

REVISTA DE LA ACADEMIA COLOMBIANA de Ciencias Exactas, Físicas y Naturales

LA ACADEMIA ES ÓRGANO CONSULTIVO DEL GOBIERNO NACIONAL

VOLUMEN XXVIII

MARZO DE 2004

NÚMERO 106

DIRECTOR DE LA REVISTA: SANTIAGO DÍAZ-PIEDRAHITA



Matemáticas y Estadística (Mathematics and Statistics)

Edición en homenaje a Jairo Charris Castañeda

CONTENIDO - CONTENTS

	Pág.		Pág.
- Semblanza del profesor Jairo Charris Castañeda [Portrait of Professor Jairo Charris Castañeda] <i>Jaime Lesmes Camacho</i>	5	- Mariposas y 3-variedades [Butterflies and 3-manifolds] <i>M. H. Hilden, J. M. Montesinos, D. M. Tejada & M. M. Toro</i>	71
- El teorema de Hahn-Banach como principio de elección [Hahn-Banach Theorem as a choice principle] <i>X. Caicedo & G. Enciso</i>	11	- Nudos combinatorios y mariposas [Combinatorial knots and butterflies] <i>Margarita María Toro Villegas</i>	79
- Aplicación de un principio de minimización [Application of a minimization principle] <i>Jorge Cossio & Germán Jiménez</i>	21	- Una lógica modal para la geometría plana de Lobachevski [A modal logic for Lobachevski's plane geometry] <i>Alfonso Rider Moyano & Rafael María Rubio Ruiz</i>	87
- Potencia de pruebas de rachas para alternativa de tendencia [Power of run tests for trend alternative] <i>E. Gómez Torres, Raydonal Ospina & Jimmy Corzo</i>	27	- p -adic oscillatory integrals and Newton polyhedra [Integrales oscilatorias p -ádicas y poliedros de Newton] <i>W. A. Zuniga-Galindo</i>	95
- Leverrier-Fadeev algorithm and classical orthogonal polynomials [Algoritmo de Leverrier-Fadeev y polinomios ortogonales clásicos] <i>J. Hernández, F. Marcellán & C. Rodríguez</i>	39	- Overview on models in homotopical algebra [Un panorama de los modelos de álgebra homotópica] <i>Roberto Ruiz S.</i>	101
- Precondicionamiento de métodos iterativos [Iterative methods preconditioning] <i>Guillermo Cervantes & Carlos Mejía S.</i>	49	- Geometría de variedades bandera [Geometry of flag manifolds] <i>Marlio Paredes & Sofía Pinzón</i>	123
- La notación diagramática de C. S. Peirce para los conectivos proposicionales binarios [C. S. Peirce's diagrammatic notation for the binary propositional connectives] <i>Arnold Oostra</i>	57	- Contribución al estudio de las ecuaciones diferenciales parciales de tipo elíptico [Contribution to the study of partial differential equations of elliptic type] <i>Jorge Cossio</i>	135
Constitución de la Academia (Members of the Academy)			147

SEMBLANZA DEL PROFESOR JAIRO CHARRIS CASTAÑEDA

por

Jaime Lesmes Camacho¹

Resumen

Lesmes Camacho, Jaime: Semblanza del profesor Jairo Charris. Rev. Acad. Colomb. Cienc. **28** (106): 5–10, 2004. ISSN 0370-3908.

Se da una semblanza del profesor Jairo A. Charris Castañeda, fallecido el 17 de julio de 2003.

Palabras clave: Jairo A. Charris Castañeda, biografía, bibliografía.

Abstract

A portrait of Professor Jairo A. Charris Castañeda who passed away on July the 17th, 2003 is given.

Key words: Jairo A. Charris Castañeda, Biography, Bibliography.

Ante todo, agradezco a los organizadores de este Festival Académico, en especial a la Academia Colombiana de Ciencias, el haberme confiado la honrosa y para mí, aunque mezclada con un sentimiento de tristeza, placentera tarea de hacer la semblanza del Profesor Jairo Charris, a quien me unió una gran amistad desde nuestra época de estudiantes, hace más de 40 años, y por quien profesé una profunda admiración, debida a su excelente calidad matemática, a su producción científica,

a su meritoria y fructífera labor de profesor y educador y a sus altísimas cualidades humanas.

Jairo Charris Castañeda nació en Ciénaga, departamento del Magdalena, el 21 de noviembre de 1939; fue el mayor de siete hermanos. Sus primeros estudios los realizó en el Instituto San Juan del Córdoba, de su ciudad natal. Se graduó de bachiller en el Colegio Mayor de Nuestra Señora del Rosario en Bogotá, en 1957, e

¹Departamento de Matemáticas, Universidad de los Andes, Bogotá. email: jlesmes@uniandes.edu.co

* Versión modificada de la conferencia pronunciada en el “Festival Académico Dedicado a la Memoria del Señor Profesor Don Jairo Charris Castañeda”, que se realizó en Bogotá el día 6 de agosto de 2003

AMS Classification 2000: 01A60, 01A90.

ingresó al año siguiente a la entonces Facultad de Ingeniería Química de la Universidad Nacional de Colombia.²

Durante sus estudios de Ingeniería Química, en los que está por demás decir que fue sobresaliente, siempre se interesó fuertemente por las matemáticas y la física, y hacia el final de ellos participó en cursos y seminarios de la entonces incipiente carrera de Matemáticas, a la cual ingresó formalmente en 1962. Obtuvo el grado de Matemático en la Universidad Nacional, en 1967, con la tesis titulada “*Sobre el completado de un espacio uniforme y el compactado de Stone-Čech*”, dirigida por Carlos Lemoine. El año siguiente viajó, como becario de la Fundación Ford, a la Universidad de Chicago, donde trabajó bajo la orientación del matemático hindú Raghavan Narasimhan y recibió el título de Master en Matemáticas en 1969, con un trabajo titulado: “*Sobre ciertos espacios de funciones y sus aplicaciones*”. Regresó a Colombia en 1971 y durante diez años realizó valiosa labor en la Universidad Nacional. En esa época contrajo matrimonio con María Victoria Castañeda y nacieron sus dos hijas, Marcela y Ana María.

En 1981 se trasladó con su familia a Tempe, para proseguir estudios de doctorado en la Arizona State University, donde obtuvo el Doctorado en Matemáticas en 1984, bajo la dirección del profesor Mourad Ismail. Su tesis versó sobre el tema de los polinomios de Pollaczek; los desarrollos posteriores de la teoría de los polinomios ortogonales han mostrado la calidad y profundidad de este trabajo. Posteriormente precisaré esta aseveración.

Jairo Charris había iniciado en 1962, como Instructor, su vinculación con el Departamento de Matemáticas de la Universidad Nacional, vinculación que continuó, con dedicación ejemplar, hasta su jubilación en 1998, siendo ya Profesor Titular desde varios años atrás. Luego de su retiro continuó colaborando con la Universidad Nacional y también con la Universidad Sergio Arboleda. Los campos de trabajo del profesor Charris eran: Análisis Complejo, Topología General y, principalmente, Funciones Especiales y Polinomios Ortogonales.

Durante su permanencia en la Universidad de Chicago, donde, como mencioné atrás, trabajó bajo la orientación de R. Narasimhan, adelantó Jairo Charris interesantes investigaciones en Análisis Complejo de una y varias variables, cuyos principales resultados fueron posteriormente publicados en dos artículos de la Revista

Colombiana de Matemáticas, en 1974 y 1979. En el primero de ellos [2]³ se introducen ciertos espacios de funciones holomorfas en polidiscos, semejantes a los espacios de Sobolev sobre conjuntos abiertos de \mathbb{R}^n , a fin de determinar si el espacio de soluciones holomorfas en un abierto conexo de \mathbb{C}^n de un tipo especial de ecuaciones diferenciales lineales homogéneas, es de dimensión finita. En el segundo artículo [4] se usan esos espacios (en dimensión uno) y técnicas homológicas de aproximación para clasificar como operadores de Fredholm a una amplia clase de operadores diferenciales complejos que actúan sobre un abierto conexo del plano, y establecer explícitamente una fórmula para el índice de tales operadores.

En la Universidad Nacional Charris dirigió seis tesis de maestría sobre temas de Análisis Complejo, las tres primeras de ellas en la década de los 70, y durante su vida profesional dictó frecuentemente cursos y ciclos de conferencias sobre temas de esa área. Recientemente, en el año 2000, la Academia Colombiana de Ciencias publicó su magistral tratado “*Fundamentos del Análisis Complejo de una Variable*”, escrito en colaboración con Rodrigo De Castro y Juan Varela [33].

Me referiré ahora al principal campo de trabajo de Jairo Charris.

Los polinomios ortogonales están relacionados con numerosas áreas de la Matemática y poseen importantes aplicaciones en Física y Probabilidad. Su estudio hunde sus raíces en los finales del siglo XVIII, ya que sus orígenes se remontan a los trabajos de Legendre de 1784 y 1787 sobre el movimiento planetario (ver [CoHi], pág. 71). En la segunda mitad del siglo XIX se estudiaron con gran detalle otros sistemas especiales de polinomios ortogonales, menciono los muy conocidos de Chebichev, de Jacobi, de Hermite y de Laguerre. En todos estos casos se trata de polinomios ortogonales con respecto a ciertas funciones de peso, sobre diversos intervalos. Estos polinomios cumplen relaciones de recurrencia, de las cuales puede partirse para su definición, son soluciones de ecuaciones diferenciales tipo Sturm-Liouville, pueden obtenerse como coeficientes de funciones generatrices y están relacionados con casos especiales de fracciones continuas (ver [CoHi]). Según G. Szegő [Sz], los orígenes de la teoría general de los polinomios ortogonales se encuentran en la investigación sistemática de un cierto tipo de fracciones continuas llevada a cabo

²Gran parte de los datos biográficos sobre Jairo Charris han sido tomados de la entrevista [Sa].

³Los números entre corchetes remiten a la lista de publicaciones de J. Charris, y las letras entre corchetes a la bibliografía de otros autores, al final del artículo.

principalmente por Stieltjes (hacia 1890), aunque este enfoque fue abandonándose gradualmente para tomar como propiedad básica la ortogonalidad misma.

Tras la publicación del clásico tratado “*Orthogonal Polynomials*” de Szegő [Sz] en 1939, el interés en los polinomios ortogonales empezó a declinar, a medida que los matemáticos mostraban cada vez más tendencias hacia una mayor abstracción. Sin embargo, de unas tres décadas para acá el interés por esta área ha revivido grandemente, tal vez, como dice T. S. Chihara en el prólogo de su libro [Ch], debido a la revolución computacional y a la mayor actividad en teoría de la aproximación y en análisis numérico.

Devolvámonos a la última década del siglo XIX, cuando el matemático inglés L. C. Rogers ([Ro1], [Ro2], [Ro3]) introdujo los polinomios q -ultraesféricos continuos $\{C_n(x; \beta|q)\}$, que proporcionaron la clave para la prueba de las famosas identidades de Rogers-Ramanujan en 1919 [RoRa]. Rogers mismo no fue conciente de que estos polinomios eran ortogonales, esto se conoció hacia 1940, cuando fueron “redescubiertos” y se verificó que cumplen una relación recurrente de tres términos. Sin embargo, la función de peso solo vino a calcularse explícitamente mucho más tarde (ver [Is]). Cuando $\beta = q^\lambda$ y $q \rightarrow 1$, se obtienen los polinomios ultraesféricos (o de Gegenbauer) $\{C_n^\lambda(x)\}$, de donde proviene su nombre. En un artículo de W. Al-Salam, W. R. Allaway y R. Askey, de 1984 [AAA] se consideran otros casos límites para los q -polinomios de Rogers, a saber, cuando q tiende a una raíz k -ésima ($k \geq 2$) primitiva de la unidad. Se tiene entonces que las relaciones de recurrencia que caracterizan el sistema de polinomios ortogonales obtenidos en este proceso de paso al límite, se rompen en bloques de k relaciones cada uno. Al-Salam, Allaway y Askey llaman a éste un *proceso de cribación* y a los polinomios así obtenidos, *polinomios cribados*.

Tal vez sea éste el punto de decir que el director de la tesis de doctorado de Jairo Charris, Mourad Ismail, fue discípulo de Waleed Al-Salam.

Actualmente se conocen q -versiones, a veces varias, de casi todos los sistemas clásicos de polinomios ortogonales, que se “recuperan” cuando q tiende a 1 de manera apropiada. Un sistema de q -polinomios, cuando se hace tender q a $\exp(2\pi i/k)$, $k \geq 2$, da origen a sistemas de polinomios cribados de diferentes clases, según la manera como tenga lugar el paso al límite.

En el artículo de Al-Salam, Allaway y Askey las propiedades básicas de los polinomios cribados, incluida la

ortogonalidad, se deducen formalmente de las propiedades de los q -polinomios que les dan origen. Los sistemas de polinomios ultraesféricos cribados son sistemas particulares de polinomios cribados de caminos aleatorios. En el primero de una serie de artículos conjuntos, Charris e Ismail [5] desarrollan un estudio completo de los sistemas generales de polinomios cribados de caminos aleatorios, partiendo directamente de las relaciones de recurrencia: obtienen funciones generatrices y dan un teorema que describe cómo computar la transformada de Stieltjes de la medida de distribución (medida con respecto a la cual se tiene la ortogonalidad) a partir del comportamiento asintótico de los polinomios y sus duales.

