

Consideraciones metodológicas sobre la enseñanza y el aprendizaje de la Teoría de la Computabilidad

A. J. PÉREZ

Departamento de Ciencias de la Computación e I.A.
Universidad de Sevilla

pjimenez@us.es

Resumen

En este artículo nos centramos en ciertas tendencias metodológicas con las que pueden abordarse la enseñanza y aprendizaje de un primer curso sobre *teoría de la computabilidad* en las Escuelas de Ingeniería Informática. Tras justificarse el método de presentación matemático como un método adecuado para la exposición de dicha teoría, abordamos los aspectos pedagógicos relacionados con la metodología escolar, decantándonos por métodos activos, como el basado en el razonamiento plausible de George Polya, en oposición a los métodos tradicionales de exposición y ejercicios.

1 Introducción

En su informe¹ de 1991 la ACM² y la IEEE-CS³ afirman que la Informática puede considerarse como una disciplina que conjuga tres componentes: una matemática, otra científica y una tercera técnica. Por eso, el desarrollo de las diferentes áreas emplea, básicamente, tres *metodologías* de trabajo o *procesos* en investigación, desarrollo y aplicaciones: **teoría**, **modelización** y **diseño**.

El primer proceso, la **teoría**, se utiliza para el desarrollo de teorías matemáticas que se usan en Ciencias de la Computación y consta de cuatro fases: *Definiciones y axiomas, teoremas, demostraciones e interpretación de resultados*.

El segundo proceso, **modelización**, se base en el método científico experimental y se divide en las siguientes fases: *Recogida de datos y elaboración*

¹**Computing Curricula 1991**. Report of the ACM/IEEE-CS Joint Curriculum Task Force. December 17, 1990. (Apartado 5.2). <http://www.computer.org/education/cc1991/>

²Association for Computing Machinery

³Institute of Electrical and Electronics Engineers-Computer Society

de hipótesis, construcción de modelos y realización de predicciones, diseño de experimentos y análisis de resultados.

El tercer proceso, **diseño**, tiene su base en la ingeniería y se usa en el desarrollo de un sistema para resolver un problema. Consta también de cuatro fases: *Estado de necesidades, estado de especificaciones, diseño e implementación del sistema, chequeo y análisis del sistema.*

2 Objetivos generales

La Informática tiene consideración de disciplina científica y técnica a la vez y, como tal, a la luz de los elementos señalados en los **CC1991** para cada uno de los procesos que la componen, puede distinguirse los siguientes objetivos generales a conseguir en la formación de los estudiantes:

- En relación con la componente **teórica**:
 - Caracterizar los objetos de estudio (definiciones, axiomas).
 - Promover la formulación de conjeturas plausibles (relaciones entre objetos de la teoría).
 - Ver con claridad que las conjeturas plausibles deben ser justificadas (teoremas).
 - Estimular el pensamiento crítico y el rigor científico.
 - Fomentar la capacidad de análisis y síntesis.
 - Impulsar la aptitud para resolver problemas.
 - Interpretar los resultados obtenidos en un marco teórico.
- En relación con la componente de **modelización**:
 - Recoger datos, organizándolos para su estudio.
 - Aprender a distinguir lo esencial de lo accesorio en una situación genérica.
 - Elaborar hipótesis a la luz de los datos de que se dispone.
 - Potenciar la búsqueda de modelos científicos asociados a situaciones concretas y hacer predicciones.
 - Estar en disposición para diseñar un experimento.
 - Analizar los resultados.
 - Potenciar la búsqueda de modelos científicos asociados a situaciones concretas.
- En relación con la componente de **diseño**:
 - Detectar las necesidades reales en un contexto concretos.

- Saber diseñar experiencias y ser capaz estudiar experiencias ya diseñadas.
- Elaborar métodos de análisis capaz de detectar las cualidades del sistema diseñado.
- Desarrollar los mecanismos que permiten aplicar los conceptos y resultados teóricos a situaciones prácticas.

Las asignaturas que versan sobre teoría de la computabilidad e impartidas en las titulaciones de Ingeniería Informática, pretenden precisamente *ser soporte de la formación teórica* de los alumnos, con lo que los objetivos a conseguir son en esencia, desde nuestro punto de vista, los enunciados anteriormente dentro del apartado de **teoría**.

3 Consideraciones metodológicas

Por un lado, los **CC1991** señala a las Matemáticas como uno de los componentes básicos de las Ciencias de la Computación. Por otro, en los **CC2001**⁴ se afirma que un estudiante de Informática necesita adquirir una serie de herramientas que arrojen su formación técnica, haciéndose notar además que es necesario que los estudiantes adquieran “ un cierto nivel *de sofisticación en matemáticas*, familiaridad con los métodos científicos y un cierto sentido de cómo la computación es aplicada en la práctica y habilidad para trabajar en equipo”. Pone el acento este documento en la necesidad de un rigor matemático en la formación que incluya el manejo de técnicas matemáticas y razonamientos formales matemáticos pues estos elementos, se afirma, están detrás de muchas áreas de la Informática.

No nos cabe duda, y las opiniones anteriores lo refuerzan, de que las asignaturas sobre teoría de la computabilidad se caracterizan precisamente por ser de naturaleza matemática y/o lógica, lo que justifica que la metodología que defendamos para las exposiciones sea precisamente *una metodología de presentación de las matemáticas*.

La metodología que ordinariamente se sigue en las clases tienen como base el *método deductivo* (o lógico-deductivo o axiomático-deductivo). Dicho método se traduce principalmente por los siguientes elementos: construcción de un corpus teórico a partir de un sistema de *definiciones, axiomas o postulados* desde los que se llega, mediante demostraciones (*proceso de deducción* que da nombre al método), a las *propiedades y teoremas* de la teoría.

Pero en la metodología escolar de cualquier materia, aparte del método de presentación científico elegido, está la vertiente pedagógica que condiciona la manera en que esa presentación es traducida en una clase concreta. Pues bien, esta componente pedagógica suele consistir en el *método tradicional expositivo* o *magistral*: el profesor explica la materia, ilustra con ejemplos y propone ejercicios que, en la clase siguiente, suele resolver el propio profesor.

⁴**Computing Curricula 2001**. Informe de la ACM/IEEE-CS - Diciembre, 2001. (Apartado 9.1) <http://www.computer.org/education/cc2001/>

Aunque el método deductivo de presentación de las matemáticas pueda ser desarrollado en clase con otros elementos pedagógicos distintos a los señalados, digamos ya que no es usual; sea por razones históricas, sea porque es “lo natural”, dicho método suele ir unido al expositivo tradicional. En adelante, al hablar de método deductivo lo consideraremos íntimamente unido en lo que concierne al desarrollo en clase, al método expositivo tradicional. Utilizaremos la nomenclatura de *metodología deductiva* o *deductiva-magistral*.

• **Metodología deductiva. Ventajas e inconvenientes.**

El método deductivo, como método para la presentación matemática, se remonta a Euclides. Sus *Elementos* son el primer paradigma de aplicación de tal método. Pero es a raíz de la crisis de fundamentación de las matemáticas, a principios del siglo XX, cuando se considera el método axiomático deductivo como una herramienta esencial en el proceso de construcción matemática y en el de la enseñanza de la misma. Este proceso que llega a su cumbre con las tesis bourbakistas, es revalidado por una corriente que se inicia a finales de los años 50, que pone el énfasis en una enseñanza axiomática y deductiva de las matemáticas.

El profesor Jean Dieudonné, muy preocupado por la enseñanza y que, por su gran personalidad, puede ser considerado como uno de los principales líderes de las llamadas *matemáticas modernas*, afirma en el prólogo de su libro *Álgebra lineal y geometría elemental* que

Es necesario que el alumno -de secundaria- tome conciencia de la necesidad de un tratamiento axiomático de las matemáticas.

Más adelante se muestra en contra de determinados tratamientos intuitivos:

Es deseable liberar al alumno cuanto antes de la camisa de fuerza de las “figuras” tradicionales(...) Asimismo conviene ciertamente enseñar a los alumnos el arte de las construcciones geométricas, pero hay que huir como de la peste de lo que es, sin duda, el mayor fastidio de la enseñanza clásica: la limitación de los instrumentos de dibujo a la regla y el compás.

Destacaré tres razones que suelen aducirse en defensa del método magistral-deductivo.

La primera es de tipo conservador y se basa en que las pautas de estudio de los alumnos están ya marcadas por este método; cambiarlo no debe hacerse en un contexto aislado y deberían reunirse determinadas condiciones (generalidad, número no elevado de alumnos, actitud, convencimiento y preparación amplia del profesor) para cambiar a una metodología más activa.

La segunda razón radica en la eficacia. Cualquier metodología activa, cualquier método heurístico es lento. Si ya de por sí los programas son bastante ajustados al tiempo que le corresponde, la adecuación a un cuatrimestre hace que el profesor tenga menos “margen de maniobra” para diseñar determinadas estrategias de actividades.

La tercera razón, y más poderosa, es la masificación en clase. Cuando el número de alumnos es muy elevado, y este es el caso actualmente en los estudios de ingeniería informática, no parece haber otra manera de proceder más que con una exposición de tipo magistral, siendo muy difícil articular cualquier forma en la que el alumno tenga un papel más participativo (trabajo en grupo, realización y exposición de trabajos, etc). Ni siquiera en las clases de problemas suele estar prevista una disminución del número de alumnos por grupo. Esta disminución sólo se contempla en las clases prácticas de Laboratorio, cuando las hay.

En estas circunstancias parecería que la metodología escolar obligada es la deductiva-magistral, el método tradicional de siempre.

Esta metodología presenta sin dudas ciertas ventajas tales como exposición clara y simple; economía de medios (en los símbolos y en el pensamiento); fácil adecuación del ritmo al programa de la asignatura; impecable si se mira desde el rigor científico, y meridianamente clara desde el punto de vista de la transmisión de conocimientos. Pero a pesar de estas ventajas y a pesar de ser un método consolidado al lo largo de muchos años, posee una serie de graves inconvenientes que hemos de señalar:

- Es una metodología impositiva y, por ello, no motivadora.
- Fomenta la actitud pasiva del alumno al convertirlo en mero receptor de las explicaciones del profesor.
- Es lineal en la presentación teórica, en aras del rigor, sin matizaciones acerca de lo que es más o menos pertinente, sin consideraciones sobre aspectos culturales.
- No tiene en cuenta la historia ni la génesis de los problemas.
- Apenas considera la intuición del alumnado, al centrarse casi de manera exclusiva en desarrollos formales.

Pensamos que en unos estudios de ingeniería, con una fuerte componente de experimentación y diseño, parece conveniente unas metodologías más intuitivas, más dinámicas y más participativas, incluso -o quizás más aún por ello- a la hora de abordar asignaturas con una fuerte componente teórica.

• Metodologías inductivas. El razonamiento plausible

En el siglo XX prolifera una rica variedad de tendencias en la enseñanza en general y, en particular, en la enseñanza de la matemáticas. Muchos matemáticos se muestran partidarios de una enseñanza más intuitiva, contrarios, en más de un caso, a una enseñanza puramente deductiva. Citemos a algunos. Poincaré, por ejemplo, se pregunta⁵:

⁵H. Poincaré, *La Ciencia y la Hipótesis*; Espasa Calpe, Madrid, 1963.; pág. 15

¿Cuál es la naturaleza del razonamiento matemático? ¿Es realmente deductivo como ordinariamente se cree? Un análisis profundo nos muestra que no es así; que participa en una cierta medida de la naturaleza del razonamiento inductivo, y que por eso es fecundo. Pese a ello no pierde su carácter de rigor absoluto.

En otra ocasión, es contundente⁶:

¿Es posible entender una teoría si desde el primer momento se le da la forma de definitiva que impone una lógica rigurosa, sin mencionar para nada el camino por el que ha llegado a adoptar esta forma? No, realmente no es posible entenderla; incluso resulta imposible retenerla si no es de memoria.

Poincaré acaba planteando una ponderación entre la lógica y la intuición al afirmar⁷:

La lógica y la intuición tienen cada una un papel necesario; ambas son indispensables. La lógica, que puede por sí misma dar la certeza, es el instrumento de la demostración; la intuición es el instrumento de la invención.

Courant y Robbins en el conocido libro *¿Qué es la Matemática?* se rebelan contra el método axiomático deductivo en los siguientes términos:

Parece existir un grave peligro en el excesivo predominio del carácter axiomático deductivo de las matemáticas(...) Si la forma deductiva cristalizada es la meta, la intuición y la construcción son, cuando menos, las fuerzas directrices. Una amenaza sería para la vida de las ciencias aparece contenida en la afirmación de que la matemática no es más que un sistema de conclusiones derivadas de definiciones, reglas y silogismos, sin meta ni motivos algunos

Y en un célebre manifiesto⁸ de 1962, firmado en EE.UU. por más de setenta profesores⁹ y publicado en las revistas *American Mathematical Monthly* y *The Mathematics Teacher*, se afirma:

El pensamiento matemático no es tan sólo razonamiento deductivo; no consiste simplemente en demostraciones formales. (...) sin alguna experiencia en procesos "informales" de pensamiento, el estudiante no puede comprender el verdadero papel de la demostración formal y rigurosa (...) el objeto del rigor es "confirmar y legitimar las conquistas de la intuición.." [Hadamard] (...) Hay varios niveles de rigor. El estudiante debería aprender a comprender, buscar y criticar las demostraciones al nivel correspondiente a su experiencia y formación.

⁶H. Poincaré, *El valor de la Ciencia*; Espasa Calpe, Madrid, 1963.; pág. 15

⁷H. Poincaré, *El valor de la Ciencia*; Espasa Calpe, Madrid, 1963.; pág. 15

⁸*On the mathematics curriculum for the High School* [10]

⁹entre otros: Courant, Coxeter, Kline, Polya, Stoker, Weil

Una personalidad como Felix Klein, que encabeza, a comienzos del s. XX un proceso de reforma para la enseñanza de la matemáticas (en secundaria principalmente) y que en 1908 crea, en el seno de la Unión Matemática Internacional (IMU), la Comisión Internacional para la Instrucción Matemática (ICMI), afirma que enseñar científicamente no consiste en enfrentar a los alumnos con "fríos sistemas científicamente pulidos" sino inducirles a pensar científicamente.

Desde otros ámbitos (pedagogía, psicología) se propugnan alternativas metodológicas, sobre todo en la enseñanza secundaria. Se considera que frente a la mera actitud de enseñar, hay que *fortalecer el aprendizaje* pues el objetivo es formar al alumno a quien, por tanto, hay que ceder el protagonismo de la enseñanza. Son las *metodologías activas*.

Una corriente extrema en este sentido es la corriente *neoconductista* que encabeza el psicólogo B.F. Skinner¹⁰ quien considera que el alumno puede, con la ayuda de una máquina, desarrollar gran parte de su aprendizaje. Su método de *enseñanza programada* consiste, en esencia, en dar los conocimientos al alumno a través de pequeñas dosis de información (items) de tal manera que cuando el alumno yerra vuelve hacia atrás con preguntas alternativas que pretenden reconducirle a la solución correcta del ítem no asimilado.

Otras metodologías menos radicales se abren paso; son las metodologías *activas*, basadas en la psicológica *genética* de J. Piaget. Estas metodologías afirman que la esencia de la enseñanza está en la *acción* sobre objetos. Por acción se entiende desde la manipulación y observación de objetos reales hasta la interacción con objetos mentales. En los años cincuenta y gracias a la acción de un grupo de trabajo internacional e interdisciplinar, la CIEAM¹¹, formado entre otros por pedagogos como el profesor inglés Caleb Gattegno; psicólogos como el suizo Jean Piaget, maestros como la italiana Emma Castelnuovo; y matemáticos como el francés Jean Dieudonné o el español Pedro Puig Adam, impulsan un movimiento en torno a la enseñanza activa de la matemática conocido como "enseñanza dinámica". Estos métodos que pretenden dar más protagonismo al aprendizaje frente al mero hecho de enseñar, van abriéndose camino con pujanza.

Dentro de estas corrientes de *enseñanza activa*, habría que situar la tendencia encabezada por George Polya conocida como *Resolución de problemas*. Importa mucho señalarla, además, porque va dirigida principalmente al mundo universitario. Frente a la metodología deductiva, Polya propone una metodología inductiva y heurística que, en su obra *Cómo resolverlo it*¹² plasma en cuatro fases:

1. Comprender el problema.
2. Concebir un plan.
3. Ejecutar el plan.

¹⁰, *Tecnología de la enseñanza*, Ed. Labor, Madrid, 1970

¹¹Comisión Internacional por l'Enseignement et l'Amelioration de la Mathematique

¹²*Comment Poser et Résoudre un Problème*. Ed. Dunod. París, 1965

4. Examinar la solución obtenida.

Descompone, a su vez, cada fase en preguntas y sugerencias que son, en realidad, estrategias encaminadas a recorrer un camino que conduzca a la solución del problema. Así, se plantea:

1. ¿Cuáles son las incógnitas?
2.
 - He aquí un problema parecido y resuelto; ¿podría utilizarlo?
 - ¿Podría enunciarlo de otra forma?
 - ¿Conoce un problema análogo?
 - Particularice.
 - Generalice.
 - Divida el problema en partes.
3. Compruebe todos los pasos; ¿ve claramente que todos son correctos?
4. ¿Puede verificar el resultado? ¿Puede obtenerlo de forma distinta?

Fundamenta esta metodología inductiva a través de su teoría del *razonamiento plausible*, que justifica en su obra “Matemáticas y razonamiento plausible”. Según Polya¹³, “nuestro conocimiento matemático se asegura mediante el *razonamiento demostrativo*, mientras que las conjeturas se afianzan mediante el *razonamiento plausible*.... En el razonamiento estricto lo principal es distinguir una prueba de una intuición, una demostración válida de un intento sin validez. En el razonamiento plausible lo importante es distinguir entre intuiciones, unas más y otras menos razonables”.

Polya considera que en la enseñanza debe tenerse en cuenta este método inductivo como elemento que genera un proceso creador. Y aboga por que, incluso en la formación del estudiante matemático, entre también esta metodología del “razonamiento plausible”:

I. Lakatos¹⁴, alumno y seguidor de las teorías de Polya, afirma:

[las matemáticas] no se desarrollan mediante el aumento del número de teoremas indubitavelmente establecidos sino mediante la incesante mejora de las conjeturas, gracias a la especulación y a la crítica, siguiendo la lógica de las pruebas y refutaciones

En España, el profesor P. Puig Adam -profesor de metodología en la Universidad de Madrid en los años 40 y 50- creó escuela con el método heurístico. Puig Adam afirmaba que es¹⁵ “absurdo el método lógico, hipotético-deductivo

¹³G. Polya *Matemáticas y Razonamiento Plausible*. Editorial Tecnos, Madrid, 1966, págs. 13-14

¹⁴Su obra, *Pruebas y refutaciones*- Alianza Editorial, Madrid, 1978- retoma un problema, el de la definición de polígono, allá donde la evidencia plausible de Polya lo ha dejado

¹⁵P. Puig Adam *La Matemática y su enseñanza actual*. Publicaciones de la Revista de Enseñanza Media. Madrid, 1960

(...) igual desde Euclides(...) olvidando el largo proceso previo de edificación de categorías mentales” y, en su conferencia de 1953, *El Movimiento Didáctico Renovador* afirmaba que “el movimiento heurístico tiende a acercar el proceso de génesis de los conocimientos y el de transmisión de los mismos”.

Pero cuando gran parte de la comunidad educativa, principalmente en los niveles primario y secundario, estaba sensibilizándose a favor de los métodos activos, en la comunidad matemática se produce una rebelión contra esa corriente considerada “pedagoga”. Lidera esa rebelión, entre otros, por el profesor J. Dieudonné, quien en un Congreso celebrado en Royaumont, en 1959, lanzó su célebre proclama “Abajo Euclides”, que se considera como la señal, el punto de arranque, de una corriente pedagógica conocida como *matemáticas modernas*. Esta corriente, si bien supuso una renovación muy importante en los contenidos de los programas de matemáticas, fue un considerable paso atrás en la renovación metodológica de esta disciplina.

El manifiesto antes citado fue una reacción de más de setenta matemáticos contra esta irrupción de la matemática moderna:

Los matemáticos, reaccionando frente al control de la educación por los educadores profesionales, quienes quizás han hecho hincapié en la pedagogía a expensas del contenido, pueden ahora acentuar el contenido a expensas de la pedagogía de forma igualmente estéril (...) La necesidad de hoy de aprender más matemáticas que en el pasado puede llevarnos a buscar atajos que, sin embargo, podrían ser más dañinos que beneficiosos

4 Consideración final

Por todo lo dicho queremos concluir haciendo hincapié en nuestra convicción de que las metodologías activas y, en particular, una metodología heurística como la de Polya basada en el razonamiento plausible, son más adecuadas para enseñanzas de disciplinas teóricas, fuertemente impregnadas por las matemáticas. Aparte de razones intrínsecas, queremos volver a poner el acento en que la formación de un informático tiene unas componentes de modelización y diseño en las que puede colaborar, sin lugar a dudas, una metodología heurística.

Creemos, finalmente, que la incorporación de los nuevos medios (tecnologías de la información y comunicación principalmente), la paulatina disminución del número de alumnos por profesor y grupo y determinadas actuaciones oficiales, como las tutorías, la participación institucional de los alumnos, etc., son elementos que deben aprovecharse para que conduzcan de manera natural a una implicación mucho más importante del alumno en su propio aprendizaje (y no sólo en los aspectos “políticos y de gestión” de su enseñanza).

Referencias

- [1] ACM/IEEE-CS, *Computing Curricula 2001*, ACM/IEEE-CS, 2001.

- [2] ACM/IEEE-CS, *Computing Curricula 1991*, ACM/IEEE-CS, 1990.
- [3] COURANT R. Y ROBBINS H., *¿Qué es la Matemática?*, Ed. Aguilar, Madrid, 1971 .
- [4] DIEUDONNÉ, J. *Álgebra Lineal y Geometría Elemental*, Secciones Científicas, Madrid, 1971.
- [5] KLEIN, F. *Enseñanza elemental desde un punto de vista superior*, Ed. Ibero-Americana; Buenos Aires, 1948.
- [6] KLINE, M. *El fracaso de la Matemática Moderna*, Ed. Siglo XXI, 1984.
- [7] KLINE, M., *La pérdida de la certidumbre*, Ed. Siglo XXI, 1985.
- [8] LAKATOS, I *Pruebas y refutaciones* Alianza Editorial, Madrid, 1978.
- [9] PÉREZ JIMÉNEZ, A. *Tendencias actuales en la enseñanza de las Matemáticas*. Actas IV Jornadas Andaluzas de Educación Matemática. SAEM Thales, Málaga, 1989.
- [10] PIAGET, J., CHOQUET, G., DIEUDONNÉ, J., THOM, R. Y OTROS, *La enseñanza de las matemáticas modernas*, Alianza Editorial, Madrid, 1980.
- [11] POINCARÉ, H., *La Ciencia y la Hipótesis*; Espasa Calpe, Madrid, 1963.
- [12] POINCARÉ, H., *El valor de la Ciencia*; Espasa-Calpe, Madrid, 1963.
- [13] POLYA, G., *Matemáticas y Razonamiento Plausible*, Editorial Tecnos, Madrid, 1966.
- [14] POLYA, G., *Comment Poser et Résoudre un Problème*. Ed. Dunod. París 1965.
- [15] PUIG ADAM, P., *La Matemática y su enseñanza actual*. Publicaciones de la Revista de Enseñanza Media. Madrid, 1960.
- [16] SKINNER, B.F., *Tecnología de la enseñanza*, Ed. Labor, Madrid, 1970.

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:

Nos dirigimos de nuevo a todos vosotros para hacer una breve presentación de un nuevo número del boletín de nuestra Sociedad. En esta ocasión, se incluye un artículo de opinión, cuatro de Investigación y uno de Educación Matemática.

Queremos destacar también en esta presentación, que recientemente han sido fallados los premios SĒMA al Joven Investigador y el de Divulgación de la Matemática Aplicada. El VII Premio al Joven Investigador ha sido concedido a Marco Antonio Fontelos López, Profesor Titular de la Universidad Rey Juan Carlos. Asimismo, el ganador del V Premio de Divulgación de la Matemática Aplicada ha sido Pablo Fernández Gallardo, Profesor del Departamento de Matemáticas de la Universidad Autónoma de Madrid. El título del trabajo premiado es: *El secreto de Google y el Álgebra Lineal*. Desde aquí queremos transmitir nuestra más sincera felicitación a los premiados.

Un cordial saludo,

Grupo Editor
boletin_sema@orion.ciencias.uniovi.es

La Matemática Aplicada entre lo divino y lo humano

L. L. BONILLA, A. LIÑÁN Y J. M. VEGA

Departamento de Matemáticas, Escuela Politécnica Superior,
Universidad Carlos III de Madrid ;
Departamento de Motopropulsión y Termofluidodinámica,
Universidad Politécnica de Madrid ;
Departamento de Fundamentos Matemáticos de la Ingeniería
Aeronáutica, Universidad Politécnica de Madrid

bonilla@ing.uc3m.es, alinan@aero.upm.es,
vega@fmetsia.upm.es

España es la décima potencia mundial si consideramos el número de artículos de Matemáticas publicados en un año con un autor español por lo menos ¹. Cuando estos artículos se ponderan con su índice de impacto y se toma la media mundial como cero, el índice español en Matemáticas es -13 ². Esta situación es particularmente grave en el área de Matemática Aplicada. Según el informe Andradás-Zuazua (que expurgó los artículos matemáticos aparecidos en revistas relevantes de Física, Mecánica de Medios Continuos, Mecánica de Fluidos, etc. e incluso en la revista SIAM Journal on Applied Mathematics), la revista en que más publican los “matemáticos aplicados” españoles es Proceedings of the American Mathematical Society, una revista que se encuentra año tras año en el último cuarto de la lista de Matemática Aplicada del índice de impacto JCR. Anecdóticamente, la ANECA reconoce esta situación usando baremos más bajos para Matemáticas que para las otras disciplinas científicas.

Esta situación de inflación en la producción española de Matemática Aplicada, y de baja relevancia de la misma, hunde sus raíces en la formación que reciben los licenciados en Matemáticas. *Nihil ex nihilo*. Si el matemático aplicado español no ha estudiado con suficiente profundidad una disciplina científica y, por tanto, ignora qué considera un científico o ingeniero problema relevante en su área y qué considera una solución admisible a dicho problema, el trabajo de este matemático aplicado en áreas aplicadas será probablemente un ejercicio en escolasticismo. Por tanto, consideramos que una parte apreciable de los créditos docentes a incluir en los nuevos planes de estudio de Matemáticas debe dedicarse a lograr que los futuros matemáticos, y muy especialmente los aplicados, aprendan suficientemente una disciplina científica. Deben entender el

¹datos de 2002, fuente: ISI

²datos del cuatrienio 1998-2002, fuente: ISI

papel primordial que tiene contrastar los resultados teóricos con la realidad a través de experimentos, y que los resultados de un buen experimento sirven para decidir qué teorías descartar y cuáles seguir (con independencia de su posible valor estético). Si los nuevos planes de estudio de Matemáticas se elaboran desde esta perspectiva, con un interés genuino por transmitir en la enseñanza la potencia y versatilidad de las Matemáticas, se conseguirá, sin duda, que los licenciados tengan acceso, de un modo natural, a puestos de trabajo que les están vedados en la actualidad. Si, en cambio, se ignoran estas realidades o, aún peor, se buscan soluciones tramposas, se perderá una magnífica ocasión de evitar que a las facultades de Matemáticas sólo accedan, y en cantidad decreciente, alumnos con una clara vocación por la docencia o la investigación en Matemáticas, o alumnos sin calidad suficiente para acceder a estudios más estimulantes o competitivos.

Una parte sustancial de los resultados matemáticos no triviales se generan en el modelado y formulación de problemas matemáticos en Ciencias Físicas, Biológicas, etc. y en Tecnologías. Para llevar a cabo esta labor, se requieren conocimientos y formación tanto en Matemáticas como en el campo de las aplicaciones de que se trate. Sólo así se puede aportar algo relevante al modelado y a la resolución teórica (simulación, análisis, etc.) de un problema surgido en un contexto científico o tecnológico y, posiblemente, a las propias Matemáticas. La aportación científica o tecnológica se debe dar a conocer en aquellas publicaciones que los científicos o ingenieros usan en su trabajo, es decir, en las revistas de referencia del campo. Las revistas estrictamente matemáticas informarán de su trabajo a otros matemáticos, pero no a los destinatarios naturales de los resultados de la Matemática Aplicada. Algunas de estas revistas contienen siempre o con frecuencia resultados que son genuinamente de Matemática Aplicada. Unos pocos ejemplos de estas revistas: *SIAM Journal on Applied Mathematics*, *Physica D*, *Physical Review A-E*, *Journal of the Mechanics and Physics of Solids* (la revista de referencia para Mecánica de Sólidos), *Proceedings of the Royal Society of London A*, *Journal of Fluid Mechanics* (la revista de referencia para Fluidos), *Combustion Theory and Modelling*, *Journal of the Optical Society of America* (la revista de referencia para Óptica, incluyendo óptica no lineal, propagación de luz en medios dispersivos, etc.).

Demostrar teoremas no es, ni de lejos, el único producto de la actividad en Matemática Aplicada. Con frecuencia, tal limitación puede conducir a bendecir matemáticamente resultados conocidos desde hace décadas y, a menudo, obsoletos, renunciando a atacar mediante métodos no rigurosos los problemas realmente candentes. Como ejemplo, los métodos de perturbaciones singulares, cuya justificación rigurosa sólo se ha podido hacer en algunos casos particulares, sirven para simplificar modelos matemáticos (modelizar matemáticamente) en límites realistas, y para obtener aproximaciones de las soluciones, que conducen a una descripción cuantitativa muy útil de su estructura espacio-temporal. Muchas veces estos métodos aclaran cuál es el comportamiento cualitativo de las soluciones del problema y cuáles son los ingredientes fundamentales del modelo. Estos métodos se aplican en situaciones en que aparecen parámetros grandes o

pequeños asociados a una disparidad de escalas espaciales, temporales o espacio-temporales. En estos casos, la simulación numérica directa es costosa o incluso problemática (problemas stiff, problemas en que no se sabe a priori cuáles son los efectos más importantes que hay que incluir en la formulación de modelo). Y hay en España grupos competitivos internacionalmente que utilizan con éxito las técnicas de perturbaciones singulares (cuyo uso raramente es rutinario) como herramienta esencial.

La situación de penuria de la Matemática Aplicada española tiene su fiel reflejo en el reciente Plan Nacional de Matemáticas. Ésto salta a la vista hojeando el desglose de temas de investigación por áreas de la Matemática recogidas en la primera convocatoria de proyectos de investigación encuadrados en el Plan Nacional. Mientras que las áreas tradicionales de la Matemática se desglosan de manera reconocible y entendible por cualquier matemático, los epígrafes de la Matemática Aplicada, o bien ya están incluidos en Ecuaciones Diferenciales (caso de las ecuaciones diferenciales ordinarias o en derivadas parciales), o incluyen al mismo nivel general casos particulares de métodos de perturbaciones singulares (homogeneización y métodos multiescala). Además, se resalta solamente un área de las aplicaciones (Mecánica de Medios Continuos), y se añade “Aplicaciones”, por si acaso hay en el mundo otras áreas que se les hayan podido olvidar a los redactores del Plan.

Dada la escasa formación de muchos matemáticos en campos donde se aplica la Matemática, la evaluación de proyectos en Matemática Aplicada constituye un problema serio al que no se ha prestado la debida atención. En la situación actual, un proyecto original de Matemática Aplicada, que contribuya de forma notable a resolver problemas científicos mediante el modelado, resolución numérica, análisis asintótico, etc. corre el riesgo de ser pospuesto, o incluso rechazado, por no tener como meta demostrar teoremas y porque el evaluador matemático típico no sabe valorar la aportación original del proyecto. Por tanto, cada matemático aplicado que redacta un proyecto debe sopesar cuidadosamente si no le conviene más enviar su proyecto al área científica o tecnológica correspondiente para que se evalúe adecuadamente la aportación real del proyecto. En muchos casos estas consideraciones hacen que proyectos de Matemática Aplicada se encaucen a través de otras disciplinas. Si, en cambio, se incluyera un informe científico o tecnológico (escrito por un profesional de la ciencia o ingeniería pertinente) de cada proyecto de Matemática Aplicada, se tendría una opinión relevante sobre la viabilidad e interés del proyecto y, seguramente, se conseguiría que más proyectos verdaderamente aplicados se dirigieran al área de Matemática Aplicada, y que proyectos cuyos aspectos aplicados son inviables o inexistentes se evaluaran por su mérito matemático intrínseco, sin hojas de parra de “interés para las aplicaciones”.

Un problema grave de la Matemática Aplicada española es que el área de conocimiento del mismo nombre es excesivamente numerosa. Y ha servido en ciertos casos para “colocar” a numerosos matemáticos puros en Escuelas de Ingeniería, travestidos como “matemáticos aplicados”. Ésto causa distorsiones, tanto en los aspectos investigadores comentados antes como en la docencia que se imparte. Por buena voluntad que muchos de estos profesionales “aplicados”

pongan, no saber para qué sirven los conocimientos que se imparten a los ingenieros da lugar a numerosas paradojas, tanto en el nivel de rigor exigido como en el tratamiento u omisión de temas importantes. Por ejemplo, se puede divagar sobre el pecado que se comete al hablar de *el arco seno* en lugar del más correcto *un arco seno*, siendo la función arco seno multivaluada, y omitir explicar la transformación conforme y sus aplicaciones en la resolución de problemas potenciales planos, y otros temas de interés para un alumno de ingeniería industrial, aeronáutica o de telecomunicación. La falta de un lenguaje y cultura comunes con los ingenieros no ayuda precisamente a que estos problemas se detecten y resuelvan. Por otro lado, el “matemático aplicado” insiste en usar como vara de medir para resolver plazas de Matemática Aplicada (a nivel de oposición o de contratos) la existencia de teoremas en los trabajos de los solicitantes (aunque dichos teoremas puedan ser sólo variantes conocidas de otros demostrados muchos años antes) y, desde luego, no la relevancia de estos trabajos en el ámbito de la Matemática Aplicada o de la disciplina aplicada de que se trate. Esta mentalidad realimenta la situación de la Matemática Aplicada española antes mencionada y hace que sea muy difícil que se consoliden grupos verdaderamente aplicados en España. Cuesta bastante tiempo y esfuerzo formar a un matemático aplicado, bastante más que a un matemático puro, porque el aplicado debe entender también cómo se modelan y resuelven problemas en otras disciplinas lo que, en gran medida, se aprende por experiencia. Estas carencias en la formación inicial sólo pueden compensarse con un esfuerzo considerable, como el que ya han hecho algunos matemáticos españoles, en cuyo trabajo se aprecia un esfuerzo genuino por las aplicaciones.

El sistema actual de acceso al funcionariado mediante habilitación es muy dudoso que resuelva nada por dos razones. La primera, que el número de habilitados que aparece tras el parto de los montes de cada proceso de habilitación es ridículo por lo exiguo. La segunda razón es que el resultado de dichas habilitaciones es aleatorio por la fragmentación del área y la disparidad de criterios entre los profesores adscritos a la misma. El resultado de estos procesos es oneroso para las arcas públicas y para las finanzas de los participantes. También puede ser muy negativo para el trabajo y la carrera científica de los participantes, porque la confusión reinante en el área puede dar lugar a decisiones basadas en criterios subjetivos, ideológicos, de pertenencia a familias matemáticas, etc., ajenos a los logros científicos en el ámbito de la Matemática Aplicada. Por ejemplo, una posibilidad muy real, dado el sistema actual y la composición del profesorado del área, es que ningún habilitado en los concursos de Matemática Aplicada en un futuro previsible sea un matemático aplicado capaz de decir algo relevante en algún campo de las aplicaciones.

A menos que el número de habilitados supere con creces al de plazas que se ofrecen, y se habilite a un número suficiente de matemáticos aplicados verdaderos, no se creará un mercado en el cual las universidades puedan elegir según sus necesidades. ¿Podrá entonces extrañar a alguien que no se oferten plazas de nueva creación en número razonable?. Se ha tratado de remediar la endogamia de la Universidad legislando sin pararse a pensar en las consecuencias. Hoy día, no se le ocurre a nadie empezar a fabricar un motor

o construir un avión sin entender primero cómo funciona, y sin modelar, simular numéricamente y experimentar a pequeña y después a gran escala. Pues bien, el diseño de la Universidad, del cual depende en buena medida el futuro de nuestra sociedad, parece estar haciendo sin prever efectos indeseables, sin simulaciones o experimentos a pequeña escala. Para erradicar la endogamia, resultaría más efectivo diseñar mecanismos de financiación competitivos que penalicen gravemente tener personal incompetente en los departamentos universitarios.

Open problems related to the p -Laplacian*

J. ČEPIČKA, P. DRÁBEK AND P. GIRG

Centre of Applied Mathematics, University of West Bohemia,
Czech Republic

cepicka@kma.zcu.cz, pdrabek@kma.zcu.cz,
pgirg@kma.zcu.cz

Abstract

Quasilinear elliptic equations play a key role in contemporary mathematical modelling in science and industry. This paper concentrates on open problems connected with one of their most popular examples — equations having the p -Laplacian $u \mapsto \Delta_p u := \operatorname{div}(|\nabla u|^{p-2} \nabla u)$ as their main part. The p -Laplacian is rather popular among nonlinear analysts not only due to the fact that it is general enough to exhibit a behaviour typical for singular and/or degenerated quasilinear differential operators, but also because its definition is simple enough in order to reduce significantly technicalities in the proofs.

Key words: *the p -Laplacian, Fredholm alternative, bifurcation, nonlinear eigenvalue problem*

AMS subject classifications: *35P30, 47J10, 35J20, 49J35*

1 Introduction

In this paper we want to discuss some properties of the quasilinear partial differential operator

$$\Delta_p u := \operatorname{div}(|\nabla u|^{p-2} \nabla u),$$

where $p > 1$ is a real number and $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^N . This operator is called the p -Laplacian and has been studied intensively during the last two decades. In one dimension ($N = 1$), it reduces to the ordinary differential operator

$$u \mapsto (|u'|^{p-2} u')'.$$

*Research supported by the Grant Agency of the Czech Republic, grant no. 201/03/0671 and the Research Plan # MSM 235200001.

Fecha de recepción: 1 de marzo de 2004

It is clear that for $p = 2$ it is a linear operator which is nothing but the Laplacian. Since for $p \neq 2$ only the homogeneity of this operator is preserved while the additivity is lost, the equations involving the p -Laplacian with $p \neq 2$ are sometimes called “half-linear” equations.

The p -Laplacian started to be an object of intensive study of mathematicians as soon as it was recognized as a principal part in some mathematical models from glaciology, non-Newtonian fluid mechanics, nonlinear diffusion, climatology, non-conformal geometry, etc. Let us mention the book of DÍAZ [10] from middle eighties as an example of extensive study of the p -Laplacian with applications to fluid mechanics.

Mathematicians concentrated on proving some fundamental properties of this nonlinear operator, namely on those which were known for the linear Laplace operator. So, the strong maximum principle was proved by VÁZQUEZ [29] and OTANI [26], a strong comparison principle was proved by CUESTA and TAKÁČ [5], the simplicity and isolatedness of the first eigenvalue of the p -Laplacian subject to the homogeneous boundary conditions was established by ANANE [1], etc.

Many other properties of the p -Laplacian were discovered in connection with the solvability of the equation

$$-\Delta_p u = f(x, u, \nabla u) \tag{1}$$

on various types of domains and subject to various types of boundary conditions. The authors of this paper would guess that 95% of the papers published on quasilinear problems of type (1) generalize the results which have been already known for the semilinear case (with $p = 2$). Sometimes new results appear directly for the quasilinear problem and the semilinear problem is considered as a special case. But there are not too many papers which show the striking difference between the linear Laplacian and the quasilinear p -Laplacian with $p \neq 2$. This is the situation we want mainly to concentrate on in this paper.

2 The eigenvalue problem

Let us assume that $\Omega \in \mathbb{R}^N$ is a bounded domain with smooth boundary $\partial\Omega$ for $N \geq 2$ and Ω is a bounded interval, say $\Omega = (0, 1)$, for $N = 1$. In many situations the study of problem

$$\begin{cases} -\Delta_p u = f(x, u) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{2}$$

with a nonlinear function f satisfying certain asymptotic hypotheses, requires the knowledge of the set of all real numbers λ for which the homogeneous (*eigenvalue*) problem

$$\begin{cases} -\Delta_p u - \lambda|u|^{p-2}u = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{3}$$

admits a nontrivial (nonzero) solution. Such values of λ are called the *eigenvalues* of (3) and the corresponding nontrivial solutions are called the *eigenfunctions* of (3). The linear eigenvalue problem (3) with $p = 2$ is well understood and the complete set of eigenvalues,

$$0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \lambda_4 \leq \dots \rightarrow \infty,$$

can be characterized by the *Courant-Weinstein variational principle* (see e.g. COURANT and HILBERT [2]). The question arises whether there is a similar tool available also in the general (nonlinear) case $p \neq 2$. The answer is positive but it is not known if the situation is exactly the same as in the linear case. The discussion of this problem will result in our first open problems.

A standard approach to the way to get a sequence of eigenvalues of (3) relies on the notion of the *Ljusternik-Schnirelmann category*. Another one relies on the notion of the *Krasnoselskij genus* and there are also other techniques. We shall explain here this approach in more detail.

Let us consider the even functional

$$I(u) := \frac{\int_{\Omega} |\nabla u|^p dx}{\int_{\Omega} |u|^p dx}, \quad u \in W_0^{1,p}(\Omega) \setminus \{0\}$$

and the symmetric manifold

$$\mathcal{B} := \{u \in W_0^{1,p}(\Omega) : \int_{\Omega} |u|^p dx = 1\}.$$

Clearly, the eigenvalues and eigenfunctions of (3) correspond to the critical values and critical points of $I|_{\mathcal{B}}$. Using a standard compactness argument, it is possible to prove that $I|_{\mathcal{B}}$ satisfies the Palais-Smale condition (see e.g. DRÁBEK and ROBINSON [17]). Now, for any nonempty symmetric closed subset $\mathcal{A} \subset \mathcal{B}$ we define

$$\gamma(\mathcal{A}) := \begin{cases} \inf\{m \in \mathbb{N} : \exists \text{ a continuous odd map of } \mathcal{A} \text{ into } \mathbb{R}^m \setminus \{0\}\}, \\ \infty \text{ if } \{m \in \mathbb{N} : \exists \text{ a continuous odd map of } \mathcal{A} \text{ into } \mathbb{R}^m \setminus \{0\}\} = \emptyset, \end{cases}$$

$$\mathcal{F}_k^* := \{\mathcal{A} \subset \mathcal{B} : \bar{\mathcal{A}} = \mathcal{A}, -\mathcal{A} = \mathcal{A}, \gamma(\mathcal{A}) \geq k\},$$

$$\lambda_k^* := \inf_{\mathcal{A} \in \mathcal{F}_k^*} \sup_{u \in \mathcal{A}} I(u). \quad (4)$$

Then, it is well known that $\{\lambda_k^*\}$ is a sequence of eigenvalues of (3) such that $\lambda_k^* \rightarrow \infty$ as $k \rightarrow \infty$ (see e.g. GARCIA-AZORERO and PERAL [22]).

In [17], a useful variant of this approach is explored. For $k \in \mathbb{N}$ let us set

$$\mathcal{F}_k := \{\mathcal{A} \subset \mathcal{B} : \mathcal{A} \text{ is the image of a continuous odd function } h : \mathcal{B}^{k-1} \rightarrow \mathcal{B}\},$$

where \mathcal{B}^{k-1} represents the unit sphere in \mathbb{R}^k . Define

$$\lambda_k := \inf_{\mathcal{A} \in \mathcal{F}_k} \sup_{u \in \mathcal{A}} I(u). \quad (5)$$

It is then straightforward to check that $\{\lambda_k\}_{k=1}^\infty$ is also a sequence of eigenvalues of (3). Moreover, since $\mathcal{F}_k \subset \mathcal{F}_k^*$ we have that $\lambda_k \geq \lambda_k^*$, so $\lambda_k \rightarrow +\infty$ as $k \rightarrow +\infty$.

It is also clear that $\lambda_1 = \lambda_1^*$, and it is proved in [17] that $\lambda_2 = \lambda_2^*$.

OPEN PROBLEM 2.1. *Let $N \geq 2$. Prove that $\lambda_k = \lambda_k^*$ for any $k \geq 3$ or prove that there exists $k \geq 3$ such that $\lambda_k > \lambda_k^*$.*

In one dimension, i.e. $N = 1$, $\Omega = (0, 1)$, we have $\lambda_k = \lambda_k^*$ for all $k = 1, 2, \dots$. It is also known in this case that the “infsup” formulæ (4) (or (5)) yield the complete set of eigenvalues of (3) (see e.g. CUESTA [3], DRÁBEK and MANÁSEVICH [16]). In more dimensions this is not known. Namely, we have:

OPEN PROBLEM 2.2. *Let $N \geq 2$. Prove that the sequence $\{\lambda_k^*\}_{k=1}^\infty$ (or $\{\lambda_k\}_{k=1}^\infty$) accounts for the entire spectrum of (3) (i.e. the set of all eigenvalues of (3)), or find Ω and $p > 1$ for which there is an eigenvalue of (3) which does not allow this characterization.*

It is well known that for $\lambda_1 = \lambda_1^*$ and $\lambda_2 = \lambda_2^*$ (given by (4) and (5) for $k = 1$ and $k = 2$), there is no other eigenvalue of (3) in the interval (λ_1, λ_2) . On the other hand, the properties of the spectrum beyond λ_2 are not known. For instance, the existence of a whole interval of eigenvalues of (3) or the existence of a sequence of eigenvalues of (3) approaching λ_2 from the right have not been excluded.

As already mentioned in the Introduction, it was proved in [1] that λ_1 is a simple eigenvalue and the corresponding eigenfunction φ_1 can be taken positive in Ω . Let us define a *nodal domain* of $u \in C(\overline{\Omega})$ as a maximal connected open subset of $\{x \in \Omega : u(x) \neq 0\}$. Then φ_1 has exactly one nodal domain. It is known that any eigenfunction associated with λ_2 has exactly two nodal domains (see CUESTA, DE FIGUEIREDO and GOSSEZ [4] or DRÁBEK and ROBINSON [18]). The situation is more complicated for higher eigenfunctions, except in the one dimensional case $N = 1$. For $\Omega = (0, 1)$, one can prove an analogue of the Sturm-Liouville theorem (see e.g. [16]). More precisely, every eigenvalue λ_n of (5) is simple and any corresponding eigenfunction has exactly n nodal domains (i.e. $n - 1$ zero points between 0 and 1).

Let λ_n be given by (5) for $k = n$. Then the following three statements are proved in [18]:

- (i) *Suppose $\lambda < \lambda_{n+1}$ is an eigenvalue and u_λ is an eigenfunction associated to λ . Then u_λ has at most n nodal domains.*
- (ii) *Let u_{λ_n} be an eigenfunction associated with λ_n . Then u_{λ_n} has at most $2n - 2$ nodal domains.*

The last statement can be strengthened if $-\Delta_p$ satisfies the following *unique continuation property* (UCP): “If u_λ is a nontrivial eigenfunction of Δ_p , then the set $\{x \in \Omega : u_\lambda(x) = 0\}$ has an empty interior.” Then we also have:

(iii) Suppose that $-\Delta_p$ satisfies (UCP) and suppose that u_{λ_n} is an eigenfunction associated with λ_n . Then u_{λ_n} has at most n nodal domains.

The last statement is a generalization of the *Courant nodal domain theorem* (known for $p = 2$). So, we arrived at the following

OPEN PROBLEM 2.3. Prove that $-\Delta_p$ satisfies (UCP), or else find a counterexample.

3 The Fredholm alternative

Let us consider the problem

$$\begin{cases} -\Delta_p u - \lambda |u|^{p-2} u = f(x) & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \quad (6)$$

with $f \in L^\infty(\Omega)$ and a spectral parameter $\lambda \in \mathbb{R}$. It is an easy application of the Schauder's fixed point theorem to prove that (6) has at least one (weak) solution for any $f \in L^\infty(\Omega)$ provided that λ is not an eigenvalue of (3).

If λ is an eigenvalue and $p = 2$, then the *linear Fredholm alternative* states that (6) is solvable if and only if

$$\int_{\Omega} f u_{\lambda} dx = 0$$

for all eigenfunctions u_{λ} associated to λ .

If λ is an eigenvalue of (3) and $p \neq 2$ the situation is far from being well understood. There are several reasons for this claim. First of all, the structure of the set of eigenvalues of (3) is not known if $N \geq 2$ (cf. Section 2). But an analogue of the Fredholm alternative is not available even in one dimension $N = 1$, although we know that the eigenvalues form a sequence $\{\lambda_k\}_{k=1}^{\infty}$, with λ_k given by (4) and (5). However, some progress has been made recently in the case of the first eigenvalue λ_1 .

Let us formulate here the *nonlinear Fredholm alternative* at the first eigenvalue of the p -Laplacian for $p > 1$, $p \neq 2$. It concerns the solvability of (6) with $\lambda = \lambda_1$. Let us split $f \in L^\infty(\Omega)$ as $f = f^\top + \bar{f}$, where $\bar{f} \in \mathbb{R}$ and $\int_{\Omega} f^\top \varphi_1 dx = 0$. Our problem then takes the form

$$\begin{cases} -\Delta_p u - \lambda_1 |u|^{p-2} u = f^\top(x) + \bar{f} & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \quad (7)$$

The following assertions hold:

Let $p \neq 2$, $p > 1$, $f^\top \not\equiv 0$. There exist real numbers $a_1 = a_1(f^\top)$ and $a_2 = a_2(f^\top)$ such that $a_1 < 0 < a_2$ and problem (7) has

- (i) at least one solution if $\bar{f} \in [a_1, a_2]$;
- (ii) at least two distinct solutions if $\bar{f} \in (a_1, 0) \cup (0, a_2)$;

(iii) *no solution if $\bar{f} \notin [a_1, a_2]$.*

These statements are proved in DRÁBEK [11] but there were several previous results in this direction: the one-dimensional case is treated in DEL PINO, DRÁBEK and MANÁSEVICH [8], DRÁBEK, GIRG and MANÁSEVICH [13], the special case $1 < p < 2$ is studied in DRÁBEK and HOLUBOVÁ [15] and the general case is also investigated in TAKÁČ [27], [28].

