2 \cong a^\varepsilon(u, u) \quad \text{for } u \in X_\varepsilon^{1/2}.$$

This imply that for the scales of spaces we shall adopt the notation

$$\|\cdot\|_{\alpha,\varepsilon} \quad \text{and} \quad \|\cdot\|_{1/2,0,\varepsilon} = \|\cdot\|_{X_\varepsilon^{1/2} \times X_\varepsilon^0}.$$

We also will drop the parameter index when the space is standard Sobolev for example $\|\cdot\|_{1/2} = \|\cdot\|_{H^{1/2}(\Omega)}$. Finally in the sequel, for simplicity we shall use the same notation $C \geq 0$ to denote any generic constant independent of $\varepsilon > 0$. In the following subsection, we formulate the initial and boundary value wave problem (1)-(2) in an evolution sense of the above functional spaces, which incorporate the boundary conditions in the equation (1).

2.1 The evolution equation.

Let us consider the unknown wave function $u(x, t)$ to our problem (1)- (2) as functions parameterized in time as follows $(0, T) \ni t \mapsto u(t) \in X$, where X is a Banach space. Then, we can rewrite the wave equations (1)-(2) in the form

$$u_{tt} + \beta u_t + L^\varepsilon u = f(u), \quad (u(0), u_t(0)) = (u_0^\varepsilon, u_1^\varepsilon) \in X_\varepsilon^{1/2} \times X_\varepsilon^0 \quad (11)$$

where the equality is understood in the sense of functions of the space $H^{-1}(\Omega)$ and almost everywhere on $(0, T)$.

Intuitively, from the computations leading to (9) it should have been guessed that

$$\sigma(\varepsilon) \stackrel{def}{=} \inf_{\Omega} \{d^\varepsilon(x)\} \rightarrow \infty \quad \text{as } \varepsilon \rightarrow 0. \quad (12)$$

In fact, this guess is correct and is the hypothesis that define the limit case of large diffusion at $\varepsilon = 0$ in the equation. But before this situation can occur and for the meantime corresponding to physical relevant cases, see for example references in [4, 26], we shall suppose the data of the evolution equation (11) is boundedness in norm for all $\varepsilon > 0$ and that only L^1 convergence as $\varepsilon \rightarrow 0$ is known.

We are now, in position to study the asymptotic limit problem to a linear equation associated with the semilinear wave equation (1)-(2).

3 The linear evolution equation

In this section, we study the asymptotic limit problem as $\varepsilon \rightarrow 0$ under model hypotheses following (11) in the linear evolution equation,

$$u_{tt} + L^\varepsilon u = f^\varepsilon(t), \quad u^\varepsilon(0) = u_0^\varepsilon \in X_\varepsilon^{1/2}, \quad u_t^\varepsilon(0) = u_1^\varepsilon \in X_\varepsilon^0, \quad (13)$$

with no damping term. Here we shall assume $f^\varepsilon \in L^1(0, T, L^2(\Omega))$ is well behaved for all $\varepsilon > 0$ and has spatial average satisfying

$$f_\Omega^\varepsilon(t) \rightarrow f_\Omega(t) \quad \text{weakly in } L^1(0, T) \text{ as } \varepsilon \rightarrow 0.$$

Thither, the notation $\psi_X(t) = \psi(x, t)|X|^{-1}\chi_X \in L^1$ for any real function ψ defined on $X \times (0, T)$, where χ is the characteristic function of the set X denote the spatial average in X . Now, regarding the solvability of the equations in (13) we have the following theorem.

Theorem 1 *The linear evolution equation of (13) has a unique weak solution $(u^\varepsilon, u_t^\varepsilon) \in Y_T = C([0, T], X_\varepsilon^{1/2} \times X_\varepsilon^0)$ and the energy identity*

$$\frac{1}{2}E^\varepsilon(t) - \frac{1}{2}E^\varepsilon(s) = \int_s^t \langle f^\varepsilon, u_t^\varepsilon \rangle d\sigma, \quad \forall t \geq s \geq 0 \quad (14)$$

where $E^\varepsilon(t) = \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2$ is satisfied.

Proof. The proof of the first part of the theorem is standard and can be found in [20, 25]. It remains only to show that the energy identity (14) holds, but this is obtained via a density argument similar to the one used below in the semilinear case. \square

Let us comment that the proof given in [25], is simple and makes use of the standard Galerkin's technique. This technique consist formally in deriving an energy inequality from (14) which yields among other things an *a priori* estimate for the solution, expressing as well a continuous dependence relation of this with respect to the data of the problem. Then use an approximation scheme and uniform energy estimate of the approximating sequence of solutions to associated finite dimensional problems, for which existence is known *a priori*, to obtain a solution to (13) as a weak limit. Since the weak solution obtained by this process depends on the choice of the approximation scheme, the energy estimate yields uniqueness only if more regularity on the data is assumed.

Thus to circumvent this difficulty the authors in [25] derive an energy equation for time - integrated quantities which together with the well known Gronwall's Lemma conclude the desired uniqueness.

We now turn back to our main goal of the section and we have the following asymptotic behaviour of solutions to the equation (13).

Theorem 2 *Consider the differential equations*

$$\ddot{u}_\Omega + \left(\frac{|\Gamma|}{|\Omega|} \int_\Gamma b + \lambda \right) u_\Omega = f_\Omega(t), \quad (u_\Omega(0), \dot{u}_\Omega(0)) = (u_\Omega^0, u_\Omega^1) \in \mathbb{R}^2 \quad (15)$$

Then, the equations in (15) are strongly limiting for those of (13) in $\varepsilon = 0$ in the sense of convergence in the space $L^2(0, T, X_\varepsilon^{1/2})$ of solutions, and moreover

$$\lim_{\varepsilon \rightarrow 0} \int_{Q_T} d^\varepsilon(x) |\nabla u^\varepsilon|^2 = 0. \quad (16)$$

In addition, if the data in the linear equations of (13) converge strongly as $\varepsilon \rightarrow 0$. Then

$$(u^\varepsilon, u_t^\varepsilon) \rightarrow (u_\Omega, \dot{u}_\Omega) \quad \text{strongly in } Y_T, \quad (17)$$

as $\varepsilon \rightarrow 0$ given that assumption (12) is satisfied.

Proof. Let us consider the energy identity in the equation (14) for $s = 0$, that is

$$\frac{1}{2}E^\varepsilon(t) = \frac{1}{2}E^\varepsilon(0) + \int_0^t \langle f^\varepsilon, u_t^\varepsilon \rangle d\sigma, \quad \forall t \geq 0. \quad (18)$$

Then applying Hölder's inequality and the following particular inequality

$$ab \leq \eta a^2 + \frac{1}{\eta} b^2, \quad a, b \geq 0, \eta = 1/4 \quad (19)$$

of the standard due to Young, succesively in the order the one stated, we estimate in the identity of (18) above to obtain

$$E^\varepsilon(t) \leq 16(E^\varepsilon(0) + \|f^\varepsilon\|_{L^1(0,T,L^2(\Omega))}^2). \quad (20)$$

By hypotheses, we have (20) is uniformly bounded in $\varepsilon > 0$. Thus, $(u^\varepsilon, u_t^\varepsilon)$ is bounded in norm of $L^\infty(0, T, X_\varepsilon^{1/2} \times X_\varepsilon^0)$ for all $\varepsilon > 0$. As a result, passing to subsequences if necessary we obtain

$$(u^\varepsilon, u_t^\varepsilon) \rightarrow (v, v_t) \quad \text{weak-* in } L^\infty(0, T, X_\varepsilon^{1/2} \times X_\varepsilon^0) \quad (21)$$

as $\varepsilon \rightarrow 0$. Since (12) is satisfied, it follows for all $t \geq 0$ that

$$\lim_{\varepsilon \rightarrow 0} \int_\Omega |\nabla u^\varepsilon|^2 = 0. \quad (22)$$

Thus, the lower semicontinuity of the $H^1(\Omega)$ norm, imply the limit v is a constant in Ω for all $t \in (0, T)$. In particular, from the strong compactness in X_ε^0 for all $t \in (0, T)$ and (22) imply $u^\varepsilon(t) \rightarrow v(t)$ strongly in $L^2(0, T, X_\varepsilon^{1/2})$ as $\varepsilon \rightarrow 0$. In addition, we have using the Poincare's inequality that

$$\int_{Q_T} |u^\varepsilon - \bar{u}^\varepsilon|^2 \leq C \int_{Q_T} |\nabla u^\varepsilon|^2 \rightarrow 0$$

as $\varepsilon \rightarrow 0$ where \bar{u}^ε as usual is the spatial average function. It follows from (3) and (22) that we conclude

$$\int_0^T \|u^\varepsilon - \bar{u}^\varepsilon\|_{L^{q_0}(\Gamma)}^2 \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \quad (23)$$

Now, let $\psi(t) \frac{1}{|\Omega|} \chi_\Omega(x)$ with $\psi \in C^\infty[0, T]$ satisfying $\psi(T) = 0$ and χ_Ω denoting the characteristic function of the domain Ω , be a test function in

(13). Then integrating by parts, using the continuity of the boundary tri-linear form and (23) we obtain

$$-\int_0^T v_t \dot{\psi} + \int_0^T \left(\frac{|\Gamma|}{|\Omega|} \int_{\Gamma} b + \lambda \right) v \psi = v_1 \psi(0) + \int_0^T \left(\int_{\Omega} f \right) \psi, \quad (24)$$

as $\varepsilon \rightarrow 0$. Thus, in a distributional sense we find

$$\ddot{v} + \left(\frac{|\Gamma|}{|\Omega|} \int_{\Gamma} b + \lambda \right) v = \int_{\Omega} f \quad \text{on } (0, T). \quad (25)$$

Moreover, if we multiple in (25) with $\psi \in C^\infty[0, T]$ verifying $\psi(T) = 0$. Then integrate by parts the first term and comparing the result of the whole process with the identity in (24) we conclude that $\dot{v}(0) = v_1$. Repeating the above limiting process, in a second integration by parts of the term with the time derivative we find $v(0) = v_0$ on using the initial data conditions. Hence, by uniqueness of the limit, we must have $(v, v_t) = (u_\Omega, \dot{u}_\Omega)$.

Now, let u^ε as a test function in (13) and integrate in time to obtain

$$\int_{Q_T} u_{tt}^\varepsilon u^\varepsilon dx dt + \int_0^T a^\varepsilon(u^\varepsilon, u^\varepsilon) dt = \int_{Q_T} f^\varepsilon u^\varepsilon dx dt. \quad (26)$$

It is easy to see in the above identity that, we can pass to the limit as $\varepsilon \rightarrow 0$. Thus using (25) if we let v as a test function, integrate by parts and compare the result with that of passing to the limit in (26) as $\varepsilon \rightarrow 0$, we get that (16) is verified and the first part of the theorem is proved.

To conclude the proof of the theorem, we find the evolution equation defined by $\varphi^\varepsilon = u^\varepsilon - u_\Omega$ and use the energy inequality (20) to obtain

$$\begin{aligned} \|(\varphi^\varepsilon, \varphi_t^\varepsilon)\|_{1/2,0,\varepsilon}^2 &\leq C \left(\|(\varphi_0^\varepsilon, \varphi_1^\varepsilon)\|_{1/2,0,\varepsilon}^2 + \right. \\ &\quad \left. + \|f^\varepsilon(t) - h_\Omega(t)\|_{L^1(0,T,L^2(\Omega))}^2 \right) \rightarrow 0 \end{aligned}$$

as $\varepsilon \rightarrow 0$. The rest is evident, and the proof of the theorem is complete. \square

Let us note that, the convergence in (16) can be strengthened, if in addition it is assumed that (16) is verified by the initial data u_0^ε . This, yield in the Theorem 2 instead of (16) the convergence that is uniform in positive time of the integrated quantity. In fact, the additional hypothesis we have cited is not a restrictive one. Since, when solving the elliptic problem for the asymptotic behaviour in question is natural, see for example [4].

4 The semilinear evolution equation

We are now in position to study the existence and uniqueness of solutions to the semilinear problem (1)-(2). Throughout this section we shall concentrate only on the case $N \geq 3$ and we remark that the argument in the remaining cases is easily adaptable with minor modifications. Thus to initiate our study let us introduce the following concept of a weak solution to the problem.

Definition 1 Assume in (1)-(2) that $d^\varepsilon, b^\varepsilon \in C^1$. Then, the pair $(u^\varepsilon, u_t^\varepsilon) \in Y_T$ is said to be a weak solution to the problem (1)-(2) \iff (11), if there exists a sequence of regular data (u_0^n, u_1^n) , $n = 1, 2, \dots$ such that

$$(u_0^n, u_1^n) \rightarrow (u_0^\varepsilon, u_1^\varepsilon) \in X_\varepsilon^{1/2} \times X_\varepsilon^0, \quad \text{and} \quad (u^n, u_t^n) \rightarrow (u^\varepsilon, u_t^\varepsilon) \in Y_T \quad \text{as } n \rightarrow \infty,$$

where (u^n, u_t^n) , $n = 1, 2, \dots$ is a unique sequence of strong solutions to (11) corresponding to the above regular initial data.

With respect to the solvability of (1)-(2) we state

Theorem 3 The semilinear evolution wave problem (11) has a unique solution $(u^\varepsilon, u_t^\varepsilon) \in Y_T$ and the energy identity

$$\frac{1}{2}E^\varepsilon(t) - \frac{1}{2}E^\varepsilon(s) = \int_s^t \langle f(u), u_t \rangle d\sigma, \quad \forall t \geq s \geq 0 \quad (27)$$

holds, where $E^\varepsilon(t) = \|(u, u_t)\|_{1/2,0,\varepsilon}^2 + \beta \int_s^t |u_t|^2$.

Proof. Assume (27) holds and define a nonlinear mapping $\mathcal{F} : Y_T \mapsto Y_T$ such that if $(u, u_t) \in Y_T$ then $(v, v_t) = \mathcal{F}(u, u_t)$ solves the problem

$$v_{tt} + \beta v_t + L^\varepsilon v = f(u), \quad v^\varepsilon(0) = v_0^\varepsilon \in X_\varepsilon^{1/2}, \quad v_t^\varepsilon(0) = v_1^\varepsilon \in X_\varepsilon^0. \quad (28)$$

Next fix $\rho > 0$ and consider the bounded subset of Y_T ,

$$U = \left\{ (\varphi, \varphi_t) \in Y_T : \sup_{0 \leq t \leq T} \|(\varphi, \varphi_t)\|_{1/2,0,\varepsilon} \leq \rho \right\}.$$

If we set $\|(v_0, v_1)\|_{1/2,0,\varepsilon} \leq \frac{\rho}{4}$ then multiplying in (28) by v_t we find

$$\begin{aligned} \frac{1}{2} \|(v, v_t)\|_{1/2,0,\varepsilon}^2 + \beta \int_0^t \int_\Omega |v_t|^2 dx d\sigma \\ \leq \frac{1}{2} \|(v_0, v_1)\|_{1/2,0,\varepsilon}^2 + \int_0^t |\langle f(u), v_t \rangle| d\sigma. \end{aligned}$$

Since $\beta \int_0^t \int_\Omega |v_t|^2 \geq 0$, it follows using the growth conditions on the nonlinear term that

$$\frac{1}{2} \|(v, v_t)\|_{1/2,0,\varepsilon}^2 \leq \frac{\rho}{8} + CT|\Omega|^{1/2}(\rho^p + 1) \sup_{0 \leq t \leq T} \|v_t\|_0,$$

from which the Young's inequality (19) imply

$$\|(v, v_t)\|_{1/2,0,\varepsilon}^2 \leq \frac{\rho}{2} + 4(CT|\Omega|^{1/2}(\rho^p + 1))^2.$$

Consequently, if $\rho \gg 1$ is sufficiently large and we choose

$$T \leq (2C|\Omega|^{1/2}(\rho^p + 1))^{-1}$$

we obtain $\|(v, v_t)\|_{1/2,0,\varepsilon} \leq \rho$, hence $(v, v_t) \in U$ and \mathcal{F} maps U onto itself.

In continuation, notice that for $(u, u_t), (w, w_t) \in U$ if we set $(v, v_t) = \mathcal{F}(u, u_t)$ and $(\psi, \psi_t) = \mathcal{F}(w, w_t)$ so that $(\varphi, \varphi_t) = (v - \psi, v_t - \psi_t)$ solves

$$\varphi_{tt} + \beta\varphi_t + L^\varepsilon\varphi = f(u) - f(w),$$

and taking the inner product with φ_t we have

$$\begin{aligned} \|(\varphi, \varphi_t)\|_{1/2,0,\varepsilon} &\leq \int_0^T |\langle f(u) - f(w), \varphi_t \rangle| d\sigma \\ &\leq 3CT\rho^p \sup_{0 \leq t \leq T} \|(\varphi, \varphi_t)\|_{1/2,0,\varepsilon} \end{aligned}$$

where again we have used the fact that $\beta \int_0^t \int_\Omega |v_t|^2 dx d\sigma \geq 0$. Thus for $\rho \gg 1$ sufficiently large, if we choose $T \leq (6C\rho^p)^{-1}$ we obtain that \mathcal{F} is a strict contractive mapping and thanks to Banach's fixed point theorem, there exists a unique $(u, u_t) = \mathcal{F}(u, u_t)$ that solves (11).

Now assume $d^\varepsilon, b^\varepsilon \in C^1$ and for $n = 1, 2, \dots$ let $v^n \in C((0, T), X_\varepsilon^1) \cap C^1((0, T), X_\varepsilon^{1/2}) \cap C^2((0, T), X_\varepsilon^0)$ be a regular sequence of solutions to (28) with $v_0^n \rightarrow v_0 \in X_\varepsilon^{1/2}, v_1^n \rightarrow v_1 \in X_\varepsilon^0$. Then (v^n, v_t^n) is Cauchy in $C([0, T], X_\varepsilon^{1/2} \times X_\varepsilon^0)$ and the limit (v, v_t) solves (28) in the sense given by

$$\frac{d}{dt} \langle v_t, \varphi \rangle + \beta \langle v_t, \varphi \rangle + \langle L^\varepsilon v, \varphi \rangle = \langle f(u), \varphi \rangle,$$

for almost everywhere on $[0, T]$. Thus $(v, v_t) = \mathcal{F}(u, u_t)$ is a weak solution to (11) and taking $\varphi = v_t$ integrating in time for $t \geq s \geq 0$ we conclude (27) holds. This concludes the proof of the theorem. \square

Let us note that, under sufficient regularity assumptions on the data of problem (11) it is usual to prove the well posedness via abstract semigroup methods [3, 16]. Often, in this case one reads the evolution problem in the form

$$\partial_t U + \mathcal{A}^\varepsilon U = F(U), \quad U_0 = (u_0^\varepsilon, u_1^\varepsilon)^\top$$

where

$$\mathcal{A}^\varepsilon = \begin{pmatrix} 0 & -I \\ L^\varepsilon & \beta \end{pmatrix}, \quad U = (u, u_t)^\top, \quad \text{and} \quad F(U) = (0, f(u))^\top.$$

Further, the nonlinear map in the proof of Theorem 3 is given by the variation of constants formula

$$\mathcal{F}(U)(t) = e^{-\mathcal{A}^\varepsilon t} U_0 + \int_0^t e^{-\mathcal{A}^\varepsilon(t-s)} F(U(s)) ds$$

in appropriate functional spaces.

5 Global Existence and boundedness of solutions

We now study the global existence and boundedness of the local dynamical system given by the equations (1)-(2). Here our arguments use the same technique as that found in [14]. We would like to point out that we were not able to extend the method to cover the case of zero damping, that is $\beta = 0$ in the semilinear wave equation (1)-(2).

In what follows, we assume the *dissipative condition*

$$\limsup_{|u| \rightarrow \infty} \frac{f(u)}{u} < 0, \quad (29)$$

holds. Now we can state the following theorem.

Theorem 4 *Consider the semilinear wave problem (1)-(2) and suppose in (4) that $p < 1 + \frac{2}{N}$. Then, there exists a non negative constant $C \geq 0$ such that if $\|(u_0^\varepsilon, u_1^\varepsilon)\|_{1/2,0,\varepsilon} \leq \rho$ for some $\rho > 0$ we have $\|(u^\varepsilon(t), u_t^\varepsilon(t))\|_{1/2,0,\varepsilon} \leq C$ for all $t \geq t_0(\rho)$. In other words, the semilinear problem is bounded dissipative and we also have*

$$\int_0^\infty \|u_t^\varepsilon\|_0^2 \leq C.$$

Proof. Let us consider, for all $t \geq 0$ the functional

$$J^\varepsilon(\psi, \psi_t) = \frac{1}{2} \|(\psi, \psi_t)\|_{1/2,0,\varepsilon}^2 + \frac{b}{2} \int_\Omega \psi \psi_t - \int_\Omega F(\psi), \quad \text{for } 0 < b < 1 \quad (30)$$

where $F(\psi) = \int_0^\psi f(s) ds$. Finding the time derivative of J^ε for $(\psi, \psi_t) = (u^\varepsilon, u_t^\varepsilon)$ solution to (11) we have

$$\begin{aligned} \frac{dJ^\varepsilon(u^\varepsilon, u_t^\varepsilon)}{dt} &= \frac{d}{2dt} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2,0,\varepsilon}^2 + \frac{b}{2} \int_\Omega |u_t^\varepsilon|^2 + \frac{b}{2} \int_\Omega u^\varepsilon u_{tt}^\varepsilon - \int_\Omega f(u^\varepsilon) u_t^\varepsilon \\ &= -\beta \int_\Omega |u_t^\varepsilon|^2 + \frac{b}{2} \int_\Omega |u_t^\varepsilon|^2 - \frac{b\beta}{2} \int_\Omega u^\varepsilon u_t^\varepsilon - \frac{b}{2} \|u^\varepsilon\|_{1/2,\varepsilon}^2 + \frac{b}{2} \int_\Omega f(u^\varepsilon) u^\varepsilon \\ &\leq -\frac{\beta}{2} \int_\Omega |u_t^\varepsilon|^2 - \frac{b}{2} \int_\Omega u^\varepsilon u_t^\varepsilon - \frac{b}{2} \|u^\varepsilon\|_{1/2,\varepsilon}^2 + \frac{b}{2} \int_\Omega f(u^\varepsilon) u^\varepsilon \\ &= -\frac{b}{2} \left(\int_\Omega |u_t^\varepsilon|^2 + \int_\Omega u^\varepsilon u_t^\varepsilon + \|u^\varepsilon\|_{1/2,\varepsilon}^2 \right) + \frac{b}{2} \int_\Omega f(u^\varepsilon) u^\varepsilon, \quad (31) \end{aligned}$$

for any $t \geq 0$. After, noticing the result of multiplying in (11) by u_t^ε and in the third term of the first line in the right hand side substituting the expression by its equivalent following a multiplication in (11) by u^ε , in both cases we take the integration by parts on Ω .

Next we observe that the Lemma 2.1 in [14] imply

$$\frac{1}{2} \left(\int_\Omega |u_t^\varepsilon|^2 + \int_\Omega u^\varepsilon u_t^\varepsilon + \|u^\varepsilon\|_{1/2,\varepsilon}^2 \right) \geq \frac{1}{4} \left(\int_\Omega |u_t^\varepsilon|^2 + \|u^\varepsilon\|_{1/2,\varepsilon}^2 \right),$$

hence we have in (31) for $t \geq 0$ that

$$\frac{dJ^\varepsilon(u^\varepsilon, u_t^\varepsilon)}{dt} \leq -\frac{b}{4} \left(\int_{\Omega} |u_t^\varepsilon|^2 + \|u^\varepsilon\|_{1/2, \varepsilon}^2 \right) + \frac{b}{2} \int_{\Omega} f(u^\varepsilon) u^\varepsilon.$$

Thanks to the dissipative hypothesis (29), for all $\eta > 0$ there exists $C_\eta \geq 0$ such that

$$\forall \varphi \in \mathbb{R} \Rightarrow f(\varphi)\varphi \leq \eta\varphi^2 + C_\eta. \quad (32)$$

Therefore, if $\eta = b/8$ we have

$$\frac{dJ^\varepsilon(u^\varepsilon, u_t^\varepsilon)}{dt} \leq -\frac{b}{8} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2 + C, \quad \forall t \geq 0. \quad (33)$$

On the other hand, (29) also imply that

$$\forall \varphi \in \mathbb{R} \Rightarrow F(\varphi) \leq \eta\varphi^2 + C_\eta$$

hence

$$\begin{aligned} J^\varepsilon(u^\varepsilon, u_t^\varepsilon) &\geq \frac{1}{2} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2 - \frac{b}{2} \|u^\varepsilon\|_0 \|u_t^\varepsilon\|_0 - \frac{1}{8} \int_{\Omega} |u^\varepsilon|^2 - C \\ &\geq \frac{1}{8} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2 - C \end{aligned} \quad (34)$$

for any $t \geq 0$. Analogously, we estimate (30) above to obtain

$$\begin{aligned} J^\varepsilon(u^\varepsilon, u_t^\varepsilon) &\leq \frac{1}{2} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2 + \frac{b}{2} \|u^\varepsilon\|_0 \|u_t^\varepsilon\|_0 - \int_{\Omega} F(u^\varepsilon) \\ &\leq \frac{3}{4} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2 + C \int_{\Omega} (|u^\varepsilon|^{p+1} + |u^\varepsilon|^2 + |u^\varepsilon|) \end{aligned} \quad (35)$$

$$\leq C \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2, 0, \varepsilon}^2 + C, \quad \forall t \geq 0. \quad (36)$$

This follows easily, on using the inequalities of Hölder, an adequate of Young and [3, 16] of Nirenberg-Gagliardo

$$\|u\|_{L^{p+1}(\Omega)} \leq C \|u\|_{1/2}^\alpha \|u\|_0^{(1-\alpha)} \quad \text{with} \quad \alpha = \frac{N}{2} - \frac{N}{p+1}. \quad (37)$$

In fact, since $p < 1 + \frac{2}{N}$ this implies $\alpha(p+1) < 1$. Hence, we can raise the above inequality (37) to the power $p+1$ on both sides. Then, using Young's inequality

$$ab \leq \frac{1}{s} a^s + \frac{1}{s'} b^{s'} \quad \text{with} \quad a, b \geq 0, s = \frac{2}{\alpha(p+1)} > 1 \text{ such that } \frac{1}{s} + \frac{1}{s'} = 1$$

it is estimated the first term in the second sum of (35). But since this is not as immediate because $s' \frac{(1-\alpha)(p+1)}{2} \neq 1$, we have to choose some $\vartheta > 1$ such that

$\frac{\alpha(p+1)}{2} + \frac{(1-\alpha)(p+1)}{2\vartheta} = 1$ and $\frac{(1-\alpha)(p+1)}{2\vartheta} s' = 1$, which yields $\vartheta = \frac{(1-\alpha)(p+1)}{2-\alpha(p+1)}$ thus after expressing

$$\int_{\Omega} |u|^{p+1} \leq C \left(\int_{\Omega} |\nabla u|^2 + \int_{\Omega} |u|^2 \right)^{\frac{\alpha(p+1)}{2}} \left(\int_{\Omega} |u|^{\frac{2}{\vartheta}} \right)^{\frac{(1-\alpha)(p+1)\vartheta}{2}}$$

it is possible to apply Young's inequality successively to furnish

$$\int_{\Omega} |u|^{p+1} \leq 2 \left(\int_{\Omega} |\nabla u|^2 + \int_{\Omega} |u|^2 \right) + C.$$

The following term of the second expression in (35) need not be estimated, while the last estimates easily as

$$\int_{\Omega} |u^\varepsilon| \leq |\Omega|^{1/2} \|u^\varepsilon\|_0 \leq \frac{1}{2} \|u^\varepsilon\|_0^2 + \frac{|\Omega|}{2}$$

by virtue of Hölder's and Young's inequalities. Finally, in the last estimate (36) the constant $C \geq 0$ is the maximum of the resulting constants of the computations following from (35).

Now with (36) in (33) we obtain

$$\frac{dJ^\varepsilon(u^\varepsilon, u_t^\varepsilon)}{dt} \leq -\frac{bC}{8} J^\varepsilon(u^\varepsilon, u_t^\varepsilon) + C.$$

Consequently, solving the differential inequality we find for $t \geq 0$ that

$$J^\varepsilon(u^\varepsilon, u_t^\varepsilon) \leq e^{-\frac{bC}{8}t} J^\varepsilon(u_0^\varepsilon, u_1^\varepsilon) + C(1 - e^{-\frac{bC}{8}t}).$$

It follows, again by (36) and the hypotheses on the initial data, taking into account (34) that

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2,0,\varepsilon}^2 \\ & \leq \limsup_{t \rightarrow \infty} \left\{ e^{-\frac{bC}{8}t} J^\varepsilon(u_0^\varepsilon, u_1^\varepsilon) + C(1 - e^{-\frac{bC}{8}t}) \right\} + C \leq C \end{aligned}$$

and the first assertion is proved.

To complete the demonstration, we observe in (30) that when $b = 0$ we have the classical Lyapunov's functional. Moreover, from the second estimate in (31) we readily see for any solution $(u^\varepsilon, u_t^\varepsilon)$ of (11) through $(u_0^\varepsilon, u_1^\varepsilon)$ that

$$\frac{dJ^\varepsilon(u^\varepsilon, u_t^\varepsilon)}{dt} = -\beta \int_{\Omega} |u_t^\varepsilon|^2 \leq 0 \Rightarrow J^\varepsilon(u^\varepsilon, u_t^\varepsilon) \leq J^\varepsilon(u_0^\varepsilon, u_1^\varepsilon)$$

for all $t \geq 0$. Next using (36) we find

$$\int_0^\infty \|u_t^\varepsilon\|_0^2 \leq 2J^\varepsilon(u_0^\varepsilon, u_1^\varepsilon) \leq C\rho + C,$$

and the demonstration of the theorem is complete. \square

Our next result is derived from Theorem 4. It is deduced similarly as results in [9, 12, 14] and references therein and its proof will resemble one found in [9]. The advantage of the method used in [9] is that we obtain compactness of the attractor by means of a density argument. Moreover, in our case, since in (1)-(2) we are considering a linear damping term is even much simpler.

Corollary 5 *The dynamical system generated by the problem (1)-(2) has a global compact attractor $\mathcal{A}^\varepsilon \subset X_\varepsilon^{1/2} \times X_\varepsilon^0$.*

Proof. Let us consider the semilinear wave equation (11) for sufficiently regular data $d^\varepsilon, b^\varepsilon \in C^1$, $(u_0^\varepsilon, u_1^\varepsilon) \in \mathcal{D} = C_0^\infty(\Omega)$, and denote by $u^\varepsilon = v^\varepsilon + w^\varepsilon$ the unique strong solution to the equation.

Assume the solution components satisfy

$$\begin{aligned} v_{tt}^\varepsilon + \beta v_t^\varepsilon + L^\varepsilon v^\varepsilon &= 0 & v^\varepsilon(0) &= u_0^\varepsilon, & v_t^\varepsilon(0) &= u_1^\varepsilon \\ w_{tt}^\varepsilon + \beta w_t^\varepsilon + L^\varepsilon w^\varepsilon &= f(u^\varepsilon), & w^\varepsilon(0) &= w_t^\varepsilon(0) = 0. \end{aligned} \quad (38)$$

Then, in the setting of Theorem 4 consider the equation in w^ε . Let

$$p = \frac{r}{q} > 1 \quad \text{for some } r \geq \frac{2N+4}{N} \text{ fixed.}$$

Since the inclusions $H^1(\Omega) \hookrightarrow L^s(\Omega)$ are compact, for $1 \leq s < \frac{2N}{N-2}$ then [20] Chapter I.5 the imbeddings

$$L^\infty(0, T, H^1(\Omega)) \cap W^{1,\infty}(0, T, L^2(\Omega)) \hookrightarrow L^m(0, T, L^s(\Omega)) \quad (39)$$

for any $1 < m < \infty$ are also compact. Thus, the nonlinearity $f : L^r(Q_T) \mapsto L^q(Q_T)$ is well defined and compact in $L^{\frac{Nr}{N+2}}(Q_T)$. Further, from the energy associated we have the mapping

$$L^{\frac{Nr}{N+2}}(Q_T) \ni f(u^\varepsilon) \mapsto (w^\varepsilon, w_t^\varepsilon) \in X_\varepsilon^{1/2} \times X_\varepsilon^0$$

is continuous. Now, using the linear semigroup

$$(v^\varepsilon(t), v_t^\varepsilon(t)) = T^\varepsilon(t)(v_0^\varepsilon, v_1^\varepsilon), \quad t \geq 0$$

we obtain

$$\|(v^\varepsilon, v_t^\varepsilon)\|_{1/2,0,\varepsilon} \leq M e^{-\frac{\beta}{2}t} \|(v_0^\varepsilon, v_1^\varepsilon)\|_{1/2,0,\varepsilon}, \quad \forall t \geq 0,$$

where

$$\frac{\beta}{2} = \operatorname{Re} \left(\frac{\beta \pm \sqrt{\beta^2 - 4\mu_1^\varepsilon}}{2} \right) \quad (40)$$

with $\mu_1^\varepsilon \in \sigma(L^\varepsilon)$ the first eigenvalue to the problem $L^\varepsilon \psi = \mu \psi$, $\psi \in H^1(\Omega)$.