En un segundo artículo conjunto [7], consideran los llamados polinomios cribados de Pollaczek; éstos se construyen a partir de q -versiones de los polinomios de Pollaczek. Los polinomios de Pollaczek son generalizaciones de polinomios de Legendre y de polinomios de Jacobi, consideradas por Pollaczek en la década de 1950 ([Po1], [Po2], [Po3]). Charris e Ismail estudian esos polinomios cribados partiendo directamente de las relaciones de recurrencia. Entre otros resultados, dan el comportamiento asintótico y calculan la medida de distribución. Ésta es muy interesante, pues puede tener infinitos puntos de masa, junto con la parte absolutamente continua. R. Askey, en la reseña que hace de este artículo en la Zentralblatt für Mathematik [As], dice: “La integral y/o suma para la medida es sorprendente. No sé cómo evaluar estas integrales y sumas directamente - sin la utilización de la maquinaria utilizada por los autores. Ésta usa el comportamiento asintótico de los polinomios ortogonales, que se encuentra a partir de una función generatriz”.

En la nota de condolencia enviada por Ismail con motivo del fallecimiento de Charris, dice literalmente: “Recientemente Barry Simon y su grupo en Caltech encontraron en [el tema de] los polinomios ortogonales y encontraron que lo que él [Charris] hizo en su tesis sobre polinomios de Pollaczek fue típico de la teoría general, así que nuestro artículo conjunto sobre el tema recibió mucha publicidad. A él le hubiera gustado saber esto y me siento mal, pues quería contárselo”.

A partir de estas dos publicaciones, que provienen de su tesis, es claro el enfoque que Charris daría a su trabajo de investigación en esta área: estudio de los sistemas de polinomios ortogonales definidos por relaciones de recurrencia en bloques. Aquí están comprendidos también

sistemas que resultan de aplicaciones polinomiales, introducidos por Geronimo y Van Assche en 1988 [GeVA] y otros casos más generales. Los métodos desarrollados por Charris y sus colaboradores mostraron ser muy poderosos y poseer un amplio rango de aplicación.

En dos artículos más escritos en coautoría con Ismail ([17], [20]) y en numerosas publicaciones elaboradas en colaboración con sus estudiantes de la Universidad Nacional (ver la lista completa de sus publicaciones al final de este artículo), Charris fue desarrollando, perfeccionando, ampliando y extendiendo esta línea de investigación, en la cual dirigió 9 tesis de maestría, codirigió una más (en el primer semestre de este año) y dirigió una tesis de doctorado, esta última en 1999.

Con su labor en este campo Charris ganó reconocimiento internacional y creó en la Universidad Nacional una verdadera escuela de investigación, caracterizada por la seriedad, profundidad, continuidad y calidad del trabajo, que le hicieron acreedor a un profundo respeto y aprecio por parte de la comunidad académica.

Charris, además de excelente matemático, era un gran maestro. Dirigió 2 tesis de doctorado, 18 tesis de maestría, 2 tesis de especialización y 4 trabajos de pregrado. La alta calidad de sus cursos, en todos los niveles y en casi todas las ramas de la matemática, es bien reconocida y fue paradigmático su ejemplo de dedicación, rigor, dignidad y honestidad, cualidades aunadas a su inmensa generosidad intelectual y dedicación a sus estudiantes. La generosidad con que compartía sus conocimientos y sus capacidades se extendía también, y en gran medida, a sus colegas, desde el nivel de la charla informal hasta el de la ayuda grande y a veces decisiva en la elaboración de publicaciones o en la solución de problemas, como lo podemos atestiguar muchos de los presentes. Más de una vez inspiró a algún colega a hacer algo que este último no hubiera pensado que podría hacer.

Otro aspecto de su actividad matemática en el cual Charris dio gran muestra de su capacidad, altruísmo y espíritu de cooperación fue en su labor como editor. Formó parte del Comité Editorial de la Revista Colombiana de Matemáticas desde que apareció el Volumen I, en 1967, habiendo actuado como uno de los Editores desde 1994. Igualmente, fue miembro del Comité Editorial de Lecturas Matemáticas desde 1994. Tanto en la Revista como en Lecturas, su colaboración fue invaluable.

Charris poseía además una clara visión de cómo debería desarrollarse nuestra ciencia en Colombia y siempre expresó y defendió abiertamente sus opiniones, con entereza y honestidad.

Charris tiene más de 30 publicaciones, de las cuales más de la mitad son artículos en revistas de circulación internacional. Fue profesor visitante en la Universidad del Estado de Arizona y en la Universidad del Sur de la Florida, fue conferencista invitado en varias reuniones internacionales en EE. UU. y en el Segundo Coloquio Latinoamericano de Análisis, que se reunió en Bogotá en 1992, y participó activamente, como conferencista o dictando cursos, en al menos 8 congresos nacionales, 12 coloquios distritales y 5 jornadas nacionales de matemáticas. En numerosas ocasiones fue invitado a dictar conferencias en diversas universidades e instituciones. Vale decir que era un excelente expositor.

En 1990 recibió el Premio Nacional de Matemáticas, otorgado por la Sociedad Colombiana de Matemáticas, y en 1991, el Premio de la Academia Colombiana de Ciencias a la Vida y Obra de un Científico. La Universidad Nacional de Colombia lo nombró en 1989 Profesor Emérito, y en 1999, Profesor Honorario.

Fue elegido Miembro Correspondiente de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales en 1998, y Miembro de Número en 2002. Pertenecía a la Sociedad Colombiana de Matemáticas, de la cual fue Presidente en dos ocasiones, de 1972 a 1973 y de 1987 a 1988, a la American Mathematical Society, y a la Society for Industrial and Applied Mathematics (SIAM).

No quiero terminar sin mencionar que además de sus cualidades como matemático y como académico a las que me he referido aquí, Jairo Charris era un hombre de gran cultura y extraordinaria curiosidad intelectual. Conversar con él sobre ciencia, historia, literatura, filosofía o música (clásica o popular) era un gran placer y una experiencia muy instructiva. Tenía especial aprecio y conocimiento de la cultura de la región caribe.

Este evento lleva oficialmente el nombre de “Festival”. Un poco paradójico al considerar la tristeza que sentimos. Sin embargo, me vinieron a la mente las palabras que en la primera conferencia en memoria del gran matemático húngaro Paul Erdős dijera su amigo y colaborador Béla Bollobás: “Ésta debería ser una ocasión alegre, pues nos hemos reunido para celebrar su vida” [Ho]. En el caso de Jairo, una vida verdaderamente ejemplar.

Muchas gracias.

Publicaciones de Jairo Charris

1. *Álgebra moderna*. Public. Dpto. Mat. y Est., Univ. Nacional de Colombia, Bogotá, 1970.
2. *Sobre ciertos espacios de funciones holomorfas y sus aplicaciones I*. Rev. Colombiana Mat. **8** (1974), 111-138.
3. *Series de Fourier*. Boletín de Matemáticas, Suplemento, Sociedad Colombiana de Matemáticas y Dpto. Mat. y Est., Univ. Nacional de Colombia, Bogotá, 1975 (con Joaquín Bustoz).
4. *Sobre ciertos espacios de funciones holomorfas II*. Rev. Colombiana Mat. **13** (1979), 245-263.
5. *On sieved orthogonal polynomials II : Sieved random walk polynomials*. Canadian J. Math. **38** (1986), 397 - 415 (con Mourad E. H. Ismail).
6. *Teoría de Cauchy y Fundamentos del Análisis Complejo*. Public. VI Coloquio Distrital de Matemáticas y Estadística, Universidad Distrital Francisco José de Caldas, Bogotá, 1986. (con Juan C. Quintero y Guillermo Rodríguez-Blanco).
7. *On sieved orthogonal polynomials V : Pollaczek polynomials*. SIAM Journal Math. Analysis **18** (1987), 1177 - 1218. (con Mourad E. H. Ismail).
8. *On the orthogonality measure of the q -Pollaczek polynomials*. Rev. Colombiana Mat. **21** (1987), 301-316 (con Luis A. Gómez).
9. *Orthogonal polynomials, functional analysis and a theorem of Markov*. Rev. Colombiana Mat. **22** (1988), 77 - 128 (con Luis A. Gómez).
10. *On some spaces of analytic functions and their duality relations*. Rev. Colombiana Mat. **22** (1988), 129 - 148 (con Ruth S. Huérfano).
11. *Análisis complejo y escisión en la homología de Artín*. Boletín de Matemáticas **21** (1988), 53 - 98 (con Carlos A. Ortiz).
12. *Orthogonal polynomials with inner and end-point masses*. Rev. Colombiana Mat. **24** (1990), 153 - 177 (con Guillermo Rodríguez - Blanco).
13. *Una nota sobre el significado de las tablas de verdad en la descripción de la matemática formal de N. Bourbaki*, Lecturas Matemáticas, **11** (1990), 1 - 8 (con Juan Varela).
14. *Sobre algunos sistemas de polinomios relacionados con problemas espectrales*. Rev. Colombiana Mat. **25** (1991), 35 - 80 (con Gustavo Salas y Victoria Silva).
15. *On two systems of orthogonal polynomials related to the Pollaczek polynomials*. Rev. Colombiana Mat. **26** (1992), 101 - 146 (con Claudia P. Gómez y Guillermo Rodríguez-Blanco).
16. *Sobre la ortogonalidad distribucional de sistemas de polinomios*. Lecturas Matemáticas, **13** (1992), 53 - 59.
17. *On sieved orthogonal polynomials VII : Generalized polynomial mappings*. Trans. Amer. Math. Soc. **340** (1993), 71 - 93 (con Mourad E. H. Ismail).
18. *Dos sistemas adicionales de polinomios ortogonales relacionados con los polinomios de Pollaczek*. Rev. Colombiana Mat. **27** (1993), 157-186 (con Claudia P. Gómez y Guillermo Rodríguez-Blanco).
19. *Ortogonalidad compleja de sistemas de polinomios*. Revista Integración **11** (1993), 5 - 24.
20. *On sieved orthogonal polynomials X : General blocks of recurrence relations*. Pacific J. Math. **163** (1994), 237 - 267 (con Mourad E. H. Ismail y Sergio Monsalve).
21. *Sobre el teorema integral de Cauchy*. Boletín de Matemáticas **1** (1994), 1 - 8 (con Guillermo Rodríguez-Blanco).
22. *Arco-conexión local en las compactaciones de Alexandroff*. Revista Integración **12** (1994), 69- 79 (con Martha P. Dussán).
23. *Sobre la fracción continua de los polinomios cribados de Pollaczek*. Rev. Fac. Ciencias Univ. Nacional Medellín **4** (1994), 125 - 154 (con Yadira L. Prieto y Félix H. Soriano).
24. *On distributional representations of moment functionals: Sieved Pollaczek polynomials*. Rev. Acad. Colombiana de Ciencias, **19** (1994), 305 - 315 (con Yadira L. Prieto).
25. *Arco-conexión local en las compactaciones por finitos puntos*. Boletín de Matemáticas, **2** (1994), 1 - 16 (con Martha P. Dussán).
26. *Representación de funcionales de momentos por distribuciones*. Memorias II Escuela de Verano, J. Cossio, Editor, Universidad Nacional, Medellín, 1994, 237 - 247.
27. *Complex and distributional weights for sieved ultraspherical polynomials*. Int. J. Math. and Math. Sciences **19** (1996), 229-242 (con Félix H. Soriano).
28. *On the distributional orthogonality of the general Pollaczek polynomials*. Int. J. Math. and Math. Sciences **19** (1996), 417 - 426 (con Félix H. Soriano).
29. *On block recursions, Askey's sieved Jacobi polynomials and two related systems*. Colloquium Math. **78** (1998), 57-91 (con Bernarda Aldana y Oriol Mora-Valbuena).

30. *Some remarks on the local path-connectedness of infinite point compactifications.* Rev. Acad. Colombiana de Ciencias, **22** (1998), 221-228 (con Carmenza Moreno y Oriol Mora Valbuena).
31. *On block recursions and the determination of spectral measures from continued fractions.* Int. J. Appl. Math. **1**, N° 6 (1999), 635-688 (con Oriol Mora-Valbuena).
32. *Some further remarks on the local path-connectedness of compactifications.* Rev. Acad. Colombiana de Ciencias, **23** (1999), 538-586 (con Carmenza Moreno-Roa).
33. *Fundamentos del Análisis Complejo de una Variable.* Academia Colombiana de Ciencias Exactas, Físicas y Naturales. Colección Julio Carrizosa Valenzuela, N° 8. Bogotá, 2000 (con Rodrigo de Castro y Januario Varela).
34. *Fixed points for ω -contractive or ω -expansive maps in uniform spaces: toward a unified approach.* Southwest J. Pure Appl. Math.(2001), 93-101 (electronic only). \equiv Reporte Interno N°74 (2000), Dpto. de Matemáticas, Universidad Nacional de Colombia (con J. Rodríguez-Montes).
35. *Sobre las relaciones de recurrencia, las fracciones continuas y la determinación de las propiedades espectrales de los sistemas ortogonales de polinomios.* Rev. Acad. Colombiana de Ciencias, **27** (2003), 381-421 (con Bernarda Aldana y Germán Preciado).