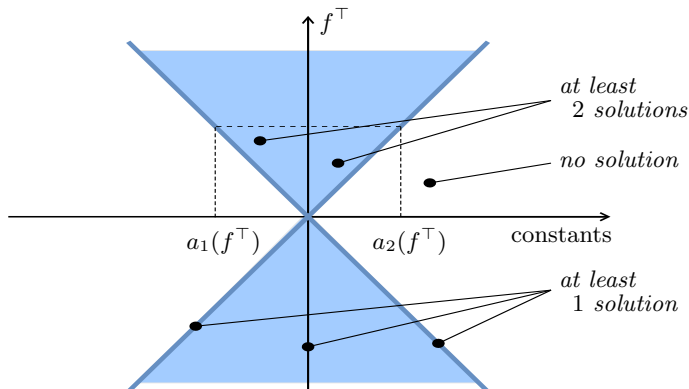


Figure 3.1: Nonlinear Fredholm alternative at the first eigenvalue.

Let us point out that every “slice” of the space of right hand sides f determined by a given f^\top and a one-dimensional subspace of constant functions can be depicted as in Fig. 3.1. In fact, even more information is hidden in the proof of the nonlinear Fredholm alternative and one can find it in the papers mentioned above. Namely, if we denote $F := \{f \in L^\infty(\Omega) : (7) \text{ has a solution}\}$ and $S_f := \{u \in W_0^{1,p}(\Omega) : u \text{ is a solution to (7) for given } f \in F\}$, then the following statements hold:

- (iv) $\text{int } F \neq \emptyset$ (more precisely, given $f \in F$, with $\int f \varphi_1 dx \neq 0$, there exists $\rho > 0$ such that $\|h - f\|_{L^\infty} < \rho$ implies $h \in F$);
- (v) For a given $f \in F$ with $f \not\equiv 0$ there exists $R = R(f) > 0$ such that for any $u \in S_f$ we have $\|u\|_{W_0^{1,p}} \leq R$.

In particular, the last two statements (iv) and (v) exhibit the striking difference between the linear ($p = 2$) and the quasilinear ($p \neq 2$) case.

To the best of the authors’ knowledge the unique attempt to analyze the situation at the second and higher eigenvalues has been MANÁSEVICH and TAKÁČ [25]. But only the one-dimensional case is treated there and the conditions are far from being necessary and sufficient. All these facts result in our next open problem:

OPEN PROBLEM 3.1. *Let λ be an eigenvalue of (3), $\lambda \neq \lambda_1$. Find necessary and sufficient conditions on $f \in L^\infty(\Omega)$ which guarantee that problem (6) has at least one solution (or more solutions).*

4 The variational approach

Let us note that

$$E_\lambda(u) = \frac{1}{p} \int_\Omega |\nabla u|^p dx - \frac{\lambda}{p} \int_\Omega |u|^p dx - \int_\Omega f u dx, \quad u \in W_0^{1,p}(\Omega),$$

is the energy functional associated with problem (6). Clearly, *critical points* of E_λ are in one-to-one correspondence with *weak solutions* of (6). The properties of the energy functional E_λ are well recognized in the case $p = 2$:

- (a) E_λ is coercive and strictly convex (i.e. E'_λ is strictly monotone) for $\lambda < \lambda_1$ and for arbitrary $f \in L^\infty(\Omega)$; for any given $f \in L^\infty(\Omega)$ there is a unique critical point of E_λ ;
- (b) If $\lambda > \lambda_1$, $\lambda \neq$ eigenvalue of (3), then E_λ has a geometry of saddle point and for any given $f \in L^\infty(\Omega)$ there is unique critical point of E_λ ;
- (c) If $\lambda =$ eigenvalue of (3), there are either no critical points of E_λ (if $\int_\Omega f u_\lambda dx \neq 0$ for some eigenfunction u_λ associated to λ), or infinitely many critical points of E_λ (if $\int_\Omega f u_\lambda dx = 0$ for all eigenfunctions u_λ associated to λ). All critical points then form a linear manifold which has the same dimension as the eigenspace corresponding to λ ;
- (d) For $\lambda = \lambda_1$ and $f \in L^\infty(\Omega)$, $\int_\Omega f \varphi_1 dx = 0$ (φ_1 is the first eigenfunction, cf. Section 2), E_λ is bounded from below.

Very different properties of the energy functional E_λ if $p \neq 2$ were observed in several papers: DEL PINO, ELGUETA and MANÁSEVICH [9], FLECKINGER et al. [20], DRÁBEK and TAKÁČ [19] and also in the above mentioned papers [8], [11], [15], etc. To point out the differences with respect to (a)–(d) we stress some properties of E_λ , $p \neq 2$, as follows:

- (α) E_λ is coercive and strictly convex for $\lambda \leq 0$, for $0 < \lambda < \lambda_1$ it is only coercive but not convex in general. Accordingly, the critical points of E_λ are not unique, in general: one can find $f \in L^\infty(\Omega)$ for which besides a global minimizer of E_λ there also exist saddle points;
- (β) If $\lambda > \lambda_1$, $\lambda \neq$ eigenvalue of (3), one can also find $f \in L^\infty(\Omega)$ for which E_λ has more than one critical point;
- (γ) If $\lambda = \lambda_1$, $f \in L^\infty(\Omega)$, $f = f^\top + \bar{f}$, $\int_\Omega f^\top \varphi_1 dx = 0$, $\bar{f} \in \mathbb{R}$, the statements (i)–(iii) in Section 3 formulated for the solutions of (7) apply for the critical points of E_{λ_1} ; moreover, for any given $f \in L^\infty(\Omega)$ all critical points of E_{λ_1} are bounded in the norm of $W_0^{1,p}(\Omega)$;
- (δ) Let $f \in L^\infty(\Omega)$, $\int_\Omega f \varphi_1 dx = 0$, $f \neq 0$, then for $p > 2$ the functional E_{λ_1} is bounded from below in $W_0^{1,p}(\Omega)$ while, for $1 < p < 2$, E_{λ_1} is unbounded from below in $W_0^{1,p}(\Omega)$.

Regarding the geometry of the energy functional there are still many open questions. Let us formulate some of them.

OPEN PROBLEM 4.1. For a given $f \in L^\infty(\Omega)$, $\int_\Omega f \varphi_1 dx = 0$, study the level sets of the functional E_{λ_1} , i.e. its asymptotic properties, etc.

OPEN PROBLEM 4.2. For given $f \in L^\infty(\Omega)$ and $\lambda = \text{eigenvalue of (3)}$, find suitable linking geometrical properties of E_λ which imply the existence of critical points.

IMPORTANT OBSERVATION! It is interesting to point out that the classical result of LANDESMAN and LAZER [24] for (perturbed) semilinear problems

$$\begin{cases} -\Delta u - \lambda_k u + g(u) &= f(x) & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega, \end{cases} \quad (8)$$

where λ_k is an eigenvalue and $g : \mathbb{R} \rightarrow \mathbb{R}$ is a bounded, continuous function with limits $g(\pm\infty) = \lim_{s \rightarrow \pm\infty} g(s)$, $g(-\infty) \neq g(+\infty)$, appeared about 70 years after I. FREDHOLM established (in 1900) the necessary and sufficient condition which characterizes the solvability of (8) with $g \equiv 0$. On the other hand, it was proved in DRÁBEK and ROBINSON [17] that the result in [24] extends to

$$\begin{cases} -\Delta_p u - \lambda |u|^{p-2} u + g(u) &= f & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega, \end{cases} \quad (9)$$

where $\lambda = \text{eigenvalue of (3)}$ (variational, i.e. provided by a variational characterization like (4) or (5), or not) and $p > 1$ is arbitrary. The variational approach was employed here and the structure of variational eigenvalues $\{\lambda_k\}_{k=1}^\infty$ given by (5) played an essential role. So, the situation in the quasilinear case ($p \neq 2$) seems to be the opposite to the linear case: the perturbed problem (9) was solved before an unperturbed problem (7) and, moreover, it was in a certain sense easier to solve (9) than (7), at least from the variational point of view.

5 The bifurcation approach

In this section we want to address another approach to problem (6) based on *bifurcation theory*. Let us start with a transformation which sends “large” solutions of (6) to “small” solutions of a “new” problem where local bifurcation results can be applied. First, we set

$$(J(u), w) = \int_\Omega |\nabla u|^{p-2} \langle \nabla u, \nabla w \rangle dx,$$

$$(S(u), w) = \int_\Omega |u|^{p-2} u w dx,$$

$$(F, w) = \int_{\Omega} fw \, dx$$

for $u, w \in W_0^{1,p}(\Omega)$. Here (\cdot, \cdot) is the duality pairing between $W_0^{1,p}(\Omega)$ and its dual space and $\langle \cdot, \cdot \rangle$ is the scalar product in \mathbb{R}^N . We will denote by

$$\|u\| = \left(\int_{\Omega} |\nabla u|^p \, dx \right)^{1/p}$$

the usual norm in $W_0^{1,p}(\Omega)$.

Clearly, the operator equation

$$J(u) - \lambda S(u) = F, \quad u \in W_0^{1,p}(\Omega), \quad (10)$$

stands for the weak formulation of (6). Now, for $u \neq 0$ let $v := u/\|u\|^2$ and

$$G(v) := \begin{cases} \|v\|^{2(p-1)}F & \text{for } v \neq 0, \\ 0 & \text{for } v = 0. \end{cases}$$

Then (10) is equivalent to

$$J(v) - \lambda S(v) = G(v) \quad v \in W_0^{1,p}(\Omega). \quad (11)$$

Since

$$\lim_{\|v\| \rightarrow 0} \frac{G(v)}{\|v\|^{p-1}} = 0,$$

standard bifurcation results can be applied to (11); see DRÁBEK [12] for the details. Actually, it is proved in [12] that $(\lambda_1, 0)$ is a global bifurcation point for (11). Going back to the original equation (10), we obtain that (λ_1, ∞) is an *asymptotic bifurcation point* for (10) (i.e. there exists a sequence $\{(\lambda_n, u_n)\}_{n=1}^{\infty} \subset \mathbb{R} \times W_0^{1,p}(\Omega)$ such that (10) holds with $\lambda = \lambda_n$ and $u = u_n$, and $(\lambda_n, \|u_n\|) \rightarrow (\lambda_1, \infty)$) and, moreover, there are maximal sets $\mathcal{C}^+, \mathcal{C}^- \subset \mathbb{R} \times W_0^{1,p}(\Omega)$ of solutions to (10) such that there exist sequences of pairs $(\mu_n, u_n) \in \mathcal{C}^+$ and $(\hat{\mu}_n, \hat{u}_n) \in \mathcal{C}^-$ with $\|\hat{u}_n\| \rightarrow \infty$, together with $u_n/\|u_n\| \rightarrow \varphi_1/\|\varphi_1\|$ and $\hat{u}_n/\|\hat{u}_n\| \rightarrow -\varphi_1/\|\varphi_1\|$. However, this information is not enough to get reasonable existence and multiplicity results for (10).

On the other hand, asymptotic bifurcation combined with suitable asymptotics yield *a priori* estimates of the solutions to (10) and, consequently, more accurate information about the sets \mathcal{C}^{\pm} for $\|u\|$ “large”. These results are published in DRÁBEK, GIRG, TAKÁČ and ULM [14] and the bifurcation diagrams in Fig. 5.1 to Fig. 5.4 depict the character of \mathcal{C}^{\pm} with $\|u\|$ “large”. We shall split $f \in L^{\infty}(\Omega)$ as $f = f^{\top} + a\varphi_1$, $\int_{\Omega} f^{\top} \varphi_1 \, dx = 0$, $a \in \mathbb{R}$, and write (6) in the form

$$\begin{cases} -\Delta_p u - \lambda|u|^{p-2}u & = f^{\top} + a\varphi_1 & \text{in } \Omega, \\ u & = 0 & \text{on } \partial\Omega. \end{cases} \quad (12)$$

Solutions to (12) will be written in the form $u = c\varphi_1 + v^{\top}$, with $\int_{\Omega} v^{\top} \varphi_1 \, dx = 0$.

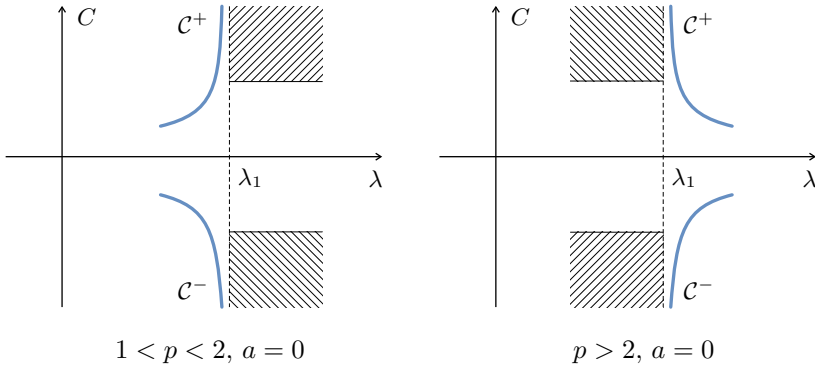


Figure 5.1

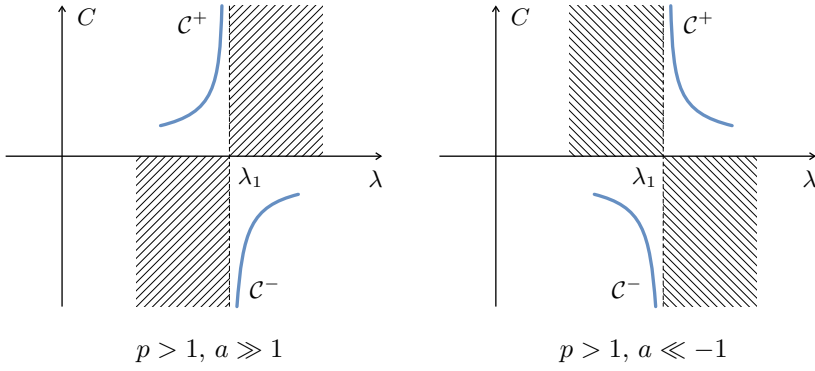


Figure 5.2

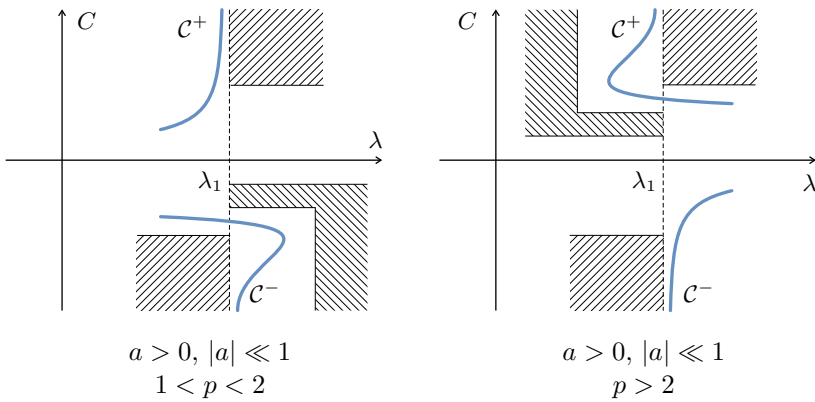


Figure 5.3

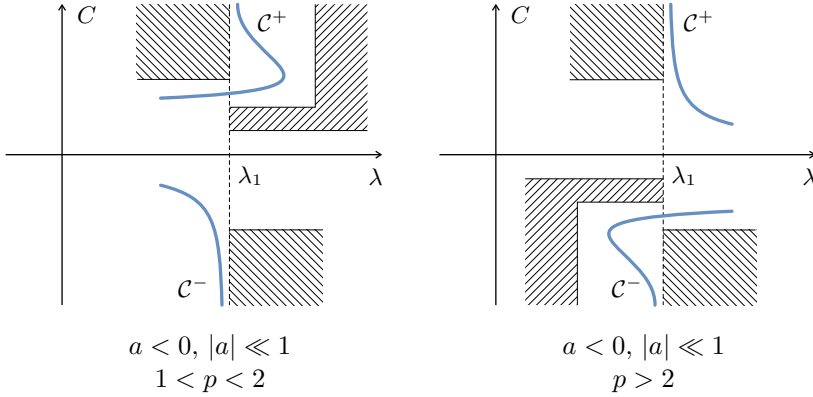


Figure 5.4

Let us explain how these results relate e.g. to the solvability of (12) with $\lambda = \lambda_1$, $a = 0$ and $1 < p < 2$. Due to the isolatedness of λ_1 and *a priori* estimates (see Fig. 5.1), there is a sequence $\{\mu_k\}_{k=1}^\infty$, $\mu_k > \lambda_1$, $\mu_k \rightarrow \lambda_1$ as $k \rightarrow \infty$ and a sequence $\{u_k\}_{k=1}^\infty \subset W_0^{1,p}(\Omega)$ such that $\|u_k\| \leq \text{const.}$ (independent of k) and (μ_k, u_k) solves (12) with $a = 0$ (the existence of u_k follows from the fact that the μ_k are not equal to the eigenvalue of (3)). The existence of a solution to (12) with $\lambda = \lambda_1$, $a = 0$ now follows from a standard compactness argument and passage to the limit as $k \rightarrow \infty$.

Combined with the method of upper and lower solutions, asymptotic estimates and the bifurcation approach provide alternative proofs of statements similar to (i)-(v) in Section 3 and, moreover, lead to the following new existence and multiplicity results:

- (I) Let $f^\top \neq 0$ be given. There exists $\varepsilon > 0$ such that
- (i) For every $\varepsilon' \in (0, \varepsilon)$ there exists $\nu = \nu(f^\top, \varepsilon, \varepsilon') > 0$ such that $\varepsilon' < |a| < \varepsilon$ and $\lambda \in (\lambda_1 - \nu, \lambda_1) \cup (\lambda_1, \lambda_1 + \nu)$ imply that (12) has at least three distinct solutions of which at least one is positive and at least one is negative;
 - (ii) $\lambda = \lambda_1$ and $0 < |a| < \varepsilon$ imply that problem (12) has at least two solutions, of which at least one is negative if $(p - 2)a < 0$, and at least one is positive if $(p - 2)a > 0$.

Moreover, there exists $\hat{\nu} = \hat{\nu}(f^\top) > 0$ such that, for $a = 0$, problem (12) has at least three distinct solutions (among them at least one is positive and one is negative) provided either $1 < p < 2$ and $\lambda \in (\lambda_1 - \hat{\nu}, \lambda_1)$ or $p > 2$ and $\lambda \in (\lambda_1, \lambda_1 + \hat{\nu})$.

- (II) Let $f^\top \neq 0$ be given. There exist $a_0 = a_0(f^\top) > 0$ and $\delta = \delta(f^\top) > 0$ such that
- (i) If either $\lambda \in (\lambda_1 - \delta, \lambda_1)$ and $a \geq a_0$, or $\lambda \in (\lambda_1, \lambda_1 + \delta)$ and $a \leq -a_0$, then (12) can only have positive solutions;

- (ii) If either $\lambda \in (\lambda_1 - \delta, \lambda_1)$ and $a \leq -a_0$, or $\lambda \in (\lambda_1, \lambda_1 + \delta)$ and $a \geq a_0$, then (12) can only have negative solutions.

6 The p -Laplacian in one dimension versus higher dimension

It is of no surprise that the case of the one-dimensional zero-Dirichlet p -Laplacian is much easier to handle from the technical point of view and, therefore, one can prove more fine results using bifurcation theory. In particular, in the one-dimensional case ($\Omega = (a, b)$, $a < b$), the continua \mathcal{C}^+ and \mathcal{C}^- become continuous curves in $W_0^{1,p}(a, b)$ in a sufficiently small neighbourhood of the point (λ_1, ∞) . We will explain this with more details, since it leads to interesting open problems.

Let us consider the following boundary value problem

$$\begin{cases} -(|u'|^{p-2}u')' - \lambda|u|^{p-2}u = f(x) & \text{in } (0, 1), \\ u(0) = u(1) = 0, \end{cases} \quad (13)$$

where $p > 1$, λ is a real number (the bifurcation parameter) and $f \in L^\infty(0, 1)$. In particular, we are interested in the values $\lambda < \lambda_2$, where λ_2 is the second eigenvalue of the associated problem

$$\begin{cases} -(|u'|^{p-2}u')' - \lambda|u|^{p-2}u = 0 & \text{in } (0, 1), \\ u(0) = u(1) = 0. \end{cases} \quad (14)$$

Note that the set of all eigenvalues of (14) can be written as a sequence $\{\lambda_k\}_{k=1}^\infty$, with

$$\lim_{k \rightarrow \infty} \lambda_k = +\infty, \quad (15)$$

each eigenvalue being simple and the corresponding eigenfunction φ_k changing sign exactly $(k-1)$ -times.

In GARCÍA-MELIÁN and SABINA DE LIS [23], bifurcation from zero at the radial eigenvalues of

$$\begin{cases} -\Delta_p u - \lambda|u|^{p-2}u = g(\lambda, u, x) & \text{in } B_r(0), \\ u = 0 & \text{on } \partial B_r(0) \end{cases}$$

was studied, where $B_r(0)$ denotes a ball with radius r centered at the origin. Using the Implicit Function Theorem, a quasilinear variant of Crandall-Rabinowitz Local Bifurcation Theorem (see CRANDALL and RABINOWITZ [7]) was proved in [23]. In particular, it was proved that a curve of radially symmetric solutions emanates at each eigenvalue with a radially symmetric eigenfunction.

If we restrict ourselves to one-dimensional case, any ball is an interval and *viceversa*. Therefore, all solutions to (13) are radially symmetric in one-dimensional case. By using the standard techniques in the previous section, we can transform bifurcation from (λ_k, ∞) to bifurcation from $(\lambda_k, 0)$. In

particular, we obtain that the sets \mathcal{C}^+ and \mathcal{C}^- become *curves* for λ close to λ_1 and $\|u\|_{W_0^{1,p}}$ large enough. More precisely, we get the following assertion:

There exist $\varepsilon > 0$ and $\delta > 0$ (small enough) and continuous mappings $\mu : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ and $\mathcal{V}^\top : (-\varepsilon, \varepsilon) \rightarrow C_0^1(0, 1) \cap W^{1,p}(0, 1)^\top$, here $W^{1,p}(0, 1)^\top := \{u \in W^{1,p}(0, 1) : \int_0^1 u \varphi_1 dx = 0\}$, such that

$$\begin{aligned} & \mathcal{C}^+ \cap \{(\lambda, u) \in \mathbb{R} \times W_0^{1,p}(0, 1) : \lambda_1 - \delta < \lambda < \lambda_1 + \delta, \quad \|u\|_{W_0^{1,p}} > 2/\varepsilon\} \\ &= \{(\lambda(t), u(t)) : \lambda = \lambda_1 + \mu(t), \quad u(t) = t^{-1}(\varphi_1 + \mathcal{V}^\top(t)), \quad t \in (0, \varepsilon/2)\}, \\ & \mathcal{C}^- \cap \{(\lambda, u) \in \mathbb{R} \times W_0^{1,p}(0, 1) : \lambda_1 - \delta < \lambda < \lambda_1 + \delta, \quad \|u\|_{W_0^{1,p}} > 2/\varepsilon\} \\ &= \{(\lambda(t), u(t)) : \lambda = \lambda_1 + \mu(t), \quad u(t) = t^{-1}(\varphi_1 + \mathcal{V}^\top(t)), \quad t \in (-\varepsilon/2, 0)\}, \end{aligned}$$

$\lim_{t \rightarrow 0} \mathcal{V}^\top(t) = 0$ in $C^{1,\beta}([0, 1])$ for some $\beta \in (0, 1)$, $\lim_{t \rightarrow 0} |t|^{-p} t \mathcal{V}^\top(t) = V^\top$ strongly in \mathcal{D}_{φ_1} for $p > 2$ and strongly in $W_0^{1,2}(0, 1)$ if $1 < p < 2$ and

$$\begin{aligned} \mu(t) &= -|t|^{p-2} t \int_0^1 f \varphi_1 dx \\ &+ \frac{(p-2)}{2} |t|^{2(p-1)} \int_0^1 \left[\left| \frac{d\varphi_1}{dx} \right|^{p-2} \left(\frac{dV^\top}{dx} \right)^2 - \lambda_1 (p-1) \varphi_1^{p-2} (V^\top)^2 \right] dx \\ &+ (p-1) |t|^{2(p-1)} \left(\int_0^1 f \varphi_1 dx \right) \left(\int_0^1 \varphi_1^{p-1} V^\top dx \right) \\ &+ o\left(|t|^{2(p-1)}\right). \end{aligned} \tag{16}$$

Here, \mathcal{D}_{φ_1} is the completion of $C_0^\infty(0, 1)$ with respect to the norm

$$\|u\|_{\mathcal{D}_{\varphi_1}} = \left(\int_0^1 \left| \frac{d\varphi_1}{dx} \right|^{p-2} \left(\frac{du}{dx} \right)^2 dx \right)^{1/2}$$

and the limit function V^\top is a solution of

$$\begin{aligned} \frac{d}{dx} \left(\left| \frac{d\varphi_1}{dx} \right|^{p-2} \frac{dV^\top}{dx} \right) - \lambda_1 \varphi_1^{p-2} V^\top &= \frac{1}{p-1} \left(f - \varphi_1^{p-1} \int_0^1 (f \varphi_1) dx \right) \quad \text{in } (0, 1), \\ V^\top(0) &= 0, \quad V^\top(1) = 0, \\ \int_0^1 V^\top \varphi_1 dx &= 0, \end{aligned}$$

which is nothing but the linearization of the boundary value problem (13) at $(\lambda, u) = (\lambda_1, \varphi_1)$ written in the divergence form.

It is natural to ask if this result extends also to similar problems in higher dimensions.

OPEN PROBLEM 6.1. *Do the sets \mathcal{C}^+ and \mathcal{C}^- form a one dimensional continuum “near” the point (λ_1, ∞) ? More precisely, does there exist a constant $R > 0$ such that $(\lambda, u) \in \mathcal{C}^+(\mathcal{C}^-)$, $(\lambda, v) \in \mathcal{C}^+(\mathcal{C}^-)$, $\|u\|_{W_0^{1,p}} > R$ and $\|v\|_{W_0^{1,p}} > R$ imply that $u = v$ and the set $\mathcal{C}^+ \cap \{u \in W_0^{1,p}(\Omega) : \|u\|_{W_0^{1,p}} < R\}$ ($\mathcal{C}^- \cap \{u \in W_0^{1,p}(\Omega) : \|u\|_{W_0^{1,p}} < R\}$) is connected?*

It is well known from the strict monotonicity of $\mathcal{I} - \lambda\mathcal{S}$ for $\lambda < 0$ (see e.g. [12]) that for each $\lambda < 0$ there exists exactly one $u_\lambda \in W_0^{1,p}(\Omega)$ such that (λ, u_λ) is a solution to (6). A natural question arises then, *whether the map U defined by $U : \mathbb{R}^- \cup \{0\} \mapsto W_0^{1,p}(\Omega)$, $U(\lambda) = u_\lambda$, is continuous.* If so, there is a connected set of solutions “arriving” from $-\infty$ in the λ -direction, which invokes another question in turn: *What happens toh this set when we cross $\lambda = 0$ and let increase the spectral parameter λ ?*

For convenience, we shall now define the sets

$$\begin{aligned}\mathcal{E}_\lambda^0 &:= \{(\nu, u) \in \mathbb{R} \times W_0^{1,p}(\Omega) : \nu = \lambda\}, \\ \mathcal{E}_\lambda^- &:= \{(\nu, u) \in \mathbb{R} \times W_0^{1,p}(\Omega) : \nu < \lambda\}, \\ \mathcal{E}_\lambda^+ &:= \{(\nu, u) \in \mathbb{R} \times W_0^{1,p}(\Omega) : \nu > \lambda\}.\end{aligned}$$

Global behaviour of the continua \mathcal{C}^+ and \mathcal{C}^- is another very important issue. Let us introduce

$$\mathcal{S} := \{(\lambda, u) \in \mathbb{R} \times W_0^{1,p}(\Omega) : J(u) - \lambda S(u) = F\}.$$

Then the following variant of the Leray continuation theorem was proved in [6]:

The set \mathcal{S} contains a continuum \mathcal{T} (a connected set that is maximal in the sense of inclusions) which satisfies

- (i) $(\mathcal{S} \cap \mathcal{E}_0^-) \subset (\mathcal{T} \cap \mathcal{E}_0^-)$, i.e. $(\mathcal{S} \cap \mathcal{E}_0^-) = (\mathcal{T} \cap \mathcal{E}_0^-)$;
- (ii) *Either $(\mathcal{T} \cap \mathcal{E}_0^+)$ is unbounded in the λ -direction, or there exists $\lambda_e \in \mathbb{R}$ such that for all $\varepsilon > 0$ the set $\mathcal{T} \cap \mathcal{E}_{\lambda_e - \varepsilon}^+ \cap \mathcal{E}_{\lambda_e + \varepsilon}^-$ is unbounded in $\mathbb{R} \times W_0^{1,p}(\Omega)$, λ_e being an eigenvalue of $-\Delta_p$.*

Let us define

$$\mathcal{C}_0^+ := \mathcal{E}_0^+ \cap \mathcal{T}, \quad \mathcal{C}_0^- := \mathcal{E}_0^- \cap \mathcal{T}.$$

Combining this result and result II(i) (or II(ii)) in p. 23, we immediately have that $\mathcal{C}_0^+ \cap \mathcal{C}^+ \neq \emptyset$ (or $\mathcal{C}_0^+ \cap \mathcal{C}^- \neq \emptyset$); see Fig. 6.1. Cases II(i) (i.e. $a < -a_0$) and II(ii) (i.e. $a > a_0$) are extremal. It is not clear, however, how does the picture evolve when a moves in the interval $(-a_0, a_0)$. We will return to this question in the next section.

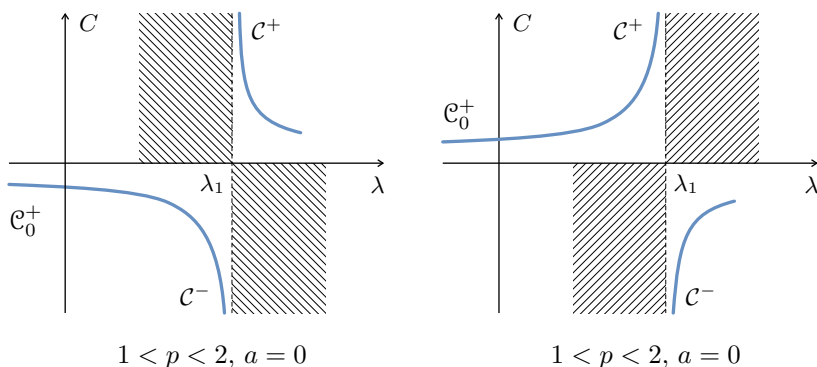


Figure 6.1: Cases II(i), i.e. $a < -a_0$ and II(ii), i.e. $a > a_0$, $p > 1$. Here the relation between \mathcal{C}_0^+ , \mathcal{C}^+ , \mathcal{C}^- is known as depicted here.

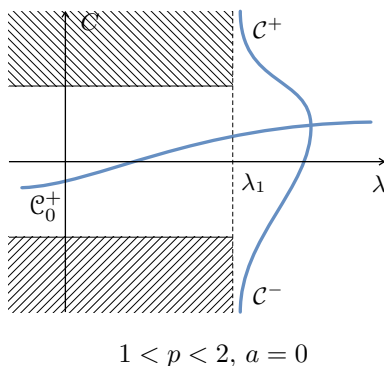


Figure 6.2: It is clear that for $p > 2$ and $a = 0$, $\mathcal{C}_0^+ \cap (\{0\} \times W_0^{1,p}(\Omega)) \neq \{0\}$.

OPEN PROBLEM 6.2. *Is the intersection $\mathcal{C}^+ \cap \mathcal{C}^- \cap \mathcal{C}_0^+$ nonempty for some $p \neq 2$ and $a \in \mathbb{R}$?*

OPEN PROBLEM 6.3. *Assuming that the intersection $\mathcal{C}^+ \cap \mathcal{C}^- \cap \mathcal{C}_0^+$ is nonempty for some $p \neq 2$ and $a \in \mathbb{R}$, do we then have a pitchfork bifurcation as indicated in Fig. 6.3 ?*

Many other questions arise in the context of the global behaviour of the continua \mathcal{C}^+ , \mathcal{C}^- , \mathcal{C}_0^+ . We kindly refer the reader to his own fantasy.

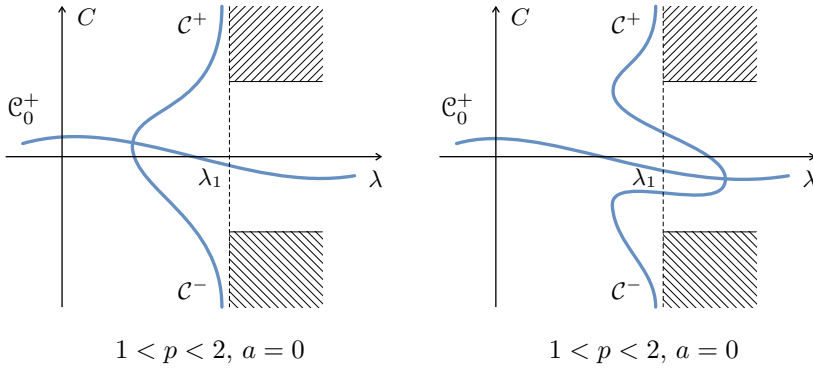


Figure 6.3: Some possible pitchfork bifurcation diagrams for $1 < p < 2$ and $a = 0$. There is no solution in the shaded regions (owing to a priori bounds).

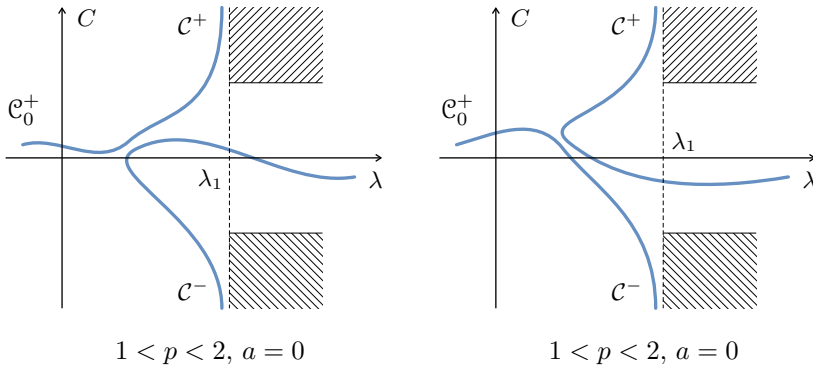


Figure 6.4: Some possible imperfect pitchfork bifurcation diagrams for $1 < p < 2$ and $a = 0$. There is no solution in the shaded regions (owing to a priori bounds).

7 Numerical experiments

There are many properties of problems involving the p -Laplacian which are difficult “to guess”, i.e. one does not know *a priori* whether to look for a rigorous proof or a counterexample (cf. for instance Open Problem 2.2) From this point of view, *numerical experiments* are an important tool in order to understand better “what is going on”. Some numerical experiments for the one-dimensional p -Laplacian are addressed in ČEPIČKA, DRÁBEK AND GIRG [6].

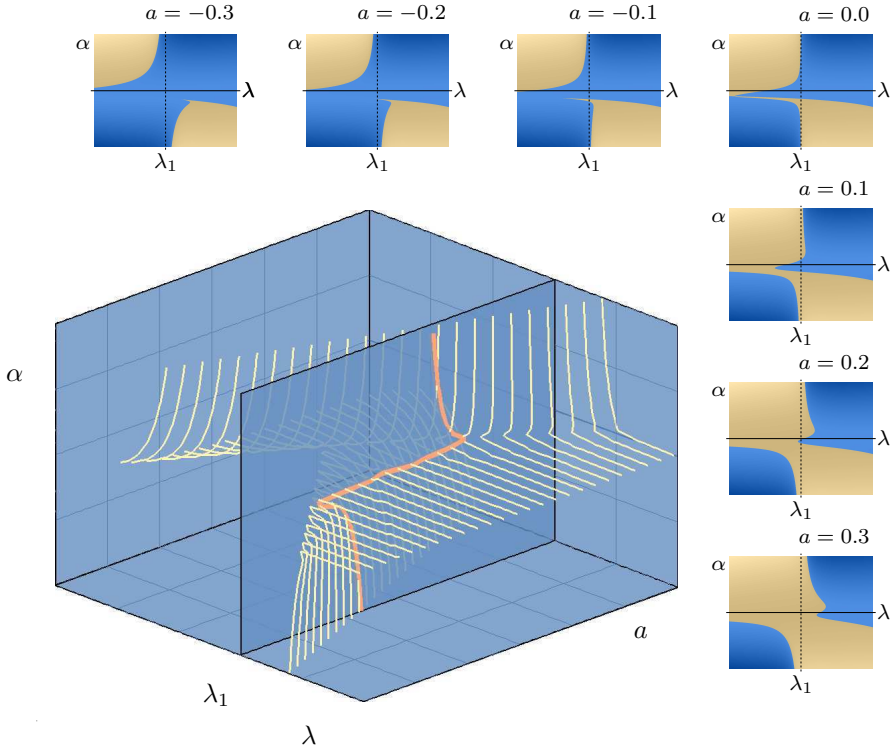


Figure 7.1: Bifurcation diagrams for the Dirichlet boundary value problem $(|u'|^{p-2}u')' + \lambda|u|^{p-2}u = a \sin \pi t + \sin 2\pi t$, $t \in (0, 1)$, $u(0) = u(1) = 0$ and $a \in \mathbb{R}$. The vertical plane ϱ located approximately in the middle of the scene is determined by the equation $\lambda = \lambda_1$. We hope that the intersections of ϱ and the continua \mathcal{C}_a^+ and \mathcal{C}_a^- which dive into the half space $\lambda < \lambda_1$ form a curve. This curve — carefully calculated by continuation methods — is depicted on ϱ .

Let us consider the one-dimensional p -Laplace equation

$$\begin{cases} -(|u'|^{p-2}u')' - \lambda|u|^{p-2}u = f, & t \in (0, 1), \\ u(0) = 0, & u(1) = 0, \end{cases} \quad (17)$$

where $p > 1$, $\lambda > 0$ and $f \in L^\infty(0, 1)$. We apply a shooting method to solve the boundary value problem (17) and define a function $M_f : \mathbb{R}^3 \mapsto \mathbb{R}$ as follows: $M_f(\alpha, \lambda, p) = u_\alpha(1)$, where $u_\alpha \in L^\infty(0, 1)$, $u_\alpha(0) = 0$ and $u'_\alpha(0) = \alpha$ is a solution of the associated initial value problem corresponding to the equation in (17). The *bifurcation diagram* is a set of zero points of M_f . In Fig. 7.2 we explain how to construct the *bifurcation diagram*. In all numerical experiments, we have considered $f(t) = \sin(2\pi t) + a\sin(\pi t)$. Note that while $\int_0^1 \sin(2\pi t) \varphi_1(t) dt = 0$, because of the symmetry of φ_1 with respect to $1/2$, the function $\sin(\pi t)$ contains nontrivial components orthogonal to φ_1 . This is

however negligible from the point of view of numerical computations and we therefore use $\sin(\pi t)$ instead of φ_1 .

The bifurcation diagrams corresponding to the boundary value problem

$$\begin{cases} -(|u'|^{p-2}u)' - \lambda|u|^{p-2}u &= f^\top + a \sin(\pi t) & \text{in } (0, 1), \\ u(0) = u(1) &= 0 \end{cases} \quad (18)$$

with $f^\top \in L^\infty(0, 1)$ satisfying $\int_0^1 f \varphi_1 dx = 0$, the spectral parameter $\lambda \in \mathbb{R}$ and various choices of the parameter $a \in \mathbb{R}$ can be collected together in one 3D-scene as we can see in Fig. 7.1. Each point in this 3D-scene is determined by the coordinates (a, λ, α) , with $\alpha = u'(0)$. Note that this scene can easily be transformed into one with coordinates (a, λ, c) , where $c = \int_0^1 u \varphi_1 dx$. Then, this scene suggests immediately the following question:

OPEN PROBLEM 7.1. *Let for each $a \in \mathbb{R}$, $\mathcal{C}_a^+, \mathcal{C}_a^- \subset \mathbb{R} \times \mathbb{R} \times W_0^{1,p}(\Omega)$ represent continua of solutions to*

$$\begin{cases} -\Delta_p u - \lambda|u|^{p-2}u &= f^\top + a\varphi_1 & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega \end{cases} \quad (19)$$

respectively emanating from $+\infty$ and $-\infty$; here again $f \in L^\infty(\Omega)$, $\int_\Omega f \varphi_1 dx = 0$, $a \in \mathbb{R}$ and $\lambda \in \mathbb{R}$ is a spectral parameter. Let moreover $\varrho \subset \mathbb{R} \times \mathbb{R} \times W_0^{1,p}(\Omega)$ be the hyperplane given by $\lambda = \lambda_1$. Does the set

$$((\{a\} \times \mathcal{C}_a^+) \cup (\{a\} \times \mathcal{C}_a^-)) \cap \varrho$$

contain a continuum $\xi \subset \varrho$ such that $\text{proj}_c(\xi) = \mathbb{R}$? Here, $\text{proj}_c : \mathbb{R} \times \mathbb{R} \times W_0^{1,p}(\Omega) \mapsto \mathbb{R}$, with $\text{proj}_c[(a, \lambda, u)] = \int_\Omega u \varphi_1$.

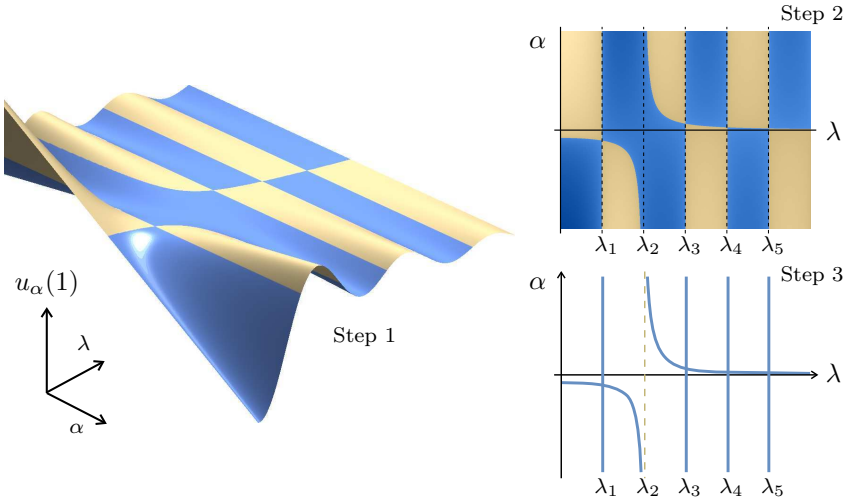


Figure 7.2: Given $p \in \mathbb{R}$, we compute the graph of the function M_f in Step 1. Level sets of $M_f(\alpha, \lambda, p)$ are displayed in Step 2. Finally, we extract solutions $(\alpha, \lambda) \in \mathbb{R}^2$ of the nonlinear equation $M_f(\alpha, \lambda, p) = 0$ in Step 3.

Note that in TAKÁČ [28] *a priori* asymptotic estimates of large solutions (λ, u) to (19) with $\lambda = \lambda_1$ were established.

We would like to close our paper with an exposition of several bifurcation diagrams which demonstrate the complexity of the behaviour of the solution set to the boundary value problem

$$\begin{cases} -(|u'|^{p-2}u')' - \lambda|u|^{p-2}u + g(u) = f, & t \in (0, 1), \\ u(0) = 0, & u(1) = 0. \end{cases} \quad (20)$$

In the following figures we consider problem (20) with some special f, g and p .

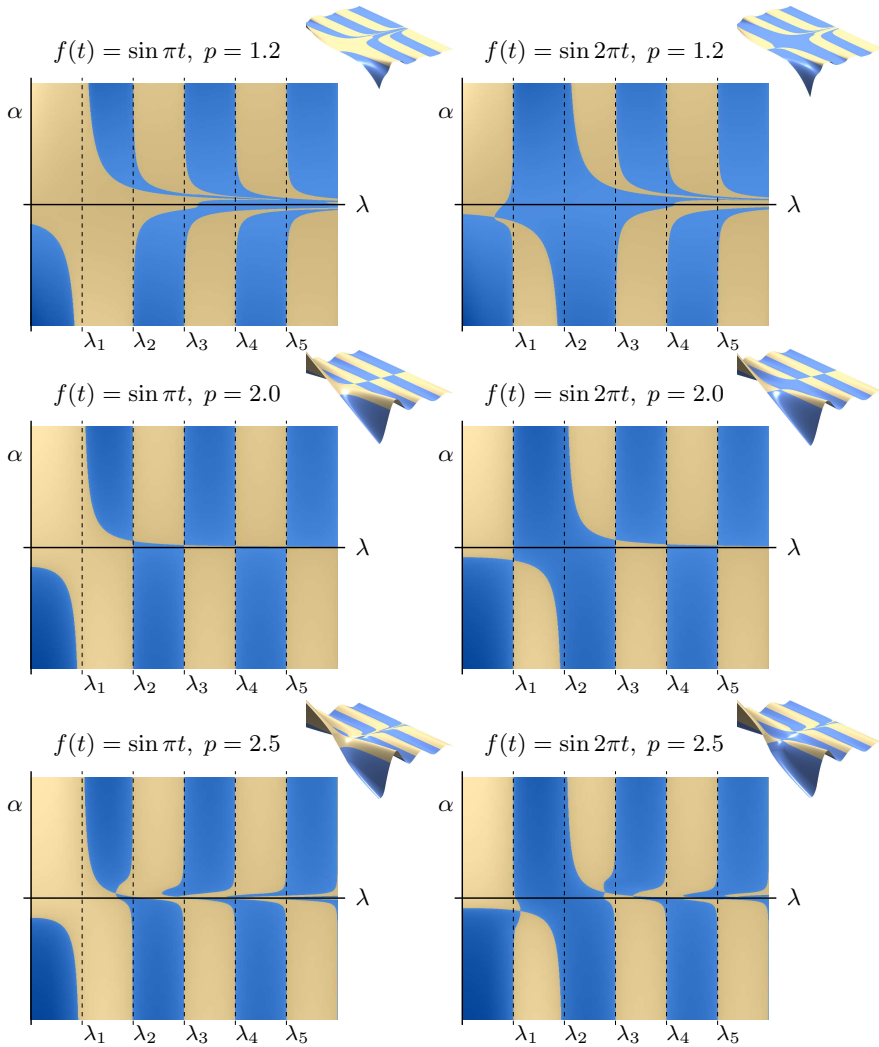


Figure 7.3: Bifurcation diagrams for equation (20) with $g(u) \equiv 0$.

See <http://drift.cam.zcu.cz> (global server for numerical experiments computed in Centre of Applied Mathematics) for these and many other related numerical results.

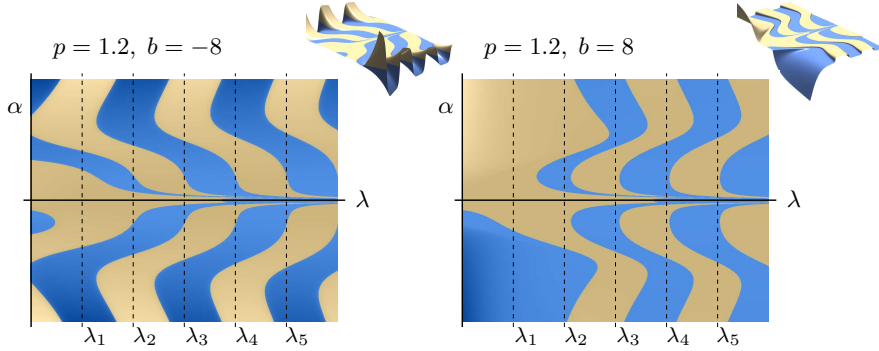


Figure 7.4: Bifurcation diagrams for the Dirichlet boundary value problem $(|u'|^{p-2}u')' + \lambda|u|^{p-2}u - b \operatorname{atan} u = \sin \pi t$, $t \in (0, 1)$, $u(0) = u(1) = 0$ and $b \in \mathbb{R}$.

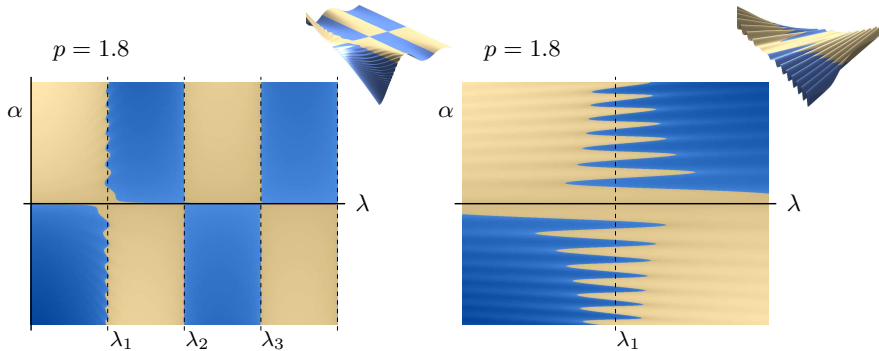


Figure 7.5: Bifurcation diagrams for the Dirichlet boundary value problem $(|u'|^{p-2}u')' + \lambda|u|^{p-2}u - |u|^{p-2}u \sin u = \sin \pi t$, $t \in (0, 1)$ and $u(0) = u(1) = 0$.

References

- [1] A. ANANE, *Simplicité et isolation de la première valeur propre du p -laplacien avec poids*, Comptes Rendus Acad. Sc. Paris Série I, **305** (1987), pp. 725–728.
- [2] R. COURANT and D. HILBERT, *Methods of Mathematical Physics*, Vol. I, Interscience Publ. INC, New York - London, 1953.
- [3] M. CUESTA, *On the Fučík spectrum of the Laplacian and the p -Laplacian*, in: Proceedings of Seminar in Differential Equations (P. Drábek Ed.), University of West Bohemia, Pilsen 2000, pp. 67–96.