Next let us consider

$$\mathcal{A}^\varepsilon = \bigcap_{t \geq 0} \overline{T^\varepsilon(t)\mathcal{D}} \quad \text{with } T^\varepsilon(t)\mathcal{D} \stackrel{def}{=} (w^\varepsilon(t), w_t^\varepsilon(t))$$

being the strong solution of (11) for regular data $(u_0^\varepsilon, u_1^\varepsilon) \in \mathcal{D}$. Then, clearly \mathcal{A}^ε is a closed set in $X_\varepsilon^{1/2} \times X_\varepsilon^0$, moreover from the above is compact. Thus by density we have a global compact attractor $\mathcal{A}^\varepsilon = \bigcap_{t \geq 0} \overline{T^\varepsilon(t)\mathcal{B}}$ in $X_\varepsilon^{1/2} \times X_\varepsilon^0$ here \mathcal{B} denote an absorbing set for the nonlinear semigroup generated by (11). This way we have complete the proof of the corollary. \square

6 Limite problem of large diffusion

In this section, we complete the study to our problem. We shall prove as consequence of the limiting problem for large diffusivity of (1)-(2) we obtain the existence of a limiting attractor in the topology of the energy space $X_\varepsilon^{1/2} \times X_\varepsilon^0$. In this way we are able to relate the dynamical structures of the geometry of solutions of the equations. Now, we shall prove the following theorem.

Theorem 6 *Let $(u^\varepsilon, u_t^\varepsilon)$ denote the solution of the semilinear wave problem (1)-(2). Also let $(u_\Omega, \dot{u}_\Omega)$ denote the solution of*

$$\ddot{u}_\Omega + \beta \dot{u}_\Omega + \left(\frac{|\Gamma|}{|\Omega|} \int_\Gamma b + \lambda \right) u_\Omega = f_\Omega(u), \quad (u_\Omega^0, u_\Omega^1) \in \mathbb{R}^2. \quad (41)$$

Assume, the hypotheses in Theorem 4 are satisfied. In addition, let there be $u_1^\varepsilon \rightarrow u_1^1$ strongly in $L^2(\Omega)$ as $\varepsilon \rightarrow 0$. Then, for all $t \geq 0$ we have

$$(u^\varepsilon, u_t^\varepsilon) \rightarrow (u_\Omega, \dot{u}_\Omega) \quad \text{strongly in } X_\varepsilon^{1/2} \times X_\varepsilon^0 \quad (42)$$

as $\varepsilon \rightarrow 0$. In particular, the family of attractors $\{\mathcal{A}^\varepsilon \cup \mathcal{A}\}_{\varepsilon > 0}$, where \mathcal{A} is a global attractor for the equation (41) verify

$$\lim_{\varepsilon \rightarrow 0} \sup_{(u^\varepsilon, u_t^\varepsilon) \in \mathcal{A}^\varepsilon} \inf_{(u_\Omega, \dot{u}_\Omega) \in \mathcal{A}} \|(u^\varepsilon, u_t^\varepsilon) - (u_\Omega, \dot{u}_\Omega)\|_{1/2,0,\varepsilon} = 0 \quad (43)$$

in other words, is upper semicontinuous in $\varepsilon = 0$.

Proof. Consider the Energy identity

$$\begin{aligned} \frac{1}{2} \|(u^\varepsilon, u_t^\varepsilon)\|_{1/2,0,\varepsilon}^2 + \beta \int_{Q_T} |u_t^\varepsilon|^2 \\ = \int_{Q_T} f(u^\varepsilon) u_t^\varepsilon + \frac{1}{2} \|(u_0^\varepsilon, u_1^\varepsilon)\|_{1/2,0,\varepsilon}^2. \end{aligned} \quad (44)$$

Since

$$f(u^\varepsilon) \in L^q(Q_T) \quad \text{and} \quad u_t^\varepsilon \in L^\infty(0, T, L^2(\Omega)) \subset L^\infty(0, T, L^{q'}(\Omega))$$

we have using Hölder's inequality and Young's (19) that

$$\left| \int_{Q_T} f(u^\varepsilon) u_t^\varepsilon \right| \leq 4C \|f(u^\varepsilon)\|_{q,Q_T}^2 + \frac{1}{4} \sup_{0 \leq t \leq T} \|u_t^\varepsilon\|_2^2.$$

Consequently, in (44) we obtain

$$\begin{aligned} \|(u^\varepsilon, u_i^\varepsilon)\|_{1/2,0,\varepsilon}^2 + \beta \int_{Q_T} |u_i^\varepsilon|^2 \\ \leq C \left(\|f(u^\varepsilon)\|_{q,Q_T}^2 + \|(u_0^\varepsilon, u_1^\varepsilon)\|_{1/2,0,\varepsilon}^2 \right). \end{aligned} \quad (45)$$

Let $r \geq \frac{2N+4}{N}$, since (9) imply $\|u^\varepsilon\|_{1/2,\varepsilon}^2 \rightarrow \infty$ as $\varepsilon \rightarrow 0$ then (37) conclude $f(u^\varepsilon) \in L^q(Q_T)$ is bounded in norm for all $\varepsilon > 0$. Therefore, from (45) we have the convergence including (22). Moreover, using (39) we obtain

$$f(u^\varepsilon) \rightarrow f(v) \quad \text{strongly in } L^{\frac{Nr}{N+2}}(Q_T) \text{ as } \varepsilon \rightarrow 0.$$

It now suffices to observe that by hypothesis on the initial data we have

$$u_0^\varepsilon \rightarrow u_\Omega^0 \quad \text{strongly in } X_\varepsilon^{1/2} \text{ as } \varepsilon \rightarrow 0$$

and we can apply the second part of Theorem 2 to obtain,

$$\text{the limit } (v, \dot{v}) = (u_\Omega, \dot{u}_\Omega) \text{ is strong in } Y_T.$$

Moreover, it verify (41) in the sense of distributions for all $t \geq 0$, this proves the first part of the theorem (42). It remains to prove its particular conclusion.

To do this, we observe that the dissipative condition (29) also holds in the vector field of the equation (41). This imply the semigroup generated by the finite dimensional equation is bounded and point dissipative. Thus [12] conclude the equation has a global compact attractor $\mathcal{A} \subset \mathbb{R}^2$. Since, the above limiting process remains true on the family of attractors $\{\mathcal{A}^\varepsilon \cup \mathcal{A}\}_\varepsilon$. Then, using [12] Sect. 4.10.2 p.165 we find that the orbits on these attractors satisfy (43). It follows now we have our theorem is proved. \square

7 Conclusive Remarks

For a good theory on second order ordinary differential equations we cite as reference the text [11]. Now, in a backward analysis of our results on the wave equation (1)-(2) with large diffusion in the above Theorem 6, we conclude from (43) that the long time dynamics generated by its solutions on the attractors in the limit (12) of large diffusion, coincide with the dynamics of the solutions on the attractor of the finite dimensional differential equations in (41). Further in retrospect, it should be noted that, the regularity of the data to our problem (1)-(2) does not allow us to obtain a fine exponential decay estimate of (16) as in the parabolic case studied in [7]. The quantity (16) tells us that in the given model of the wave equation (1)-(2) spatial homogenization of solutions to a constant function in the domain occur, uniformly in time if the additional hypothesis following in the comments after the proof of Theorem 2 is verified, and is sufficiently rapid.

In other words, if we view $E^\varepsilon(t), t \geq 0$ given as in Theorem 3 as a function representing the spatial heterogeneities (see for example [23]) of the solution $u^\varepsilon(x, t)$ to the nonlinear wave problem (1)-(2) in model hypotheses for varied $\varepsilon \rightarrow 0$, then (16) imply that these are eliminated much quicker in the sense of an integral in the domain, than the diffusion tends to infinity. In this sense the physical meaningful hypotheses on the data although being mild ones, may be seen as restricting the effects of large diffusion in the limit case. Thus apparently the results become natural. However, embodied as we appreciate in much refined estimates and arguments in the proofs of our theorems asserting the limiting problems that the boundedness assumptions in norms for all $\varepsilon > 0$ on the data can be lifted and our results still remain true.

The preceding in consequence, establish that the asymptotic limiting wave equation is described by those processes defined in the limit of large diffusion of the spatial average solutions to the original semilinear wave equation (1)-(2) under only model convergence without boundedness in norms of the data uniformly in $\varepsilon > 0$ hypotheses as given after hypothesis (12). This thus extend the given case we have had in consideration following the physical relevance of the asymptotic limit problem.

7.1 Generalized equation and boundary conditions.

As we have already mentioned in the introductory section of this paper, that it is only a provision of the initial step and a revisit to the study of the semilinear wave equation. In its subsequent one, we intend to give an analogue dynamical and large diffusion limit asymptotic study, which is further to include the study of the blow-up phenomenon of the solutions, to the following nonlinear equation

$$u_{tt} + \beta(u_t) - \text{Div}(d^\varepsilon(x)\nabla u) + (V^\varepsilon(x) + \lambda)u = f(u) \quad \text{in } Q_T, \quad (46)$$

where the damping term is nonlinear $\beta \in C^1(\mathbb{R})$, $\beta(0) = 0$ is strictly increasing, in addition satisfy given growth conditions and the potential

$$V^\varepsilon \in L^{p_0}(\Omega) \quad \text{with } p_0 \geq 1 \text{ such that similar exponents relation as in (3)}$$

are satisfied for Sobolev inclusions into the spaces $L^r(\Omega)$. The equation (46) will be supplemented by lateral boundary surface and initial data conditions of the form

$$\sigma u_t + \partial_{\bar{n}_\varepsilon} u + b^\varepsilon u = g(u) \quad \text{on } \Sigma_T, \quad (u(0), u_t(0)) = (u_0^\varepsilon, u_1^\varepsilon) \quad \text{in } \bar{\Omega}. \quad (47)$$

Here either $\sigma = 0$ or exclusive $\sigma \geq 1$ and the boundary nonlinearity of at least class C^1 to be assumed to verify certain growth behaviors. It is worthwhile observing that in the equation (47) if we consider the case $\sigma = 0$ then the conditions are standard nonlinear for the equation (46) on the lateral surface of the cylinder, whereas if $\sigma \geq 1$ they take into account a boundary damping in the equation and are less familiar in the literature and in research notes.

The extensiveness of the analytic details needed for a complete self-contained study of the above nonlinear wave equation (46) - (47) in both its linear

nonhomogeneous counterpart or in its given formulated contexts, is far above the scope of the actual paper and for this reason we have decided to reserve its treatment in an independent future study.

8 Acknowledgements

The author would like to thank Professor A. Rodríguez Bernal (U.C.M. Mat.Apl.Spain) for introducing the problem to him and for providing him with the reference [6]. He truly would like to thank the referees for useful stimulating comments and suggestions. More important for recommending a detailed introduction. Without which, the paper would not have assumed its final state. The author is an A.E.C.I. Spain Fellow.

References

- [1] J.Arrieta, A.R. Bernal and A.N. Carvalho. Upper Semicontinuity of Attractors for Parabolic problems with Localized large diffusion and Nonlinear Boundary Conditions. *Journal of Differential Equations* 168: 33-59, 2000.
- [2] A.V. Babin & M.L. Vishik. Regular Attractors of Semigroups and Evolution Equations. *Jrnl. Math. Pures et Appl.* 62: 441-291, 1983.
- [3] A.R. Bernal. *Ecuaciones de Evolución Semilinear. Curso 96-97*. Universidad Complutense de Madrid, Matemática Aplicada, 28040 Madrid.
- [4] A.R. Bernal and R.Willie. Singular Large Diffusivity and Spatial Homogenization in Linear Parabolic Problem. In review *Series B. Journal of Discrete and Continuous Dynamical Systems*: July, 2002.
- [5] H.Brézis. *Analisis Funcional, Alianza Universidad Textos*. Masson Editeur, de Paris. 1983.
- [6] A.N. Carvalho. Spatial Homogeneity in Damped Hyperbolic Equations. Georgia Institute of Technology. CDSNS preprint series: 93-145, 1991.
- [7] E.Conway, D. Hoff and J. Smoller. Large Time Behaviour of Solutions of Systems of Nonlinear Reaction - Diffusion Equations. *Siam. Jrnl. Appl. Math.* 35(1): 1-16, 1978.
- [8] S.Cox & E. Zuazua. The Rate at which Energy decays in a Damped Wave String. *Comm. Partial Diff. Eqs.* 19: 213-243, 1994.
- [9] E.Feireisl. Global Attractors for Semilinear Damped Wave Equations with Supercritical Exponent. *Journal of Differential Equations* 116: 431-447, 1995.

- [10] V. Georgiev and G. Todorova, Existence of a Solution of the Wave Equation with Nonlinear Damping and Source Terms. *Journal of Differential Equations* 109: 295-308, 1994.
- [11] J.K.Hale. *Ordinary Differential Equations*. 2nd Kreiger, 1978.
- [12] J.K.Hale. *Asymptotic Behaviour of Dissipative Systems*. Mathematical Monographs, Vol 25, American Mathematics Society. Providence, RI 1988.
- [13] J.K.Hale. Large Diffusivity and Asymptotic Behavior in Parabolic Systems. *Jrnl. Math. Anal. & Appl.* 118: 455-466, 1986.
- [14] J.K. Hale and G. Raugel. Upper Semicontinuity of the Attractors for a Singularly Perturbed Hyperbolic Equation. *Journal of Differential Equations* 73: 197-214, 1988.
- [15] J. K. Hale and C. Rocha. Varying Boundary Conditions with Large Diffusivity. *Jrnl. Math. Pures et Appl.* 66: 139-158, 1987.
- [16] D.Henry. *Geometric Theory of Semilinear Parabolic Equations*. Lecture Notes in Mathematics 840 Springer- Verlag 1981.
- [17] F. John. *Partial Differential Equations*, Springer-Verlag. New York 1982.
- [18] J. López Gómez. On the Linear Damped Wave Equation. *Journal of Differential Equations* 134: 26-45, 1997.
- [19] I. Lasiecka, J.L. Lions and R. Triggiani. Non Homogenous Boundary Value Problems for Second order Hyperbolic Operators. *Jrnl. Math. Pures et Appl.* 65: 149-192, 1986.
- [20] J.L. Lions. *Quelques méthodes de résolution des problèmes aux limites nonlinéaires*. Dunod Gáuthier -Villars, Paris 1969.
- [21] J.L. Lions and E. Magenes. *Non Homogeneous Boundary Value Problems and Applications I*, Springer Verlag. 1961.
- [22] J.S. Morlaes and M. Valencia. Trend of Spatial Homogeneity for Solutions of a Semilinear Damped Wave Equations. *Sección de Matemática, Universidad Autonoma de Barcelona, Preprint*. 1986.
- [23] J.D.Murray. *Mathematical Biology*. Bio-Maths.Texts 19, Springer-Verlag, 1993.
- [24] J. Smoller. *Shock Wave and Reaction -Diffusion Equations*. Springer Verlag, New York, 1983.
- [25] M. Renardy and R.C. Rogers. *An Introduction to Partial Differential Equations*. Springer-Verlag, Texts in Applied Mathematics 13, 1993.

- [26] R. Willie. *Ecuaciones de Reacción-Difusión con Difusión Alta*. Tesis Doctoral, Universidad Complutense de Madrid, Matemática Aplicada, 28040 Madrid. 2000.
- [27] R. Willie. A Semilinear System of Reaction and Diffusion Equations with Nonlinear Boundary Conditions and Large Diffusion. *To appear in Journal of Dynamics and Differential equations*. 2004.
- [28] R. Willie. Upper Semi-continuity of Attractors for Reaction and Diffusion System of equations in Sup- Norm. *Abstract and Applied Analysis. Hindawi Publishing Corporation. Galley proofs details-* 2003(00): 1-25 2003.
- [29] R. Willie. The Inertial Manifold for A Nonlinear Reaction-Diffusion System with Large Diffusion. Submitted in *Jrnl. Diff. Eqs:* April, 2003.
- [30] R. Willie. Large Diffusivity Finite Dimensional Asymptotic Behaviour of a Semilinear Wave Equation. *Journal of Applied Mathematics. Hindawi publishing corporation*. 8 (2003): 409-427, 2003.
- [31] E. C. Zachmanoglou and D. W. Thoe. *Introduction to Partial Differential Equations with Applications*, The Williams & Wilkins company. 1976.

Autómatas celulares elementales reversibles *

A. MARTÍN

Departamento de Matemática Aplicada, Universidad de Salamanca

delrey@usal.es

Resumen

En este trabajo se resuelve el problema de la reversibilidad para el autómata celular elemental con número de Wolfram 150; esto es, se demuestra que la reversibilidad de dicho autómata depende del número de células que posea y se calcula explícitamente en dichos casos el autómata celular inverso.

Palabras clave: *Autómatas Celulares, Reversibilidad*

Clasificación por materias AMS: *68Q80, 68Q70*

1 Introducción

Un autómata celular booleano unidimensional (AC para abreviar) es un tipo particular de sistema dinámico discreto formado por un número finito de objetos idénticos entre sí, denominados células, que se encuentran dispuestos uno a continuación de otro a modo de cadena lineal y que poseen un determinado estado: 0 ó 1, que va cambiando con el discurrir discreto del tiempo. El estado de cada célula en un determinado instante de tiempo se calcula a partir de una regla de transición local cuyas variables son los estados de un conjunto de células (denominado vecindad de la célula considerada) en el instante de tiempo inmediatamente anterior. Un tipo particularmente interesante de AC son los denominados autómatas celulares elementales (ACE), en los que la vecindad de cada célula está formada por ella misma y las dos células adyacentes a la dada (a derecha e izquierda).

*Este trabajo ha sido parcialmente subvencionado por la Consejería de Educación y Cultura de la Junta de Castilla y León, bajo el proyecto SA052/03 y por el Ministerio de Educación y Ciencia bajo el proyecto SEG2004-02418. El autor quiere agradecer a Y. Cortés Gómez el trabajo realizado en la elaboración de las figuras.

Fecha de recepción: 4/11/04

Un AC se dice que es reversible cuando existe otro autómatas celular, denominado inverso, que realiza la computación inversa al primero. Este tipo de sistemas dinámicos discretos preservan la información dada por los estados iniciales a lo largo de toda su evolución y son particularmente interesantes en lo relativo a sus aplicaciones a la criptografía, la computación, etc.

Una cuestión muy interesante es saber qué tipo de reglas de transición dan lugar a ACs reversibles. En este sentido, existe un algoritmo para decidir tal cuestión para ACs unidimensionales ([1]), pero no existe ningún procedimiento que nos permita calcular explícitamente el AC inverso. Hay que indicar que en el caso de considerar AC bidimensionales o de dimensión superior, no es posible encontrar un algoritmo semejante ([6, 7]).

El principal objetivo de este trabajo es el estudio de la reversibilidad de uno de los AC elementales más importantes: el ACE con número de Wolfram 150. Su importancia radica en sus múltiples aplicaciones a los más diversos campos de la ciencia: criptografía y teoría de códigos ([2]), generación de secuencias pseudoaleatorias ([4, 5]), estudio del ADN ([10]), etc.

El resto del trabajo se organiza como sigue: en la Sección 2 se introducen los principales conceptos y resultados relacionados con los autómatas celulares elementales reversibles; en la Sección 3, se realiza la caracterización de la reversibilidad del ACE con número de Wolfram 150; en la Sección 4 se muestran algunos ejemplos y, finalmente, en la Sección 5 se presentan las conclusiones.

2 Autómatas celulares elementales reversibles

Los autómatas celulares elementales ([9]) son sistemas dinámicos discretos formados por un conjunto de n objetos, C , indistinguibles entre sí, denominados células:

$$C = \{\langle i \rangle, 0 \leq i \leq n - 1\},$$

las cuales están dispuestas uniformemente una a continuación de otra, y tales que están dotadas de un estado de entre un conjunto finito: $S = \mathbb{Z}_k$. En este artículo trabajaremos en el caso $k = 2$, con lo que $S = \mathbb{Z}_2$ y nos encontraremos así ante los llamados autómatas celulares elementales. Además denotaremos por $a_i^{(t)}$ al estado de la célula $\langle i \rangle$ en el instante de tiempo t .

Cada célula cambia su estado en pasos discretos de tiempo según una regla determinista denominada función de transición local. De esta forma, el estado de la célula $\langle i \rangle$ en el instante de tiempo $t + 1$, dependerá de los estados en el instante de tiempo t de un conjunto de células denominado vecindad:

$$V = \{\langle i - 1 \rangle, \langle i \rangle, \langle i + 1 \rangle\}.$$

A este tipo de vecindad se la denomina vecindad simétrica de radio 1; en general, una vecindad simétrica de radio r viene dada por:

$$V = \{\langle i - r \rangle, \dots, \langle i \rangle, \dots, \langle i + r \rangle\}.$$

Consecuentemente, la función de transición local de todo ACE será:

$$a_i^{(t+1)} = f \left(a_{i-1}^{(t)}, a_i^{(t)}, a_{i+1}^{(t)} \right), \quad 0 \leq i \leq n-1.$$

El vector $C^{(t)} = \left(a_0^{(t)}, \dots, a_{n-1}^{(t)} \right) \in \mathbb{Z}_2^n$ se denomina configuración en el instante t del ACE.

Como estamos suponiendo que el número de células del ACE es finito, entonces es necesario indicar una serie de condiciones de contorno para que la dinámica de la evolución sea correcta. Así, estableceremos las denominadas condiciones de contorno periódicas, según las cuales si $i \equiv j \pmod{n}$, entonces $a_i^{(t)} = a_j^{(t)}$, con $0 \leq i, j \leq n-1$.

Obsérvese que como la vecindad está formada por 3 células, entonces existen 8 posibles configuraciones de tales vecindades según los estados que adopten dichas células. Consecuentemente, se pueden construir $2^{2^3} = 256$ funciones de transición local, cada una de las cuales tiene asociado un número entero denominado número de Wolfram, $w = \sum_{i=0}^7 f_i \cdot 2^i$, donde

$$\begin{aligned} f_0 &= f(0, 0, 0), \quad f_1 = f(0, 0, 1), \quad f_2 = f(0, 1, 0), \quad f_3 = f(0, 1, 1), \\ f_4 &= f(1, 0, 0), \quad f_5 = f(1, 0, 1), \quad f_6 = f(1, 1, 0), \quad f_7 = f(1, 1, 1). \end{aligned}$$

Un ACE se dice que es lineal cuando su función de transición local es lineal, esto es, de la forma:

$$a_i^{(t+1)} = \alpha a_{i-1}^{(t)} + \beta a_i^{(t)} + \gamma a_{i+1}^{(t)} \pmod{2}, \quad 0 \leq i \leq n-1, \quad (1)$$

donde $\alpha, \beta, \gamma \in \mathbb{Z}_2$. Es inmediato ver que, en función de los valores que toman dichos parámetros, existen $2^3 = 8$ ACE lineales, que son dados en la siguiente tabla:

α	β	γ	w
0	0	0	0
1	1	0	60
1	0	1	90
0	1	1	102
1	1	1	150
0	0	1	170
0	1	0	204
1	0	0	240

La característica fundamental de los ACE lineales es que pueden ser interpretados en términos del Álgebra Lineal (ver, por ejemplo [2, 3, 8]). Así, consideremos el ACE lineal cuya función de transición local es la dada por la ecuación (1), entonces su evolución viene dada por

$$C^{(t+1),T} = M_n \cdot C^{(t),T} \pmod{2}, \quad (2)$$

donde $C^{(t),T}$ es la matriz traspuesta de $C^{(t)}$, y

$$M_n = \begin{pmatrix} \beta & \gamma & 0 & \cdots & 0 & \alpha \\ \alpha & \beta & \gamma & \cdots & 0 & 0 \\ 0 & \alpha & \beta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta & \gamma \\ \gamma & 0 & 0 & \cdots & \alpha & \beta \end{pmatrix} \quad (3)$$

se denomina matriz de transición del ACE.

Además podemos modelizar la dinámica de los ACE en términos de polinomios. En este sentido, la configuración de un ACE lineal en el instante t , $C^{(t)}$, puede ser representada por el denominado polinomio característico en el instante t , $P^{(t)}(x)$, cuya expresión es la siguiente:

$$P^{(t)}(x) = \sum_{i=0}^{n-1} a_i^{(t)} x^i,$$

donde $C^{(t)} = (a_0^{(t)}, \dots, a_{n-1}^{(t)}) \in \mathbb{Z}_2^n$.

Las condiciones de contorno periódicas son implementadas reduciendo el polinomio $P^{(t)}(x)$ módulo $x^n - 1$ en cada paso de tiempo. Es más, la evolución temporal puede ser representada multiplicando el polinomio característico por un dipolinomio fijo, $T(x)$, denominado dipolinomio de transición:

$$\begin{aligned} P^{(t+1)}(x) &= T(x) \cdot P^{(t)}(x) \pmod{(x^n - 1)}, \\ T(x) &= \alpha x + \beta + \gamma x^{-1}, \quad \alpha, \beta, \gamma \in \mathbb{Z}_2, \end{aligned}$$

donde la aritmética es realizada en \mathbb{Z}_2 .

En este trabajo consideraremos el ACE lineal con número de Wolfram 150, para el cual $\alpha = \beta = \gamma = 1$, y consecuentemente su función de transición local es:

$$a_i^{(t+1)} = a_{i-1}^{(t)} + a_i^{(t)} + a_{i+1}^{(t)} \pmod{2}, \quad 0 \leq i \leq n-1.$$

Consecuentemente su dipolinomio de transición es $T(x) = x^{-1} + 1 + x$.

Se llama función de transición global de un ACE a la transformación lineal

$$\begin{aligned} \Phi: \mathbb{Z}_2^n &\rightarrow \mathbb{Z}_2^n \\ C^{(t)} &\mapsto \Phi(C^{(t)}) = C^{(t+1)} \end{aligned}$$

Usando esta función podemos representar todo ACE mediante un grafo dirigido no conexo con 2^n vértices. Los vértices representarán las distintas configuraciones del autómata celular y entre dos de ellos, C y C' , existirá una arista (de C a C') cuando $\Phi(C) = C'$. En este caso, se dirá que C es un ancestro de C' y que C' es un sucesor de C .

En la Figura 1 se muestra el grafo asociado al ACE lineal con $w = 150$ y $n = 5$.

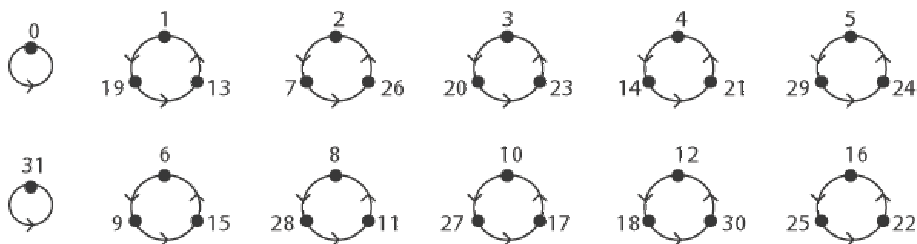


Figura 1: Grafo asociado al ACE $w = 150$ con $n = 5$

Por otro lado, en la Figura 2 podemos ver el grafo asociado al mismo autómata celular pero con $n = 6,5$ células.

Obsérvese que para simplificar la notación, cada configuración, $C = (c_0, \dots, c_{n-1}) \in \mathbb{Z}_2^n$, es denotada por un número entero $0 \leq z_C \leq 2^n - 1$, cuyo código binario es la secuencia (c_0, \dots, c_{n-1}) , esto es:

$$z_C = c_0 \cdot 2^{n-1} + c_1 \cdot 2^{n-2} + \dots + c_{n-2} \cdot 2 + c_{n-1}.$$

Una propiedad muy importante de los autómatas celulares en general, y de los ACE en particular, es la reversibilidad, particularmente cuando el ACE es utilizado para la modelización de fenómenos físicos. Así, cuando la función de transición global de un ACE, Φ , es biyectiva, existe otro autómata celular (que, en general, no es elemental), denominado inverso, cuya función de transición global es Φ^{-1} y que, obviamente, realiza la evolución inversa al primero ([11]). En este caso, al primer ACE se le llama reversible. Obsérvese que la reversibilidad de un ACE implica que en su grafo dirigido asociado, todo vértice no tiene más de un ancestro. Como ejemplo se puede observar que el ACE lineal $w = 150$ de $n = 5$ células es reversible ya que en su grafo asociado (véase Figura 1) todo vértice posee como mucho un ancestro; por su parte, el mismo ACE con $n = 6$ células no es reversible ya que en su grafo asociado (véase Figura 2), existen vértices con cuatro ancestros (las configuraciones representadas por los enteros $z_C = 0, 9, 14, 18, 21, 27, 35, 36, 42, 45, 54, 56, 63$).

La noción de reversibilidad de un ACE lineal está íntimamente relacionada con el carácter de su matriz de transición. Así, si dicha matriz de transición es no singular, entonces el correspondiente ACE lineal es reversible (ver [2]). Además, en dicho caso, el autómata celular inverso será también lineal aunque en general, no va a ser un autómata celular elemental: las vecindades no se conservarán.