Bibliografía de otros autores

- [**AAA**] Al-Salam, W., Allaway, W. R. and Askey, R.- *Sieved ultraspherical polynomials.* Trans. Amer. Math. Soc. **284** (1984), 39-55.
- [**As**] Askey, R.- Reseña de [7] en Zbl 0645.33018.
- [**Ch**] Chihara, T.- *An Introduction to Orthogonal Polynomials.* Gordon and Breach, New York, N. Y., 1978.
- [**CoHi**] Courant, R. und Hilbert, D.- *Methoden der Mathematischen Physik I*, 3.Auflage. Springer-Verlag, Berlin-Heidelberg-New York, 1968.
- [**GeVA**] Geronimo, J. and Van Assche, W.- *Orthogonal polynomials on several intervals via a polynomial mapping.* Trans. Amer. Math. Soc. **308** (1988), 559-581.
- [**Ho**] Hoffman, P.- *The Man Who Loved Only Numbers: The Story of Paul Erdős and the Search for Mathematical Truth.* Hyperion, New York, N. Y., 2000.
- [**Is**] Ismail, M.- Reseña de [**AAA**] en MR 85j:33005.
- [**Po1**] Pollaczek, F.- *Sur une généralisation des polynômes de Legendre.* C. R. Acad. Sci. Paris **228** (1949), 1363-1365.
- [**Po2**] Pollaczek, F.- *Sur une famille de polynômes orthogonaux à quatre paramètres.* C. R. Acad. Sci. Paris **230** (1950), 2254-2256.
- [**Po3**] Pollaczek, F.- *Sur une généralisation des polynômes de Jacobi.* Memor. Sci. Mathématiques **131** (1956), Gauthier-Villars, Paris.
- [**Ro1**] Rogers, L. J.- *On the expansion of some infinite products.* Proc. London Math. Soc. **24** (1892), 337-352.
- [**Ro2**] Rogers, L. J.- *Second memoir on the expansion of some infinite products.* Proc. London Math. Soc. **25** (1894), 318-342.
- [**Ro3**] Rogers, L. J.- *Third memoir on the expansion of some infinite products.* Proc. London Math. Soc. **26** (1895), 15-32.
- [**RoRa**] Rogers, L. J. and Ramanujan, S.- *Proof of certain identities in combinatory analysis.* Proc. Cambridge Philos. Soc. **19** (1919), 211-216.
- [**Sa**] Sánchez, C. H.- *Jairo Antonio Charris Castañeda, Premio Sociedad Colombiana de Matemáticas 1990.* Matemáticas, Enseñanza Universitaria **1** (2ªSerie), N°1 (1990), 5-11.
- [**Sz**] Szegő, G.- *Orthogonal Polynomials*, 4th. Ed. Colloquium Publications, Vol.XXII. Amer. Math. Soc., Providence, R. I., 1975.

EL TEOREMA DE HAHN-BANACH COMO PRINCIPIO DE ELECCIÓN

por

Xavier Caicedo ¹ & Germán Enciso ²

A Jairo Charris, in memoriam

Resumen

Caicedo, X. & G. Enciso. El teorema de Hahn-Banach como principio de elección. Rev. Acad. Colomb. Cienc. **28** (106): 11–20, 2004. ISSN 0370-3908.

El teorema de Hahn–Banach implica el axioma de elección para familias de conjuntos convexos cerrados en espacios reflexivos y para familias más generales de convexos en espacios localmente convexos. Es, en efecto, equivalente a varias formas de elección coherente en familias inversamente dirigidas de convexos y transformaciones continuas afines. Lo anterior es consecuencia de algunos resultados relacionados con baricentros de medidas finitamente aditivas y compacidad convexa. Dos caracterizaciones de la reflexividad de espacios normados en términos de estos últimos conceptos se siguen de Hahn–Banach.

Palabras clave: Teorema de Hahn–Banach, axioma de elección, medidas finitamente aditivas, baricentros, compacidad convexa.

Abstract

The Hahn–Banach theorem implies the axiom of choice for families of closed convex sets in normed reflexive spaces and for more general families of convex sets in locally convex spaces. It is, in fact, equivalent to several forms of coherent choice in inversely directed families of convex sets and affine continuous transformations. This is consequence of some results about convex compactness and baricenters of finitely additive measures. Two characterizations of reflexive normed spaces in terms of these last concepts follow from Hahn–Banach.

Key words: Hahn–Banach theorem, axiom of choice, finitely additive measures, baricenters, convex compactness.

¹Departamento de Matemáticas, Universidad de los Andes, A.A. 4976, Bogotá, E-mail: xcaicedo@uniandes.edu.co
Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá.

²Department of Mathematics, Rutgers University, New Brunswick, NJ 08854–9019, EE. UU., E-mail: enciso@eden.rutgers.edu
AMS 2000 Mathematics Subject Classification 46A03, 46A12, 03E25, 52A07, 46G12.

1. Introducción

El famoso teorema sobre de extensión de funcionales que hoy llamamos de Hahn–Banach, pieza fundamental de análisis funcional, se debe independientemente al matemático austriaco **Hans Hahn** (1927) y al polonés **Stefan Banach** (1929), quienes aparentemente generalizaron ideas del también austriaco **Edward Helly** (1912). El teorema se presenta en la literatura en formas diversas, tanto analíticas como geométricas. Tomaremos como básica para este trabajo la siguiente:

HB. *Todo funcional lineal f definido en un subespacio E de un espacio vectorial real V y tal que $f(x) \leq p(x)$, donde p es un funcional sublineal definido en V , puede extenderse a un funcional lineal \hat{f} de V que sigue siendo acotado por p .*

Recuérdese que $p : V \rightarrow \mathbb{R}$ es un *funcional sublineal* si $p(x + y) \leq p(x) + p(y)$ y $p(\alpha x) = \alpha p(x)$, para todo $x, y \in V$ y $\alpha \in \mathbb{R}^+$. Del caso real se obtiene fácilmente el teorema para espacios complejos con respecto a una seminorma p y con el acotamiento expresado en la forma $|f(x)| \leq p(x)$.

Por simplicidad consideraremos solamente espacios vectoriales reales.

Si $\dim(V/E)$ es finita, la extensión \hat{f} puede construirse explícitamente por inducción en la dimensión. En particular, la demostración del teorema para espacios finito-dimensionales no requiere el uso de principios no constructivos como el Axioma de Elección (AE). Sin embargo, en el caso de dimensión infinita la existencia de \hat{f} requiere alguna forma de elección.³ Comúnmente se utiliza toda la fuerza de este axioma, en la forma del Lema de Zorn, para demostrar el teorema.

Por otra parte, HB funciona en muchas circunstancias como un buen sustituto de AE. El mismo **Banach** (1932) lo utilizó para demostrar la existencia de extensiones totales finitamente aditivas e invariantes bajo isometrías de la medida de Lebesgue en \mathbb{R} o en \mathbb{R}^2 .

Posteriormente **Luxemburg** (1969) demostró que HB es equivalente a la existencia de medidas finitamente aditivas en cualquier álgebra booleana. Además de sus consabidas aplicaciones al análisis funcional, HB permite demostrar algunas famosas consecuencias de AE que han sido utilizadas para cuestionar la validez de este último axioma, como son la existencia de conjuntos no Lebesgue-medibles en la recta (**Foreman** y **Wehrung**,

1991) o la paradoja de Banach-Tarski en la esfera (**Pawiikowski**, 1991). Como es bien sabido, de ésta última se desprende la imposibilidad de extender la medida de Lebesgue en \mathbb{R}^n , $n \geq 3$, a una medida total finitamente aditiva e invariante bajo isometrías (cf. **Wagon**, 1994).

Se sabe, sin embargo, que HB es estrictamente más débil que AE, pues se sigue del principio de existencia de ultrafiltros (**Łoś** y **Ryll-Nardzewski**, 1954):

UF. *Todo filtro propio sobre un conjunto puede extenderse a un ultrafiltro,*

principio que no implica AE, como fuera demostrado por **Halpern** y **Levy** (1971). Se sabe incluso que HB es estrictamente más débil que UF (**Pincus**, 1974). Tenemos pues, en la teoría axiomática de conjuntos sin el axioma de elección las implicaciones no reversibles:

$$\text{AE} \Rightarrow \text{UF} \Rightarrow \text{HB}$$

Cabe preguntar entonces cuál es el alcance de HB como un principio de elección, pregunta que intentaremos responder en esta nota.

Mostraremos que HB implica el axioma de elección para familias de convexos cerrados en espacios de Banach reflexivos y para familias más generales de convexos en espacios localmente convexos, y que es en efecto equivalente a ciertas formas de elección coherente en familias inversamente dirigidas de convexos. En particular, HB resulta equivalente a la afirmación de que el límite proyectivo de una familia inversamente dirigida de simplejos finito-dimensionales y funciones simpliciales es siempre no vacío. En el curso del trabajo presentamos algunos resultados relacionados con baricentros de medidas finitamente aditivas y compacidad convexa, incluyendo dos caracterizaciones de la reflexividad de espacios normados en términos de estos conceptos.

Presuponemos familiaridad con las bases del análisis funcional (véase, por ejemplo, **Berberian** 1974, **Horváth** 1966 o **Rudin** 1973). Entre las diversas consecuencias del teorema de Hahn–Banach, utilizaremos más adelante las dos siguientes. La primera se obtiene directamente de HB tomando como funcional sublineal la función $p(x) = \|f\| \|x\|$ y es, como veremos, equivalente a HB.

HBN (Hahn–Banach para espacios normados). *Cualquier funcional lineal continuo f definido en un subespacio de un espacio normado tiene una extensión continua \hat{f} a todo el espacio tal que $\|\hat{f}\| = \|f\|$.*

³Aunque para algunas familias especiales, por ejemplo espacios de Hilbert, la prueba puede hacerse constructivamente.

Para la segunda, recuérdese que un espacio vectorial topológico es *localmente convexo* si tiene una base de vecindades convexas, como es el caso de los espacios normados.

HBG (Hahn–Banach geométrico). *Sea $v \notin C$, donde C es un subconjunto cerrado convexo de un espacio localmente convexo V . Entonces existe $\alpha \in \mathbb{R}$ y un funcional continuo f de V tal que $f(v) < \alpha < f(x)$ para toda $x \in C$.*

El anterior resultado se sigue de HB sin ninguna intervención auxiliar de AE, pues depende de la regularidad de los espacios vectoriales topológicos y de las propiedades del funcional de Minkowski en espacios localmente convexos, las cuales pueden demostrarse constructivamente. Examínese, por ejemplo, la demostración del Th. 3.4(b) en **Rudin** (1973). HBG es una de los muchos teoremas de separación derivados de HB, y vale también si reemplazamos $\{v\}$ por un convexo compacto disjunto de C . Garantiza, entre otras cosas, que los funcionales lineales continuos de un espacio localmente convexo de Hausdorff separan puntos.

2. Medidas finitamente aditivas

Una *medida finitamente aditiva* en una álgebra booleana B no trivial (en adelante simplemente, una *medida*) será una función $m : B \rightarrow \mathbb{R}$ que cumple: $m(1_B) = 1$, $m(a) \geq 0$ y $m(a \vee b) = m(a) + m(b)$ si $a \wedge b = 0$. Esto implica $m(a) = 0$ y $m(a) \leq m(b)$ cuando $a \leq b$.

Luxemburg (1969) demostró la equivalencia entre HB y la existencia de medidas en cualquier álgebra booleana no trivial, utilizando productos reducidos e ideas del análisis no-estándar. Damos aquí una nueva demostración de este resultado. Comenzamos mostrando a partir de HBN un teorema más general sobre extensión de medidas, debido originalmente a **Horn** y **Tarski** (1948), del cual se sigue el de existencia. La otra dirección de la equivalencia se obtendrá más adelante. A lo largo del trabajo, indicaremos entre paréntesis las hipótesis de las que cada proposición depende. ZF denotará la teoría de conjuntos sin el axioma de elección.