- [4] M. CUESTA, D. G. DE FIGUEIREDO and J.-P. GOSSEZ, *A nodal domain property for the p -Laplacian*, C. R. Acad. Sci., Paris, Ser I Math. **330**, No. 8 (2000), 669-673.
- [5] M. CUESTA and P. TAKÁČ, *A strong comparison principle for positive solutions of degenerate elliptic equations*, Diff. and Int. Equations **13**, (2000), 721-746.
- [6] J. ČEPIČKA, P. DRÁBEK and P. GIRG, *Quasilinear boundary value problems: existence and multiplicity results*, Clay Math. Proceedings, to appear.
- [7] M. G. CRANDALL and P. RABINOWITZ, *Bifurcation from simple eigenvalues*, J. Funct. Anal. **8** (1971), pp. 321-340.
- [8] M. DEL PINO, P. DRÁBEK and R. F. MANÁSEVICH, *The Fredholm alternative at the first eigenvalue for the one-dimensional p -Laplacian*, J. Differential Equations, **151** (1999), pp. 386-419.
- [9] M. DEL PINO, M. ELGUETA and R. MANÁSEVICH, *A homotopic deformation along p of a Leray-Schauder degree result and existence for $(|u'|^{p-2}u')' + f(t, u) = 0, u(0) = u(T) = 0, p > 1$* , J. Differential Equations **80** (1989), 1-13.
- [10] J. I. DÍAZ, *Nonlinear Partial Differential Equations and Free Boundary Value Problems*, Vol. I Elliptic Equations, Pitman Research Notes in Mathematics, Longman 1985.
- [11] P. DRÁBEK, *Geometry of the energy functional and the Fredholm alternative for the p -Laplacian in more dimensions*, Electronic J. Diff. Equations, Conference **08** (2002), pp. 103-120.
- [12] P. DRÁBEK, *Solvability and Bifurcations of Nonlinear Equations*, Pitman Research Notes in Mathematics Series, Vol. 264, Longman Scientific & Technical, Essex, 1992.
- [13] P. DRÁBEK, P. GIRG and R. F. MANÁSEVICH, *Generic Fredholm alternative - type results for the one dimensional p -Laplacian*, Nonlin. Diff. Equations and Applications **8** (2001), pp. 285-298.
- [14] P. DRÁBEK, P. GIRG, P. TAKÁČ and M. ULM, *The Fredholm alternative for the p -Laplacian: bifurcation from infinity, existence and multiplicity of solutions*, Indiana Univ. Math. J., to appear.
- [15] P. DRÁBEK, G. HOLUBOVÁ, *Fredholm alternative for the p -Laplacian in higher dimensions*, J. Math. Anal. Appl. **263** (2001), pp. 182-194.
- [16] P. DRÁBEK and R. MANÁSEVICH, *On the closed solution to some nonhomogeneous eigenvalue problems with p -Laplacian*, Differential and Integral Equations **12** (6) (1999), pp. 773-788.

- [17] P. DRÁBEK and S. ROBINSON, *Resonance problems for the p -Laplacian*, J. Funct. Anal. **169** (1999), pp. 189–200.
- [18] P. DRÁBEK and S. ROBINSON, *On the generalization of the Courant Nodal Domain Theorem*, Journal of Differential Equations **181** (2002), pp. 58–71.
- [19] P. DRÁBEK, P. TAKÁČ, *A counterexample to the Fredholm alternative for the p -Laplacian*, Proc. Amer. Math. Soc. **127** (1999), pp. 1079–1087.
- [20] A. ELBERT, *A half-linear second order differential equations*, in: Qualitative theory of differential equations, Coll. Math. Soc. J. Bolyai **30** (Szeged, 1979), pp. 153–179.
- [21] J. FLECKINGER, J. HERNANDEZ, P. TAKÁČ and F. DE THÉLIN, *Uniqueness and positivity for solutions of equations with the p -Laplacian*, in: G. Caristi and E. Mitidieri; eds., “Proceedings of the Conference on Reaction-Diffusion Equations”, 1995, Trieste, Italy. Lecture Notes in Pure and Applied Math., Vol **194**, pp. 141–155.
- [22] J. P. GARCÍA-AZORERO and I. PERAL, *Existence and nonuniqueness for the p -Laplacian: nonlinear eigenvalues*, Comm. Partial Differential Equations **12** (1987), pp. 1389–1430.
- [23] J. P. GARCÍA-MELIÁN and J. SABINA DE LIS, *A Local Bifurcation Theorem for Degenerated Elliptic Equations With Radial Symmetry*, J. Differential Equations **179** (2002), pp. 27–43.
- [24] E. M. LANDESMAN, A. C. LAZER, *Nonlinear perturbations of linear elliptic boundary value problems at resonance*, J. Math. Mech. **19** (1970), pp. 609–623.
- [25] R. F. MANÁSEVICH, P. TAKÁČ, *On the Fredholm alternative for the p -Laplacian in one dimension*, Proc. London Math. Soc. **84** (2002), pp. 324–342.
- [26] M. OTANI and T. TESHIMA, *On the first eigenvalue of some quasilinear elliptic equations*, Proc. Japan Acad. **64**, sèr. A (1998), pp. 8–10.
- [27] P. TAKÁČ, *On the Fredholm alternative for the p -Laplacian at the first eigenvalue*, Indiana Univ. Math. J. **51** (2002), pp. 187–237.
- [28] P. TAKÁČ, *On the number and structure of solutions for a Fredholm alternative with the p -Laplacian*, J. Differ. Equations **185** (2002), pp. 306–347.
- [29] J. L. VÁZQUEZ, *A strong maximum principle for some quasilinear elliptic equations*, Appl. Math. Optim. **12**, (1984), 191–192.

Codificación de información mediante códigos bidimensionales*

F. J. ESPINOSA¹, L. HERNÁNDEZ²
Y A. MARTÍN³

¹BearingPoint Software Solutions, S.L, Madrid

²Departamento de Tratamiento de la Información y Codificación,
Instituto de Física Aplicada, CSIC

³Departamento de Matemática Aplicada,
Universidad de Salamanca

javier.espinosa@bearingpoint.com
luis@iec.csic.es
delrey@usal.es

Resumen

En este artículo se presentan diferentes sistemas de codificación bidimensional para el tratamiento informático de determinado tipo de información. De forma concreta, se presentan las características y se muestra cómo puede elaborarse el código bidimensional llamado PDF417, posiblemente uno de los más utilizados hoy en día. Se incluyen y comentan los procesos matemáticos que se emplean para elaborar dicho código.

Palabras clave: *Codificación de la información, códigos bidimensionales, códigos correctores de errores, códigos de Reed-Solomon, PDF417.*

Clasificación por materias AMS: *94A05, 94A15, 94B60*

1 Introducción

Desde hace algunos años, las aplicaciones de la Informática y de las Matemáticas a la codificación se han extendido de forma creciente. Hoy en día se codifican los cheques bancarios y cheques de viaje para su control y para la detección

* Parcialmente subvencionado por el Ministerio de Ciencia y Tecnología, TIC2001–0586.
Fecha de recepción: 11 de febrero de 2004

de posibles falsificaciones; se catalogan libros mediante el código ISBN ([12]); se imprimen códigos de barras en los artículos de venta al por menor (códigos EAN13 [8], UPC [19], Codabar [1], entre otros), etc. Todos estos códigos tienen la posibilidad de corregir errores, caso de que se hayan producido bien a la hora de ser impresos, bien cuando son leídos por un escáner lineal, debido al dígito de control que incluyen como parte del propio código.

Por otra parte, la necesidad de incrementar la información a codificar sin aumentar la superficie ocupada por el código, ha propiciado la aparición de los denominados *códigos bidimensionales*.

Esta nueva simbología difiere sustancialmente de la de los códigos de barras, permitiendo representar en el mismo espacio que el ocupado por un código de barras, mucha más información, como se verá más adelante (para más detalles sobre los códigos de barras, véase, por ejemplo, [10]). Además, la lectura de un código de barras puede hacerse mediante un escáner lineal, no así los códigos bidimensionales, que requieren de un escáner que sea capaz de leer toda la superficie que ocupa el código y no sólo una línea transversal al mismo. Por otra parte, la lectura de los códigos de barras proporciona una serie de letras y dígitos (los representados por dichas barras) que deben ser interpretados a través de su correspondiente entrada en una base de datos.

Debido a las grandes posibilidades de codificación, esta nueva simbología se emplea en numerosos sectores, como en el de ventas (tanto al por mayor como al por menor), en envíos postales y por mensajero; en sanidad (para almacenar la información de análisis clínicos o historiales médicos); en aplicaciones gubernamentales (documentación de vehículos, carnés de conducir, pasaportes), etc. En el caso particular de España, las aplicaciones más extendidas de estos códigos bidimensionales utilizan el código PDF417 (que se estudiará con detalle a lo largo del presente trabajo) y se puede ver en el etiquetado de artículos de unos grandes almacenes (El Corte Inglés), y en determinada documentación del Ministerio de Hacienda y del de Sanidad y Consumo.

Concretamente, desde el año 1997, el Ministerio de Hacienda utiliza el código PDF417 en la Declaración sobre la Renta de las Personas Físicas, cuando ésta se cumplimenta mediante el programa de ayuda que proporciona la Agencia Tributaria (el propio programa PADRE genera e imprime este código en la declaración). Por su parte, el Ministerio de Sanidad y Consumo utiliza el mismo código PDF417 en las consultas médicas, ya sea para la prescripción de medicamentos, petición de consultas al especialista, partes de enfermedad, etc. En la Figura 1 pueden verse dos ejemplos de este código, uno para la prescripción de medicamentos y el otro para una petición de consulta especializada. Como se puede observar, en la etiqueta aparece, además del PDF417, información del paciente y del médico. El objetivo de ambos Ministerios es el de agilizar y mejorar el control de sus documentos.

Todos los códigos que se presentan a lo largo de este trabajo son *no significativos*, es decir, no contienen más información que la que se ha codificado y que va incluida en el propio código, por tanto no tienen relación con protocolos criptográficos ni es su intención ocultar información.

Este artículo, en el que se presentan algunos sistemas para la codificación de

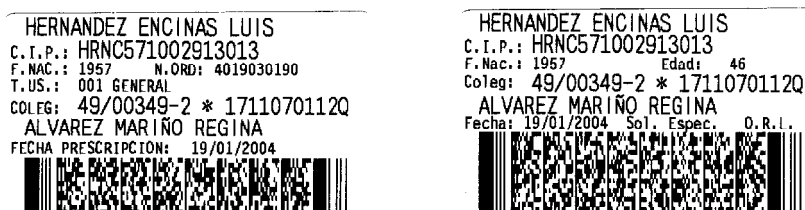


Figura 1: Ejemplos de Códigos del Ministerio de Sanidad y Consumo

información mediante códigos bidimensionales, está organizado de la siguiente manera: en la sección 2 se incluyen las principales definiciones matemáticas acerca de la teoría de códigos, que serán de utilidad a lo largo de este trabajo. En la sección 3 se comentan brevemente los códigos bidimensionales más empleados y se incluyen algunos ejemplos. Finalmente, el código PDF417, uno de los más extendidos, se analiza con detalle en la sección 4.

2 Códigos lineales y de Reed-Solomon

Se llama *alfabeto* a un conjunto finito de q símbolos, $\Sigma_q = \{s_1, s_2, \dots, s_q\}$, de modo que los elementos o vectores de Σ_q^n se conocen como *palabras de longitud n* . Se llama *código de longitud n* a todo subconjunto $\mathcal{C} \subset \Sigma_q^n$. Los elementos de \mathcal{C} se conocen como *palabras codificadas* o, siguiendo la terminología anglosajona, *codewords*. En el caso particular en que $q = 2$, los códigos se llaman *códigos binarios* (para un estudio más detallado véanse, por ejemplo, [3, 5, 9, 15, 22]).

Se define la *distancia de Hamming* entre dos vectores dados $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ de Σ_q^n , como el número de posiciones en las que difieren dichos vectores y puede expresarse como sigue:

$$d(x, y) = |\{i : 1 \leq i \leq n \text{ y } x_i \neq y_i\}|. \quad (1)$$

Dado un código \mathcal{C} , se define su *distancia mínima*, d , mediante la expresión

$$d = \min\{d(x, y) : x, y \in \mathcal{C} \text{ y } x \neq y\}.$$

Si \mathcal{C} tiene k elementos, es decir, $|\mathcal{C}| = k$, el código \mathcal{C} suele llamarse un (n, k, d) -código. Los códigos con distancia mínima $d \geq 2$ pueden ser empleados para detectar $d - 1$ errores (se conocen como *códigos detectores de errores*). Sea $d = 2e + 1 \geq 3$, entonces \mathcal{C} puede utilizarse para detectar errores (llamado *código corrector de errores*), de modo que puede corregir $e = (d - 1)/2$ errores. Si $d = 2e \geq 4$, el código puede corregir $e - 1 = (d - 2)/2$ errores. La capacidad de detectar y de corregir errores depende del valor de d y del algoritmo que se utilice. Así, por ejemplo, un código con distancia $d = 5$ puede utilizarse para detectar 4 errores, para corregir 2 errores, o para detectar 3 y corregir 1, etc.

La *tasa de información* de un código \mathcal{C} se define como la razón entre la cantidad de información significativa de cada palabra y la longitud de cada

palabra, es decir, es una medida del coste que supone transmitir un exceso de información:

$$R = \log_q |\mathcal{C}|/n. \quad (2)$$

Se llama *peso de Hamming* de un vector $x \in \Sigma_q^n$, representado por $w(x)$, al número de elementos no nulos que contiene, es decir, es la distancia del vector x al vector 0. Dado que la distancia de Hamming entre dos vectores —véase la expresión (1)—, x, y , es el número de posiciones en las que ambos difieren, se verifica que $d(x, y) = w(x - y)$. El *peso mínimo* de un código es el mínimo de los pesos de Hamming de las palabras no nulas de dicho código.

Los ejemplos de alfabetos más empleados son $\Sigma_q = \mathbb{Z}_q$ y, en el caso que q sea la potencia de un número primo, el cuerpo $\Sigma_q = \mathbb{F}_q$ (recuérdese que \mathbb{Z}_q es un cuerpo si y sólo si q es primo). Un código \mathcal{C} se dice un *código lineal* sobre \mathbb{F}_q si es un subespacio vectorial de \mathbb{F}_q^n , y se llama un $[n, k]$ -*código* si dicho código tiene dimensión k . En el caso en que, además, su mínima distancia sea d , se llama un $[n, k, d]$ -*código*. Es inmediato que la tasa de información de un $[n, k, d]$ -código es $R = k/n$. Por otra parte, como los códigos lineales son subespacios vectoriales, resulta que este peso mínimo coincide con la distancia mínima del código.

Toda matriz G cuyas filas sean vectores de una base del código lineal \mathcal{C} se dice que es una *matriz generadora* de \mathcal{C} . Dos códigos se dicen *equivalentes* si uno se obtiene a partir del otro cambiando, mediante la misma permutación para todas las palabras, el orden en el que aparecen las letras de cada palabra. De modo que, sin pérdida de generalidad, se puede suponer que la matriz generadora G de \mathcal{C} tiene la *forma estándar*: $G = (I_k P)$, siendo P una matriz $k \times (n - k)$, sin más que elegir en lugar de \mathcal{C} un código equivalente a él.

Si \mathcal{C} es un $[n, k]$ -código sobre \mathbb{F}_q , su *código dual* se define por:

$$\mathcal{C}^\perp = \{b \in \mathbb{F}_q^n : \forall a \in \mathcal{C}, \langle a, b \rangle = 0\},$$

siendo $\langle a, b \rangle$ el producto interior de a por b . Es claro que \mathcal{C}^\perp es un $[n, n - k]$ -código. Además, una matriz generadora, H , para \mathcal{C}^\perp , se llama *matriz de control* para \mathcal{C} . En este caso, el código \mathcal{C} se puede representar como sigue:

$$\mathcal{C} = \{a \in \mathbb{F}_q^n : Ha^T = 0^T\}.$$

Además, si la matriz generadora de \mathcal{C} es $G = (I_k P)$, entonces $H = (-P^T I_{n-k})$.

Si \mathcal{C} es un código lineal con longitud n y distancia d , eliminando las $d - 1$ últimas coordenadas de cada codeword, se obtiene un código de longitud $n - d + 1$, en el que todas las palabras son distintas, por lo que se tiene que $|\mathcal{C}| \leq q^{n-d+1}$; como consecuencia, para los $[n, k, d]$ -códigos se verifica que (*cota de Singleton*)

$$d \leq n - k + 1. \quad (3)$$

Los códigos que verifican la igualdad de la condición (3) se llaman *códigos de máxima distancia de separación* o *códigos MDS*. Si G es la matriz generadora de un código MDS, entonces cualesquiera k columnas de G son linealmente independientes. Como además, G es la matriz de control del código dual, éste tiene distancia mínima mayor que k , y dado que tiene dimensión $n - k$, de la

expresión dada en (3) se sigue que su distancia no puede exceder de $k + 1$. Por tanto, el código dual de un código MDS es también un código MDS.

Un código lineal se dice *cíclico* si para todo $(a_0, \dots, a_{n-1}) \in \mathcal{C}$ se tiene que $(a_{n-1}, a_0, \dots, a_{n-2}) \in \mathcal{C}$. A partir de ahora supondremos que $\text{mcd}(n, q) = 1$ e identificaremos las palabras $(a_0, a_1, \dots, a_{n-1}) \in \mathbb{F}_q^n$ con los polinomios de grado menor que n con coeficientes en \mathbb{F}_q , $a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in \mathbb{F}_q[x]/(x^n - 1)$. Dado que la multiplicación por x en este anillo se identifica con un cambio cíclico de las palabras, todo código cíclico se corresponde con un ideal de $\mathbb{F}_q[x]/(x^n - 1)$. Además, como este anillo lo es de ideales principales, un código cíclico está completamente descrito por un generador del ideal correspondiente. Si la descomposición de $(x^n - 1)$ en factores irreducibles sobre \mathbb{F}_q es

$$x^n - 1 = f_1(x) \cdots f_r(x),$$

entonces hay 2^r códigos cíclicos (no necesariamente no equivalentes) de longitud n , dado que $g(x)$ debe ser un divisor de $(x^n - 1)$. Si $x^n - 1 = g(x)h(x)$ con $g(x) = g_0 + g_1x + \dots + g_{n-k}x^{n-k}$ y $h(x) = h_0 + h_1x + \dots + h_kx^k$, entonces existe una matriz generadora, G , del código cíclico \mathcal{C} , con generador $g(x)$, que tiene k cambios cíclicos de $(g_0, \dots, g_{n-k}, 0, \dots, 0)$ como sus filas y una matriz de control, H , que tiene $n - k$ cambios cíclicos de $(0, \dots, 0, h_k, \dots, h_0)$ como filas. Se sigue que el código cíclico de generador $h(x)$ es equivalente a \mathcal{C}^\perp .

Sea \mathcal{C} un código cíclico de generador $g(x) = f_1(x) \cdots f_t(x)$. El cuerpo de descomposición de $g(x)$, es decir, la menor extensión de \mathbb{F}_q donde $g(x)$ descompone en factores lineales, es un cuerpo finito, \mathbb{F}_{q^m} , llamado *cuerpo localizador de errores*, por razones que se verán posteriormente. Sea β_i , $1 \leq i \leq t$, un cero de $f_i(x)$ en \mathbb{F}_{q^m} , entonces \mathcal{C} queda completamente determinado por este conjunto de ceros. Cualquiera de estos conjuntos, Z , que determinan el código se llama *conjunto de ceros* de \mathcal{C} , es decir,

$$(a_0, \dots, a_{n-1}) \in \mathcal{C} \Leftrightarrow \forall \beta \in Z, a_0 + a_1\beta + \dots + a_{n-1}\beta^{n-1} = 0. \quad (4)$$

Dado que los elementos de \mathbb{F}_{q^m} se pueden representar como vectores columna en \mathbb{F}_q^m , la matriz H de tamaño $t \times n$ y con filas $(1, \beta_i, \dots, \beta_i^{n-1})$, se puede interpretar como una matriz $tm \times n$ sobre \mathbb{F}_q ; y de la condición dada en (4), se sigue que a es un codeword si y sólo si $Ha^T = 0^T$. Si las filas de H no son linealmente independientes, la matriz de control para el código se puede obtener de H sin más que eliminar algunas de sus filas.

Sea β una raíz primitiva n -ésima de la unidad en una extensión de \mathbb{F}_q . El código definido por el conjunto de ceros dado por $Z = \{\beta^m, \beta^{m+1}, \dots, \beta^{m+d-2}\}$ se llama un *código BCH* (código de Bose-Chaudhuri-Hocquenghem) con distancia asignada d . La distancia mínima de un código BCH con distancia asignada d , es, al menos, d .

Un *código de Reed-Solomon* (*RS* para abreviar) es un código BCH de longitud $n = q - 1$ sobre \mathbb{F}_q . Como las raíces de $x^{q-1} - 1$ son precisamente los elementos no nulos de \mathbb{F}_q , resulta que todos los polinomios minimales son lineales y, para estos códigos, el cuerpo localizar de errores coincide con el alfabeto \mathbb{F}_q .

El generador de un código de RS corrector de e errores es un polinomio de grado $2e$ y tiene la siguiente forma:

$$g(x) = \prod_{i=1}^{2e} (x - \alpha^i), \quad (5)$$

siendo α un elemento primitivo de \mathbb{F}_q , es decir, un generador del grupo multiplicativo \mathbb{F}_q^* , y $d = 2e + 1$ la distancia asignada. Dado que la dimensión de un código de RS con distancia asignada d es $k = n - d + 1$, los códigos de RS son $[n, n - d + 1]$ -códigos y su distancia mínima es, por (3), exactamente d . Así pues, los códigos de RS son códigos MDS. Estas características son las principales razones para utilizar los códigos de RS, dado que para n y k fijos, no existen códigos que tengan una mínima distancia mayor que la de un código de RS. De ahí que los códigos de RS se usen para la corrección de errores en múltiples aplicaciones, como en la lectura de discos compactos (CD de audio y de datos, DVD, etc.), discos duros, imágenes enviadas por sondas espaciales (Voyager 1 y 2), etc. (ver [17, 22]).

De forma muy resumida, el proceso de corrección de errores se basa en los siguientes hechos. Si se transmite el polinomio $c(x)$ asignado a un codeword y lo que se recibe es el polinomio $v(x) = c(x) + e(x) \in \mathbb{F}_q[x]$, al polinomio $e(x) = v(x) - c(x)$ se le llama *polinomio de error*. El polinomio $v(x)$ puede ser evaluado sobre los elementos de \mathbb{F}_q , en particular sobre los ceros del polinomio generador $g(x)$: $\gamma_1, \dots, \gamma_r$. Dado que $c(\gamma_i) = 0$, $1 \leq i \leq r$, se tiene el siguiente conjunto de ecuaciones:

$$v(\gamma_i) = e(\gamma_i) = \sum_{j=0}^{n-1} e_j \gamma_i^j, \quad 1 \leq i \leq r,$$

en el que sólo intervienen las componentes del error y no las de los codewords. (Las expresiones $S_i = v(\gamma_i)$ se conocen como *síndromes*.) Si este conjunto de ecuaciones se puede resolver para los coeficientes e_j , el error puede ser determinado y corregido. Ahora bien, considerando como ceros del polinomio generador $g(x)$ los valores $\alpha, \alpha^2, \dots, \alpha^{2e}$, siendo α un elemento primitivo de \mathbb{F}_q (caso de los códigos de RS), resulta que los e errores pueden ser determinados de modo efectivo a partir de los síndromes S_1, S_2, \dots, S_{2e} (ver [3, §7.2]).

La combinación de dos códigos con distancia 2 permite obtener códigos con distancia mayor, que pueden corregir errores. Por ejemplo, esto sucede con algunos códigos de barras, como el código Postnet del Servicio Postal americano o como el código de barras EAN13 (véase [10]). Esta capacidad para corregir errores se debe a que son códigos detectores de errores (gracias a que su distancia mínima es 2 o a que utilizan un dígito de control) y a que es posible conocer el lugar que ocupa el dígito erróneo debido a la distribución de las barras que representan a los dígitos. A partir de estos dos datos, pueden recuperarse determinados errores en la lectura del código. Sin embargo, esta capacidad no suele incluirse en las aplicaciones informáticas que manipulan tales códigos.

3 Códigos bidimensionales

En los últimos años se ha desarrollado una serie de códigos, llamados *códigos bidimensionales*, que permiten la codificación de información en menos espacio que el que requeriría un código de barras unidimensional, para codificar la misma cantidad de información. El hecho de que un código de barras bidimensional pueda contener gran cantidad de información ocupando una pequeña superficie, ha favorecido su rápida aceptación en muchos ámbitos empresariales. Estos códigos se llaman bidimensionales porque para ser leídos requieren de lectores ópticos que puedan leer toda la superficie delimitada por el código y no sólo una parte transversal del mismo, dado que cada sección transversal es diferente de las restantes. En general, estos nuevos códigos han sido desarrollados por empresas privadas, que necesitaban un sistema de codificación que fuera capaz de almacenar gran cantidad de información en una superficie limitada. A continuación se describen, de forma resumida, los códigos bidimensionales más utilizados en la actualidad (existen más de veinte, ver [2]) y que suelen ser clasificados en dos grupos, dependiendo del aspecto de su representación gráfica: apilados y matriciales. Los primeros se asemejan a códigos de barras lineales dispuestos unos encima de otros y están formados por barras y espacios; mientras que los segundos se parecen más a matrices de puntos. Además de esta diferencia en su aspecto, los códigos matriciales tienen mayor capacidad de codificación, no necesitan de una zona muda a su alrededor para ser leídos correctamente por un escáner y pueden ser escalados puesto que todos sus elementos tienen el mismo tamaño. En contraposición, los códigos apilados pueden ser leídos por un escáner bidimensional, mientras que los matriciales necesitan de una cámara CCD (Charge Coupled Device), mucho más costosa. De ahí que el código bidimensional más extendido sea el PDF417, que entra dentro del grupo de los apilados.

El *Código 49* ([1], propuesto por Intermec Corporation [11] en 1987) está formado por una serie de códigos de barras apilados (ver Figura 2(a)) y es capaz de codificar hasta 170 caracteres alfanuméricos en el área de una pulgada cuadrada ($6,45 \text{ cm}^2$). El *Código 16K* ([1], desarrollado en 1989 por Laserlight Systems, Inc.) está basado en el código de barras Code 128 (de hecho ambos fueron inventados por T. Williams y el nombre de 16K procede del hecho que $128^2 = 16 \cdot 1024 = 16Kb$). Cada código contiene entre 2 y 16 filas y puede codificar hasta 77 caracteres ASCII o 154 dígitos (ver Figura 2(b)).

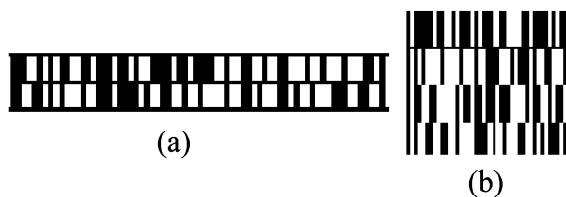


Figura 2: Ejemplos de Códigos 49 y 16K

En cuanto a los códigos matriciales, de entre los de dominio público cabe destacar al *Código Uno* (también inventado por T. Williams, en 1992), que contiene módulos de datos cuadrados (blancos y negros) en la zona superior e inferior, y varias líneas paralelas en la parte central. Puede tener diferentes formas y hasta 8 tamaños (ver Figura 3(a)). El más pequeño (1A) puede codificar hasta 13 caracteres alfanuméricos (incluyendo datos para la corrección de errores) o 22 dígitos; mientras que el mayor (1H) llega hasta 2218 caracteres o 3550 dígitos. El *MaxiCode* fue desarrollado por la empresa UPS (United Parcel Service [20]) en 1992. Este código se identifica fácilmente debido al “ojo” que aparece en el centro del código (ver Figura 3(b)), además de por los módulos blancos y negros hexagonales que contiene. MaxiCode puede codificar alrededor de 100 caracteres en una pulgada cuadrada y su capacidad para corregir errores permite que se recupere toda la información que contiene aunque se haya perdido el 25 % del código original.

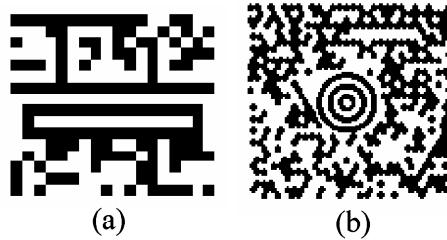


Figura 3: Ejemplos de Código Uno y Maxicode

La simbología *DataMatrix* es muy eficaz (fue desarrollada por CiMatrix [7]) y es una de las más utilizadas, sobre todo en los mercados de circuitos integrados e impresos, debido a que es capaz de codificar hasta 50 caracteres en $3mm^2$. *DataMatrix* contiene módulos de datos blancos y negros cuadrados y dos líneas blancas y otras dos negras en su perímetro (ver Figura 4(a)). Puede codificar hasta 3116 dígitos o 2335 caracteres alfanuméricos en un símbolo cuadrado de 144 módulos, empleando códigos de Reed-Solomon para corregir errores. Finalmente, el *Código Azteca* (desarrollado en 1995 por Welch Allyn, Inc. [21]) contiene, como el anterior, módulos cuadrados blancos y negros y varios cuadrados concéntricos en el centro (ver Figura 4(b)). Existen 32 tamaños diferentes, desde 15×15 módulos, que codifican 12 caracteres o 13 dígitos; hasta 151×151 módulos, que almacenan 3067 caracteres, 3832 dígitos o 1914 bytes. Además, el usuario puede seleccionar porcentualmente, entre el 5 % y el 95 %, el nivel de corrección de errores mediante códigos de Reed-Solomon.

4 El código bidimensional PDF417

Posiblemente el código bidimensional más utilizado sea el denominado *PDF417* (otras de sus aplicaciones, además de las ya comentadas de los Ministerios de

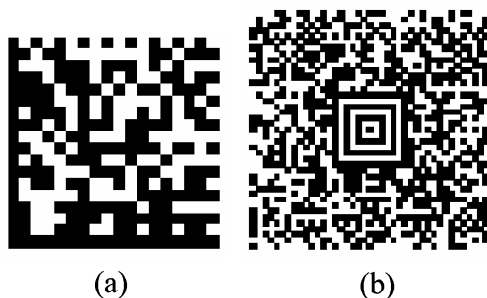


Figura 4: Ejemplos de DataMatrix y Código Azteca

Hacienda y Sanidad y Consumo españoles, pueden verse en [13, 14]). Este código fue desarrollado por Symbol Technologies ([18]) en 1990, y responde a las iniciales de Portable Data File. El significado de los números 417 se aclarará más adelante, cuando se vea cómo se genera esta simbología. Todo código PDF417 (ver Figura 5) está formado por diferentes partes: un *separador de inicio*, que indica dónde comienza el código, y un *separador de fin*, que señala dónde termina. Adyacentes a ambos separadores se sitúan el *indicador izquierdo* y el *indicador derecho*, que contienen información de carácter general sobre el contenido del código. Finalmente, entre los indicadores va la información codificada, que se distribuye en *filas* (entre 3 y 30) y en *columnas* (entre 1 y 30). Todas las filas de los separadores son idénticas (por tal motivo tienen el aspecto de barras), no así las del resto de la información del código.

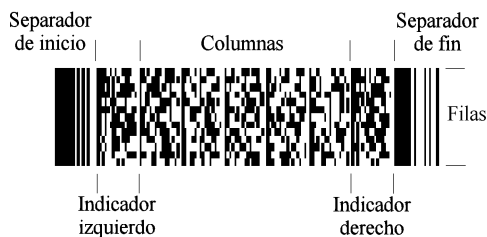


Figura 5: Descripción de la partes de un Código PDF417

Un código PDF417 puede contener hasta 1850 caracteres de texto, hasta 2710 dígitos o hasta 1108 bytes. Es posible codificar en un mismo PDF417 los tres tipos de datos. Por otra parte, es posible seleccionar entre 9 posibles niveles de corrección de errores mediante códigos de Reed-Solomon; que permiten recuperar la información perdida en un código por deterioro, desgaste e incluso rotura. Existen dos variantes de este código: el *Micro PDF417*, que es un PDF417 al que se le han suprimido el indicador derecho y el separador de fin, con el objetivo de ahorrar espacio, y el *Macro PDF417*, que está formado

por la sucesión de varios códigos PDF417 distintos. Este Macro PDF417 se emplea cuando la cantidad de información a codificar es elevada y no cabe en un único PDF417. En este caso, cada código incorpora información adicional, como el número de orden que hace en la serie de los códigos, etc., de modo que es posible leer los códigos en un orden diferente al natural.

4.1 Formato e indicadores

Los datos a incluir en un PDF417 (texto, numérico, byte o código corrector de errores) se transforman en la unidad básica de información del PDF417: el *codeword*. Existen 929 codewords diferentes y dado que 929 es un número primo, los codewords —nombrados desde el 0 al 928— se supondrán pertenecientes al cuerpo \mathbb{F}_{929} . Los primeros 900 codewords —del 0 al 899— se utilizan para la codificación de datos, mientras que los restantes 29 codewords —del 900 al 928— tienen asignadas funciones específicas dentro del código, como señalar los cambios entre los diferentes tipos de información codificada, el número de orden que corresponde a dicho símbolo si se trata de un Macro PDF417, etc.

Para elaborar el símbolo que representa de forma gráfica a un código PDF417, que corresponde a determinada información, cada codeword se representa gráficamente mediante barras y espacios verticales, alternados, de diferente anchura, de modo que si la anchura básica es el *módulo*, cada codeword consta de 17 módulos, divididos en 4 barras (*b*) y 4 espacios (*e*). El módulo toma el valor 0 (\square) si se refiere a un espacio y 1 si hace referencia a una barra (\blacksquare). El ancho del módulo se denomina *dimensión-x* y su altura *dimensión-y*. A la hora de elaborar el símbolo del código PDF417, se pueden elegir diferentes proporciones entre la *dimensión-x* y la *dimensión-y*, dependiendo del tamaño que se quiera dar al PDF417 y de la capacidad de lectura del lector óptico. Lo más habitual es utilizar una proporción de 3 a 1, es decir, $y = 3x$.

A modo de ejemplo, si el orden de barras y espacios es: *bebebebe*, entonces el código (31314122) está formado alternativamente por barras de 3, 3, 4 y 2 módulos; y por espacios de 1, 1, 1 y 2 módulos:

$$(31314122) \rightarrow \blacksquare\blacksquare\blacksquare\square\blacksquare\blacksquare\blacksquare\square\blacksquare\blacksquare\blacksquare\square\blacksquare\blacksquare\square\blacksquare\blacksquare\square$$

Ahora pueden entenderse los dígitos 4-17 del nombre PDF417: 4 barras y 4 espacios que suman 17 módulos. Dado que se trata de un código estándar, los separadores de inicio y de fin son siempre los mismos y corresponden a las siguientes secuencias gráficas:

$$\begin{aligned} \textit{inicio} : (81111113) &\rightarrow \blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\square\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\square\square\square \\ \textit{fin} : (711311121) &\rightarrow \blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\square\blacksquare\square\square\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare\blacksquare \end{aligned}$$

Esta representación obliga a que la de los codewords empiece siempre por una barra y termine por un espacio: *bebebebe*. Además, para evitar confusiones con los separadores, ninguna barra o espacio de un codeword puede tener más de 6 módulos. Nótese que el número de módulos del separador de fin es 18 y que contiene 5 barras y 4 espacios. La razón se debe a que el código debe acabar en una barra y no en un espacio para señalar el final del símbolo.

Cada uno de los 929 codewords puede ser representado gráficamente de tres maneras distintas, dependiendo de la fila en la que se encuentre, de modo que cada fila usa sólo uno de tres conjuntos de representaciones diferentes, los cuales se repiten secuencialmente cada tres filas. Dichos conjuntos se llaman *patrones*: \mathcal{P}_0 , \mathcal{P}_3 y \mathcal{P}_6 . El patrón para la fila 0 (primera fila) es \mathcal{P}_0 , para la 1 es \mathcal{P}_3 , para la 2 es \mathcal{P}_6 , para la fila 3 es de nuevo \mathcal{P}_0 , y así sucesivamente. Es decir, si \mathcal{P}_p denota el patrón, entonces se tiene

$$p = (\text{número de la fila} \bmod 3) \cdot 3. \quad (6)$$

Por otra parte, los indicadores izquierdo, L_i , y derecho, R_i , son diferentes para cada PDF417 dado que dependen de la fila en la que están situados, i , del número de filas totales, f , del número de columnas de datos, c , y del nivel de corrección de errores que se vaya a emplear, s . Teniendo en cuenta la expresión (6), que indica el patrón a utilizar, el cálculo de los codewords para los indicadores izquierdo y derecho se efectúa como sigue:

$$L_{i,p} = \begin{cases} 30 \lfloor \frac{i-1}{3} \rfloor + \lfloor \frac{f-1}{3} \rfloor, & \text{si } p = 0 \\ 30 \lfloor \frac{i-1}{3} \rfloor + 3s + ((f-1) \bmod 3), & \text{si } p = 3 \\ 30 \lfloor \frac{i-1}{3} \rfloor + c - 1, & \text{si } p = 6 \end{cases}$$

$$R_{i,p} = \begin{cases} 30 \lfloor \frac{i-1}{3} \rfloor + c - 1, & \text{si } p = 0 \\ 30 \lfloor \frac{i-1}{3} \rfloor + \lfloor \frac{f-1}{3} \rfloor, & \text{si } p = 3 \\ 30 \lfloor \frac{i-1}{3} \rfloor + 3s + ((f-1) \bmod 3), & \text{si } p = 6 \end{cases}$$

donde $\lfloor x \rfloor$ denota la parte entera de x .

Si se consideran $c = 3$, $f = 7$ y $s = 2$, los valores de los indicadores son

$$L = [2, 6, 2, 2, 36, 32, 32], \quad R = [2, 0, 6, 2, 30, 36, 32]. \quad (7)$$

Como ejemplo de distribución de los codewords para elaborar un PDF417, supóngase que la información codificada ha proporcionado 14 codewords de datos: c_1, \dots, c_{14} , y 4 codewords adicionales para corregir errores: e_1, \dots, e_4 . La estructura del código requiere añadir a estos 18 codewords uno más que indique el número total de codewords de datos: d_{15} . Los 19 codewords obtenidos podrían distribuirse en 10 filas y 2 columnas, pero como el número de celdas sería de 20, habría que añadir un dato de relleno: d_0 . Así pues, salvo los patrones de inicio y fin, este PDF417 tendría el siguiente formato:

L_1	d_{15}	c_{14}	R_1
L_2	c_{13}	c_{12}	R_2
L_3	c_{11}	c_{10}	R_3
\vdots	\vdots	\vdots	\vdots
L_7	c_3	c_2	R_7
L_8	c_1	d_0	R_8
L_9	e_4	e_3	R_9
L_{10}	e_2	e_1	R_{10}

Separador de inicio Separador de fin (8)

El formato presentado en (8) es el estándar que corresponde al *nivel alto de codificación*, es decir, a la conversión de los datos en codewords. Estos codewords se colocan de izquierda a derecha y de arriba a abajo, de modo que el primero de ellos es el descriptor de la longitud del símbolo. Este descriptor codifica el número total de codewords de datos del símbolo, es decir, incluye el propio descriptor, los codewords de datos y los codewords de relleno, y excluye el número de codewords correctores de errores. A continuación del descriptor, se distribuyen los codewords de datos, ordenados de modo que el codeword más significativo sea el adyacente al descriptor de la longitud del símbolo (de ahí que aparezcan escritos en el orden inverso al señalado por los subíndices); siguen los codewords de relleno y los codewords correctores de errores. El segundo paso, el *nivel bajo de codificación*, consiste en transformar dichos codewords a barras y espacios, dando lugar al símbolo correspondiente, para su impresión definitiva.

A continuación se verán los diferentes procedimientos que permiten codificar cada uno de los tipos de datos (texto, números y bytes) a codewords.

4.2 Codificaciones

El modo texto, que incluye los caracteres ASCII internacionales (9, 10, 13 y del 32 al 127), está dividido en 4 submodos, cada uno de los cuales contiene 30 codewords. Estos submodos se corresponden, básicamente, a las letras mayúsculas, las minúsculas, los 10 dígitos y símbolos aritméticos, y los símbolos de puntuación. La transformación de un carácter cualquiera se lleva a cabo mediante un cambio entre la base 30 (número de caracteres de cada grupo) y la base 900 (número de codewords de datos). Dado que $900 = 30^2$, la proporción de caracteres de texto por codewords es, aproximadamente, de 2 caracteres de texto por cada codeword. De hecho, el mayor valor que podrían alcanzar dos caracteres sería el de números de dos cifras en base 30: $29 \cdot 30 + 29 = 899$, que es menor que el número de posibles codewords. Por tanto, el proceso de codificación consiste en expresar en base 900 cada pareja de caracteres y tomar como codeword el valor resultante. Si $A \mapsto 0$, $B \mapsto 1$, etc., para el par “EJ” se tiene

$$EJ = 49_{(30)} = 4 \cdot 30 + 9 = 129.$$

Así, los codewords del texto “EJEMPLO DE DEMOSTRACION”¹ son:

$$[129, 132, 461, 446, 94, 783, 132, 438, 587, 2, 254, 419]. \quad (9)$$

Antes de llevar a cabo la codificación de dígitos se debe establecer la proporción más adecuada para transformar dígitos a codewords. En el caso del modo texto era sencillo porque 2 caracteres en base 30 están acotados por el máximo valor de un codeword: $30^2 = 900$. Sin embargo, en esta ocasión se deben buscar dos números enteros d y c de modo que d dígitos puedan codificarse con

¹No se ha utilizado la letra mayúscula acentuada “Ó” en la palabra “DEMOSTRACION” dado que tanto las vocales acentuadas como la letra “ñ” son propias del alfabeto español y no pertenecen a los caracteres ASCII internacionales, por lo que no se pueden codificar en el modo texto.

c codewords, teniendo en cuenta que hay 10 dígitos y 900 codewords. Así pues, inicialmente se trataría de encontrar enteros d y c , no demasiado grandes —en cualquier caso menores que $q - 1 = 900$, que es la longitud del código de Reed-Solomon que se empleará para corregir errores— de modo que $10^d \leq 900^c$, es decir, verificando

$$\frac{d}{c} \leq \log_{10} 900 = 2,9542. \quad (10)$$

Se podrían considerar diferentes parejas de enteros (d, c) que cumplieran la condición dada en (10), por ejemplo,

$$(44, 15), (59, 20), (369, 125), \dots \quad (11)$$

De este modo las proporciones de codificación serían de 44 dígitos por cada 15 codewords; 59 dígitos por cada 20 codeword, etc. Ahora bien, como se verá en el algoritmo que permite codificar dígitos como codewords, los valores que se deben tomar son $d = 44$ y $c = 15$. Este algoritmo es el siguiente:

- (i) Se hacen grupos de 44 dígitos (el último grupo podría contener menos dígitos y se codificaría con menos de 15 codewords).
- (ii) Se añade a cada grupo determinado en (i) el 1 como primer dígito (en la decodificación este 1 es eliminado) para evitar comenzar por ceros y considerar cada grupo como un entero g_i ; de tal manera que, por ejemplo, los grupos 0, 00 y 000 puedan considerarse diferentes, al igual que los grupos 12, 012 y 0012, etc.
- (iii) Para cada uno de los enteros $g = g_i \neq 0$ se determina su expresión en base 900 y se utilizan los dígitos obtenidos como codewords, es decir,

1. $c_j \leftarrow g \pmod{900}$
2. $g \leftarrow \lfloor \frac{g}{900} \rfloor$.
3. $j \leftarrow j + 1$.

El algoritmo termina cuando $g = 0$, con lo que el número de codewords para un g dado puede ser menor que 15, como se verá en el ejemplo que se presenta más adelante.

Queda por justificar la elección de los valores $d = 44$ y $c = 15$, que permiten la codificación de 44 dígitos mediante 15 codewords. Esta justificación es inmediata teniendo en cuenta que en el paso (ii) del algoritmo anterior se consideran, después de añadir un 1 como primer dígito, grupos de 45 dígitos, con lo que el mayor valor entero que puede tomar uno de estos grupos es un uno seguido de $d = 44$ nueves, valor que es menor que el número de codewords elevado a $c = 15$, es decir, se verifica que

$$19 \cdot \overset{44}{\dots} 9 = 2 \cdot 10^{44} - 1 < 900^{15},$$

condición que no cumplen las restantes parejas dadas en (11). Así pues, la desigualdad dada en (10) debe modificarse por esta otra:

$$\frac{d + \log_{10} 2}{c} \leq \log_{10} 900 = 2,9542.$$

(Nótese que $(59 + 0,3010)/20 = 2,9650 > 2,9542$ y que $(369 + 0,3010)/125 = 2,9544 > 2,9542$).

Se consideran, a modo de ejemplo, los 100 dígitos siguientes

$$[1, 2, 3, 4, 5, 6, 7, 8, 9, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, \dots, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0].$$

Sus correspondientes codewords, según el algoritmo anterior, son:

$$[491, 81, 137, 450, 302, 67, 15, 174, 492, 862, 667, 475, 869, 12, 434, 685, 326, 422, 57, 117, 339, 377, 238, 839, 698, 145, 870, 348, 517, 378, 2, 808, 3, 153, 190].$$

Para la codificación de bytes se procede de forma análoga a como se hizo para números. Como hay 256 bytes (código ASCII extendido), deben elegirse enteros b y c de modo que b bytes puedan codificarse con c codewords, es decir, deben determinarse dos enteros de modo que $256^b \leq 900^c$. Se tiene entonces, que

$$\frac{b}{c} \leq \frac{\log_{10} 900}{\log_{10} 256} = 1,227.$$

Tomando como valores $b = 6$ y $c = 5$, resulta que $b/c = 6/5 = 1,2$, y la codificación resultante es de 5 codewords por cada 6 bytes.

La transformación de una cadena de bytes, que podría ser un texto con caracteres del código ASCII extendido, como el siguiente “Demostración del modo byte para PDF417”, proporcionaría los siguientes codewords:

$$[114, 558, 291, 773, 560, 191, 613, 775, 591, 258, 54, 255, 340, 3, 737, 186, 607, 60, 228, 873, 195, 52, 341, 605, 302, 162, 690, 251, 210, 444, 49, 55].$$

4.3 PDF417, códigos binarios y particiones de 17

Se pueden justificar de forma matemática algunas de las decisiones tomadas por los inventores del PDF417 a la hora de elaborar el código tal y como se ha señalado en §4.1. Así, se justificará el uso de tres patrones para representar cada codeword según la fila a la que pertenezca y se razonará la necesidad de utilizar 17 módulos (suma de barras y espacios) para representar gráficamente cada codeword.

El hecho de utilizar tres patrones está íntimamente relacionado con la forma en que los escáneres leen las barras de un código. Supongamos que las barras y espacios de la representación gráfica de un codeword es la siguiente, donde los \bar{x}_i son el número de módulos de cada barra o espacio:



Cuando un escáner lee una serie de barras y espacios, en lugar de determinar la secuencia $(\bar{x}_1, \dots, \bar{x}_8)$, lo que en realidad calcula son las distancias entre los bordes similares, lo que permite compensar las posibles deficiencias de la impresión, es decir, calcula los valores de

$$\bar{t}_i = \bar{x}_i + \bar{x}_{i+1}, \quad 1 \leq i \leq 7.$$

Por otra parte, el escáner también puede proporcionar la longitud, l , del codeword sin más que calcular la distancia que separa la primera de las barras de la quinta —que es la primera del codeword adyacente—. De este modo se tiene $l = \sum_{i=1}^8 \bar{x}_i$. Como se sabe que en un PDF417 el número de módulos es 17, resulta que la dimensión- x del símbolo que se está leyendo es $x = l/17$. A continuación, cada valor \bar{t}_i se divide entre x y se redondea el cociente para obtener una secuencia normalizada, es decir, se calcula $t_i = \lceil \bar{t}_i/x \rceil$, $1 \leq i \leq 7$, donde los t_i son una secuencia de valores que debería corresponder a las sumas de cada dos módulos consecutivos de un codeword. El valor

$$t = t_1 + t_3 + t_5 + t_7 = \sum_{i=1}^8 x_i \approx 17,$$

obtenido a partir de esta secuencia normalizada se compara con el correspondiente valor obtenido de la lectura de los separadores de inicio y fin. Si coinciden, el proceso de lectura continúa; en caso contrario, el escáner procede a leer de nuevo el codeword.

Por ejemplo, si los valores de las \bar{x}_i fueran 43, 19, 31, 25, 55, 40, 80, y 47; los valores proporcionados por el escáner serían $\bar{t}_1 = 62$, $\bar{t}_2 = 50$, $\bar{t}_3 = 56$, $\bar{t}_4 = 80$, $\bar{t}_5 = 95$, $\bar{t}_6 = 120$, y $\bar{t}_7 = 127$; $l = 340$ y $x = 340/17 = 20$. A partir de estos valores se obtienen: $t_1 = t_2 = t_3 = 3$, $t_4 = 4$, $t_5 = 5$, $t_6 = t_7 = 6$; de donde resulta que $t = 17$.

Por otra parte, como todo codeword se representa gráficamente mediante 4 barras y 4 espacios alternados (*bebebebe*) de diferente longitud, resulta que cada codeword del PDF417 se puede expresar mediante unos (barras) y ceros (espacios). Así pues, cada codeword no es sino un vector binario de longitud n . Además, todo vector binario de longitud n que contenga de forma alternada ocho grupos de, a lo más, 6 unos y 6 ceros consecutivos puede representarse de la siguiente forma:

$$a(n) = \left(1, \overset{x_1}{\dots}, 1, 0, \overset{x_2}{\dots}, 0, \dots, 1, \overset{x_7}{\dots}, 1, 0, \overset{x_8}{\dots}, 0\right), \quad (12)$$

siendo $n = \sum_{i=1}^8 x_i$, y $1 \leq x_i \leq 6$.

Con el fin de dar una interpretación del vector $a(n)$ en términos matemáticos conocidos, recordamos que, dado un número entero n , se definen las *particiones de n como suma de m partes* a las diferentes formas de escribir n como suma no ordenada de m enteros positivos, es decir, $(\alpha_1, \dots, \alpha_m)$ es una partición de n si $\sum_{i=1}^m \alpha_i = n$. Se denotará por $P(n, m)$ a tal conjunto. Dado que el orden de los α_i no importa, éstos se suelen considerar bien en orden creciente o decreciente.