La propiedad de la reversibilidad puede ser expresada en términos de polinomios como sigue:

Proposición 1 ([8]) *Un ACE con dipolinomio de transición $T(x)$ es reversible si y sólo si existe otro AC lineal booleano, con dipolinomio de transición $\tilde{T}(x)$, tal que*

$$T(x) \cdot \tilde{T}(x) = 1 \pmod{(x^n - 1)}.$$

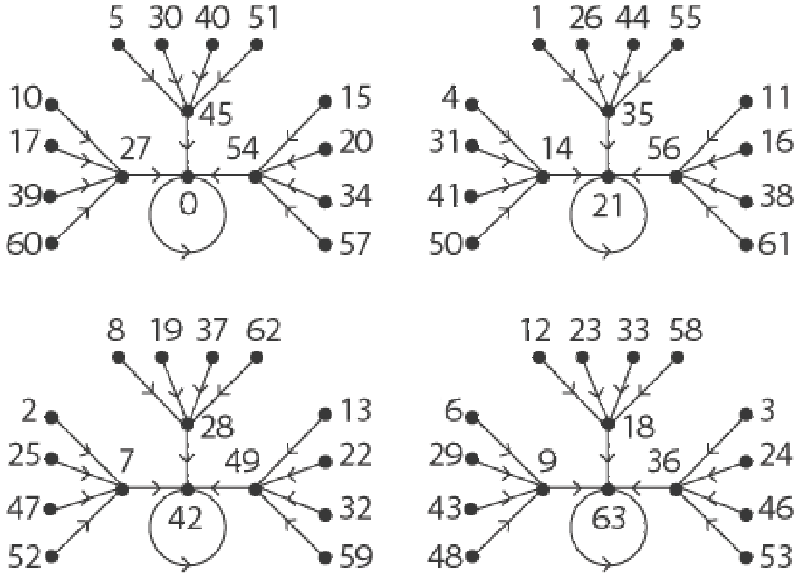


Figura 2: Grafo asociado al ACE $w = 150$ con $n = 6$

3 Caracterización de la reversibilidad

La reversibilidad del ACE lineal con número de Wolfram 150 depende del número de células que lo forman. Así, se tiene el siguiente resultado:

Teorema 2 *El ACE lineal con número de Wolfram $w = 150$ es reversible si y sólo si $n \not\equiv 0 \pmod{3}$, siendo $n \geq 3$.*

Demostración. La evolución de dicho ACE viene dada por (2) donde M_n es la matriz (3) con $\alpha, \beta, \gamma = 1$. Consideremos sus vectores fila:

$$\begin{aligned}
 u &= (1, 1, 0, \overset{(n-3)}{\dots}, 0, 1), \\
 e_i &= (0, \overset{(i-1)}{\dots}, 0, 1, 1, 1, 0, \overset{(n-2-i)}{\dots}, 0), \quad 1 \leq i \leq n-2, \\
 v &= (1, 0, \overset{(n-3)}{\dots}, 0, 1, 1).
 \end{aligned}$$

Es fácil ver que los vectores $\{e_1, \dots, e_{n-2}\}$ son linealmente independientes y que, además:

$$u = e_1 + \sum_{i=1}^{n/3-1} (e_{3i+1} + e_{3i}), \quad v = \sum_{i=0}^{n/3-2} (e_{3i+1} + e_{3i+2}) + e_{n-2},$$

lo cual ocurre si y sólo si $n \equiv 0 \pmod{3}$. □

Además podemos determinar explícitamente los autómatas celulares inversos en términos de su dipolinomio de transición. Así, y antes de ver dicha

caracterización, una sencilla comprobación nos permite obtener el siguiente resultado:

Lema 3 Consideremos para cada $m \in \mathbb{Z}^+$ el dipolinomio:

$$D_m(x) = 1 + \sum_{j=1}^m (x^{-j} + x^j).$$

Las siguientes propiedades se satisfacen en \mathbb{Z}_2 :

- (i) $(x^{-1} + x) D_m(x) = x^{-(m+1)} + x^{-m} + x^m + x^{m+1}.$
- (ii) $D_m(x) + D_{m+1}(x) = x^{-(m+1)} + x^{m+1}.$

Consecuentemente:

Teorema 4 Consideremos el ACE lineal con número de Wolfram $w = 150$, tal que $n \not\equiv 0 \pmod{3}$ y $n \geq 3$. Entonces la función de transición local del AC inverso es:

- (i) Si $n \equiv 1 \pmod{3}$ con $n = 3k + 1$, $k \in \mathbb{Z}^+$,

$$a_i^{(t+1)} = \sum_{j \in I} a_{i+j}^{(t)} \pmod{2}, \quad 0 \leq i \leq n-1, \quad (4)$$

donde I es el conjunto de índices siguiente:

$$I = \left\{ 0, \pm \left\lfloor \frac{3}{2} \right\rfloor, \pm \left\lfloor \frac{6}{2} \right\rfloor, \dots, \pm \left\lfloor \frac{3(k-1)}{2} \right\rfloor, \pm \left\lfloor \frac{3k}{2} \right\rfloor \right\}. \quad (5)$$

- (ii) Si $n \equiv 2 \pmod{3}$ con $n = 3k + 2$, $k \in \mathbb{Z}^+$,

$$a_i^{(t+1)} = \sum_{j \in I} a_{i+j}^{(t)} \pmod{2}, \quad 0 \leq i \leq n-1, \quad (6)$$

donde I es el siguiente conjunto de índices:

$$I = \left\{ 0, \pm \left\lfloor \frac{4}{2} \right\rfloor, \pm \left\lfloor \frac{7}{2} \right\rfloor, \dots, \pm \left\lfloor \frac{3k-2}{2} \right\rfloor, \pm \left\lfloor \frac{3k+1}{2} \right\rfloor \right\} \quad (7)$$

Demostración. Sea $T(x) = x^{-1} + 1 + x$ el dipolinomio de transición del ACE lineal con $w = 150$. Teniendo en cuenta la Proposición 1, sólo necesitamos probar que $T(x) \cdot \tilde{T}(x) = 1 \pmod{(x^n - 1)}$, donde $\tilde{T}(x)$ es el dipolinomio de transición del autómata celular dado por (4) ó (6).

- (i) Si $n = 3k + 1$ con $k \in \mathbb{Z}^+$, entonces el dipolinomio de transición del ACE dado por (4) es $\tilde{T}(x) = \sum_{j \in I} x^j$, donde I es el conjunto de subíndices definido por (5). Un simple cálculo nos muestra que:

$$\tilde{T}(x) = D_{\lfloor \frac{n-1}{2} \rfloor}(x) + \sum_{i=1}^{\lfloor \frac{n-1}{6} \rfloor} \left(x^{-(3i-1)} + x^{3i-1} \right),$$

y teniendo en cuenta el apartado (i) del Lema 3, obtenemos:

$$\begin{aligned} T(x) \cdot \tilde{T}(x) &= T(x) D_{\lfloor \frac{n-1}{2} \rfloor}(x) + T(x) \sum_{i=1}^{\lfloor \frac{n-1}{6} \rfloor} \left(x^{-(3i-1)} + x^{3i-1} \right) \\ &= (x^{-1} + x) D_{\lfloor \frac{n-1}{2} \rfloor}(x) + D_{\lfloor \frac{n-1}{2} \rfloor}(x) \\ &\quad + \sum_{i=1}^{\lfloor \frac{n-1}{6} \rfloor} \left(x^{-3i} + x^{-3i+1} + x^{-3i+2} + x^{3i-2} + x^{3i-1} + x^{3i} \right) \\ &= x^{-\lfloor \frac{n-1}{2} \rfloor} + x^{-\lfloor \frac{n-1}{2} \rfloor - 1} + x^{\lfloor \frac{n-1}{2} \rfloor} + x^{\lfloor \frac{n-1}{2} \rfloor + 1} \\ &\quad + D_{\lfloor \frac{n-1}{2} \rfloor}(x) + D_{3\lfloor \frac{n-1}{6} \rfloor}(x) + 1. \end{aligned} \tag{8}$$

Ahora podemos distinguir dos subcasos: si k es par o impar.

En el primer caso, es decir cuando $k = (n-1)/3$ es par, entonces n es impar y las siguientes igualdades se satisfacen:

$$\begin{aligned} \left\lfloor \frac{n-1}{2} \right\rfloor &= \frac{n-1}{2}, \\ x^{\frac{n+1}{2}} &\equiv x^{-\frac{n-1}{2}} \pmod{x^n - 1}, \\ x^{-\frac{n+1}{2}} &\equiv x^{\frac{n-1}{2}} \pmod{x^n - 1}. \end{aligned}$$

Consecuentemente, teniendo en cuenta estas tres igualdades y (8), se verifica que

$$\begin{aligned} T(x) \cdot \tilde{T}(x) &= x^{-\frac{n+1}{2}} + x^{-\frac{n-1}{2}} + x^{\frac{n-1}{2}} + x^{\frac{n+1}{2}} + 1 \\ &\equiv 1 \pmod{x^n - 1}, \end{aligned}$$

con lo que se concluye.

Por otro lado, cuando $k = (n-1)/3$ es impar, entonces n es par y las siguientes igualdades se satisfacen:

$$\begin{aligned} \left\lfloor \frac{n-1}{2} \right\rfloor &= \frac{n-2}{2}, \\ x^{\frac{n}{2}} &\equiv x^{-\frac{n}{2}} \pmod{x^n - 1}. \end{aligned}$$

Consecuentemente, aplicando el apartado (ii) del Lema 3, la fórmula (8) es

$$\begin{aligned} T(x) \cdot \tilde{T}(x) &= x^{-\frac{n}{2}} + x^{-\frac{n-2}{2}} + x^{\frac{n-2}{2}} + x^{\frac{n}{2}} \\ &\quad + D_{\lfloor \frac{n-4}{2} \rfloor + 1}(x) + D_{\frac{n-4}{2}}(x) + 1 \\ &= x^{-\frac{n}{2}} + x^{\frac{n}{2}} + 1 \\ &\equiv 1 \pmod{(x^n - 1)}, \end{aligned}$$

y acabamos.

- (ii) Finalmente, si $n = 3k+2$ con $k \in \mathbb{Z}^+$, entonces el dipolinomio de transición asociado al ACE definido por (6) es $\tilde{T}(x) = \sum_{j \in I} x^j$, donde I es el conjunto de subíndices dado por (7). En consecuencia, es sencillo ver que:

$$\tilde{T}(x) = D_{\lfloor \frac{n-1}{2} \rfloor}(x) + \sum_{i=1}^{\lfloor \frac{n+2}{6} \rfloor} x^{-(3i-2)} + x^{3i-2},$$

luego aplicando el Lema 3, se tiene:

$$\begin{aligned} T(x) \cdot \tilde{T}(x) &= (x^{-1} + x) D_{\lfloor \frac{n-1}{2} \rfloor}(x) + D_{\lfloor \frac{n-1}{2} \rfloor}(x) \\ &\quad + \sum_{i=1}^{\lfloor \frac{n+2}{6} \rfloor} (x^{-3i+1} + x^{-3i+2} + x^{-3i+3} \\ &\quad + x^{3i-3} + x^{3i-2} + x^{3i-1}) \\ &= x^{-\lfloor \frac{n-1}{2} \rfloor - 1} + x^{-\lfloor \frac{n-1}{2} \rfloor} + x^{\lfloor \frac{n-1}{2} \rfloor + 1} + x^{\lfloor \frac{n-1}{2} \rfloor} \\ &\quad + D_{\lfloor \frac{n-1}{2} \rfloor}(x) + D_{3\lfloor \frac{n+1}{6} \rfloor - 1}(x) + 1. \end{aligned} \tag{9}$$

Como en el caso anterior, podemos distinguir dos subcasos dependiendo de si k es par o impar.

Así, cuando $k = (n-2)/3$ es par, entonces n es también par y las siguientes dos igualdades se satisfacen:

$$\begin{aligned} \left\lfloor \frac{n-1}{2} \right\rfloor &= 3 \left\lfloor \frac{n+1}{6} \right\rfloor = \frac{n-2}{2}, \\ x^{\frac{n}{2}} &\equiv x^{-\frac{n}{2}} \pmod{(x^n - 1)}. \end{aligned}$$

Consecuentemente, aplicando el apartado (ii) del Lema 3, la fórmula (9) da lugar a:

$$\begin{aligned} T(x) \cdot \tilde{T}(x) &= x^{-(\frac{n-2}{2}+1)} + x^{-\frac{n-2}{2}} + x^{\frac{n-2}{2}+1} + x^{\frac{n-2}{2}} \\ &\quad + D_{\frac{n-2}{2}}(x) + D_{\frac{n-2}{2}-1}(x) + 1 \\ &= x^{-\frac{n}{2}} + x^{\frac{n}{2}} + x^{-\frac{n-2}{2}} + x^{\frac{n-2}{2}} + x^{-\frac{n-2}{2}} + x^{\frac{n-2}{2}} + 1 \\ &\equiv 1 \pmod{(x^n - 1)}, \end{aligned}$$

con lo que se termina.

Por otra parte, si $k = (n - 2) / 3$ es impar, entonces n es también impar y se verifica que:

$$\begin{aligned} \left\lfloor \frac{n-1}{2} \right\rfloor &= 3 \left\lfloor \frac{n+1}{6} \right\rfloor - 1 = \frac{n-1}{2}, \\ x^{\frac{n+1}{2}} &\equiv x^{-\frac{n-1}{2}} \pmod{x^n - 1}, \\ x^{-\frac{n+1}{2}} &\equiv x^{\frac{n-1}{2}} \pmod{x^n - 1}. \end{aligned}$$

Así pues

$$\begin{aligned} T(x) \cdot \tilde{T}(x) &= x^{-\frac{n+1}{2}} + x^{\frac{n+1}{2}} + x^{-\frac{n-1}{2}} + x^{\frac{n-1}{2}} \\ &\quad + D_{\frac{n-1}{2}}(x) + D_{\frac{n-1}{2}}(x) + 1 \\ &\equiv 1 \pmod{x^n - 1}, \end{aligned}$$

con lo que se concluye.

□

Obsérvese que los autómatas celulares inversos obtenidos no son elementales pues las vecindades que poseen son simétricas de radio $r > 1$.

4 Ejemplos

Ejemplo 1 Si consideramos el ACE lineal con $w = 150$ y $n = 5 \equiv 2 \pmod{3}$, entonces la función de transición local del AC inverso es:

$$a_i^{(t+1)} = a_{i-2}^{(t)} + a_i^{(t)} + a_{i+2}^{(t)} \pmod{2}, \quad 0 \leq i \leq 5.$$

Por lo tanto su matriz de transición local viene dada por:

$$M_5^{-1} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Ejemplo 2 Consideremos el ACE lineal con $w = 150$ y $n = 10 \equiv 1 \pmod{3}$. Entonces, la función de transición local del AC inverso viene dada por la siguiente expresión booleana:

$$a_i^{(t)} = a_{i-4}^{(t)} + a_{i-3}^{(t)} + a_{i-1}^{(t)} + a_i^{(t)} + a_{i+1}^{(t)} + a_{i+3}^{(t)} + a_{i+4}^{(t)} \pmod{2}, \quad 0 \leq i \leq 10.$$

Consecuentemente su matriz de transición es:

$$M_{10}^{-1} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

5 Conclusiones

En este trabajo se ha resuelto completamente el problema de la reversibilidad para el autómatas celular elemental lineal con número de Wolfram 150. Se ha demostrado que dicha reversibilidad depende del número de células, n , que formen el ACE; concretamente, es reversible cuando $n \not\equiv 0 \pmod{3}$. Además para cada n satisfaciendo dicha condición se construye de manera explícita la función de transición local del autómatas celular inverso.

Referencias

- [1] S. Amoroso y Y.N. Patt. Decision procedures for surjectivity and injectivity of parallel maps for tessellations structures. *J. Comput. Syst.*, 6: 448-464, 1972.
- [2] P. Chaudhuri, D. Chowdhury, S. Nandi y S. Chattopadhyay. *Additive cellular automata. Theory and Applications. Volume 1* IEEE Computer Society Press, Los Alamitos, 1997.
- [3] W. Chin, B. Cortzen y J. Goldman. Linear cellular automata with boundary conditions. *Linear Algebra Appl.* 322: 193–206, 2001.
- [4] R. Díaz, A. Hernández, L. Hernández, S. Hoya, A. Martín, G. Rodríguez e I. Visus. Wolfram cellular automata and their cryptographic use as pseudorandom bit generators. *Internat. J. Pure Appl. Math.*, 4: 87–103, 2003.
- [5] C. Fraile, L. Hernández, S. Hoya, A. Martín y G. Rodríguez. The use of linear hybrid cellular automata as pseudorandom bit generators in cryptography. *Neural Parallel Sci. Comput.* , 12: 175–192, 2004.
- [6] J. Kari. *Reversibility and surjectivity of cellular automata*. Licenciante's Thesis, University of Turku, 1989.

- [7] J. Kari, Reversibility of 2D cellular automata is undecidable. *Physica D*, 45: 379-385, 1990.
- [8] O. Martin, A. M. Odlyzko y S. Wolfram. Algebraic properties of cellular automata. *Comm. Math. Phys.*, 93: 219-258, 1984.
- [9] L. Hernández, A. Martín y G. Rodríguez. Aplicaciones de los autómatas celulares a la generación de bits. *Bol. Soc. Esp. Mat. Apl.*, 21: 65-87, 2002.
- [10] G.Ch. Sirakoulis, I. Karafyllidis, Ch. Mizas, V. Mardiris, A. Thanailakis y Ph. Tsalides, A cellular automaton model for the study of DNA sequence evolution. *Comput. Biol. Med.*, 33: 439-453, 2003.
- [11] T. Toffoli y N. Margolus, Invertible cellular automata: A review. *Physica D*, 45:229-253, 1990.

El secreto de Google™ y el Álgebra lineal

P. FERNÁNDEZ

Departamento de Matemáticas, Universidad Autónoma de Madrid

pablo.fernandez@uam.es

Resumen

En este artículo analizaremos algunos de los ingredientes matemáticos que fundamentan el algoritmo (PageRank) con el que Google ordena los resultados de las búsquedas: un sabroso cóctel de Álgebra lineal, Teoría de Grafos y Probabilidad que nos facilita la vida.

Clasificación por materias AMS: 05C50, 15A48, 60J20.

Palabras clave: Google, algoritmos de ordenación, matrices positivas, teoría de Perron-Frobenius.

1 Introducción

Los editoriales de los principales periódicos del mundo se ocupan estos días de la noticia de la salida a Bolsa de Google, no sólo por el volumen de negocio que supone¹, sino por lo que de simbólico tiene el ser la primera operación de este tipo en una empresa de tal envergadura desde la “exuberancia irracional” de las llamadas “*puntocom*” de los años 90.

Pero hay algo más que explica este destacado interés, y tiene que ver con las características propias de la empresa. Ya han pasado varios lustros desde que se produjo una revolución en el campo de la tecnología y de la información — y quizás (¿o sin duda?), una revolución cultural, sociológica, etc.—, como

Fecha de recepción: 09/11/04

¹Se pondrán a la venta acciones por valor de 2700 millones de dólares. Quizás por el recuerdo de los excesos y escándalos que rodearon la burbuja financiera de las empresas tecnológicas en años pasados, la oferta se articulará en forma de subasta *on line*, de manera que muchos inversores tengan similares oportunidades de adquirir acciones. Se trata así de evitar grandes movimientos especulativos... pero nunca se sabe.

fue la generalización del acceso, uso y participación en la red de redes, Internet. La aparición del buscador Google ha supuesto una revolución equiparable a la anterior, al convertirse en una herramienta capaz de poner orden en todo ese (antes inabarcable) universo de información.

El diseño de un buscador en la red es un problema de *ingeniería matemática*. Nótese el adjetivo. Se necesita primero un buen conocimiento del contexto, que permita traducirlo a modelos, a Matemáticas. Pero tras ese proceso de abstracción, de matematización, y una vez extraídas las conclusiones pertinentes, se requiere también una cuidadosa y eficiente implementación, un detallado diseño de las cuestiones computacionales inherentes al problema.

2 El buscador Google

El origen del buscador Google es ya bien conocido. Fue diseñado en 1998 por Sergei Brin y Lawrence Page, dos estudiantes de doctorado en Informática de la Universidad de Stanford: Brin se había graduado en Matemáticas, y Page en Informática. Los vemos en las fotografías de la derecha². Dos jóvenes que hoy, en la frontera de la treintena, se han convertido en multimillonarios.



Brin (a la izquierda) y Page

El curioso nombre del buscador es una variación sobre el término *googol*, que alguien³ inventó para referirse al apabullante número 10^{100} . Uno de esos números que los matemáticos manejamos con comodidad pero que, quizás, sea mayor que el número de partículas del Universo.

Aunque sin llegar a esos extremos, las escalas de la cuestión que nos interesa son también gigantescas. En 1997, cuando Brin y Page empezaban a trabajar en el diseño de Google, había censadas en torno a los 100 millones de páginas *web*. Altavista, el buscador más popular por entonces, atendía 20 millones de consultas diarias. Hoy, esas cifras se han multiplicado: el propio buscador Google atiende 200 millones de consultas diarias e indexa varios miles de millones de páginas *web*.

Así que el diseño de un buscador ha de resolver con eficacia ciertas cuestiones computacionales, como la manera en que se almacena toda esa información, cómo se actualiza, cómo se pueden gestionar las peticiones, cómo buscar en las gigantescas bases de datos, etc., problemas en los que, sin duda, las Matemáticas tienen mucho que decir.

²Podemos jurar que no tenemos nada que ver con esa suerte de bolitas navideñas que adornan las fotografías. Están extraídas, tal cual, de la página *web* de Google.

³Se dice que un sobrino del matemático Edward Kasner. Kasner se animó y también definió el googolplex, que es 10^{googol} . ¡Ooooh!

Pero sobre ninguna de estas cuestiones, con ser interesantes, reflexionaremos aquí. El asunto que nos interesa se puede formular de manera sencilla: supongamos que, a resultas de una consulta, tras el proceso de búsqueda en las bases de datos, hemos determinado que, digamos, unas cien páginas contienen información que, de una manera u otra, puede resultar relevante para el usuario.

Ahora,

¿en qué orden mostramos esos resultados?

El objetivo, como de manera explícita se plantearon⁴ Brin y Page (véase [6]), es que, en un número suficientemente grande de los casos, al menos una de, digamos, las diez primeras páginas que se muestren contenga información útil para el que realiza la consulta.

Pedimos ahora al lector (que en un alto porcentaje será sin duda *googleadicto*) que decida, a partir de su experiencia personal, si Google cumple este objetivo. Estamos seguros de que la respuesta general es que sí, y de manera asombrosa. Parece magia⁵. . . pero son Matemáticas.

Matemáticas que, además, no requieren conocimientos que vayan más allá de cursos básicos de licenciatura, como veremos a lo largo de estas páginas.

Para la tarea que queremos abordar, necesitamos un **criterio** de ordenación. Obsérvese (y aquí ya nos hemos puesto en modo matemático) que si etiquetamos con los símbolos P_1, \dots, P_n cada una de las páginas de la red, todo lo que nos hace falta es asignar a cada P_j un número x_j , que llamaremos su **importancia**. Estos números podrían ser, por ejemplo, números entre 0 y 1.

Supongamos entonces que, tras un censo de los sitios de la red, hemos conseguido construir la lista de páginas *web*, y que le hemos asignado a cada una de ellas, de la manera que sea, una importancia. Esta ordenación queda a nuestra disposición para ser utilizada cada vez que contestemos a una determinada consulta: las páginas seleccionadas se mostrarán en el orden que indique dicha lista.

Queda, por supuesto, explicar cómo se construye esa lista. Vamos con ello.

⁴No era el único objetivo que se propusieron. También pretendían que el buscador fuera “resistente” a intentos de manipulación, a estrategias que, por ejemplo por intereses comerciales, pretendieran situar ciertas páginas en los primeros puestos de la lista. Curiosamente, hoy en día se ha hecho popular un “deporte”, el *Google bombing*, que consiste precisamente en eso, en conseguir colocar una página en esas primeras posiciones, aunque generalmente con fines más lúdicos. Pruebe por ejemplo el lector a buscar en Google los términos “miserable fracasado”. El resultado le sorprenderá (y quizás regocijará).

⁵Por no hablar de la capacidad de la que hace gala el buscador para “corregir” los términos de la búsqueda y sugerir, ¡sí!, la palabra que uno en realidad estaba pensando teclear. Lo que le lleva a uno en ocasiones a imaginar intervenciones sobrenaturales. . . En fin, dejémoslo estar.

NOTA. Para completar la descripción, y aunque no entraremos en los detalles, deberíamos decir que hay un par de ingredientes que, en cada consulta particular, se combinan en el buscador Google con el criterio general que veremos aquí:

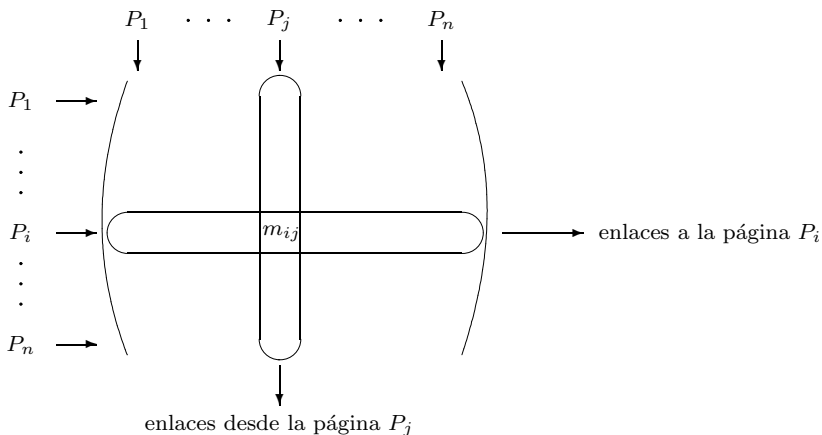
- por un lado, y como resulta razonable, Google no “puntuá” por igual el que un cierto término aparezca en el título de la página, en negrita, en un tipo de letra pequeña, etc.
- Para búsquedas combinadas, tampoco es lo mismo que los términos buscados aparezcan, dentro del documento, “cerca” o “lejos” unos de otros.

3 El modelo

El primer paso consiste en describir, en forma manejable, los elementos relevantes para la cuestión que interesa, la asignación de importancias. Digamos que hemos recogido toda la información sobre la red (sitios, contenidos, enlaces de unas páginas a otras, etc.). De toda esta información, y en este primer paso de la modelización, vamos a rescatar únicamente la referida a cada una de las páginas, a la que asignamos etiquetas P_1, \dots, P_n , y los enlaces entre ellas.

En estos términos, la red puede ser descrita mediante un **grafo** (dirigido) G . Cada página P_j de la red es un vértice del grafo, y hay una arista (dirigida) entre los vértices P_i y P_j si desde la página P_i hay un enlace a la página P_j . Es un grafo gigantesco, abrumador, sobre cuya estructura real reflexionaremos más adelante (véase la Sección 8).

Cuando tratamos con grafos, nos gusta siempre recurrir a los dibujos en el papel, en los que los vértices son puntos del plano, mientras que las aristas son flechas que unen esos puntos. Pero, a los efectos que aquí perseguimos, conviene considerar una interpretación alternativa, matricial. Vamos, pues, a formar una matriz \mathbf{M} , de dimensiones $n \times n$, cuyas filas y columnas van etiquetados con los símbolos P_1, \dots, P_n , y cuyas entradas son ceros y unos. La entrada m_{ij} de la matriz será un *uno* si es que hay un enlace de la página P_j a la página P_i ; y un *cero* en caso contrario:



La matriz \mathbf{M} es, salvo una trasposición, la matriz de adyacencia (o de vecindades) del grafo. Nótese que no tiene por qué ser simétrica, pues nuestro grafo es dirigido. Obsérvese también que, tal como se señala en el dibujo, la suma de las entradas correspondientes a la columna de P_j es el número de enlaces que salen de la página P_j , mientras que la suma de los registros de una fila coincide con el número de enlaces entrantes.

Vamos a postular que la importancia de una cierta página de la red P_j “tiene que ver” con las páginas desde las que hay enlaces a ella. Esto es muy razonable: si muchas páginas enlazan con P_j , será porque la información que ésta contiene ha sido considerada por muchos participantes de la red como “recomendable”.

Ese “tiene que ver” queda por ahora algo difuso. Un **primer intento**, quizás algo ingenuo, consiste en suponer que la importancia x_j de cada P_j es *proporcional al número de páginas desde las que hay enlaces a P_j* . Observemos que, una vez que disponemos de la matriz \mathbf{M} , el cálculo de cada uno de los x_j es sencillo, pues basta sumar las entradas correspondientes a la fila de P_j .

Pero este modelo no recoge adecuadamente una situación que nos gustaría tener en cuenta, y es aquélla en la que una cierta página es citada desde unas pocas páginas, pero muy relevantes. Digamos, por ejemplo, que desde www.microsoft.com, o desde www.amazon.com, etc. El procedimiento anterior le asignaría baja importancia, algo que no parece razonable. Así que necesitamos afinar nuestro modelo de manera que asignemos alta importancia

- tanto a páginas muy citadas;
- como a páginas poco citadas, pero desde sitios “importantes”.

El **segundo intento**, ya en esta línea, consiste en decidir que la importancia x_j de cada página P_j es *proporcional a la suma de las importancias de las páginas que enlazan con P_j* . Nótese que hemos cambiado, con respecto a nuestro primer intento, “número de páginas que enlazan con P_j ” por “suma de las importancias de las páginas que enlazan con P_j ”. Esta ligera variación cambia por completo, como veremos, las características del problema.

Supongamos, por ejemplo, que la página P_1 es citada desde las páginas P_2 , P_{25} y P_{256} , que P_2 sólo se cita desde P_1 y P_{256} , etc., mientras que, digamos, hay enlaces a la última página, P_n , desde P_1, P_2, P_3, P_{25} y P_{n-1} . En nuestra asignación anterior, x_1 debería ser proporcional a 3, x_2 lo sería a 2, etc., mientras que x_n habría de ser proporcional a 5.

Pero ahora nuestra asignación x_1, \dots, x_n debe cumplir que

$$\begin{aligned} x_1 &= K(x_2 + x_{25} + x_{256}), \\ x_2 &= K(x_1 + x_{256}), \\ &\vdots \\ x_n &= K(x_1 + x_2 + x_3 + x_{25} + x_{n-1}), \end{aligned}$$

donde K es una cierta constante de proporcionalidad. Nos encontramos así con un enorme sistema de ecuaciones, cuyas soluciones son las posibles asignaciones x_1, \dots, x_n .

Escribamos el sistema anterior en términos más manejables, matriciales:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = K \begin{pmatrix} P_1 & P_2 & & P_{25} & & P_{256} & & P_{n-1} \\ \downarrow & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Hagamos ahora un alarde simbólico al que tan aficionados somos los matemáticos: llamemos \mathbf{x} al vector de importancias. La matriz de dimensiones $n \times n$ del sistema es, justamente, la \mathbf{M} asociada al grafo. Así que podremos escribir que la asignación de importancias que perseguimos es una solución de

$$\mathbf{M}\mathbf{x} = \lambda\mathbf{x}.$$

Ya hemos escrito la constante de proporcionalidad con el símbolo λ . Y es que, como cualquier estudiante de un curso de Álgebra lineal reconoce al instante, la cuestión se ha transformado en un problema de **autovalores** y **autovectores**: nuestro anhelado vector de importancias \mathbf{x} no es sino un autovector de la matriz \mathbf{M} que, recordemos, recogía la información sobre la estructura de la red (vértices y relaciones de adyacencia).

Salvo por la cuestión, casi sentimental, de comprobar que uno de los ingredientes básicos de los cursos de Álgebra lineal permite describir una cuestión tan “aplicada”, es posible que esto todavía no haya despertado el entusiasmo del lector. Sí, de acuerdo, un autovector, ¿pero cuál? Hay tantos... Y además, ¿cómo podemos calcularlo?, la matriz es inimaginablemente grande: recordemos, varios miles de millones de filas (o columnas).

Paciencia.

Por ahora, parece razonable exigir que las entradas del vector (las importancias de los páginas de la red) sean todas no negativas (o, al menos, del mismo signo). Esto lo escribiremos, y perdónenos el lector el abuso de notación, como $\mathbf{x} \geq 0$.

Pero además, si pretendemos que el método sea útil, necesitaríamos que ese hipotético autovector de entradas no negativas fuera **único** (si hubiera más de uno, ¿con cuál nos quedaríamos?).

4 El surfista aleatorio

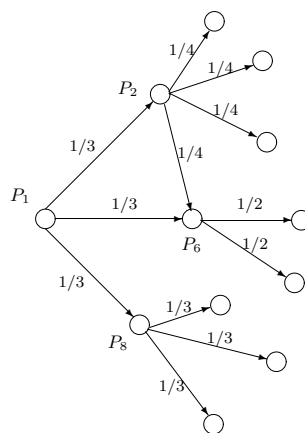
A modo de interludio, vamos a enfocar el problema desde un nuevo punto de vista. Imaginemos que un surfista se dedica a navegar por la red. En un cierto instante de tiempo está en una página, por ejemplo en P_1 . En el instante siguiente, algo aburrido de los contenidos de P_1 , va a saltar a una de las páginas a las que enlaza P_1 (digamos que hay N_1 de ellas). Pero, ¿exactamente a cuál?

Nuestro intrépido navegante es un surfista (además de rubio y bronceado) *aleatorio*. Así que, para decidir a cuál de ellas va, efectúa un sorteo. Y lo

hace de la manera más simple posible: utilizando un dado regular (y virtual, suponemos), que tenga tantas caras como enlaces salgan de la página P_1 . En términos más técnicos, la elección de la página de destino sigue una distribución de probabilidad *uniforme* (discreta) en $[1, N_1]$.

Nuestro modelo no es *determinista*, sino *probabilístico*: no sabemos dónde estará un instante de tiempo después, pero sí *con qué probabilidad* estará en cada uno de los posibles destinos. Y además es un modelo *dinámico*, porque el mismo argumento se podría aplicar al segundo movimiento, y al tercero, etc. Nuestro surfista sigue lo que se denomina un **paseo aleatorio** por el grafo.