Teorema 1 (ZF + HBN). *Cualquier medida definida en una subálgebra de una álgebra booleana puede extenderse a toda el álgebra.*

Demostración. Dada una álgebra booleana $B \neq \{0_B\}$, una *partición* de B es un subconjunto finito $P \subseteq B - \{0\}$ tal que $\bigvee P = 1_B$ y $p \wedge p' = 0$ para todo $p, p' \in P$,

$p \neq p'$. Las particiones están parcialmente ordenadas por la siguiente relación (*Q es más fina que P*):

$$Q \leq P \text{ si y solamente si } \forall q \in Q \exists p \in P \\ \text{(necesariamente \u00fanico) tal que } q \leq p,$$

equivalentemente, todo elemento de P es un supremo de elementos de Q . Con este orden las particiones forman un semiret\u00edculo inferior con

$$P \wedge Q = \{p \wedge q : p \in P, q \in Q\} - \{0\}.$$

Toda funci\u00f3n $f : P \rightarrow \mathbb{R}$ puede “extenderse” a $Q \leq P$, definiendo $f_Q(q) = f(p)$ para el \u00fanico $p \in P$ tal que $q \leq p$. El conjunto $S(B) = \{f : P \rightarrow \mathbb{R} : P \text{ partici\u00f3n de } B\} / \sim$, donde \sim es la relaci\u00f3n de equivalencia:

$$f \sim g \text{ si y solamente si } f_{P \wedge Q} = g_{P \wedge Q},$$

forma un espacio vectorial normado con producto por escalar $\alpha f \sim (\alpha f) \sim$, suma $f \sim + g \sim = (f_{P \wedge Q} + g_{P \wedge Q}) \sim$, y norma $\|f \sim\| = \max\{|f(p)| : p \in P\}$. Cada $a \in B$ tiene una *funci\u00f3n caracter\u00edstica* $\chi_a \in S(B)$:

$\chi_a = f \sim$, donde $f : \{a, a^c\} \rightarrow \mathbb{R}$, $f(a) = 1$, $f(a^c) = 0$, si $a, a^c \neq 0_B$,

$$\chi_{1_B} = f \sim, \text{ donde } f : \{1_B\} \rightarrow \mathbb{R}, f(1_B) = 1,$$

$$\chi_{0_B} = f \sim, \text{ donde } f : \{1_B\} \rightarrow \mathbb{R}, f(1_B) = 0.$$

Se verifica f\u00e1cilmente que χ_{0_B} es el cero del espacio vectorial y que si $a \wedge b = 0$ entonces $\chi_{a \vee b} = \chi_a + \chi_b$. Ahora, si A es una sub\u00e1lgebra de B en la cual est\u00e1 definida una medida m , considere el siguiente subespacio de $S(B)$: $E = \{f \sim : \text{dom } f \subseteq A\}$. Entonces la funci\u00f3n $\varphi : E \rightarrow \mathbb{R}$ dada por

$$\varphi(f \sim) = \sum_{p \in \text{dom}(f)} f(p)m(p)$$

es un funcional lineal bien definido. Claramente $\|\varphi\| = 1$, y por HBN podemos extender φ a todo $S(B)$ manteniendo la misma norma. Entonces $m(a) = \varphi(\chi_a)$ es la medida buscada pues $\varphi(\chi_{1_B}) = 1$, $\varphi(\chi_a) = 1 \cdot m(a) + 0 \cdot m(a^c) = m(a)$ para todo $a \in A$, $\varphi(\chi_{a \vee b}) = \varphi(\chi_a + \chi_b) = \varphi(\chi_a) + \varphi(\chi_b)$ si $a \wedge b = 0$, y adem\u00e1s $\varphi(\chi_a) \geq 0$, pues de lo contrario se tendr\u00eda $1 = \varphi(\chi_a) + \varphi(\chi_{a^c}) < \varphi(\chi_{a^c}) \leq \|\varphi\| \|\chi_{a^c}\| = \|\chi_{a^c}\| \leq 1$. \square

Recuérdese que un subconjunto S de una \u00e1lgebra booleana tiene la propiedad de intersecciones finitas (p.i.f.) si para todo $x_1, \dots, x_n \in S$ se tiene $x_1 \wedge \dots \wedge x_n \neq 0$.

Corolario 1 (ZF + HBN). *En toda \u00e1lgebra booleana no trivial B existe una medida m . Adem\u00e1s, si $S \subseteq B$ tiene*

la p.i.f., ésta puede escogerse de manera que $m(b) = 1$ para todo $b \in S$.

Demostración. Para obtener la existencia de una medida en B , extienda la medida obvia en la subálgebra $\{0_B, 1_B\}$. Ahora, si el subconjunto S tiene la p.i.f. considere el homomorfismo natural $\eta : B \rightarrow B/F$, donde F es el filtro generado por S en B . Como F es propio, B/F es no trivial y tiene una medida μ ; luego $m = \mu \circ \eta$ es la medida deseada. \square

Una primera ilustración del poder de elección de HB es que permite escoger una medida para cada álgebra de una familia dada de álgebras booleanas, sin tener que apelar al axioma de elección, observación debida a Pawiikowski (1991).

Corolario 2 (ZF + HB). *Para toda familia $\{B_i\}_{i \in I}$ de álgebras booleanas no triviales existe una familia $\{m_i\}_{i \in I}$ tal que m_i es una medida en B_i .*

Demostración. Sea B el coproducto de las B_i y sean $h_i : B_i \rightarrow B$, $i \in I$, los homomorfismos canónicos. Como éstos son siempre inyectivos, B es no trivial y por tanto soporta una medida m . Defina $m_i = m \circ h_i$. \square

Una medida en el álgebra de partes $\mathcal{P}(X)$ de un conjunto X se llamará una medida *sobre* X . Las siguientes afirmaciones y construcciones, que generalizan hechos bien conocidos acerca de medidas σ -aditivas definidas en σ -álgebras de conjuntos, pueden obtenerse sin utilizar HB ni forma alguna de AE, como puede verificarlo el lector.

Dado un conjunto X sea $A(X)$ el espacio vectorial de las funciones reales acotadas en X con la norma del supremo, y sea $S(X)$ el subespacio de las funciones en X de rango finito (*funciones simples*). $S(X)$ es efectivamente denso en $A(X)$, es decir, para cada $f \in A(X)$ puede construirse explícitamente una sucesión de Cauchy s_n en $S(X)$ tal que $s_n \rightarrow f$. Tómese, por ejemplo, $r_n = (-1 + \frac{k}{n})\|f\|$, $k = 0, 1, \dots, 2n$, y defínase:

$$s_n(x) = r_k \text{ si } x \in f^{-1}([r_k, r_{k+1})), k = 0, 1, \dots, 2n - 1.$$

$$s_n(x) = r_{2n} \text{ si } x \in f^{-1}(\{r_{2n}\}).$$

(en general, la densidad de un espacio métrico en otro no implica la existencia de tales sucesiones, a no ser que se utilice el axioma de elección enumerable.)

Dada una medida m sobre X puede definirse una integral en las funciones simples de la manera obvia:

$$\int_X s \, dm = \sum_{i=1}^n a_i m(s^{-1}(a_i))$$

donde $\{a_1, \dots, a_n\}$ es el rango de s , y se verifica fácilmente que $\int_X (rs + t) dm = r \int_X s \, dm + \int_X t \, dm$, y $|\int_X s \, dm| \leq \|s\|$. Esto significa que $\int_X (\cdot) dm$ es un funcional lineal continuo en $S(X)$. Como la continuidad es necesariamente uniforme, este funcional puede extenderse unívocamente a todo $A(X)$, definiendo

$$\int_X f \, dm = \lim_n \int_X s_n \, dm.$$

La linealidad y el acotamiento se heredan automáticamente. Otras propiedades que se verifican fácilmente, primero en funciones simples y luego en funciones acotadas, son las siguientes:

I-1. Si $f \leq g$ en X , se tiene $\int_X f \, dm \leq \int_X g \, dm$.

I-2. $\int_X c \, dm = c$, para c constante.

I-3. $|\int_X f \, dm| \leq \|f\|_{A(X)} = \sup_{x \in X} |f(x)|$.

Si $C \subseteq X$ es tal que $m(C) = 1$, entonces m puede considerarse como una medida sobre C y se tiene, escribiendo $\int_C f \, dm$ por $\int_C f|_C \, dm$:

I-4. $\int_C f \, dm = \int_X f \, dm$.

Una función cualquiera $h : X \rightarrow Y$ induce una medida $\mu(S) = m(h^{-1}(S))$ sobre Y , y se tiene para toda $f \in A(Y)$:

I-5. $\int_Y f \, d\mu = \int_X f \circ h \, dm$.

Verificamos esta última propiedad: vale claramente para funciones simples pues $\int_Y s \, d\mu = \sum_{i=1}^n a_i \mu(s^{-1}(a_i)) = \sum_{i=1}^n a_i m((s \circ h)^{-1}(a_i)) = \int_A s \circ h \, dm$. Ahora, si $s_n \in S(X)$ converge a $f \in A(X)$, entonces $s_n \circ h$ converge a $f \circ h$ y por lo tanto, $\int_Y f \, d\mu = \lim_n \int_Y s_n \, d\mu = \lim_n \int_A s_n \circ h \, dm = \int_B f \circ h \, dm$.

3. Baricentros

En adelante, V denotará un espacio vectorial topológico (e.v.t.) y V' denotará el espacio de funcionales lineales continuos de V , es decir su *espacio dual*.

Recuérdese que si V es normado entonces también lo es V' en forma natural, con la norma $\|f\| = \sup\{|f(x)| : \|x\| \leq 1\}$. Para cada $x \in V$ la *evaluación* $\mathcal{E}_x : V' \rightarrow \mathbb{R}$, $\mathcal{E}_x(f) = f(x)$ es un funcional lineal continuo, es decir, $\mathcal{E}_x \in V''$. La transformación lineal $J : V \rightarrow V''$ dada por

$J(x) = \mathcal{E}_x$ es continua, y es una inyección isométrica si V' separa puntos de V . Este será siempre el caso si suponemos HB. El espacio V se dice *reflexivo* si $J : V \rightarrow V''$ es una biyección, es decir todo funcional $\varphi \in V''$ tiene la forma $\varphi(f) = f(x)$ para algún $x \in V$.

Consideraremos también en V' la topología *débil estrella*, es decir, la topología inicial inducida por las evaluaciones \mathcal{E}_x . Esta topología es de Hausdorff, localmente convexa, y en general más débil que la de la norma.

Definición. Sea $X \neq \emptyset$ un subconjunto de un e.v.t. V tal que todo $f \in V'$ es acotado en X y sea m una medida sobre X . Entonces para todo $f \in V'$ existe la integral $\varphi_m(f) = \int_X f dm = \int_X f \upharpoonright X dm$, así que φ_m constituye un funcional lineal de V' . Diremos que $b \in V$ es un *baricentro* de m si $\int_X f dm = f(b)$ para toda $f \in V'$.⁴

Si X es acotado en un espacio normado, se tiene $|\varphi_m(f)| \leq \sup_{x \in X} |f(x)| \leq \|f\|(\sup_{x \in X} \|x\|)$, y por tanto la integral φ_m es continua para la topología de la norma en V' , es decir $\varphi_m \in V''$. Luego, por definición de reflexividad tenemos:

Lema 1 (ZF). *Toda medida sobre un subconjunto acotado de un espacio normado reflexivo tiene baricentro.*

En particular, en espacios de Hilbert y en espacios finito-dimensionales el baricentro existe y coincide con el familiar *centro de masa* (ver **Enciso**, 2000).

Para acotados de espacios normados no reflexivos no necesariamente existe el baricentro (véase el Teorema 2 en esta sección). Sin embargo, una medida en un subconjunto acotado del dual V' de un espacio normado V siempre tiene baricentro si se considera la topología débil estrella en lugar de la topología de la norma, independientemente de que V' sea reflexivo o no.

Lema 2 (ZF). *Sean V normado y $F \subseteq V'$ acotado para la norma inducida en V' . Entonces toda medida m sobre F tiene baricentro con respecto a la topología débil estrella de V' .*

Demostración. Se verifica fácilmente que $b_m(x) = \int_F \mathcal{E}_x dm$ es lineal en $x \in V$, y además $|b_m(x)| \leq \sup_{f \in F} |f(x)| \leq (\sup_{f \in F} \|f\|)\|x\|$, es decir, $b_m \in V'$. Como $\int_F \mathcal{E}_x dm = \mathcal{E}_x(b_m)$ para cada x por definición, y los funcionales continuos de V' para la topología débil estrella son precisamente las evaluaciones \mathcal{E}_x (**Rudin**,

Th. 3.10), se tiene que b_m es el baricentro de m con respecto a esta topología. \square

HB garantiza que el baricentro de una medida sobre X en un espacio de Hausdorff localmente convexo es único y pertenece a la clausura de la envolvente convexa de X ($\text{conv}X$), además de que es preservado por transformaciones lineales continuas. Más precisamente,

Lema 3 (ZF + HB). *Sea V un espacio localmente convexo de Hausdorff y suponga que m es una medida en $X \subseteq V$ con baricentro b . Entonces:*

- i. b es único.
- ii. $b \in \overline{\text{conv}C}$, para todo $C \subseteq X$ tal que $m(C) = 1$.
- iii. Si $t : X \rightarrow Y$ es la restricción de una transformación lineal continua de V a un espacio localmente convexo W y μ es la medida inducida por m en Y entonces $t(b)$ es baricentro de μ en W .

Demostración. (i) Si $x \neq b$ por HBG existe $f \in V'$ tal que $f(x) \neq f(b)$, luego $\int_X f dm = f(b) \neq f(x)$.

(ii) Consideremos a m como una medida en C . Si $b \notin \overline{\text{conv}C}$ entonces por HBG existen $\alpha \in \mathbb{R}$ y $f \in V'$ tales que $f(b) < \alpha \leq f(x)$ para toda $x \in \overline{\text{conv}C}$. Luego $f(b) < \alpha = \int_C \alpha dm \leq \int_C f dm = \int_X f dm$, por las propiedades I-1,2,4 en la sección anterior, una contradicción.