Aquí se supondrá que los α_i se ordenan de forma creciente. Por otra parte, se llama *composición de un entero n en m partes* a cada una de las formas de escribir el entero n como suma de exactamente m enteros positivos, de modo que dos composiciones son diferentes si cambia el orden de los enteros que la forman. El conjunto de todas las composiciones de n en m partes se denota por $\mathcal{C}(n, m)$ ([4, Capítulo 5.3], [6, Capítulo 2.7]).

Por tanto, dar el vector binario $a(n)$, definido en (12), es equivalente a dar una composición del entero n como suma de 8 números positivos menores o iguales que 6, es decir,

$$a(n) \Leftrightarrow (x_1, \dots, x_8) : n = \sum_{i=1}^8 x_i, \quad 1 \leq x_i \leq 6. \quad (13)$$

Veamos ahora que para el caso del PDF417 no puede tomarse $n \leq 16$.

Si $n = 16$, se tiene que $|\mathcal{C}(16, 8)| = 6435$. El subconjunto $\mathcal{C}_{\leq 6}(16, 8) \subset \mathcal{C}(16, 8)$ formado por las composiciones de enteros menores o iguales que 6 tiene cardinal 6147. Los posibles patrones, \mathcal{P}_i , a utilizar en este caso (mencionados en §4.1) se pueden determinar, a partir de los datos t_i normalizados o de los valores de x_i dados en (12), sin más que calcular

$$p = (t_1 - t_2 + t_5 - t_6) \pmod 9 = (x_1 - x_3 + x_5 - x_7) \pmod 9. \quad (14)$$

Los resultados obtenidos en este caso son

$p = 0$	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 8$
945	867	710	554	470	470	554	710	867

No obstante, ninguno de los patrones anteriores permite codificar los 929 codewords, dado que sólo uno de ellos (el correspondiente a $p = 0$) contiene más de tal número. Así pues, no pueden considerarse vectores de longitud 16 o menor, es decir, cada codeword tendrá, al menos, $n = 17$ módulos.

Sea ahora $n = 17$. Se verifica entonces que $|\mathcal{C}(17, 8)| = 11440$, y $|\mathcal{C}_{\leq 6}(17, 8)| = 10480$. Para las particiones correspondientes se tiene que $|\mathcal{P}(17, 8)| = 29$ y $|\mathcal{P}_{\leq 6}(17, 8)| = 22$. Estas últimas particiones son las siguientes:

[1, 1, 1, 1, 1, 1, 5, 6], [1, 1, 1, 1, 1, 2, 4, 6], [1, 1, 1, 1, 1, 3, 3, 6], [1, 1, 1, 1, 2, 2, 3, 6],
 [1, 1, 1, 2, 2, 2, 2, 6], [1, 1, 1, 1, 1, 2, 5, 5], [1, 1, 1, 1, 1, 3, 4, 5], [1, 1, 1, 1, 2, 2, 4, 5],
 [1, 1, 1, 1, 2, 3, 3, 5], [1, 1, 1, 2, 2, 2, 3, 5], [1, 1, 2, 2, 2, 2, 2, 5], [1, 1, 1, 1, 1, 4, 4, 4],
 [1, 1, 1, 1, 2, 3, 4, 4], [1, 1, 1, 2, 2, 2, 4, 4], [1, 1, 1, 1, 3, 3, 3, 4], [1, 1, 1, 2, 2, 3, 3, 4],
 [1, 1, 2, 2, 2, 2, 3, 4], [1, 2, 2, 2, 2, 2, 2, 4], [1, 1, 1, 2, 3, 3, 3, 3], [1, 1, 2, 2, 2, 3, 3, 3],
 [1, 2, 2, 2, 2, 2, 3, 3], [2, 2, 2, 2, 2, 2, 2, 3].

Para determinar los codewords que pueden codificarse con los posibles patrones, se determinan, como en (14), los siguientes valores de p :

$p = 0$	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 8$
1484	1392	1202	1002	902	902	1002	1202	1392

Ahora bien, con el fin de minimizar la probabilidad de error, sólo se tienen en cuenta los siguientes tres casos: $p = 0, 3, 6$, que son los que dan lugar a los tres patrones ya mencionados en la expresión (6): $\mathcal{P}_0, \mathcal{P}_3$ y \mathcal{P}_6 .

Dado que hay 929 codewords, se eligen y ordenan de los conjuntos correspondientes a $p = 0, 3, 6$, las 929 composiciones necesarias, dando lugar a los tres patrones: $\mathcal{P}_0, \mathcal{P}_3$ y \mathcal{P}_6 . De esta forma, es posible asignar a cada codeword una partición dentro de cada patrón, según la fila a la que pertenezca. Ahora bien, por la equivalencia señalada en (13) entre particiones y vectores binarios, ahora sí de longitud 17, estos patrones pueden identificarse como códigos binarios, $\mathcal{C}_0, \mathcal{C}_3, \mathcal{C}_6 \in \mathbb{F}_2^{17}$, cada uno con 929 codewords y de longitud 17. La tasa de información de estos (17, 929)-códigos, no lineales, según la expresión (2), es

$$R = \frac{\log_2 929}{17} = 0,5799.$$

Nótese que diferentes elecciones y ordenaciones de los elementos de cada uno de los patrones o códigos binarios supone elaborar códigos diferentes al estándar PDF417.

4.4 Códigos correctores de errores

De la misma manera que otros códigos bidimensionales utilizan códigos correctores de errores para recuperar información codificada, aunque el código esté deteriorado por manchas o roturas, el código PDF417 utiliza códigos de Reed-Solomon sobre \mathbb{F}_{929} . En este caso, cada palabra del código de RS corresponde a un codeword del código PDF417. Existen 9 niveles de corrección de errores —del 0 al 8— de modo que todos los niveles, a excepción del 0, no sólo detectan errores sino que son capaces de recuperar codewords erróneos o perdidos. El nivel de corrección de errores a utilizar depende de la cantidad de codewords de datos que se prevea puedan perderse, aunque las recomendaciones al uso siguen el Cuadro 1.

Número de codewords	1 a 40	41 a 160	161 a 320	321 a 863
Mínimo nivel de corrección	2	3	4	5

Cuadro 1: Niveles de corrección de errores recomendados

Por otra parte, debe tenerse en cuenta que cuanto mayor sea el nivel de seguridad, más codewords para la corrección de errores deben utilizarse y, por tanto, menor cantidad de información se puede codificar. El número de codewords necesarios para la corrección de errores para un nivel de seguridad s es $c = 2^{s+1} = d - 1$, siendo d la distancia del código de RS. El valor de c no depende del número de letras de información y se pueden corregir así hasta $(d-1)/2 = 2^s$ errores, es decir, para un nivel 0 se utilizan 2 codewords, mientras que para un nivel 8 se requieren 512 codewords.

Por otra parte, el cálculo de estos codewords necesita, como ya se indicó en §2, del cómputo previo de los coeficientes del polinomio generador de los códigos

de Reed-Solomon (ver [3, 5, 9, 15, 22]). Dado que estos coeficientes son siempre los mismos, se pueden precomputar y hacer uso de ellos cada vez que sean necesarios ([16]). A partir de estos coeficientes y de los codewords de datos a codificar se determinan los codewords que se incluirán en el código PDF417 como códigos correctores de errores.

El polinomio generador de un código de RS de distancia asignada $d = c + 1$ es, según (5),

$$g(x) = \prod_{i=1}^c (x - \alpha^i) = x^c + a_{c-1}x^{c-1} + \dots + a_1x + a_0,$$

siendo α un elemento primitivo del cuerpo \mathbb{F}_q . En este caso, como el cuerpo es \mathbb{F}_{929} , se toma $\alpha = 3$. Los coeficientes, a_i , de $g(x)$, $0 \leq i \leq c - 1$, para $s = 2$ son:

$$237, 308, 436, 284, 646, 653, 428, 379.$$

Una vez determinados los coeficientes del polinomio generador del código de RS, se puede proceder al cálculo de los códigos correctores de errores de los codewords de datos. Para ello se incluye como primer dato el descriptor de la longitud del símbolo, es decir, el número que indica el número total de codewords y que incluye al propio descriptor, los codewords de datos y los datos de relleno. De este modo, cuando un escáner lee el símbolo del PDF417, conoce cuántos son los codewords de datos y cuántos los codewords correctores de errores (y a la vez el nivel de seguridad que tiene el código PDF417 que se está leyendo). Así pues, el número de datos a codificar aumenta en 1.

Los códigos correctores de errores de los 12 datos dados en la expresión (9), a los que se añade el número total de datos: 13, para $s = 2$ son:

$$[473, 783, 151, 898, 416, 869, 102, 473]. \quad (15)$$

4.5 Elaboración del PDF417

Para elaborar gráficamente el código PDF417 se utilizan los codewords correspondientes a la información a codificar y los relativos a los códigos correctores de errores. Luego se decide el aspecto que tendrá el código, es decir, el número de filas y de columnas (lo que puede obligar a considerar datos de relleno). Si se utiliza el ejemplo tratado en §4.2, a los 12 codewords de datos de (9) hay que añadir como primer dato (tal y como se hizo en §4.4 para calcular los códigos correctores de errores) el número de los mismos, que pasa a ser un nuevo dato, por lo que los codewords de datos son:

$$[13, 129, 132, 461, 446, 94, 783, 132, 438, 587, 2, 254, 419],$$

mientras que los codewords correctores son los de (15), por lo que el número de codewords totales es de 21, que se pueden distribuir en 7 filas y 3 columnas (en este ejemplo no se precisan datos de relleno). Utilizando los indicadores a izquierda y derecha determinados en (7), resulta que el código

PDF417, escrito en codewords, correspondiente a la información “EJEMPLO DE DEMOSTRACION”, con un nivel de seguridad 2, es el siguiente:

	002	013	129	132	002	
	006	461	446	094	000	
	002	783	132	438	006	
Separador de inicio	002	587	002	254	002	Separador de fin
	036	419	473	783	030	
	032	151	898	416	036	
	032	869	102	473	032	

Si estos codewords se escriben como particiones, esto es, como elementos de P_p , $p = 0, 3, 6$, según la fila a la que pertenezcan, se tiene el Cuadro 2. Finalmente, se transforma cada dígito de este Cuadro en una barra o un espacio cuya anchura es el valor del propio dígito, es decir, se considera el código de C_p que le corresponda.

81111113	51111152	31112144	31151231	11151413	51111152	711311121
81111113	51111323	43111412	11245211	31123223	41111216	711311121
81111113	11111246	13411214	12111245	12242141	21111353	711311121
81111113	21114251	13161113	51111152	32131331	21114251	711311121
81111113	21113414	12144113	33113123	32324111	41113232	711311121
81111113	51114122	12113162	23521112	12232142	41114411	711311121
81111113	31123151	14222411	41141141	21313223	31123151	711311121

Cuadro 2: Expresión de un PDF417 mediante particiones de 17

En general y dado que lo que se va a codificar suele proceder de un procesador de textos, se suelen utilizar dos métodos para generar el código e imprimirlo.

El primero de ellos consiste en generar el dibujo (mediante programación y en formato BMP, TIF, etc.) a partir del Cuadro 2 y luego insertarlo como tal gráfico en el procesador de textos para su posterior impresión.

El segundo método consiste en definir una fuente de letra especial sin recurrir a procedimientos gráficos. Para ello se procede de la siguiente manera. Una vez sustituidos los codewords por sus correspondientes particiones y éstas por sus vectores binarios, el código PDF417 puede verse como una tabla de f filas y $c + 4$ columnas, siendo f el número de filas y c el número de columnas de datos. Es decir, la tabla contiene en cada fila $c' = 17 \cdot (c + 4) + 1$ dígitos binarios. A continuación las filas se leen de tres en tres (caso de que éstas no fueran múltiplo de tres, se añadirían las filas que fuera necesario, formadas por vectores nulos), de modo que cada una de las columnas de las tres filas se interpreta como la expresión en base 2 de un dígito entre 0 y 7, siendo el bit menos significativo el de la fila inferior. De este modo, la tabla de códigos del PDF417 se convierte en una tabla que tiene $\lceil f/3 \rceil$ filas (donde $\lceil x \rceil$ denota la parte entera por exceso de x) y c' columnas formadas por los dígitos $0, \dots, 7$. Si ahora se define un tipo

de letra de modo que asigne a cada dígito $0, \dots, 7$, un símbolo que tenga zonas negras en los lugares correspondientes al 1 y blancas en los lugares donde aparece el 0, se podrá transformar la tabla de códigos PDF417 en un código de barras bidimensional sin más que cambiar de tipo de letra. En la Figura 6 se muestran los símbolos que corresponderían a esta fuente de letra.



Figura 6: Fuente de letra para PDF417

A modo de ejemplo, si las particiones de las posiciones $(1, 3)$, $(2, 3)$ y $(3, 3)$ del Cuadro 2, se escriben como vectores binarios y se determina el dígito cuya expresión en base 2 es la que se lee en cada columna, se tienen los 17 dígitos de la última fila del Cuadro 3.

31112144	→	11101011011110000
43111412	→	11110001010000100
13411214	→	10001111010010000
		76625157074450200

Cuadro 3: Los 17 dígitos de tres filas y una columna

Finalmente, repitiendo este proceso para todos los códigos y posiciones mostrados en el Cuadro 3, y utilizando el tipo de letra de la Figura 6, se obtiene la representación gráfica del código PDF417 que se muestra en el Figura 7.



Figura 7: Ejemplo de PDF417

Agradecimientos. Los autores agradecen a la Dra. R. Álvarez Mariño las facilidades prestadas para obtención de la Figura 1 y al recensor anónimo sus sugerencias para la mejora en la versión final de este artículo.

Referencias

- [1] American National Standard Intitute (ANSI), <http://www.ansi.org/>
- [2] D. Barnes, J. Bradshaw, L. Day, T. Schott, and R. Wilson, Two dimensional bar coding, Tech 621, Purdue University, Spring 1999.

- [3] R.E. Blahut, *Theory and practice of error control codes*, Addison-Wesley Publishing Company, Reading, MA, 1983.
- [4] C. Chuan Chong and K. Khee Meng, *Principles and techniques in combinatorics*, World Scientific, Singapore, 1992.
- [5] G.C. Clark, Jr. and J.B. Cain, *Error-correction coding for digital communications*, Plenum Press, New York, 1981.
- [6] G.M. Constantine, *Combinatorial theory and statistical design*, John Wiley & Sons, New York, 1987.
- [7] CiMatrix, <http://www.cisworldservices.org/cimatrix.html>
- [8] European Article Numbering, <http://www.ean-int.org/index800.html>
- [9] D. Hankerson, D.R. Hoffman, D.A. Leonard, C.C. Lindner, K.T. Phelps, C.A. Rodger, and J.R. Wall, *Coding theory and cryptography, The essentials*, 2nd edition, Marcel Dekker, Pure and Applied Mathematics 234, 2000.
- [10] L. Hernández Encinas y A. Martín del Rey, Codificación de información mediante códigos de barras, aceptado para su publicación en Bol. Soc. Esp. Mat. Apl.
- [11] Intermec Corporation, http://www.intermec.net/intermec_corporation.htm
- [12] The International ISBN Agency, <http://www.isbn-international.org/>
- [13] PDF417, Aplicaciones, <http://www.pdf417.com/applications.htm>
- [14] PDF417, Soluciones, <http://www.symbol.com/solutions/>
- [15] W.W. Peterson and E.J. Weldon, Jr., *Error-correcting codes*, The MIT Press, 2nd edition, Cambridge, 1972.
- [16] www.pdf417.com/PDF417_tutorial/coefficient_tables_for_pdf417_ecl.htm
- [17] L. Steen (Ed.), *Las Matemáticas en la vida cotidiana*, Addison Wesley, 3ª edición, 1999.
- [18] Symbol Technologies, <http://www.symbol.com/>
- [19] Uniform Code Council (UCC), <http://www.uc-council.org/>
- [20] United Parcel Service (UPS), <http://www.ups.com/>
- [21] Welch Allyn, Inc., <http://www.welchallyn.com/>
- [22] S.B. Wicker and V.K. Bhargava, *Reed-Solomon codes and their applications*, IEEE Press, Piscataway, NJ, 1994.

Modelos de computación celular con membranas

M.J. PÉREZ, A. ROMERO Y F. SANCHO¹

Dpto. Ciencias de la Computación e Inteligencia Artificial
E.T.S. Ingeniería Informática. Universidad de Sevilla

{marper,alvaro,fsancho}@us.es

Resumen

En este trabajo se presenta un nuevo modelo de computación no convencional (los sistemas celulares con membranas), inspirado en la estructura y funcionamiento de las células de los organismos vivos, y se desarrolla una teoría de la complejidad computacional en dicho modelo que permite dar una nueva caracterización de la relación $\mathbf{P} \neq \mathbf{NP}$.

Palabras clave: *Computación celular, Sistemas P, Sistemas celulares, Modelo no convencional, Complejidad computacional, Clases P y NP.*

Clasificación por materias AMS: *68Q05, 68Q10, 68Q15, 92B20*

1 Introducción

La búsqueda de procedimientos sistemáticos (*algoritmos*) que proporcionen soluciones de problemas, ha sido una constante del hombre a lo largo de la historia. G.W. Leibniz formuló la necesidad de disponer de un lenguaje universal (*lingua characteristic*) en el que poder expresar cualquier idea, así como la conveniencia de *mecanizar* cualquier tipo de razonamiento (*calculus ratiocinator*). En 1928, en el Congreso Internacional de Matemáticos, celebrado en Bolonia, D. Hilbert formuló tres cuestiones acerca de las Matemáticas que marcarían su devenir: *¿Son completas? ¿Son consistentes? ¿Son decidibles?*

En 1931, K. Gödel presentó sus teoremas de incompletitud, mostrando que era irrealizable el Programa de Hilbert (que trataba de justificar el uso de la metodología transfinita y, al mismo tiempo, proporcionar una formalización completa de las Matemáticas). K. Gödel consideró un sistema axiomático que contenía la Aritmética elemental y construyó una proposición verdadera que no se podía probar en dicho sistema. Además, demostró que suponiendo la consistencia del mismo y la verificación de una serie de propiedades básicas, el sistema no era capaz de probar su propia consistencia. La esperanza de

Fecha de recepción: 31 de diciembre de 2003

responder afirmativamente a la tercera cuestión planteada por D. Hilbert quedó truncada en 1936 cuando A. Church y A. Turing dieron una respuesta negativa al *Entscheidungsproblem* (problema de decisión relativo a un sistema axiomático: decidir, para cualquier aserto, si el sistema puede probarlo o no).

A lo largo del siglo XX se han desarrollado modelos abstractos para máquinas que tratan de implementar procedimientos de cálculo (máquinas de Turing, RAM, URM, etc.), para lenguajes formales (regulares, libres de contexto, recursivamente enumerables, etc.), para sistemas de programas (compiladores, sistemas operativos, etc.), para estructuras de datos (pilas, colas, montículos, etc.) y para bases de datos (relacionales, orientadas a objetos, etc.).

Definir un *modelo de computación* consiste, básicamente, en formalizar el concepto de *procedimiento mecánico*. Para ello, es necesario precisar sintácticamente qué es un tal procedimiento y definir con rigor la semántica del mismo. La semántica de un modelo de computación debe de ir acompañada del concepto de *resolución* de un problema, a través de distintos *modos* de computación (determinista, no determinista, secuencial, paralelo, aproximado, etc.).

Los trabajos de K. Gödel, A. Church, S. Kleene y A. Turing, entre 1931 y 1936, proporcionan las primeras formalizaciones del concepto de *algoritmo*, dando lugar a la aparición de los primeros *modelos de computación*. Concretamente, en 1931 K. Gödel define el concepto de *relación recursiva* e introduce la clase de funciones que denominó *recursivas* (y que hoy se conocen como *primitivas recursivas*). Posteriormente, en 1934, K. Gödel y S. Kleene extienden la clase anterior a las funciones *general recursivas* (que son las que conocemos hoy con el nombre de *recursivas*). En 1931, A. Church y S. Kleene desarrollan el concepto de λ -cálculo relacionándolo directamente con el concepto intuitivo de *función computable*. En 1936, A. Turing utiliza por primera vez el concepto abstracto de *máquina* a fin de formalizar la idea intuitiva de algoritmo.

Hacia mediados de la década de los treinta, A. Church y A. Turing formulan de manera independiente que las definiciones que han dado del concepto de *procedimiento mecánico* capturan completamente la idea intuitiva del mismo (*tesis de Church-Turing*). En 1936, A. Turing establece la equivalencia de su modelo y el de A. Church. Así, la tesis citada puede expresarse así: *toda función computable de manera efectiva puede ser calculada por una máquina de Turing*.

En 1936, A. Church proporciona el primer ejemplo de un problema irresoluble algorítmicamente (según la formalización del λ -cálculo): el problema de *decisión* relativo a la *lógica de primer orden*, respondiendo negativamente a la tercera cuestión formulada por D. Hilbert. Pocos meses después, A. Turing establece de manera independiente el mismo resultado probándolo en su modelo.

El auge y desarrollo de dispositivos computacionales teóricos contribuirían de manera decisiva a la construcción de los primeros ordenadores electrónicos a finales de los cuarenta del pasado siglo, basados en el modelo conceptual de J. von Neumann, A. W. Burks y H. H. Goldstine descrito en 1946 [4], en donde se formulan las bases teóricas de los futuros ordenadores que son susceptibles de manipular *programas almacenados*.

Y ya en los comienzos se empezaron a utilizar algunas ideas de la biología en el marco computacional (por ejemplo, a través de las redes neuronales

artificiales de McCulloch y Pitts, en 1943). Desde entonces se han realizado grandes esfuerzos para entender, desde una óptica computacional, los procesos que se dan en la naturaleza, encaminados a la obtención de algoritmos más eficientes o incluso para el diseño de nuevos tipos de ordenadores.

La *Bioinformática* es una disciplina que estudia la aplicación de herramientas y técnicas de la información para la manipulación de datos biológicos. Entre sus objetivos destacamos el desarrollo de una base de datos específica que permita almacenar y actualizar la gran cantidad de datos que han sido y están siendo generados de manera continua e ininterrumpida; el diseño de algoritmos eficientes para la comparación de sucesiones de cadenas; la investigación de métodos de predicción de la estructura tridimensional de las moléculas orgánicas, en general, y de las proteínas, en particular; y, finalmente, la definición de modelos físicos y/o matemáticos de sistemas biológicos. y Durante la última década, el campo de investigación denominado *Computación Natural* ha experimentado un enorme desarrollo. Se trata de una disciplina cuyo objetivo es la simulación e implementación de procesos dinámicos que se dan en la naturaleza y que son susceptibles de ser interpretados como procedimientos de cálculo. Esta disciplina surge de la investigación de modelos matemáticos y de las herramientas tecnológicas necesarias para la implementación de la computación bio-inspirada. Entre las áreas que se enmarcan dentro de la Computación Natural, destacamos las siguientes:

1. Los *algoritmos genéticos* (o más en general, la *computación evolutiva*), introducidos por J. Holland [7] en 1975, que hacen uso de algunas operaciones inspiradas en la evolución y en la selección natural a fin de encontrar una *buena* solución a partir de una gran cantidad de posibles soluciones candidatas.
2. Las *redes neuronales artificiales*, introducidas por W.S. McCulloch y W. Pitts [11] en 1943, que están inspiradas en las interconexiones y en el funcionamiento de las neuronas en el cerebro.
3. La *computación molecular*, cuyo objetivo consiste en usar moléculas orgánicas (ADN, ARN, proteínas, etc.) como hardware biológico que permite realizar computaciones. Esta disciplina nace a finales de 1994 con los trabajos de L. Adleman [1], si bien tiene precedentes en un trabajo de T. Head [6] en el que formula el *sistema splicing* (modelo teórico de procesamiento de moléculas de ADN con la participación de enzimas).
4. La *computación celular con membranas*, introducida por Gh. Păun [15] en 1998, que está inspirada en la estructura y el funcionamiento de las células de los organismos vivos, en cuanto a su capacidad para procesar y generar información.

Los algoritmos genéticos y las redes neuronales artificiales han sido implementadas a través de programas en ordenadores electrónicos convencionales. La computación molecular basada en ADN ha sido implementada en medios bio-químicos (el experimento de L. Adleman permitió resolver una instancia con-

creta del problema del camino hamiltoniano, en su versión dirigida y con dos nodos distinguidos, a través de la manipulación de moléculas de ADN en el laboratorio). En cambio, la computación celular con membranas aún no ha sido implementada ni en medios electrónicos ni bioquímicos.

Recientemente, en octubre de 2003, el Institute for Scientific Information (ISI, USA), ha designado a la *computación celular con membranas* como *Fast Emerging Research Front* en el área de *Computer Science*.

2 Una visión computacional de las células

La célula es la unidad fundamental de todo organismo vivo. Posee una estructura compleja y, a la vez, muy organizada que permite la ejecución simultánea de un gran número de reacciones químicas.

Existen dos tipos de células: las *procariotas* (propias de los organismos unicelulares, como las bacterias) que carecen de un núcleo bien definido, y las *eucariotas* (específicas de los animales y plantas) que poseen un núcleo rodeado por una doble membrana. Ambas realizan de manera similar una serie de procesos que son esenciales para la vida, tales como la replicación del ADN, la producción de energía, la síntesis de proteínas y los procesos metabólicos.

En un primer análisis podemos distinguir tres partes claramente diferenciadas en una célula: una especie de piel (*membrana plásmica*) que delimita a la célula de su entorno; el corazón propiamente dicho (*núcleo*), que almacena la información genética a través de moléculas de ADN y de ARN; y el resto de la célula, denominado *citoplasma*.

A su vez, en el citoplasma de las células eucariotas podemos destacar las siguientes componentes fundamentales: la *mitocondria*, encargada de la producción de energía para la célula; el *aparato de Golgi*, que viene a ser una fábrica de proteínas y juega un papel esencial en el metabolismo de la célula; el *retículo endoplásmico*, que es una red de membranas interconectadas estructurada en dos partes: una que facilita el paso del ARN mensajero del núcleo a fin de producir la síntesis proteínica, y otra que se encarga del transporte de moléculas y la comunicación entre las componentes de la célula; y, finalmente, los *lisosomas*, que constan de una única membrana y se encargan de digerir sustancias que proceden del exterior y de ingerir restos celulares que han llegado a ser obsoletos para el organismo.

A lo largo de este trabajo se van a estudiar dispositivos computacionales inspirados en la estructura y funcionamiento de las células eucariotas.

2.1 Las membranas biológicas

Las *membranas biológicas* son estructuras dinámicas básicas para la célula y juegan un papel esencial a la hora de definir el fenómeno que usualmente denominamos como *vida*.

Una membrana (denominada *plasma*) separa el espacio interno de la célula protegiéndolo de su entorno. Las restantes membranas (denominadas *internas*) proporcionan la estructura jerarquizada propia de la célula y ofrecen la

adecuada protección al núcleo, que es el depositario de la información genética. Las membranas están involucradas de manera decisiva en la mayoría de reacciones químicas que tienen lugar en los compartimentos celulares y pueden considerarse como barreras semipermeables que, o bien permiten el paso de algunas sustancias químicas en cualquiera de los dos sentidos (dentro-fuera o fuera-dentro), o bien bloquean el paso de sustancias. Las membranas se comportan como canales selectivos de comunicación para la transferencia de compuestos químicos entre distintos compartimentos, así como entre la propia célula y su entorno, controlando así un cierto flujo de datos (es decir, de información).

El premio Nobel de química del año 2003, ha sido otorgado a los científicos P. Agre y R. MacKinnon por sus descubrimientos relacionados con los canales de las membranas celulares, que han supuesto un gran avance en los estudios de la química celular e ilustran su importancia en los procesos vitales.

Cada compartimento celular (delimitado y separado del resto por una membrana) puede considerarse como una unidad computacional (un *procesador*) con sus propios datos (compuestos químicos) y su propio programa local (reacciones químicas), de tal manera que el conjunto de compartimentos, considerado como una unidad global (la célula), puede ser interpretado como un *modelo de computación no convencional*.

A lo largo de la breve historia de la teoría de la computación, muchos avances, tanto en el plano teórico como en el de aplicaciones prácticas, se han producido bajo la inspiración de los procesos que se dan en la naturaleza. ¿Qué puede proporcionar la célula como fuente de inspiración computacional? ¿Se puede abstraer una especie de modelo de computación a partir de las células vivas? ¿Es posible implementar computaciones a través de las células, con la esperanza de disponer algún día de un *computador celular*, de propósito general?

En las ideas originales de Gh. Păun, los sistemas celulares con membranas no fueron introducidos propiamente para modelizar completamente la estructura y funcionamiento de una célula, sino más bien para analizar algunos hechos computacionalmente relevantes que pueden ser abstraídos de las mismas.

Los sistemas P (dispositivos computacionales de los modelos celulares) constituyen un modelo teórico de computación inspirados en la realidad de las células vivas. No obstante, existe la esperanza de que la computación celular con membranas proporcione un marco de simulaciones de fenómenos moleculares. En este contexto, se está trabajando en el desarrollo del hardware y simulaciones biológicas. Así el proyecto E-CELL, iniciado en 1996, tiene como objetivo la modelización y simulación de procesos celulares tales como trayectorias metabólicas, síntesis de proteínas y transporte a través de las membranas, con el fin de predecir el comportamiento de las células vivas.

Uno de los retos más importantes planteado en este nuevo modelo consiste en encontrar una posible implementación biológica del mismo.

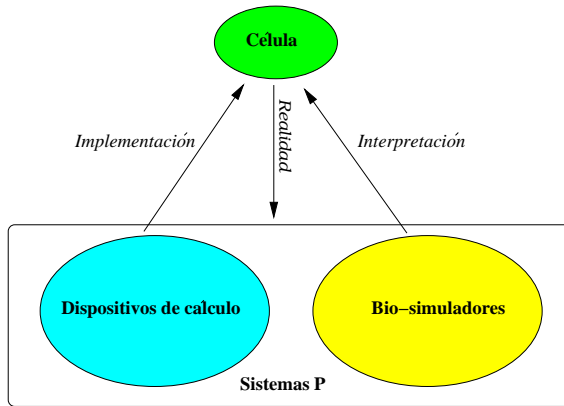


Figura 1: Sistemas celulares

3 Sistemas celulares: Una descripción informal

En este trabajo se trata de modelizar las células vivas y su funcionamiento a través de los *sistemas celulares con membranas*. En esta sección vamos a dar una descripción informal de los sistemas celulares básicos de transición.

Una *estructura de membranas* consta de un conjunto finito de membranas que se encuentran dentro de una membrana principal (denominada *piel*) y delimita una serie de *regiones* o compartimentos formados por el espacio comprendido entre una membrana y las situadas inmediatamente dentro de ella (si las hay). Una membrana es *elemental* si no contiene ninguna membrana.

Una estructura de membranas puede ser descrita como una cadena sobre un alfabeto finito y se puede representar mediante diagramas de Venn a través una familia de subconjuntos de un conjunto base, tal que dos conjuntos cualesquiera de la familia o son disjuntos, o uno de ellos está contenido en el otro. También se puede considerar como un árbol cuya raíz es la membrana *piel* y las hojas son las membranas elementales.

Las sustancias químicas contenidas en las distintas regiones de las membranas celulares se representan mediante símbolos de un alfabeto finito. Por tanto, los objetos son símbolos atómicos, indivisibles, sin *estructura interna*, pero que pueden *evolucionar* a lo largo del tiempo.

Teniendo presente que las diferentes sustancias químicas pueden aparecer en una región en diferentes cantidades, representaremos el contenido de cada región a través de *multiconjuntos* de objetos (conjuntos cuyos elementos pueden aparecer repetidos, es decir, con multiplicidades asociadas a cada uno de ellos).

Las reacciones químicas que se producen en las regiones se representan a través de *reglas de evolución* que, básicamente, son reglas de reescritura de un sistema formal que actúan sobre los objetos modificándolos y/o trasladándolos

de un compartimento a otro de acuerdo con una determinada semántica.

Una *configuración* de un sistema celular consta de una estructura de membranas y una familia de multiconjuntos de objetos asociados a cada región. En la descripción de un tal sistema, aparece siempre una configuración denominada *inicial*. En algunas variantes existe una *membrana de entrada* que permite añadir un multiconjunto de objetos antes de que se produzcan evoluciones.

Las reglas de evolución serán aplicadas de manera *no determinista, paralela y maximal*; es decir, las reglas a ser usadas y los objetos involucrados en la misma serán escogidos de tal manera que en cada paso de computación todos los objetos que *puedan evolucionar tienen* que evolucionar. Los objetos pueden pasar de una región a otra a través de las membranas y éstas pueden ser disueltas, creadas o divididas. Así se produce una *transición* de una configuración a otra.

Una *computación* en un sistema celular es una sucesión (finita o no) de configuraciones tal que la primera de ellas es una configuración inicial y toda configuración de la sucesión que no sea la inicial se obtiene de la anterior mediante un paso de transición. Se dice que una configuración es de *parada* si no existe ninguna regla de evolución que sea aplicable a dicha configuración. Se dice que una *computación* es de *parada* si contiene una configuración de parada, y suele tener asociado un *resultado* codificado en una membrana elemental distinguida (denominada de *salida*) de la correspondiente configuración de parada.

Así pues, una computación en un sistema celular parte de una configuración y mediante sucesivas transiciones, o bien se llega a una situación de parada (en tanto que el sistema no puede seguir evolucionando), o bien, el sistema nunca se detiene (en tanto que sigue evolucionando de manera indefinida).

Admitiremos que toda computación ejecuta un proceso sincronizado; es decir, se supone que hay una especie de *reloj universal* que marca las actuaciones de todos los elementos que integran el sistema celular. Hay que destacar la existencia de dos niveles de paralelismo en la ejecución: por una parte, las reglas asociadas a una membrana son aplicadas de manera simultánea; y, por otra, estas operaciones son realizadas en el mismo instante en todas las membranas que vertebran el sistema. Es decir, en un paso de reloj se ejecutarán todas las reglas que pueden ser aplicadas (de manera maximal), y en todas las membranas.

Los sistemas celulares con membranas constituyen un modelo de computación que conjuga la elegancia y sencillez del modelo, con su proximidad a la realidad de la célula, desde un punto de vista biológico, abstrayendo de la estructura y funcionamiento de las células una serie de ideas computacionalmente relevantes. Son modelos matemáticos con propiedades atractivas desde el punto de vista computacional: algunos sistemas celulares son *completos* (tienen la misma potencia computacional que las máquinas de Turing) y *eficientes* (capaces de proporcionar soluciones polinomiales de problemas NP-completos).

Aunque los sistemas celulares tienen una inspiración biológica, debe resaltarse el hecho de que constituyen igualmente un buen modelo teórico de computación distribuida, en donde distintas unidades de cálculo trabajan independientemente pero estructuradas en una cierta jerarquía vertical. Por ejemplo, la jerarquización usada en las conexiones establecidas en redes de ordenadores como Internet puede ser representada como una estructura de membranas, en

la que los nodos de la red se interpretan como las membranas que forman el sistema, y el flujo de información entre nodos conectados se interpreta como las componentes químicas que dichas membranas generan e intercambian.

Asimismo, sistemas dinámicos complejos, como aquellos que surgen en el estudio de la dinámica de poblaciones, pueden ser también interpretados como sistemas celulares en los que los individuos de las distintas especies interactúan entre sí dentro de hábitats que permiten el traspaso de los mismos.

Como primera nota diferenciadora con respecto a otros modelos de computación natural (por ejemplo, los modelos de computación molecular basados en ADN) conviene hacer notar que no está descrito a través de un lenguaje de programación; es decir, no proporciona una serie de operaciones básicas susceptibles de ser secuenciadas sobre un dato de entrada para obtener un resultado final (solución del problema que se pretende resolver), sino que se generan *dispositivos* (al modo de las máquinas de Turing) cuya ejecución modifica el contenido de las distintas componentes que los integran hasta llegar, en su caso, a un estado de parada (en el que la máquina deja de funcionar).

3.1 Estructuras de membranas

Recordemos que un *alfabeto*, A , es un conjunto no vacío, y una *cadena* sobre A es una sucesión finita de elementos de A . Notaremos por A^* el conjunto de todas las cadenas sobre A . Un *lenguaje* sobre A es un subconjunto de A^* .

Para introducir las *estructuras de membranas*, en primer lugar se define recursivamente el lenguaje MS sobre el alfabeto $\{[,]\}$, como sigue: es el menor lenguaje L sobre $\{[,]\}$ que verifica las condiciones siguientes:

1. $[] \in L$.
2. Si $\mu_1, \dots, \mu_n \in L$, entonces $[\mu_1 \dots \mu_n] \in L$.

Intuitivamente las estructuras de membranas se describen a través de pares de corchetes *coincidentes*, imponiendo que dentro de las membranas, el orden de aparición no sea relevante. Por ello, se define una relación, \sim , en MS de forma que estén relacionadas aquellas cadenas que reflejen la misma estructura de corchetes (salvo el orden): $x \sim y \Leftrightarrow x = \mu_1\mu_2\mu_3\mu_4 \wedge y = \mu_1\mu_3\mu_2\mu_4 \wedge \mu_1\mu_4, \mu_2, \mu_3 \in MS$. Si notamos por \sim^* la clausura transitiva y reflexiva de la relación \sim , entonces una *estructura de membranas* es una clase de equivalencia de MS con respecto a la relación \sim^* . Dada una estructura de membranas, a cada par de *corchetes coincidentes* se le denomina *membrana*. El número de membranas de una estructura de membranas, μ , se denomina *grado* de la estructura. A la estructura *más externa* se le llama *piel*, y aquellas membranas que no contienen otras membranas en su interior se denominan *elementales*.

Se puede representar gráficamente la estructura de membranas como un conjunto de curvas simples cerradas (a modo de *diagrama de Venn*) manteniendo la jerarquía impuesta por la cadena que la representa en el conjunto cociente $\overline{MS} = MS / \sim^*$ (figura 2).

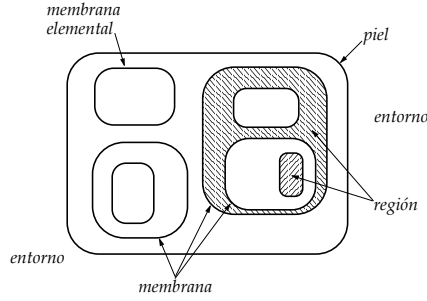


Figura 2: Representación de una estructura de membranas.

A partir de esta representación surge el concepto intuitivo de *región* delimitada por una membrana, como el espacio comprendido entre dicha membrana y las inmediatamente interiores. Existe una biyección natural entre las membranas de una estructura y las regiones que delimitan, siendo habitual identificar unas con otras.

3.2 Sistemas celulares de transición

A continuación se introducen los sistemas P de transición siguiendo las ideas del trabajo fundacional de Gh. Păun[15].

Definición 1 *Un sistema P de transición de grado $p \geq 1$ es una tupla*

$$\Pi = (\Gamma, \mu_{\Pi}, \mathcal{M}_1, \dots, \mathcal{M}_p, (R_1, \rho_1), \dots, (R_p, \rho_p), i_0)$$

donde:

- Γ es un alfabeto finito (alfabeto de trabajo).
- μ_{Π} es una estructura de membranas (que consta de p membranas). Las membranas de μ_{Π} están etiquetadas de forma unívoca usando los números naturales desde 1 hasta p .
- \mathcal{M}_i ($1 \leq i \leq p$) es un multiconjunto finito sobre Γ asociado a la membrana i del sistema.
- R_i ($1 \leq i \leq p$) es un conjunto finito de reglas de evolución de transición asociado a la membrana i del sistema. Una regla de evolución de transición es un par (u, v) , habitualmente representado $u \rightarrow v$, donde u es una cadena sobre Γ y $v = v'$ o $v = v'\delta$, siendo v' una cadena sobre $\Gamma \times (\{here, out\} \cup \{in_i : i = 1, \dots, p\})$.
- ρ_i ($1 \leq i \leq p$) es un orden parcial estricto sobre R_i , que establece una relación de prioridad entre las reglas de R_i .

- i_0 es un número natural entre 1 y p (membrana de salida del sistema).

Para describir de manera informal la semántica del modelo se introduce el concepto de configuración, de donde seguirá, una vez establecida la forma en que se aplican las reglas de transición, la noción de computación del sistema.

Una *configuración* de Π es una tupla $(\mu, M_{i_1}, \dots, M_{i_q})$ tal que μ es la estructura de membranas que se obtiene de μ_Π al eliminar las membranas distintas de i_1 a i_q (la membrana piel no se puede eliminar, por lo que coincide en ambas estructuras); y M_{i_j} es un multiconjunto finito sobre Γ , para cada $j = 1, \dots, q$.

La *configuración inicial* de Π es la tupla $(\mu_\Pi, \mathcal{M}_1, \dots, \mathcal{M}_p)$.

La *ejecución de una regla* $u \rightarrow v$ asociada a una membrana i presente en una configuración se realiza como sigue: los objetos en u se eliminan de la membrana i (ésta debe contener, por tanto, suficientes objetos para que la regla se pueda aplicar); entonces, para cada $(a, out) \in v$ un objeto $a \in \Gamma$ se incluye en la membrana padre de la membrana i (o abandona el sistema si la membrana i es la piel); para cada $(a, here) \in v$ un objeto $a \in \Gamma$ se añade a la membrana i ; para cada $(a, in_j) \in v$ un objeto $a \in \Gamma$ se introduce en la membrana j (teniendo en cuenta que si la membrana j no es hija de la membrana i , entonces la regla no se puede aplicar); finalmente, si $\delta \in v$, entonces la membrana i se disuelve; es decir, se elimina de la estructura de membranas, pasando sus objetos a su membrana padre y desapareciendo las reglas de evolución y relaciones de prioridad asociadas a ella (la membrana piel no se puede disolver).

Dada una regla $u \rightarrow v$, la longitud de la cadena u se denomina *radio* de la regla. Se dice que un sistema celular es *cooperativo* si posee, al menos, una regla de evolución de radio mayor que uno.

Las *relaciones de prioridad* dadas sobre las reglas se interpretarán como sigue: si la regla r_1 tiene mayor prioridad que la regla r_2 y se puede aplicar r_1 , entonces no se aplicará r_2 , aunque fuera factible. Una posible interpretación de esta versión se encuentra en el consumo de energía: en cada paso de transición tenemos una cantidad fija de energía para poder aplicar las reglas, de tal manera que las reglas de prioridad superior consumen la suficiente energía como para que no quede energía para reglas de prioridad inferior.

Dadas dos configuraciones, C y C' , de Π , diremos que C' se obtiene de C en un *paso de transición*, y notaremos $C \Rightarrow_\Pi C'$, si C' es el resultado de aplicar a C , en *paralelo* y para todas las membranas al mismo tiempo, las reglas de evolución contenidas en un determinado (multi)conjunto de reglas asociadas a las membranas que aparecen en la estructura de membranas de C . En dicho (multi)conjunto se indica el número de veces que se aplica cada una de las reglas y debe ser *maximal*, en el sentido de que, tras la aplicación de las reglas, en ninguna membrana permanezcan objetos sin evolucionar que puedan activar alguna de las reglas asociada a la membrana. Puesto que para una configuración usualmente existe más de un (multi)conjunto de reglas aplicable, los pasos de transición se realizan de manera no determinista.

Una *computación* \mathcal{C} de Π es una sucesión (finita o no) de configuraciones, $\{C^i\}_{i < r}$, tal que C^0 es la configuración inicial de Π ; $C^i \Rightarrow_\Pi C^{i+1}$, para todo $i < r - 1$; y, o bien r es un número natural no nulo y no existe una regla que se

pueda aplicar en ninguna de las membranas de C^{r-1} (en cuyo caso se dice que \mathcal{C} es de parada, ha realizado $r - 1$ pasos y C^{r-1} es su *configuración de parada*), o bien $r = \infty$ (en cuyo caso se dice que \mathcal{C} no es de parada).

Diremos que una *computación* \mathcal{C} es *exitosa* si es una computación de parada y, además, la membrana de salida i_0 aparece en C^{r-1} como membrana elemental. A cada computación exitosa de Π , se le puede asociar una salida (por ejemplo, el tamaño del multiconjunto asociado a la membrana de salida del sistema en su configuración de parada). Entonces, un sistema P de transición es una máquina que genera el conjunto de números naturales formado por las salidas de todas las computaciones exitosas de dicho sistema.

A partir del modelo de computación celular básico se han diseñado múltiples variantes. A continuación hacemos un breve análisis de algunas de las variantes que se han considerado hasta la actualidad.

3.3 Variantes en los objetos

En los sistemas celulares de transición vistos hasta ahora se puede interpretar que no se hace uso de estructuras de datos para codificar la información: los números son codificados a través de la cardinalidad de multiconjuntos; por consiguiente, están representados en base 1. Esto puede ser adecuado para una implementación bioquímica, pero parece claramente insuficiente desde un punto de vista clásico. Más aún, por esta vía no se podrá trabajar (directamente, sin una codificación numérica) con computaciones simbólicas.

Por ello, se trata de usar sistemas que representen la información a través de una estructura de datos de tipo estándar: las cadenas. Así, en lugar de considerar objetos de tipo atómico se pueden usar objetos que pueden ser descritos por cadenas sobre un alfabeto finito prefijado. En tal situación, la evolución de un objeto correspondería propiamente a una transformación de la cadena.

3.3.1 Sistemas celulares con reescritura.

En estos sistemas celulares se procesan las cadenas de objetos a través de reglas del tipo $r \equiv X \rightarrow (v; tar)$, donde X es un símbolo, v es una cadena de símbolos y tar es la membrana destino de dicha cadena (eventualmente, la cadena v puede ir seguida de un símbolo δ , de disolución). La forma de considerar la ejecución de este tipo de reglas (que actúan en el *interior* de las cadenas) es la habitual en sistemas P (la aplicación de r sobre una cadena w reemplaza una ocurrencia de X por la cadena v , disolviendo la membrana si así se indica en r), con la siguiente puntualización: todas las cadenas son procesadas en paralelo, pero sobre cada cadena presente en la membrana sólo actúa una regla [9].

3.3.2 Sistemas celulares con replicación de cadenas.

Este tipo de sistemas es una extensión del procedimiento de reescritura anterior: las reglas combinan la posibilidad de actuar en el interior de las cadenas con la capacidad de replicación de las mismas [10]. De esta forma, una regla adquiere

la expresión general $r \equiv X \rightarrow (u_1, tar_1) || \dots || (u_n, tar_n)$, y su aplicación sobre una cadena w produce la aparición de n cadenas, w_1, \dots, w_n , de modo que w_i se obtiene de w reemplazando una aparición de X por u_i , y su destino es tar_i .

Esta variante de los sistemas P de transición ha sido usada para dar soluciones deterministas, y con coste en tiempo polinomial, de problemas **NP**-completos clásicos, como son el problema **SAT** y el problema **HPP** del camino hamiltoniano [25]. No obstante, la parte exponencial del proceso queda encubierta en la aplicación de reglas que permiten la replicación de las cadenas.

3.4 Variantes en las membranas

Otra posibilidad de mejorar las características de los sistemas celulares es añadir funcionalidad y dinamismo a la estructura de membranas, permitiendo la creación y división de membranas por medio de reglas.

3.4.1 División de Membranas.

Este tipo de variante es conocida por el nombre de *sistemas celulares con membranas activas*. Se trata de un dispositivo útil para conseguir una cantidad exponencial de espacio, en tiempo polinomial, partiendo de una estructura de membranas descrita en forma polinomial (en función del dato de entrada).

En esta variante se permite el uso de reglas que implementan la división (análoga a la producida en la naturaleza) de membranas elementales (siempre que no sea la piel) en otras dos que contienen un duplicado del contenido de la membrana original (salvo el objeto que dispara la regla, que evoluciona a otro objeto en las nuevas membranas), y tienen asociadas el mismo conjunto de reglas. Se admite disolución, pero no cooperación ni prioridad entre reglas.

Si en esta variante se permite la división de membranas no elementales, el sistema celular se denomina *extendido*. Se prueba [13] que el problema **HPP** puede ser resuelto en tiempo cuadrático y el problema **SAT** en tiempo lineal, por medio de sistemas P extendidos con membranas activas.

Este resultado ha sido mejorado por C. Zandron, C. Ferretti, G. Mauri [26], y por M.J. Pérez Jiménez, A. Romero Jiménez y F. Sancho Caparrini [23] que han obtenido soluciones lineales a través de sistemas celulares que sólo usan reglas de división para las membranas elementales.

3.4.2 Creación de membranas.

La creación de membranas es un proceso habitual en la biología. En relación con los sistemas celulares, se considera la posibilidad de crear nuevas membranas por la acción de ciertos objetos. En este sentido, se pueden introducir reglas en las que un determinado multiconjunto de objetos produce, en el interior de la membrana a la que pertenece, una nueva membrana con un cierto contenido. Este tipo de creación de membranas puede ser aplicado tanto a sistemas que usan objetos atómicos como a aquellos que usan objetos de tipo cadena.

Esta variante de sistemas fue considerada por M. Ito, C. Martín-Vide y Gh. Păun [8] para conseguir resultados de universalidad en los que se permitía un

uso bastante moderado de esta nueva operación. También puede ser usada para la obtención de soluciones polinomiales de problemas **NP**-completos.

3.5 Variantes en las comunicaciones

Hay un tercer tipo de variantes que estudian el papel que juega la comunicación entre las membranas (por medio de los elementos que son transferidos de unas a otras). Para ello, se intenta debilitar alguna capacidad inherente a los sistemas básicos de transición con el fin de estudiar cuáles son las características mínimas necesarias para seguir disponiendo de un modelo universal (lo que podría ser útil si se consigue una implementación práctica de estos modelos).

Una primera variante consiste en cambiar el destino in_j en las reglas de los sistemas celulares por un destino más general, in , que no especifique a qué membrana han de ir los objetos producidos, de modo que dicha elección es no determinista entre las posibles membranas existentes que sean hijas de la membrana a la que está asociada la regla. Se debe hacer notar que muchos de los resultados de universalidad obtenidos para variantes mejoradas se siguen manteniendo para esta variante débil de sistemas P, debido a que en dichas pruebas se hace uso de estructuras de membranas de grado 2, en las que no es necesario especificar qué membrana hija es la que recibe los objetos.