En el dibujo de la derecha mostramos una posible situación: de P_1 parten tres aristas, a los vértices P_2 , P_6 y P_8 . Así que el navegante sortea entre estas tres páginas, con probabilidad $1/3$ para cada una de ellas. Si, por ejemplo, el resultado del sorteo hubiera sido la página P_2 , entonces volvería a sortear, pero esta vez asignando probabilidad $1/4$ a cada una de los cuatro posibles destinos desde P_2 .

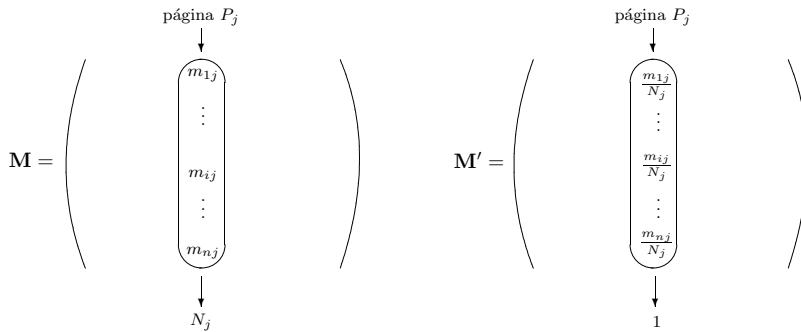


La formulación es muy sugerente, pero no está claro cómo podemos formalizarla: cómo podemos calcular, por ejemplo, la probabilidad de que, partiendo de la página P_1 , el navegante esté en la página P_{17} tras cinco instantes de tiempo. Más aún, todavía ni sospechamos qué puede tener que ver con nuestro problema de asignación de importancias.

Vamos poco a poco. Recordemos que \mathbf{M} es la matriz del grafo de la red, y que sus entradas m_{ij} son ceros o unos. Llamemos N_j al número de enlaces desde la página P_j (esto es, la suma de las entradas de cada columna de \mathbf{M}). Vamos ahora a construir una nueva matriz \mathbf{M}' a partir de la \mathbf{M} original sustituyendo cada entrada m_{ij} por

$$m'_{ij} = \frac{m_{ij}}{N_j}$$

Los registros de la nueva matriz \mathbf{M}' son números no negativos (entre 0 y 1) tales que la suma de los registros de *cada columna* vale 1. En la siguiente figura se muestra, para una cierta columna, el resultado de la transformación:



La matriz M' así construida es lo que se llama una **matriz estocástica** (o de Markov).

Supongamos que el surfista se encuentra, en el primer instante de tiempo, en la página (o vértice) P_k . Para poner en términos probabilísticos esta situación de por sí tan determinística, digamos que está en la página P_k con una probabilidad del 100 %. Vamos a representar esta situación inicial con el vector cuyas entradas son todas cero, excepto la de la posición k , que lleva un uno. El navegante sortea entre los N_k destinos, asignando probabilidad $1/N_k$ a cada uno de ellos.

Pero si multiplicamos la matriz M' por este vector inicial obtenemos

$$\begin{pmatrix} \dots & \dots & m'_{1k} & \dots \\ \vdots & \ddots & \vdots & \vdots \\ \dots & \dots & m'_{kk} & \dots \\ \vdots & \vdots & \vdots & \ddots \\ \dots & \dots & m'_{nk} & \dots \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} m'_{1k} \\ \vdots \\ m'_{kk} \\ \vdots \\ m'_{nk} \end{pmatrix}.$$

El resultado sigue siendo un vector cuyas entradas (números entre 0 y 1, pues los m'_{jk} son, o bien 0, o bien $1/N_k$) suman 1, porque hay justamente N_k entradas no nulas. Pero más aún: el vector que obtenemos describe exactamente con qué probabilidad estará el surfista, tras una unidad de tiempo, en cada una de las páginas de la red, si es que ha partido de P_k .

Lo que hace especialmente manejable este modelo es que, para conocer con qué probabilidad estará en cada una de las páginas de la red tras *dos* instantes de tiempo, basta repetir la operación. Esto es, multiplicar $(M')^2$ por el vector inicial. Y el análogo para el transcurrir de tres instantes de tiempo, cuatro, etc.

Siguiendo la nomenclatura habitual, vamos a considerar una serie de **estados**, que en este caso son los vértices del grafo G . La matriz M' recibe el (apropiado) nombre de **matriz de transición** del sistema: cada entrada m'_{ij} es la probabilidad de pasar del estado (vértice) P_j al estado (vértice) P_i . Y los registros de las sucesivas potencias de la matriz, las probabilidades de pasar de cada P_i a cada P_j tras varios instantes de tiempo.

Esto resuelve la primera de las cuestiones: cómo calcular las probabilidades de transición entre vértices para el transcurso de varias unidades de tiempo. Pero todavía no hemos descubierto la relación con el problema de ordenación de Google. Esto lo dejamos para más adelante (véase la subsección 9.1). Aunque

avanzamos, para aquel lector especialmente impaciente y versado en estas cuestiones, que el vector de entradas no negativas que llevamos persiguiendo ya un buen rato resulta ser, justamente, el estado estacionario de la cadena de Markov que acabamos de construir.

NOTA. Podría ocurrir que alguna página no tuviera enlaces salientes (su columna tuviera sólo ceros). No sería una matriz estocástica. La solución de Google: la sustituimos por una columna con valores $1/n$. Así que si el surfista llega a una página de la que (antes) “no podía salir”, ahora sortea (uniformemente) entre *todas* las de la red. Más sobre esto, en la Sección 8.

5 La clasificación para las eliminatorias por el título

Con tanta reinterpretación y reformulación, nos estamos olvidando de la cuestión que nos interesa: el sistema de ordenación de Google. Disciplinémonos y volvamos a ella. Empecemos precisando que las ideas que utiliza Google para ordenar los resultados provienen de unos algoritmos desarrollados por Kendall y Wei en los años 50 (uups, ¡del siglo pasado!, que siempre nos olvidamos de recalcarlo), véanse [11] y [16].

Vamos a plantear una cuestión muy similar a la que nos ha ocupado hasta ahora, en un contexto algo más lúdico, que ilustraremos con un ejemplo sencillo que podremos resolver sin ayuda de grandes herramientas computacionales.

La estructura de los campeonatos deportivos, por ejemplo de baloncesto, en Estados Unidos es algo diferente a la que suele seguirse en los europeos. En las ligas europeas, cada equipo juega un cierto número de partidos (dos, generalmente) contra cada uno de los demás equipos de la competición. Al final de lo que se conoce con el (curioso) nombre de temporada regular, se disputan unas eliminatorias por el título, en las que se intervienen los que han conseguido mejores clasificaciones (por ejemplo, los ocho primeros).

Pero en Estados Unidos, debido a las grandes distancias que separan las distintas ciudades, y para evitar un número excesivo de viajes, la estructura no es tan homogénea.

Lo vamos a ilustrar con la liga profesional de baloncesto, aunque la cuestión que vamos a explicar podría aplicarse también a la Liga universitaria. En la NBA, los 29 equipos están divididos en dos conferencias, cada una de las cuales está a su vez formada por dos divisiones:

- Conferencia Este:
 - División Atlántico;
 - División Central;
- Conferencia Oeste:
 - División Medio Oeste⁶;

⁶Aquí andan los Memphis Grizzlies y los Utah Jazz, donde juegan nuestros Pau Gasol y Raúl López, respectivamente.

– División Pacífico,

siguiendo (más o menos) criterios geográficos. Todos los equipos juegan el mismo número de partidos, 82, pero no disputan *el mismo número* contra cada equipo; lo habitual es que se jueguen más partidos contra los de la propia conferencia.

Se plantea así la siguiente cuestión: acabada la temporada regular, ¿qué 16 equipos han de clasificarse para las eliminatorias? El sistema actual computa, simplemente, el número de victorias obtenidas por cada equipo, que determinan las posiciones finales. Pero cabría preguntarse⁷ si éste es un sistema “justo”. Al fin y al cabo, podría ocurrir que un equipo hubiera acumulado muchas victorias precisamente por estar encuadrado en una conferencia muy “débil”. ¿Qué debe valer más, el número de victorias, o la “calidad” de las victorias obtenidas? De nuevo, la dicotomía ante la que nos encontrábamos en el diseño de Google.

Digamos que en la competición hay n equipos, E_1, \dots, E_n . Los resultados de la temporada regular los registramos en una matriz \mathbf{A} cuyas entradas a_{ij} son las victorias obtenidas por cada equipo:

$$a_{ij} = \frac{\# \text{victorias de } E_i \text{ sobre } E_j}{\# \text{partidos de } E_i}.$$

Como medida de normalización, dividimos por ejemplo por el número total de partidos que juega cada equipo.

Nuestro objetivo es asignar a cada equipo E_j un número x_j , que por analogía denominaremos importancia, de manera que la lista x_1, \dots, x_n determine la clasificación final.

El primer modelo, en el que x_j es proporcional al número total de victorias conseguido por el equipo E_j , no recoge adecuadamente nuestro interés por valorar la calidad de las victorias.

En el segundo modelo, decidimos que la importancia x_j sea proporcional al número de victorias obtenidas, pero *ponderadas* con la importancia que asignemos a los demás equipos:

$$x_j = K \sum_{k=1}^n a_{jk} x_k.$$

Lo que nos conduce, otra vez, a

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

El interés, de nuevo, radica en encontrar un autovector de \mathbf{A} con entradas no negativas (y único, a poder ser). Veamos un ejemplo concreto.

5.1 Para cálculos serios: el ordenador

Tenemos seis equipos, E_1, \dots, E_6 : los tres primeros forman una conferencia, los siguientes la otra. Cada equipo juega 21 partidos en total: 6 contra los de su propia conferencia, 3 contra cada uno de los de la otra. La información sobre las victorias conseguidas está contenida en la siguiente tabla:

⁷Véanse las reflexiones al respecto en [10].

	E_1	E_2	E_3	E_4	E_5	E_6	
E_1	–	3/21	0/21	0/21	1/21	2/21	→ 6/21
E_2	3/21	–	2/21	2/21	2/21	1/21	→ 10/21
E_3	6/21	4/21	–	2/21	1/21	1/21	→ 14/21
E_4	3/21	1/21	1/21	–	2/21	2/21	→ 9/21
E_5	2/21	1/21	2/21	4/21	–	2/21	→ 11/21
E_6	1/21	2/21	2/21	4/21	4/21	–	→ 13/21

A la derecha de la tabla hemos sumado el número de victorias que ha conseguido cada escuadra. Ese balance de victorias y derrotas sugiere que la ordenación adecuada es

$$E_3 \rightarrow E_6 \rightarrow E_5 \rightarrow E_2 \rightarrow E_4 \rightarrow E_1.$$

Pero observemos que, por ejemplo, el líder de la competición, E_3 , ha acumulado muchas victorias contra E_1 , que es el peor equipo.

¿Qué ocurre con el segundo enfoque? Tenemos que calcular, y en serio. Incluso en un ejemplo tan sencillo como éste, necesitamos la ayuda del ordenador. Así que recurrimos a un paquete matemático de cálculo, como por ejemplo⁸ **Maple**, para realizar los cálculos. Al usuario experto, los comandos que siguen le resultarán familiares. Al novicio tampoco le costará mucho trabajo entenderlos. Primero, cargamos un paquete necesario:

```
> restart:with(linalg):
```

```
Warning, the protected names norm and trace have been redefined and unprotected
```

Tras esta concisa (y quizás profunda) respuesta del programa, lo ponemos a trabajar en serio. ¡**Maple** en acción!:

```
> A:=matrix(6,6,[0,3/21,0,0,1/21,2/21,3/21,0,2/21,2/21,2/21,1/21,6/21,
4/21,0,2/21,1/21,1/21,3/21,1/21,1/21,0,2/21,2/21,2/21,1/21,2/21,4/21,0,
2/21,1/21,2/21,2/21,4/21,4/21,0]);
```

$$A := \begin{bmatrix} 0 & \frac{1}{7} & 0 & 0 & \frac{1}{21} & \frac{2}{21} \\ \frac{1}{7} & 0 & \frac{2}{21} & \frac{2}{21} & \frac{2}{21} & \frac{1}{21} \\ \frac{2}{7} & \frac{4}{21} & 0 & \frac{2}{21} & \frac{1}{21} & \frac{1}{21} \\ \frac{1}{7} & \frac{1}{21} & \frac{1}{21} & 0 & \frac{2}{21} & \frac{2}{21} \\ \frac{2}{21} & \frac{1}{21} & \frac{2}{21} & \frac{4}{21} & 0 & \frac{2}{21} \\ \frac{1}{21} & \frac{2}{21} & \frac{2}{21} & \frac{4}{21} & \frac{4}{21} & 0 \end{bmatrix}$$

```
> autovects:=[evalf(eigenvalues(A))];
```

⁸Sin ningún afán propagandístico. Matlab, Mathematica, Derive o incluso Excel también valdrían.

```

autovecs := [[0.012, 1., {[-0.062, -0.916, -2.131, 0.873, 0.731, 1.]},
              [0.475, 1., {[0.509, 0.746, 0.928, 0.690, 0.840, 1.]},
              [-0.111 + 0.117 I, 1., {[-0.151 - 0.901 I, 0.123 + 0.451 I, -0.727 + 0.728 I,
              -0.435 - 0.128 I, 0.192 + 0.377 I, 1.]},
              [-0.126, 1., {[0.008, -0.685, 1.434, -1.071, 0.032, 1.]},
              [-0.139, 1., {[1.785, -3.880, 3.478, -5.392, 4.415, 1.]},
              [-0.111 - 0.117 I, 1., {[-0.151 + 0.901 I, 0.123 - 0.451 I, -0.727 - 0.728 I,
              -0.435 + 0.128 I, 0.192 - 0.377 I, 1.]}}]]

```

Una vez acostumbrados a la manera en que **Maple** muestra los resultados (en cada línea, primero el autovalor, luego la multiplicidad, y luego los autovectores asociados), comprobamos que la matriz **A** tiene seis autovalores distintos, dos complejos (conjugados) y cuatro reales. Si calculamos sus módulos,

```

> autovals:=[evalf(eigenvalues(A))]: modulos:=[seq(abs(autovals[j]),j=1..6)];
modulos := [0.012, 0.475, 0.161, 0.126, 0.139, 0.161]

```

observamos que el mayor autovalor (en módulo) es $\lambda = 0.475$, cuyo autovector asociado,

$$\mathbf{x} = (0.509, 0.746, 0.928, 0.690, 0.840, 1),$$

es el **único** cuyas entradas son todas números reales y no negativos.

Ya tenemos la respuesta que buscábamos: el orden que sugiere este cálculo (véanse las entradas del autovector **v**) es

$$E_6 \rightarrow E_3 \rightarrow E_5 \rightarrow E_2 \rightarrow E_4 \rightarrow E_1,$$

que difiere del anterior en los dos primeros (ahora E_6 es el mejor equipo).

Recapitemos. En esta matriz particular (que bien pudiera servir de versión a pequeña escala de la de Internet), cuyas entradas son todas no negativas, se ha dado la mejor de las situaciones posibles: hay un único autovector cuyas entradas son todas no negativas. Este autovector, que además está asociado al autovalor de módulo máximo, nos sirve como solución al problema de ordenación que nos habíamos planteado.

¿Ha sido casualidad? ¿Ha sido simplemente un ardid de quien esto escribe, una artera elección de matriz que busca convencer a los incautos de que las cosas funcionan realmente?

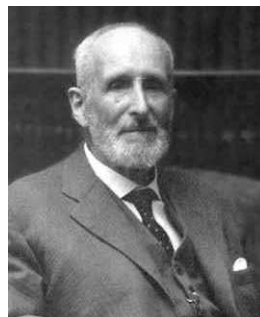
El lector, que seguramente de incauto no tiene nada, y sí mucho de curioso, sentirá la necesidad de dar contundente respuesta a esta cuestión. Y sabrá, sin duda, que ha llegado el momento de dar paso a un nuevo actor en esta obra.

6 Las Matemáticas entran en escena

Es la hora de las Matemáticas, esa Ciencia que se ocupa de realidades abstractas, virtuales, para, a través de ellas, entender realidades concretas. Pongámonos, pues, matemáticos, y destilemos la esencia común a todas las cuestiones de las que nos hemos ocupado hasta ahora.

Si lo hacemos, descubriremos que la única propiedad que comparten todas las matrices que nos conciernen (sean estocásticas o no) es que sus entradas son **no negativas**. No parece mucha información. Las matrices no tienen por qué ser simétricas, ni definidas positivas, ni...

Sin embargo, como mostraría Perron⁹ a principios del siglo XX, es suficiente para empezar a extraer buenos réditos:



Oskar Perron

Teorema 1 (Perron, 1907) *Sea \mathbf{A} una matriz (cuadrada) con entradas positivas, $\mathbf{A} > 0$. Entonces,*

- existe un autovalor (simple) $\lambda > 0$ tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, donde el autovector correspondiente es $\mathbf{v} > 0$;*
- este autovalor es mayor, en módulo, que todos los demás autovalores;*
- cualquier otro autovector positivo de \mathbf{A} es un múltiplo de \mathbf{v} .*

Paremos un instante y reflexionemos: *ahora* tenemos un *teorema*. Maravillémonos: si usted logra comprobar que las hipótesis se satisfacen (y aquí no admitiremos medias tintas, han de satisfacerse, sin excusas, sin excepciones), entonces la conclusión es inapelable.

No albergue dudas, sea cual sea la matriz, mientras que tenga entradas no negativas, existirá un autovalor simple tal que...y un único autovector con... ¡Qué cosas!



F.G. Frobenius

Tras este raptó casi místico, que el lector sabrá disculpar, volvamos a la realidad y analicemos el resultado de Perron. Apunta en la dirección que buscamos, pero no es suficiente, porque las matrices que manejamos pueden tener ceros en ciertas posiciones. Así que necesitamos algo más.

El siguiente acto de esta obra lo escribe Frobenius¹⁰, unos años después, cuando se ocupa

⁹El alemán Oskar Perron (1880-1975), todo un ejemplo de longevidad matemática, trabajó en campos tan diversos como el Análisis, las Ecuaciones Diferenciales, el Álgebra, la Geometría, la Teoría de Números, etc., en los que publicó algunos textos que se convertirían en clásicos.

¹⁰Ferdinand Georg Frobenius (1849-1917) es uno de los más destacados integrantes de la Escuela de Berlín, que lideró el mundo de las matemáticas durante el final del siglo XIX y principios del XX, junto con personajes de la talla de Kronecker, Kummer o Weierstrass (quien fue su director de tesis). Estricta escuela prusiana: rigor, tradición, un matemático "puro", sin concesiones a la Matemática aplicada. El caprichoso curso de la Historia haría, sin embargo, que muchas de sus ideas sobre representaciones de grupos finitos acabaran sirviendo como base de la Mecánica Cuántica. Es sobre todo conocido por sus aportaciones a la Teoría de Grupos, y sus trabajos sobre las matrices no negativas pertenecen a la última etapa de su vida.

del caso más general de las matrices no negativas. Frobenius observa que si únicamente tenemos que $\mathbf{A} \geq 0$ entonces, aunque sigue habiendo un autovalor $\lambda > 0$ *dominante* (de valor absoluto máximo) asociado a un autovector $\mathbf{v} \geq 0$, puede haber otros autovalores del mismo “tamaño”. Ahí va su resultado:

Teorema 2 (Frobenius, 1908-1912) *Sea \mathbf{A} una matriz (cuadrada) con entradas no negativas, $\mathbf{A} \geq 0$. Si \mathbf{A} es irreducible, entonces*

(a) *existe un autovalor (simple) $\lambda > 0$ tal que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, donde el autovector es $\mathbf{v} > 0$. Además, $\lambda \geq |\mu|$, para cualquier otro autovalor μ de \mathbf{A} .*

(b) *Cualquier autovector ≥ 0 es un múltiplo de \mathbf{v} .*

(c) *Si hay k autovalores de módulo máximo, entonces son las soluciones de la ecuación $x^k - \lambda^k = 0$.*

NOTA. ¿Qué quiere decir que una matriz \mathbf{A} de dimensiones $n \times n$ sea *irreducible*? Hay varias maneras de entenderlo.

1. No existe ninguna permutación (de filas y columnas) que transforma \mathbf{A} en una matriz del tipo

$$\left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{0} & \mathbf{A}_{22} \end{array} \right),$$

donde \mathbf{A}_{11} y \mathbf{A}_{22} son matrices cuadradas.

2. La matriz $(\mathbf{I} + \mathbf{A})^{n-1}$, donde \mathbf{I} es la identidad $n \times n$, tiene todas sus entradas positivas.
3. Si \mathbf{A} es la matriz de adyacencia de un grafo, entonces el grafo está *fuertemente* conectado (véase la Sección 8).

Observemos primero que el teorema de Frobenius generaliza realmente el de Perron, pues si $\mathbf{A} > 0$, entonces es $\mathbf{A} \geq 0$ e irreducible.

Segundo, si la matriz \mathbf{A} es irreducible (y sobre si éste es el caso o no de la matriz de la red reflexionaremos más adelante), entonces la cuestión queda completamente resuelta: existe un único autovector con entradas no negativas, que además está asociado al autovalor positivo de módulo máximo (en un momento entenderemos las ventajas que supone esto último).

Estos resultados, a los que a partir de ahora nos referiremos conjuntamente como teorema de Perron-Frobenius, tienen una amplia utilidad en diversos contextos (véase la Sección 9). Tanto es así que hay quien habla de la “teoría de Perron-Frobenius”, del que este teorema es el resultado fundamental.

Su demostración es algo complicada, y aquí nos limitaremos a pergeñar un argumento que contiene los ingredientes fundamentales de la prueba.

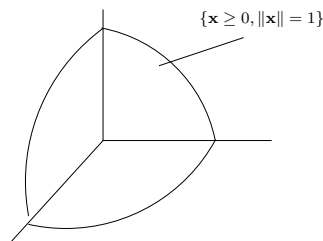
Una “demostración” del teorema de Perron-Frobenius (ilustrada en el caso 3×3)

Como \mathbf{A} tiene entradas no negativas, si partimos de un vector $\mathbf{x} \geq 0$, entonces el vector $\mathbf{A}\mathbf{x}$ sigue teniendo entradas no negativas.

En términos geométricos, la matriz \mathbf{A} aplica el octante positivo en sí mismo. Consideramos ahora la aplicación α dada por

$$\alpha(\mathbf{x}) = \frac{\mathbf{A}\mathbf{x}}{\|\mathbf{A}\mathbf{x}\|}.$$

Obsérvese que $\alpha(\mathbf{x})$ es siempre un vector de longitud 1. Por la observación anterior, α envía el conjunto $\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \geq 0, \|\mathbf{x}\| = 1\}$, esto es, la porción de la esfera unidad que dibujamos a la derecha, en sí mismo.



Ahora, aplicando el teorema del punto fijo de Brouwer, existe un cierto $\tilde{\mathbf{x}}$ tal que $\alpha(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}$. Así que

$$\alpha(\tilde{\mathbf{x}}) = \frac{\mathbf{A}\tilde{\mathbf{x}}}{\|\mathbf{A}\tilde{\mathbf{x}}\|} = \tilde{\mathbf{x}} \implies \mathbf{A}\tilde{\mathbf{x}} = \|\mathbf{A}\tilde{\mathbf{x}}\| \tilde{\mathbf{x}}.$$

En resumen, $\tilde{\mathbf{x}}$ es un autovector de \mathbf{A} con entradas no negativas asociado a un autovalor > 0 .

Quedan algunos detalles, como demostrar que ese autovector es esencialmente único, y los demás apartados del teorema. Remitimos al lector interesado en estas cuestiones a las referencias [1], [4], [13] y [14].

7 ¿Y la cuestión computacional?

El lector crítico estará planteando ya una objeción (seria): el teorema de Perron-Frobenius garantiza la existencia del autovector que es la respuesta a nuestro problema de ordenación, pero *nada nos dice* sobre cómo calcularlo. Obsérvese que la demostración que hemos exhibido no es constructiva. Así que cabe todavía la posibilidad de que estos resultados no constituyan una respuesta satisfactoria. Recordemos que la matriz que maneja Google es gigantesca. El cálculo de nuestro anhelado autovector, ¿es realmente una cuestión abordable desde el punto de vista computacional?

Vamos a suponer que estamos en el mejor escenario posible, en las condiciones que aseguran la existencia de un autovalor λ_1 positivo *estrictamente mayor* (en módulo) que todos los demás autovalores. Llamemos \mathbf{v}_1 al autovector (positivo) asociado.

NOTA. Una matriz $\mathbf{A} \geq 0$ es **primitiva** si tiene un único autovalor dominante (mayor, en módulo, que todos los demás). Esto ocurre cuando, por ejemplo, si, para cierto entero positivo k , la matriz \mathbf{A}^k tiene todas sus entradas positivas.

Podemos, por supuesto, calcular *todos* los autovectores y quedarnos con el que nos interesa. Pero incluso aunque utilizáramos métodos eficaces para calcular autovalores, la tarea sería excesiva.

Pero de nuevo la propia estructura del problema sale en nuestra ayuda para facilitarnos el cálculo. Una feliz coincidencia. Todo arranca de la observación, hasta ahora inocente, de que el autovector que buscamos está asociado al autovalor de módulo máximo.

Supongamos, por simplificar el argumento, que \mathbf{A} es diagonalizable. Los autovectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, numerados de manera que los autovalores correspondientes vayan en orden decreciente de tamaños,

$$\lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|,$$

son una base de \mathbb{R}^n . Partimos, por ejemplo, de un $\mathbf{v}_0 \geq 0$, que escribimos como

$$\mathbf{v}_0 = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n,$$

donde los números c_1, \dots, c_n son las coordenadas de \mathbf{v}_0 en la base considerada. En lo que sigue *no va a ser necesario* calcularlos explícitamente, y de hecho no lo haremos. Nos basta con saber que tales números existen. Ahora multiplicamos el vector \mathbf{v}_0 por la matriz \mathbf{A} , para obtener

$$\mathbf{A} \mathbf{v}_0 = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + \dots + c_n \lambda_n \mathbf{v}_n,$$

puesto que los vectores $\mathbf{v}_1, \dots, \mathbf{v}_n$ son autovectores de \mathbf{A} . Sólo *calculamos* el producto de la izquierda; de lo de la derecha nos basta, para nuestros argumentos, con *saber* que ocurre. Perdone el lector la insistencia.

Cada vez más animados, repetimos la operación:

$$\mathbf{A}^2 \mathbf{v}_0 = c_1 \lambda_1^2 \mathbf{v}_1 + c_2 \lambda_2^2 \mathbf{v}_2 + \dots + c_n \lambda_n^2 \mathbf{v}_n.$$

Y lo hacemos muchas más veces, digamos k de ellas:

$$\mathbf{A}^k \mathbf{v}_0 = c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + \dots + c_n \lambda_n^k \mathbf{v}_n$$

Supongamos que $c_1 \neq 0$. Entonces,

$$\frac{1}{\lambda_1^k} \mathbf{A}^k \mathbf{v}_0 = c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{v}_n.$$

Pero como $|\lambda_j/\lambda_1| < 1$ para cada $j = 2, \dots, n$ (recordemos que λ_1 era el autovalor dominante),

$$\frac{1}{\lambda_1^k} \mathbf{A}^k \mathbf{v}_0 \xrightarrow{k \rightarrow \infty} c_1 \mathbf{v}_1.$$

De manera que, al multiplicar reiteradamente el vector inicial por la matriz \mathbf{A} , vamos determinando, cada vez con mayor precisión, la *dirección* que nos interesa, la que determina el vector \mathbf{v}_1 . Este método numérico es el bien conocido **método de las potencias**, y su velocidad de convergencia depende del cociente entre el primer y el segundo autovalor (véase en [8] una estimación para el caso de la matriz de Google).

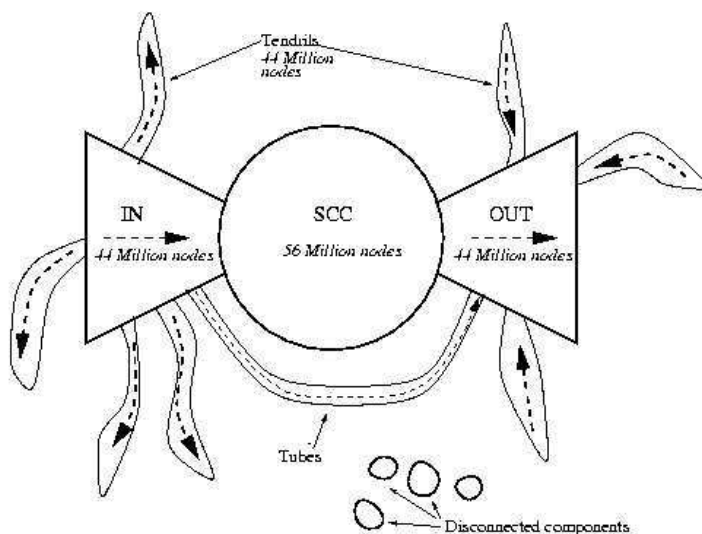
Con esta observación, nuestro problema queda resuelto, al menos si estamos en las mejores condiciones posibles (una matriz de entradas no negativas e irreducible). La respuesta que buscábamos existe, es única, y además disponemos de un método eficaz¹¹ para calcularla. Pero...

8 ¿Estamos realmente en una situación ideal?

Para que todo funcione, necesitamos que la matriz \mathbf{M} asociada al grafo G de la red sea irreducible. En otras palabras, que G sea fuertemente conexo.

NOTA. Consideremos un grafo dirigido G (un conjunto de vértices y un conjunto de aristas a las que asociamos un sentido). Si dados cualesquiera dos vértices u y v del mismo podemos encontrar una sucesión de aristas que lleven de uno a otro, entonces el grafo es *fuertemente* conexo. Si la misma conclusión se obtiene cuando “borramos” los sentidos de las aristas, entonces se dice que G es *débilmente* conexo. Por supuesto, un grafo dirigido fuertemente conexo es débilmente conexo, pero no (necesariamente) al revés.

Como el lector podrá sospechar, no es éste el caso. Un estudio realizado en 1999 (véase [7]) concluía que, de las 203 millones de páginas censadas, el 90% de ellas estaba en una gigantesca componente (débilmente) conexas, que a su vez tenía una estructura interna compleja (más, quizás, de lo que uno podría sospechar), como se muestra en el siguiente dibujo, extraído del citado [7]:



Es una estructura peculiar, que nos recuerda a un organismo biológico, una suerte de gigantesca ameba. La pieza central (SCC, *strongly connected component*) es, como el nombre indica, una componente fuertemente conexa.

¹¹Según información extraída de la propia página de Google, es un cálculo que no lleva más que unas horas de trabajo computacional.

Junto a ella aparecen otras dos piezas¹²: la IN está formada por las páginas que tienen enlaces hacia las de SCC, y la OUT está constituida por las páginas a las que apuntan los de SCC. Además aparecen una especie de dendritas, que contienen páginas desde las que no se puede llegar a los nodos de SCC, ni tampoco son accesibles desde ellos (que, en ocasiones, y para aumentar la complejidad, pueden llegar a formar tubos).

Obsérvese, de todas formas, que la organización de la red es dinámica, y cambia con el tiempo. Y no está claro si se habrá mantenido, en líneas generales, la estructura¹³ que aquí exhibimos.

¿Qué podemos hacer (o más bien, ¿qué hace Google?) ante esto? Un truco muy habitual: se trata de conseguir (de una manera razonable) estar en la mejor situación posible. Por ejemplo, *añadiendo* toda una serie de probabilidades de transición (de salida) a todos los vértices. Esto es, considerando la matriz

$$\mathbf{M}'' = c\mathbf{M}' + (1 - c) \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} (1, \dots, 1),$$

donde p_1, \dots, p_n es una distribución de probabilidad ($p_j \geq 0$, $\sum_j p_j = 1$) y c es un cierto parámetro entre 0 y 1 (en los cálculos de Google, del orden de 0.85).