(iii) Para cualquier $g \in W'$ se tiene $\int_Y g d\mu = \int_X g \circ t dm = g(t(b))$ por la propiedad I-5 de la sección anterior. *square*

Observación. En el caso de espacios de Hilbert o espacios finito-dimensionales, el lema anterior depende solamente de ZF, pues los usos de HBG en la prueba pueden obtenerse constructivamente sin invocar HB.

Como hemos dicho, el baricentro no siempre existe en espacios normados no reflexivos. En efecto, suponiendo HB, la existencia de baricentros caracteriza los espacios normados reflexivos.

Teorema 2 (ZF + HB). *Un espacio normado es reflexivo si y solamente si toda medida sobre un subconjunto acotado (o simplemente sobre la bola unitaria cerrada) tiene baricentro para la topología de la norma.*

Demostración. Suponga que V es normado no reflexivo, y sean B y B'' las bolas unitarias cerradas de V y V'' , respectivamente. Entonces existe $d \in B'' \setminus J(B)$, de lo contrario V sería reflexivo. Sea \mathcal{F} el filtro de vecindades

⁴La terminología utilizada en la literatura no es homogénea, **Bourbaki** llama baricentro al funcional φ_m mismo.

cerradas convexas de d para la topología débil estrella de V'' . Como $J(B)$ es densa en B'' para esta topología (la prueba de esta afirmación puede hacerse constructivamente, véase Th. 45.3 en **Berberian**, 1974), el filtro $\mathcal{G} = \{J^{-1}(N) \cap B : N \in \mathcal{F}\}$ tiene la p.i.f. Tome una medida m en B que dé valor 1 a los elementos de \mathcal{G} (Corolario 1) y suponga que tiene baricentro $b \in V$ con respecto a la topología de la norma, entonces $b \in B$ por el Lema 1-ii. Como $J : V \rightarrow V''$ es continua para la topología débil estrella en V'' entonces $J(b)$ debe ser el baricentro de la medida inducida en B'' con respecto a esta topología por el Lema 3-iii. Como las vecindades cerradas convexas de d en B'' para esta topología tienen medida 1, $J(b)$ pertenece a su intersección (Lema 1-ii). Pero por convexidad local y la propiedad de Hausdorff de la topología débil estrella en V'' , la intersección de estas vecindades debe ser $d \notin J(B)$, lo que da una contradicción. \square

4. Compacidad convexa

Diremos que un subconjunto convexo C de un espacio vectorial topológico V es *convexamente compacto* (c -compacto) si toda colección de conjuntos convexas cerrados en C con la propiedad de intersecciones finitas (p.i.f.) tiene intersección no vacía. **Kothe** (1969) considera una versión débil de esta propiedad.

Por supuesto, todo convexo compacto es c -compacto, pero no vale el recíproco. Por ejemplo, la bola unitaria cerrada de un espacio de Hilbert de dimensión infinita es c -compacta (ver Corolario 3), pero es bien sabido que no puede ser compacta. En espacios finito-dimensionales c -compacto y compacto coinciden. Los c -compactos comparten muchas de las agradables propiedades de los compactos si nos restringimos a funciones lineales, o más generalmente a funciones afines. Una función $f : X \rightarrow Y$ entre convexas se dice *afín* si $f(\alpha x + (1 - \alpha)y) = \alpha f(x) + (1 - \alpha)f(y)$, para todo $x, y \in X$ y $\alpha \in \mathbb{R}^+$. Claramente, la imagen directa o inversa de una función afín preserva convexidad.

Lema 4 (ZF). *Sea C un conjunto convexo c -compacto en un e.v.t. V . Entonces:*

- i. *Toda función real afín continua definida en C es acotada.*
- ii. *La imagen de C bajo cualquier transformación afín continua es c -compacta.*
- iii. *Si V es de Hausdorff, C es cerrado.*

Demostración. (i) Si $f \in V'$ no fuera acotado en C entonces los convexas cerrados de $C : C_n = C \cap f^{-1}([n, \infty))$ tendrían la p.i.f. y por c -compacidad existiría $x \in C$ con $f(x) \geq n$ para toda n . (ii) La prueba es la misma que para compacidad ordinaria, pues los conjuntos convexas son preservados por imágenes inversas de funciones afines. (iii) Si $x \in \overline{C}$, entonces $C \cap N \neq \emptyset$ para toda vecindad convexa N de x . Por c -compacidad, $C \cap \bigcap_{x \in N} \overline{N} \neq \emptyset$, y por convexidad local, regularidad y la propiedad de Hausdorff: $\bigcap_{x \in N} \overline{N} = \{x\}$, luego $x \in C$. \square

Se desprende del lema anterior que si m es una medida en un convexo c -compacto C entonces $\int_C f dm$ está definida para toda función real afín continua f .

El siguiente teorema generaliza un resultado clásico sobre existencia de baricentros para medidas σ -aditivas en convexas compactos de espacios localmente convexas (cf. **Bourbaki**, 1955, o **Fonf-Lindenstrauss-Phelps**, 2001), y generaliza el hecho de que en este caso el baricentro representa también la integral de las funciones reales afines continuas.

Teorema 3 (ZF). *Toda medida finitamente aditiva sobre un conjunto convexo c -compacto C de un espacio vectorial topológico tiene un baricentro b en C . Además se tiene $\varphi(f) = \int_C f dm = f(b)$ para toda función real afín continua definida en C .*

Demostración. Por c -compacidad, es suficiente demostrar que la familia de conjuntos convexas cerrados $H_f = C \cap f^{-1}(\varphi(f))$ con f real afín continua en C tiene la p.i.f., pues un elemento $b \in \bigcap_{f \in V'} H_f$ será necesariamente un baricentro de m en C con la propiedad adicional requerida. Dadas funciones reales afines continuas f_1, \dots, f_n en C , sea $T : V \rightarrow \mathbb{R}^n$ el operador afín continuo $T(v) = (f_1(v), \dots, f_n(v))$. Entonces $T(C)$ es convexo, c -compacto por el Lema 4-ii, y cerrado por el Lema 4-iii. Suponga $t = (\varphi(f_1), \dots, \varphi(f_n)) \notin T(C)$, entonces por HBG en \mathbb{R}^n (que no requiere invocar HB, por ser \mathbb{R}^n finito-dimensional), existen un funcional $h(y_1, \dots, y_n) = \sum_i a_i y_i$ en \mathbb{R}^n y un real α tales que $h(t) > \alpha \geq h(x)$, para todo $x \in T(C)$. Es decir, $\sum_i a_i \varphi(f_i) > \alpha \geq \sum_i a_i f_i(v)$ para toda $v \in C$. Tomando la función afín $g = \sum_i a_i f_i$, y usando la linealidad de la integral, esto significa $\varphi(g) > \alpha \geq g(v)$ para todo $v \in C$; lo cual es imposible pues la segunda desigualdad implica $\varphi(g) = \int_C g dm \leq \int_C \alpha dm = \alpha$. Concluimos que $(\varphi(f_1), \dots, \varphi(f_n)) \in T(C)$ y por tanto $(\varphi(f_1), \dots, \varphi(f_n)) = (f_1(b), \dots, f_n(b))$ para algún $b \in C$, es decir, $b \in H_{f_1} \cap \dots \cap H_{f_n}$. \square

Si suponemos HB, la existencia de baricentros caracteriza los c -compactos convexos en espacios localmente convexos de Hausdorff.

Lema 5 (ZF + HB). *Un convexo X en un espacio localmente convexo de Hausdorff es c -compacto si y solamente si es cerrado y toda medida en X tiene baricentro en X .*

Demostración. Una dirección la dan el Lema 4-iii y el teorema anterior. Para la otra suponga que X es cerrado y toda medida tiene baricentro en X , y sea $\{C_i\}_{i \in I}$ una familia de subconjuntos convexos cerrados de X con la p.i.f. Sin pérdida de generalidad podemos suponer que la familia es cerrada bajo intersecciones finitas. Por HB existe una medida m sobre X tal que $m(C_i) = 1$ para toda $i \in I$. Si b es el baricentro de m , entonces $b \in C_i$ para todo i por el Lema 3-ii. \square

De los Lemas 1, 2 y 5 obtenemos:

Corolario 3 (ZF + HB). *Todo convexo cerrado y acotado de un espacio normado reflexivo, o del dual de un espacio normado con la topología débil estrella, es c -compacto.*

Finalmente, del Teorema 2 y el Corolario 3 obtenemos otra caracterización de los espacios normados reflexivos.

Corolario 4 (ZF + HB). *Un espacio normado es reflexivo si y solamente si su bola unitaria es c -compacta para la topología de la norma.*

5. Elección en convexos

Tenemos ya los elementos para obtener las formas de elección explícita que HB provee. Una *elección* para una familia $\{C_i\}_{i \in I}$ de conjuntos no vacíos será una familia $\{x_i\}_{i \in I}$ tal que $x_i \in C_i$ para todo $i \in I$. El axioma de elección afirma, por supuesto, que una tal elección siempre existe.

Teorema 4 (ZF). *HB implica el axioma de elección para familias $\{C_i\}_{i \in I}$ de conjuntos c -compactos en espacios localmente convexos de Hausdorff. Igualmente, para familias de convexos cerrados en espacios normados reflexivos, o en duales de normados con la topología débil estrella.*

Demostración. Por el Corolario 2, existe una familia de medidas m_i sobre C_i , que por el Teorema 3 y el Lemas 3-ii y tienen baricentro único $b_i \in C_i$. La familia $\{b_i\}_{i \in I}$ es la elección requerida. Para la segunda afirmación, si los C_i son convexos cerrados pero no acotados, sea $\delta_i =$

$\inf\{\|x\|_{V_i} : x \in C_i\}$ y aplique lo anterior a los conjuntos $C_i^* = \overline{N_{\delta_i+1}(0)} \cap C_i$, que son convexos cerrados y acotados por construcción, y por lo tanto c -compactos por el Corolario 3. \square

Debemos observar que en el caso de espacios de Hilbert la elección en una familia de convexos cerrados no vacíos $\{C_i\}_{i \in I}$ puede obtenerse explícitamente sin apelar al teorema de Hahn–Banach, pues en este caso existe un único $x_i \in C_i$ tal que $\|x_i\| = \inf\{\|x\| : x \in C_i\}$. Sin embargo, esta última afirmación no es válida para espacios de Banach reflexivos que no sean de Hilbert, ni siquiera en el caso finito-dimensional.

Sea $\{C_i, f_{ij}\}_{i \leq j \in I}$ una familia inversamente dirigida de conjuntos y funciones, donde (I, \leq) es un conjunto parcialmente ordenado e inversamente dirigido (es decir, para todo $i, j \in I$ existe $k \in I$ tal que $k \leq i, j$). Una elección $\{x_i\}_{i \in I}$ de $\{C_i\}_{i \in I}$ se dice *coherente* si

$$f_{ij}(x_i) = x_j \text{ para todo } i \leq j.$$

En otras palabras, $\{x_i\}_{i \in I}$ pertenece al límite inverso o proyectivo de la familia $\{C_i, f_{ij}\}_{i \leq j \in I}$.

Una elección coherente no siempre existe, ni siquiera suponiendo el axioma de elección. Considere, por ejemplo, la cadena de inclusiones $\cdots \rightarrow (0, \frac{1}{3}) \rightarrow (0, \frac{1}{2}) \rightarrow (0, 1)$ en \mathbb{R} . Sin embargo, como consecuencia del Teorema de Tichonoff, si los C_i son compactos de Hausdorff no vacíos y cada f_{ij} es continua entonces el límite proyectivo es no vacío y por lo tanto existe una elección coherente. Ésta es realmente una consecuencia de UF, pues requiere solamente el Teorema de Tichonoff para espacios compactos de Hausdorff, principio equivalente a UF (cf. **Levy**, 1979). El teorema de Hahn–Banach nos proporciona una versión lineal de este resultado.