Otra variante introducida en esta dirección consiste en dotar de una carga eléctrica a los objetos y membranas del sistema P (con tres cargas posibles: $+$, $-$, 0), de tal modo que se elige *a priori* una carga en las membranas (en la definición del sistema celular) y los objetos pueden modificar sus cargas de acuerdo con las reglas de evolución que se aplican sobre ellos. Así, la relación entre las cargas de los objetos y las membranas determina si un objeto puede atravesar una membrana. Para más detalles acerca de estas dos variantes, ver [18], capítulo 3, sección 3.6.1.

Por otra parte, es interesante observar que las membranas biológicas tienen una permeabilidad que es variable y, por tanto, se puede considerar un modelo en el que se controle el paso de objetos de una membrana a otra mediante un parámetro de *permeabilidad de una membrana*, como valor numérico que indica su proximidad a un estado que permite su disolución. En dichos sistemas [17], se puede modificar esa permeabilidad por medio de reglas, de tal manera que la ejecución de una regla puede hacer que la membrana engrose o adelgace. Si una membrana tiene un grosor muy elevado, se vuelve *impermeable*, impidiendo la transferencia de objetos a través de ella. Así por ejemplo, el control de la permeabilidad de las membranas se puede implementar introduciendo una acción, que notaremos τ , y de tal manera que es opuesta a δ , en el sentido de que *engrosa* la membrana, la aleja de la disolución.

En los *sistemas celulares con transportadores (carriers)* [25], los objetos no son transformados, sino que únicamente son transmitidos entre las membranas de acuerdo con ciertas reglas de transferencia. Usualmente, en esta variante se permite la transmisión con objetos del *entorno* del sistema, con el fin de poder disponer de una cantidad ilimitada de los mismos.

Una variante interesante de los sistemas con transportadores son los *sistemas*

celulares con porteros (porters) [14], en los que se restringe la capacidad en la transmisión de objetos, reduciendo el *radio* de las reglas.

4 Una formalización de los sistemas P de transición

El problema de describir una formalización completa de los sistemas celulares, así como de sus computaciones y los resultados de las mismas, fue formulado explícitamente por Gh. Păun en [16]. En este trabajo se presenta una formalización con una orientación distinta de las propuestas parciales dadas por A. V. Baranda y otros en [2] y por A. Obtulowicz en [12]. Asimismo es ligeramente diferente de la presentada por M. J. Pérez Jiménez y F. Sancho en [20].

4.1 Una sintaxis para los sistemas celulares de transición

En primer lugar, vamos a introducir una serie de conceptos que serán de utilidad para la formalización que se va a presentar.

Un *multiconjunto* sobre un conjunto, A , es una aplicación $m : A \rightarrow \mathbb{N}$, siendo \mathbb{N} el conjunto de los números naturales. Un subconjunto $B \subseteq A$ se puede identificar con el multiconjunto sobre A dado por la función característica de B (en A). Notaremos por $\mathbf{M}(A)$ el conjunto de todos los multiconjuntos sobre A . El *soporte* de un multiconjunto m es el conjunto $\{a \in A : m(a) > 0\}$. Un *multiconjunto* se dice que es *finito* si su soporte es un conjunto finito, y se dice que es vacío si lo es su soporte. El *tamaño* de un multiconjunto finito, m , se define por $|m| = \sum_{a \in A} m(a)$ (obsérvese que dicha suma es finita).

Definición 2 Sean m_1 y m_2 multiconjuntos sobre A .

- *Inclusión:* $m_1 \leq m_2 \Leftrightarrow \forall j (j \in A \rightarrow m_1(j) \leq m_2(j))$.
- *Inclusión estricta:* $m_1 < m_2 \Leftrightarrow m_1 \leq m_2 \wedge m_1 \neq m_2$.
- *Unión:* La aplicación $+$: $\mathbf{M}(A) \times \mathbf{M}(A) \rightarrow \mathbf{M}(A)$, se define como sigue: $(m_1 + m_2)(j) = m_1(j) + m_2(j)$, para cada $j \in A$.
- *Diferencia:* La aplicación $-$: $\mathbf{M}(A) \times \mathbf{M}(A) \rightarrow \mathbf{M}(A)$, se define como sigue: $(m_1 - m_2)(j) = \max\{m_1(j) - m_2(j), 0\}$, para cada $j \in A$.
- *Amplificación:* La aplicación \otimes : $\mathbb{N} \times \mathbf{M}(A) \rightarrow \mathbf{M}(A)$, se define como sigue: $(n \otimes m_1)(j) = n \cdot m_1(j)$, para cada $j \in A$.

Todo multiconjunto finito, m , con soporte $\{a_1, \dots, a_n\}$, se puede representar mediante una cadena w sobre A en donde a_i aparece exactamente $m(a_i)$ veces. Las cadenas obtenidas a partir de w permutando sus símbolos también representan a m . El multiconjunto vacío se representa por la cadena vacía, λ .

Dado un grafo no dirigido, G , diremos que un nodo v es *alcanzable* desde otro nodo u , y notaremos por $u \rightsquigarrow_G v$, si existe un camino desde u hasta v .

Sea $G = (V, E)$ un árbol enraizado de raíz r . Para cada nodo, u , que no sea la raíz, notaremos por $f(u)$ (el *padre* de u en G) el único antecesor propio de u

que verifica $\{f(u), u\} \in E$. Asimismo, notaremos por $Ch(u)$ (los hijos de u en G) el conjunto de descendientes propios de u tales que u es su padre. Además, se puede definir de manera natural una relación binaria $E^*(G)$ en V como sigue: $(x, y) \in E^*(G) \Leftrightarrow y \in Ch(x)$. Esta relación establece, de alguna forma, un direccionamiento sobre las aristas del árbol enraizado.

Definición 3 Una estructura de membranas es un árbol enraizado en el que los nodos se denominan membranas, la raíz se denomina piel, y las hojas se denominan membranas elementales.

En esta formalización las etiquetas de las membranas serán los propios vértices del grafo. Si esta formalización se extiende a otras versiones de sistemas celulares, puede ser útil un etiquetado independiente de los nodos; no obstante, para la versión que estamos formalizando esto no es necesario.

Definición 4 Una célula sobre un alfabeto, A , es un par (μ, M) , en donde μ es una estructura de membranas y M es una aplicación de $V(\mu)$ en $\mathbf{M}(A)$, por la que a cada nodo del árbol se le asocia un multiconjunto sobre el alfabeto base.

Si $C = (\mu, M)$ es una célula sobre A , denotaremos $\mu = (V(\mu), E(\mu))$; es decir, $V(\mu)$ y $E(\mu)$ son los conjuntos de vértices y aristas, respectivamente, del árbol enraizado etiquetado, μ , que determina la célula. El número total de membranas (o nodos) de la célula, se denomina *grado* de dicha célula.

Definición 5 Sea (μ, M) una célula sobre un alfabeto A . Una regla (de evolución) asociada a dicha célula es una tupla $r = (d_r, v_r, \delta_r, x_r)$, en donde: d_r es un multiconjunto sobre A ; v_r es una función de dominio $V(\mu) \cup \{here, out\}$ y rango contenido en $\mathbf{M}(A)$, en donde $here, out \notin V(\mu)$ y $here \neq out$; $\delta_r \in \{\neg\delta, \delta\}$, con $\neg\delta, \delta \notin A$ y $\neg\delta \neq \delta$; y x_r es un nodo de μ .

Definición 6 Sea $C = (\mu, M)$ una célula sobre un alfabeto y $x \in V(\mu)$. Sea $r = (d_r, v_r, \delta_r, x_r)$ una regla asociada a C . Diremos que r está asociada a x si se verifica que $x_r = x$.

Finalmente, para dar la sintaxis completa del sistema P se necesita asignar a cada membrana de la estructura un conjunto de reglas que serán las que se puedan aplicar a los elementos presentes en esa membrana.

Definición 7 Sea $C = (\mu, M)$ una célula sobre un alfabeto A . Una colección R de reglas asociada a C es una función con dominio $V(\mu)$ tal que para cada membrana $x \in V(\mu)$, $R(x)$ (que notaremos R_x) es un conjunto finito (posiblemente vacío) de reglas asociadas a x .

Definición 8 Sea $C = (\mu, M)$ una célula sobre un alfabeto, A , y sea R una colección de reglas asociada a C . Una relación de prioridad sobre R es una función, ρ , con dominio en $V(\mu)$ tal que para cada membrana $x \in V(\mu)$, $\rho(x)$ (que notaremos ρ_x) es un orden parcial estricto (posiblemente vacío) sobre R_x .

El orden parcial estricto, ρ_x , sobre R_x , se interpretará como sigue: $(r', r) \in \rho_x$ significa que la regla r' tiene una prioridad estrictamente mayor que la regla r .

Definición 9 *Un sistema P de transición de grado $p \geq 1$ es una tupla, $\Pi = (A, C_0, \mathcal{R}, i_0)$, en donde:*

- A es un conjunto finito no vacío (alfabeto de trabajo).
- $C_0 = (\mu_\Pi, M_0)$ es una célula sobre A , de grado p .
- \mathcal{R} es un par ordenado (R, ρ) donde R es una colección de reglas (de evolución) asociadas a C_0 , y ρ es una relación de prioridad sobre R .
- i_0 es un nodo del árbol enraizado μ_Π , que especifica la membrana de salida del sistema P de transición Π .

4.2 Una semántica para los sistemas celulares de transición

La semántica de un sistema P de transición trata de describir la forma en que funciona como dispositivo computacional. Para definir formalmente la semántica se necesita el concepto de *configuración*, que intenta capturar cada uno de los estados puntuales, instantáneos, que puede alcanzar un sistema a lo largo de su ejecución/evolución. Un sistema P consta de una estructura de membranas con objetos en su interior, con reglas de evolución propias para cada membrana y con especificaciones de entrada–salida. Un sistema celular tiene como únicos elementos variables, la propia estructura de membranas (debido a la operación de disolución) y los objetos contenidos en cada una de las membranas. Por tanto, los elementos básicos para describir una configuración de un sistema P en un instante de su evolución serán las estructuras de membranas y los objetos. Intuitivamente, las configuraciones contienen una descripción completa del estado actual de la ejecución; por ello, van a jugar un papel fundamental en el análisis y discusión de las computaciones de los sistemas celulares.

Definición 10 *Sea $\Pi = (A, C_0, \mathcal{R}, i_0)$ un sistema P de transición, con $C_0 = (\mu_\Pi, M_0)$. Una configuración de Π es una célula $C = (\mu, M)$ sobre A , donde $V(\mu) \subseteq V(\mu_\Pi)$, y de tal manera que los árboles μ y μ_Π tienen la misma raíz.*

De la definición anterior resulta que C_0 es, en particular, una configuración de Π , denominada *configuración inicial* del sistema. Para cada $i \in V(\mu)$, diremos que $\mathcal{M}_i = M_0(i)$ es el *multiconjunto inicial* asociado a la membrana i .

A continuación estudiamos cómo decidir si una regla puede ser aplicada o no en un instante determinado. Informalmente, una regla de evolución, $r = (d_r, v_r, \delta_r, x_r)$ de una membrana $x \in V(\mu)$ se puede expresar como sigue:

$$a_1 \dots a_m \rightarrow (b_1, in_{j_1}) \dots (b_n, in_{j_n})(c_1, out) \dots (c_k, out)(h_1, here) \dots (h_l, here)s$$

con $j_1, \dots, j_n \in V(\mu)$; $a_i, b_i, c_i, h_i \in A$; y $s = \delta$ ó $s = \lambda$.

En esta formalización $d_r = a_1 \dots a_m$; para cada $y \in V(\mu)$, $v_r(y)$ es el multiconjunto formado por los elementos que aparecen en la regla bajo la

forma (b, in_y) ; $v_r(out) = c_1 \dots c_k$, y $v_r(herc) = h_1 \dots h_l$. La componente δ_r (δ o $\neg\delta$) representa si $s = \delta$ o $s = \lambda$. Por último, la componente x_r representa la membrana a la que está asociada la regla.

Entonces

- Esta regla puede ser aplicada sólo cuando hay suficientes copias de los objetos a_1, \dots, a_m en la membrana x y, además, ninguna regla de R_x con prioridad superior es aplicable.
- El resultado de usar esta regla viene determinado por su lado derecho:
 - Los objetos h_1, \dots, h_l van a parar a la misma región x .
 - Los objetos c_1, \dots, c_k van a parar a la región inmediatamente exterior a x (es decir, al entorno, si que x sea la piel, o bien al padre de x , en caso contrario).
 - Cada objeto b_i pasará a formar parte del multiconjunto asociado a la membrana j_i , en el caso en que j_i fuese un hijo de la membrana x . Caso contrario, la regla no sería aplicable.
 - Si el símbolo δ aparece en la regla (es decir, si $s = \delta$), entonces la membrana x se elimina (se *disuelve*) y el conjunto de reglas R_x (junto con su relación de prioridad asociada) no volverá a ser usado. El multiconjunto asociado a x se añade a la región inmediatamente exterior a ella (su padre). Una regla que contenga el símbolo δ no puede ser aplicada en la membrana piel.

Notaremos por $\mathbf{R} = \bigcup_{x \in V(\mu_0)} R_x$, al conjunto de todas las reglas de evolución asociadas a la configuración inicial (es decir, al sistema P de transición).

Definición 11 Sea $C = (\mu, M)$ una configuración de un sistema P de transición $\Pi = (A, C_0, \mathcal{R}, i_0)$, con $C_0 = (\mu_\Pi, M_0)$. Diremos que la regla (de evolución) $r = (d_r, v_r, \delta_r, x_r) \in \mathbf{R}$ es aplicable a C si:

- La membrana x_r a la que está asociada existe en la célula C : $x_r \in V(\mu)$.
- No intenta disolver el nodo raíz: si x_r es el nodo raíz de μ , $\delta_r = \neg\delta$.
- La membrana x_r tiene los objetos necesarios: $d_r \leq M(x_r)$.
- Los nodos a los que la regla intenta enviar objetos son hijos de la membrana a la que pertenece la regla: $\forall j \in V(\mu)(v_r(j) \neq \emptyset \rightarrow j \in Ch(x_r))$.
- No hay otras reglas aplicables a C con prioridad mayor: no existe una regla r' tal que $\rho_{x_r}(r', r) \wedge r'$ aplicable a C .

A continuación, se define el índice de aplicabilidad de una regla r en una configuración C , que notaremos $N_{Ap}(r, C)$, y que nos va a indicar el máximo número de veces que la regla r puede ser aplicada a C . Es decir,

$$N_{Ap}(r, C) = \begin{cases} 0 & , \text{ si } r \text{ no es aplicable a } C \\ \text{máx}\{n : n \otimes d_r \leq M(x_r)\} & , \text{ en otro caso.} \end{cases}$$

Seguidamente, se introduce un concepto que representa un multiconjunto sobre \mathbf{R} de reglas que indica, con la multiplicidad asociada a cada regla, cuántas veces ha de aplicarse dicha regla a fin de realizar un paso de transición en el sistema.

Definición 12 Sea $C = (\mu, M)$ una configuración de un sistema P de transición $\Pi = (A, C_0, \mathcal{R}, i_0)$. Diremos que un multiconjunto m sobre \mathbf{R} es de aplicabilidad sobre C , y lo notaremos por $m \in \mathbf{M}_{\mathbf{Ap}}(C)$, si :

- $\forall r \in \mathbf{R} (m(r) \leq N_{\mathbf{Ap}}(r, C))$.
- $\forall x \in V(\mu_0) (\sum_{r \in R_x} m(r) \otimes d_r \leq M(x))$.
- Es maximal, en el sentido siguiente: no existe un multiconjunto m' sobre \mathbf{R} tal que $m < m' \wedge m' \in \mathbf{M}_{\mathbf{Ap}}(C)$.

Para determinar la configuración siguiente de una configuración, C , por un multiconjunto, m , de aplicabilidad sobre C , se procede en dos etapas. En la primera, las reglas se ejecutan sin atender a la disolución. En la segunda, se disuelve cada membrana que se tenga que disolver, se distribuye su contenido a su primer antecesor no disuelto, y se actualiza el árbol enraizado que resulta.

La primera etapa (aplicación de las reglas sin atender a la disolución) produce la configuración $C'' = (\mu, M'')$, en donde:

$$M''(x) = M(x) - \sum_{r \in R_x} m(r) \otimes d_r + \sum_{r \in R_x} m(r) \otimes v_r(\text{here}) + \sum_{r \in R_{f(x)}} m(r) \otimes v_r(x) + \sum_{y \in Ch(x)} \sum_{r \in R_x} m(r) \otimes v_r(\text{out})$$

Para llevar a cabo la segunda etapa (disolución y distribución de los contenidos) necesitamos caracterizar, previamente, los nodos que deben disolverse por la ejecución de un multiconjunto de aplicabilidad.

Definición 13 Sea $C = (\mu, M)$ una configuración de un sistema P de transición Π , y $m \in \mathbf{M}_{\mathbf{Ap}}(C)$ un multiconjunto de aplicabilidad sobre C . Se define el conjunto: $\Delta(m, C) = \{x_r : r \in \mathbf{R} \wedge m(r) > 0 \wedge \delta_r = \delta\}$.

Es decir, $\Delta(m, C)$ representa el conjunto de nodos de la estructura de membranas de C que deben ser disueltos tras la aplicación del multiconjunto m .

Hay que tener presente que en un paso del sistema celular de transición, debido al paralelismo, se puede disolver más de una membrana de la estructura. Por tanto, será conveniente determinar para cada membrana, x , que permanezca en la próxima configuración, cuál es el conjunto de membranas que se disuelven por la aplicación de un cierto multiconjunto de aplicabilidad y añaden su contenido a x (tales membranas se llamarán *donantes* de x).

Definición 14 Sea $C = (\mu, M)$ una configuración de un sistema P de transición Π . Sea $m \in \mathbf{M}_{\mathbf{Ap}}(C)$ un multiconjunto de aplicabilidad sobre C .

Para cada nodo $x \in V(\mu)$, se define el conjunto $Don(x, m, C)$ de los donantes de x para C por la aplicación de m , como sigue:

$$\begin{cases} \emptyset & , \text{ si } x \in \Delta(m, C) \\ \{y \in V(\mu) : y \in \Delta(m, C) \wedge x \rightsquigarrow_{\mu} y \wedge \\ \wedge \forall z \in V(\mu)(x \rightsquigarrow_{\mu} z \rightsquigarrow_{\mu} y \rightarrow z \in \Delta(m, C))\} & , \text{ si } x \notin \Delta(m, C) \end{cases}$$

Es decir, si un nodo x no se disuelve, entonces el nodo y es donante de x si y se disuelve por la aplicación de m , así como todo nodo que sea descendiente propio de x y antecesor propio de y .

A continuación se define la ejecución de un multiconjunto de aplicabilidad sobre una configuración dada.

Definición 15 Sea $C = (\mu, M)$ una configuración de un sistema P de transición Π . Sea $m \in \mathbf{M}_{\mathbf{Ap}}(C)$ un multiconjunto de aplicabilidad sobre C . Se define la ejecución de m sobre C , y se notará por $m(C)$, como la configuración $C' = (\mu', M')$ de Π , en donde:

- $\mu' = (V(\mu'), E(\mu'))$ es el árbol enraizado obtenido de μ como sigue:
 - $V(\mu') = V(\mu) - \Delta(m, C)$
 - Si $x, y \in V(\mu')$, entonces $(x, y) \in E^*(\mu')$ si y sólo si existen $x_0, \dots, x_n \in V(\mu)$ tales que $(x_1, \dots, x_{n-1}) \in \Delta(m, C) \wedge x_0 = x \wedge x_n = y \wedge \forall i (0 \leq i < n \rightarrow (x_i, x_{i+1}) \in E^*(\mu))$
- $M'(x) = \begin{cases} M''(x) \cup \bigcup_{y \in Don(x, m, C)} M''(y) & , \text{ si } x \notin \Delta(m, C) \\ \emptyset & , \text{ si } x \in \Delta(m, C) \end{cases}$

Es decir, μ' es el árbol enraizado obtenido de μ eliminando los nodos que se disuelven por la aplicación de m , y restaurando las conexiones en la forma correcta. Si un nodo y es disuelto, el nodo se elimina, junto con todas las aristas adyacentes a él, y se añaden aristas entre su primer antecesor no disuelto y los primeros descendientes de y no disueltos. Si un nodo no se disuelve, debemos añadir a su contenido el de todos sus donantes por la aplicación de m .

A continuación, formalizamos la ejecución de un paso en un sistema celular; es decir, precisamos cómo se produce la transición de una configuración del sistema P a otra configuración tras ejecutar un paso. Se trata, pues, de precisar cómo calcula, cómo evoluciona un sistema P a lo largo del tiempo.

Definición 16 Se dice que una configuración C_2 de un sistema P de transición, Π , deriva de una configuración C_1 por una transición en un paso de Π (o tras la ejecución de un paso), y se nota por $C_1 \Rightarrow_{\Pi} C_2$, si existe un multiconjunto, m , de aplicabilidad sobre C_1 tal que $m \neq \emptyset$ y $m(C_1) = C_2$.

Intuitivamente, $C_1 \Rightarrow_{\Pi} C_2$ significa que en un paso de computación del sistema Π , de la configuración C_1 resulta la configuración C_2 .

Una vez definida la relación de transición en un paso entre configuraciones de un sistema celular, se define la relación de transición como su clausura transitiva.

Definición 17 Sean C y C' dos configuraciones de un sistema Π y $k \in \mathbb{N}$. Diremos que C' se obtiene de C en k pasos de transición (es decir, al ejecutarse k pasos), si existen configuraciones C_1, \dots, C_{k+1} tales que

$$C_1 = C \wedge C_{k+1} = C' \wedge \forall i (1 \leq i \leq k \rightarrow C_i \Rightarrow_{\Pi} C_{i+1})$$

Se dice que C' se obtiene de C en una computación de Π , y se nota por $C \Rightarrow_{\Pi}^* C'$, si existe $k \in \mathbb{N}$ tal que C' se obtiene de C en k pasos de transición.

A partir de la configuración inicial de un sistema celular se puede construir el árbol de computación asociado al mismo.

Definición 18 El árbol de computación de un sistema P de transición, Π , es el árbol enraizado etiquetado maximal definido como sigue: la raíz es la configuración inicial de Π ; los hijos de un nodo son las configuraciones que se obtienen a partir de él por un paso de transición; los nodos y las aristas están etiquetadas por configuraciones y multiconjuntos de aplicabilidad, respectivamente, de tal forma que dos nodos etiquetados, C y C' , son adyacentes por medio de una arista etiquetada por m , si y sólo si $m \in \mathbf{M}_{\mathbf{Ap}}(C) - \{\emptyset\} \wedge C' = m(C)$. Las ramas maximales se denominan computaciones de Π . Diremos que una computación de Π es de parada si es una rama finita. Las configuraciones C que verifican $\mathbf{M}_{\mathbf{Ap}}(C) = \{\emptyset\}$ se denominan de parada.

Es decir, una computación, \mathcal{C} , de un sistema P de transición, $\Pi = (A, C_0, \mathcal{R}, i_0)$, es una sucesión (finita o no) de configuraciones del sistema $C_0 \Rightarrow_{\Pi} C_1 \Rightarrow_{\Pi} C_2 \Rightarrow_{\Pi} \dots \Rightarrow_{\Pi} C_r$, con $r \in \mathbb{N} \cup \{\infty\}$. En tal situación, diremos que r es la longitud de \mathcal{C} y notaremos $|\mathcal{C}| = r$. Además, si $r \in \mathbb{N}$ entonces diremos que C_r es la configuración final de la computación \mathcal{C} .

Definición 19 Sea Π un sistema P de transición. Diremos que una computación de Π , es exitosa si i_0 es una hoja del árbol enraizado de la configuración final asociada (en este caso, diremos que dicha configuración es exitosa).

Es decir, una computación es exitosa si es de parada y, además, la membrana de salida es elemental en la configuración final de dicha computación. Es claro que se puede construir sistemas celulares en los que exista una configuración que posea distintos multiconjuntos de aplicabilidad. En ese momento de la computación, el sistema tendrá varias posibilidades para continuar, apareciendo el no determinismo al que se ha hecho mención con anterioridad.

4.3 Sistemas celulares con entrada

Definición 20 Un sistema P de transición con entrada es una tupla (Π, Σ, i_{Π}) , tal que:

- Π es un sistema P de transición, con alfabeto de trabajo Γ , con p membranas etiquetadas por $1, \dots, p$, y multiconjuntos iniciales $\mathcal{M}_1, \dots, \mathcal{M}_p$ asociados a cada una de ellas.

- Σ es un alfabeto (denominado de entrada) estrictamente contenido en Γ .
- $\mathcal{M}_1, \dots, \mathcal{M}_p$ son multiconjuntos sobre $\Gamma - \Sigma$.
- i_{Π} es la etiqueta de una membrana distinguida (membrana de entrada).

Si m es un multiconjunto sobre Σ , entonces la configuración inicial de (Π, Σ, i_{Π}) con entrada m es $(\mu, \mathcal{M}_1, \dots, \mathcal{M}_{i_{\Pi}} + m, \dots, \mathcal{M}_p)$.

Las computaciones en un sistema celular con entrada asociada a un cierto multiconjunto de entrada, se definen de manera análoga a como se hizo en el caso general, con la única diferencia de que la configuración inicial está asociada a ese multiconjunto de entrada.

4.4 Sistemas celulares con salida externa

Una primera clasificación de los sistemas celulares se puede realizar en función de dónde se recoge la salida de las computaciones. Los *sistemas celulares de transición con salida externa* se caracterizan porque la salida de las computaciones se recoge en el entorno de la estructura de membranas.

La membrana piel de una estructura de membranas (nos referiremos a ella usando la meta-etiqueta *skin*) aísla a la estructura de su entorno (al que nos referiremos con la meta-etiqueta *env*). A continuación vamos a asociar a cada estructura de membrana una estructura *con entorno*.

Definición 21 Sea $\mu = (V(\mu), E(\mu))$ una estructura de membranas. La estructura de membranas con entorno asociada a μ es el árbol enraizado $Ext(\mu)$ donde: $V(Ext(\mu)) = V(\mu) \cup \{env\}$; $E(Ext(\mu)) = E(\mu) \cup \{\{env, skin\}\}$; y la raíz del árbol es el nodo *env*, que se denomina entorno de la estructura μ .

Obsérvese que lo único que se hace es añadir un nuevo nodo que representa al entorno y que, por tanto, sólo es adyacente a la piel, mientras que la estructura de membranas original permanece inalterada.

Definición 22 Un sistema P de transición con salida externa es un par (Π, \mathcal{M}_e) , en donde Π es un sistema P de transición y $\mathcal{M}_e = \emptyset$ es el multiconjunto vacío de objetos (que representa el contenido del entorno).

A continuación vamos a detallar de qué forma evoluciona un sistema P de transición con salida externa atendiendo a los multiconjuntos de objetos contenidos en cada una de las membranas de su estructura de membranas, así como a las reglas de evolución asociadas a ellas.

Definición 23 Sea Π un sistema P de transición con salida externa. Una configuración de Π es un par $C = (Ext(\mu), M)$ tal que:

- $\mu = (V(\mu), E(\mu))$ es una estructura de membranas.
- $Ext(\mu)$ es la estructura de membranas con entorno asociada a μ .

- El conjunto $V(\mu)$ de nodos de μ está incluido en el conjunto $V(\mu_{\Pi})$ de nodos de μ_{Π} , y contiene la raíz de μ_{Π} .
- Las raíces de ambas estructuras de membranas coinciden.
- M es una aplicación de dominio $V(\text{Ext}(\mu))$ y rango contenido en $\mathbf{M}(\Gamma)$.

Notaremos por $C = (\mu, M_{env}, M_{i_1}, \dots, M_{i_q})$ una configuración de Π , donde $V(\mu) = \{i_1, \dots, i_q\}$, $M_{env} = M(env)$ es el multiconjunto asociado al entorno de μ , y $M_{i_j} = M(i_j)$ es el multiconjunto asociado a la membrana i_j de μ , para cada $j = 1, \dots, q$.

A la hora de definir las configuraciones que especifican el estado inicial de un sistema P (es decir, sus configuraciones iniciales) debemos tener en cuenta si dicho sistema posee o no membrana de entrada.

Definición 24 Sea (Π, \mathcal{M}_e) un sistema P de transición con salida externa, cuyos multiconjuntos iniciales son $\mathcal{M}_1, \dots, \mathcal{M}_p$. Si Π no posee membrana de entrada, entonces existe una única configuración inicial del sistema, a saber $C0 = (\mu_{\Pi}, \emptyset, \mathcal{M}_1, \dots, \mathcal{M}_p)$. Si Π posee membrana de entrada, entonces existe una configuración inicial para cada multiconjunto $m \in \mathbf{M}(\Sigma)$, a saber $C0(m) = (\mu_{\Pi}, \emptyset, \mathcal{M}_1, \dots, \mathcal{M}_{im} + m, \dots, \mathcal{M}_p)$

A partir de aquí, se define de manera usual el concepto de paso de transición y computación del sistema celular. Además, una de las características de estos sistemas radica en que el resultado de cualquier computación de parada se *recoge* en el entorno, de acuerdo con una cierta función de salida.

4.5 Sistemas celulares reconocedores

Definición 25 Un sistema P de decisión es un par (Π, F_{Π}) tal que Π es un sistema P y F_{Π} es una función total booleana sobre el conjunto de las computaciones de parada de Π . Si \mathcal{C} es una computación de parada y $F_{\Pi}(\mathcal{C}) = 1$ (respectivamente, $F_{\Pi}(\mathcal{C}) = 0$), entonces diremos que \mathcal{C} es de aceptación (respectivamente, de rechazo).

A continuación se introduce un caso particular de sistemas de decisión: aquellos que pueden ser considerados como dispositivos reconocedores de lenguajes.

Definición 26 Un sistema P reconocedor es un sistema P de transición con entrada y con salida externa tal que:

- El alfabeto de trabajo tiene dos elementos distinguidos: *yes*, *no*.
- Todas las computaciones del sistema son de parada.
- Si \mathcal{C} es una computación del sistema, entonces o bien algún objeto *yes* o bien algún objeto *no* (pero no ambos) son enviados al entorno (y, únicamente, en el último paso de la computación).

Sea Π un sistema P reconocedor. Definimos la función *Output* sobre el conjunto de las computaciones de Π . Si \mathcal{C} es una computación de Π y M_{env} es el multiconjunto asociado al entorno en la configuración de parada, entonces

$$Output(\mathcal{C}) = \begin{cases} yes & , \text{ si } yes \in M_{env} \\ no & , \text{ si } yes \notin M_{env} \end{cases}$$

Diremos que una computación \mathcal{C} es de *aceptación* (respectivamente, de *rechazo*) si el objeto *yes* (respectivamente, *no*) aparece en el entorno asociado a la correspondiente configuración de parada de \mathcal{C} .

Notaremos por \mathcal{LR} a la clase de los sistemas P reconocedores de lenguajes.

5 Una Teoría de la Complejidad en sistemas P

Siempre que se introduce un nuevo modelo de computación es necesario desarrollar una teoría formal de la Complejidad Computacional, a fin de cuantificar la cantidad de recursos necesarios para resolver problemas dentro del modelo, y para poder comparar esa cantidad con la que se necesita en otros modelos de computación. La primera teoría de la complejidad en sistemas celulares con membranas ha sido desarrollada en [22] y [23]

En lo que sigue, supondremos que \mathcal{D} es una clase genérica de sistemas P de decisión, $g : \mathbb{N}^+ \rightarrow \mathbb{N}^+$ una función total computable y $X = (I_X, \theta_X)$ un problema de decisión (en donde \mathbb{N}^+ es el conjunto $\mathbb{N} - \{0\}$, I_X es un lenguaje sobre un alfabeto finito, el conjunto de *instancias* de X , y θ_X es un predicado sobre dicho conjunto).

5.1 Clases de complejidad en sistemas P

Los primeros resultados sobre resolución eficiente de problemas **NP**-completos por medio de sistemas de computación celular con membranas se han obtenido usando variantes de sistemas celulares que carecían de membrana de entrada. Por ello, en las pruebas de dichos resultados para *cada instancia* del problema, se construye un sistema que la decide en tiempo polinomial.

Definición 27 Diremos que un problema de decisión, X , pertenece a la clase $MC_{\mathcal{D}}^f(g)$ si existe una familia de sistemas celulares, $\Pi = (\Pi(w))_{w \in I_X}$, verificando las siguientes propiedades:

1. La familia Π es consistente, respecto de la clase \mathcal{D} ; es decir, todo sistema de la familia pertenece a la clase \mathcal{D} .
2. Para cada $w \in I_X$, el sistema $\Pi(w)$ carece de membrana de entrada.
3. La familia Π es polinomialmente uniforme, por máquinas de Turing; es decir, existe una tal máquina que, dado $w \in I_X$, construye el sistema $\Pi(w)$ en tiempo polinomial en el tamaño de w .

4. La familia Π está acotada, respecto del problema X y la función g ; es decir, para cada $w \in I_X$ se tiene que toda computación del sistema $\Pi(w)$ es de parada y, además, realiza a lo sumo $g(|w|)$ pasos.
5. La familia Π es adecuada, respecto del problema X ; es decir, para cada $w \in I_X$ se tiene que si existe una computación del sistema $\Pi(w)$ que es de aceptación, entonces $\theta_X(w) = 1$.
6. La familia Π es completa, respecto del problema X ; es decir, para cada $w \in I_X$ se tiene que si $\theta_X(w) = 1$, entonces toda computación del sistema $\Pi(w)$ es de aceptación.

Si $X \in MC_{\mathcal{D}}^f(g)$, entonces diremos que el problema X es resoluble por una familia de sistemas celulares sin membrana de entrada, pertenecientes a \mathcal{D} , en tiempo acotado por g .

Si en la definición anterior consideramos como cotas los polinomios, se obtiene la clase de complejidad polinomial para familias de sistemas P sin entrada.

Definición 28 La clase de problemas de decisión resolubles en tiempo polinomial por una familia de sistemas celulares sin membrana de entrada, pertenecientes a la clase \mathcal{D} , es la clase

$$\mathbf{PMC}_{\mathcal{D}}^f = \bigcup_{k>0} MC_{\mathcal{D}}^f(n^k)$$

Esta clase es cerrada bajo reducibilidad en tiempo polinomial [24].

Proposición 1 Sean X e Y problemas de decisión tales que X es reducible a Y en tiempo polinomial. Si $Y \in \mathbf{PMC}_{\mathcal{D}}^f$, entonces $X \in \mathbf{PMC}_{\mathcal{D}}^f$.

Además, se tiene que la clase $\mathbf{PMC}_{\mathcal{D}}^f$ es cerrada bajo complementario.

Por otra parte, si trabajamos con sistemas con membrana de entrada podemos considerar familias tales que cada uno de los sistemas que pertenezcan a ellas decida instancias de *tamaño equivalente*.

Definición 29 Diremos que un problema de decisión X pertenece a la clase $MC_{\mathcal{D}}^{fi}(g)$ si existe una familia de sistemas celulares, $\Pi = (\Pi(n))_{n \in \mathbb{N}^+}$, verificando las siguientes propiedades:

1. La familia Π es consistente, respecto de la clase \mathcal{D} .
2. Para cada $n \in \mathbb{N}^+$, el sistema $\Pi(n)$ tiene membrana de entrada.
3. La familia Π es polinomialmente uniforme, por máquinas de Turing.
4. Existen dos funciones, $\text{cod} : I_X \rightarrow \bigcup_{n \in \mathbb{N}^+} I_{\Pi(n)}$ y $s : I_X \rightarrow \mathbb{N}^+$, computables en tiempo polinomial, tales que:
 - Para todo $w \in I_X$, $\text{cod}(w) \in I_{\Pi(s(w))}$.

- La familia Π está acotada, respecto del problema X , la función cod , la función s y la función g ; es decir, para cada $w \in I_X$ se tiene que toda computación del sistema $\Pi(s(w))$ con entrada $cod(w)$ es de parada y, además, realiza a lo sumo $g(|w|)$ pasos.
- La familia Π es adecuada, respecto del problema X , la función cod y la función s ; es decir, para cada $w \in I_X$ se tiene que si existe una computación del sistema $\Pi(s(w))$ con entrada $cod(w)$ que es de aceptación, entonces $\theta_X(w) = 1$.
- La familia Π es completa, respecto del problema X , la función cod y la función s ; es decir, para cada $w \in I_X$ se tiene que si $\theta_X(w) = 1$, entonces toda computación del sistema $\Pi(s(w))$ con entrada $cod(w)$ es de aceptación.

Si $X \in MC_{\mathcal{D}}^{fi}(g)$, entonces diremos que el problema X es resoluble por una familia de sistemas celulares con membrana de entrada, pertenecientes a la clase \mathcal{D} , en tiempo acotado por g .

Por otra parte, en las condiciones de la definición anterior diremos que el par (cod, s) es una *codificación polinomial* de X en Π .

Definición 30 La clase de problemas de decisión resolubles en tiempo polinomial por una familia de sistemas de computación celular con membrana de entrada, pertenecientes a la clase \mathcal{D} , es la clase

$$\mathbf{PMC}_{\mathcal{D}}^{fi} = \bigcup_{k>0} MC_{\mathcal{D}}^{fi}(n^k)$$

Esta clase es cerrada bajo reducibilidad en tiempo polinomial[24].

Proposición 2 Sean X e Y problemas de decisión tales que X es reducible a Y en tiempo polinomial. Si $Y \in \mathbf{PMC}_{\mathcal{D}}^{fi}$, entonces $X \in \mathbf{PMC}_{\mathcal{D}}^{fi}$.

Además, se tiene que la clase $\mathbf{PMC}_{\mathcal{D}}^{fi}$ es cerrada bajo complementario.

Obsérvese que en las definiciones anteriores suponemos una *confluencia* de los sistemas de la familia; es decir, los sistemas proporcionan la misma respuesta para todas las computaciones que partan de la misma configuración inicial.

A continuación vamos a caracterizar la relación $\mathbf{P} \neq \mathbf{NP}$ mediante la irresolubilidad en tiempo polinomial de problemas \mathbf{NP} -completos a través de sistemas \mathbf{P} de transición.

5.2 Simulación de máquinas de Turing por sistemas celulares

Comenzamos observando que es posible asociar un problema de decisión a cada máquina de Turing, lo que nos permitirá establecer con precisión qué significa que una tal máquina sea simulada por sistemas celulares.

Definición 31 Sea TM una máquina de Turing, como dispositivo de decisión de lenguajes, y con alfabeto de entrada Σ_{TM} . El problema de decisión asociado a TM es el problema $X_{TM} = (I_{TM}, \theta_{TM})$, siendo $I_{TM} = \Sigma_{TM}^*$; y para cada cadena $w \in \Sigma_{TM}^*$, $\theta_{TM}(w) = 1$ si y solo si TM acepta w .

Diremos que la máquina TM es simulada en tiempo polinomial por una familia de sistemas en la clase \mathcal{D} si $X_{TM} \in \mathbf{PMC}_{\mathcal{D}}^{fi}$ o $X_{TM} \in \mathbf{PMC}_{\mathcal{D}}^f$, según si los sistemas poseen o no membrana de entrada.

Consideramos las máquinas de Turing como dispositivos de decisión de lenguajes; es decir, dichas máquinas paran sobre cualquier dato de entrada, siendo el estado de parada igual al de aceptación, en el caso en que la cadena pertenezca al lenguaje decidido, o bien igual al estado de rechazo, en el caso en que la cadena no pertenezca a ese lenguaje.

Proposición 3 *Sea TM una máquina de Turing determinista que trabaja en tiempo polinomial. Entonces $X_{TM} \in \mathbf{PMC}_{\mathcal{LR}}^{fi}$.*

La prueba del resultado consiste en asociar a la máquina de Turing TM una familia de sistemas celulares de tal forma que cada uno de ellos decida las cadenas de entrada de un tamaño determinado. En la figura 3 se proporciona un diagrama de flujo que muestra el funcionamiento de estos sistemas una vez recibida la cadena de entrada, codificada como un multiconjunto de objetos del sistema (para detalles de la prueba, véase [24]).

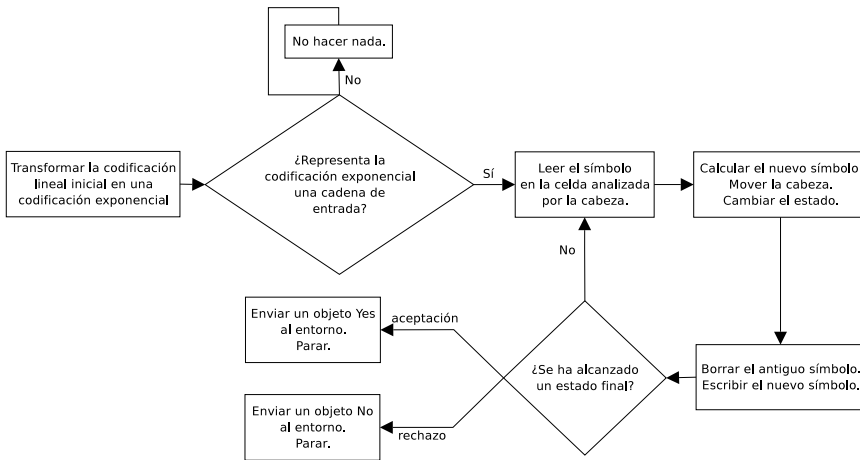


Figura 3: Simulación de máquinas de Turing por sistemas celulares

Seguidamente, se establece una condición *suficiente* para que se verifique la relación $\mathbf{P} \neq \mathbf{NP}$, en términos de irresolubilidad en tiempo polinomial de problemas \mathbf{NP} -completos por sistemas \mathbf{P} de transición.

Proposición 4 *Si existe un problema \mathbf{NP} -completo irresoluble en tiempo polinomial por una familia de sistemas celulares reconocedores, entonces $\mathbf{P} \neq \mathbf{NP}$.*

Demostración. Sea X un problema \mathbf{NP} -completo tal que $X \notin \mathbf{PMC}_{\mathcal{LR}}^{fi}$. Supongamos que $\mathbf{P} = \mathbf{NP}$. Entonces $X \in \mathbf{P}$. Luego existe una máquina de Turing determinista, TM , que resuelve el problema X en tiempo polinomial.

Por la proposición 3, $X_{TM} \in \mathbf{PMC}_{\mathcal{LR}}^{fi}$. Por tanto, existe una familia $\mathbf{\Pi}_{TM} = (\mathbf{\Pi}_{TM}(k))_{k \in \mathbb{N}^+}$ de sistemas reconocedores de lenguajes que simula TM en tiempo polinomial.

Consideremos la función $cod_X : I_X \rightarrow \bigcup_{k \in \mathbb{N}^+} I_{\mathbf{\Pi}_{TM}(k)}$, dada por $cod_X(w) = cod_{TM}(w)$, y la función $s_X : I_X \rightarrow \mathbb{N}^+$, dada por $s_X(w) = |w|$. Entonces se prueba que la familia $\mathbf{\Pi}_{TM}$ es adecuada y completa, respecto de X , la función cod_X y la función s_X .

En consecuencia, se tendría que $X \in \mathbf{PMC}_{\mathcal{LR}}^{fi}$. Lo que es una contradicción. \square

5.3 Simulación de sistemas celulares por máquinas de Turing

A continuación vamos a probar que si un problema de decisión se puede resolver en tiempo polinomial por una familia de sistemas reconocedores, también se puede resolver en tiempo polinomial por una máquina de Turing determinista.

Para el diseño de la máquina de Turing nos hemos inspirado en un trabajo de C. Zandron, C. Ferretti y G. Mauri [26], aunque dicho artículo se enmarca dentro del estudio de los sistemas celulares que usan membranas activas.

Proposición 5 *Si $X \in \mathbf{PMC}_{\mathcal{LR}}^{fi}$, entonces existe una máquina de Turing determinista que resuelve el problema X en tiempo polinomial.*

Demostración. Sea $\mathbf{\Pi} = (\mathbf{\Pi}(n))_{n \in \mathbb{N}^+}$ una familia de sistemas de la clase \mathcal{LR} que es polinomialmente uniforme por máquinas de Turing; y para la que existe una codificación polinomial (cod, s) de X en $\mathbf{\Pi}$ tal que dicha familia está polinomialmente acotada, es adecuada y completa respecto de X , la función cod y la función s .

A continuación asociamos al sistema $\mathbf{\Pi}(n)$ una máquina de Turing determinista, $TM(n)$, con múltiples cintas, tal que, dado un multiconjunto de entrada, $m \in I_{\mathbf{\Pi}(n)}$, la máquina reproduce una determinada computación, fijada por el proceso de funcionamiento de la máquina, de $\mathbf{\Pi}(n)$ sobre m .

El alfabeto de entrada de la máquina $TM(n)$ coincide con el del sistema $\mathbf{\Pi}(n)$. El alfabeto de trabajo contiene, además, un símbolo por cada etiqueta asignada a las membranas de dicho sistema; los símbolos 0 y 1, que permitirán operar con números representados en base 2; tres símbolos que indicarán si una membrana no se ha disuelto, ha de disolverse o se ha disuelto; y tres símbolos que indicarán si una regla está en espera, es aplicable o no es aplicable.

El proceso de simulación, como veremos, se puede estructurar en varias fases realizan distintos bucles en los que se recorren las membranas de la estructura de membranas original del sistema y las reglas asociadas a ellas. Así, el conjunto de estados de la máquina de Turing contendrá estados relacionados con dichas membranas y reglas para cada una de las fases. Estos estados determinarán en qué momento de la simulación nos encontramos en cada paso del funcionamiento de la máquina.

Seguidamente, detallamos las cintas de las que consta esta máquina.

- Una *cinta de entrada*, que guarda una cadena representando el multiconjunto de entrada recibido.
- Por cada membrana del sistema tenemos:
 - Una *cinta de estructura*, que guarda en la segunda celda la etiqueta del padre de la membrana, y en la tercera uno de los tres símbolos que indican si la membrana no se ha disuelto, si la membrana ha de disolverse, o si la membrana se ha disuelto.
 - Para cada objeto del alfabeto de trabajo del sistema:
 - * Una *cinta principal*, que guarda la multiplicidad del objeto, en base 2, en el multiconjunto contenido en la membrana.
 - * Una *cinta auxiliar*, que guarda resultados temporales, también en base 2, de aplicar las reglas asociadas a la membrana.

Tras la simulación de cada paso del sistema P, el contenido de estas dos cintas coincidirá.

 - Una *cinta de reglas*, en la que cada celda a partir de la segunda corresponde a una regla asociada a la membrana (suponemos que el conjunto de dichas reglas está ordenado), y guarda uno de los tres símbolos que indican si dicha regla está en espera, si es aplicable, o si no es aplicable.
- Para cada objeto del alfabeto de salida tenemos:
 - Una *cinta de entorno*, que guarda la multiplicidad del objeto, en base 2, en el multiconjunto asociado al entorno externo.

A continuación describimos sucintamente los pasos realizados por la máquina $\Pi(n)$ para simular un paso de transición del sistema celular. Los algoritmos que implementan cada uno de los pasos pueden encontrarse en [24].

- *Inicialización del sistema.* En la primera fase del proceso de simulación seguido por la máquina de Turing se colocan en las cintas correspondientes los símbolos necesarios para que quede reflejada la configuración inicial de la computación con entrada el multiconjunto m que se va a simular.
- *Determinar las reglas aplicables.* Para simular un paso del sistema celular, lo primero que debe hacer la máquina de Turing es determinar el conjunto de reglas que son aplicables (cada una de ellas de forma independiente) a la configuración considerada en las membranas a las que están asociadas.
- *Aplicación de las reglas.* Una vez determinadas las reglas aplicables, éstas se aplican de forma maximal a las membranas a las que están asociadas. El hecho de que las reglas se consideren en un cierto orden (usando maximalidad local para cada regla, según dicho orden) determina un multiconjunto de reglas aplicable particular, fijando así la computación del sistema que simula la máquina de Turing. No obstante, de la definición de

clases de complejidad que hemos realizado se deducirá que la computación elegida no es relevante para la prueba del resultado que pretendemos demostrar, debido a la *confluencia* del sistema.

- *Actualización de los multiconjuntos.* Una vez aplicadas las reglas a las membranas, las cintas auxiliares de éstas guardan los resultados obtenidos, por lo que dichos resultados deben ser trasladados a las cintas principales correspondientes.
- *Disolución de las membranas.* Para terminar la simulación de un paso de la computación del sistema P es necesario disolver las membranas según indiquen las reglas que se han aplicado en la fase anterior y reestructurar el árbol de membranas, redefiniendo para ello la función padre de dicho árbol.
- *Comprobar si la simulación ha terminado.* Finalmente, tras terminar la simulación de un paso de la computación del sistema $\Pi(n)$, la máquina de Turing debe comprobar si dicha computación ha llegado a una configuración de parada y, en su caso, si la computación es de aceptación o de rechazo.

Finalmente, consideremos la siguiente máquina de Turing determinista, TM_{Π} :

Entrada: $w \in I_X$

- Calcular $s(w)$
- Construir $TM(s(w))$
- Calcular $cod(w)$
- Reproducir el funcionamiento de $TM(s(w))$ con dato de entrada $cod(w)$

Entonces, la máquina TM_{Π} resuelve el problema X en tiempo polinomial. \square

Seguidamente vamos a establecer una condición *necesaria* para que se verifique la relación $P \neq NP$ a través de sistemas reconocedores.

Proposición 6 *Si $P \neq NP$, entonces ningún problema NP -completo puede ser resuelto en tiempo polinomial por una familia de sistemas celulares reconocedores.*

Demostración. Supongamos que existe un problema NP -completo, X , tal que $X \in \mathbf{PMC}_{\mathcal{LR}}^{fi}$. Entonces, por la proposición 5 existe una máquina de Turing que resuelve el problema X en tiempo polinomial. Por tanto, $X \in P$. Luego $P = NP$, lo que lleva a una contradicción. \square

De las proposiciones 4 y 6, se deduce la siguiente caracterización de la relación $P \neq NP$ a través de la irresolubilidad de problemas NP -completos por familias de sistemas P reconocedores.

Teorema 7 *Las proposiciones siguientes son equivalentes:*

1. $P \neq NP$.
2. $\exists X (X \text{ es un problema de decisión } NP\text{-completo} \wedge X \notin \mathbf{PMC}_{\mathcal{LR}}^{fi})$.

3. $\forall X (X \text{ problema de decisión NP-completo} \rightarrow X \notin \mathbf{PMC}_{\mathcal{LR}}^{fi})$.