Por ejemplo, podríamos tomar una distribución uniforme, $p_j = 1/n$ para cada $j = 1, \dots, n$ (con lo que la matriz ya tendría todas sus entradas positivas). Pero podríamos elegir otras, y este nuevo grado de libertad permitiría hacer búsquedas “personalizadas”.

En términos del surfista aleatorio, estamos añadiendo la posibilidad de que (con probabilidad $1 - c$) se “aburra” de seguir los enlaces y opte por saltar a otras páginas (con arreglo a cierta distribución de probabilidad).

¹²Los autores del estudio aventuran algunas explicaciones sobre estas partes: el conjunto IN podría estar formado por páginas de nueva creación, tanto, que no habría dado tiempo a que los sitios del núcleo central hubieran establecido vínculos hacia ellas. Las de la componente OUT podrían ser páginas de *webs* corporativas, que sólo contienen vínculos internos.

¹³Hay muchas otras preguntas interesantes que podemos hacernos sobre la estructura del grafo de la red. Por ejemplo, cuál es el número medio de enlaces que tiene una página, cuál es la distancia media entre dos páginas o cuál es la probabilidad $\mathbf{P}(k)$ de que un página elegida al azar tenga exactamente k enlaces (por ejemplo, entrantes). Si el grafo de la red fuera *aleatorio* (en el sentido preciso que establecieron Erdős y Rényi), entonces esperaríamos que fuera una distribución binomial (o, en el límite, una distribución de Poisson). Y que, en media, la mayor parte de las páginas tuviera un número similar de enlaces. Sin embargo, estudios empíricos (véase, por ejemplo, [2]) sugieren que el decaimiento de la distribución de probabilidad no es exponencial, sino como el de una potencia (*power law*), $k^{-\beta}$, donde β es del orden de 2 y pico. Lo que implicaría, por ejemplo, que hubiera una gran mayoría de páginas con muy pocos enlaces, al tiempo que una pequeña minoría (pero muy significativa) tendría un gran número de enlaces. Más aún, si consideramos la red como un sistema en evolución, al que sucesivamente se van añadiendo nuevas páginas, la conclusión es que esta tendencia se reforzaría. En términos coloquiales, “los ricos se vuelven cada vez más ricos”. Conclusión habitual en sistemas competitivos (y, no nos engañemos, en la vida real). Recomendamos aquí la lectura de la referencia [3].

9 Matrices no negativas en otros contextos

No crea el lector que los resultados sobre matrices no negativas que hemos ido repasando en las páginas anteriores tienen relevancia únicamente en el problema de ordenación que nos incumbe. Su campo de aplicación es mucho más amplio. Las siguientes dos observaciones pueden explicar su ubicuidad:

- En las situaciones “reales”, las interacciones que se miden son, muy frecuentemente, positivas, o al menos no negativas. Y los objetos con que codificamos estas medidas son matrices cuyas entradas son no negativas.
- Por otro lado, muchos modelos son procesos iterativos simples: de un estado inicial \mathbf{x}_0 pasamos a uno general dado por $\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0$. La convergencia del método depende del tamaño de los autovalores de \mathbf{A} , o más bien de la razón entre los tamaños de los autovalores (en particular, del más grande a los demás). Y aquí, por supuesto, el teorema de Perron-Frobenius tiene mucho que decir, si es que la matriz es no negativa.

En lo que sigue revisaremos brevemente algunos modelos en los que resultan imprescindibles (el lector podrá encontrar otras aplicaciones en [13]). En la subsección 9.4 mencionaremos algunas otras extensiones.

9.1 Modelos de evolución probabilística

Recordemos que una matriz \mathbf{A} es estocástica (o de Markov) si $\mathbf{A} \geq 0$ y, para cada columna, la suma de las entradas es 1. Para este tipo de matrices, $\lambda = 1$ es siempre un autovalor.

NOTA. Razón: las columnas de $\mathbf{A} - \mathbf{I}$ suman 0. Luego al sumar todas las filas de $\mathbf{A} - \mathbf{I}$ obtenemos el vector cero, así que las filas son linealmente dependientes. Esto quiere decir que $\mathbf{A} - \mathbf{I}$ es singular. Por tanto, $\det(\mathbf{A} - \mathbf{I}) = 0$. Lo que supone, finalmente, que $\lambda = 1$ sea un autovalor.
Además, no puede haber autovalores de módulo > 1 , pues \mathbf{A} transforma vectores cuyas entradas sumen 1 (vectores de probabilidad) en vectores del mismo tipo.

Las cadenas de Markov son un modelo probabilístico bien conocido, aplicable a situaciones muy diversas: la vista de Google es un buen ejemplo, pero también se utiliza como modelo de migraciones, de propagación de epidemias, etc.

Consideremos otra situación, que tal vez no le resulte tan familiar al lector. Supongamos que tenemos algo de dinero y que hemos decidido invertirlo para obtener un cierto rendimiento. Quizás debido a alguna experiencia reciente algo desagradable, optamos por no invertirlo en Bolsa. En este plan decididamente conservador, buscamos una inversión “segura”: le prestaremos dinero al Estado. Podría ser al Estado español, quizá a la Reserva Federal estadounidense, pero también, ¿por qué no?, al Corte Inglés, a Repsol, etc. Hay multitud de formas de hacerlo: Bonos del Tesoro, Obligaciones de empresas, emisiones de deuda, etc., aunque alguna de ellas no está al alcance del inversor particular. Es posible que obtengamos un rendimiento medio menor que en la Bolsa, pero a cambio tenemos la seguridad de que no perderemos, en ningún caso, el valor de nuestra

inversión. Y aunque no tengamos tanto pánico a la Bolsa, hemos oído por ahí que la regla básica del inversor es la de “diversificar”, y ésta parece una buena opción.

Pero no, las cosas no son tan de color de rosa. Háblenles de la seguridad de este tipo de inversiones a aquéllos que en su momento prestaron dinero al Estado argentino, al ruso o, en casos más recientes, a los que aventuraron su dinero en lo que parecían sólidas¹⁴ compañías como Enron, Parmalat o Swiss Air (si nos fallan hasta los suizos, ¡a dónde iremos a parar!).

No existen inversiones “seguras”, aunque algunas lo son más que otras. Así que el mercado se ocupa de determinar, con la máxima precisión posible, la **solvencia** de cada una de las empresas (entendidas con la mayor generalidad posible, también valen Estados) en las que uno puede invertir. Si una empresa es muy segura, en el sentido de que con altísima probabilidad devolverá el dinero (junto con los intereses) que los inversores le han prestado, entonces podrá permitirse ofrecer una baja rentabilidad. Pero si otra está catalogada como “dudosa” y pretende captar inversiones, entonces deberá necesariamente ofrecer intereses más altos que la hagan atractiva. Queda luego al arbitrio del poseedor del dinero elegir una u otra, en función de sus características (su aversión al riesgo, en términos técnicos, su vena especuladora, en lenguaje llano).

Éste es todo un mundo, el llamado *riesgo de crédito*, en el que los modelos matemáticos, algunos muy sofisticados, desempeñan un papel fundamental. Describamos aquí uno de ellos, el modelo (de evolución probabilística) de la solvencia de las empresas.

Tenemos n estados de solvencia, digamos S_1, \dots, S_n . En la jerga de las empresas que se dedican a estos menesteres clasificatorios (las agencias de calificación, Standard & Poor’s, Moody’s, etc.), estos estados son AAA, BBB+, CC, etc. (con un cierto soniquete de calificaciones escolares anglosajonas).

Tras un estudio detallado, se determina que la probabilidad de pasar del estado S_i al estado S_j es el número a_{ij} . Toda esta información se recoge en una matriz \mathbf{A} , cuyas entradas, una vez más, son no negativas. Además, es una matriz estocástica (por filas). Suele incluirse un estado especial, D (de *default*), que en la jerga de las cadenas de Markov se dice que es *absorbente* (todos ceros, menos el elemento de la diagonal). La siguiente matriz \mathbf{A} es un ejemplo¹⁵:

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	90,58	8,36	0,84	0,09	0,13	0,00	0,00	0,00
AA	0,84	90,04	8,11	0,71	0,12	0,15	0,03	0,00
A	0,14	2,69	89,94	5,92	0,85	0,32	0,04	0,10
BBB	0,03	0,32	5,94	87,41	4,92	1,04	0,13	0,21
BB	0,02	0,12	0,62	7,49	81,93	7,88	0,92	1,02
B	0,00	0,09	0,29	0,65	6,96	81,60	4,43	5,98
CCC	0,17	0,00	0,36	1,09	2,35	10,37	64,06	21,60
D	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100,0

Que nos dice, por ejemplo, que el 90.58% de las empresas con la máxima calificación siguen en la misma categoría al año siguiente.

¹⁴Sólidas, pero con productos líquidos (Enron y Parmalat) ¡o gaseosos! (Swiss Air).

¹⁵Es la matriz de Standard & Poor’s correspondiente a las probabilidades de transición, en tantos por ciento, entre estados de *rating* con los datos del periodo 1981-96.

Miramos la situación hoy, la proporción de empresas que están en cada uno de los estados de solvencia, para formar el vector inicial $\mathbf{z}^{(0)} = (z_1^{(0)}, \dots, z_n^{(0)})$. Según este modelo, el vector de proporciones vendrá dado, tras k años, por $\mathbf{z}^{(k)} = \mathbf{z}^{(0)} \mathbf{A}^k$.

Puede interesarnos estudiar, por ejemplo si queremos hacer o valorar inversiones a muy largo plazo (muchos de los bonos y obligaciones tienen horizontes temporales de 20 años o más), el *comportamiento asintótico* del sistema. Esto es, calcular el vector de proporciones cuando $k \rightarrow \infty$, lo que llamaríamos $\mathbf{z}^{(\infty)}$.

Lo sorprendente del caso es que, si $\lambda = 1$ es el único autovalor dominante (es decir, si la matriz es primitiva), entonces, *sean cuales sean las proporciones iniciales*, el estado estacionario $\mathbf{z}^{(\infty)}$ es el autovector correspondiente a $\lambda = 1$. Autovector, por cierto, que es el que la teoría de Perron-Frobenius nos ofrecía.

Como estas cosas más vale verlas antes de creerlas, recurrimos de nuevo a Maple para los cálculos, con la matriz estocástica (por columnas) siguiente:

$$A := \begin{pmatrix} 0.4 & 0.1 & 0.1 & 0.2 \\ 0.4 & 0.6 & 0.3 & 0.1 \\ 0.15 & 0.2 & 0.4 & 0 \\ 0.05 & 0.1 & 0.2 & 0.7 \end{pmatrix}$$

Ahora, amablemente, solicitamos a Maple que nos calcule los autovalores y los autovectores:

```
> eigenvectors(A);
[1.0000000, 1, {[−0.37822566, −0.78653745, −0.35673556, −0.56304047]}],
[0.4999999982, 1, {[−0.2962451280, 0.4443676920, 0.4443676932, −0.5924902550]}],
[0.2292893218, 1, {[0.2060703913, −0.7035683205, .6432117000, −.1457137708]}],
[0.3707106784, 1, {[−1.587596362, 0.9299924275, 1.780204107, −1.122600167]}]
```

A estas alturas, el resultado ya no nos sorprende: el autovector asociado al autovalor $\lambda = 1$ es el único que tiene todas las entradas con el mismo signo. Maple nos ofrece una versión del autovector cuyas entradas son todas negativas, pero esto no nos importa mucho, porque basta cambiarlas de signo (y seguiremos teniendo un autovector).

Para la comparación que haremos en un momento, conviene reescribir el autovector de manera que todas sus entradas sean positivas y que además sumen 1. El resultado es el siguiente:

$$\mathbf{v} = (0.1814432986, 0.3773195881, 0.1711340205, 0.2701030929).$$

Vamos ahora con la evolución probabilística. Partimos de un cierto vector \mathbf{v}_0 (que incluso podríamos sortear) y dejamos que la matriz haga el resto:

```
> v0:=matrix(4,1,[0.2,0.2,0.4,0.2]);
```

$$v_0 := \begin{bmatrix} 0.2 \\ 0.2 \\ 0.4 \\ 0.2 \end{bmatrix}$$

```
> seq(evalm(A^n &* v0), n=1..20);
```

$$\begin{bmatrix} 0.18 \\ 0.34 \\ 0.230 \\ 0.250 \end{bmatrix}, \begin{bmatrix} 0.1790 \\ 0.3700 \\ 0.1870 \\ 0.2640 \end{bmatrix}, \begin{bmatrix} 0.18010 \\ 0.37610 \\ 0.175650 \\ 0.268150 \end{bmatrix}, \begin{bmatrix} 0.1808450 \\ 0.3772100 \\ 0.1724950 \\ 0.2694500 \end{bmatrix}, \begin{bmatrix} 0.18119850 \\ 0.37735750 \\ 0.171566750 \\ 0.269877250 \end{bmatrix},$$

$$\begin{bmatrix} 0.1813472750 \\ 0.3773516500 \\ 0.1712779750 \\ 0.2700231000 \end{bmatrix}, \begin{bmatrix} 0.1814064925 \\ 0.3773356025 \\ 0.1711836113 \\ 0.2700742937 \end{bmatrix}, \begin{bmatrix} 0.1814293771 \\ 0.3773264713 \\ 0.1711515388 \\ 0.2700926128 \end{bmatrix}, \begin{bmatrix} 0.1814380744 \\ 0.3773223566 \\ 0.1711403163 \\ 0.2700992527 \end{bmatrix},$$

$$\begin{bmatrix} 0.1814413475 \\ 0.3773206638 \\ 0.1711363090 \\ 0.2701016795 \end{bmatrix}, \begin{bmatrix} 0.1814425722 \\ 0.3773199980 \\ 0.1711348585 \\ 0.2701025713 \end{bmatrix}, \begin{bmatrix} 0.1814430288 \\ 0.3773197423 \\ 0.1711343288 \\ 0.2701029001 \end{bmatrix}, \begin{bmatrix} 0.1814431987 \\ 0.3773196455 \\ 0.1711341344 \\ 0.2701030214 \end{bmatrix},$$

$$\begin{bmatrix} 0.1814432617 \\ 0.3773196092 \\ 0.1711340626 \\ 0.2701030663 \end{bmatrix}, \begin{bmatrix} 0.1814432852 \\ 0.3773195957 \\ 0.1711340361 \\ 0.2701030830 \end{bmatrix}, \begin{bmatrix} 0.1814432937 \\ 0.3773195905 \\ 0.1711340263 \\ 0.2701030891 \end{bmatrix}, \begin{bmatrix} 0.1814432970 \\ 0.3773195887 \\ 0.1711340226 \\ 0.2701030913 \end{bmatrix},$$

$$\begin{bmatrix} 0.1814432982 \\ 0.3773195881 \\ 0.1711340214 \\ 0.2701030923 \end{bmatrix}, \begin{bmatrix} 0.1814432987 \\ 0.3773195878 \\ 0.1711340209 \\ 0.2701030926 \end{bmatrix}, \begin{bmatrix} 0.1814432988 \\ 0.3773195877 \\ 0.1711340207 \\ 0.2701030927 \end{bmatrix}.$$

Sólo hemos incluido 20 pasos porque, como el lector puede comprobar, la convergencia al vector \mathbf{v} es vertiginosa. Una última prueba, por si alguien sigue desconfiando. Partimos ahora de otro vector inicial (muy poco “probabilístico”, en verdad):

```
> v0:=matrix(4,1,[1,0,0,0]);
```

$$v0 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

```
> seq(evalm(A^n &* v0), n=1..20);
```

$$\begin{bmatrix} 0.4 \\ 0.4 \\ 0.15 \\ 0.05 \end{bmatrix}, \begin{bmatrix} 0.225 \\ 0.450 \\ 0.200 \\ 0.125 \end{bmatrix}, \begin{bmatrix} 0.1800 \\ 0.4325 \\ 0.20375 \\ 0.18375 \end{bmatrix}, \begin{bmatrix} 0.172375 \\ 0.411000 \\ 0.195000 \\ 0.221625 \end{bmatrix}, \begin{bmatrix} 0.1738750 \\ 0.3962125 \\ 0.18605625 \\ 0.24385625 \end{bmatrix}, \begin{bmatrix} 0.17654812 \\ 0.38748000 \\ 0.17974625 \\ 0.25622562 \end{bmatrix},$$

$$\begin{bmatrix} 0.17858700 \\ 0.38265368 \\ 0.17587671 \\ 0.26288259 \end{bmatrix}, \begin{bmatrix} 0.17986435 \\ 0.38007828 \\ 0.17366947 \\ 0.26638787 \end{bmatrix}, \begin{bmatrix} 0.18059809 \\ 0.37873234 \\ 0.17246310 \\ 0.26820645 \end{bmatrix}, \begin{bmatrix} 0.18100007 \\ 0.37803822 \\ 0.17182142 \\ 0.26914027 \end{bmatrix}, \begin{bmatrix} 0.18121405 \\ 0.37768341 \\ 0.17148622 \\ 0.26961630 \end{bmatrix},$$

$$\begin{bmatrix} 0.18132584 \\ 0.37750316 \\ 0.17131328 \\ 0.26985770 \end{bmatrix}, \begin{bmatrix} 0.18138352 \\ 0.37741199 \\ 0.17122482 \\ 0.26997965 \end{bmatrix}, \begin{bmatrix} 0.18141302 \\ 0.37736601 \\ 0.17117985 \\ 0.27004110 \end{bmatrix}, \begin{bmatrix} 0.18142801 \\ 0.37734288 \\ 0.17115710 \\ 0.27007199 \end{bmatrix}, \begin{bmatrix} 0.18143560 \\ 0.37733126 \\ 0.17114562 \\ 0.27008750 \end{bmatrix},$$

$$\begin{bmatrix} 0.181439432 \\ 0.377325439 \\ 0.171139842 \\ 0.270095285 \end{bmatrix}, \begin{bmatrix} 0.181441358 \\ 0.377322518 \\ 0.171136939 \\ 0.270099184 \end{bmatrix}, \begin{bmatrix} 0.181442325 \\ 0.377321054 \\ 0.171135483 \\ 0.270101136 \end{bmatrix}, \begin{bmatrix} 0.181442811 \\ 0.377320321 \\ 0.171134753 \\ 0.270102113 \end{bmatrix}.$$

No hay trampas, ¡sale siempre el mismo resultado! Magia matemática.

9.2 Modelos de dinámica de poblaciones

En una cierta especie, los individuos se agrupan por grupos de edad C_1, \dots, C_n . La población inicial es $\mathbf{z}^{(0)} = (z_1^{(0)}, \dots, z_n^{(0)})$. Planteamos las siguientes hipótesis:

- cada individuo pasa al siguiente grupo en cada unidad de tiempo.
- en la etapa i , cada individuo da lugar a m_i descendientes.
- s_i es la fracción que sobrevive de la edad $i - 1$ a la edad i .

Este modelo, muy utilizado en Biología, es una generalización del modelo de cunicultura (vamos, la cría de conejos) que se planteó en su día Fibonacci (y lo que han dado de sí los dichosos conejos¹⁶).

La dinámica de la población viene determinada por el siguiente sistema matricial:

$$\begin{pmatrix} s_1 m_1 & s_1 m_2 & \cdots & s_1 m_{n-1} & s_1 m_n \\ s_2 & 0 & \cdots & 0 & 0 \\ 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_n & 0 \end{pmatrix}^k \begin{pmatrix} z_1^{(0)} \\ \vdots \\ z_n^{(0)} \end{pmatrix} = \begin{pmatrix} z_1^{(k)} \\ \vdots \\ z_n^{(k)} \end{pmatrix}.$$

Si la matriz anterior (conocida como **matriz de Leslie**) es “buena”, entonces

$$\mathbf{z}^{(k)} \approx K \lambda_1^k \mathbf{v}_1 \quad \text{cuando } k \rightarrow \infty,$$

donde λ_1 es “el” autovalor dominante y \mathbf{v}_1 es su autovector asociado. De donde se deduce que el comportamiento asintótico (crecimiento, extinción o quizá oscilación) de la población depende de si λ_1 es > 1 , < 1 ó $= 1$.

9.3 Modelos económicos

Una economía (muy simplificada) consta de tres sectores, digamos agricultura, industria y servicios, cuyas producciones respectivas son x_1 , x_2 y x_3 (en las unidades que sean).

La hipótesis fundamental es que el consumo que de la producción x_i hace el sector j es proporcional a x_j (la producción de j). Éste es un modelo bien conocido en Economía, el **modelo input-output de Leontief**¹⁷.

¹⁶No los conejos en sí, sino la sucesión de números de Fibonacci que surge del análisis de la cuestión. Tanto es así que, quizá como caso único en las Matemáticas, existe una revista, *Fibonacci Quarterly*, íntegramente dedicada a la sucesión y sus conexiones.

¹⁷En la exposición de motivos del Premio Nobel en Economía concedido a Leontief en 1973 se lee: “[...] por el desarrollo del método *input-output* y por sus aplicaciones a importantes problemas económicos”. El modelo que Leontief utilizó para analizar la economía de Estados Unidos en 1958 constaba de 81 sectores.

Podríamos estar, por ejemplo, en una situación como la que se describe en la siguiente tabla:

	Agricultura	Industria	Servicios	Consumidor	total producido
Agricultura	$0.3 x_1$	$0.2 x_2$	$0.3 x_3$	4	x_1
Industria	$0.2 x_1$	$0.4 x_2$	$0.3 x_3$	5	x_2
Servicios	$0.2 x_1$	$0.5 x_2$	$0.1 x_3$	12	x_3

De las x_1 unidades producidas por el sector agrario, el 30% son “autoconsumidas”, $0.2 x_2$ utilizadas por la industria, $0.3 x_3$ por el sector de servicios, mientras que 4 unidades lo son por los consumidores finales.

Ahora, la cuestión, escrita en términos matriciales, adopta la forma

$$\mathbf{Ax} + \mathbf{b} = \mathbf{x}.$$

Y la pregunta es: dado un vector de consumo $\mathbf{b} \geq 0$, ¿tiene el sistema anterior una solución $\mathbf{x} \geq 0$?

Desde luego, será el caso si la matriz $\mathbf{I} - \mathbf{A}$ tiene una inversa no negativa, pues entonces

$$\mathbf{x} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} \geq 0.$$

Una condición suficiente para garantizar la existencia de esta inversa no negativa es que el autovalor dominante de \mathbf{A} sea¹⁸ < 1 .

9.4 Extensiones de la teoría de Perron-Frobenius

La clave del teorema de Perron-Frobenius reside en que una matriz $n \times n$ de entradas por ejemplo positivas conserva el “octante” \mathbb{R}_+^n . Una versión más general de este resultado trata con *conos* en \mathbb{R}^n . El lector con especiales inclinaciones hacia la geometría del espacio euclídeo n -dimensional disfrutará leyendo las definiciones de la siguiente nota y comparando el enunciado habitual del teorema de Perron-Frobenius con la generalización que escribimos más adelante.

NOTA. Un conjunto $C \subset \mathbb{R}^n$ es un *cono* si $a\mathbf{x} \subseteq C$ para todo $\mathbf{x} \in C$ y para todo escalar $a \geq 0$. Será un *cono convexo* si $\lambda\mathbf{x} + \mu\mathbf{y} \in C$ para cualesquiera $\mathbf{x}, \mathbf{y} \in C$ y $\lambda, \mu \geq 0$. Por último, diremos que el cono es *propio* si (a) $C \cap (-C) = \{0\}$, (b) $\text{int}(C) \neq \emptyset$; y (c) $\text{span}(C) = \mathbb{R}^n$.

Teorema 3 (Perron-Frobenius, versión general) Sean A una matriz $n \times n$ y C un cono convexo propio. Si $AC \subseteq C$ (esto es, si A conserva el cono C), entonces

- a) el radio espectral $\rho(A)$ es un autovalor de A , de grado mayor al de los autovalores de igual módulo.
- b) C contiene un autovector de A correspondiente al autovalor $\rho(A)$.

¹⁸La razón es que, en este caso, $\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots = (\mathbf{I} - \mathbf{A})^{-1}$.

c) Si además $A(C \setminus \{0\}) \subseteq \text{int}(C)$, entonces $\rho(A)$ es un autovalor simple de A , mayor en módulo que cualquier otro, y su autovector asociado está en $\text{int}(C)$.

Aquellos otros lectores que gusten más del análisis funcional y la teoría espectral apreciarán la siguiente versión en dimensión infinita debida a Krein y Rutman ([12]), que aquí enunciamos tal como aparece en los comentarios al capítulo VI de [5]:

Teorema 4 (Krein-Rutman) Sea E un espacio de Banach y C un cono convexo de vértice 0. Supongamos que C es cerrado, $\text{int}(C) \neq \emptyset$ y $C \cap (-C) = \{0\}$. Sea $T : E \rightarrow E$ un operador compacto tal que

$$T(C \setminus \{0\}) \subset \text{int}(C).$$

Entonces existe un autovalor $\lambda > 0$ de T simple, de módulo máximo y cuyo autovector asociado está en C .

Pero si su especialidad son las ecuaciones en derivadas parciales, le seducirá el siguiente desafío: sea Ω un abierto regular, conexo y acotado de \mathbb{R}^n y consideremos el problema de Dirichlet para el laplaciano:

$$\begin{cases} -\Delta u = f \in L^2(\Omega), \\ u \in H_0^1(\Omega). \end{cases}$$

Compruébese¹⁹, utilizando el teorema de Krein-Rutman, que la primera autofunción del laplaciano $-\Delta$ en el problema de Dirichlet (la asociada al menor autovalor, que resulta ser simple) es *positiva*. Para los detalles, véase el apéndice del capítulo VIII de [9].

Coda

De un reto tecnológico formidable, como es el diseño de un buscador en la red, descubrimos que, al final, la clave son las Matemáticas: la astuta aplicación de teoremas y el cuidadoso análisis de la convergencia de los algoritmos. Una nueva comprobación de la *unreasonable effectiveness* de las Matemáticas en las Ciencias Naturales, que decía Eugene Wigner, y en tantos campos, que podríamos añadir nosotros.

Confiamos en que estas páginas, además de haber servido como entretenimiento para el lector, le animen a indagar por su cuenta en los numerosos problemas que aquí apenas hemos esbozado.

Y nos despedimos con especial cariño del teorema de Perron-Frobenius, que desempeña un papel tan destacado en tan variadas cuestiones. Lo hacemos con una coplilla, manriqueña y jocosa:

¹⁹ Algunas pistas: el cono que conviene considerar es

$$C = \{u \in L^2(\Omega) : u \geq 0 \text{ en casi todo punto}\}.$$

El operador T es $(-\Delta)^{-1}$, que es compacto de $L^2(\Omega)$ en $L^2(\Omega)$ y conserva el cono C . La clave es que si $f \geq 0$ en casi todo punto, entonces también $u \geq 0$ en casi todo punto.

Un hermoso resultado
 que además se nos revela
 indiscreto;
 y un tanto desvergonzado,
 porque de Google desvela
 ... su secreto.

Para saber más

- [1] BAPAT, R. B. Y RAGHAVAN, T. E. S.: *Nonnegative matrices and Applications*. Encyclopedia of Mathematics and its Applications **64**. Cambridge University Press, 1997.
- [2] BARABÁSI, A.-L.: *The Physics of the web*. Physics World (julio 2001). Se puede obtener en la página web del autor, <http://www.nd.edu/~alb>.
- [3] BARABÁSI, A.-L.: *Linked, the new Science of Networks. How everything is connected to everything else and what it means*. Plume Books, 2003.
- [4] BERMAN, A. Y PLEMMONS, R. J.: *Nonnegative matrices in the Mathematical Sciences*. Classics in Applied Mathematics **9**. Academic Press, 1979.
- [5] BRÉZIS, H.: *Análisis funcional. Teoría y aplicaciones*. Alianza Editorial, Madrid, 1984.
- [6] BRIN, S. Y PAGE, L.: *The anatomy of a large-scale hypertextual web search engine*. Accesible en: www-db.stanford.edu/~sergey/
- [7] BRODER, A. ET AL.: *Graph structure in the web*. Accesible en la página web: www9.org/w9cdrom/160/160.html.
- [8] HAWELIWALA, T. S. Y KAMVAR, S. D.: *The second eigenvalue of the Google matrix*. Se puede obtener en la página web: www.stanford.edu/~taherh/papers/secondeigenvalue.pdf.
- [9] DAUTRAY, R. Y LIONS, J.-L.: *Mathematical Analysis and Numerical Methods for Science and Technology. Volumen 3: Spectral Theory and Applications*. Springer-Verlag, Berlin, 1990.
- [10] KEENER, J. P.: *The Perron-Frobenius Theorem and the ranking of football teams*. SIAM Review **35** (1993), no. 1, 80–93.
- [11] KENDALL, M. G.: *Further contributions to the theory of paired comparisons*. Biometrics **11** (1955), 43–62.
- [12] KREIN, M. G. Y RUTMAN, M. A.: *Linear operators leaving invariant a cone in a Banach space*. Uspehi Matem. Nauk (N. S.) **3** (1948), no. 1, 3–95 [ruso]. Traducido al inglés en Amer. Math. Soc. Transl. **10** (1962), 199–325.

- [13] MACLAUER, C. R.: *The Many Proofs and Applications of Perron's Theorem*. SIAM Review **42** (2000), no. 3, 487–498.
- [14] MINC, H.: *Nonnegative matrices*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, New York, 1988.
- [15] MOLER, C.: *The world's largest matrix computation*. Se puede obtener en la siguiente página web: www.mathworks.com/company/newsletter/clevescorner/oct02_cleve.shtml.
- [16] WEI, T. H.: *The algebraic foundations of ranking theory*. Cambridge University Press, London, 1952.
- [17] WILF, H. S.: *Searching the web with eigenvectors*. Accesible en la dirección: www.cis.upenn.edu/~wilf/.

Formación de singularidades y problemas de frontera libre en mecánica de fluidos

M.A. FONTELOS

Departamento de Matemáticas, Universidad Autónoma de Madrid

mafontel@escet.urjc.es

Resumen

En este artículo exponemos algunos de los resultados más relevantes obtenidos por el autor en torno a problemas que involucran la evolución de los fluidos y de las superficies que los separan del medio externo o de otro fluido. El interés esencial es el de entender los diversos tipos de singularidad que pueden aparecer debido al contacto con superficies sólidas, al cambio de topología del dominio ocupado por el fluido o a la formación de frentes u otro tipo de estructuras en el seno del mismo.

Palabras clave: *Ecuaciones en derivadas parciales, Problemas de frontera libre, Mecánica de fluidos, Singularidades.*

Clasificación por materias AMS: *35Q30, 35R35, 74H35, 35Q35*

1 Introducción

Entendemos como singularidad, en el contexto de la mecánica de fluidos, la existencia o aparición espontánea de valores no acotados (en algún punto del espacio) para las soluciones de los sistemas de ecuaciones diferenciales que modelan la mecánica de fluidos.

Así formulado, el tema de investigación es vastísimo. Esto se debe principalmente al hecho de que, tradicionalmente, los problemas de mecánica de fluidos han sido uno de los centros de atención fundamentales en la comunidad de la matemática aplicada. No es de extrañar, ya que la no linealidad inherente a las ecuaciones de Navier-Stokes —las que modelan los fluidos Newtonianos incompresibles— y todas las relacionadas supone el obstáculo teórico fundamental en su tratamiento y análisis por parte de físicos e ingenieros. Es imposible exagerar la importancia que el tratamiento matemático ha tenido en el avance de la comprensión del comportamiento de los fluidos y sus consecuentes aplicaciones. Citemos como ejemplo claro la importancia que la

Fecha de recepción: 09/11/04

teoría de capa límite desarrollada por von Karman y Prandtl ha tenido en la comprensión del vuelo y la repercusión práctica que ello tiene en la aeronáutica. Tampoco es posible exagerar la importancia de los problemas que aún quedan abiertos. Recordemos, por ejemplo, el célebre problema de Leray (ver [41]) sobre la existencia de soluciones que desarrollen singularidades espontáneas para el sistema de Navier-Stokes, que lleva abierto 70 años y ha atraído la atención de muchos de los matemáticos más importantes del siglo XX.