Teorema 5 (ZF). *HB implica elección coherente para familias inversamente dirigidas de convexos c -compactos no vacíos en espacios localmente convexos de Hausdorff y transformaciones afines continuas.*

Demostración. Una familia inversamente dirigida

$$\{C_i \xrightarrow{t_{ij}} C_j : i, j \in I, i \leq j\},$$

donde cada C_i yace en un espacio localmente convexo de Hausdorff y t_{ij} es una transformación afín, induce una familia directamente dirigida de álgebras booleanas y homomorfismos booleanos con respecto al orden opuesto de (I, \leq) :

$$\cdots P(C_i) \xleftarrow{t_{ij}^{-1}} P(C_j) \cdots$$

Sea $B = \varinjlim P(C_i)$ el límite directo (o inductivo) de esta familia, y sean $P(C_i) \xrightarrow{\rho_i} B$ los homomorfismos canónicos. Como $C_i \neq \emptyset$, entonces cada $P(C_i)$ es una álgebra booleana no trivial y por tanto B es no trivial (el 0 y el 1 de las álgebras no son identificados en ningún $P(C_i)$, luego no son identificados en el límite). Por HB (Corolario 1-i), existe una medida m en B , que induce a su vez una medida $m_i = m \circ \rho_i : P(C_i) \rightarrow \mathbb{R}$, para cada i . Entonces se tiene para toda $j \geq i$

$$m_j = \mu \circ \rho_j = \mu \circ \rho_i \circ t_{ij}^{-1} = m_i \circ t_{ij}^{-1}.$$

Es decir, m_j es la medida inducida por m_i y t_{ij} sobre C_j . Por el Teorema 3 y el Lema 3-i, m_i tiene un único baricentro $b_i \in C_i$. Por la afinidad y continuidad de t_{ij} y el Lema 3-iii, que vale para t afín por el mismo Teorema 3, tenemos para toda $j \geq i$ que $t_{ij}(b_i)$ es baricentro de m_j , y por unicidad de b_j concluimos que $t_{ij}(b_i) = b_j$. Es decir, $\{b_i\}_{i \in I}$ es una elección coherente. \square

Por el Corolario 3, tenemos entonces elección coherente para familias inversamente dirigidas de convexos cerrados acotados no vacíos en espacios reflexivos o duales de normados con la topología débil estrella, y funciones lineales o afines continuas. En particular, para convexos cerrados y acotados en espacios finito-dimensionales.

Un espacio finito-dimensional V posee una única topología que lo hace espacio vectorial topológico de Hausdorff, pues dadas dos tales topologías, la escogencia de una base produce un isomorfismo con \mathbb{R}^n que es un homeomorfismo para cada una de ellas. Fijando una base se ve fácilmente que toda transformación lineal es continua para esta topología. Además cualquier norma en el espacio la induce, y un subconjunto es compacto para esta única topología si y solamente es cerrado y acotado bajo dicha norma (la prueba de estos hechos no requiere forma alguna de AE). La noción de compacto es entonces intrínseca a la estructura lineal de los espacios finito-dimensionales. Por otra parte, los usos de HB en la demostración del teorema anterior (Lema 3-i, iii), aparte de la existencia de medidas, son eliminables en el caso finito-dimensional por la observación que sigue al Lema 3. Si llamamos EM a la existencia de medidas en álgebras booleanas, tenemos entonces:

Corolario 5 (ZF). *EM implica elección coherente para familias inversamente dirigidas de convexos compactos finito-dimensionales y restricciones de funciones lineales.*

Elección coherente en convexos compactos finito-dimensionales permite demostrar a su vez el Teorema de Hahn–Banach, en forma bastante natural.

Teorema 6. (ZF) *Elección coherente para convexos compactos finito-dimensionales y restricciones de funciones lineales implica HB.*

Demostración. Sea p un funcional sublineal en V , E un subespacio de V , y $f : E \rightarrow \mathbb{R}$ un funcional lineal tal que $f \leq p$ en E . Para cada subespacio finito-dimensional W de V , defina

$$C_W = \{g : g \in W', g \supseteq f \upharpoonright W \cap E, g(x) \leq p(x) \text{ para toda } x \in W\},$$

y para $W \supseteq Z$ sea $t_{WZ} : W' \rightarrow Z'$ el operador lineal de restricción. Obviamente C_W es convexo, y es no vacío ya que HB vale en dimensión finita. Si fijamos una base en W y tomamos por norma $\| \cdot \|$ la suma del valor absoluto de las coordenadas, se verifica fácilmente por la sublinealidad y homogeneidad de p que $\sup_{\|x\| \leq 1} p(x) = M < \infty$. Además, $\| \cdot \|$ induce una norma $\|g\| = \sup_{\|x\| \leq 1} |g(x)|$ en W' para la cual C_W es obviamente acotado por M , y con la cual es fácil probar que C_W es cerrado. Esto significa que C_W es compacto para la única topología de Hausdorff de W' (también finito-dimensional). Finalmente, por ser lineal, cada t_{WZ} es continua. Entonces $\{C_W, t_{WZ} \upharpoonright C_W\}_{W \supseteq Z}$ forma un sistema inversamente dirigido de convexos compactos en espacios finito-dimensionales y por hipótesis existe una elección coherente $g_W \in C_W$. Así, $g = \bigcup g_W$ es un funcional que cumple los requisitos de HB. \square

De los corolarios 1, 5 y el Teorema 6 obtenemos la dirección que nos faltaba del teorema de Luxemburg, y también la equivalencia entre HB y HBN.

Corolario 6 (ZF). *HB, HBN y EM son equivalentes.*

Formas aparentemente más débiles de elección coherente ya implican HB. Recordemos que un *simplejo geométrico* es la envolvente convexa de un conjunto finito de vectores (en posición general) en algún espacio vectorial. Una *función simplicial* entre simplejos geométricos, que pueden yacer en distintos espacios, es la extensión convexa de una función entre vértices de los simplejos.

Claramente, todo simplejo es compacto en el espacio finito-dimensional que genera y toda función simplicial es afín y continua.

Teorema 7 (ZF). *Elección coherente para simplejos geométricos y funciones simpliciales implica HB.*

Demostración. Por el corolario anterior, es suficiente mostrar EM. Dada una álgebra booleana B y una subálgebra finita no trivial D , el conjunto de sus átomos, A_D , es necesariamente no vacío. Es fácil ver que si $D' \leq D \leq B$ son subálgebra finitas no triviales de B entonces A_D y $A_{D'}$ son particiones de B (en el sentido de la sección 2) con A_D más fina que $A_{D'}$, luego la función $f_{DD'} : A_D \rightarrow A_{D'}$ dada por

$$f_{DD'}(a) = \text{único } a' \in A_{D'} \text{ tal que } a \leq a'$$

esta bien definida. Sea $\Delta(A_D)$ el simplejo geométrico generado por A_D en \mathbb{R}^{A_D} , y $\bar{f}_{DD'}$ la extensión convexa de $f_{DD'}$ a $\Delta(A_D)$. Entonces la familia

$$\bar{f}_{DD'} : \Delta(A_D) \rightarrow \Delta(A_{D'}), \quad D \geq D'$$

subindicada por las subálgebras finitas de B forma un sistema inversamente dirigido de simplejos y funciones simpliciales. Sea $v_D = \sum_{a \in A_D} v_D(a)a \in \Delta(A_D)$ una elección coherente en este sistema, entonces

$$\bar{f}_{DD'}(v_D) = \sum_{a \in A_D} v_D(a)f_{DD'}(a) = \sum_{a' \in A_{D'}} v_{D'}(a)a',$$

lo que implica, por definición de $f_{DD'}$ que $v_{D'}(a) = \sum\{v_D(a) : f_{DD'}(a) = a'\} = \sum\{v_D(a) : a \leq a', a \in A_D\}$. Defina en D la medida:

$$\mu_D(b) = \sum_{a \in A_D, a \leq b} v_D(a).$$

Entonces $\mu_{D'} = \mu_D \upharpoonright D'$ para $D' \leq D$, pues si $b \in D'$ se tiene $\mu_{D'}(b) = \sum_{a' \in A_{D'}, a' \leq b} \sum\{v_D(a) : a \leq a', a \in A_D\} = \sum_{a \leq b, a \in A_D} v_D(a) = \mu_D(b)$. Así que la medida deseada es $\mu_B = \bigcup\{\mu_D : D \text{ subálgebra finita de } B\}$. \square

Resumiendo los resultados relacionados con elección coherente, tenemos:

Teorema 8 (ZF). *HB es equivalente a cualquiera de los siguientes casos de elección coherente en familias inversamente dirigidas:*

1. *Convexos c -compactos en espacios localmente convexos de Hausdorff y funciones afines continuas.*
2. *Convexos cerrados acotados en espacios normados reflexivos o en duales de normados con la topología débil estrella, y funciones afines continuas.*
3. *Convexos compactos en espacios finito-dimensionales y funciones lineales.*
4. *Simplejos geométricos y funciones simpliciales.*

Demostración. $\text{HB} \Rightarrow 1$ (Teorema 5), $\text{HB} \Rightarrow 2$ (1 y el Corolario 3), $1 \text{ o } 2 \Rightarrow 3 \Rightarrow 4$ (trivialmente), $4 \Rightarrow \text{HB}$ (Teorema 7). \square

6. Limitaciones de HB

Algunos importantes resultados de análisis funcional requieren mucho más que HB para su demostración. Por ejemplo el teorema de Alaoglu (**AL**) que afirma: *La bola unitaria cerrada del dual de un espacio normado es compacta para la topología débil estrella*, no es deducible de HB pues es equivalente a UF. Sin embargo, por el Corolario 3, HB implica una versión débil de AL: *La bola unitaria cerrada del dual de un espacio normado es c -compacta para la topología débil estrella*. Esta forma de AL es equivalente a HB (**Luxemburg**, 1969).

Tampoco implica HB el teorema de Kreĭn-Mil'man (**KM**): *Todo compacto convexo no vacío en un espacio localmente convexo de Hausdorff tiene un punto extremo*, pues se sabe que

$$\text{AL} + \text{KM} \Leftrightarrow \text{AE}$$

(**Bell y Fremlin**, 1972), por lo que ni siquiera UF implica KM, de lo contrario implicaría AE.

Finalmente observamos que, aunque la existencia de puntos extremos en simplejos finito-dimensionales es trivialmente demostrable en ZF, el teorema de Hahn-Banach no garantiza la elección coherente de puntos extremos en familias inversamente dirigidas de simplejos y funciones simpliciales, pues ello implicaría elección coherente en conjuntos finitos, la cual es equivalente a UF.

Bibliografía

- [1] **S. Banach** (1929), *Sur les fonctionelles lineaires, I, II*. Studia Math. **1**, 211–216, 223–239.
- [2] **S. Banach** (1932), *Théorie des opérations linéaires*. Monografie Matematyczne, Vol. 1, Warszawa.
- [3] **J.L. Bell & D.H. Fremlin** (1972), *A geometric form of the axiom of choice*. Fundamenta Mathematicae **LXXVII**, 167–170.
- [4] **S. K. Berberian** (1974), *Lectures in Functional Analysis and Operator Theory*. Springer Verlag, Berlin.
- [5] **N. Bourbaki** (1955), *Éléments de Mathématique, Livre VI, Intégration*. Hermann, Paris.
- [6] **G. Enciso** (2000), *La fuerza de elección del Teorema de Hahn-Banach*. Tesis de pregrado, Universidad de los Andes, Bogotá.
- [7] **V. P. Fonf, J. Lindenstrauss, & R.R. Phelps** (2001), *Infinite dimensional convexity*. In *Handbook of the Geometry of Banach Spaces*. North Holland-Elsevier, pp. 599–670.
- [8] **M. Foreman & F. Wehrung** (1991), *The Hahn-Banach theorem implies the existence of a non Lebesgue measurable set*. Fundamenta Mathematicae **138**, 13–19.

- [9] **H. Hahn** (1926), *Über lineare Gleichungssysteme in linearen Räumen*. Journal für die reine und angewandte Mathematik **157**, 214–229.
- [10] **J.D. Halpern & A. Lévy** (1971), *The boolean prime ideal theorem does not imply the axiom of choice*. AMS Proceedings in Axiomatic Set Theory, 83–134.
- [11] **E. Helly** (1912), *Über lineare Funktionaloperationen*. Sitzgsber. Akad. Wiss. Wien Math.-Kl. **121**, 265–297.
- [12] **A. Horn & A. Tarski** (1948), *Measures in boolean algebras*. Trans. Amer. Math. Soc. **64**, 467–497.
- [13] **J. Horváth** (1966), *Topological vector spaces and distributions*, Vol. I. Addison-Wesley.
- [14] **G. Köthe** (1969), *Topological vector spaces, I*. Springer Verlag, Berlin.
- [15] **A. Levy** (1979), *Basic Set Theory*. Springer Verlag, Berlin.
- [16] **J. Łoś & C. Ryll-Nardzewski** (1954), *On the applications of Tychonoff's theorem in mathematical proofs*. Fundamenta Mathematicae **41**, 49–65.
- [17] **W. A. J. Luxemburg** (1969), *Reduced powers of the real number system*. In: *Applications of model theory to algebra, analysis and probability*, Holt, Rinehart and Winston, New York.
- [18] **J. Pawlikowski** (1991), *The Hahn–Banach theorem implies the Banach-Tarski paradox*. Fundamenta Mathematicae **138**, 21–22.
- [19] **D. Pincus** (1974), *The strength of the Hahn–Banach theorem*. Victoria Symposium on Non-Standard Analysis. Lect. Notes in Math. 369. Springer Verlag, Berlin, 203–248.
- [20] **W. Rudin** (1973), *Functional Analysis*, McGraw-Hill, New York.
- [21] **S. Wagon** (1994), *The Banach-Tarski Paradox*. Cambridge University Press, Cambridge.

APLICACIÓN DE UN PRINCIPIO DE MINIMIZACIÓN*

por

Jorge Cossio[†] & Germán Jiménez[‡]

Dedicado a la memoria del profesor Jairo Charris C.

Resumen

Cossio, Jorge & Germán Jiménez: Aplicación de un principio de minimización. Rev. Acad. Colomb. Cienc. **28** (106): 21–25, 2004. ISSN 0370-3908.