Finalmente, vamos a establecer una descripción de la clase \mathbf{P} de los problemas *tratables* en modo determinista, a través de la clase de complejidad polinomial relativa a sistemas celulares reconocedores de lenguajes.

Teorema 8 *Se verifica que $\mathbf{P} = \mathbf{PMC}_{\mathcal{LR}}^{fi}$.*

Demostración. Si $X \in \mathbf{P}$, entonces de la proposición 3, se deduce que $X \in \mathbf{PMC}_{\mathcal{LR}}^{fi}$.

Por otra parte, si X es un problema de decisión resoluble en tiempo polinomial por una familia de sistemas \mathbf{P} reconocedores de lenguajes, entonces de la proposición 5, resulta que $X \in \mathbf{P}$. \square

Referencias

- [1] L.M. Adleman. Molecular Computations of Solutions to Combinatorial Problems. *Science*, 226, (1994), 1021–1024.
- [2] A.V. Baranda, J. Castellanos, F. Arroyo, R. Gonzalo. Towards an electronic implementation of membrane computing: A formal description of nondeterministic evolution in transition \mathbf{P} systems. En *Proceedings of the 7th International Meeting on DNA Based Computers* (N. Jonoska, N.C. Seeman, eds.), Tampa, Florida, USA, (2001), 273–282.
- [3] G. Berry, G. Boudol. The chemical abstract machine. *Theoretical Computer Science*, 96, (1992), 217–248.
- [4] A.W. Burks, H.H. Goldstine, J. von Neumann. Preliminary discussion of the logical design of an electronic computing instrument. En *Collected Works of John von Neumann* (A. H. Taub, ed.), vol. 5, 34–79, The Macmillan Company, New York, 1963. Taken from report to U. S. Army Ordnance Department, 1946.
- [5] S. Cook. The \mathbf{P} versus \mathbf{NP} Problem. *Manuscript prepared for the Clay Mathematics Institute for the Millennium Prize Problems* (revised November, 2000).
- [6] T.J. Head. Formal language theory and DNA: an analysis of the generative capacity of specific recombinant behaviours. *Bulletin of Mathematical Biology*, 49, (1987), 737–759.
- [7] J.H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.
- [8] M. Ito, C. Martín-Vide, Gh. Păun. A characterization of Parikh sets of ETOL languages in terms of \mathbf{P} systems. En *Words, Semigroups, and Transductions* (M. Ito, G. Paun, S. Yu, eds.), World Scientific, Singapore, (2001), 239–253.

- [9] S.N. Krishna, R. Rama. On the power of P systems with sequential and parallel rewriting. *International Journal of Computer Mathematics*, 77, 1-2 (2000), 1–14.
- [10] S.N. Krishna, R. Rama. P systems with replicated rewriting. *Journal of Automata, Languages and Combinatorics*, 6, 3 (2001), 345–350.
- [11] W.S. McCulloch, W. Pitts. A logical calculus of the ideas inmanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, (1943), 115–133.
- [12] A. Obtulowicz. Membrane computing and one-way functions. *International Journal of Foundations of Computer Science*, 12, 4 (2001), 551–558.
- [13] A. Păun. On P systems with membrane division. En *Unconventional Models of Computation* (I. Antoniou, C.S. Calude, M.J. Dinnen, eds.), Springer-Verlag, London (2000), 187–201.
- [14] A. Păun, Gh. Păun. The power of communication: P systems with Symporters-Antiporters. *New Generation Computing*, 20, 3 (2002).
- [15] Gh. Păun. Computing with membranes. *Journal of Computer and System Sciences*, 61, 1 (2000), 108–143, and *Turku Center for Computer Science-TUCS Report No 208* (1998) (www.tucs.fi).
- [16] Gh. Păun. Further research topics about P systems. *Pre-Proceedings of Workshop on Membrane Computing*, Curtea de Arges, Rumania, Agosto 2001, Technical Report 17/01 of Research Group on Mathematical Mathematical Linguistics, Rovira i Virgili University, Tarragona, España, (2001), 243–250.
- [17] Gh. Păun. Computing with membranes. A variant. *International Journal of Foundations of Computer Sciences*, 11, 1 (2000), 167–182.
- [18] Gh. Păun. *Membrane Computing. An introduction*, Springer-Verlag, Berlin, 2002.
- [19] Gh. Păun, G. Rozenberg. A guide to membrane computing. *Theoretical Computer Science*, 287 (2002), 73–100.
- [20] M.J. Pérez Jiménez, F. Sancho Caparrini. A formalization of transition P systems. *Fundamenta Informaticae*, 49 (2002), 261–272.
- [21] M.J. Pérez Jiménez, F. Sancho Caparrini. Verifying a P system generating squares. *Romanian Journal of Information Science and Technology*, 5, 2–3 (2002), 181–191.
- [22] M.J. Pérez Jiménez, A. Romero Jiménez, F. Sancho Caparrini. *Teoría de la Complejidad en modelos de computación celular con membranas*. Ed. Kronos, 2003.

- [23] M.J. Pérez Jiménez, A. Romero Jiménez, F. Sancho Caparrini. Complexity classes in models of cellular computing with membranes. *Natural Computing*, 2, 3 (2003), 265–285.
- [24] A. Romero Jiménez. *Complejidad y universalidad en modelos de computación celular*. Tesis doctoral. Universidad de Sevilla, 2003.
- [25] C. Martín-Vide, V. Mitrană, Gh. Păun. On the power of P systems with valuations. *Computación y sistemas*, 5, 2, 120-127.
- [26] C. Zandron, G. Mauri, C. Ferreti. Solving NP-complete problems using P systems with active membranes. En *Unconventional Models of Computation* (I. Antoniou, C.S. Calude, M.J. Dinnen, eds.), Springer-Verlag, London, (2000), 289–301.

Modelos de Localización Continua

J. PUERTO¹ Y A. M. RODRÍGUEZ-CHÍA²

¹Departamento de Estadística e Investigación Operativa.
Universidad de Sevilla.

²Departamento de Estadística e Investigación Operativa.
Universidad de Cádiz.

puerto@us.es antonio.rodriguezchia@uca.es

Resumen

En este trabajo se revisan tres modelos de problemas de localización continua: (1) un problema general de localización con respecto a regiones de demanda; (2) el problema de la mediana ordenada y (3) un problema de localización multiobjetivo. Con ellos se pretende dar una amplia muestra de los problemas que aparecen en el ámbito de la Teoría de Localización continua, así como estudiar propiedades que permitan caracterizar los conjuntos de soluciones. El trabajo incluye una larga lista de referencias que facilitarán al lector adentrarse más profundamente en éstos y otros modelos de la Teoría de Localización.

Palabras clave: *Continuous location theory, ordered median function.*

Clasificación por materias AMS: *90-B85, 52B55, 90C25*

1 Introducción

Un problema de localización consiste en determinar la posición de uno o más servidores que optimicen alguna *medida de efectividad* con respecto a un conjunto de ubicaciones conocidas “a priori”. Estas ubicaciones son usualmente denominadas centros de demanda.

La Teoría de Localización, como cualquier otra disciplina de la Investigación Operativa, desarrolla modelos matemáticos que reflejan con la máxima fidelidad posible la situación real que se está considerando, con soluciones adecuadas al problema real que se modela.

Fecha de recepción: 19 de febrero de 2004

Aunque a veces resulta difícil realizar una clasificación de los problemas considerados dentro de un campo de investigación, los problemas de localización pueden agruparse en tres familias: (1) de localización continua; (2) de localización discreta y (3) de localización en redes.

La localización continua considera problemas en los cuales las ubicaciones se consideran inmersas en un espacio *continuo*, típicamente en un espacio Euclídeo. La localización discreta impone que el conjunto de lugares candidatos para ubicar el servicio o servicios es finito. Finalmente, cuando los puntos de demanda están sobre un grafo y los servicios deben posicionarse sobre los nodos o aristas del mismo se denominan problemas de localización en redes.

Por motivos de extensión, este trabajo se restringe a considerar algunos modelos de localización continua. Excelentes referencias que cubren todas las vertientes de la Teoría de Localización son [9, 10, 16, 24, 32].

El origen de la Teoría de Localización se puede atribuir a la Grecia antigua, donde se estudiaron por primera vez las propiedades de puntos singulares dependiendo de la configuración de unos puntos dados. Se cree que, cuando la armada Sarracena en el 638 destruyó la Biblioteca de Alejandría, ésta contenía al menos tres soluciones diferentes a lo que hoy se denomina *problema de Fermat* (1601–1665). Éste, al final de su célebre trabajo sobre máximos y mínimos, enuncia el siguiente problema, al que no encuentra solución:

“Dados tres puntos A, B y C en el plano, encontrar un cuarto punto P tal que la suma de las distancias a los otros tres sea mínima”.

Este problema fue “resuelto” geoméricamente por Torricelli (1608–1647). Su procedimiento consiste en trazar un triángulo equilátero a partir de cada lado del triángulo que se obtiene al unir los puntos dados y trazar después las circunferencias que circunscriben a cada uno de ellos. El punto donde intersecan estas tres circunferencias, llamado punto de Torricelli, es la solución.

En 1834, Heinen probó que la solución propuesta por Torricelli no era correcta de forma general. En el caso de que uno de los ángulos que forman los tres puntos iniciales sea igual o superior a 120° , el vértice correspondiente a este ángulo es la solución del problema.

En su *Doctrine and Application of Fluxions* (1750), Simpson generalizó este problema para así obtener el punto que minimiza la suma ponderada de las distancias a tres puntos dados.

En 1857, Sylvester consideró un problema equivalente de localización bajo el criterio *minimax*, donde se buscaba el menor círculo que contenía un conjunto de puntos. Una solución geométrica, atribuida a Pierce, fue dada en 1860.

A principios del siglo XX, A. Weber [46] propuso como criterio para la localización de las industrias la minimización de los costes de transporte o, equivalentemente, las distancias de viaje. Es por ello que el problema consistente en minimizar la suma de distancias se conoce hoy día como problema de Fermat-Weber.

Debido a que estos autores no disponían de ordenador, sus métodos de resolución no podían ser iterativos sino geométricos. Uno de los primeros

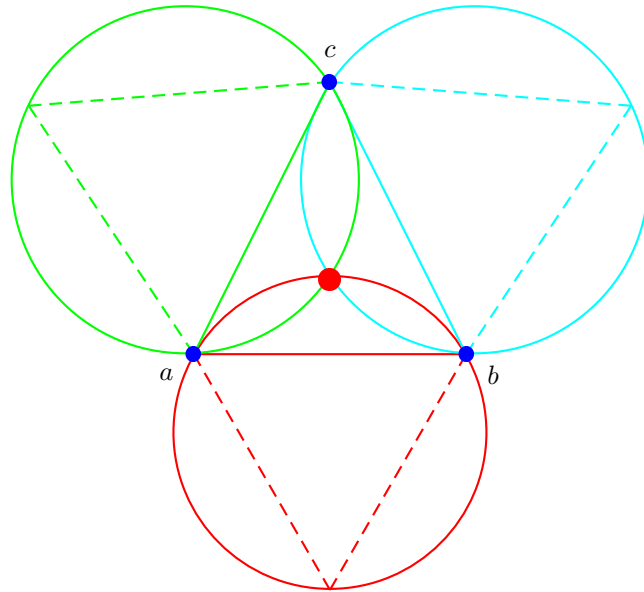


Figura 1: Punto de Torricelli

métodos analíticos para resolver el problema de Fermat-Weber se debe a Weiszfeld y se encuentra en su artículo “*Sur le point le quel la somme des distances de n points donnés est minimum*” en *Tohoku Mathematics Journal*. Se trata de un procedimiento iterativo para localizar un nuevo servicio que minimiza la suma ponderada de las distancias Euclídeas a los puntos de demanda. Este artículo escrito en francés, enviado desde Praga y publicado en una revista japonesa en 1937, fue virtualmente desconocido hasta finales de los sesenta. Por esta razón, el método fue redescubierto por otros investigadores a finales de los cincuenta y principios de los sesenta.

A finales de los años sesenta la Teoría de Localización se diversificó y comenzó a interesar a arquitectos, economistas, ingenieros, geógrafos, especialistas en ciencias de la empresa, investigadores operativos, analistas de sistemas, planificadores urbanos, etcétera, convirtiéndose así en una cuestión multidisciplinar. Aunque el interés por este tema ha venido creciendo desde los años sesenta, ha sido en las dos últimas décadas cuando el número de trabajos publicados en revistas científicas ha aumentado considerablemente. Así desde 1991 pueden encontrarse en el *Mathematics Subject Classification* los epígrafes 90B85 Localización continua y 90B80 Localización discreta entre los 19 títulos de la sección 90Bxx de *Operations Research and Management Science*.

Para obtener una mejor comprensión de la estructura de los problemas de localización, a continuación se describen brevemente los elementos comunes a todos ellos.

El espacio solución:

Es el espacio soporte donde se define el problema. La elección de un espacio solución apropiado es crucial, porque determina aspectos tan importantes como la precisión o la resolubilidad del modelo. Entre los espacios más usuales se encuentran:

- *Espacios discretos*: Cuando existe un número finito de candidatos a solución.
- *Redes*: Los candidatos a solución se encuentran sobre un grafo, usualmente representando una red de comunicación. Los nodos correspondientes representan elementos importantes, como ciudades o cruces. Los arcos representan conexiones entre nodos, como carreteras, calles, conexiones de luz, etc.
- *Espacio Euclídeo \mathbb{R}^n* : Se emplea cuando el problema presenta aspectos regionales que no pueden ser discretizados. Además, puede ser utilizado para aproximar redes cuando el número de nodos y arcos es muy grande. Los casos $n = 2$ y $n = 3$ tienen un claro significado físico. Los casos en que $n \geq 4$ han sido usados para modelar y resolver problemas de estimación estadística.
- *Esfera*: Es útil para aquellas situaciones reales que hacen aparecer distancias de grandes magnitudes.
- *Superposición de un grafo sobre un espacio continuo*: Éste es un espacio solución donde una red, que representa conexiones de alta velocidad, se superpone sobre un espacio Euclídeo o una esfera. Está motivado por el auge de las Nuevas Tecnologías en la sociedad actual y ha recibido gran atención recientemente.

Las ubicaciones existentes:

En términos de Teoría de Localización, las *ubicaciones existentes* representan los usuarios que requieren ser servidos. Por ello, se les denomina puntos de demanda. Usualmente se modelan mediante un conjunto A y una *función de intensidad* asociada que es una medida de probabilidad sobre A .

Existen dos formas diferentes de representar la demanda: por un conjunto finito de puntos o por regiones del espacio solución.

En el primer caso, se considera un conjunto de puntos $A = \{a_1, \dots, a_M\}$ y un conjunto de *pesos* $\{w_1, \dots, w_M\}$ que representan la importancia (o intensidad) de la demanda generada en cada punto.

En el modelo regional, la demanda se representa mediante una región \mathcal{R} (no necesariamente conexas) incluida en el espacio solución y una medida de probabilidad con soporte en \mathcal{R} . Ésta representa la importancia de cada subconjunto medible de \mathcal{R} .

La nueva ubicación:

La localización de la nueva ubicación es la incógnita o variable de decisión del problema general de localización. Esta variable está caracterizada por:

- a) El número y la cualidad de los servicios que se ubican. Si es más de uno, será necesario especificar las características de cada uno de ellos. Cuando son idénticos, como por ejemplo buzones de correo, el modelo que se considera es un modelo multiservicio. Por otro lado, como en el caso de servicios de salud (centros de emergencia, centros de salud, hospitales, etc.), se pueden considerar problemas de localización jerarquizados.
- b) La naturaleza del servicio. Se puede hacer una distinción entre servicios *atractivos* y *repulsivos*.

Un servicio es atractivo cuando todos los puntos de demanda valoran positivamente la proximidad al nuevo servidor. Se pueden considerar centros atractivos aquéllos que produzcan beneficios de tipo económico, cultural, etc. a los puntos de demanda, como por ejemplo una biblioteca o un hospital.

Sin embargo, desde la segunda mitad del siglo XX han ido apareciendo diferentes centros o lugares de almacenamiento de distintos tipos que pueden causar daños o perjuicios a la salud, que deterioran el medio ambiente o que constituyen una amenaza a nuestro bienestar o estilo de vida. En definitiva, centros de actividad que se prefiere tener lo más lejos posible. Así ocurre con la ubicación de centrales nucleares, cementerios de residuos radiactivos, fábricas que producen contaminación ambiental por gases, ruidos, radiaciones, plantas de tratamiento de desechos de los centros urbanos (basureros), etc. A este tipo de servicios se les denomina repulsivos o semi-repulsivos.

La función objetivo:

Debido a que el estudio de la Teoría de Localización que se lleva a cabo en este trabajo es de tipo cuantitativo, los problemas que se consideran tienen la siguiente formulación común:

$$\text{opt}_{x=(x_1, \dots, x_N) \in S} F(d(a, (x_1, \dots, x_N))_{a \in A})$$

donde

F es una función globalizadora (no necesariamente escalar),

“opt” significa optimizar, es decir, minimizar o maximizar,

S es el espacio solución,

x es la nueva ubicación (variable de decisión),

A es el conjunto de ubicaciones existentes,

a es una ubicación existente,

$d(\cdot, \cdot)$ es una medida de la distancia de una ubicación existente a una nueva ubicación.

La elección de la función objetivo a utilizar es a veces difícil. Evidentemente, la solución final depende fuertemente de esta elección (ver [4]), por lo que resulta esencial cuidar esta parte del proceso de modelado.

La propuesta más aceptada para su selección consiste en establecer unos principios de racionalidad (axiomas) que debe verificar la solución y entonces intentar obtener funciones objetivo a través de un proceso basado en esa axiomática.

Diferentes funciones objetivo han dado lugar a los problemas clásicos considerados en Teoría de Localización:

1. *Problema de Fermat-Weber o modelo "minisum"*. El problema consiste, como hemos indicado anteriormente, en la minimización de la distancia esperada a una nueva ubicación x :

$$\inf_{x \in S} \int_A d(x, a) d\nu(a),$$

donde $d\nu(a)$ es una medida de probabilidad sobre A . Cuando A es finito, $A = \{a_1, \dots, a_M\}$, el problema es

$$\inf_{x \in S} \sum_{i=1}^M \mu(a_i) d(x, a_i),$$

donde los $\mu(a_i) > 0$ y $\sum_{i=1}^M \mu(a_i) = 1$.

El modelo "minisum" tiene sentido cuando intentamos hacer valer un criterio económico, en el sentido de minimizar el coste promedio que soporta la localización de un nuevo servicio.

2. *Problema del centro o modelo "minimax"*. Es similar al problema anterior, pero ahora, minimizando la máxima distancia ponderada por μ , es decir,

$$\inf_{x \in X} \left(\sup_{a \in A} \mu(a) d(x, a) \right).$$

El modelo minimax corresponde a aplicar un criterio de calidad del servicio proporcionado.

3. *Problema "cent-dian"*, [4, 17]. Dado $\lambda \in (0, 1)$, la función objetivo es ahora la correspondiente combinación convexa de las funciones objetivo suma y máximo. Es decir, el problema es

$$\inf_{x \in X} \left\{ \lambda \int_A d(x, a) d\mu(a) + (1 - \lambda) \sup_{a \in A} \mu(a) d(x, a) \right\}.$$

El modelo “cent-dian” corresponde a un compromiso entre los modelos “minimax” y “minisum”. Está justificado por el hecho de que estos criterios son contrapuestos en la mayoría de los casos.

4. *Problema de la mediana ordenada.* Dado un número finito de ubicaciones existentes $A = \{a_1, \dots, a_M\}$ y dados unos pesos $\lambda_1, \dots, \lambda_M$, ordenados en orden no decreciente, el problema consiste en minimizar la siguiente suma de distancias ordenadas:

$$\inf_{x \in X} \sum_{i=1}^M \lambda_i d_{(i)}(x).$$

Aquí, $d_{(i)}(x) = d(x, a_{\sigma_i})$ es el i -ésimo elemento en la lista ordenada de distancias, es decir,

$$d(x, a_{\sigma_1}) \leq \dots \leq d(x, a_{\sigma_M}),$$

donde σ es una permutación de $\{1, \dots, M\}$. Nótese que este problema incluye como casos particulares todos los problemas mencionados anteriormente. Hablaremos de este problema en la sección 4.

5. *Problema multiobjetivo.* Los problemas anteriores establecen “a priori” el criterio bajo el cual se localizarán los servicios. Sin embargo, existen situaciones reales donde sería razonable utilizar simultáneamente varios criterios. Esto implica encontrar soluciones que sean óptimas frente a varios criterios al mismo tiempo. Este tipo de problema se denomina problema de localización multiobjetivo. La solución ideal rara vez existe, ya que los objetivos suelen ser contrapuestos, así pues, el concepto de solución ideal es poco operativo. Esto impone la necesidad de introducir soluciones de compromiso y de aceptar como solución el concepto de *elemento no dominado*. Con las soluciones de compromiso se asegura que no existen soluciones mejores (en todos los criterios) que las que se han encontrado. Para una discusión detallada sobre la naturaleza multiobjetivo del problema de localización, véase [27, 36]. Volveremos sobre los problemas multiobjetivo en la sección 5.

Las distancias:

La ubicación de un nuevo centro va a depender de la posición relativa de éste, a través de la noción de distancia, con respecto a los puntos de demanda.

La distancia (o métrica) en un espacio métrico X , es una función que asigna un valor real no negativo $d(x, y)$ a cada par (x, y) . Recuérdese que debe verificar las propiedades siguientes:

1. $d(x, y) = 0$ si y sólo si $x = y$.
2. $d(x, y) = d(y, x)$ (simetría).
3. $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in X$ (desigualdad triangular).

Cuando el objetivo perseguido es resolver problemas de optimización, como es nuestro caso, X es generalmente un espacio vectorial y se suelen exigir a $d(\cdot, \cdot)$ algunas propiedades adicionales, como la convexidad, que tenga expresión analítica o que sea lineal a trozos. De forma que, para su resolución, se puedan usar herramientas del Análisis Convexo, del Análisis Matemático o de la Programación Lineal, respectivamente.

Un calibrador $\gamma(\cdot)$ es una función definida sobre un espacio vectorial X que toma valores reales y que verifica las siguientes propiedades;

1. $\gamma(x) \geq 0 \quad \forall x \in X$.
2. $\gamma(x) = 0$ si y sólo si $x = 0$.
3. $\gamma(\lambda x) = \lambda \gamma(x)$ cuando $\lambda > 0$.
4. $\gamma(x + y) \leq \gamma(x) + \gamma(y) \quad \forall x, y \in X$.

Se dice que una distancia procede de un calibrador γ , cuando para todo $x, y \in X$ se verifica que $d(x, y) = \gamma(y - x)$. Dado un calibrador $\gamma(\cdot)$ sobre X , la bola unidad asociada es el conjunto $B_\gamma = \{x \in X : \gamma(x) \leq 1\}$. Si B_γ es simétrica (i.e. $-B_\gamma = B_\gamma$), entonces $\gamma(\cdot)$ es una norma en X .

Dado un espacio vectorial X , toda función $d(\cdot, \cdot)$ definida sobre $X \times X$ que toma valores reales y verifica las propiedades de una métrica (excepto quizá la simetría) y además es convexa, procede de un calibrador.

Entre las distancias más usuales destacan la distancia Euclídea, que procede de la norma Euclídea ($\gamma(x) = \sqrt{\sum_{i=1}^n x_i^2}$) y tiene la ventaja de tener como bola unidad el círculo, que permanece invariante bajo rotaciones.

En segundo lugar, destaca la *distancia rectangular* o *distancia de Manhattan* derivada de la norma rectangular (o norma l_1) $\gamma(x) = \sum_{i=1}^n |x_i|$. Se usa para modelar situaciones en redes ortogonales o movimientos de *robots*, que generalmente tienen sus desplazamientos restringidos a dos direcciones ortogonales. En estos casos, cada movimiento está generado por motores independientes que no pueden trabajar simultáneamente y que tienen velocidades constantes e iguales.

También resulta interesante citar la *distancia de Tchebycheff*, que procede de la norma del máximo o norma l_∞ , es decir, $\gamma(x) = \max_{i=1, \dots, n} |x_i|$ y cuya bola unidad en \mathbb{R}^2 es la resultante de aplicar un giro de 45° más una homotecia a la bola unidad de la distancia rectangular.

Como es bien sabido, las distancias mencionadas anteriormente son casos particulares de las distancias l_p , que proceden de la norma $\gamma_p(x) = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$.

Finalmente, mencionamos explícitamente las distancias que proceden de calibradores poliédricos y que fueron introducidas en [49]. Sea $X = \mathbb{R}^n$ y sea $\gamma(\cdot)$ un calibrador en \mathbb{R}^n . Se dice que $\gamma(\cdot)$ es poliédrico si la correspondiente bola unidad B_γ es un politopo convexo (envolvente convexa de un número finito de puntos, ver figura 2). Como casos particulares de estos calibradores volvemos a encontrar las normas l_1 y l_∞ .

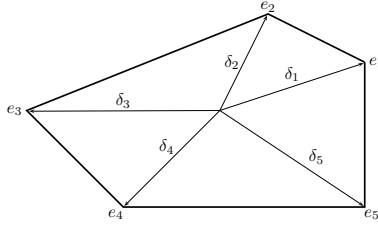


Figura 2: Ejemplo de bola unidad de calibrador poliédrico.

A continuación y para finalizar esta introducción, se describe el esquema general del resto del artículo. Los contenidos han sido divididos en seis secciones. En la segunda sección se recuerdan algunos resultados, principalmente de Análisis Convexo, que se van a utilizar a lo largo del trabajo. En las secciones tercera, cuarta y quinta, se analizan tres modelos generales de localización continua, obteniéndose caracterizaciones de sus respectivos conjuntos de soluciones. La sexta sección está dedicada a conclusiones, finalizando el trabajo con una lista de referencias bibliográficas que permitirán al lector profundizar en los contenidos.

2 Preliminares

Esta sección se dedica a preliminares matemáticos. Así, se revisan una serie de definiciones y resultados que facilitarán el entendimiento del resto del trabajo. Por simplicidad, los resultados analizados se han planteado en \mathbb{R}^n , aunque en su mayoría sigan siendo válidos con mucha mayor generalidad.

Sea γ un calibrador con bola unidad $B \subset \mathbb{R}^n$, siendo B un conjunto convexo y compacto que contiene al origen en su interior.

La bola polar de B , denominada B^0 , viene dada por

$$B^0 := \{p \in \mathbb{R}^n : \langle p, x \rangle \leq 1 \quad \forall x \in B\}.$$

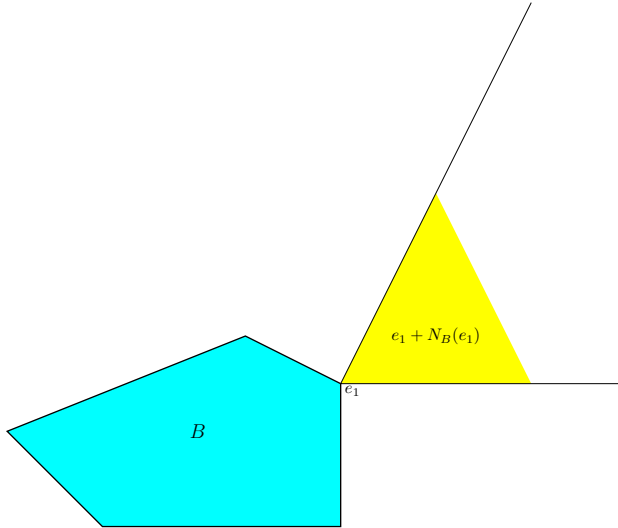
Dado un calibrador poliédrico γ cuya bola unidad es B , se denota por $Ext(B) := \{e_1, \dots, e_G\}$ al conjunto de puntos extremos de B y se llaman direcciones fundamentales de B , $\delta_1, \dots, \delta_G$, a las direcciones definidas por 0 y los e_1, \dots, e_G (véase figura 2). La bola polar asociada es entonces también un polítopo, ver [45, 12].

El cono normal a B en x viene dado por (véase figura 3)

$$N_B(x) := \{p \in \mathbb{R}^n : \langle p, y - x \rangle \leq 0 \quad \forall y \in B\}. \quad (1)$$

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa. Un vector p es un subgradiente de f en x si

$$f(y) \geq f(x) + \langle p, y - x \rangle \quad \forall y \in \mathbb{R}^n.$$

Figura 3: Cono normal a B en e_1 .

Evidentemente, si f es, por ejemplo, continuamente diferenciable en un abierto $U \subset \mathbb{R}^n$, para cada $x \in U$ el único subgradiente de f en x es $p = f'(x)$. El conjunto de todos los subgradientes de f en x se llama subdiferencial de f en x y se denota $\partial f(x)$.

Dado un conjunto cerrado C , se denota por $I_C(\cdot)$ su función *indicatriz*, es decir,

$$I_C(x) = \begin{cases} 0 & \text{si } x \in C \\ +\infty & \text{c.c.} \end{cases}$$

y se denota por $\sigma_C(\cdot)$ la función soporte del conjunto C :

$$\sigma_C(p) = \sup_{x \in C} \langle p, x \rangle \quad \forall p \in \mathbb{R}^n.$$

Se verifica que $\gamma(x) = \sigma_{B^0}(x)$. Además, se tiene que

$$\partial \gamma(x) = \begin{cases} B^0 & \text{si } x = 0 \\ \{p \in B^0 : \langle p, x \rangle = \gamma(x)\} & \text{si } x \neq 0 \end{cases} \quad (2)$$

$$\partial I_C(x) = N_C(x) \quad \forall x \in C, \quad (3)$$

(véase [21] para una prueba de estas igualdades). Sea f una función de \mathbb{R}^n en $\mathbb{R} \cup \{+\infty\}$ no idénticamente igual a $+\infty$ y minimizada por una función afín. La conjugada f^* de f es la función definida por

$$f^*(p) = \sup\{\langle p, x \rangle - f(x) : x \in \mathbb{R}^n\} \quad \forall p \in \mathbb{R}^n.$$

Es conocido que

$$u \in \partial f(x) \iff x \in \partial f^*(u).$$

Asimismo, se verifica la siguiente igualdad entre la función conjugada de la función indicatriz y la función soporte:

$$I_C^*(p) = \sigma_C(p) \quad \forall p \in \mathbb{R}^n. \quad (4)$$

El análisis de problemas de localización de naturaleza regional que se llevará a cabo en secciones posteriores hace aconsejable que a continuación, analicemos algunas propiedades de la función inf-distancia, es decir, la función que calcula la mínima distancia de un punto a un conjunto. Para ello, se consideran un conjunto convexo $A_i \subset \mathbb{R}^n$ y un calibrador arbitrario γ_i . La distancia de un punto $x \in \mathbb{R}^n$ al conjunto A_i con el calibrador γ_i viene definida por

$$d_i(x, A_i) = \inf\{\gamma_i(x - a_i) : a_i \in A_i\}.$$

Al conjunto $\text{proj}_{A_i}(x) := \{a_i \in A_i : d_i(x, A_i) = \gamma_i(x - a_i)\}$ se le denomina proyección de x sobre A_i con el calibrador γ_i . Nótese que este conjunto no es necesariamente un punto, e incluso puede ser vacío cuando A_i no es cerrado. Para asegurar que la proyección es un conjunto no vacío, impondremos desde ahora que el conjunto A_i sea compacto (obsérvese que todo calibrador γ es una función continua, ya que es convexa).

La función inf-distancia se puede expresar como sigue:

$$d_i(x, A_i) = \inf_{a_i \in A_i} \gamma_i(x - a_i) = \inf_{y \in \mathbb{R}^n} \{I_{A_i}(y) + \gamma_i(x - y)\} =: (I_{A_i} * \gamma_i)(x) \quad (5)$$

(usualmente a $(I_{A_i} * \gamma_i)(x)$ se le llama producto de inf-convolución de las funciones I_{A_i} y γ_i). Gracias a esta representación, es fácil comprobar que $d_i(\cdot, A_i)$ es una función convexa (puesto que es el resultado de una inf-convolución de dos funciones convexas). Además, por el corolario VI.4.5.5 de [21], se obtiene la siguiente expresión del subdiferencial de $d_i(\cdot, A_i)$

$$\partial d_i(\cdot, A_i)(x) = \partial I_{A_i}(a_i) \cap \partial \gamma_i(x - a_i) \quad \forall a_i \in \text{proj}_{A_i}(x).$$

Nótese que, cuando $x \in A_i$, $\text{proj}_{A_i}(x) = \{x\}$ y puesto que $\partial \gamma_i(0) = B_i^\circ$, se tiene que $\partial d_i(\cdot, A_i)(x) = N_{A_i}(x) \cap B_i^\circ$ si $x \in A_i$. Mientras que, en general, usando (2) y (3), se obtiene que

$$\partial d_i(\cdot, A_i)(x) = N_{A_i}(a_i) \cap \{p_i \in B_i^\circ : \langle p_i, x - a_i \rangle = \gamma_i(x - a_i)\} \quad \forall a_i \in \text{proj}_{A_i}(x). \quad (6)$$

Denotemos d_i^* la función conjugada de $d_i(\cdot, A_i)$. La representación (5) nos permite además calcular $\partial d_i^*(p_i)$. En efecto, puesto que $d_i(x, A_i) = (I_{A_i} * \gamma_i)(x)$, aplicando el teorema 1 de la sección 3.4 de [22], vemos que $d_i^* = I_{A_i}^* + \gamma_i^*$. Ahora bien, usando (4), resulta que $I_{A_i}^*$ es la función soporte de A_i (es decir $I_{A_i}^* = \sigma_{A_i}$) y la función conjugada de γ_i es la función indicatriz de su bola dual, es decir, $\gamma_i^* = I_{B_i^\circ}$. Por tanto, puesto que se verifica la hipótesis de cualificación de Moreau (se requiere que una de las funciones sea continua en un punto del dominio de la otra función, ver por ejemplo [22]), se tiene:

$$\partial d_i^*(p_i) = \partial(I_{A_i}^* + \gamma_i^*)(p_i) = \partial I_{A_i}^*(p_i) + \partial \gamma_i^*(p_i) = \partial \sigma_{A_i}(p_i) + N_{B_i^\circ}(p_i) =: C_i(p_i). \quad (7)$$

Una propiedad interesante de esta familia de conjuntos $(C_i(p_i))$ es que constituyen dominios de linealidad de la función $d_i(\cdot, A_i)$. Esto es lo que se prueba en el siguiente lema:

Lema 2.1 *Para cada $p_i \in B_i^o$, $d_i(\cdot, A_i)$ es una función afín en $\partial d_i^*(p_i)$.*

Demostración:

Por la identidad de Fenchel se tiene que

$$x \in \partial d_i^*(p_i) \iff p_i \in \partial d_i(x, A_i).$$

Por tanto, aplicando (6), para cada $x \in \partial d_i^*(p_i)$ se obtiene que

$$d_i(x, A_i) = \langle p_i, x - a_i \rangle = \langle p_i, x \rangle - \langle p_i, a_i \rangle \quad \forall a_i \in \text{proj}_{A_i}(x).$$

Ahora bien, puesto que $p_i \in \partial d_i(x, A_i)$, se cumple que $p_i \in N_{A_i}(a_i) = \partial \sigma_{A_i}^*(a_i)$, siendo esto equivalente a que $a_i \in \partial \sigma_{A_i}(p_i)$. Por tanto, $\langle p_i, a_i \rangle = \sigma_{A_i}(p_i)$ para cualquier $a_i \in \partial \sigma_{A_i}(p_i)$, es decir, $\langle p_i, a_i \rangle$ es constante para cualquier $a_i \in \partial \sigma_{A_i}(p_i)$. Además, puesto que $p_i \in N_{A_i}(a_i)$ para cualquier $a_i \in \text{proj}_{A_i}(x)$, entonces $\langle p_i, a \rangle \leq \langle p_i, a_i \rangle$ para todo $a \in A_i$, es decir, $\text{proj}_{A_i}(x) \subseteq \partial \sigma_{A_i}(p_i)$. De este modo, para $x \in \partial d_i^*(p_i)$ se verifica

$$d_i(x, A_i) = \langle p_i, x \rangle - \sigma_{A_i}(p_i) = \langle p_i, x - a_i \rangle \quad \forall a_i \in \text{proj}_{A_i}(x),$$

lo que prueba el resultado. □

Indiquemos finalmente que, en el resto del trabajo, $co(A)$ denotará la envolvente convexa del conjunto A .

3 Un modelo general de localización continua

En esta sección se analizará un modelo de localización con demanda regional y función objetivo dependiendo de las distancias entre los conjuntos. Esta formulación permite caracterizar geoméricamente el conjunto solución de una amplia gama de problemas de localización que se pueden obtener como casos particulares de este modelo.

Sea $\mathcal{A} = \{A_1, \dots, A_M\}$ una familia de conjuntos en \mathbb{R}^n , donde cada A_i es un conjunto convexo y compacto. Sea $\Phi(\cdot)$ una norma monótona sobre \mathbb{R}^M . Recordemos que una norma Φ se dice que es monótona si $\Phi(u) \leq \Phi(v)$ para cada u, v verificando que $|u_i| \leq |v_i|$ para todo $i = 1, \dots, M$ (ver [2]). Con esta notación, la formulación del problema considerado en esta sección es:

$$\inf_{x \in \mathbb{R}^n} F(x) := \Phi(d(x)), \quad (P_{\Phi}(\mathcal{A}))$$

donde $d(x) = (d_1(x, A_1), \dots, d_M(x, A_M))$.

Funciones objetivo similares a ésta han sido consideradas en problemas estándar de localización, i.e. cuando los centros de demanda son puntos (ver

[6, 13, 14]). La novedad de este modelo proviene de considerar demanda regional. Se puede suponer sin pérdida de generalidad que $\cap_{i=1}^M A_i = \emptyset$ ($d(x) \neq 0$ para todo $x \in \mathbb{R}^n$). En efecto, si $\cap_{i=1}^M A_i \neq \emptyset$ entonces el conjunto solución sería $\cap_{i=1}^M A_i$ con valor objetivo cero.

Recuérdese que para un problema de minimización sin restricciones con función objetivo f convexa, \hat{x} es una solución si y sólo si $0 \in \partial f(\hat{x})$. Esto hace esencial la caracterización del subdiferencial de la función objetivo F .

Lema 3.1 *Sea $x \in \mathbb{R}^n$. El punto $q \in \partial F(x)$ si y sólo si existen $a_i \in \text{proj}_{A_i}(x)$, $p_i \in N_{A_i}(a_i) \cap B_i^o$ $1 \leq i \leq M$ y $\lambda = (\lambda_1, \dots, \lambda_M) \geq 0$ tales que*

1. $x \in \bigcap_{i=1}^M (a_i + N_{B_i^o}(p_i))$,
2. $\Phi^o(\lambda) = 1$ y $\sum_{i=1}^M \lambda_i d_i(x, A_i) = F(x)$, donde Φ^o es la norma dual de Φ ,
3. $q = \sum_{i=1}^M \lambda_i p_i$.

3.1 Caracterización geométrica del conjunto de soluciones

Para caracterizar geoméricamente el conjunto de soluciones del problema $(P_\Phi(\mathcal{A}))$, se introducen las siguientes familias:

Definición 3.1 *Dados $p = (p_1, \dots, p_M) \in (\mathbb{R}^n)^M$ con $p_i \in B_i^o$ e $I \subseteq \{1, \dots, M\}$, denotemos*

$$C_I(p) := \bigcap_{i \in I} \partial d_i^*(p_i),$$

donde d_i^* es la función conjugada de $d_i(x, A_i)$. Por otra parte, para cualquier $\lambda = (\lambda_1, \dots, \lambda_M) \geq 0$, llamaremos $D_I(\lambda)$ al conjunto siguiente:

$$D_I(\lambda) := \{x : \sum_{i \in I} \lambda_i d_i(x, A_i) = F(x)\}.$$

Nótese que $C_I(p)$ es no vacío solamente para algunas elecciones de I y p . Estos conjuntos son llamados convexos elementales cuando las funciones distancias son normas (diferentes definiciones de convexos elementales pueden verse en [12, 26, 31, 35]). Por esta razón, los conjuntos $C_I(p)$ se denominan convexos elementales generalizados (g.e.c.s.). Los g.e.c.s. son convexos, pues están definidos por intersecciones de conjuntos convexos (recuérdese que los subdiferenciales son conjuntos convexos).

En el caso de que las bolas unidad consideradas sean polítopos, entonces los g.e.c.s. pueden ser obtenidos como intersecciones de los conos generados por las direcciones fundamentales de estas bolas.

De este modo, usando el lema 2.1 y la compacidad del conjunto solución (ver [30]), el problema de Fermat-Weber con distancias ínfimas poliédricas tiene una solución en el conjunto de vértices de los g.e.c.s. Esta caracterización es una generalización de la propiedad de *puntos de intersección* obtenida en [47] para la norma l_1 en \mathbb{R}^2 , en [41] para el caso de una norma poliédrica y en [12] para el problema de Fermat-Weber con coste lineal.

Sea $M_\Phi(\mathcal{A})$ el conjunto de soluciones de $(P_\Phi(\mathcal{A}))$. Se dice que (I, λ, p) es una tripleta admisible si

1. $I \neq \emptyset, I \subseteq \{1, \dots, M\}$.
2. $\lambda = (\lambda_1, \dots, \lambda_M)$ con $\lambda_i > 0$ ($i \in I$) y $\lambda_i = 0$ ($i \notin I$) y además $\Phi^\circ(\lambda) = 1$.
3. $p = (p_1, \dots, p_M)$ con $p_i \in B_i^\circ \cap N_{A_i}(y_i)$ para cualquier $y_i \in \partial\sigma_{A_i}(p_i)$ ($i \in I$) y $\sum_{i \in I} \lambda_i p_i = 0$.

Lema 3.2 $\hat{x} \in M_\Phi(\mathcal{A})$ si y sólo si existe una tripleta admisible (I, λ, p) satisfaciendo

$$\hat{x} \in C_I(p) \cap D_I(\lambda).$$

Demostración:

Observemos que $\hat{x} \in M_\Phi(\mathcal{A})$ si y sólo si $0 \in \partial F(\hat{x})$. De este modo, aplicando el lema 3.1 y las definiciones de $C_I(p)$ y $D_I(\lambda)$, el resultado queda probado. \square

Nótese que el lema 3.2 implica que para cada $\hat{x} \in M_\Phi(\mathcal{A})$ existe un tripleta admisible tal que

$$C_I(p) \cap D_I(\lambda) \subseteq M_\Phi(\mathcal{A}).$$

El siguiente resultado prueba que en realidad la inclusión anterior es una igualdad y además dicha igualdad es independiente de la elección de la tripleta admisible (I, λ, p) siempre que $C_I(p) \cap D_I(\lambda) \neq \emptyset$:

Teorema 3.1

1. Si $M_\Phi(\mathcal{A}) \neq \emptyset$, entonces existe una tripleta admisible (I, λ, p) tal que $M_\Phi(\mathcal{A}) = C_I(p) \cap D_I(\lambda)$.
2. Para cualquier tripleta admisible (I, λ, p) tal que $C_I(p) \cap D_I(\lambda) \neq \emptyset$, se tiene que $M_\Phi(\mathcal{A}) = C_I(p) \cap D_I(\lambda)$.

Demostración:

Sea (I, λ, p) una tripleta admisible con

$$\emptyset \neq C_I(p) \cap D_I(\lambda) \subseteq M_\Phi(\mathcal{A}).$$

La existencia de esta tripleta está garantizada por el lema 3.2 cuando $M_\Phi(\mathcal{A}) \neq \emptyset$. Veamos que cualquier $\bar{x} \in M_\Phi(\mathcal{A})$ verifica $\bar{x} \in C_I(p) \cap D_I(\lambda)$.

Sea $\bar{x} \in M_{\Phi}(\mathcal{A})$ y sea $x^* \in C_I(p) \cap D_I(\lambda)$. Entonces existe $a_i(x^*) \in \text{proj}_{A_i}(x^*)$ tal que

$$F^* := F(x^*) = \sum_{i=1}^M \lambda_i \langle p_i, x^* - a_i(x^*) \rangle = - \sum_{i=1}^M \lambda_i \langle p_i, a_i(x^*) \rangle.$$

Por otro lado, puesto que F^* es el mínimo valor de la función F , se tiene que $\langle p_i, a_i(x^*) \rangle = \sup_{a_i \in A_i} \langle p_i, a_i \rangle$, es decir, $a_i(x^*) \in \partial\sigma_{A_i}(p_i)$ para todo $i \in I$.

Ahora bien, para todo $x \in \mathbb{R}^n$, se tiene que

$$\begin{aligned} F^* &= - \sum_{i=1}^M \lambda_i \langle p_i, a_i(x^*) \rangle \leq - \sum_{i=1}^M \lambda_i \langle p_i, a_i(x) \rangle \\ &= \sum_{i=1}^M \lambda_i \langle p_i, x - a_i(x) \rangle \quad \forall a_i(x) \in \text{proj}_{A_i}(x). \end{aligned}$$

Puesto que $d_i(x, A_i) = \sup_{q_i \in B_i^o} \langle q_i, x - a_i(x) \rangle = \gamma_i(x - a_i(x))$, usando que $\Phi(\cdot)$ es una norma y $\Phi^o(\lambda) = 1$, se obtiene que

$$F^* \leq \sum_{i=1}^M \lambda_i \langle p_i, x - a_i(x) \rangle \leq \sum_{i=1}^M \lambda_i d_i(x, A_i) \leq F(x). \quad (8)$$

Debido a que la desigualdad anterior es cierta para todo x , si se toma en particular $x = \bar{x} \in M_{\Phi}(\mathcal{A})$, todas las desigualdades se convierten en igualdades, es decir,

$$F^* = \sum_{i=1}^M \lambda_i \langle p_i, \bar{x} - a_i(\bar{x}) \rangle = \sum_{i=1}^M \lambda_i d_i(\bar{x}, A_i) \quad \forall a_i(\bar{x}) \in \text{proj}_{A_i}(\bar{x})$$

y

$$\sum_{i=1}^M \lambda_i \langle p_i, a_i(x^*) \rangle = \sum_{i=1}^M \lambda_i \langle p_i, a_i(\bar{x}) \rangle.$$

Esto, junto con las desigualdades existentes entre los correspondientes términos, conduce a que para todo $i \in I$ se verifica: (1) $\langle p_i, \bar{x} - a_i(\bar{x}) \rangle = d_i(\bar{x}, A_i)$ y (2) $\langle p_i, a_i(\bar{x}) \rangle = \langle p_i, a_i(x^*) \rangle$.

Por (1), se obtiene que

$$d_i(\bar{x}, A_i) = \gamma_i(\bar{x} - a_i(\bar{x})) = \langle p_i, \bar{x} - a_i(\bar{x}) \rangle \quad \forall i \in I.$$

De esta forma, $p_i \in \partial\gamma_i(\bar{x} - a_i(\bar{x}))$, lo que equivale a tener $\bar{x} - a_i(\bar{x}) \in \partial\gamma_i^*(p_i)$ para todo $i \in I$. Por (2) y puesto que $a_i(x^*) \in \partial\sigma_{A_i}(p_i)$, se deduce que $a_i(\bar{x}) \in \partial\sigma_{A_i}(p_i)$ para todo $i \in I$. Por tanto,

$$\bar{x} \in a_i(\bar{x}) + \partial\gamma_i^*(p_i) \subset \partial\sigma_{A_i}(p_i) + \partial\gamma_i^*(p_i) = C_i(p_i) \quad \forall i \in I,$$

(ver (7) para la definición de $C_i(p_i)$). Además, puesto que $\bar{x} \in M_{\Phi}(\mathcal{A})$, usando la última desigualdad en (8), se tiene que $F(\bar{x}) = \sum_{i \in I} \lambda_i d_i(\bar{x}, A_i)$, lo que implica $\bar{x} \in D_I(\lambda)$. Por tanto, $\bar{x} \in C_I(p) \cap D_I(\lambda)$ y el resultado queda demostrado. \square

3.2 Aplicación a dos problemas clásicos de localización

En esta sección se particularizan las caracterizaciones obtenidas anteriormente al problema de Fermat-Weber y al problema minimax. Para ello, al igual que antes, se considera una familia de conjuntos $\mathcal{A} = \{A_1, \dots, A_M\}$ donde cada A_i es compacto y convexo, un conjunto de pesos positivos $W = \{w_1, \dots, w_M\}$ y un conjunto de normas en \mathbb{R}^n $\{\gamma_1(\cdot), \dots, \gamma_M(\cdot)\}$, cada una con bola unidad B_i .

3.2.1 El problema de Fermat-Weber

Este problema, también denominado problema de Weber, problema de Fermat-Steiner-Weber, problema de la mediana o problema minisum, tiene sus orígenes, como se mencionó en la sección 1, en el trabajo de Pierre Fermat.

El problema de Fermat-Weber con distancias ínfimas con respecto a \mathcal{A} y W es el siguiente:

$$\inf_{x \in \mathbb{R}^n} G(x) := \sum_{i=1}^M w_i d_i(x, A_i), \quad (P_W(\mathcal{A}))$$

donde $d_i(x, A_i) = \inf_{a_i \in A_i} \gamma_i(x - a_i)$.

Nuestro principal objetivo será caracterizar el conjunto de soluciones $M_W(\mathcal{A})$ de $P_W(\mathcal{A})$. Los resultados siguientes son casos particulares del lema 3.1 y el teorema 3.1, tomando como Φ la norma rectangular en \mathbb{R}^M y tomando $\gamma'_i = w_i \gamma_i$ para todo $i \in \{1, \dots, M\}$. Por ello, las demostraciones son omitidas.

Lema 3.3 *Para cualquier $x \in \mathbb{R}^n$ se tiene que $q \in \partial G(x)$ si y sólo si existen $a_i \in \text{proj}_{A_i}(x)$, $p_i \in N_{A_i}(a_i) \cap B_i^o$ $1 \leq i \leq M$, tales que*

1. $x \in \bigcap_{i=1}^M (a_i + N_{B_i^o}(p_i))$,
2. $q = \sum_{i=1}^M w_i p_i$.

Teorema 3.2

1. Si $M_W(\mathcal{A}) \neq \emptyset$, entonces existe una tripleta admisible (I, λ, p) con $\lambda_i = w_i$ $\forall i \in I$ tal que $M_W(\mathcal{A}) = C_I(p) = \bigcap_{i \in I} (\partial \sigma_{A_i}(p_i) + N_{B_i^o}(p_i))$.
2. Si existe una tripleta admisible (I, λ, p) con $\lambda_i = w_i$ $\forall i \in I$ tal que $C_I(p) \neq \emptyset$, entonces $M_W(\mathcal{A}) = C_I(p)$.