La no linealidad de la que hemos hablado en el párrafo anterior puede, en principio, dar lugar a soluciones singulares (en el sentido definido al principio). El problema matemático es pues claro: determinar si existen o no soluciones singulares en el contexto de la mecánica de fluidos y, en caso afirmativo, hallar la estructura de las mismas. Las singularidades pueden ser de naturaleza diversa, pero nosotros, para fijar ideas, nos centraremos en tres tipos:

- 1.- Singularidades que aparecen y persisten en la interfase de separación del fluido con otro u otros medios (sólido o gaseoso).
- 2.- Singularidades espontáneas por cambio de topología de la región ocupada por el fluido.
- 3.- Singularidades espontáneas en el seno del fluido.

La experiencia del autor se centra en los dos primeros tipos de singularidades. El tercero, consiste esencialmente en el problema de Leray citado arriba, y hasta el momento solamente los hemos abordado en el contexto de modelos uni y bidimensionales análogos a los sistemas de Euler y Navier-Stokes (en un sentido que precisaremos más abajo) .

2 Singularidades en interfases y el problema de la línea de contacto

En esta sección comenzaremos con una formulación matemática general del problema. Basándonos en ella comentaremos algunos de los problemas interesantes planteados y su aparición en el contexto de las aplicaciones tanto industriales como puramente científicas.

La formulación sería la siguiente: una masa fluida ocupa una región del espacio bi- o tri-dimensional que puede ser acotada o no acotada. La frontera del dominio (que asumimos en principio simplemente conexo) ocupado por el fluido está compuesta por líneas o superficies de contacto con un sólido y por líneas o superficies de contacto con otro fluido o con el vacío (nosotros nos centraremos en este último caso). Las líneas o superficies de contacto con el sólido pueden ser fijas o móviles y pueden ser suaves o tener esquinas. Las líneas o superficies de contacto con otro fluido o el vacío son, en principio, móviles. Finalmente, los fluidos que consideramos pueden ser líquidos Newtonianos o viscoelásticos.

Como hablaremos a menudo de fluidos viscoelásticos, recordemos en este punto lo que eso significa. Los fluidos viscoelásticos son aquellos que combinan aspectos genuinos de los fluidos, como la existencia de esfuerzos viscosos, con efectos propios de los sólidos elásticos como la aparición de fuerzas de recuperación en respuesta a esfuerzos ejercidos sobre el fluido. Los fluidos

viscoelásticos entran dentro de la familia de fluidos no-Newtonianos y se pueden generar en el laboratorio mediante adición de polímeros en forma de disolución en el seno de un fluido newtoniano como el agua.

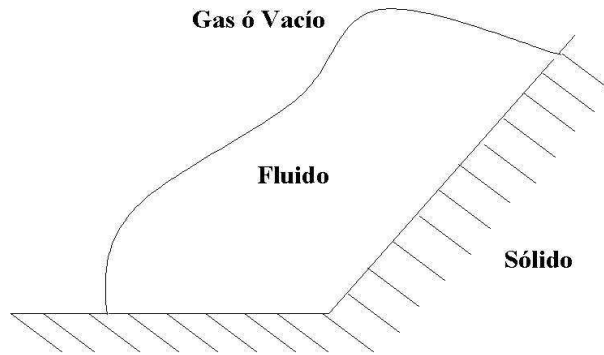


Figura 1: Esquema básico general de los problemas de interfases fluidas.

Como es usual en mecánica de fluidos, el estado dinámico se describe mediante el campo vectorial de velocidades $\vec{v}(\mathbf{r}, t)$ y el campo escalar de presiones $p(\mathbf{r}, t)$. Las ecuaciones satisfechas por estos campos son

$$\rho \left(\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} \right) = \nabla \cdot \mathbf{T} \quad (1)$$

$$\nabla \cdot \vec{v} = 0 \quad (2)$$

siendo ρ la densidad del fluido y \mathbf{T} el tensor de esfuerzos. Este tensor de esfuerzos se relaciona con \vec{v} y p mediante una relación constitutiva. En el caso de fluidos Newtonianos dicha relación constitutiva tiene componentes

$$T_{ij} = -p\delta_{ij} + \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \quad (3)$$

siendo μ la viscosidad del fluido. Si el fluido es No-Newtoniano la situación es más complicada por el hecho de que, a fecha de hoy, no hay relación constitutiva capaz de explicar toda la fenomenología asociada a este tipo de fluidos. De hecho, una de las motivaciones científicas que orientan nuestra investigación con estos tipos de fluidos es la de usar leyes constitutivas con una razonable sustentación

teórica para explicar y predecir hechos experimentales que las validen (siquiera parcialmente). Sin ánimo de entrar en una disquisición teórica sobre cuáles leyes constitutivas son buenas y cuáles no, diremos que el acuerdo casi mayoritario entre la comunidad reológica se inclina hacia las leyes constitutivas de tipo Oldroyd-B para disoluciones poliméricas diluídas, que constituyen una familia amplia dentro de los fluidos No-Newtonianos de uso más común en aplicaciones (ver [39] y [6]). Esta relación es, en los casos más habituales, de la forma

$$T_{ij} = -p\delta_{ij} + \mu_s \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + S_{ij}, \quad (4)$$

donde μ_s es la viscosidad del fluido Newtoniano en cuyo seno se hallan disueltas moléculas de polímero que hacen al fluido No-Newtoniano. S_{ij} es por tanto la componente ij del tensor \mathbf{S} que supone la contribución genuinamente No-Newtoniana, y satisface

$$\mathbf{S} + \lambda (\mathbf{S}_{(1)} + \{\mathbf{DS} + \mathbf{SD}\}) = 2\mu_p \mathbf{D}. \quad (5)$$

con λ el tiempo de relajación del polímero, μ_p la contribución del polímero a la viscosidad total de la disolución en el límite de deformaciones nulas,

$$D_{ij} = 2\mu_s \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)$$

y $\mathbf{S}_{(1)}$ definido como

$$\mathbf{S}_{(1)} \equiv \frac{\partial \mathbf{S}}{\partial t} + (\vec{v} \cdot \nabla) \mathbf{S} - (\nabla \vec{v})^T \mathbf{S} + \mathbf{S} \nabla \vec{v},$$

La inserción de (3) en (1) da lugar a un sistema de ecuaciones parabólicas no lineales para las componentes de la velocidad, mientras que (5) es una ecuación hiperbólica para \mathbf{S} y hay que tratarla en pie de igualdad con (1); es decir, formando un sistema consistente en (1), (2), (4) y (5). El alto número de ecuaciones y el carácter mixto parabólico-hiperbólico o parabólico-parabólico degenerado no hace más que añadir dificultades matemáticas al ya de por sí difícil sistema (1), (2) con \mathbf{T} dado por (3).

No existe aún una teoría general de existencia y unicidad para este sistema de ecuaciones salvo en situaciones físicas muy restrictivas. Parte de nuestro trabajo, en colaboración con A. Friedman, ha girado en torno a estas cuestiones básicas. En [20] se estudiaron flujos estacionarios de fluidos no-Newtonianos en dominios no acotados cuya forma geométrica es la de un canal (en dimensión 2) o una tubería (en dimensión 3) deformadas. El flujo satisface condiciones de contorno de no deslizamiento y se comporta como un flujo de Poiseuille en el infinito. Probamos existencia y unicidad de soluciones bajo la suposición de alta viscosidad cinemática. Los resultados se aplican a una gran familia de modelos que incluye el modelo Oldroyd-B como caso particular. Además, para los fluidos de segundo orden se prueba la existencia y unicidad de soluciones en el caso en el que el dominio tenga esquinas.

En el caso bidimensional (el tridimensional se analiza de forma análoga), la idea fundamental yace en el hecho de que, en el caso de estos fluidos, el sistema de ecuaciones se puede escribir en la siguiente forma:

$$\Delta^2\psi + \varepsilon(\vec{v} \cdot \vec{\nabla})\Delta^2\psi = \varepsilon F(D^\alpha\psi, |\alpha| \leq 4)$$

siendo ε un parámetro pequeño y F una cierta función no lineal. Esta ecuación de quinto orden puede verse como una ecuación de transporte para $\Delta^2\psi$, donde \vec{v} está relacionado con ψ por la relación $\vec{v} = (v_x, v_y) = (\psi_y, -\psi_x)$. La existencia y unicidad se prueba usando teoría clásica de regularidad para el operador biarmónico en espacios de Hölder combinada con estimaciones sobre la regularidad de las curvas características que debemos obtener en dichos espacios haciendo uso de la desigualdad de Gronwall. Un punto fijo de Banach sirve para poner todas estimaciones juntas y concluir la demostración.

En el caso de dominios con esquinas debemos usar el marco funcional introducido por Maz'ya, Plamenevskij y colaboradores (ver [37], [52]) y estimaciones de ciertas integrales mediante métodos asintóticos clásicos. Esa estrategia fue la seguida en [22] para describir el flujo alrededor de una esquina.

Si la geometría del problema involucra fronteras con el vacío debe entonces imponerse una condición de equilibrio dinámico entre los esfuerzos del fluido y la fuerza de tensión superficial. Esto se traduce en la relación:

$$T_{ij}n_j = \sigma H n_i \quad (6)$$

donde n_j es la componente j -ésima del vector normal a la interfase, H es la curvatura media de la misma y σ es el coeficiente de tensión superficial del fluido.

Finalmente, como la interfase fluido-vacío puede moverse libremente, hace falta una ecuación para su movimiento. Dicha ecuación es de la forma

$$\frac{d\mathbf{r}(t)}{dt} = \vec{v}(\mathbf{r}(t), t)$$

siendo $\mathbf{r}(t)$ la posición de un punto material en el instante t y $\vec{v}(\mathbf{r}(t), t)$ el campo de velocidades del fluido en dicho punto en el mismo instante.

Un contexto en el que aparecen singularidades es el de la extrusión de fluidos viscoelásticos. La situación es la siguiente: un chorro de material viscoelástico es obligado a salir de un tubo aplicando presión en uno de los extremos. Si la presión no es muy alta, entonces la superficie del chorro es suave, pero si la presión aplicada es suficientemente grande (mayor que un cierto valor crítico), entonces aparece una inestabilidad y la superficie se vuelve algo irregular con una perturbación de pequeña amplitud y alta frecuencia que se traduce en una geometría rugosa de la misma conocida con el término inglés *sharkskin* (*piel de tiburón*). Por encima de un segundo valor crítico aparece intermitencia entre regiones de *sharkskin* y regiones de superficie suave. Por encima de un tercer valor crítico, el flujo se vuelve completamente irregular (ver los experimentos en [35]). ¿Qué da lugar a estos fenómenos?. No hay respuesta, pero la idea

generalizada es que sea lo que sea que ocurre, debe de estar muy relacionado con el comportamiento del flujo cerca del punto triple que separa la pared del tubo, la superficie del chorro, y el aire. Esto ocurre, por ejemplo, en el caso Newtoniano donde el diámetro del chorro emergente con respecto al diámetro del tubo depende fuertemente de la estructura del flujo a la salida del tubo. En el caso viscoelástico hay otro fenómeno no entendido que es el "hinchado" anormalmente grande que sufre el chorro a la salida. Este fenómeno recibe el nombre de *die-swell* (salida-hinchado). La versión linealizada del problema es el llamado problema de "stick-slip" (pegado-deslizado), que se puede visualizar como el flujo a la salida de un conjunto infinito de canales paralelos. Nuestra contribución a este problema (ver [21]) ha consistido en el establecimiento de la existencia y unicidad de soluciones y en derivar la presencia de una característica genuinamente no-newtoniana: la aparición, cerca de los puntos de contacto entre las paredes y la interfase, de una singularidad nueva que se traduce en un término de orden inferior a la singularidad Newtoniana pero que es altamente oscilatorio. Las técnicas utilizadas son estimaciones de tipo Schauder en espacios de Hölder con pesos (cerca de las esquinas) especiales y evaluación de ciertas integrales mediante técnicas clásicas del análisis asintótico.

Otro problema que se ajusta al esquema teórico descrito es el de los flujos recubridores, que son aquellos que aparecen al depositar una película fluida sobre una superficie sólida. Pensemos por ejemplo en una capa de pintura que ha de ser depositada sobre la carrocería de un coche. El principal obstáculo con el que siempre se ha chocado para estudiar problemas relacionados (como por ejemplo el del modo de extender la capa de modo que tenga un espesor preciso y se gaste la mínima cantidad de pintura posible) es el llamado "problema de la línea de contacto". El problema tiene lugar en la línea triple que conecta las fases sólida (substrato), líquida y gaseosa (aire). Si se imponen condiciones de contorno clásicas, como la de ángulo de contacto constante, se llega a resultados paradójicos como la aparición de esfuerzos infinitos y no integrables en la vecindad de dicha línea. Las aproximaciones basadas en la teoría de la lubricación (ver [46], [2], [3], [49]) y que dan lugar a ecuaciones del tipo

$$h_t + \nabla \cdot (h^{n+2} \nabla \Delta h) = 0$$

(con $h(x, y, t)$ la altura de una película fluida y $0 \leq n \leq 1$) también tropiezan con dificultades que se traducen en la no-unicidad de soluciones. Para una descripción general del problema y los intentos de resolución propuestos referimos al artículo clásico de Ngan y Dussan (ver [47]). A pesar de la sencillez de formulación de esta clase de problemas, las dificultades matemáticas con las que uno se encuentra son enormes: falta de unicidad, elevado orden en las ecuaciones, condiciones de contorno mixtas (dependiendo de que en un punto del contorno de la masa fluida haya contacto con gas, sólido o los dos), etc. Los intentos de dar solución a los mismos han ocupado buena parte del tiempo de algunos de los grandes expertos en ecuaciones en derivadas parciales. Citemos los trabajos de Friedman y Velázquez (ver [32], [33]) y los trabajos de Solonnikov (ver [59]) en los que se demuestra la existencia de soluciones no paradójicas en el caso de interfases que avanzan sobre un substrato con ángulo de contacto nulo,

pero para un fluido de Stokes en aproximación cuasiestacionaria; es decir, con el término a la izquierda de (1) nulo. En [26] se mostró que dichas soluciones son, de hecho, estables. El resultado análogo para el sistema de Navier-Stokes no ha sido aún obtenido.

En [4] investigamos el derrame de películas fluidas de fluidos del tipo "ley de potencias", en los que la viscosidad del fluido depende de las deformaciones como una ley de potencias de exponente $\lambda - 1$ (con $\lambda = 1$ correspondiendo al fluido Newtoniano). Construimos soluciones explícitas del tipo onda viajera y autosimilares. Mostramos que si el exponente de la no linealidad λ asociada a la reología es mayor que uno, entonces la ecuación unidimensional que gobierna la evolución es:

$$h_t + (h^{\lambda+2} |h_{xxx}|^{\lambda-1} h_{xxx})_x = 0$$

y admite soluciones de soporte compacto y frentes móviles (soluciones autosimilares). También comprobamos que esta solución tiene además energía y tasa de disipación de energía acotadas. Esto implica la no existencia de la paradoja de la línea de contacto en esta clase de fluidos. El resultado es fácilmente generalizable a geometrías bidimensionales.

3 Singularidades por cambio de topología

La experiencia cotidiana demuestra que en numerosas circunstancias el dominio $\Omega(t)$ ocupado por un fluido cambia su topología. Más concretamente, el dominio $\Omega(t)$ que inicialmente es simplemente conexo puede evolucionar a un dominio múltiplemente conexo. El interés por esta cuestión es antiguo y lo podemos remontar a las observaciones experimentales de Savart (ver [57]) y a los trabajos de Plateau (ver [53]) y sobre todo de Rayleigh (ver [54]). Nosotros nos restringiremos al estudio del caso en el que los dominios $\Omega(t)$ son siempre axisimétricos. Para una revisión del trabajo hecho en torno a este problema hasta 1997 ver [19].

Supongamos que $\Omega(t)$ es un tubo deformado extremadamente fino y que se extiende hasta el infinito. Cuando un cierto parámetro adimensional que mide esencialmente el espesor del tubo frente a sus dimensiones longitudinales es pequeño, es posible deducir a partir del sistema de Navier-Stokes, mediante un análisis perturbativo, el siguiente sistema aproximado llamado sistema unidimensional:

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial z} = -\frac{\partial \kappa}{\partial z} + 3\mu \frac{1}{h^2} \frac{\partial}{\partial z} \left(h^2 \frac{\partial v}{\partial z} \right), \quad (7)$$

$$\frac{\partial h}{\partial t} + v \frac{\partial h}{\partial z} = -\frac{h}{2} \frac{\partial v}{\partial z} \quad (8)$$

donde μ es un parámetro adimensional que se definirá abajo y la coordenada z se considera a lo largo del eje del tubo y va de $-\infty$ a $+\infty$. La función $h(z, t)$ representa la distancia de un punto de la frontera libre a eje, mientras que la función $v(z, t)$ representa la velocidad del fluido a lo largo del eje. κ es la

curvatura media de la superficie dada por

$$\kappa = \frac{1}{h(1+h_z^2)^{\frac{1}{2}}} - \frac{h_{zz}}{(1+h_z^2)^{\frac{3}{2}}}. \quad (9)$$

Los parámetros originales de las ecuaciones de Navier-Stokes (densidad ρ , viscosidad cinemática ν , coeficiente de tensión superficial σ , radio característico del chorro R) se relacionan con μ mediante la relación $\mu = \nu\sqrt{\rho/(R\sigma)}$. El parámetro adimensional μ se llama número de Ohnesorge y representa el peso relativo de las fuerzas viscosas con respecto a las fuerzas capilares.

El modelo unidimensional se deduce bajo la suposición de que las longitudes características de la interfase perturbada son mucho mayores que el radio del tubo y que esta característica se mantiene válida para todo tiempo (cf. [18]). Esto equivale a decir que el parámetro R es muy pequeño comparado con la longitud de las perturbaciones de la frontera libre (suposición de "delgadez"). El sistema (7), (8) debe ser complementado con condiciones de contorno e iniciales adecuadas. Como condiciones iniciales tomaremos la geometría y velocidad iniciales a lo largo del tubo $h(z,0) = h_0(z)$, $v(z,0) = v_0(z)$. Tanto $h_0(z)$ como $v_0(z)$ se asumen suficientemente regulares. Si modelamos un tubo libre es natural imponer $h_0(z) > 0$, y $h(z,t) \rightarrow 1$, $v(z,t) \rightarrow 0$ cuando $|z| \rightarrow \infty$. Estas condiciones implican que el chorro es estacionario en el infinito. La suposición de "delgadez" se traduce, mientras es válida, en valores muy pequeños h_z y h_{zz} de modo que el término a la derecha de (9) se puede sustituir por $\frac{1}{h}$.

El parámetro μ depende linealmente de la viscosidad y puede tomar todos los valores posibles desde los muy pequeños a los muy grandes. Tomemos por ejemplo dos fluidos comunes como el agua y la glicerina y las mezclas de los mismos tal y como se hizo en los experimentos de [61]. Los valores típicos de R que se consideraron allí fueron 0,1 *cms.*, mientras que los valores de ν, ρ y σ de ambos líquidos a temperatura ambiente dan $3\mu = 2,98 \cdot 10^{-2}$ para el agua y $3\mu = 24,95$ para la glicerina. La adición de surfactantes puede hacer decrecer el valor de σ dramáticamente y, en consecuencia, incrementar el valor de μ . Valores menores de R (en el rango de los micrometros) fueron observados en [36] y una situación como esta daría lugar a valores de μ del orden de 10^3 . Por tanto, está justificado un análisis asintótico del sistema (7), (8) cuando $\mu \rightarrow 0$ y $\mu \rightarrow \infty$. De hecho, ambos límites se han considerado en la literatura (cf. por ejemplo [43]) para el caso $\mu = 0$ y [56], [50] para el caso que corresponde formalmente a $\mu = \infty$). Bajo la acción de la gravedad actuando en la dirección z (es decir, un chorro que cae), debemos restar al término a la derecha de (7) el número de Bond dado por $(\rho g R^2 / \sigma)$ que resulta ser, para los ejemplos descritos, del orden de 10^{-3} . A efectos de formación de singularidades podemos despreciarlo, aunque podría ser incluido fácilmente en los análisis.

El sistema (7), (8) es conocido desde hace bastante tiempo. Hay varios fenómenos para los que se piensa que el modelo unidimensional puede ofrecer una buena descripción. Uno de ellos es la estructura del flujo hasta tiempos próximos al tiempo de ruptura del tubo (cf. [18]), que es el tiempo en el que la función $h(z,t)$ alcanza el valor cero en algún punto por primera vez. El análisis

de la ruptura se ha llevado a cabo casi exclusivamente mediante la búsqueda de soluciones autosimilares de (7), (8). En [18] se encontró una solución autosimilar numéricamente. Otro fenómeno interesante es el de *overturning*, que consiste en la aparición espontánea de singularidades en la derivada espacial de la función $h(z, t)$. Este hecho ha sido observado numéricamente tanto en el caso no viscoso como en el de débil viscosidad. Por último, hay un fenómeno sorprendente y aún no entendido: la formación de estructuras iteradas que consisten en hilos largos y finos previos a la ruptura en el régimen muy viscoso $\mu \gg 1$ (cf. [58]).

El objeto de los artículos [27], [28] fue el estudio del sistema (7), (8) en el límite $\mu = 0$. Sorprendentemente, este sistema no-lineal permite un tratamiento analítico preciso vía un proceso de linealización consistente en dos pasos: 1) la formulación del sistema en coordenadas lagrangianas, 2) la aplicación de la transformación hodógrafa, conocida en el contexto de la dinámica de gases y que introducimos, como contribución original al campo, en nuestro problema.

Tras esta linealización, el problema se transforma en la resolución de la ecuación elíptica

$$\frac{\partial^2 t}{\partial y^2} + w^3 \frac{\partial}{\partial w} \left(\frac{1}{w^3} \frac{\partial t}{\partial w} \right) = 0 \quad (10)$$

en un dominio exterior plano con adecuadas condiciones de contorno.

La construcción de soluciones explícitas de este problema, que es relativamente sencilla dada su linealidad, se traduce en la construcción de soluciones explícitas de (7), (8). En particular:

1) Las soluciones autosimilares de (7), (8) que representan ruptura en tiempo finito son soluciones de (10) que se obtienen por separación de variables cuando se escriben en coordenadas polares ($(w, y) = (r \cos \theta, r \sin \theta)$) y que involucran potencias de r y funciones de Legendre para la dependencia en θ .

2) Soluciones que se forman por combinación lineal de infinitas de estas soluciones y que representan un chorro cilíndrico que se rompe en tiempo finito. Es remarcable que un principio de superposición lineal se pueda aplicar a un problema no lineal como el (7), (8) a través de la transformación hodógrafa.

3) Soluciones tales que h desarrolla cúspides (singularidad en la derivada) en tiempo finito.

4) las soluciones que desarrollan singularidades de tipo fractal. La existencia de este tipo de soluciones en circunstancias experimentales reales no es clara dado al hecho de que estamos despreciando en nuestro modelo algunos términos derivados de la tensión superficial que suelen "regularizar" las soluciones en este tipo de problemas. Nuestro objetivo en este artículo era el de mostrar la relación entre relaciones de dispersión de tipo catastrófico (crecientes y positivas) y la aparición de estructuras autosimilares de tipo fractal. La perspectiva es pues más amplia que la propia de los chorros viscosos y busca exponer un modelo general de formación de estas estructuras que podría ser exportado a otros problemas importantes, como Hele-Shaw y Stefan, donde sí aparecen. La idea esencial consiste en formular el problema en el plano hodógrafo con dominios limitados por curvas análogas a la curva de Koch. Esto, traducido al plano físico, da lugar a soluciones regulares que se vuelven singulares fractales en

tiempo finito.

El artículo [31] presenta un análisis matemático del sistema (7), (8) con μ finito. Tres son los resultados fundamentales:

- 1) Para datos iniciales suficientemente regulares, las soluciones de (7), (8) se pueden continuar con la misma regularidad hasta el instante en el que h se anula en algún punto.
- 2) Bajo ciertas condiciones en μ y los datos iniciales, h se anula en algún punto al menos cuando $t \rightarrow \infty$.
- 3) Para grandes valores de μ , se tiene

$$h(z, t) = \frac{1}{\sqrt{g^{-1}\left(\frac{t}{2\mu} + g(h_0^{-2}(f^{-1}(z; t)))\right)}} + \mu^{-\frac{1}{2}}A(z, t) \text{ en } [0, \mu T^*] \quad (11)$$

con

$$T^* < -2(\inf h_0(z) + \log |1 - \inf h_0(z)|) \equiv \mu^{-1}t_0 ,$$

$A(z, t)$ una función acotada en $[0, \mu T^*]$ con la cota independiente de μ , $g(u) \equiv u^{-\frac{1}{2}} + \log \left| 1 - u^{-\frac{1}{2}} \right|$ y $f(s; t) = \int_0^s g^{-1}\left(\frac{t}{2\mu} + g(h_0^{-2}(\sigma))\right)d\sigma$ (su inversa se toma para t fijo). El primer término a la derecha de (11) es tal que desarrolla un cero en algún tiempo $t = t_0$, pero previamente desarrolla un filamento muy largo y fino con radio que escala como $(t_0 - t)^{\frac{1}{2}}$, lo que está de acuerdo con las observaciones experimentales en [36].

En ausencia de aproximaciones como la unidimensional mencionada arriba, el único resultado matemático al respecto de la formación de singularidades es el presentado en [11], [12]. En estos artículos probamos la imposibilidad de ruptura de chorros fluidos a través del colapso uniforme de un filamento fluido. Probamos que el volumen encerrado por un filamento satisface la siguiente desigualdad:

$$V(t) \geq Ce^{-Ct^2}$$

de modo que la ruptura, caso de ocurrir en el sistema de Navier-Stokes bajo la acción de fuerzas de tensión superficial, debe ser puntual. La estrategia de demostración se basa en el establecimiento de una desigualdad para la masa encerrada en el tubo en términos de la norma L^∞ de la velocidad y la estimación de la misma en la región de interés mediante estimaciones de energía con pesos.

La situación cambia en el caso de fluidos viscoelásticos. Es un hecho singular y bien documentado (cf. [34]) que la adición de polímeros a un fluido Newtoniano en forma de solución es a menudo capaz de retrasar significativamente el proceso de ruptura en tubos fluidos debido a la inestabilidad de Rayleigh. Incluso pequeñas trazas de polímero pueden transformar el proceso de ruptura, que en un fluido Newtoniano dura una fracción de segundo, en un proceso de segundos e incluso minutos para la solución. El interés industrial de este hecho es considerable. La razón estriba en que el conocimiento teórico preciso de los procesos involucrados en el retraso de la ruptura puede otorgar pistas de posibles modos de controlar el momento, la posición y el número de gotas a las que da lugar. La emisión de chorros es un procedimiento usado, por ejemplo, en las

impresoras llamadas "de chorro de tinta" o en diversos procesos industriales como los de soldadura.

Durante la evolución el tubo desarrolla una estructura de gotas-en-alambre donde una secuencia de gotas espaciadas de forma más o menos aleatoria están conectadas por filamentos muy finos (ver la Figura 2). Esta estructura permanece durante un largo periodo de tiempo (del orden del tiempo de relajación del polímero) después del cual la ruptura tiene lugar. Otro interesante fenómeno descrito en [34] es el movimiento de las gotas a lo largo de la configuración gotas-en-alambre y la coalescencia de las mismas que tiene lugar.

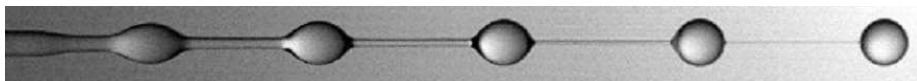


Figura 2: Estructura de gotas-en-alambre. El fluido es una disolución acuosa al 0,01 en peso de poliáciamida.

El estudio del problema de frontera libre asociado (consistente en resolver el sistema de Navier-Stokes con relaciones constitutivas y condiciones de contorno apropiadas en un dominio móvil) es prácticamente imposible desde el punto de vista analítico. Usaremos, en su lugar, el llamado modelo unidimensional análogo al descrito arriba para fluidos Newtonianos. Como ya hemos mencionado, los fluidos No-Newtonianos y, en particular, las soluciones poliméricas, se caracterizan desde el punto de vista matemático por tener relaciones constitutivas que no son simplemente una relación algebraica (y lineal) entre los tensores de deformaciones y esfuerzos, sino un sistema de ecuaciones diferenciales parciales de primer orden. Para un fluido con ley constitutiva del tipo Johnson-Segalman la aproximación unidimensional lleva al sistema:

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial z} = \frac{\partial \kappa}{\partial z} + \frac{1}{h^2} \frac{\partial}{\partial z} \left(h^2 (3\mu_s \frac{\partial v}{\partial z} + \sigma_{zz} - \sigma_{rr}) \right), \quad (12)$$

$$\frac{\partial h}{\partial t} + v \frac{\partial h}{\partial z} + \frac{1}{2} h \frac{\partial v}{\partial z} = 0, \quad (13)$$

$$\sigma_{zz} + D \left(\frac{\partial \sigma_{zz}}{\partial t} + v \frac{\partial \sigma_{zz}}{\partial z} - 2a \frac{\partial v}{\partial z} \sigma_{zz} \right) = 2\mu_p \frac{\partial v}{\partial z}, \quad (14)$$

$$\sigma_{rr} + D \left(\frac{\partial \sigma_{rr}}{\partial t} + v \frac{\partial \sigma_{zz}}{\partial z} + a \frac{\partial v}{\partial z} \sigma_{rr} \right) = -\mu_p \frac{\partial v}{\partial z}, \quad (15)$$

análogo al (7), (8) pero con dos ecuaciones adicionales (14), (15) correspondientes a las contribuciones dominantes del esfuerzo debido a la presencia del polímero en disolución. El parámetro a es tal que $-1 \leq a \leq 1$, siendo el límite $a = 1$ el correspondiente al modelo Oldroyd-B. El parámetro D (llamado número de Deborah) es la razón entre el tiempo de relajación λ

y el tiempo característico $(\rho R^3/\sigma)^{\frac{1}{2}}$. Puede ser muy grande, ya que el tiempo característico es del orden de 10^{-3} s (en un tubo de radio 0,1 mm. como los considerados en [34] por ejemplo) mientras que λ puede ser 3 ó 4 órdenes de magnitud mayor. El hecho de que D es grande lo hemos usado para deducir límites asintóticos del sistema (12)-(15).

El artículo [30] se ha dedicado a la investigación del sistema (12)-(15) en el límite asintótico $D \rightarrow \infty$. Nos hemos centrado en los casos de viscosidad del solvente muy pequeña ($\mu_s \rightarrow 0$) y viscosidad del solvente muy grande ($\mu_s \rightarrow \infty$) y hemos demostrado los siguientes resultados principales:

1) En el límite ($\mu_s \rightarrow \infty$), no hay ruptura de los tubos si $a > \frac{1}{2}$. Esto se demuestra haciendo uso de estimaciones clásicas de las ecuaciones parabólicas y el principio del máximo. Si $\frac{1}{4} < a \leq \frac{1}{2}$, entonces puede haber ruptura en forma autosimilar con exponentes dados por la raíz de una cierta ecuación integral. Si $a \leq \frac{1}{4}$, entonces la ruptura puede tener lugar en forma autosimilar con los mismos exponentes que en el caso newtoniano.

2) En el límite ($\mu_s \rightarrow 0$), se prueba la no existencia de soluciones autosimilares si $a > \frac{1}{2}$ y se hallan las posibles soluciones autosimilares en el caso $a \leq \frac{1}{2}$.

3) Se hallan ciertas soluciones autosimilares interesantes mediante el análisis de un mapa de fases: la distribución de gotas-en-alambre y las soluciones de tipo onda viajera (que modelizarían la migración de gotas sobre el filamento).

En [8] realizamos simulaciones numéricas y cálculos asintóticos formales para describir la formación y evolución de la estructura de gotas-en-alambre que aparece en los chorros viscoelásticos.

En el contexto de un puente liquido (ver [1]), tras un régimen transitorio en el que el puente se comporta como los puentes newtonianos, i.e. se estrecha a un ritmo que sigue la misma ley de potencias, entra abruptamente en un régimen puramente viscoelástico en el que se forma un filamento muy fino cuyo grosor decae exponencialmente a un ritmo

$$h_{\min} \sim Ae^{-\frac{t}{3D}}$$

siendo D el número de Deborah y A una constante. Este hecho podría constituir la base de un reómetro que sirva para medir D y con él el tiempo de relajación del polímero λ .