En este artículo se demuestra la existencia de una solución para un problema de Dirichlet no lineal cuando la derivada en el infinito de la no linealidad es menor que el primer valor propio. En la demostración se utiliza, de manera esencial, un teorema de minimización obtenido a partir del principio variacional de Ekeland.

Palabras clave: Principio variacional de Ekeland, soluciones débiles, problema no lineal de Dirichlet.

Abstract

In this paper we prove that a nonlinear Dirichlet problem has a solution when the derivative at infinity of the nonlinearity is less than the first eigenvalue. The proof uses a basic minimization theorem of functionals obtained from the Ekeland variational principle.

Key words: Ekeland variational principle, weak solution, nonlinear Dirichlet problem.

*2000 Mathematics Subject Classification: 35J65, 35J60, 35J20. Este trabajo ha sido apoyado parcialmente por Colciencias contrato 063-2002 y por la DIME proyecto No. 030802593.

[†]Escuela de Matemáticas, Universidad Nacional de Colombia, Apartado Aéreo 3840, Medellín, Colombia. email: jcossio@unalmed.edu.co

[‡]Germán Jiménez, Universidad del Norte, Barranquilla, Colombia, email: gjimenez@uninorte.edu.co

1. Introducción

Sea $f : R \rightarrow R$ una función continua tal que

$$\begin{aligned} f'(-\infty) &:= \lim_{s \rightarrow -\infty} \frac{f(s)}{s} \in R, \\ f'(+\infty) &:= \lim_{s \rightarrow +\infty} \frac{f(s)}{s} \in R. \end{aligned} \quad (1)$$

Sean $\Omega \subset R^N$ ($N \geq 3$) un dominio acotado con frontera suave y Δ el operador de Laplace. Sea $\lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_k \leq \dots$ la sucesión de valores propios de $-\Delta$ con condición de Dirichlet cero en la frontera.

En este artículo se demuestra el siguiente resultado.

Teorema 1.1. *Si $f'(-\infty) = a$ y $f'(+\infty) = b$ con a y $b \in (0, \lambda_1)$, entonces el problema*

$$\begin{cases} \Delta u + f(u) = 0 & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (2)$$

tiene una solución.

La solubilidad del problema (2) ha demostrado estar estrechamente relacionada con la posición de la derivada de la no linealidad con respecto al espectro de $-\Delta$. En efecto, A. Castro y J. Cossio en [C-C2], utilizando métodos variacionales y teoría de grado, demostraron que si el intervalo $(f'(0), f'(\infty))$ contiene los valores propios $\lambda_1, \lambda_2, \dots, \lambda_k$ y $f'(t) < \lambda_{k+1}$ para todo $t \in R$, entonces (2) tiene por lo menos cuatro soluciones no triviales. Los mismos autores en [C-C1] estudiaron la existencia de soluciones en el caso radialmente simétrico cuando el rango de la derivada de la no linealidad incluye al menos los primeros j valores propios y Ω es la bola unitaria en R^N ; allí se emplean técnicas de teoría de bifurcación y se utiliza el hecho de que la ecuación (2) en el caso radialmente simétrico se reduce a una ecuación diferencial ordinaria. K.C. Chang (véase [Ch]) se ha aproximado al problema (2) usando Teoría de Morse. La existencia de soluciones al problema (2) ha sido ampliamente estudiada por muchos autores (véase [Ad-C], [Am-Ra], [C-C-N1], [C-C-N2], [C-C-N3], [C-L1], [C-L2], [C], [E], [H], [Wa]).

El Teorema 1.1 se demuestra usando un teorema de minimización de funcionales obtenido a partir del Principio Variacional de Ekeland (véase [Ek]). En la Sección 2 se presenta la demostración de dicho teorema de minimización. En la Sección 3 se demuestra el Teorema 1.1.

Como un corolario del Teorema 1.1 se obtiene el siguiente resultado.

Teorema 1.2. *Sea $f : R \rightarrow R$ una función continua tal que $\lim_{|s| \rightarrow \infty} \frac{f(s)}{s} = a$ con $a \in (0, \lambda_1)$. Entonces el problema*

$$\begin{cases} \Delta u + f(u) = 0 & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega \end{cases} \quad (3)$$

tiene una solución.

2. Un teorema de minimización obtenido a partir del principio variacional de Ekeland

En [Ek], I. Ekeland demostró el siguiente principio variacional que ha demostrado ser una herramienta muy útil en el estudio de problemas de optimización, teoría de control, geometría diferencial y ecuaciones diferenciales.

Teorema 2.1. [Principio variacional de Ekeland] *Sean (X, d) un espacio métrico completo y $\phi : X \rightarrow R \cup \{+\infty\}$ una función semicontinua inferiormente, no idénticamente igual a $+\infty$, y acotada inferiormente. Sean $\varepsilon > 0$ y $\bar{u} \in X$ tales que*

$$\phi(\bar{u}) \leq \inf_X \phi + \frac{\varepsilon}{2}. \quad (4)$$

Entonces existe $u_\varepsilon \in X$ tal que se cumplen las siguientes condiciones:

$$\phi(u_\varepsilon) \leq \phi(\bar{u}) \quad (5)$$

$$d(u_\varepsilon, \bar{u}) \leq \sqrt{\varepsilon} \quad (6)$$

$$\phi(u_\varepsilon) < \phi(u) + \sqrt{\varepsilon}d(u_\varepsilon, u) \quad \forall u \neq u_\varepsilon \quad (7)$$

El siguiente teorema es una consecuencia del principio variacional de Ekeland y es útil en la prueba del teorema principal de esta sección.

Teorema 2.2. *Sean X un espacio de Banach y $\phi : X \rightarrow R$ una función semicontinua inferiormente, acotada inferiormente y Gateaux-diferenciable $\forall x \in X$. Entonces para cada $\varepsilon > 0$ y $\bar{u} \in X$ tales que*

$$\phi(\bar{u}) \leq \inf_X \phi + \frac{\varepsilon}{2} \quad (8)$$

existe $u_\varepsilon \in X$ tal que

$$\phi(u_\varepsilon) \leq \phi(\bar{u}), \quad (9)$$

$$\|\bar{u} - u_\varepsilon\| \leq \sqrt{\varepsilon}, \quad (10)$$

$$\|\phi'(u_\varepsilon)\|_{X^*} \leq \sqrt{\varepsilon}. \quad (11)$$

Demostración. Sean $\varepsilon > 0$ y $\bar{u} \in X$ tales que satisfacen (8). Por el Teorema 2.1 existe $u_\varepsilon \in X$ tal que se cumplen (9), (10) y

$$\phi(u_\varepsilon) < \phi(u) + \sqrt{\varepsilon} \|u - u_\varepsilon\| \quad \forall u \in X \quad u \neq u_\varepsilon. \quad (12)$$

Sean $v \in X - \{0\}$ y $t > 0$. Tomando $u = u_\varepsilon + tv$ se tiene

$$\frac{\phi(u_\varepsilon) - \phi(u_\varepsilon + tv)}{t} \leq \sqrt{\varepsilon} \|v\|. \quad (13)$$

Pasando al límite cuando $t \rightarrow 0^+$ se obtiene

$$-\langle \phi'(u_\varepsilon), v \rangle \leq \sqrt{\varepsilon} \|v\| \quad \forall v \in X. \quad (14)$$

Como $-v \in X$ y $\phi'(u_\varepsilon)$ es lineal se sigue que

$$|\langle \phi'(u_\varepsilon), v \rangle| \leq \sqrt{\varepsilon} \|v\| \quad \forall v \in X. \quad (15)$$

Por lo tanto

$$\|\phi'(u_\varepsilon)\|_{X^*} = \sup_{v \in X, v \neq 0} \frac{|\langle \phi'(u_\varepsilon), v \rangle|}{\|v\|} \leq \sqrt{\varepsilon},$$

lo cual concluye la prueba del teorema. \square

En la siguiente definición presentamos una condición de compacidad, conocida como la condición de Palais-Smale, la cual es una herramienta útil en la demostración de la existencia de puntos críticos de funcionales definidos en espacios de Banach.

Definición 2.3. Sean X un espacio de Banach y $\phi : X \rightarrow \mathbb{R}$ una función de clase C^1 . Decimos que ϕ satisface la condición de Palais-Smale si toda sucesión $(u_n)_{n \in \mathbb{N}}$ en X tal que $\phi(u_n)$ es una sucesión acotada y

$$\phi'(u_n) \rightarrow 0 \quad \text{en } X^* \quad (16)$$

tiene una subsucesión convergente en X .

A continuación se demuestra el teorema principal de esta sección, el cual establece condiciones suficientes que permiten encontrar puntos de mínimo de funcionales definidos en espacios de Banach.

Teorema 2.4. Sean X un espacio de Banach y $\phi : X \rightarrow \mathbb{R}$ un funcional de clase C^1 que satisface la condición de Palais-Smale y tal que ϕ está acotado inferiormente. Entonces existe $u_0 \in X$ tal que

$$\inf_X \phi = \phi(u_0) \quad \text{y} \quad \phi'(u_0) = 0. \quad (17)$$

Demostración. Sea $n \in \mathbb{N}$. Aplicando el Teorema 2.2 con $\varepsilon = \frac{1}{n}$, existe $u_n \in X$ tal que

$$\phi(u_n) \leq \inf_X \phi + \frac{1}{n} \quad \text{y} \quad \|\phi'(u_n)\|_{X^*} \leq \frac{1}{\sqrt{n}}. \quad (18)$$

Por lo tanto, la sucesión $\phi(u_n)$ es acotada y $\phi'(u_n) \rightarrow 0$ en X^* . Como ϕ satisface la condición de Palais-Smale, existen una subsucesión $(u_{n_j})_{j \in \mathbb{N}}$ y un $u_0 \in X$ tales que $u_{n_j} \rightarrow u_0$. Como ϕ es continua se tiene que

$$\lim_{j \rightarrow \infty} \phi(u_{n_j}) \leq \lim_{j \rightarrow \infty} \left(\inf_X \phi + \frac{1}{n_j} \right) \quad (19)$$

$$\phi(u_0) \leq \inf_X \phi. \quad (20)$$

Luego

$$\phi(u_0) = \inf_X \phi. \quad (21)$$

Además, como ϕ' es continua,

$$\lim_{j \rightarrow \infty} \|\phi'(u_{n_j})\| \leq \lim_{j \rightarrow \infty} \left(\frac{1}{\sqrt{n_j}} \right) \quad (22)$$

$$\|\phi'(u_0)\| \leq 0. \quad (23)$$

Por lo tanto, $\phi'(u_0) = 0$. \square

3. Demostración del teorema 1.1

Sea H el espacio de Sobolev $H_0^1(\Omega)$ (véase [Br]). Recordamos que una **solución débil** de (2) es una función $u \in H$ que satisface

$$\int_{\Omega} \nabla u \cdot \nabla v - \int_{\Omega} f(u) v = 0 \quad \forall v \in H. \quad (24)$$

Sea $J : H \rightarrow \mathbb{R}$ el funcional definido por

$$J(u) = \frac{1}{2} \int_{\Omega} |\nabla u|^2 - \int_{\Omega} F(u), \quad (25)$$

donde $F(s) = \int_0^s f(t) dt$.

Como $f'(-\infty) = a$ y $f'(+\infty) = b$ con a y $b \in (0, \lambda_1)$, existen $c_1 \in \mathbb{R}$ con $\max\{a, b\} < c_1 < \lambda_1$ y $s_0 > 0$ tales que

$$\left| \frac{f(s)}{s} \right| < c_1 \quad \text{para} \quad |s| > s_0. \quad (26)$$

Por la continuidad de f en $[-s_0, s_0]$ existe c_2 tal que

$$|f(s)| \leq c_2 \quad \text{para} \quad |s| \leq s_0. \quad (27)$$

Luego

$$|f(s)| \leq c_1 |s| + c_2 \quad \forall s \in \mathbb{R}. \quad (28)$$

Por lo tanto $J \in C^1(H, \mathbb{R})$ (véase [Ra]) y

$$J'(u)(v) = \int_{\Omega} \nabla u \cdot \nabla v - \int_{\Omega} f(u) v \quad \forall v \in H_0^1. \quad (29)$$

En particular $u \in H$ es solución débil de (2) si y sólo si u es un punto crítico de J .

Demostremos ahora que el funcional J satisface la condición de Palais-Smale. Para ello utilizaremos el siguiente lema y un resultado de D. G. De Figueiredo (véase [DF]).

Lema 3.1 Si $a, b \in (0, \lambda_1)$ entonces el problema

$$\begin{cases} -\Delta w = bw^+ - aw^- & \text{en } \Omega \\ w = 0 & \text{en } \partial\Omega \end{cases} \quad (30)$$

tiene como única solución $w \equiv 0$.

Demostración. Si w es una solución de (30) tal que $w > 0$ en Ω entonces $w^+ = w$ y $w^- = 0$. Luego

$$\begin{cases} -\Delta w = bw & \text{en } \Omega \\ w = 0 & \text{en } \partial\Omega, \end{cases} \quad (31)$$

por lo tanto b es un valor propio de $-\Delta$, esto es una contradicción ya que por hipótesis $b < \lambda_1$. Similarmente se prueba que si w es una solución de (30) tal que $w < 0$ en Ω entonces a es un valor propio de $-\Delta$, lo cual es una contradicción.