Ejemplo 3.1 (Ver la figura 4) Consideremos tres conjuntos en \mathbb{R}^2 , con $\gamma_i = l_1$ para cada i . Dichos conjuntos son $A_1 := \text{co}\{(0, 1), (-1, 2), (1, 2)\}$, $A_2 := \text{co}\{(2, -0'5), (2, 0'5), (3, 0'5), (3, -0'5)\}$ y $A_3 := \text{co}\{(-2, -2), (-2, -1), (-3, -1), (-3, -2)\}$. Tomamos $w_1 = w_2 = w_3 = 1$.

En este caso es fácil comprobar que los *g.e.c.s.* $C_I(p)$, son los conjuntos delimitados por las líneas dibujadas en la figura 4. Así por ejemplo para

$p = (p_1, p_2, p_3)$ con $p_1 = (0, -1)$, $p_2 = (-1, 0)$ y $p_3 = (1, 1)$, se tiene que: $C_{\{1\}}(p) = \{(0, 1 - x) : x \geq 0\}$, $C_{\{2\}}(p) = \{(x_1, x_2) : x_1 \leq 2, -0'5 \leq x_2 \leq 0'5\}$ y $C_{\{3\}}(p) = \{(x_1, x_2) : x_1 \geq -2, x_2 \geq -1\}$. Además, se tiene que el conjunto de soluciones viene dado por:

$$M_W(A) = C_{\{1,2,3\}}((p_1, p_2, p_3)) = co\{(0, -0'5), (0, 0'5)\}.$$

Obsérvese que efectivamente (I, λ, p) con $I = \{1, 2, 3\}$, $(w_1, w_2, w_3) = (1, 1, 1)$ y $p = (p_1, p_2, p_3)$ es una triplete admisible.

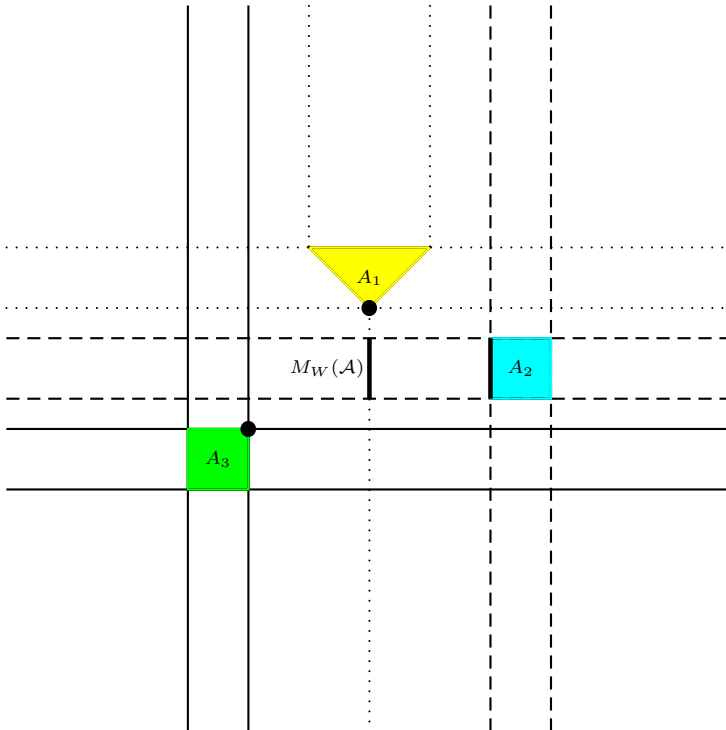


Figura 4: Ilustración del Ejemplo 3.1

El lector puede comprobar que los resultados que caracterizan el conjunto de soluciones del problema de Fermat-Weber en el caso puntual (ver [12]) pueden obtenerse como caso particular del anterior al considerar conjuntos unitarios.

Corolario 3.1 *El problema de Fermat-Weber con distancias mínimas siempre tiene una solución en el conjunto de puntos extremos de los correspondientes g.e.c.s.*

El siguiente resultado es similar al Teorema de la “Mayoría”, bien conocido en el caso puntual. Para ello, se supone que $\gamma_i = \gamma$ para todo $i = 1, \dots, M$ y que $d(x, y) = \gamma(x - y)$ (todas las distancias coinciden y son generadas por una única norma).

Teorema 3.3 *Si γ es una norma en \mathbb{R}^n , $M_W(\mathcal{A}) \neq \emptyset$ y existe $A_i \in \mathcal{A}$ tal que $w_i \geq \sum_{i \neq j} w_j$, entonces existe una solución en A_i .*

Demostración:

Sea $x^* \in M_W(\mathcal{A})$ y supongamos que $x^* \notin A_i$. Si $\hat{x} \in \text{proj}_{A_i}(x^*)$, entonces se tiene que

$$G(x^*) = \sum_{i=1}^M w_i d(x^*, A_i) \leq G(\hat{x}) = \sum_{j \neq i} w_j d(\hat{x}, A_j).$$

Ahora bien, por la desigualdad triangular, se tiene también que

$$\begin{aligned} \sum_{j \neq i} w_j d(\hat{x}, A_j) &\leq \sum_{j \neq i} w_j (\gamma(\hat{x} - x^*) + d(x^*, A_j)) \\ &\leq w_i \gamma(\hat{x} - x^*) + \sum_{j \neq i} w_j d(x^*, A_j) \\ &= G(x^*). \end{aligned}$$

Por tanto, \hat{x} es también una solución. □

El caso puntual

En el caso particular en que los conjuntos A_i son unitarios, es posible dar mejores caracterizaciones del conjunto solución. En este apartado se revisan resultados de coincidencia de soluciones con puntos de demanda, además del problema de Fermat-Weber con norma l_1 y l_2 , en este caso. Denotemos $A = \{a_1, \dots, a_M\}$.

Soluciones coincidentes con los puntos de demanda

En primer lugar se estudian los casos donde alguna solución del problema de Fermat-Weber puntual coincide con un punto de demanda. Por cuestión de simplicidad, pondremos $\varphi_i(x) = w_i \gamma_i(x - a_i)$.

Teorema 3.4 *Se verifica que $a_j \in A$ es una solución de $(P_W(A))$ si y sólo si existe $p_i \in \partial \varphi_i(a_j) \forall i \in \{1, \dots, M\} \setminus \{j\}$ tal que $\gamma_j^0(-\sum_{i \neq j} w_i p_i) \leq w_j$.*

Demostración:

Por el teorema 3.2, $a_j \in M_W(A)$ si y sólo si existen $p_i \in \partial \varphi_i(a_j)$ ($1 \leq i \leq M$), tales que $\sum_{i=1}^M w_i p_i = 0$ y $\bigcap_{i=1}^M a_i + N_{B^0}(p_i) \neq \emptyset$.

Puesto que $p_i \in \partial\varphi_i(a_j) \forall i \in \{1, \dots, M\} \setminus \{j\}$, se tiene que $\gamma_i^0(p_i) = 1$. Además, para cada a_j se cumple que $p_j \in \partial\varphi_j(a_j) = \partial\gamma_j(a_j - a_j) = \partial\gamma_j(0)$; lo que implica que $\gamma_j^0(p_j) \leq 1$.

De esta forma,

$$a_j \in M_W(A) \Leftrightarrow \exists p_i \in \partial\varphi_i(a_j) \quad (1 \leq i \leq M) \text{ tales que } w_j p_j = - \sum_{i \neq j} w_i p_i.$$

Entonces, puesto que $\gamma_j^0(w_j p_j) = \gamma_j^0(-\sum_{i \neq j} w_i p_i)$, se tiene que

$$w_j \geq w_j \gamma_j^0(p_j) = \gamma_j^0 \left(- \sum_{i \neq j} w_i p_i \right).$$

Recíprocamente, si se supone que $w_j \geq \gamma_j^0(-\sum_{i \neq j} w_i p_i)$, se tiene que $-\sum_{i \neq j} w_i p_i \in w_j B_j^0$. De esta forma, existe $p_j \in B_j^0$ tal que $w_j p_j = -\sum_{i \neq j} w_i p_i$, con lo que el resultado queda probado. \square

El teorema anterior permite reducir el problema clásico de Fermat a un simple ejercicio.

Ejemplo 3.2 (Problema de Fermat en el plano; véase la figura 1). Sean a , b y c tres puntos de demanda no alineados en el plano, donde las distancias vienen medidas con la norma Euclídea. El problema clásico de Fermat se formula así:

$$\min_{x \in \mathbb{R}^2} \|x - a\|_2 + \|x - b\|_2 + \|x - c\|_2.$$

En primer lugar, se establecen las condiciones necesarias y suficientes para que a , b ó c sea solución del problema anterior. Sin pérdida de generalidad, se obtiene que c es óptima si y sólo si:

$$\left\| \frac{c - a}{\|c - a\|_2} + \frac{c - b}{\|c - b\|_2} \right\|_2 \leq 1,$$

lo que ocurre si y sólo si

$$\left\| \frac{c - a}{\|c - a\|_2} + \frac{c - b}{\|c - b\|_2} \right\|_2^2 \leq 1.$$

Se denota por $u^t v$ el producto escalar de los vectores u y v . Desarrollando el lado izquierdo de la desigualdad anterior se obtiene que

$$2 + 2 \left(\frac{c - a}{\|c - a\|_2} \right)^t \left(\frac{c - b}{\|c - b\|_2} \right) \leq 1,$$

de donde se deduce que

$$c \text{ es óptimo si y sólo si } \left(\frac{c - a}{\|c - a\|_2} \right)^t \left(\frac{c - b}{\|c - b\|_2} \right) \leq -\frac{1}{2}.$$

Esto quiere decir que el ángulo de vértice c asociado al triángulo T de vértices a , b y c tiene más de 120° .

En segundo lugar, se supone que ningún ángulo es de más de 120° , es decir, la solución no está en ningún vértice. Entonces la solución se encuentra en un punto de diferenciabilidad de la función objetivo. Además, se sabe que la solución está en el punto donde se anula el gradiente.

Si se definen los vectores

$$p_a(x) = \frac{x - a}{\|x - a\|_2}, \quad p_b(x) = \frac{x - b}{\|x - b\|_2}, \quad p_c(x) = \frac{x - c}{\|x - c\|_2},$$

se tiene que

$$\hat{x} \text{ es óptimo si y sólo si } p_a(\hat{x}) + p_b(\hat{x}) + p_c(\hat{x}) = 0.$$

Multiplicando por $p_a(\hat{x})$, $p_b(\hat{x})$ y $p_c(\hat{x})$, se obtiene que:

$$\begin{cases} p_a(\hat{x})^t p_b(\hat{x}) + p_a(\hat{x})^t p_c(\hat{x}) = -1 \\ p_b(\hat{x})^t p_a(\hat{x}) + p_b(\hat{x})^t p_c(\hat{x}) = -1 \\ p_c(\hat{x})^t p_a(\hat{x}) + p_c(\hat{x})^t p_b(\hat{x}) = -1 \end{cases}$$

de donde

$$p_a(\hat{x})^t p_b(\hat{x}) = p_a(\hat{x})^t p_c(\hat{x}) = p_b(\hat{x})^t p_c(\hat{x}) = -\frac{1}{2}.$$

Puesto que $p_a(\hat{x})$, $p_b(\hat{x})$ y $p_c(\hat{x})$ son vectores unitarios, la igualdad anterior implica que el punto \hat{x} forma ángulos de 120° con los vértices del triángulo T . En consecuencia, \hat{x} ha de ser el punto de Torricelli del triángulo determinado por a , b y c (véase la sección 1).

El caso n -dimensional con norma $\|\cdot\|_1$

El problema de Fermat-Weber con distancia rectangular es un modelo con muchas aplicaciones (localización en zonas urbanas, localización de componentes trasladados por “robots”, etc.). Además de su aplicabilidad, destaca su fácil resolución, ya que

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^M w_i \|x - a_i\|_1 = \sum_{k=1}^n \min_{x_k \in \mathbb{R}} \sum_{i=1}^M w_i |x_k - a_{ik}|.$$

Es decir, la resolución del problema de Fermat-Weber con norma $\|\cdot\|_1$ en \mathbb{R}^n es equivalente a resolver n problemas de Fermat-Weber en \mathbb{R} con la norma habitual.

Aunque el procedimiento anterior proporciona un método de resolución eficiente, es posible desarrollar un método alternativo para resolver este problema que además se puede extender a cualquier calibrador poliédrico. En primer lugar, se observa que

$$B_1 = \{x \in \mathbb{R}^n : \|x\|_1 \leq 1\} = \text{co}\{(1, 0), (0, 1), (-1, 0), (0, -1)\},$$

y que su bola dual, B_1^0 , viene dada por:

$$B_1^0 = \{(u_1, u_2) : -1 \leq u_1 \leq 1, -1 \leq u_2 \leq 1\}.$$

De esta forma, $\|\cdot\|_1$ tiene como expresión alternativa

$$\begin{aligned} \|x\|_1 &= \max\{\langle u, x \rangle : u \in \text{ext}(B_1^0)\} \\ &= \max\{x_1 + x_2, -x_1 + x_2, -x_1 - x_2, x_1 - x_2\}. \end{aligned}$$

Por tanto, el problema original se puede reformular como

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^M w_i \|x - a_i\|_1 = \min_{x \in \mathbb{R}^n} \sum_{i=1}^M w_i \max_{u \in \text{ext}(B_1^0)} \{\langle u, x - a_i \rangle\}$$

y éste a su vez puede ser formulado como un problema de Programación Lineal. Para ello, se introducen unas variables auxiliares z_i tales que $\max\{\langle u, x - a_i \rangle : u \in \text{ext}(B_1^0)\} \leq z_i$. Entonces el problema anterior es equivalente a:

$$\begin{aligned} \min \quad & \sum_{i=1}^M w_i z_i \\ \text{sujeto a} \quad & \langle (1, 1), x - a_i \rangle \leq z_i \quad \forall i \in \{1, \dots, M\} \\ & \langle (-1, 1), x - a_i \rangle \leq z_i \quad \forall i \in \{1, \dots, M\} \\ & \langle (-1, -1), x - a_i \rangle \leq z_i \quad \forall i \in \{1, \dots, M\} \\ & \langle (1, -1), x - a_i \rangle \leq z_i \quad \forall i \in \{1, \dots, M\} \\ & x \in \mathbb{R}^n, \quad z_i \in \mathbb{R}^n. \end{aligned}$$

El lector puede comprobar que el problema de Fermat-Weber con cualquier otro calibrador poliédrico también se transforma en un problema de Programación Lineal (de forma general, con mayor número de restricciones).

El problema de Fermat-Weber con distancia Euclídea

Sin duda la norma Euclídea es la que podría considerarse como norma natural en las aplicaciones del problema de Fermat-Weber. Por ello, esta sección está dedicada a analizar este problema con la norma Euclídea, $\|\cdot\|_2$. Es decir,

$$\min_{x \in \mathbb{R}^n} F(x) := \sum_{i=1}^M w_i \|x - a_i\|_2.$$

Para resolver este problema, se calcula el subdiferencial de dicha norma. Denotando $\varphi_i(x) = w_i \|x - a_i\|_2$, se obtiene que

$$\partial\varphi_i(x) = \begin{cases} w_i \frac{x - a_i}{\|x - a_i\|_2} & \text{si } x \neq a_i \\ a_i + B_2 & \text{si } x = a_i. \end{cases}$$

Ahora bien, utilizando el teorema 3.4, se obtiene que $a_j \in A$ es solución si y sólo si para cada $i \neq j$ existe $p_i \in \partial\varphi_i(a_j)$ tal que

$$\left\| \sum_{i \neq j} w_i p_i \right\|_2 \leq w_j,$$

es decir si y sólo si

$$\left\| \sum_{i \neq j} w_i \frac{a_j - a_i}{\|a_j - a_i\|_2} \right\|_2 \leq w_j.$$

Por otro lado, suponiendo que la solución no está sobre ningún punto de demanda, la condición necesaria de optimalidad (en este caso, debido a la convexidad, también suficiente) viene dada por

$$\sum_{i=1}^M w_i \frac{x - a_i}{\|x - a_i\|_2} = 0$$

o, equivalentemente,

$$x = \frac{\sum_{i=1}^M w_i \frac{a_i}{\|x - a_i\|_2}}{\sum_{i=1}^M w_i \frac{1}{\|x - a_i\|_2}}.$$

una ecuación de punto fijo. A partir de ella, queda sugerido el siguiente proceso iterativo

$$x^{(l+1)} = \frac{\sum_{i=1}^M w_i \frac{a_i}{\|x^{(l)} - a_i\|_2}}{\sum_{i=1}^M \frac{w_i}{\|x^{(l)} - a_i\|_2}} = x^{(l)} - \frac{\sum_{i=1}^M w_i \frac{x^{(l)} - a_i}{\|x^{(l)} - a_i\|_2}}{\sum_{i=1}^M \frac{w_i}{\|x^{(l)} - a_i\|_2}}.$$

Esto es conocido en la literatura como el algoritmo de Weiszfeld. Del argumento precedente se deduce que es un algoritmo de descenso y se puede demostrar que converge a una solución, supuesto que ésta no es un punto de demanda.

Este algoritmo se puede extender al caso de normas l_p con $p \in [1, 2]$ (ver [24]).

3.2.2 El problema min max con inf-distancias.

El otro problema clásico de localización es la minimización de la distancia máxima o problema de círculo de menor radio (Sylvester 1857). La versión de este problema con inf-distancias consiste en hallar el mínimo círculo que incluya las regiones de demanda y su formulación es:

$$\min_{x \in \mathbb{R}^n} H(x) := \max_{1 \leq i \leq M} w_i d_i(x, A_i), \quad (9)$$

donde $d_i(x, A_i) = \inf_{a \in A_i} \gamma_i(x - a)$ para cada i . Denotamos $M_W^{l_\infty}(\mathcal{A})$ el conjunto de soluciones de (9) y definimos para cada $I \subseteq \{1, \dots, M\}$ y cada $\alpha \geq 0$ el siguiente conjunto:

$$AS_I(\alpha) = \{x \in \mathbb{R}^n : w_i d_i(x, A_i) = \alpha \ \forall i \in I, \quad w_i d_i(x, A_i) < \alpha \ \forall i \notin I\}.$$

El siguiente teorema proporciona una caracterización de $M_W^{l_\infty}(\mathcal{A})$.

Teorema 3.5

1. Si $M_W^{l_\infty}(\mathcal{A}) \neq \emptyset$, entonces existen una tripleta admisible (I, λ, p) y un número $\alpha \geq 0$ tales que

$$M_W^{l_\infty}(\mathcal{A}) = C_I(p) \cap AS_I(\alpha).$$

2. Para cualquier tripleta admisible (I, λ, p) y cualquier $\alpha \geq 0$ tales que $C_I(p) \cap AS_I(\alpha) \neq \emptyset$, se verifica

$$M_W^{l_\infty}(\mathcal{A}) = C_I(p) \cap AS_I(\alpha).$$

Demostración:

La demostración consiste en aplicar el teorema 3.1 con $\Phi = \|\cdot\|_\infty$ en \mathbb{R}^M y $\gamma'_i = w_i \gamma_i$ para todo i .

Puesto que $\Phi = \|\cdot\|_\infty$, esto implica $\Phi^o = \|\cdot\|_1$. De este modo, $\gamma^o(\lambda) = 1$ si y sólo si existe $I \subseteq \{1, \dots, M\}$ tal que $\sum_{i \in I} \lambda_i = 1$. Por tanto, decir que $\alpha = H(x) = \max_{1 \leq i \leq M} w_i d_i(x, A_i) = \sum_{i \in I} \lambda_i w_i d_i(x, A_i)$ con $\sum_{i \in I} \lambda_i = 1$ es equivalente a decir que $w_i d_i(x, A_i) = \alpha$ para cualquier $i \in I$ y $w_i d_i(x, A_i) < \alpha$ para cualquier $i \notin I$. En otras palabras, $x \in D_I(\lambda)$ si y sólo si $x \in AS_I(\alpha)$ para $\alpha = H(x)$ y la demostración queda concluida. \square

Nota 3.1 El valor de α que define el conjunto solución $AS_I(\alpha)$ en el teorema anterior, es el valor óptimo del problema (9). \square

Ejemplo 3.3 (Ver la figura 5) Consideremos un problema en \mathbb{R}^2 con:

1. $\Phi(x_1, x_2, x_3) = \max\{|x_1|, |x_2|, |x_3|\}$.
2. Conjuntos de demanda $\mathcal{A} = \{A_1 := \text{co}\{(5, 1), (5, -1), (3, -1), (3, 1)\}, A_2 := \text{co}\{(1, -3), (0, -5), (-1, -3)\}, A_3 := \text{co}\{(-3, 3), (-3, 5), (-5, 4)\}$.
3. Conjunto de pesos $W = \{1, 1, 1\}$ y $\gamma_i = l_\infty$ para $i = 1, 2, 3$.

Por tanto, la formulación como problema min max viene dada por:

$$\min_{x \in \mathbb{R}^2} H(x) := \max_{i=1,2,3} d_i(x, A_i).$$

Tomando $I = \{2, 3\}$, $p_2 = (0, 1)$, $p_3 = (0, -1)$, se sigue que:

$$C_I(p) = \{x \in \mathbb{R}^2 : x_2 \geq -3, x_1 + x_2 \leq 0, x_1 + x_2 \geq -4, x_1 - x_2 \geq -6, x_1 - x_2 \leq 4\}.$$

Para $\alpha = 3$, se tiene que $AS_I(3) = \{(0, 0)\}$, que coincide con $D_I((0, 0'5, 0'5))$ (obsérvese que para $\lambda = (0, 0'5, 0'5)$ se tiene que $\Phi^o(\lambda) = 1$). De hecho, este conjunto se define como

$$\begin{aligned} D_I((0, 0'5, 0'5)) &= \{x : \max_{1 \leq i \leq 3} d_i(x, A_i) = \frac{1}{2} \langle (0, 1), x - a_2 \rangle + \frac{1}{2} \langle (0, -1), x - a_3 \rangle\} \\ &= \{(0, 0)\}, \end{aligned}$$

donde $a_2 = (0, -3)$ y $a_3 = (-3, 3)$. Por tanto, el conjunto de soluciones es

$$M_W^{l\infty}(\mathcal{A}) = C_I(p) \cap AS_I(\alpha) = \{(0, 0)\}.$$

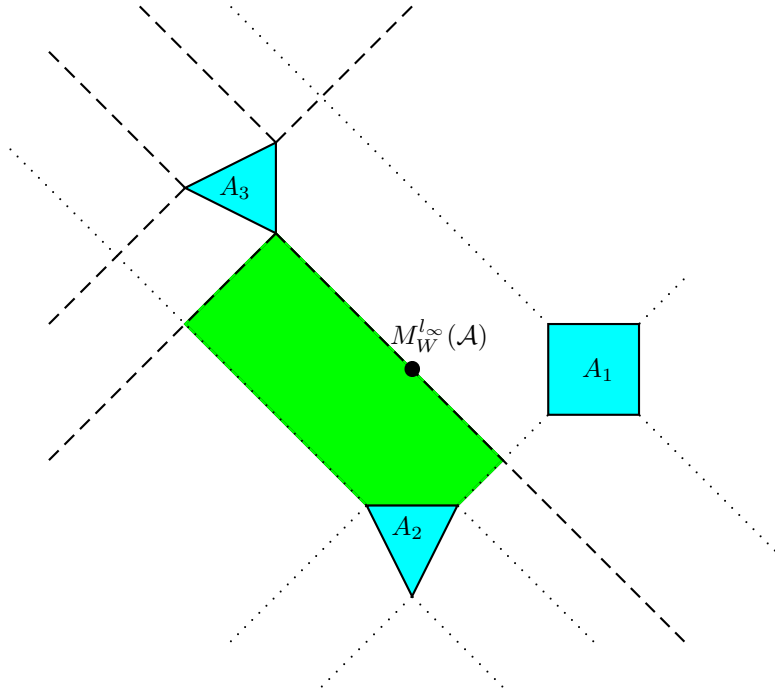


Figura 5: Ilustración del ejemplo 3.3

4 El problema de mediana ordenada

Esta sección está dedicada al estudio de otro problema genérico dentro de la Teoría de Localización. El modelo corresponde a una formulación relativamente reciente de una familia de problemas [31, 35] cuya principal característica es permitir, en cierto modo, un tratamiento algebraico unificado.

Por simplicidad en la exposición, se considera el modelo puntual en espacios de dimensión finita. El lector interesado puede consultar [35, 37] para un tratamiento más general.

Sean $A = \{a_1, \dots, a_M\}$ un conjunto de puntos de demanda y $W = \{w_1, \dots, w_M\}$ y $\Lambda = \{\lambda_1, \dots, \lambda_M\}$ dos conjuntos de escalares no negativos tales que el elemento w_i de W representa la importancia del punto de demanda a_i y los elementos de Λ diseñan la función objetivo. Para facilitar la comprensión de este modelo, se usará un sólo calibrador para medir las distancias, pudiéndose extender todos los resultados obtenidos al caso de calibradores diferentes.

Dada una permutación σ del conjunto $\{1, \dots, M\}$ verificando

$$w_{\sigma_1} \gamma(x - a_{\sigma_1}) \leq w_{\sigma_2} \gamma(x - a_{\sigma_2}) \leq \dots \leq w_{\sigma_M} \gamma(x - a_{\sigma_M}),$$

se denota por

$$d_{(i)}(x) = w_{\sigma_i} \gamma(x - a_{\sigma_i}),$$

el elemento i -ésimo de la lista ordenada de distancias ponderadas.

Como indicamos en la sección 1, la formulación del problema de mediana ordenada es la siguiente:

$$\min_{x \in \mathbb{R}^n} F(x) = \sum_{i=1}^M \lambda_i d_{(i)}(x), \quad (10)$$

donde $\lambda_1 \leq \dots \leq \lambda_M$.

Obsérvese que esta función objetivo, dependiendo del punto evaluado, puede tener diferentes expresiones como suma ponderada de las distancias. En otras palabras, dado que σ depende de x , su forma funcional está definida puntualmente (hay en cierto modo dependencia “no local”). Esto significa que la manipulación de F va a tener una dificultad añadida que se ve compensada por su flexibilidad, esto es, por la gran cantidad de modelos que se pueden considerar como casos particulares. Este hecho queda reflejado en la observación siguiente.

Nota 4.1

- i) El problema de Fermat-Weber es un problema de mediana ordenada donde $\lambda_1 = \dots = \lambda_M = 1$.
- ii) El problema cent-dian es un problema de mediana ordenada donde $\lambda_1 = \dots = \lambda_{M-1} = \lambda$ y $\lambda_M = 1 - \lambda$.
- iii) El problema minimax es un problema de mediana ordenada donde $\lambda_1 = \dots = \lambda_{M-1} = 0$ y $\lambda_M = 1$.

- iv) El problema del k -centrum es el problema de mediana ordenada con $\lambda_1 = \dots = \lambda_{M-(k+1)} = 0$ y $\lambda_{M-k} = \dots = \lambda_M = 1$.
- v) El problema *lexicográfico* es el problema de mediana ordenada con $\lambda_M \gg \lambda_{M-1} \gg \dots \gg \lambda_1$ ($a \gg b$ significa que a es “mucho mayor” que b).

Por tanto, a través del estudio del problema de mediana ordenada se puede:

- a) Caracterizar el conjunto de soluciones de varios modelos simultáneamente.
- b) Desarrollar un algoritmo eficiente único que resuelva todos estos problemas. \square

En la subsección que sigue, se presenta el caso en que las distancias son medidas a través de calibradores poliédricos. Abordaremos a continuación el caso general a través de un método de aproximación basado en el caso poliédrico.

4.1 El problema de mediana ordenada con calibradores poliédricos

En primer lugar, queremos insistir en que, a pesar de que la función objetivo del problema (10) es similar a la del problema de Fermat-Weber, no existe una representación unificada en todo el espacio como suma de distancias. En este apartado, se caracterizan los conjuntos donde la representación como suma de distancias permanece invariante.

Definición 4.1 Sea $i \neq j$. Al conjunto $B_\gamma(a_i, a_j)$, formado por los puntos $\{x : w_j \gamma(x - a_j) = w_i \gamma(x - a_i)\}$, se le llama *mediatriz* de a_i y a_j con respecto a γ .

Como ilustración de la definición 4.1, se puede ver en la figura 6 la mediatriz para los puntos $a_1 = (1, 4)$ y $a_2 = (5, 1)$ con norma l_1 y pesos $w_1 = w_2 = 1$.

Seguidamente se analizan algunas propiedades que permiten describir geoméricamente las mediatrices.

Proposición 4.1 La mediatriz de a_i y a_j es un conjunto de puntos que verifican una ecuación lineal dentro de cada convexo elemental.

Demostración:

En un convexo elemental, $\gamma(x - a_i)$ y $\gamma(x - a_j)$ pueden ser escritos como $l_i(x - a_i)$ y $l_j(x - a_j)$ respectivamente, donde $l_i(\cdot)$ y $l_j(\cdot)$ son funciones lineales. De esta forma, decir que $w_i \gamma(x - a_i) = w_j \gamma(x - a_j)$ es equivalente a decir que $w_i l_i(x - a_i) = w_j l_j(x - a_j)$, de donde se sigue el resultado. \square

En el caso bidimensional esta descripción puede ser más precisa.

Proposición 4.2 La mediatriz de a_i y a_j con respecto a un calibrador poliédrico con G puntos extremos contiene a lo más $O(G)$ subconjuntos diferentes definidos por ecuaciones lineales.

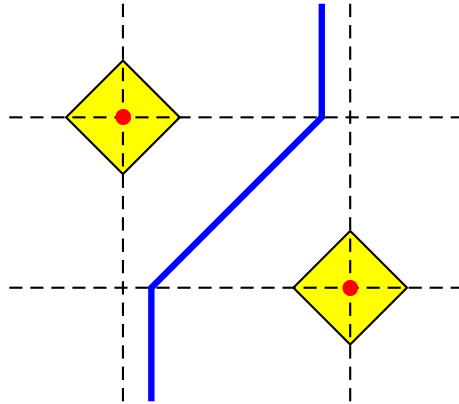


Figura 6: Mediatriz

La demostración de este resultado puede verse en [37].

Este resultado implica que el número de expresiones lineales que definen cualquier mediatriz es también lineal en G (el número de direcciones fundamentales). Obsérvese que no hay ninguna razón por la cual la mediatriz debe tener interior vacío. Ver por ejemplo la figura 7, donde se muestra la mediatriz definida por los puntos $(0,0)$ y $(4,0)$ con la norma de Tchebychev.

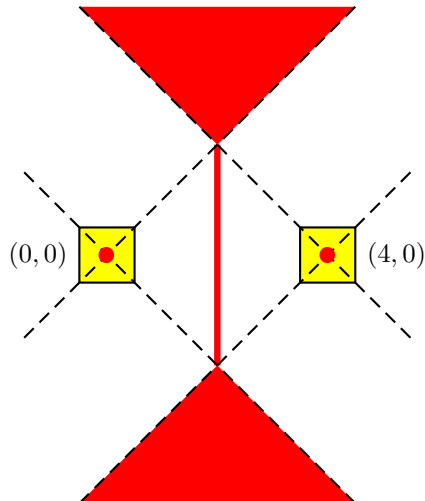


Figura 7: Mediatriz

Cuando el orden de la secuencia ordenada de distancias no varía, la expresión de la función objetivo del problema de mediana ordenada como suma de distancias ponderadas no cambia. Con el objetivo de describir esas regiones

donde la expresión de la función objetivo es invariante, se da la siguiente definición.

Definición 4.2 *Dada una permutación σ del conjunto $\{1, 2, \dots, M\}$, la región ordenada O_σ es el conjunto*

$$O_\sigma = \{x \in \mathbb{R}^2 : w_{\sigma_1} \gamma(x - a_{\sigma_1}) \leq \dots \leq w_{\sigma_M} \gamma(x - a_{\sigma_M})\}.$$

Nota 4.2 Nótese que estas regiones no son necesariamente convexas, véase la figura 8. □

La importancia de las regiones ordenadas radica en que el problema de mediana ordenada dentro de cada región se reduce a un problema de Fermat-Weber.

Definición 4.3 *Un convexo elemental ordenado es cualquier conjunto resultante de la intersección de un convexo elemental y una región ordenada. Los vértices de los convexos elementales ordenados se denominan puntos de intersección ordenados.*

Nota 4.3 Nótese que estos conjuntos son convexas, ya que las mediatrices tienen un comportamiento lineal en cada convexo elemental. Por tanto, los convexos elementales ordenados son el resultado de aplicar restricciones lineales a un conjunto convexo. □

La importancia de los convexos elementales ordenados viene dada por el comportamiento lineal de la función objetivo dentro de cada uno de ellos. Ello se debe a que resultan de la intersección de una región ordenada (donde el comportamiento del problema de mediana ordenada es igual al de un problema de Fermat-Weber) y un convexo elemental (donde el comportamiento del problema de Fermat-Weber es lineal).

Para concluir con la descripción de las regiones ordenadas, se puede observar que dichas regiones están definidas a lo más por $M - 1$ mediatrices.

Sin embargo, la principal desventaja de tratar con estas regiones es su complejidad. Un análisis simplista, podría conducir a concluir que el número de regiones ordenadas es $M!$, lo que haría que el problema fuera intratable. Afortunadamente, cuando la dimensión del espacio es fija, se puede obtener una cota polinomial del número de regiones ordenadas. En el caso más simple de \mathbb{R}^2 , el resultado es el siguiente:

Teorema 4.1 *Una cota superior sobre el número de regiones ordenadas es $O(M^4 G^2)$.*

Demostración:

Dadas dos mediatrices con $O(G)$ tramos lineales, el máximo número de intersecciones es $O(G^2)$. El número total de mediatrices entre M puntos es $\binom{M}{2}$, entonces, el máximo número de intersecciones entre ellas es

$$O(G^2 \binom{M}{2}).$$

Puesto que el número de intersecciones tiene la misma complejidad que el número de regiones, una cota superior para el número de regiones ordenadas es $O(M^4 G^2)$. \square

Un análisis detallado de este teorema muestra que esta cota no es demasiado elevada, a pesar de tener orden $O(M^4 G^2)$. Se observa que el número de mediatrices entre los puntos de A es $\binom{M}{2}$, que es de orden $O(M^2)$. Por tanto, incluso en el caso más favorable en que las mediatrices son líneas rectas, el número de regiones en el análisis del peor caso es de orden $O(\binom{M}{2}^2)$, es decir, de orden $O(M^4)$. Debido a que cada mediatriz podría tener G tramos lineales, esta cota es bastante ajustada.

Ejemplo 4.1 La figura 8 muestra las regiones ordenadas determinadas por los puntos $a_1 = (0, 11)$, $a_2 = (3, 0)$ y $a_3 = (16, 8)$ con la norma hexagonal $\gamma(\cdot)$, cuyo conjunto de puntos extremos es $Ext(B) = \{(2, 0), (1, 2), (-1, 2), (-2, 0), (-1, -2), (1, -2)\}$ y con pesos $w_1 = w_2 = w_3 = 1$. Por ejemplo, la región $O_{(3,2,1)}$ es el conjunto de puntos

$$\{x \in \mathbb{R}^2 : \gamma(x - a_3) \leq \gamma(x - a_2) \leq \gamma(x - a_1)\}.$$

Veremos a continuación que, a pesar de que no existe una representación global única de la función objetivo, ésta es convexa. Usando las herramientas del Análisis Convexo junto con las descripciones geométricas introducidas anteriormente, podremos así desarrollar un algoritmo eficiente para resolver el problema.

Lema 4.1 *Se tiene que*

$$F(x) = \sum_{i=1}^M \lambda_i d_{(i)}(x) = \max_{\sigma \in P(M)} \sum_{i=1}^M \lambda_i w_{\sigma_i} \gamma(x - a_{\sigma_i}) \quad \forall x \in \mathbb{R}^n,$$

siendo $P(M)$ el conjunto de permutaciones de $\{1, \dots, M\}$.

La demostración de este resultado es una consecuencia del teorema 386 de [20].

Nótese que el lema precedente implica en particular que F es convexa, por ser máximo de funciones convexas.

A continuación se muestra un problema similar al de mediana ordenada, pero con pesos no ordenados, para el que la función objetivo no es convexa. En este problema, el conjunto solución no es convexo.

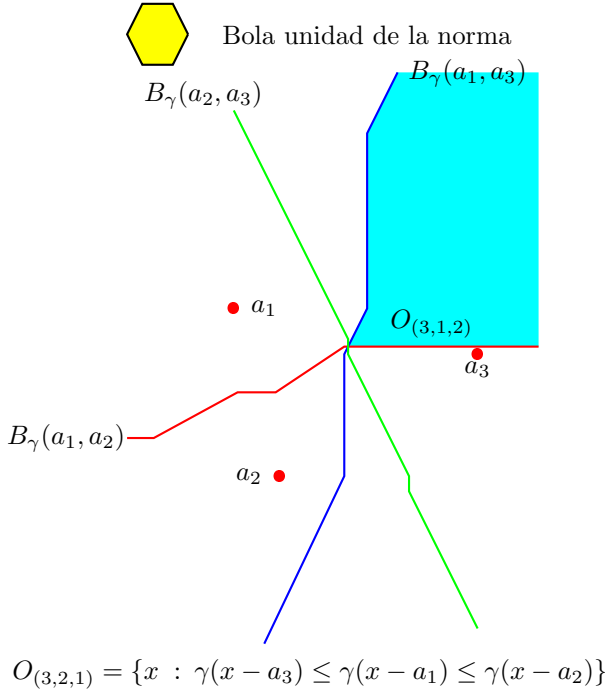


Figura 8: Regiones Ordenadas

Ejemplo 4.2 Sea $A := \{a_1 = (3, 0), a_2 = (16, 8)\}$ y sea $\gamma(\cdot)$ la norma hexagonal definida en el ejemplo 4.1. La función

$$G(x) = 2d_{(1)}(x) + d_{(2)}(x)$$

no es una función convexa porque, en la región ordenada $O_{(1,2)}$,

$$G(x) = 2\gamma(x - (3, 0)) + \gamma(x - (16, 8))$$

y, en $O_{(2,1)}$,

$$G(x) = 2\gamma(x - (16, 8)) + \gamma(x - (3, 0)).$$

Por tanto, usando el Teorema de la “Mayoría” se obtiene que las soluciones de estos problemas son $(3, 0)$ y $(16, 8)$ respectivamente. Es decir, el conjunto solución no es convexo, lo que implica que G no es convexa. En la figura 9, se presenta una ilustración gráfica de este ejemplo.

En [31, 35] se prueba que el conjunto solución del problema (10) coincide con un convexo elemental ordenado. Sin embargo, la existencia de un gran número de convexos elementales ordenados requiere un procedimiento de enumeración a partir del cual se pueda deducir un algoritmo que nos permita la resolución de dicho problema.

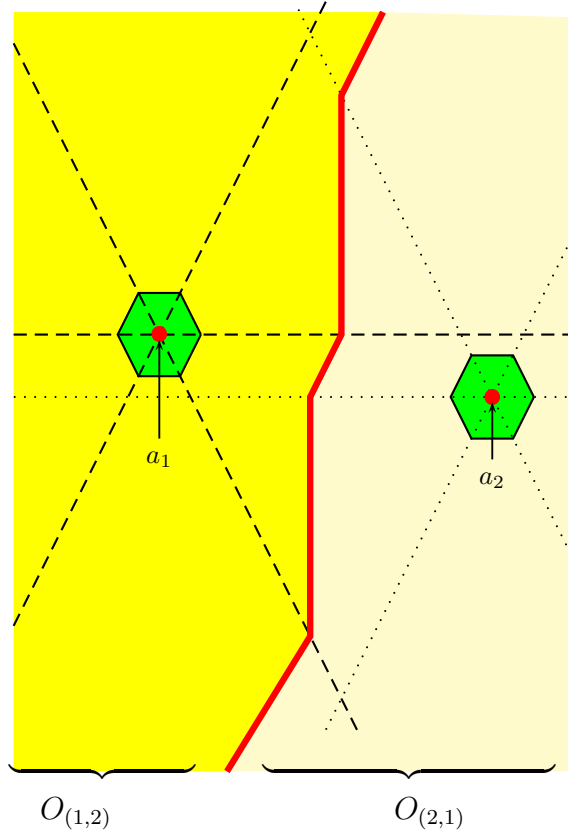


Figura 9: No convexidad de la función objetivo.

Esto es posible en el caso de normas poliédricas. En efecto, dentro de cada región ordenada, O_σ , se considera el siguiente problema de programación lineal:

$$\begin{aligned} \min & \quad \sum_{i=1}^M \lambda_i w_{\sigma_i} z_{\sigma_i} \\ \text{sujeto a} & \quad w_i \langle b_g^0, x - a_i \rangle \leq z_i \quad b_g^0 \in B^0, \quad 1 \leq i \leq M, \\ & \quad z_{\sigma_i} \leq z_{\sigma_{i+1}} \quad 1 \leq i \leq M - 1, \end{aligned} \quad (P_\sigma)$$

donde las b_g^0 son las direcciones fundamentales de B^0 .

Lema 4.2 Sea x^* una solución de (P_σ) .

- i) Si $x^* \in O_\sigma$, entonces x^* es también una solución del problema de la mediana ordenada restringido a O_σ .
- ii) Si $x^* \in O_{\sigma'}$ donde $O_{\sigma'} \neq O_\sigma$, entonces la solución del problema de la mediana ordenada restringido a $O_{\sigma'}$ es mejor (i.e. el valor, en ella de la función objetivo es menor) que la solución del problema de la mediana ordenada restringido a O_σ .

Demostración:

Si $x^* \in O_\sigma$ se tiene que

$$w_i \langle b_{g_i}^o, x - a_i \rangle = z_i, \quad \text{para algún } g_i = 1, \dots, G \text{ y para todo } i = 1, 2, \dots, M$$

lo que significa que $z_i = w_i \gamma(x - a_i)$ para todo $i = 1, 2, \dots, M$, de donde se obtiene el resultado.

Si $x^* \in O_{\sigma'}$ ($O_{\sigma'} \neq O_\sigma$), se tiene que

$$w_i \langle b_{g_i}^o, x - a_i \rangle < z_i \quad \text{para todo } g_i = 1, \dots, G \text{ y para al menos un } i.$$

Esto significa que la función objetivo del problema (P_σ) en x^* es menor que el valor objetivo de ($P_{\sigma'}$) en x^* . \square

A continuación, usando los lemas previos y el hecho de que la función objetivo es globalmente convexa, se desarrollará un algoritmo que resuelve el problema de la mediana ordenada.

Para cada región ordenada, O_σ , se resolverá el problema de programación lineal (P_σ), pudiendo darse dos casos: (1) la solución está en el interior de la región considerada (en \mathbb{R}^n), lo que por convexidad implica que esta solución es el óptimo global; (2) la solución está sobre la frontera o en otra región ordenada; entonces se tiene que hacer una nueva búsqueda local en alguna región vecina a la que pertenece este punto.

ALGORITMO 4.1

PASO 1: ELEGIR una región ordenada O_σ .

PASO 2: RESOLVER el problema (P_σ) y elegir una solución y^ .*

PASO 3:

- *SI $y^* \notin O_\sigma$, ENTONCES $\sigma := \sigma'$ donde $y^* \in O_{\sigma'}$ e IR AL PASO 2.*
- *SI $y^* \in \text{int}(O_\sigma)$, ENTONCES $x^* := y^*$ e IR AL PASO 4.*
- *SI y^* pertenece a la frontera de O_σ , ELEGIR una región adyacente no visitada $O_{\sigma'}$ tal que $y^* \in O_{\sigma'}$ e ir al PASO 2 con $\sigma := \sigma'$. SI todas las regiones adyacentes han sido visitadas, ENTONCES $x^* := y^*$ e IR AL PASO 4.*

PASO 4: SALIDA x^ .*

El algoritmo anterior es eficiente en el sentido de que está polinomialmente acotado. Una vez que se ha fijado la dimensión del problema, su complejidad está dominada por la complejidad de resolver un problema lineal en cada región ordenada. El número de regiones ordenadas está polinomialmente acotado y la resolución de un problema de programación lineal se puede hacer en tiempo polinomial mediante algoritmos de punto interior, ver [42]. Por tanto, el Algoritmo 4.1 es polinomial.

Otros algoritmos más eficientes que hacen uso de técnicas propias de Geometría Computacional pueden hallarse en [23].

Ejemplo 4.3 Para $A = \{(3, 0), (0, 11), (16, 8)\}$, $w_i = 1$, $i = 1, \dots, M$, $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 3$, el problema ordenado tiene la siguiente fomulación:

$$\min_{x \in \mathbb{R}^2} \left(d_{(1)}(x) + 2d_{(2)}(x) + 3d_{(3)}(x) \right),$$

donde $\gamma_B(\cdot)$ es la norma hexagonal definida en el Ejemplo 4.1.

La figura 10 muestra los convexos elementales ordenados para este problema. Nótese que las líneas gruesas representan las mediatrices entre los puntos de A , mientras las líneas delgadas son las direcciones fundamentales de la norma sobre cada punto de demanda. Se resuelve el problema usando el Algoritmo 4.1. Comenzando con el punto $x^0 = (0, 11)$ se obtiene la solución en dos iteraciones. En la primera, se obtiene el punto $x^1 = (6'5, 8)$ con valor objetivo 26'25. En la segunda iteración, se obtiene $x^2 = (7, 8)$ con valor objetivo 26. Este punto no puede ser mejorado en su entorno, con lo cual es la solución del problema. Las iteraciones del algoritmo para este ejemplo se muestran en la figura 11.

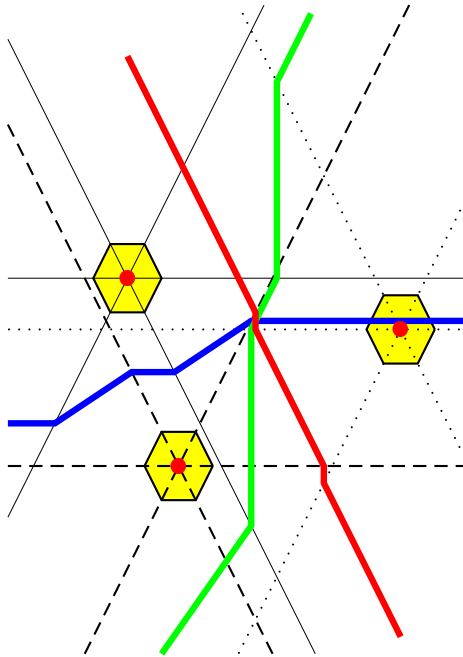


Figura 10: Convexos elementales generalizados

4.2 Caso no poliédrico

En los apartados anteriores se ha desarrollado una metodología para resolver el problema de la mediana ordenada en el caso de calibradores poliédricos. En esta sección, se resuelve el problema con un calibrador general, no necesariamente poliédrico. Para ello se prueba que las soluciones de estos problemas pueden

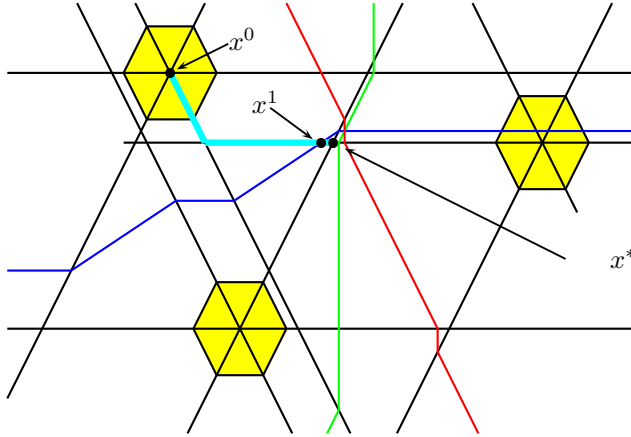


Figura 11: Solución óptima

ser arbitrariamente aproximadas por una sucesión de soluciones de problemas con calibradores poliédricos cuyas bolas unidad convergen, en la métrica de Hausdorff, ver [39], a la bola unidad del problema no poliédrico.

Sea B la bola unidad del calibrador $\gamma_B(\cdot)$, $\{B_m\}$ una sucesión creciente de polígonos convexos incluidos en B y $\{B^m\}$ una sucesión decreciente de polígonos convexos que incluyen a B :

$$B_m \subset B_{m+1} \subset B \subset B^{m+1} \subset B^m.$$

Sean $\gamma_{B_m}(\cdot)$ y $\gamma_{B^m}(\cdot)$ los calibradores cuyas bolas unidad son B_m y B^m , respectivamente.

Proposición 4.3 *Si $B_m \subset B \subset B^m$ se tiene que*

$$\gamma_{B_m}(x - y) \geq \gamma_B(x - y) \geq \gamma_{B^m}(x - y) \quad \forall x, y \in \mathbb{R}^n.$$

La demostración se sigue directamente de la definición de calibrador.

Proposición 4.4 *Sea K un conjunto compacto. Si B_m converge a B y B^m converge a B en el sentido de la métrica de Hausdorff, entonces para todo $\varepsilon > 0$ existe n_0 tal que, para todo $m \geq n_0$,*

$$\begin{aligned} \max_{x \in K} |F_m(x) - F_B(x)| &< \varepsilon, \\ \max_{x \in K} |F^m(x) - F_B(x)| &< \varepsilon, \end{aligned}$$

siendo $F_m(x) := \sum_{i=1}^M \lambda_i d_{(i)}^{B_m}(x)$, $F^m(x) := \sum_{i=1}^M \lambda_i d_{(i)}^{B^m}(x)$ y $F_B(x) := \sum_{i=1}^M \lambda_i d_{(i)}(x)$.

Demostración:

Sólo se probará la primera desigualdad. La segunda se demuestra de forma análoga.