Hemos usado, como primera aproximación, el sistema unidimensional (12)-(15). Nuestro trabajo se ha dividido esencialmente en tres partes:

1.- Construcción de un algoritmo numérico eficiente capaz de resolver la evolución de (12)-(15) a lo largo de grandes periodos de tiempo. La evidencia experimental demuestra que la ruptura se ve inhibida por largos periodos de tiempo, después de los cuales tiene lugar la fenomenología de la que hemos hablado en los párrafos anteriores. Es por ello que nos ha interesado minimizar los errores de integración numérica que, como es sabido, tienden a propagarse y crecer con el tiempo. Un problema añadido está en el hecho de que estructuras casi-singulares (como aparentes ondas de choque) aparecen a lo largo de la evolución y hay que tratarlas de una forma numéricamente correcta. El

algoritmo que desarrollamos está basado en las diferencias finitas, es implícito y adaptativo (para resolver las casi-singularidades).

2.- Hemos entendido teóricamente, mediante el uso de desarrollos asintóticos acoplados (matched asymptotic expansions), la naturaleza de las estructuras casi-singulares que aparecen y que se traducen en elevados gradientes de h que, aparentemente, tienen la propiedad de propagarse cual onda de choque. Esto es importante desde el punto de vista fundamental (cosas así se han visto en experimentos). También desde el punto de la validez del modelo unidimensional, que podría perder precisión en cuando se forman tales estructuras.

3.- Finalmente, los resultados de nuestro análisis han sido contrastados con experimentos diseñados al efecto y el acuerdo con los mismos ha sido excelente.

En estos trabajos hemos estudiado el modelo de Oldroyd que asume implícitamente una posible extensibilidad infinita del polímero. Físicamente esta es una suposición razonable durante largos periodos de tiempo, pero siempre llega un momento en el cuál, debido al efecto de la anisotropía en el arrastre hidrodinámico experimentado por las moléculas o a su extensibilidad finita, el polímero alcanza un estado de máxima extensión más allá del cual ésta no es posible. Los dos modelos que tienen en cuenta estos efectos son el modelo de Giesekus:

$$\sigma_{zz} + D \left(\frac{\partial \sigma_{zz}}{\partial t} + v \frac{\partial \sigma_{zz}}{\partial z} - 2 \frac{\partial v}{\partial z} \sigma_{zz} \right) + \alpha \sigma_{zz}^2 = 2\mu_p \frac{\partial v}{\partial z}, \quad (16)$$

$$\sigma_{rr} + D \left(\frac{\partial \sigma_{rr}}{\partial t} + v \frac{\partial \sigma_{rr}}{\partial z} + \frac{\partial v}{\partial z} \sigma_{rr} \right) + \alpha \sigma_{rr}^2 = -\mu_p \frac{\partial v}{\partial z}, \quad (17)$$

donde $\mu_p = \nu_p \left(\frac{\rho}{\gamma R} \right)^{\frac{1}{2}}$ en que aparece el parámetro α , factor de movilidad (si $\alpha = 0$ se recupera el modelo Oldroyd-B) y el modelo FENE en el que

$$\sigma = \frac{\mu_p}{D} f(R)(\mathbf{A} - \mathbf{I}),$$

donde $R = \text{traza}(\mathbf{A})$, y

$$f(R) = \frac{1}{1 - R/L^2}. \quad (18)$$

con L siendo el factor de extensibilidad y A_{zz} y A_{rr} satisfaciendo

$$f(R)A_{zz} - 1 + D \left(\frac{\partial A_{zz}}{\partial t} + v \frac{\partial A_{zz}}{\partial z} - 2 \frac{\partial v}{\partial z} A_{zz} \right) = 0, \quad (19)$$

$$f(R)A_{rr} - 1 + D \left(\frac{\partial A_{rr}}{\partial t} + v \frac{\partial A_{rr}}{\partial z} + \frac{\partial v}{\partial z} A_{rr} \right) = 0. \quad (20)$$

Notar que, en el modelo FENE, si se hace $L = \infty$ recobramos el modelo Oldroyd-B. En [29] mostramos que estos modelos presentan ruptura del tubo en tiempo finito con escalas espaciales y temporales características análogas a las de los fluidos Newtonianos. Esto encaja con las observaciones experimentales en las que se constata como, después de largos tiempos tras la aparición de los filamentos

característicos de la estructura de "gotas-en-alambre", tiene lugar una ruptura puntual de los mismos.

Por último, en [42] hemos entendido teóricamente la génesis de los comportamientos que hemos bautizado como "elásticos". Hemos sido capaces de obtener soluciones explícitas que representen el colapso de gotas que migran en los filamentos una vez que se forma la estructura de "gotas-en-alambre", así como soluciones en las que unas gotas "absorben" el fluido de otras. Además presentamos en este artículo la explicación, que creemos definitiva desde que se planteó el problema en 1969, de estos fenómenos como el resultado de una interacción entre fuerzas elásticas y de tensión superficial. Esta interacción obedece lo que hemos bautizado como "4 times rule" y establece que las fuerzas elásticas en un filamento cuyo perfil varía lentamente son 4 veces mayores y de sentido opuesto que las de tensión superficial.

4 Análisis de la formación de singularidades en análogos de Navier-Stokes en dimensiones 1 y 2

Parte de la actividad desarrollada por el autor en los dos últimos años en colaboración con A. Córdoba y D. Córdoba trata del estudio de las soluciones de la ecuación cuasigeoestrófica bidimensional y de algunos análogos unidimensionales. La ecuación cuasigeoestrófica es a su vez un modelo para los sistemas de Euler y Navier-Stokes tal y como comentaremos más abajo. La ecuación se puede derivar a partir de principios provenientes de la geofísica y tiene la siguiente forma:

$$(\partial_t + \vec{v} \cdot \nabla) \theta = -k(-\Delta)^\gamma \theta \quad (21)$$

$$\vec{v} = \nabla^\perp \psi \quad y \quad \theta = -(-\Delta)^{\frac{1}{2}} \psi \quad (22)$$

donde θ representa el campo de temperaturas, \vec{v} el campo de velocidades y k el coeficiente de viscosidad. Tiene aplicaciones a la meteorología y a la oceanografía, y es un caso especial del sistema cuasigeoestrófico más general que modela el flujo en el océano cuando los números de Rossby y Ekman son pequeños. A pesar de estas aplicaciones potenciales, nosotros estamos más interesados en los aspectos teóricos y analíticos de la ecuación debido a las propiedades tan notables que presenta.

En los últimos años ha habido un intenso interés científico en la comprensión del comportamiento de las soluciones de la ecuación cuasigeoestrófica como un posible modelo para explicar la formación de frentes singulares que separan regiones con distintos valores de θ (que se pueden interpretar como regiones con masas de aire caliente y frío). Por otra parte, Constantin, Majda y Tabak [10] mostraron que en el caso $k = 0$ la ecuación cuasigeoestrófica es un modelo bidimensional para la ecuación de Euler tridimensional, ya que comparten muchas de las dificultades analíticas que han hecho imposible, hasta la fecha, comprender completamente el comportamiento de las soluciones de ambas. En particular, es esencial conocer si las soluciones desarrollan singularidades en

tiempo finito o no debido a las consecuencias directas que una situación u otra tendría en Física e Ingeniería. Recordemos que el sistema de Navier-Stokes se puede escribir en su forma más general como

$$\vec{v}_t + \vec{v} \cdot \nabla \vec{v} = -\nabla P - k(-\Delta)^\gamma \vec{v} \quad (23)$$

$$\nabla \cdot \vec{v} = 0 \quad (24)$$

donde $u = (v_1(x_1, x_2, x_3, t), v_2(x_1, x_2, x_3, t), v_3(x_1, x_2, x_3, t))$ es el campo de velocidades en el seno del fluido, la presión viene dada por $P = P(x_1, x_2, x_3, t)$ y k es el coeficiente de viscosidad. Desde el punto de vista físico es razonable asumir que el flujo tiene una energía cinética inicial finita:

$$\int_{\Omega} |\vec{v}(x, t)|^2 dx < \infty.$$

Ladyzhenskaya fue capaz de probar existencia global (i.e., ausencia de singularidades) para $\gamma \geq \frac{5}{4}$. El valor más destacable de γ , desde el punto de vista físico, es $\gamma = 1$.

El vector vorticidad se define como $\vec{\omega} = \nabla \times \vec{v}$. Las ecuaciones (23), (24) se pueden escribir en términos de ω como

$$\vec{\omega}_t + \vec{v} \cdot \nabla \vec{\omega} = (\nabla \vec{v}) \vec{\omega} - k(-\Delta)^\gamma \vec{\omega}$$

$$\nabla \cdot \vec{v} = \nabla \cdot \vec{\omega} = 0.$$

Usando la ley de Biot-Savart, la velocidad se puede obtener a partir de la vorticidad como

$$\vec{v}(x, t) = \frac{1}{4\pi} \int_{\Omega} \frac{y \times \vec{\omega}(x + y, t)}{|y|^3} dy.$$

Análogamente, la ecuación cuasigeoestrófica se puede escribir en la forma

$$(\partial_t + \vec{v} \cdot \nabla) \nabla^\perp \theta = (\nabla \vec{v}) \cdot \nabla^\perp \theta - k(-\Delta)^\gamma (\nabla^\perp \theta).$$

$$\nabla \cdot \vec{v} = 0$$

donde la función de corriente ψ se puede obtener de θ a partir de

$$\psi(x, t) = - \int_{\Omega} \frac{\theta(x + y, t)}{|y|} dy,$$

y por tanto

$$\vec{v}(x, t) = - \int_{\Omega} \frac{\nabla^\perp \theta(x + y, t)}{|y|} dy$$

La analogía entre $\nabla^\perp \theta = (-\theta_{x_2}, \theta_{x_1})$ y $\vec{\omega} = \nabla \times \vec{v}$ es clara ya que ambos satisfacen la misma ecuación. Por otra parte, el operador $\nabla \vec{v}$ es una integral singular en dos dimensiones con respecto a $\nabla^\perp \theta$ para la ecuación cuasigeoestrófica mientras que es un operador integral singular en tres dimensiones con respecto a la vorticidad en Navier-Stokes. En [48] se ha observado numéricamente que

en el caso $\gamma = 1$ hay solamente un crecimiento exponencial de las derivadas de θ , mientras que en [14] se demostró que en escenarios plausibles en los que se conjeturó la existencia de singularidades (puntos hiperbólicos) tales singularidades no pueden existir.

La ecuación cuasigeométrica no viscosa puede escribirse, en términos de la transformada de Riesz (ver [60]), como

$$\theta_t + \operatorname{div}[(R^\perp \theta)\theta] = 0.$$

Con objeto de construir un modelo unidimensional, en [7] consideraremos como incógnita una función $\theta(x, t)$ definida en $(x, t) \in \mathbb{R} \times \mathbb{R}_+$, y reemplazaremos la transformada de Riesz $R^\perp(\cdot)$ por su análogo unidimensional, la transformada de Hilbert. Finalmente, reemplazamos $\operatorname{div}(\cdot)$ por ∂_x . Entonces el problema resulta ser

$$\begin{aligned} \theta_t + (H(\theta)\theta)_x &= 0, \\ \theta(x, 0) &= \theta_0(x). \end{aligned} \tag{25}$$

Siguiendo [9] transformamos (25) introduciendo variables y funciones complejas y teniendo en cuenta las siguientes relaciones algebraicas para la transformada de Hilbert:

$$H(Hf) = -f \tag{26}$$

$$H(fHg + gHf) = (Hf)(Hg) - fg \tag{27}$$

$$(Hf)_x = H(f_x). \tag{28}$$

Aplicando H a la primera ecuación en (25) se tiene

$$(H\theta)_t + \frac{1}{2}((H\theta)^2 - (\theta)^2)_x = 0. \tag{29}$$

Entonces, introduciendo la función compleja

$$z(x, t) = H\theta(x, t) + i\theta(x, t), \quad z_0(x) = H\theta_0(x) + i\theta_0(x), \tag{30}$$

las ecuaciones (25) y (29) son respectivamente la parte imaginaria y real de la siguiente ecuación de Burgers compleja:

$$\begin{aligned} z_t + zz_x &= 0, \\ z(x, 0) &= z_0(x). \end{aligned} \tag{31}$$

Usando esta representación demostramos que las soluciones desarrollan genéricamente singularidades de tipo "cúspide" en tiempo finito.

En una dimensión el término de viscosidad en el caso crítico $-\kappa(-\Delta)^{\frac{1}{2}}\theta$ es equivalente a $-\kappa(H\theta)_x$. En este artículo probamos también que las soluciones de esta ecuación desarrollan singularidades con el mismo dato que el caso no viscoso si $\kappa < \|\theta_0\|_{L^\infty}$. Cuando $\kappa \geq \|\theta_0\|_{L^\infty}$ la solución es regular para todo tiempo.

Un modelo similar a (25) es el dado por la siguiente ecuación de tipo Burgers con velocidad nlocal:

$$\begin{aligned}\theta_t + H(\theta)\theta_x &= 0, \\ \theta(x, 0) &= \theta_0(x).\end{aligned}\tag{32}$$

Esta ecuación guarda relación con la ecuación de Birkhoff-Rott para la evolución de una hoja de vorticidad en dos dimensiones (ver [13]). En [13] demostramos que las soluciones de (32) pueden desarrollar singularidades en forma de cúspide en tiempo finito tal y como se conjeturó para (32) y para la propia ecuación de Birkhoff-Rott. Dichas singularidades se regularizan tras la adición de términos de viscosidad.

5 Bifurcaciones de ruptura de simetría en fluidos conductores cargados

Consideremos ahora una gota de fluido conductor con carga total Q , en un medio de constante dieléctrica ε_0 , y con radio R_0 . Por debajo de un cierto valor del coeficiente de tensión superficial γ la gota puede dejar de ser esférica. De hecho, la pérdida de esfericidad ocurre cuando un cierto parámetro adimensional

$$X \equiv \frac{Q^2}{32\gamma\pi^2\varepsilon_0 R_0^3}$$

es mayor que un cierto valor crítico X_c .

Este problema fue estudiado por primera vez por Lord Rayleigh (cf. [55]) quién mostró que una gota de fluido cargado se puede volver inestable cuando X es suficientemente grande. Trabajo numérico debido a Miksis [45] muestra que cuando X excede X_c las gotas se vuelven inestables tomando una forma aproximada de esferoide prolato para $X - X_c$ pequeño; ver también las referencias en [45]. La evolución posterior una vez la gota se ha vuelto inestable es un problema difícil de estudiar tanto numérica como experimentalmente. La primera observación experimental del proceso completo de evolución de una gota levitante (en la que se equilibra la gravedad con una fuerza de sustentación electrostática) cuyo radio decrece lentamente debido a la evaporación ha sido muy reciente (ver Duft et al. [17]). Una vez que el radio es suficientemente pequeño y se alcanza el valor crítico de X , la gota adquiere la forma de un esferoide prolato que, tras un breve periodo de tiempo, se vuelve muy excéntrico y emite dos chorros extremadamente finos por sus polos. Estos chorros llevan consigo aproximadamente un 30 % de la carga de la gota, pero solamente un 1 % de la masa. A continuación los chorros se desintegran en cientos de pequeñas gotas, la gota madre vuelve a evaporarse y el proceso se repite cuando su radio es suficientemente pequeño. El resultado final es la formación de multitud de pequeñas gotas. Se conjetura que podría ser este proceso de formación de cientos de gotitas en torno a las que pueden nuclear gotas mayores lo que desencadena las tormentas de lluvia.

Puesto que estamos considerando un fluido conductor, toda la carga eléctrica se reparte en la superficie de la gota. Fuera del dominio ocupado por la gota (que llamaremos Ω) habrá un campo eléctrico con un potencial V que decae en el infinito y es solución de la ecuación de Laplace. En la superficie de la gota, al ser conductora, habrá un potencial constante así como un equilibrio entre diferencia de presión (entre el interior y el exterior), fuerzas de tensión superficial proporcionales a la curvatura media y fuerzas de repulsión electrostática de cargas. Las fuerzas electrostáticas son proporcionales a la densidad superficial de carga y a la componente normal del campo eléctrico (que es igual a la derivada normal de V). Se puede demostrar fácilmente que en la superficie de un conductor las cargas se distribuyen proporcionalmente a la derivada normal del potencial. Poniendo todos estos hechos juntos se obtiene el sistema:

$$\Delta V = 0 \text{ en } \mathbb{R}^3 \setminus \Omega, \quad (33)$$

$$V = C \text{ sobre } \partial\Omega, \quad (34)$$

$$V(\mathbf{r}) \rightarrow 0 \text{ cuando } |\mathbf{r}| \rightarrow \infty, \quad (35)$$

$$\delta p = \gamma\kappa - \frac{\varepsilon_0}{2} \left(\frac{\partial V}{\partial n} \right)^2 \text{ sobre } \partial\Omega \quad (36)$$

donde δp es la diferencia de presión entre el fluido dentro de Ω y el fluido fuera de Ω , y κ es la curvatura media de la superficie $\partial\Omega$ ($\kappa > 0$ si Ω es una esfera). La constante C en (34) se tiene que elegir de tal manera que se satisfaga la condición (35).

El sistema (33)-(36) tiene una solución explícita en la que Ω es una esfera de radio R_0 ,

$$V(\mathbf{r}) = \frac{Q}{4\pi\varepsilon_0 r}, \quad (37)$$

$$\delta p = \frac{\gamma}{R_0} - \frac{\varepsilon_0}{2} \left(\frac{Q}{4\pi\varepsilon_0 R_0^2} \right)^2, \quad (38)$$

y

$$C = \frac{Q}{4\pi\varepsilon_0 R_0}. \quad (39)$$

Las soluciones no esféricas que se han observado experimentalmente podrían aparecer como ramas bifurcadas de la rama principal de soluciones esféricas. Eso es precisamente lo que demostramos en [24] (ver también [23] donde se estudian problemas similares que aparecen en el contexto de la biología matemática). Escogemos como parámetro de bifurcación el coeficiente de tensión superficial. Consideramos los parámetros Q , ε_0 y R_0 fijos y entonces la constante C en (39) es también fija y la diferencia de presiones δp se toma como la misma de una gota esférica, es decir,

$$\delta p = \delta p_0(\gamma) \equiv \frac{\gamma}{R_0} - \frac{\varepsilon_0}{2} \left(\frac{Q}{4\pi\varepsilon_0 R_0^2} \right)^2.$$

Probamos que existe una sucesión de ramas de bifurcación con

$$\gamma = \gamma_l + \varepsilon\gamma_{l1} + \varepsilon^2\gamma_{l2} + \dots \quad (l = 2, 3, \dots)$$

y frontera libre

$$r = R_0 + x(\theta) = R_0 + \varepsilon Y_{l,0}(\theta) + \varepsilon^2 \Lambda_{l2}(\theta) + \dots$$

donde $Y_{l,0}(\theta)$ son los armónicos esféricos $Y_{l,m}$ con $m = 0$, y $\gamma_2 < \gamma_3 < \dots$

Como el primer punto de bifurcación $\gamma = \gamma_2$ es el más relevante desde el punto de vista físico (es el genérico), es de interés conocer la forma de la curva de bifurcación $\gamma = \gamma(\varepsilon)$ cerca de $\gamma = \gamma_2, \varepsilon = 0$. Hemos calculado γ_{21} y concluído que

$$\gamma_{21} > 0 \tag{40}$$

de modo que para $l = 2, \gamma'(\varepsilon) > 0$ cerca de $\varepsilon = 0$.

Probamos además que las gotas esféricas son asintóticamente estables si $\gamma > \gamma_2$ mientras que en la rama bifurcada con $\gamma(\varepsilon) \sim \gamma_2$ y $\varepsilon > 0$ (formas prolatas) las gotas son inestables tal y como se muestra en la Figura 3 y son estables ante perturbaciones axisimétricas si $\varepsilon < 0$ (formas oblatas). Estos resultados han sido obtenidos usando la potente teoría de bifurcaciones en espacios de Banach de Crandall y Rabinowitz (ver [15] y [16]) adaptada por nosotros al estudio de problemas de frontera libre. Bajo esta perspectiva, las bifurcaciones de ruptura de simetría en gotas cargadas pueden clasificarse como bifurcaciones transcíticas.

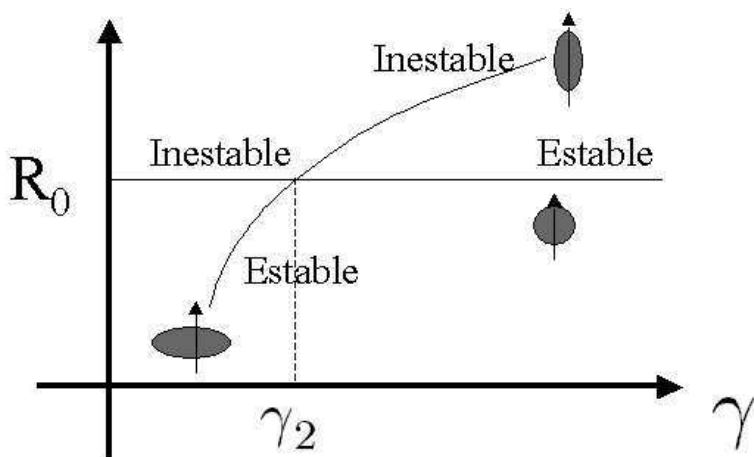


Figura 3: Representación gráfica de las ramas de bifurcación. La forma de la gota es axisimétrica en todos los casos y representamos el eje de simetría con una flecha.

Los problemas de frontera libre involucrando fluidos cargados y la acción de fuerzas de tensión superficial están comenzando a atraer una gran atención

por parte de amplias comunidades de físicos e ingenieros. La razón estriba en la posibilidad de control del comportamiento de los fluidos a muy pequeña escala mediante la acción de campos eléctricos y magnéticos y las potenciales aplicaciones que ello tendría en niveles micro y nanoscópicos. Nuestra contribución en este campo se limita, por el momento, al estudio del derrame de gotas conductoras cargadas (ver [5]). El desarrollo de herramientas numéricas apropiadas basadas en la formulación de las ecuaciones de fluidos de Stokes mediante integrales de contorno y su combinación con el uso de técnicas asintóticas será el arma básica con el que pretendemos atacar muchos de estos problemas en el futuro próximo.

Referencias

- [1] Y. Amarouchene, D. Bonn, J. Meunier, H. Kellay, Inhibition of the finite-time singularity during droplet fission of a polymeric fluid, *Physical Review Letters*, Vol. 86, no. 16 (2001), 3558-3561.
- [2] F. Bernis, A. Friedman, Higher order nonlinear degenerate parabolic equations, *J. Differ. Equations* 83, 1 (1990), 179-206.
- [3] A. L. Bertozzi, M. Pugh, The lubrication approximation for thin viscous films: Regularity and long-time behavior of weak solutions. *Commun. Pure Appl. Math.* 49, 2 (1996), 85-123.
- [4] S. Betelú, M. A. Fontelos, Capillarity driven spreading of power-law fluids, *Applied Mathematics Letters*, 16 (8) (2003), 1315-1320.
- [5] S. Betelú, M. A. Fontelos, Spreading of charged microdroplets, aparecerá en *Physica D*, 2005.
- [6] R. B. Bird, R. C. Armstrong, O. Hassager. *Dynamics of polymeric liquids*, Vol. 1. Wiley, New York, 1977.
- [7] D. Chae, A. Córdoba, D. Córdoba, M. A. Fontelos. Finite time singularities in a 1D model of the quasi-geostrophic equation, aparecerá en *Adv. Math.*, 2005.
- [8] C. Clasen, J. Eggers, M. A. Fontelos, J. Li, G. H. McKinley, The beads-on-string structure of viscoelastic threads, aparecerá en *J. of Fluid Mech.*, 2005.
- [9] P. Constantin, P.D. Lax, A. J. Majda, A simple one-dimensional model for the three-dimensional vorticity equation, *Comm. Pure Appl. Math* 38 (1985), 715-724.
- [10] P. Constantin, A. J. Majda, E. Tabak, Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar, *Nonlinearity*, 7 (1994), 1495-1533.

- [11] A. Córdoba, D. Córdoba, C. Fefferman, M. A. Fontelos, A geometrical constraint for capillary jet breakup, *Adv. Math.* 187 (2004), no. 1, 228–239.
- [12] A. Córdoba, D. Córdoba, C. Fefferman, M. A. Fontelos, Drops: the collapse of capillary jets, *Proc. Natl. Acad. Sci. USA* 99 (2002), no. 17, 11006–11007.
- [13] A. Córdoba, D. Córdoba, M. A. Fontelos, Formation of singularities for a transport equation with nonlocal velocity, *aparecerá en Ann. of Math.*, 2005.
- [14] D. Cordoba, Nonexistence of simple hyperbolic blow-up for the quasi-geostrophic equation, *Ann. of Math.*, 148 (1998), 1135–1152.
- [15] M. G. Crandall, L. H. Rabinowitz, Bifurcation from simple eigenvalues, *J. Functional Analysis*, 8 (1971), 321–340.
- [16] M. G. Crandall, L. H. Rabinowitz, Bifurcation, perturbation of simple eigenvalues and linearized stability, *Arch. Rat. Mech. Anal.*, 52 (1973), 161–180.
- [17] D. Duft, T. Achtzehn, R. Müller, B. A. Huber, T. Leisner, Rayleigh jets from levitated microdroplets, *Nature*, vol. 421, 9 Enero 2003, pg. 128.
- [18] J. Eggers, Universal pinching of 3D axisymmetric free-surface flow. *Phys. Rev. Lett.* 71 (1993), 3458.
- [19] J. Eggers, Nonlinear dynamics and breakup of free surface flows, *Rev. Modern Phys.*, 69 (1997), 865–929.
- [20] M. A. Fontelos, A. Friedman, Stationary non-Newtonian fluid flows in channel-like and pipe-like domains, *Arch. Ration. Mech. Anal.* 151 (2000), no. 1, 1–43.
- [21] M. A. Fontelos, A. Friedman, Analysis of the stick-slip problem for non-Newtonian flows, *Comm. Partial Differential Equations* 26 (2001), no. 3–4, 461–536.
- [22] M. A. Fontelos, A. Friedman, The flow of a class of Oldroyd fluids around a re-entrant corner, *J. Non-Newtonian Fluid Mech.* 95 (2000), no.2–3, 185–198.
- [23] M. A. Fontelos, A. Friedman, Symmetry-breaking bifurcations of free boundary problems in three dimensions, *Asymptotic Analysis* 35 (3–4) (2003), 187–206.
- [24] M. A. Fontelos, A. Friedman, Symmetry-breaking bifurcations of charged drops, *Arch. Ration. Mech. Anal.*, 172 (2004) , no. 2, 267–294.

- [25] M. A. Fontelos, A. Friedman, B. Hu, Mathematical analysis of a model for the initiation of angiogenesis, *SIAM J. Math. Anal.* 33 (2002), no. 6, 1330–1356.
- [26] M. A. Fontelos, J. J. L. Velázquez, A free boundary problem for the Stokes system with contact lines, *Comm. Partial Differential Equations* 23 (1998), no. 7-8, 1209–1303.
- [27] M. A. Fontelos, J. J. L. Velázquez, On some breakup and singularity formation mechanisms for inviscid liquid jets, *SIAM J. Appl. Math.*, 59-6 (1999), 2274-2300.
- [28] M. A. Fontelos, J. J. L. Velázquez, Fractal-like singularities for an inviscid one-dimensional model of fluid jets, *Euro Jnl. of Applied Mathematics*, 11 (2000), 29-60.
- [29] M. A. Fontelos, J. Li, On the evolution and rupture of filaments in Giesekus and FENE models, *J. Non-Newtonian Fluid Mech.*, 118 (2004), no. 1, 1-16.
- [30] M. A. Fontelos, Break-up and no break-up in some models for the evolution of viscoelastic jets, *Z. Angew. Math. Phys.* 54 (2003), 84-111.
- [31] M. A. Fontelos, Analysis of the evolution of thin viscous jets: filament formation, *Mathematical Methods in the Applied Science*, 27 (2004), no. 10, 1197-1220.
- [32] A. Friedman, J. J. L. Velázquez, The analysis of coating flows near the contact line, *Journal of Differential Equations* 119, 1 (1995), 137-208.
- [33] A. Friedman, J. J. L. Velázquez, The analysis of coating flows in a strip, *Transactions of the AMS* 349, 8 (1995), 134-182.
- [34] M. Goldin, J. Yerushalmi, R. Pfeffer, R. Shinnar, Breakup of a laminar capillary jet of a viscoelastic fluid, *J. Fluid Mech.* 38-4 (1969), 689-711.
- [35] D. S. Kalika, M. M. Denn, Wall slip and extrudate distortion in linear low-density polyethylene, *J. Rheology* 31 (1997), 815-834.
- [36] T.A. Kowalewski, On the separation of droplets from a liquid jet, *Fluid Dyn. Res.* 17 (1996), 121.
- [37] V. A. Kozlov, V. G. Maz'ya, J. Rossmann, Elliptic boundary problems in domains with point singularities, *Mathematical surveys and monographs*, Vol. 52, American Mathematical Society, Providence R. I. (1997).
- [38] O.A. Ladyzhenskaya, *The mathematical theory of viscous Incompressible Flow*, 2nd ed., Gordon and Breach New York-London-Paris, 1969.
- [39] R. G. Larson, Instabilities in viscoelastic flows, *Rheologica Acta* 31 (1992), 213-263.

- [40] J. Leray, Etude de diverses équations nonlinéaires et de quelques problèmes que pose l'hydrodynamique, Jour. Math. Pures Appl., 12 (1933), 1-82.
- [41] J. Leray, Sur le mouvement d'un fluide visqueux emplissant l'espace, Acta Math. 63 (1934), 193-248.
- [42] J. Li, M. A. Fontelos, Drop dynamics on the beads-on-string structure for viscoelastic jets: a numerical study, Physics of fluids, 15 (2003), no. 4, 922-937.
- [43] Lu Ting, J.B. Keller. Slender jets and thin sheets with surface tension. SIAM J. Appl. Math., 50,6 (1990), 1533-1546.
- [44] A. Majda, E. Tabak, A two-dimensional model for quasi-geostrophic flow: comparison with the two-dimensional Euler flow, Physica D , 98 (1996), 515-522.
- [45] M. J. Miksis, Shape of a drop in an electric field, Phys. of Fluids, 24-11 (1981), 1967-1972.
- [46] T. G. Myers, Thin films with high surface tension, SIAM Rev. 40, 3 (1998), 441-462.
- [47] C. C. Ngan, E. B. Dussan, On the dynamics of liquid spreading on solid surfaces, J. Fluid Mech. 209 (1989), 191-226.
- [48] K. Ohkitani, M. Yamada, Inviscid and inviscid-limit behavior of a surface quasi-geostrophic flow, Phys. Fluids, 9 (1997), 876-882.
- [49] F. Otto, Lubrication approximation with prescribed nonzero contact angle, Comm. Partial Differential Equations 23 (1998), no. 11-12, 2077-2164.
- [50] D. T. Papageorgiou, On the breakup of viscous liquid jets, Physics of fluids, 7, (1995), 1529-1544.
- [51] J. Pedlosky, Geophysical fluid dynamics, Springer-Verlag New york, 1987.
- [52] B. A. Plamenevskij, Elliptic boundary value problems in domains with piecewise smooth boundary, in Encyclopedia of Mathematical Sciences, Vol. 79, Springer-Verlag Berlin (1997).
- [53] M. T. Plateau, Smithsonian Report 250 (1863).
- [54] Rayleigh, Lord (J.W. Strutt), On the instability of jets. Proc. Lond. Math. Soc. 10 (1879), 4-13.
- [55] Rayleigh, Lord (J.W. Strutt), On the equilibrium of liquid conducting masses charged with electricity, Phil. Mag. 14 (1882), 184-186.
- [56] M. Renardy, A numerical study of the asymptotic evolution and breakup of Newtonian and viscoelastic jets. J. Non-Newtonian Fluid Mechanics, 59 (1995), 267-282.

- [57] F. Savart, Mémoire sur la Constitution des veines liquides lancées par des orifices circulaires en mince paroi. *Ann. Chim. Phys.* 53, 337 (1833), 337-386.
- [58] X. D. Shi, M. P. Brenner, S. R. Nagel, A cascade of Structure in a drop falling from a faucet. *Science*, Vol. 265 (1994), 219-222.
- [59] V. A. Solonnikov, On some free boundary problems for the Navier-Stokes equations with moving contact points and lines, *Mathematische Annalen* 302, 4 (1995), 743-772.
- [60] E. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press. Princeton, NJ, 1970.
- [61] E. D. Wilkes, S. D. Phillips and O. A. Basaran, Computational and experimental analysis of the dynamics of drop formation. *Physics of Fluids*, Vol. 11, no. 12 (1999), 3577-3598.

Modelado matemático en Biología: redes genéticas y relojes biológicos

J.D. SALAZAR

salazar@maths.warwick.ac.uk

La aplicación de las Matemáticas a la Biología tiene ya una larga historia. Algunas de las áreas más conocidas incluyen (véase [1] y [2]): la dinámica de poblaciones, la teoría de epidemias, la genética clásica, los modelos evolutivos, morfogénesis (tanto el mecanismo de Turing como los modelos mecánicos), los modelos de propagación de impulsos nerviosos, los modelos neuronales y la aplicación de la mecánica de medios continuos a los fluidos y membranas biológicos. A estas áreas más estudiadas se ha venido a sumar recientemente el modelado de las redes genéticas: cascadas bioquímicas que regulan las funciones celulares en los seres vivos.

Redes Genéticas

Para entender lo que es una red genética hay que empezar mencionando el llamado *Dogma Fundamental de la Biología Molecular*: la prescripción de cómo el flujo de información ocurre en una célula. La información necesaria para realizar todas las labores celulares se encuentra almacenada en el ADN (ácido desoxirribonucleico), en donde se encuentra dividida en capítulos llamados genes. Cada gen es leído cuando le corresponde y transcrito en una cadena relativamente pequeña de ARN (ácido ribonucleico). El ARN es una molécula muy similar al ADN que realiza varias funciones en la célula, incluyendo el transporte de información. La cadena de ARN transcrito es procesada posteriormente y usada para crear la proteína correspondiente. Las proteínas son los auténticos obreros de la célula; realizan casi todas las funciones celulares. Podemos entonces resumir el dogma fundamental de la Biología Molecular en el siguiente esquema:



Las cosas sin embargo no son nunca tan claras y simples en Biología. Las proteínas son sustancias químicas extremadamente versátiles y además de participar en muchas reacciones, también son capaces de actuar como agentes activadores o represores de la transcripción de genes. Con lo cual es

perfectamente posible que la transcripción de un gen de lugar a la activación o inhibición de otros genes, o incluso de sí mismo. Éste es el mecanismo de las redes genéticas, que resultan en complicados diagramas de activación y represión de genes y que frecuentemente incluyen bucles (véase la figura 2).

Debido al grado de complejidad de estas redes genéticas y al hecho de que las distintas reacciones involucradas ocurren en distintos momentos y lugares, no es sorprendente la necesidad de usar matemáticas avanzadas para obtener una buena comprensión de su funcionamiento.

Afortunadamente, en los últimos años se han desarrollado técnicas experimentales capaces de obtener mediciones múltiples del estado de cientos (y a veces, miles) de genes en una célula dada. Otros avances experimentales han permitido asociar un marcador visible a la expresión de un gen, lo que permite estudiar la evolución temporal de parte de una red genética sin tener que realizar mediciones destructivas. Finalmente, es posible hoy en día alterar el ADN de una bacteria, animal o planta y producir mutantes artificiales, que cuando son viables permiten comprender lo que ocurre en la ausencia o sobreabundancia de un proteína particular.

Estos avances recientes han revolucionado la Biología Molecular y han llevado a una acumulación de datos y conocimientos parciales sobre el funcionamiento de las redes genéticas. Pero la utilización que la células hacen de estas redes depende de su dinámica global (cómo vamos a ver después) y esto convierte al modelado matemático en una herramienta imprescindible.

Para una discusión más detallada, nos vamos a centrar en uno de los tipos de redes genéticas más curiosos e interesantes desde el punto de vista matemático: los relojes biológicos.

Relojes Biológicos

La rotación del planeta Tierra alrededor del Sol y sobre sí mismo introduce dos ritmos muy importantes en la actividad de casi todas las criaturas vivas: el año y el día. Es obvia la influencia que el paso de las estaciones tiene en los animales y las plantas. También son obvios los ritmos de sueño y actividad de los animales, sincronizados con los periodos de luz y oscuridad. Por cierto que el efecto del desarrollo tecnológico e industrial tiene en los ritmos de sueño de los humanos da lugar a algunos ejemplos espectaculares de lo que sucede cuando se perturban estos ritmos: el famoso “jet lag” que ocurre cuando viajamos por avión y cambiamos a una zona horaria diferente y el incremento en el número de accidentes de tráfico y laborales que ocurre en ciertas horas debido a la conducción nocturna o al trabajo por turnos.

Sin embargo, estas observaciones no prueban la existencia de una reloj interno, sino solo la sincronización de las actividades de los seres vivos con los ritmos exteriores. Aparentemente, la primera persona en darse cuenta de la existencia de estos relojes internos fue el astrónomo francés Jean Jacques d’Ortous de Mairan en 1729. Intrigado por la observación de que ciertas plantas abren sus hojas durante el día y las cierran durante la noche, llevó un ejemplar a su sótano y observó que la planta continuaba con este ritmo en la oscuridad

permanente.

En los años 1950, Gustav Kramer y Klaus Hoffmann, mediante un aparato de su diseño, observaron que los pájaros migratorios usan el Sol como referencia, aunque el Sol se mueve durante el día. De este hecho dedujeron que el reloj interno del animal orienta su brújula para compensar por el movimiento del Sol. Además, demostraron que estos pájaros eran capaces de ajustar sus relojes a la hora local.

Otros experimentos realizados en la misma década por Colin Pittendrigh usando moscas, pusieron de relieve que el momento de eclosión de la larva está controlado por un reloj interno que se ajusta a los ciclos de luz y oscuridad. Muy interesantemente, si se daban a las larvas pulsos de luz u oscuridad en determinados momentos del ciclo, se podía alterar la fase de eclosión. Esto llevó a la creación del concepto de *curvas de repuestas de fase*, que representan el número de horas que la fase de una actividad biológica determinada se adelanta o retrasa cuando se le da a la criatura un pulso de luz o de oscuridad en un momento concreto. Desde entonces, estas curvas han sido calculadas en muchas otras especies.

En aquellos años se establecieron otros dos hechos muy interesantes:

- El periodo de oscilación libre suele ser cercano a veinticuatro horas pero nunca es exactamente veinticuatro horas: puede variar de veintiuna horas en el hongo *Neurospora crassa* hasta más de veinticinco horas en la planta *Arabidopsis thaliana*. Debido a este hecho, se decidió denominar a estas oscilaciones con el nombre de *ritmos circadianos* (“casi un día”).
- Cuando uno de estos organismos se encuentra en la oscuridad o en iluminación continua, sus oscilaciones pueden ser sincronizadas con pulsos de temperatura. Sin embargo, el periodo de oscilación libre es independiente de cambios sostenidos de temperatura, como las diferencias entre verano e invierno. En otras palabras, estas oscilaciones son a la vez sensibles a pulsos de temperatura pero insensibles a cambios sostenidos de temperatura. Estas propiedades aparentemente contradictorias presentan un desafío a la teoría de osciladores, que se ha desarrollado principalmente estudiando ejemplos físicos.

En términos de las curvas de respuesta de fase se puede dar una formulación matemática del concepto de sincronización: sea $\Phi(\phi)$ el cambio de fase debido a un pulso de luz aplicado cuando el reloj se encontraba en la fase ϕ . Si el pulso de luz se aplica el día n ésimo, y la fase del reloj al alba es ϕ_n , la fase del reloj al alba del día siguiente va a ser

$$\phi_{n+1} = F(\phi_n) = \phi_n + \Phi(\phi_n) + T - \tau,$$

donde T es la duración del día y τ es el periodo libre del reloj. Si F posee un punto fijo ϕ^* , que es estable y alcanzable desde una fase de partida ϕ_0 , entonces será posible sincronizar el reloj con una cierta duración del día T . En la fase de sincronización tenemos que,

$$\Phi(\phi^*) = \tau - T.$$

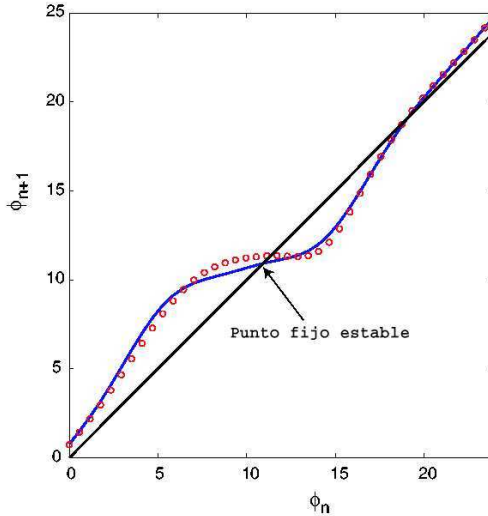


Figure 1: Curva de respuesta de fase correspondiente al diagrama de la figura 2, cuando T es 9 horas. La línea de puntos corresponde a la curva de respuesta de fase calculada directamente del modelo y la línea sólida corresponde a nuestra CRI.

Dado que $|F'(\phi^*)| < 1$ para que haya estabilidad, podemos concluir que

$$-2 < \Phi'(\phi^*) < 0.$$

Finalmente, obsérvese que para que el reloj sea fácilmente sincronizable, la curva de respuesta de fase $\Phi(\phi)$ debe tener una amplitud apreciable (véase la figura 1).

Desde el descubrimiento de la estructura del ADN en 1953, se propuso primero y se ha comprobado después que la naturaleza de los relojes biológicos es bioquímica y celular: estos osciladores producen variaciones periódicas en el nivel de las concentraciones de ciertas proteínas, que tienen efectos observables en los animales y plantas. Estas variaciones son afectadas por factores como el número de horas de luz o la temperatura, lo que implica que hay canales sensoriales que perciben estas señales. Sin embargo, las oscilaciones continúan en ausencia de señales exteriores.

Esta naturaleza bioquímica hace aún más sorprendente la propiedad de la compensación por temperatura. Dado que la velocidad de una reacción química tiende a duplicarse cuando la temperatura ambiental sube unos diez grados centígrados, las redes genéticas que implementan los relojes biológicos deben estar diseñadas de tal manera que compensen por incrementos sostenidos de temperatura de manera que el periodo de oscilación libre no varíe.

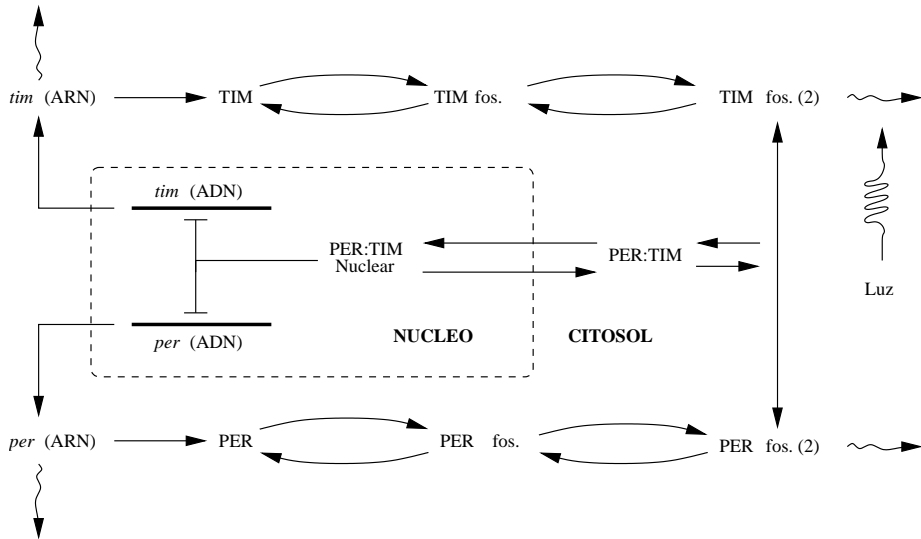


Figure 2: Red genética que implementa el reloj circadiano en la mosca de la fruta, *Drosophila melanogaster*. PER y TIM son los dos genes involucrados en esta red; la unión de sus dos proteínas en el núcleo inhibe su propia transcripción. La abreviación “fos.” se refiere a fosforilación. Las flechas ondulantes representan degradación. Obsérvese como la luz contribuye a la degradación de la proteína TIM una vez que ha sido fosforilada dos veces.

Ciclos Límite

Aún antes de desentrañar los detalles genéticos involucrados en los ritmos circadianos, se propuso modelar estas oscilaciones usando el concepto de ciclo límite. Esta idea no sólo permitió explicar la robustez de las oscilaciones sino además el mecanismo de sincronización a través de las curvas de respuesta de fase (véase [2] y [3]).

En un claro ejemplo de evolución convergente, todos los relojes biológicos conocidos están basados en bucles de transcripción y traducción: un gen particular es transcrito en una molécula de ARN mensajero que es posteriormente traducida en la proteína correspondiente. Esta proteína puede sufrir una serie de transformaciones moleculares (fosforilaciones por ejemplo), formar complejos con otras proteínas, degradarse y desaparecer o, muy interesantemente, ser afectada por condiciones ambientales: niveles de luz o de temperatura. Al final de la cadena de reacciones, emerge un factor de transcripción, esto es una proteína capaz de activar o inhibir la transcripción genética. Y este factor de transcripción es el que cierra el bucle o bucles existentes (véase figura 2).

Esta estructura de retroalimentación, con bucles positivos y negativos, produce una dinámica de un tipo muy diferente a los conocidos relojes químicos,

cuyo ejemplo más famoso es el de la reacción de Belousov-Zhabotinsky.

Los relojes biológicos controlan ritmos celulares. Espectacularmente, se han observado en bacterias con una duración de su vida menor de veinticuatro horas. En otras palabras, las bacterias pueden transmitir la fase del reloj a sus descendientes. En plantas parece haberse demostrado que hay poca comunicación entre los relojes presentes en cada célula. Por contra, en insectos y a aún más claramente en mamíferos, parece haber un oscilador central formado por un grupo de neuronas en el cerebro que recibe la señal luminosa de los ojos y que envía su señal horaria a los distintos órganos. Éstos poseen un reloj individual que se sincroniza con el reloj central.

El modelado matemático comienza con la elección de un tipo de descripción: estocástica o determinista. Cada opción tiene sus ventajas e inconvenientes, pero me voy a centrar en la descripción determinista. Ésta se basa en la aplicación de la Ley de Acción de Masa: la velocidad de una reacción química es directamente proporcional al producto de las concentraciones de los reactantes, elevadas al grado en que intervienen en la reacción. Sin embargo, dado que las reacciones involucradas suelen ser catalizadas por enzimas muy potentes y el hecho de que a menudo no se conocen todas las sustancias que participan en una reacción (ni muchos otros detalles de la misma), se suele optar por aproximaciones del tipo de Michaelis-Menten: la evolución del sustrato de la reacción enzimática sigue un perfil logístico, con un cierto nivel de saturación (véase [2]). La combinación de los términos matemáticos correspondientes a las diferentes reacciones da lugar a sistemas de ecuaciones diferenciales (ordinarias, y a veces, con retardo) no lineales muy complejos. Por ejemplo, el modelo de Leloup et al. [4] correspondiente a la figura 2 está formado por diez ecuaciones diferenciales no lineales con treinta y ocho parámetros.

En estos sistemas se suele desconocer a priori el valor de los parámetros. Al mismo tiempo, se dispone de muchas series temporales correspondientes a diferentes cadenas de ARN y proteínas, y otros datos sobre la dinámica del sistema. Se convierte entonces en un desafío matemático encontrar la manera computacionalmente más eficiente de calcular los valores óptimos de los parámetros. Como ejemplo diré que se pueden tardar varios días de cómputo en un PC para ajustar los valores de los parámetros en algunos modelos no demasiado complicados.

Complejidad y Flexibilidad

Diagramas como el de la figura 2 pueden parecer muy complejos, pero en realidad representan una simplificación considerable de la realidad biológica: casi cada mes se descubren nuevos genes que participan en la generación del ritmo circadiano en cada organismo modelo, o nuevas interacciones o pasos intermedios. Modelos recientes del reloj biológico de los mamíferos contienen hasta un centenar de ecuaciones diferenciales [5].

Pero, ¿por qué tanta complejidad? Se puede demostrar matemáticamente que un solo bucle de retroalimentación negativa bastaría para implementar un reloj eficiente. La explicación más recurrida apela a la necesidad de robustez en

el sistema: los ambientes bioquímicos en que los relojes actúan sufren cambios de todo tipo continuamente.

Pero aún así parece exagerada la complejidad observada. Esto nos ha llevado a mis colaboradores y a mi (véase [6]) a proponer que esta complejidad se debe a que los relojes deben ajustar evolutivamente muchas de sus propiedades: tanto la necesidad de compensar por cambios de temperatura sostenidos como el poder anticipar el momento del alba y del crepúsculo independientemente, al mismo tiempo que mantienen diferencias de fase entre los distintos componentes del reloj; la compensación por cambios de pH en el medio y otros factores químicos que influyen en las reacciones; etc ...

Para fundamentar esta idea, introducimos el concepto de *curvas de respuesta infinitesimales*: si el vector $k = (k_1, \dots, k_s)$ de parámetros de un modelo de un reloj circadiano es perturbado infinitesimalmente solamente cuando el reloj se encuentra entre las fases ϕ_1 y ϕ_2 , entonces cualquier característica del reloj Q_j se verá afectada de la forma:

$$\delta Q_j = \sum_{i=1}^s \delta k_i \left(\int_{\phi_1}^{\phi_2} f_{k_i, Q_j} d\phi \right).$$

Esta relación se puede establecer cualquiera que sean las fase ϕ_1 y ϕ_2 y consecuentemente define únicamente a la función f_{k_i, Q_j} , a la que llamamos curva de respuesta infinitesimal (CRI) de la característica Q_j debida al parámetro k_i .

Cuando Q_j es el periodo del ciclo límite,

$$f(\phi) = -\delta k_i \int_{\phi_1}^{\phi_2} f_{k_i, Q_j} d\phi$$

es la curva de respuesta de fase debida a una perturbación infinitesimal del parámetro k_i ocurrida entre las fases ϕ_1 y ϕ_2 . Si esa perturbación fue debida a un pulso luminoso, entonces $f(\phi)$ es la curva de respuesta de fase normal. En los modelos que estudiamos, las CRI nos dieron muy buenas aproximaciones a las curvas de respuesta de fase (véase la figura 1).

Como mencionados antes, las curvas de respuesta de fase (y consecuentemente las CRI) deben de tener una amplitud apreciable para que el reloj sea sincronizable. Si la CRI corresponde a cambios de temperatura, entonces debe de tener otra propiedad importante: su integral a lo largo de un ciclo completo tiene que ser cero para que el reloj esté compensado con respecto a cambios sostenidos de temperatura. Así las CRI nos dan una manera de caracterizar matemáticamente esta propiedad sorprendente.

Utilizando la naturaleza de ciclo límite de estas oscilaciones y la teoría clásica de dependencia con respecto de los parámetros pudimos calcular las CRI de una manera numéricamente efectiva (véase la información suplementaria en [6]). Estas ideas fueron codificadas en un programa que corre sobre Matlab (un cierto ambiente de computación matemática) y que puede ser obtenido en el ancla <http://www.maths.warwick.ac.uk/ipcr/> pulsando sobre "IRC Analysis - Software Download".

Si ahora consideramos un vector de características del reloj $Q = (Q_j)$ y sus perturbaciones posibles $\delta Q = (\delta Q_j)$, podemos escribir —dentro de los límites de nuestra teoría lineal— que la relación entre δQ y δk va a venir dada por una matriz $M = (M_{ij})$ tal que

$$\delta Q_j = \sum_i M_{ij} \delta k_i.$$

Entonces definimos $R_{\delta Q}$ como la razón entre la longitud del vector de cambios proporcionales $(\delta Q_j/Q_j)$ y el vector $(\delta k_i/k_i)$. Dado que δQ y δk están relacionados linealmente, $R_{\delta Q}$ no depende del tamaño de δQ sino solo de su dirección.

Dado un número pequeño ϵ , decimos entonces que la dirección δQ es *accesible* si $R_{\delta Q} > \epsilon$. En otras palabras, si $\epsilon = 0.01$, una variación δQ es inaccesible si requiere de un cambio porcentual en los parámetros δk mayor de cien veces el cambio porcentual en δQ . Se puede demostrar que el conjunto de vectores accesibles forman un cono con una cierta dimensión d .

Todas las características del reloj dependen de las propiedades de su ciclo límite y de su periodo libre. Así que resulta natural escoger como δQ el vector $(\delta\gamma, \delta p)$, donde γ es el ciclo límite (o una discretización adecuada del mismo) y p es el periodo libre. A la dimensión d del cono de vectores accesibles en este caso la llamamos *dimensión de flexibilidad* del reloj.

Esta dimensión se puede calcular de varias maneras (la más efectiva quizás consiste en considerar cambios aleatorios de los parámetros y calcular la descomposición en valores singulares de la matriz de las varianzas). En cada uno de los modelos de relojes publicados hasta la fecha, su dimensión de flexibilidad resultó ser un orden de magnitud inferior a su número de parámetros y creció monótonamente con el número de bucles del modelo (véase tabla 1 en [6]). Por ejemplo, para el modelo representado por la figura 2, el número de parámetros es 38 y $d = 2$ si $\epsilon = 0.05$.

Consecuentemente, parece que hacen falta varios bucles acoplados débilmente para que un reloj circadiano disfrute de suficiente flexibilidad para adaptarse a los cambios en su medio natural. Este hecho se puede motivar considerando los multiplicadores de Floquet de estos sistemas, pero hasta ahora no hemos podido construir una demostración matemática.

Incluso aceptando que esta teoría que hemos propuesto sea la correcta, quedan muchas preguntas en el aire: ¿por qué el periodo libre no es exactamente veinticuatro horas?; ¿cómo se alcanza exactamente el objetivo de la compensación por cambios sostenidos de temperatura?; dado que estos relojes controlan muchas funciones diferentes en el organismo y consecuentemente, están integrados en redes genéticas aún mayores, ¿hasta que punto existe modularidad en el diseño?; ¿cual es la naturaleza de la comunicación celular?.

La resolución de estas preguntas requerirá sin duda de la continuación de la colaboración entre biólogos, químicos, físicos y matemáticos. Esta colaboración tiene que ser directa y estrecha, debido a la gran complejidad de las técnicas experimentales y de las redes genéticas mismas. Inspirarse en la Biología para

desarrollar nuevas ideas matemáticas es fascinante (un ejemplo de Biología aplicada a las Matemáticas) pero si se desea que lo que se descubra sea relevante en un área biológica particular, es imprescindible tener un conocimiento de primera mano de esa área y poder intercambiar ideas continuamente con los especialistas. O esta es al menos la lección que yo he aprendido realizando este trabajo.

References

- [1] J. Keener and J. Sneyd. *Mathematical physiology*. Springer-Verlag, 1998.
- [2] J.D. Murray. *Mathematical biology*. Springer-Verlag, 2002.
- [3] A.T. Winfree. *The geometry of biological time*. Springer-Verlag, 1980.
- [4] J.C. Leloup, D. Gonze and A. Goldbeter. Limit cycle models for circadian rhythms based on transcriptional regulation in *Drosophila* and *Neurospora*. *Journal of biological rhythms*, 14(6):433-448, 1999.
- [5] D.B. Forger and C.S. Peskin. A detailed predictive model of the mammalian circadian clock. *PNAS*, 100(25):14806-14811, 2003.
- [6] D.A. Rand, B.V. Shulgin, J.D. Salazar and J. Millar. Design principles underlying circadian clocks. Aceptado en *Interfaces*, 2004.

Título:	MÉTODOS NUMÉRICOS PARA EL ESTUDIO DE LA DISPERSIÓN DE ONDAS TÉRMICAS.
Doctorando:	M ^a Luisa Rapún Banzo.
Director/es:	Francisco Javier Sayas González.
Defensa:	29 de junio de 2004, Zaragoza.
Calificación:	Sobresaliente cum laude.

Resumen:

El estudio de soluciones armónicas en tiempo de la ecuación del calor surge ligado a varios problemas de física e ingeniería, como son los ensayos no destructivos en materiales con incrustaciones mediante técnicas de fototerapia. Los problemas matemáticos resultantes están en el terreno de las ecuaciones de tipo Helmholtz en dominios no acotados con distintas condiciones de transmisión entre las zonas donde el material cambia de propiedades. En esta tesis doctoral se estudian analítica y numéricamente problemas de transmisión Helmholtz en un semiplano con un número finito de obstáculos internos donde los coeficientes de la ecuación son diferentes.

Para la resolución del problema bajo la hipótesis de coeficientes constantes, se proponen tres formulaciones integrales distintas mostrando la existencia y unicidad de solución y propiedades de regularidad de la misma en dos situaciones diferenciadas: problemas en los que las fronteras de los obstáculos son curvas regulares y para dominios Lipschitz. Se analizan en un nivel abstracto condiciones necesarias y suficientes de convergencia de métodos numéricos de tipo Petrov–Galerkin para tres tipos de ecuaciones muy distintos asociados a cada una de las formulaciones (un sistema elíptico, otro con estructura triangular y uno de estructura mixta). El análisis se lleva a cabo tanto en el contexto regular como en el lipschitziano. Se proponen discretizaciones con polinomios trigonométricos, con funciones spline y discretizaciones no conformes basadas en espacios de deltas de Dirac, dando en todos los casos las cotas de convergencia resultantes. También se proponen versiones completamente discretas de los mismos que preservan el orden y propiedades de convergencia puntual. Se ilustra con algunos ensayos numéricos el comportamiento los métodos propuestos.

En el caso de materiales con obstáculos interiores no homogéneos, se aborda el problema interior desde su formulación variacional y utilizando técnicas de métodos de contorno para el problema exterior. Para la resolución numérica del sistema de ecuaciones resultante se plantea un acoplamiento de elementos de contorno (escogiendo una aproximación conforme de las incógnitas de la frontera mediante splines periódicos) y elementos finitos (con constantes a trozos y elementos de Raviart–Thomas de orden cero definidos sobre una triangulación del dominio acotado). Asimismo, se plantea una discretización completa del método anterior que preserva sus propiedades de convergencia y se muestran algunos ensayos numéricos.

Título:	ANÁLISIS DE ALGUNOS MODELOS DE DINÁMICA DE POBLACIONES ESTRUCTURADOS EN EDAD CON Y SIN DIFUSIÓN.
Doctorando:	Mónica Molina Becerra.
Director/es:	Manuel Delgado Delgado y Antonio Suárez Fernández.
Defensa:	2 de Noviembre de 2004, Universidad de Sevilla .
Calificación:	Sobresaliente Cum Laude por unanimidad.

Resumen: Recientemente se han desarrollado modelos de dinámica de poblaciones en los que se tiene en cuenta que las tasas de natalidad y mortalidad dependen de la edad del individuo (modelos estructurados en edad) y además, el crecimiento de la población está limitado debido a la competición por los recursos (modelos no lineales). En este trabajo se analizan dos tipos de cuestiones relativas a ellos.

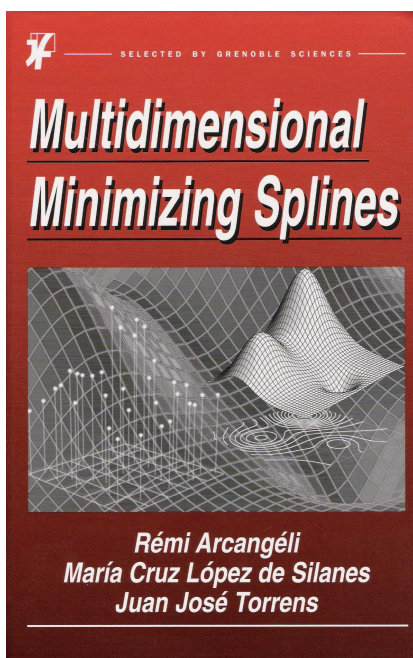
En una primera parte, analizamos un modelo presa-depredador (sin difusión), donde se supone que el comportamiento de la presa depende de la edad que tenga cada individuo y le afecta una enfermedad contagiosa, la cual es grave, en el sentido que puede causar mortalidad pero también es curable, sin causar inmunidad. Y el depredador sigue un crecimiento de tipo logístico (en ausencia de la presa), no está estructurado en edad y es inmune a la enfermedad. Primero damos resultados de existencia y unicidad de solución. A continuación, estudiamos el comportamiento asintótico de los equilibrios libres de enfermedad (aquellos donde los individuos infectados desaparecen). Y para finalizar con la primera parte, analizamos el comportamiento global del modelo con datos no dependientes en edad; comprobando que para algunos valores de la tasa de contagio, la enfermedad se vuelve endémica en ausencia del depredador y desaparece si actúa el depredador.

En una segunda parte, estudiamos modelos estructurados en edad con difusión. Primero analizamos un modelo evolutivo con un término de reacción no lineal, positivo cuando el entorno es favorable para el individuo y negativo cuando es hostil. Comprobamos que un método de sub-supersoluciones funciona, y lo aplicamos a dos modelos ecológicos: a un modelo logístico generalizado y a otro de tipo Holling-Tanner. Gracias a dicho método, analizamos el comportamiento asintótico de la solución de cada modelo. En el último Capítulo de la memoria, estudiamos el modelo estacionario (en tiempo) asociado al modelo evolutivo del Capítulo anterior. Mediante un método de sub-supersoluciones y el estudio del problema de autovalores asociado, se estudian resultados de existencia y de no existencia de solución para un modelo logístico generalizado; comprobando que la condición no local que aparece por el proceso de nacimiento hace variar completamente los resultados que se obtienen en el caso de un problema parabólico clásico.

Multidimensional Minimizing Splines. Theory and Applications

Rémi Arcangéli, María Cruz López de Silanes, Juan José Torrens
Kluwer Academic Publishers. 2004.

ISBN: Edición en papel 1-4020-7786-6 (xv + 261 páginas) Edición electrónica 1-4020-7786-4



Por M.C. López de Silanes

Este libro está dedicado a la teoría de los D^m -splines y a algunas de sus aplicaciones en la aproximación de superficies. Los D^m -splines sobre un subconjunto abierto Ω de \mathbf{R}^n son splines de minimización multidimensionales, es decir, funciones definidas sobre Ω , sujetas a condiciones de interpolación o de ajuste, y que minimizan un funcional de tipo “energía” que contiene derivadas de orden m . Cuando el conjunto Ω es \mathbf{R}^n , los D^m -splines son justamente las funciones spline introducidas por J. Duchon, denominadas asimismo splines de placa fina o, más generalmente, splines poliarmónicos. En su estudio, abordado en la primera parte del libro, se adopta un punto de vista distinto del de J. Duchon, ya que se prescinde de la noción de núcleo reproductor. Sin embargo, este enfoque permite no sólo obtener resultados ya conocidos de J. Duchon, sino también establecer resultados complementarios relativos a la convergencia y a es-

timaciones del error. Se consideran a continuación los D^m -splines sobre un dominio acotado y sus aproximantes de tipo elementos finitos, llamados D^m -splines discretos. Se sigue para ello un procedimiento análogo al usado para los D^m -splines sobre \mathbf{R}^n . Por último, el libro trata varios problemas de ajuste de superficies que aparecen en la investigación petrolera y en ciencias geofísicas, prestando especial atención a la aproximación de superficies con fallas. Mediante detalladas consideraciones teóricas y ejemplos numéricos, se muestra cómo

los D^m -splines discretos pueden ser utilizados satisfactoriamente para resolver tales problemas.

Este libro está dirigido a matemáticos, geólogos, ingenieros y, en general, investigadores y estudiantes de doctorado que estén interesados en la teoría de funciones spline, problemas de aproximación de superficies o métodos variacionales.