Puesto que B_m converge a B , $B_m \subset B_{m+1}$ para todo m y K es un conjunto compacto, dados $\varepsilon > 0$ y $a \in A$, existe n_a tal que, si $m > n_A := \max_{a \in A} n_a$, se verifica que

$$\left| \gamma_B(x - a) - \gamma_{B_m}(x - a) \right| < \frac{\varepsilon}{\sum_{i=1}^M w_i \sum_{i=1}^M \lambda_i} \quad \forall x \in K.$$

Por continuidad se tiene que, cualesquiera i y j , y cualquier $x \in K$ verificando que $w_i \gamma_B(x - a_i) < w_j \gamma_B(x - a_j)$, existe $n_o > n_A$ tal que

$$w_i \gamma_{B_m}(x - a_i) < w_j \gamma_{B_m}(x - a_j) \quad m > n_o$$

y tal que, para cualesquiera i, j y $x \in K$ con $w_i \gamma_B(x - a_i) = w_j \gamma_B(x - a_j)$, se tiene $\gamma_B(x - A)_{(i)} = w_{\sigma_i^m} \gamma_B(x - a_{\sigma_i^m})$, donde $w_{\sigma_i^m} \gamma_{B_m}(x - a_{\sigma_i^m}) = d_{(i)}^{B_m}(x)$ para cada $m > n_o$. De aquí se deduce que, para cualquier $x \in K$ y cualquier $m > n_o$,

$$\begin{aligned} |F_m(x) - F(x)| &= \sum_{i=1}^M \lambda_i |d_{(i)}(x) - d_{(i)}^{B_m}(x)| \\ &= \sum_{i=1}^M \lambda_i w_{\sigma_i^m} |\gamma_B(x - a_{\sigma_i^m}) - \gamma_{B_m}(x - a_{\sigma_i^m})| < \varepsilon. \end{aligned}$$

□

Corolario 4.1

- i) Si B_m converge a B en el sentido de la métrica de Hausdorff, entonces F_m converge a F_B uniformemente sobre compactos. Además, la sucesión $\{F_m(x)\}$ es decreciente.
- ii) Si B^m converge a B en el sentido de la métrica de Hausdorff, entonces F^m converge a F_B uniformemente sobre compactos. Además, la sucesión $\{F^m(x)\}$ es creciente.

Este resultado será utilizado para demostrar el siguiente teorema.

Teorema 4.2

- i) Sea $\{x_m\}$ una sucesión tal que $x_m \in \arg \min F_m$ para cada $m \geq 1$. Entonces cualquier punto de acumulación de $\{x_m\}$ pertenece a $\arg \min F_B$.
- ii) Sea $\{x^m\}_{m \in \mathbb{N}}$ una sucesión tal que $x^m \in \arg \min F^m(x)$, para cada $m \geq 1$. Entonces cualquier punto de acumulación de $\{x^m\}_{m \in \mathbb{N}}$ pertenece a $\arg \min F_B$.

Demostración:

Sólo se probará la primera parte, puesto que la demostración de la segunda es análoga.

En primer lugar, debido a que la sucesión $\{F_m\}$ es decreciente, se puede aplicar el teorema 2.46 en [1], obteniéndose que $\{F_m\}$ es epi-convergente.

Además se obtiene por la proposición 2.4.8 en [1] que

$$\lim_{m \rightarrow \infty} \inf_{x \in \mathbb{R}^2} F_m(x) = \inf_{x \in \mathbb{R}^2} \lim_{m \rightarrow \infty} F_m(x) = \inf_{x \in \mathbb{R}^2} F_B(x). \quad (11)$$

Finalmente, por el teorema 2.12 de [1], deducimos que cualquier punto de acumulación de la sucesión $\{x_m\}$ es una solución de F_B . \square

El resultado anterior prueba que la solución del problema ordenado de Weber con un calibrador general puede ser arbitrariamente aproximada por las soluciones del mismo problema con calibradores poliédricos.

5 El problema multiobjetivo

En cualquier problema de localización donde el criterio para localizar un servicio sea atractivo, todos los usuarios quieren tener simultáneamente el servicio tan cercano como sea posible. Por ello, una característica intrínseca de los problemas de localización es su naturaleza multiobjetivo.

El propósito de los problemas multiobjetivo es el de optimizar simultáneamente varias funciones objetivo. Sin embargo, debido a que rara vez existe un punto simultáneamente óptimo en todas las funciones objetivo, tratamos de caracterizar el conjunto de soluciones no dominadas o, equivalentemente, el conjunto de puntos para los que no existe otro mejor en cada función objetivo considerada en el problema. En Teoría de Localización, la caracterización del conjunto de soluciones no dominadas en problemas multiobjetivo ha sido ampliamente estudiada, aunque esta cuestión ha sido sólo parcialmente resuelta. Una formulación general del problema multiobjetivo es la que viene dada por la siguiente expresión:

$$v - \min(F_1(x), \dots, F_k(x)),$$

donde $v - \min$ representa la minimización vectorial y las F_1, \dots, F_k son las funciones objetivo.

Al problema donde cada función F_i es la distancia al punto fijo a_i , medida con la misma norma, se le llama problema de localización punto-objetivo. Este problema fue tratado por primera vez en [47] para la norma l_2 . Posteriormente, Durier en [11] resolvió este problema para calibradores poliédricos. Otras referencias dedicadas a estudiar diferentes versiones del problema de localización punto-objetivo son [3, 5, 7, 15, 19, 25, 29, 44, 48].

El caso en que cada F_i es una suma ponderada de distancias medidas con la norma l_1 (el problema de Fermat-Weber con norma rectangular) fue resuelto en

[18, 33]. El problema en que las F_i son sumas ponderadas de distancias medidas con normas arbitrarias fue resuelto en [34].

Por otro lado, el caso donde existe una demanda regional y las funciones F_i para $i = 1, \dots, k$ son funciones inf-distancias (es decir, la distancia al punto más próximo de cada conjunto de demanda) fue resuelto en [40].

De esta forma, aunque el problema ha sido estudiado en numerosas ocasiones y por diferentes autores, las caracterizaciones obtenidas se aplican sólo en casos particulares. En esta sección, se proporciona una solución geométrica para esta clase de problemas, donde la única condición que se impone a las funciones objetivo consideradas es que sean convexas e inf-compactas. A pesar de la generalidad del problema considerado, los resultados son fáciles de comprender y en las demostraciones se usan, casi exclusivamente, herramientas del Análisis Convexo.

5.1 El modelo

Sean F_1, \dots, F_k funciones reales convexas e inf-compactas definidas por simplicidad en \mathbb{R}^2 que reperesentan diferentes criterios o escenarios. Recuérdese que una función f es inf-compacta si los conjuntos de nivel inferior $\{x : f(x) \leq \rho\}$ son compactos para cada $\rho \in \mathbb{R}$. Nuestro objetivo es encontrar el conjunto de puntos $x \in \mathbb{R}^2$ tales que no existe $y \in \mathbb{R}^2$ que mejore el valor de F_i en x para cada $i \in \{1, \dots, k\}$. De esta forma, la formulación del problema viene dada por:

$$v - \min_{x \in \mathbb{R}^2} (F_1(x), \dots, F_k(x)). \quad (12)$$

El conjunto solución que se va a considerar para este problema viene dado por:

$$\text{WE}(F_1, \dots, F_k) := \{x \in \mathbb{R}^2 : \forall y (\neq x) \in \mathbb{R}^2, \exists i, 1 \leq i \leq k, \text{ con } F_i(x) \leq F_i(y)\}.$$

Usualmente este conjunto es llamado conjunto de *puntos débilmente eficientes*.

Con el objetivo de facilitar el entendimiento del resto de la sección, se usa la siguiente notación. El conjunto de nivel inferior de la función F para un valor $\rho \in \mathbb{R}$ es

$$L_{\leq}(F, \rho) := \{x \in \mathbb{R}^2 : F(x) \leq \rho\}$$

y el conjunto de nivel estrictamente inferior

$$L_{<}(F, \rho) := \{x \in \mathbb{R}^2 : F(x) < \rho\}.$$

Para una función convexa e inf-compacta F_i , se usará la notación

$$\mathcal{X}^*(F_i) := \arg \min_{x \in \mathbb{R}^2} F_i(x). \quad (13)$$

Nótese que este conjunto es convexo por la convexidad de F_i .

En esta sección se estudia el problema (12), probándose que la estructura geométrica de $\text{WE}(F_1, F_2, F_3)$ viene dada por la región delimitada por las cadenas de soluciones “biobjetivo” asociadas a cada par de funciones F_i

y F_j , con $1 \leq i, j \leq 3$. Este resultado permite caracterizar el conjunto $WE(F_1, \dots, F_k)$ mediante la unión de los conjuntos solución de cada problema con tres objetivos. A continuación se enuncia una caracterización teórica del conjunto $WE(F_1, \dots, F_k)$ que será utilizada a lo largo de esta sección.

Teorema 5.1 *Se verifica que:*

$$\hat{x} \in WE(F_1, \dots, F_k) \iff \bigcap_{i=1}^k L_{<}(F_i, F_i(\hat{x})) = \emptyset.$$

La demostración de este resultado puede verse en [38].

Nota 5.1 En el caso $k = 2$, el resultado previo establece que $WE(F_1, F_2)$ coincide con el lugar geométrico de los puntos de tangencia de los conjuntos de nivel inferior de las funciones F_1 y F_2 unión con $\mathcal{X}^*(F_1) \cup \mathcal{X}^*(F_2)$. Para comprender bien esta afirmación, véase el ejemplo 5.1. \square

Ejemplo 5.1 (ver figura 12) Sean $a_1 = (0, 0)$, $a_2 = (10, 0)$, $a_3 = (14, 6)$, $F_1(x) = \|x - a_1\|_2$, $F_2(x) = \|x - a_2\|_1$ y $F_3(x) = \|x - a_3\|_1$. Por el teorema 5.1, $WE(F_1, F_2)$ es el segmento definido por a_1 y a_2 . En efecto, se trata del conjunto de puntos \hat{x} tales que no hay ningún otro punto que esté más cerca que él de a_1 en norma $\|\cdot\|_2$ y más cerca que él de a_2 en norma $\|\cdot\|_1$. Por otra parte, $WE(F_2, F_3)$ es el rectángulo sombreado con vértices opuestos a_2 y a_3 .

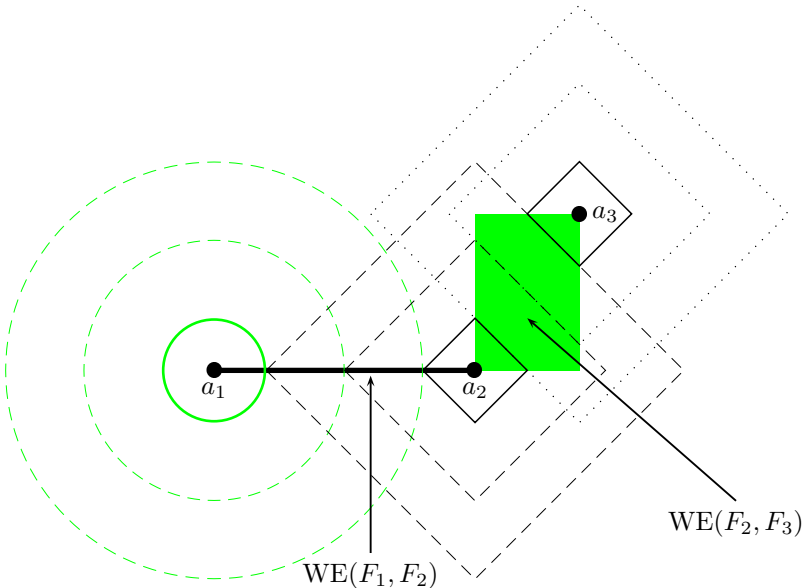


Figura 12: Ilustración del ejemplo 5.1.

Usando las caracterizaciones precedentes, se obtiene una descripción geométrica del conjunto $\text{WE}(F_1, F_2, F_3)$. El próximo resultado muestra que el conjunto solución del problema 3-criterio es una especie de envolvente definida por el conjunto de soluciones biobjetivo.

Sea

$$\text{WE}(2) := \bigcup_{\substack{i,j \in \{1,2,3\} \\ i \neq j}} \text{WE}(F_i, F_j)$$

(la unión de todas las cadenas bicriterio para los tres criterios F_1 , F_2 y F_3 considerados).

Teorema 5.2

$$\text{WE}(F_1, F_2, F_3) = \text{encl}(\text{WE}(2))$$

donde $\text{encl}(\text{WE}(2))$ es la región acotada rodeada por $\text{WE}(2)$ unión con $\text{WE}(2)$.

El lector interesado puede encontrar la demostración en [38].

Nota 5.2 Nótese que $\text{encl}(\text{WE}(2))$ está bien definida porque $\text{WE}(2)$ es conexo (ver [43]). Además, esta región puede ser equivalentemente definida como el conjunto de puntos tales que si $x \in \text{encl}(\text{WE}(2)) \setminus \text{WE}(2)$ no existe una curva continua $c(t)$, $t \in [0, \infty)$ con $c(0) = x$ y $\lim_{t \rightarrow \infty} \|c(t)\| = +\infty$, tal que $c(t) \notin \text{WE}(2)$, $\forall t \in [0, \infty)$. \square

Como ilustración de los resultados anteriores se muestra el siguiente ejemplo.

Ejemplo 5.2 Sean $a_1 = (0, 0)$, $a_2 = (3, -1)$, $a_3 = (3, 3)$ y las funciones F_1 , F_2 y F_3 tales que,

$$\begin{aligned} L_{\leq}(F_1, 1) &= \left\{ (x_1, x_2) : \frac{x_1^2}{4} + \frac{x_2^2}{9} \leq 1 \right\} \\ L_{\leq}(F_2, 1) &= \left\{ (x_1, x_2) : (x_1 - 3)^2 + (x_2 + 1)^2 \leq 1 \right\} \\ L_{\leq}(F_3, 1) &= \left\{ (x_1, x_2) : \frac{(x_1 - 3)^2}{9} + \frac{(x_2 - 3)^2}{4} \leq 1 \right\}. \end{aligned}$$

Es inmediato que estas tres funciones son convexas e inf-compactas. Por consiguiente, aplicando los resultados previos se obtiene que el conjunto $\text{WE}(F_1, F_2, F_3)$ es la región delimitada por las curvas que unen a_1 con a_3 , a_1 con a_2 y el segmento que une a_2 con a_3 en la figura 13.

Los resultados anteriores nos permiten demostrar el que sigue, que proporciona una caracterización geométrica del problema general con k objetivos.

Teorema 5.3

$$\text{WE}(F_1, \dots, F_k) = \bigcup_{i,j,\ell \in \{1, \dots, k\}} \text{WE}(F_i, F_j, F_\ell)$$

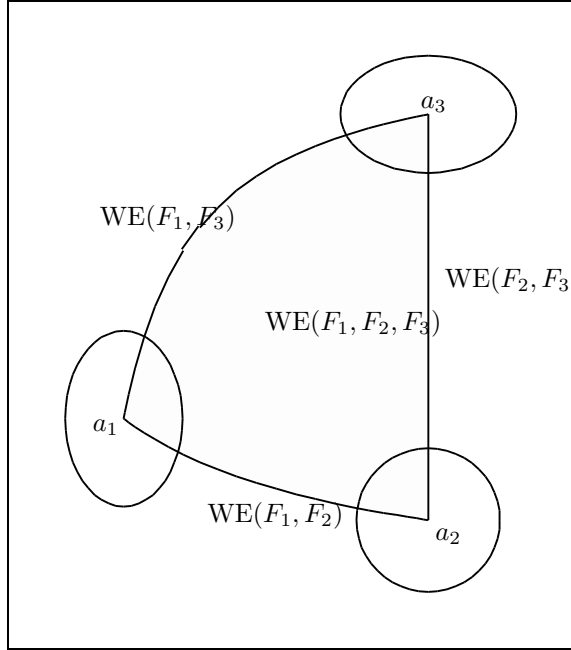


Figura 13: Ilustración del ejemplo 5.2.

Demostración:

Por el teorema 5.1, $\hat{x} \in WE(F_1, \dots, F_k)$ si y sólo si $\bigcap_{1 \leq i \leq k} L_{<}(F_i, F_i(\hat{x})) = \emptyset$. Aplicando el teorema de Helly (ver [39]), esta intersección es vacía si y sólo si existen $i, j, \ell \in \{1, \dots, k\}$ tales que $L_{<}(F_i, F_i(\hat{x})) \cap L_{<}(F_j, F_j(\hat{x})) \cap L_{<}(F_\ell, F_\ell(\hat{x})) = \emptyset$ y esto es equivalente a $\hat{x} \in WE(F_i, F_j, F_\ell)$. Esto prueba que

$$WE(F_1, \dots, F_k) \subseteq \bigcup_{i, j, \ell \in \{1, \dots, k\}} WE(F_i, F_j, F_\ell).$$

Puesto que en cualquier caso se tiene que

$$\bigcup_{i, j, \ell \in \{1, \dots, k\}} WE(F_i, F_j, F_\ell) \subseteq WE(F_1, \dots, F_k),$$

el resultado es cierto. □

Ejemplo 5.3 (Ver figura 14) Sean $a_1 = (7, 5)$, $a_2 = (18, 2'5)$, $a_3 = (22, 6'5)$, $F_1(x) = \|x - a_1\|_\infty$, $F_2(x) = \|x - a_2\|_1$ y $F_3(x) = \|x - a_3\|_1$. Por el teorema 5.1, $WE(F_1, F_2)$ es el camino en negrita que une a_1 y a_2 , $WE(F_1, F_3)$ es el camino en negrita uniendo a_1 y a_3 , y $WE(F_2, F_3)$ es el cuadrado sombreado con a_2 y a_3 como extremos opuestos. De esta forma, por el teorema 5.2, $WE(F_1, F_2, F_3)$

es la región sombreada que encierra la unión de los tres conjuntos definidos previamente. Nótese que este conjunto es unión de convexos elementales de dimensión completa.

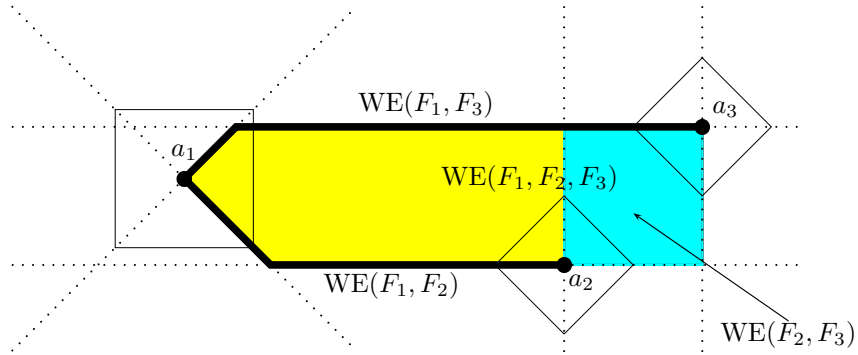


Figura 14: Ilustración del ejemplo 5.3.

6 Conclusión

En este trabajo se han revisado problemas de localización continua de tres tipos, cuya generalidad incluye como casos particulares un gran número de modelos clásicos considerados en la literatura. En cada caso se ha presentado el modelo matemático y se han estudiado propiedades que permiten caracterizar su conjunto solución. En la presentación se ha utilizado una amplia bibliografía que permitirá al lector interesado profundizar en el estudio de la Teoría de Localización.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos del Ministerio de Ciencia y Tecnología BFM2001-2378 y BFM2001-4028.

Referencias

- [1] ATTOUCH H. *Variational Convergence for Functions and Operators*, Pitman Advanced Publishing Program, 1984.
- [2] BAUER F.L., STOER J., AND WITZGALL C. Absolute and monotonic norms, *Numer. Math.*, 3:257–264, 1961.

- [3] CARRIZOSA E., CONDE E., FERNÁNDEZ F.R. AND PUERTO J. Efficiency in euclidean constrained location problems, *Oper. Res. Letters*, 14: 291–295, 1993.
- [4] E. CARRIZOSA, E. CONDE, F.R. FERNÁNDEZ, AND J. PUERTO. An axiomatic approach to the cent-dian criterion, *Location Science*, 2:165–171, 1994.
- [5] CARRIZOSA E., CONDE E., FERNÁNDEZ F.R. AND PUERTO J. Efficiency in spatial location problems with non convex constraint, *J. Global Optimization*, 6:77–86, 1995.
- [6] CARRIZOSA E. AND PUERTO J. A discretizing algorithm for location problems, *European J. Oper. Research*, 80:166–174, 1995.
- [7] CARRIZOSA E. AND PLASTRIA F. A characterization of efficient points in constrained location with regional demand, *Oper. Res. Letters*, 19, 3:129–134, 1996.
- [8] CROUZEIX J.P. AND KEBBOUR R. On the convexity of some simple functions of ordered samples, *Belgian Journal of Operations Research, Statistics and Computer Science (JORBEL)*, 36, 1:11–25, 1996.
- [9] DREZNER Z. *Facility Location. A Survey of Applications and Methods*, Springer Series in Operations Research, 1995.
- [10] DREZNER Z. AND HAMACHER H.W. *Facility Location. Applications and Theory*. Springer–Verlag, Berlin 2002.
- [11] DURIER R. On pareto optima, the Fermat-Weber problem and polyhedral gauges, *Math. Programming*, 47:65–79, 1990.
- [12] DURIER R. AND MICHELOT C. Geometrical properties of Fermat-Weber problem, *European J. Oper. Research*, 20:332–343, 1985.
- [13] DURIER R. A general framework for the one center location problem, in “*Advances in Optimization, Lecture Notes in Economics and Mathematical Systems*”, 382:441–457, Springer Verlag, Heidelberg 1992.
- [14] DURIER R. The general one center location problem, *Math. Oper. Research*, 20:400–414, 1995.
- [15] DURIER R. AND MICHELOT C. Set of efficient points in a normed space, *J. Math. Anal. and Appl.*, 117:506–528, 1986.
- [16] FRANCIS R.L., MCGINNIS L.F. AND WHITE, J.A. *Facility Layout and Location: An Analytical Approach*, Prentice Hall, Englewood Cliffs, New York 1992.
- [17] HALPERN J. The location of a centdian convex combination on a undirected tree, *J. Regional Science*, 16:237–245, 1976.

- [18] HAMACHER H.W. AND NICKEL S. Multicriteria planar location problems, *European J. Oper. Research*, 94:66–86, 1996.
- [19] HANSEN P., PERREUR J. AND THISSE J.F. Location theory, dominance and convexity: some further results, *Oper. Research*, 28, 5:1241–1250, 1980.
- [20] HARDY G. H., LITTLEWOOD J. E. AND PÓLYA, G. *Inequalities*, Cambridge University Press, Cambridge 1952.
- [21] HIRIART-URRUTY J.B. AND LEMARÉCHAL C. *Convex Analysis and Minimization Algorithms I*. Springer-Verlag, Berlin 1993.
- [22] IOFFE A.D. AND TIHOMIROV V.M. *Theory of Extremal Problems*, North-Holland, Amsterdam-New York 1979.
- [23] KALCSICS J., NICKEL S., PUERTO J. AND TAMIR, A. Algorithmic results for ordered median problems, *Oper. Res. Letters*, 30, 3:149–158, 2002.
- [24] LOVE R.F., MORRIS J.G. AND WESOLOWSKY G.O. *Facilities Location: Models and Methods*, North Holland, New York 1988.
- [25] NDIAYE M. AND MICHELOT C. Efficiency in constrained continuous location. *European J. Oper. Research*, 104, 2:288–287, 1998.
- [26] NICKEL S., PUERTO J. AND RODRÍGUEZ-CHÍA A.M. An approach to location models involving sets as existing facilities. *Mathematics of Oper. Research*, 28, 4:693–715.
- [27] NICKEL S., PUERTO J. AND RODRÍGUEZ-CHÍA A.M. *MCDM Location Problems*. Kluwer Academic Publishers, aparecerá.
- [28] NICKEL S., PUERTO J., RODRÍGUEZ-CHÍA A.M., AND WEISSLER A. *Multicriteria ordered Weber problems*, Technical Report, University of Kaiserslautern, Report in Wirtschaftsmathematik No. 53, 1999
- [29] PELEGRÍN R. AND FERNÁNDEZ F.R. Determination of efficient points in multiple-objective location problems. *Naval Research Logistics*, 35:697–705, 1988.
- [30] PLASTRIA F. Localization in single facility location. *European J. Oper. Research*, 18:215–219, 1984.
- [31] PUERTO J. AND FERNÁNDEZ F.R. The symmetrical single facility location problem, *preprint*. Universidad de Sevilla, 1995.
- [32] PUERTO J. *Lecturas en Teoría de Localización*. Secretariado de Publicaciones, Universidad de Sevilla, 1996.
- [33] PUERTO J. AND FERNÁNDEZ F.R. A convergent approximation scheme for efficient sets of the multicriteria Weber location problem, *TOP*, 6, 2:195–204, 1998.

- [34] PUERTO J. AND FERNÁNDEZ F.R. Multicriteria Weber location problems, *J. Multicriteria Decision Anal.*, 18:268–280, 1999.
- [35] PUERTO J. AND FERNÁNDEZ F.R. Geometrical properties of the symmetrical single facility location problem, *J. Nonlinear Convex Anal.*, 3:321–342, 2000.
- [36] REVELLE C., COHON J.L. AND SHOBRY D. Multiple objectives in facility location: A review, *Lecture Notes in Economics and Math. Systems*, 190:320–337, Springer-Verlag, 1981.
- [37] RODRÍGUEZ-CHÍA A.M., NICKEL S., PUERTO J., AND FERNÁNDEZ F.R. A flexible approach to location problems. *Math. Methods Oper. Research*, 51:69–89, 2000.
- [38] RODRÍGUEZ-CHÍA A.M. AND PUERTO J. Geometrical Description of the Weakly Efficient Solution Set for Multicriteria Location Problem. *Annals of Operations Research*, 111:179–194, 2002.
- [39] ROCKAFELLAR R.T. *Convex Analysis*, Princeton University Press, Princeton, 1970.
- [40] RODRÍGUEZ-CHÍA A.M. Avances del problema de localización continua de un único centro. *Tesis, Universidad de Sevilla 1998*.
- [41] THISSE J.-F. WARD J.E. AND WENDELL R.E. Some properties of location problems with block and round norms, *Oper. Research*, 32:1309–1327, 1984.
- [42] VANDERBEI R.J. *Linear Programming: Foundations and Extensions*, Kluwer Academic Publisher, Second Edition, Boston 2001.
- [43] WARBURTON A.R. Quasiconcave vector maximization: Connectedness of the sets of pareto-optimal and weak Pareto-optimal alternatives, *J. Optim. Theory Appl.*, 40, 4:537–557, 1983.
- [44] WARD, J.E. Structure of efficient sets for convex objectives, *Math. Oper. Research*, 2, 14:249–257, 1989.
- [45] WARD J.E. AND WENDELL R.E. Using block norms for location modelling, *Oper. Research*, 33:1074–1090, 1985.
- [46] WEBER A. “Über den Standort der Industrien”. *Tübingen. Theory of the location of industries*, University of Chicago Press, Chicago 1909.
- [47] WENDELL R.E. AND HURTER A.P. Location theory, dominance and convexity, *Oper. Research*, 21:314–321, 1973.
- [48] WENDELL R.E., HURTER A.P. AND LOWE T.J. Efficient points in location problems. *AIIE Transactions*, 9:238–246, 1977.
- [49] WITZGALL, C. On convex metrics. *Journal of Research of the National Bureau of Standards*, 69:175–177, 1965.

Consideraciones metodológicas sobre la enseñanza y el aprendizaje de la Teoría de la Computabilidad

A. J. PÉREZ

Departamento de Ciencias de la Computación e I.A.
Universidad de Sevilla

pjimenez@us.es

Resumen

En este artículo nos centramos en ciertas tendencias metodológicas con las que pueden abordarse la enseñanza y aprendizaje de un primer curso sobre *teoría de la computabilidad* en las Escuelas de Ingeniería Informática. Tras justificarse el método de presentación matemático como un método adecuado para la exposición de dicha teoría, abordamos los aspectos pedagógicos relacionados con la metodología escolar, decantándonos por métodos activos, como el basado en el razonamiento plausible de George Polya, en oposición a los métodos tradicionales de exposición y ejercicios.

1 Introducción

En su informe¹ de 1991 la ACM² y la IEEE-CS³ afirman que la Informática puede considerarse como una disciplina que conjuga tres componentes: una matemática, otra científica y una tercera técnica. Por eso, el desarrollo de las diferentes áreas emplea, básicamente, tres *metodologías* de trabajo o *procesos* en investigación, desarrollo y aplicaciones: **teoría**, **modelización** y **diseño**.

El primer proceso, la **teoría**, se utiliza para el desarrollo de teorías matemáticas que se usan en Ciencias de la Computación y consta de cuatro fases: *Definiciones y axiomas, teoremas, demostraciones e interpretación de resultados*.

El segundo proceso, **modelización**, se base en el método científico experimental y se divide en las siguientes fases: *Recogida de datos y elaboración*

¹**Computing Curricula 1991**. Report of the ACM/IEEE-CS Joint Curriculum Task Force. December 17, 1990. (Apartado 5.2). <http://www.computer.org/education/cc1991/>

²Association for Computing Machinery

³Institute of Electrical and Electronics Engineers-Computer Society

de hipótesis, construcción de modelos y realización de predicciones, diseño de experimentos y análisis de resultados.

El tercer proceso, **diseño**, tiene su base en la ingeniería y se usa en el desarrollo de un sistema para resolver un problema. Consta también de cuatro fases: *Estado de necesidades, estado de especificaciones, diseño e implementación del sistema, chequeo y análisis del sistema.*

2 Objetivos generales

La Informática tiene consideración de disciplina científica y técnica a la vez y, como tal, a la luz de los elementos señalados en los **CC1991** para cada uno de los procesos que la componen, puede distinguirse los siguientes objetivos generales a conseguir en la formación de los estudiantes:

- En relación con la componente **teórica**:
 - Caracterizar los objetos de estudio (definiciones, axiomas).
 - Promover la formulación de conjeturas plausibles (relaciones entre objetos de la teoría).
 - Ver con claridad que las conjeturas plausibles deben ser justificadas (teoremas).
 - Estimular el pensamiento crítico y el rigor científico.
 - Fomentar la capacidad de análisis y síntesis.
 - Impulsar la aptitud para resolver problemas.
 - Interpretar los resultados obtenidos en un marco teórico.
- En relación con la componente de **modelización**:
 - Recoger datos, organizándolos para su estudio.
 - Aprender a distinguir lo esencial de lo accesorio en una situación genérica.
 - Elaborar hipótesis a la luz de los datos de que se dispone.
 - Potenciar la búsqueda de modelos científicos asociados a situaciones concretas y hacer predicciones.
 - Estar en disposición para diseñar un experimento.
 - Analizar los resultados.
 - Potenciar la búsqueda de modelos científicos asociados a situaciones concretas.
- En relación con la componente de **diseño**:
 - Detectar las necesidades reales en un contexto concretos.

- Saber diseñar experiencias y ser capaz estudiar experiencias ya diseñadas.
- Elaborar métodos de análisis capaz de detectar las cualidades del sistema diseñado.
- Desarrollar los mecanismos que permiten aplicar los conceptos y resultados teóricos a situaciones prácticas.

Las asignaturas que versan sobre teoría de la computabilidad e impartidas en las titulaciones de Ingeniería Informática, pretenden precisamente *ser soporte de la formación teórica* de los alumnos, con lo que los objetivos a conseguir son en esencia, desde nuestro punto de vista, los enunciados anteriormente dentro del apartado de **teoría**.

3 Consideraciones metodológicas

Por un lado, los **CC1991** señala a las Matemáticas como uno de los componentes básicos de las Ciencias de la Computación. Por otro, en los **CC2001**⁴ se afirma que un estudiante de Informática necesita adquirir una serie de herramientas que arrojen su formación técnica, haciéndose notar además que es necesario que los estudiantes adquieran “ un cierto nivel *de sofisticación en matemáticas*, familiaridad con los métodos científicos y un cierto sentido de cómo la computación es aplicada en la práctica y habilidad para trabajar en equipo”. Pone el acento este documento en la necesidad de un rigor matemático en la formación que incluya el manejo de técnicas matemáticas y razonamientos formales matemáticos pues estos elementos, se afirma, están detrás de muchas áreas de la Informática.

No nos cabe duda, y las opiniones anteriores lo refuerzan, de que las asignaturas sobre teoría de la computabilidad se caracterizan precisamente por ser de naturaleza matemática y/o lógica, lo que justifica que la metodología que defendamos para las exposiciones sea precisamente *una metodología de presentación de las matemáticas*.

La metodología que ordinariamente se sigue en las clases tienen como base el *método deductivo* (o lógico-deductivo o axiomático-deductivo). Dicho método se traduce principalmente por los siguientes elementos: construcción de un corpus teórico a partir de un sistema de *definiciones, axiomas o postulados* desde los que se llega, mediante demostraciones (*proceso de deducción* que da nombre al método), a las *propiedades y teoremas* de la teoría.

Pero en la metodología escolar de cualquier materia, aparte del método de presentación científico elegido, está la vertiente pedagógica que condiciona la manera en que esa presentación es traducida en una clase concreta. Pues bien, esta componente pedagógica suele consistir en el *método tradicional expositivo* o *magistral*: el profesor explica la materia, ilustra con ejemplos y propone ejercicios que, en la clase siguiente, suele resolver el propio profesor.

⁴**Computing Curricula 2001**. Informe de la ACM/IEEE-CS - Diciembre, 2001. (Apartado 9.1) <http://www.computer.org/education/cc2001/>

Aunque el método deductivo de presentación de las matemáticas pueda ser desarrollado en clase con otros elementos pedagógicos distintos a los señalados, digamos ya que no es usual; sea por razones históricas, sea porque es “lo natural”, dicho método suele ir unido al expositivo tradicional. En adelante, al hablar de método deductivo lo consideraremos íntimamente unido en lo que concierne al desarrollo en clase, al método expositivo tradicional. Utilizaremos la nomenclatura de *metodología deductiva* o *deductiva-magistral*.

• **Metodología deductiva. Ventajas e inconvenientes.**

El método deductivo, como método para la presentación matemática, se remonta a Euclides. Sus *Elementos* son el primer paradigma de aplicación de tal método. Pero es a raíz de la crisis de fundamentación de las matemáticas, a principios del siglo XX, cuando se considera el método axiomático deductivo como una herramienta esencial en el proceso de construcción matemática y en el de la enseñanza de la misma. Este proceso que llega a su cumbre con las tesis bourbakistas, es revalidado por una corriente que se inicia a finales de los años 50, que pone el énfasis en una enseñanza axiomática y deductiva de las matemáticas.

El profesor Jean Dieudonné, muy preocupado por la enseñanza y que, por su gran personalidad, puede ser considerado como uno de los principales líderes de las llamadas *matemáticas modernas*, afirma en el prólogo de su libro *Álgebra lineal y geometría elemental* que

Es necesario que el alumno -de secundaria- tome conciencia de la necesidad de un tratamiento axiomático de las matemáticas.

Más adelante se muestra en contra de determinados tratamientos intuitivos:

Es deseable liberar al alumno cuanto antes de la camisa de fuerza de las “figuras” tradicionales(...) Asimismo conviene ciertamente enseñar a los alumnos el arte de las construcciones geométricas, pero hay que huir como de la peste de lo que es, sin duda, el mayor fastidio de la enseñanza clásica: la limitación de los instrumentos de dibujo a la regla y el compás.

Destacaré tres razones que suelen aducirse en defensa del método magistral-deductivo.

La primera es de tipo conservador y se basa en que las pautas de estudio de los alumnos están ya marcadas por este método; cambiarlo no debe hacerse en un contexto aislado y deberían reunirse determinadas condiciones (generalidad, número no elevado de alumnos, actitud, convencimiento y preparación amplia del profesor) para cambiar a una metodología más activa.

La segunda razón radica en la eficacia. Cualquier metodología activa, cualquier método heurístico es lento. Si ya de por sí los programas son bastante ajustados al tiempo que le corresponde, la adecuación a un cuatrimestre hace que el profesor tenga menos “margen de maniobra” para diseñar determinadas estrategias de actividades.

La tercera razón, y más poderosa, es la masificación en clase. Cuando el número de alumnos es muy elevado, y este es el caso actualmente en los estudios de ingeniería informática, no parece haber otra manera de proceder más que con una exposición de tipo magistral, siendo muy difícil articular cualquier forma en la que el alumno tenga un papel más participativo (trabajo en grupo, realización y exposición de trabajos, etc). Ni siquiera en las clases de problemas suele estar prevista una disminución del número de alumnos por grupo. Esta disminución sólo se contempla en las clases prácticas de Laboratorio, cuando las hay.

En estas circunstancias parecería que la metodología escolar obligada es la deductiva-magistral, el método tradicional de siempre.

Esta metodología presenta sin dudas ciertas ventajas tales como exposición clara y simple; economía de medios (en los símbolos y en el pensamiento); fácil adecuación del ritmo al programa de la asignatura; impecable si se mira desde el rigor científico, y meridianamente clara desde el punto de vista de la transmisión de conocimientos. Pero a pesar de estas ventajas y a pesar de ser un método consolidado al lo largo de muchos años, posee una serie de graves inconvenientes que hemos de señalar:

- Es una metodología impositiva y, por ello, no motivadora.
- Fomenta la actitud pasiva del alumno al convertirlo en mero receptor de las explicaciones del profesor.
- Es lineal en la presentación teórica, en aras del rigor, sin matizaciones acerca de lo que es más o menos pertinente, sin consideraciones sobre aspectos culturales.
- No tiene en cuenta la historia ni la génesis de los problemas.
- Apenas considera la intuición del alumnado, al centrarse casi de manera exclusiva en desarrollos formales.

Pensamos que en unos estudios de ingeniería, con una fuerte componente de experimentación y diseño, parece conveniente unas metodologías más intuitivas, más dinámicas y más participativas, incluso -o quizás más aún por ello- a la hora de abordar asignaturas con una fuerte componente teórica.

• Metodologías inductivas. El razonamiento plausible

En el siglo XX prolifera una rica variedad de tendencias en la enseñanza en general y, en particular, en la enseñanza de la matemáticas. Muchos matemáticos se muestran partidarios de una enseñanza más intuitiva, contrarios, en más de un caso, a una enseñanza puramente deductiva. Citemos a algunos. Poincaré, por ejemplo, se pregunta⁵:

⁵H. Poincaré, *La Ciencia y la Hipótesis*; Espasa Calpe, Madrid, 1963.; pág. 15

¿Cuál es la naturaleza del razonamiento matemático? ¿Es realmente deductivo como ordinariamente se cree? Un análisis profundo nos muestra que no es así; que participa en una cierta medida de la naturaleza del razonamiento inductivo, y que por eso es fecundo. Pese a ello no pierde su carácter de rigor absoluto.

En otra ocasión, es contundente⁶:

¿Es posible entender una teoría si desde el primer momento se le da la forma de definitiva que impone una lógica rigurosa, sin mencionar para nada el camino por el que ha llegado a adoptar esta forma? No, realmente no es posible entenderla; incluso resulta imposible retenerla si no es de memoria.

Poincaré acaba planteando una ponderación entre la lógica y la intuición al afirmar⁷:

La lógica y la intuición tienen cada una un papel necesario; ambas son indispensables. La lógica, que puede por sí misma dar la certeza, es el instrumento de la demostración; la intuición es el instrumento de la invención.

Courant y Robbins en el conocido libro *¿Qué es la Matemática?* se rebelan contra el método axiomático deductivo en los siguientes términos:

Parece existir un grave peligro en el excesivo predominio del carácter axiomático deductivo de las matemáticas(...) Si la forma deductiva cristalizada es la meta, la intuición y la construcción son, cuando menos, las fuerzas directrices. Una amenaza sería para la vida de las ciencias aparece contenida en la afirmación de que la matemática no es más que un sistema de conclusiones derivadas de definiciones, reglas y silogismos, sin meta ni motivos algunos

Y en un célebre manifiesto⁸ de 1962, firmado en EE.UU. por más de setenta profesores⁹ y publicado en las revistas *American Mathematical Monthly* y *The Mathematics Teacher*, se afirma:

El pensamiento matemático no es tan sólo razonamiento deductivo; no consiste simplemente en demostraciones formales. (...) sin alguna experiencia en procesos "informales" de pensamiento, el estudiante no puede comprender el verdadero papel de la demostración formal y rigurosa (...) el objeto del rigor es "confirmar y legitimar las conquistas de la intuición.." [Hadamard] (...) Hay varios niveles de rigor. El estudiante debería aprender a comprender, buscar y criticar las demostraciones al nivel correspondiente a su experiencia y formación.

⁶H. Poincaré, *El valor de la Ciencia*; Espasa Calpe, Madrid, 1963.; pág. 15

⁷H. Poincaré, *El valor de la Ciencia*; Espasa Calpe, Madrid, 1963.; pág. 15

⁸*On the mathematics curriculum for the High School* [10]

⁹entre otros: Courant, Coxeter, Kline, Polya, Stoker, Weil

Una personalidad como Felix Klein, que encabeza, a comienzos del s. XX un proceso de reforma para la enseñanza de la matemáticas (en secundaria principalmente) y que en 1908 crea, en el seno de la Unión Matemática Internacional (IMU), la Comisión Internacional para la Instrucción Matemática (ICMI), afirma que enseñar científicamente no consiste en enfrentar a los alumnos con "fríos sistemas científicamente pulidos" sino inducirles a pensar científicamente.

Desde otros ámbitos (pedagogía, psicología) se propugnan alternativas metodológicas, sobre todo en la enseñanza secundaria. Se considera que frente a la mera actitud de enseñar, hay que *fortalecer el aprendizaje* pues el objetivo es formar al alumno a quien, por tanto, hay que ceder el protagonismo de la enseñanza. Son las *metodologías activas*.

Una corriente extrema en este sentido es la corriente *neoconductista* que encabeza el psicólogo B.F. Skinner¹⁰ quien considera que el alumno puede, con la ayuda de una máquina, desarrollar gran parte de su aprendizaje. Su método de *enseñanza programada* consiste, en esencia, en dar los conocimientos al alumno a través de pequeñas dosis de información (items) de tal manera que cuando el alumno yerra vuelve hacia atrás con preguntas alternativas que pretenden reconducirle a la solución correcta del item no asimilado.

Otras metodologías menos radicales se abren paso; son las metodologías *activas*, basadas en la psicológica *genética* de J. Piaget. Estas metodologías afirman que la esencia de la enseñanza está en la *acción* sobre objetos. Por acción se entiende desde la manipulación y observación de objetos reales hasta la interacción con objetos mentales. En los años cincuenta y gracias a la acción de un grupo de trabajo internacional e interdisciplinar, la CIEAM¹¹, formado entre otros por pedagogos como el profesor inglés Caleb Gattegno; psicólogos como el suizo Jean Piaget, maestros como la italiana Emma Castelnuovo; y matemáticos como el francés Jean Dieudonné o el español Pedro Puig Adam, impulsan un movimiento en torno a la enseñanza activa de la matemática conocido como "enseñanza dinámica". Estos métodos que pretenden dar más protagonismo al aprendizaje frente al mero hecho de enseñar, van abriéndose camino con pujanza.

Dentro de estas corrientes de *enseñanza activa*, habría que situar la tendencia encabezada por George Polya conocida como *Resolución de problemas*. Importa mucho señalarla, además, porque va dirigida principalmente al mundo universitario. Frente a la metodología deductiva, Polya propone una metodología inductiva y heurística que, en su obra *Cómo resolverlo it*¹² plasma en cuatro fases:

1. Comprender el problema.
2. Concebir un plan.
3. Ejecutar el plan.

¹⁰, *Tecnología de la enseñanza*, Ed. Labor, Madrid, 1970

¹¹Comisión Internacional por l'Enseignement et l'Amelioration de la Mathematique

¹²*Comment Poser et Résoudre un Problème*. Ed. Dunod. París, 1965

4. Examinar la solución obtenida.

Descompone, a su vez, cada fase en preguntas y sugerencias que son, en realidad, estrategias encaminadas a recorrer un camino que conduzca a la solución del problema. Así, se plantea:

1. ¿Cuáles son las incógnitas?
2.
 - He aquí un problema parecido y resuelto; ¿podría utilizarlo?
 - ¿Podría enunciarlo de otra forma?
 - ¿Conoce un problema análogo?
 - Particularice.
 - Generalice.
 - Divida el problema en partes.
3. Compruebe todos los pasos; ¿ve claramente que todos son correctos?
4. ¿Puede verificar el resultado? ¿Puede obtenerlo de forma distinta?

Fundamenta esta metodología inductiva a través de su teoría del *razonamiento plausible*, que justifica en su obra “Matemáticas y razonamiento plausible”. Según Polya¹³, “nuestro conocimiento matemático se asegura mediante el *razonamiento demostrativo*, mientras que las conjeturas se afianzan mediante el *razonamiento plausible*.... En el razonamiento estricto lo principal es distinguir una prueba de una intuición, una demostración válida de un intento sin validez. En el razonamiento plausible lo importante es distinguir entre intuiciones, unas más y otras menos razonables”.

Polya considera que en la enseñanza debe tenerse en cuenta este método inductivo como elemento que genera un proceso creador. Y aboga por que, incluso en la formación del estudiante matemático, entre también esta metodología del “razonamiento plausible”:

I. Lakatos¹⁴, alumno y seguidor de las teorías de Polya, afirma:

[las matemáticas] no se desarrollan mediante el aumento del número de teoremas indubitavelmente establecidos sino mediante la incesante mejora de las conjeturas, gracias a la especulación y a la crítica, siguiendo la lógica de las pruebas y refutaciones

En España, el profesor P. Puig Adam -profesor de metodología en la Universidad de Madrid en los años 40 y 50- creó escuela con el método heurístico. Puig Adam afirmaba que es¹⁵ “absurdo el método lógico, hipotético-deductivo

¹³G. Polya *Matemáticas y Razonamiento Plausible*. Editorial Tecnos, Madrid, 1966, págs. 13-14

¹⁴Su obra, *Pruebas y refutaciones*- Alianza Editorial, Madrid, 1978- retoma un problema, el de la definición de polígono, allá donde la evidencia plausible de Polya lo ha dejado

¹⁵P. Puig Adam *La Matemática y su enseñanza actual*. Publicaciones de la Revista de Enseñanza Media. Madrid, 1960

(...) igual desde Euclides(...) olvidando el largo proceso previo de edificación de categorías mentales” y, en su conferencia de 1953, *El Movimiento Didáctico Renovador* afirmaba que “el movimiento heurístico tiende a acercar el proceso de génesis de los conocimientos y el de transmisión de los mismos”.

Pero cuando gran parte de la comunidad educativa, principalmente en los niveles primario y secundario, estaba sensibilizándose a favor de los métodos activos, en la comunidad matemática se produce una rebelión contra esa corriente considerada “pedagoga”. Lidera esa rebelión, entre otros, por el profesor J. Dieudonné, quien en un Congreso celebrado en Royaumont, en 1959, lanzó su célebre proclama “Abajo Euclides”, que se considera como la señal, el punto de arranque, de una corriente pedagógica conocida como *matemáticas modernas*. Esta corriente, si bien supuso una renovación muy importante en los contenidos de los programas de matemáticas, fue un considerable paso atrás en la renovación metodológica de esta disciplina.

El manifiesto antes citado fue una reacción de más de setenta matemáticos contra esta irrupción de la matemática moderna:

Los matemáticos, reaccionando frente al control de la educación por los educadores profesionales, quienes quizás han hecho hincapié en la pedagogía a expensas del contenido, pueden ahora acentuar el contenido a expensas de la pedagogía de forma igualmente estéril (...) La necesidad de hoy de aprender más matemáticas que en el pasado puede llevarnos a buscar atajos que, sin embargo, podrían ser más dañinos que beneficiosos

4 Consideración final

Por todo lo dicho queremos concluir haciendo hincapié en nuestra convicción de que las metodologías activas y, en particular, una metodología heurística como la de Polya basada en el razonamiento plausible, son más adecuadas para enseñanzas de disciplinas teóricas, fuertemente impregnadas por las matemáticas. Aparte de razones intrínsecas, queremos volver a poner el acento en que la formación de un informático tiene unas componentes de modelización y diseño en las que puede colaborar, sin lugar a dudas, una metodología heurística.

Creemos, finalmente, que la incorporación de los nuevos medios (tecnologías de la información y comunicación principalmente), la paulatina disminución del número de alumnos por profesor y grupo y determinadas actuaciones oficiales, como las tutorías, la participación institucional de los alumnos, etc., son elementos que deben aprovecharse para que conduzcan de manera natural a una implicación mucho más importante del alumno en su propio aprendizaje (y no sólo en los aspectos “políticos y de gestión” de su enseñanza).

Referencias

- [1] ACM/IEEE-CS, *Computing Curricula 2001*, ACM/IEEE-CS, 2001.

- [2] ACM/IEEE-CS, *Computing Curricula 1991*, ACM/IEEE-CS, 1990.
- [3] COURANT R. Y ROBBINS H., *¿Qué es la Matemática?*, Ed. Aguilar, Madrid, 1971 .
- [4] DIEUDONNÉ, J. *Álgebra Lineal y Geometría Elemental*, Secciones Científicas, Madrid, 1971.
- [5] KLEIN, F. *Enseñanza elemental desde un punto de vista superior*, Ed. Ibero-Americana; Buenos Aires, 1948.
- [6] KLINE, M. *El fracaso de la Matemática Moderna*, Ed. Siglo XXI, 1984.
- [7] KLINE, M., *La pérdida de la certidumbre*, Ed. Siglo XXI, 1985.
- [8] LAKATOS, I *Pruebas y refutaciones* Alianza Editorial, Madrid, 1978.
- [9] PÉREZ JIMÉNEZ, A. *Tendencias actuales en la enseñanza de las Matemáticas*. Actas IV Jornadas Andaluzas de Educación Matemática. SAEM Thales, Málaga, 1989.
- [10] PIAGET, J., CHOQUET, G., DIEUDONNÉ, J., THOM, R. Y OTROS, *La enseñanza de las matemáticas modernas*, Alianza Editorial, Madrid, 1980.
- [11] POINCARÉ, H., *La Ciencia y la Hipótesis*; Espasa Calpe, Madrid, 1963.
- [12] POINCARÉ, H., *El valor de la Ciencia*; Espasa-Calpe, Madrid, 1963.
- [13] POLYA, G., *Matemáticas y Razonamiento Plausible*, Editorial Tecnos, Madrid, 1966.
- [14] POLYA, G., *Comment Poser et Résoudre un Problème*. Ed. Dunod. París 1965.
- [15] PUIG ADAM, P., *La Matemática y su enseñanza actual*. Publicaciones de la Revista de Enseñanza Media. Madrid, 1960.
- [16] SKINNER, B.F., *Tecnología de la enseñanza*, Ed. Labor, Madrid, 1970.