Si w es una solución de (30) que cambia de signo, sea $A = \{x \in \Omega / w(x) > 0\}$. Luego

$$\begin{cases} -\Delta w = bw & \text{en } A \\ w = 0 & \text{en } \partial A. \end{cases} \quad (32)$$

Como $b < \lambda_1$ y el autovalor principal de $-\Delta$ en cualquier subregión de Ω es mayor o igual a λ_1 (véase [C-H]), se sigue que $A = \emptyset$. Similarmente se demuestra que si $B = \{x \in \Omega / w(x) < 0\}$ entonces $B = \emptyset$.

Como $w \equiv 0$ es una solución de (30), entonces tiene lugar la afirmación del lema.

Por el Lema 6.3 de De Figueiredo (véase [DF]) se sigue que J satisface la condición de Palais-Smale.

De (28) se sigue que existen c_3 y c_4 en R con $c_1 < c_3 < \lambda_1$ tales que

$$F(s) \leq \frac{1}{2}c_3s^2 + c_4 \quad \forall s \in R. \quad (33)$$

Demostremos que J está acotado inferiormente: en efecto, usando (33) se tiene que

$$\begin{aligned} J(u) &= \frac{1}{2} \int_{\Omega} |\nabla u|^2 - \int_{\Omega} F(u) \\ &\geq \frac{1}{2} \int_{\Omega} |\nabla u|^2 - \frac{1}{2}c_3 \int_{\Omega} u^2 - \int_{\Omega} c_4. \end{aligned} \quad (34)$$

Por la desigualdad de Poincaré (véase [C1]) se sigue que

$$J(u) \geq \left(\frac{\lambda_1 - c_3}{2} \right) \int_{\Omega} u^2 - \int_{\Omega} c_4 \geq -c_4|\Omega|. \quad (35)$$

Luego J está acotado inferiormente.

Hemos demostrado que J satisface las hipótesis del Teorema 2.4, por lo tanto existe $u_0 \in H$ tal que

$$\inf_H J = J(u_0) \quad \text{y} \quad J'(u_0) = 0 \quad (36)$$

Es decir u_0 es un punto crítico de J y por lo tanto una solución débil de (2). Como f es continua y sublineal, por la teoría de regularidad para operadores elípticos (véase [C-L2]) se sigue que las soluciones débiles del problema (2) son soluciones clásicas. Por lo tanto u_0 es una solución clásica de (2). \square

Como una consecuencia inmediata del Teorema 1.1 se obtiene el Teorema 1.2.

Referencias

- [Ad-C] **H. Aduén and A. Castro**, *Infinitely Many Non-radial Solutions to a Superlinear Dirichlet Problem*, Proc. Amer. Math. Soc., **131** (2003), 835–843.
- [Am-Ra] **A. Ambrosetti and P. Rabinowitz**, *Dual variational methods in critical point theory*, J. Funct. Anal., **14** (1973), 349–381.
- [Br] **H. Brezis**, *Análisis Funcional*, Alianza Editorial, Madrid, (1984).
- [C1] **A. Castro**, *Métodos de Reducción via Minimax*, Primer Simposio Colombiano de Análisis Funcional, Medellín, Colombia, (1981).
- [C-C1] **A. Castro and J. Cossio**, *Multiple Radial Solutions for a Semilinear Dirichlet Problem in a ball*, Rev. Colombiana Mat., **27** (1993), 15–24.
- [C-C2] **A. Castro and J. Cossio**, *Multiple Solutions for a Nonlinear Dirichlet Problem*, SIAM J. Math. Anal., **25** (1994), pp. 1554–1561.
- [C-C-N1] **A. Castro, J. Cossio and J. M. Neuberger**, *A Sign-Changing Solution for a Superlinear Dirichlet Problem*, Rocky Mountain J.M., **27** (1997), 1041–1053.
- [C-C-N2] **A. Castro, J. Cossio and J. M. Neuberger**, *On Multiple Solutions of a Nonlinear Dirichlet Problem*, Nonlinear Analysis, Theory, Methods & Applications, **30** (1997), 3657–3662.
- [C-C-N3] **A. Castro, J. Cossio and J. M. Neuberger**, *A Minimax Principle, Index of the Critical Point, and Existence of Sign-changing Solutions to Elliptic Boundary Value Problems*, Electronic Journal of Differential Equations, Vol. 1998 (1998), 1–18.
- [C-L1] **A. Castro and A. C. Lazer**, *Applications of a Max-min Principle*, Rev. Colombiana Mat., **10** (1976), 141–149.
- [C-L2] **A. Castro and A. C. Lazer**, *Critical Point Theory and the Number of Solutions of a Nonlinear Dirichlet Problem*, Ann. Mat. Pura Appl., **70** (1979), 113–137.
- [Ch] **K. C. Chang**, *Solutions of asymptotically linear operator equations via Morse Theory*, Comm. Pure Appl. Math., **34** (1981), 693–712.
- [C] **J. Cossio**, *Múltiples soluciones para un problema elíptico semilineal*, en *Memorias de la III Escuela de*

verano en *Geometría diferencial, ecuaciones diferenciales parciales y análisis numérico*. Academia Colombiana de Ciencias Exactas, Físicas y Naturales, Colección Memorias No. 7, 1995, 53–59.

- [C-H] **R. Courant and D. Hilbert**, *Methods of Mathematical Physics, Volume I*, New York, John Wiley (1989).
- [DF] **D. G. De Figueiredo**, *The Ekeland Variational Principle with Applications and Detour*, Tata Institute of Fundamental Research / Springer-Verlag, (1989).
- [Ek] **I. Ekeland**, *On the Variational Principle*, J. Math. Anal. Appl., **47** (1974), 324–353.

- [E] **M. Esteban**, *Multiple Solutions of Semilinear Elliptic Problems in a Ball*, J. Differential Equations, **57** (1985), 112–137.
- [H] **H. Hofer**, *Variational and Topological Methods in Partially Ordered Hilbert Spaces*, Math. Ann., **261** (1982), 493–514.
- [Ra] **P. Rabinowitz**, *Minimax methods in critical point theory and applications to differential equations*, Conference Series in Mathematics, number 65, American Mathematical society, Providence (1986).
- [Wa] **Z. Q. Wang**, *On a Superlinear Elliptic Equation*, Ann. Inst. H. Poincaré Analyse Non Linéaire **8** (1991), 43–57.

POTENCIA DE PRUEBAS DE RACHAS PARA ALTERNATIVA DE TENDENCIA

por

Emilse Gómez Torres¹, Raydonal Ospina² & Jimmy Corzo³

Dedicado al profesor Don Jairo Charris Castañeda

Resumen

Gómez Torres, E., Raydonal Ospina & Jimmy Corzo: Potencia de pruebas de rachas para alternativa de tendencia, Rev. Acad. Colomb. Cienc. **28** (106): 27–38, 2004. ISSN 0370-3908.

Se presenta un estudio de simulación de la potencia de varias pruebas de rachas para una alternativa de tendencia monótona, en muestras obtenidas de distribuciones de Laplace, logística y normal, y para funciones de tendencia monótona con comportamiento de crecimiento lento a moderado.

Palabras Clave. Simulación, tendencia monótona en localización, pruebas de rangos, hipótesis de aleatoriedad

Abstract

A simulation study with observations coming from Laplace, logistic and normal distributions is presented to estimate the power of run tests for trend alternative and to compare them with Daniels test.

Key Words and Phrases. Simulation, location trend, rank tests, hypothesis of randomness

1. Introducción

Sea Y_1, Y_2, \dots, Y_T una sucesión de variables aleatorias independientes que representan T observaciones en el tiempo de una variable Y , con funciones de distribución continuas $F_t(y \pm a(t))$ para $t = 1, \dots, T$ con varianza

constante, donde $a(t)$ es una función monótona en t , continua en el intervalo $[1, T]$ y derivable en cada punto del intervalo $(1, T)$. La hipótesis de aleatoriedad se formula como sigue

$$H_1 : F_1(x) = \dots = F_T(x) = F(x) \text{ para todo } x,$$

¹Profesora Asistente del Departamento de Estadística de la Universidad Nacional de Colombia (egomez@matematicas.unal.edu.co)

²Estadístico de la Universidad Nacional de Colombia (raydonal@cox.de.ufpe.br).

³Profesor Asociado del Departamento de Estadística de la Universidad Nacional de Colombia (jcorzo@matematicas.unal.edu.co).

lo cual equivale en términos de $a(t)$ a la hipótesis:

$$H_1 : a'(t) = 0, \text{ para todo } t \in (1, T) \quad (1)$$

Se consideran las siguientes dos alternativas de tendencia:

$$K_1 : a'(t) > 0 \text{ para todo } t \in (1, T) \quad (2)$$

tendencia monótona creciente

$$K_2 : a'(t) < 0 \text{ para todo } t \in (1, T)$$

tendencia monótona decreciente

Obsérvese que la función $a'(t)$ es responsable de la tendencia en las alternativas consideradas.⁴

2. Estadístico basado en rachas

Para la construcción de un estadístico de rachas⁵ es necesario transformar la sucesión Y_1, Y_2, \dots, Y_T es una sucesión dicótoma⁶, que contenga información sobre la tendencia de la sucesión. Para esto se define la siguiente función de cambio en la cual se compara una observación en el tiempo i con otra que se encuentra m unidades más adelante:

$$\Delta_{i,m} := Y_{i+m} - Y_i$$

donde: $i = 1, \dots, T-1$, y m es un número que depende del número máximo de comparaciones entre una observación y cada una de las posteriores a ella; este número máximo de comparaciones se denota por k y se fija de acuerdo con la cantidad de información sobre la tendencia que se quiere recuperar. Así $m = 1, \dots, \min\{k, T-i\}$, esto es $m = 1, \dots, k$ cuando $i = 1, \dots, T-k$ y $m = 1, \dots, T-i$ cuando $i = T-k+1, \dots, T-1$.

Por ejemplo, si se toma $k = 3$, es porque se quiere comparar cada observación con las tres subsiguientes a ella. En tal caso cuando $i = 1, \dots, T-3$, $m = 1, 2, 3$; cuando $i = T-2$, $m = 1, 2$ y cuando $i = T-1$, $m = 1$; a partir de los cuales se generan las siguientes $3(T-2)$ cambios: $\Delta_{1,1}, \Delta_{1,2}, \Delta_{1,3}, \Delta_{2,1}, \Delta_{2,2}, \Delta_{2,3}, \dots, \Delta_{T-3,1}, \Delta_{T-3,2}, \Delta_{T-3,3}, \Delta_{T-2,1}, \Delta_{T-2,2}, \Delta_{T-1,1}$.

⁴Aiyar, Guillier y Albers (1979) utilizan el siguiente modelo más dinámico $Y_t = \beta_T a(t) + b + \varepsilon_t$ donde el parámetro β_T depende del número de observaciones en la sucesión y b es una constante real que representa un cambio en localización, para estudiar la eficiencia relativa asintótica de ciertas pruebas de rangos para alternativas de tendencia monótona.

⁵Corresponde al término inglés *run*, el cual según Gibbons se puede definir así: Dada una secuencia de dos o más tipos de símbolos, una racha es definida como una sucesión de uno o más símbolos idénticos que son seguidos o precedidos por un símbolo diferente o por ningún símbolo (Gibbons (1971, pág. 50))

⁶Una sucesión dicótoma es un arreglo de dos símbolos distinguibles. La transformación de una sucesión de variables aleatorias a una sucesión de variables aleatorias binarias se denomina dicotomización. La correspondiente sucesión de realizaciones de las variables binarias también se denominará sucesión dicótoma.

También se define la siguiente variable que indica si el cambio $\Delta_{i,m}$ corresponde a un **estado de incremento** en el valor de Y en el tiempo $i+m$ con respecto al tiempo i o no:

$$\begin{aligned} \eta_{i,m} &:= \begin{cases} 1 & \text{si } \Delta_{i,m} > 0 \\ 0 & \text{si } \Delta_{i,m} \leq 0 \end{cases} \\ &:= \frac{1}{2} \{1 + \text{sign}(\Delta_{i,m})\} \end{aligned} \quad (3)$$

para $i = 1, \dots, T-1$ y $1 \leq m \leq T-i$, donde

$$\text{sign}(x) := \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x \leq 0 \end{cases} \quad \text{para } x \in R$$

El número de elementos incluidos en la sucesión dicótoma depende de k , por esta razón se denota por $N(k)$ el cual se obtiene sumando las $k(T-k)$ comparaciones entre k elementos sucesivos a todas las comparaciones posibles de los últimos k elementos tomados por pares, así: $N(k) = k(T-k) + k(k-1)/2 = k(T - (k+1)/2)$. Retomando el ejemplo anterior, cuando $k = 3$, $N(k)$ se obtiene sumando los $3(T-3)$ cambios de los $\Delta_{i,m}$ con $i = 1, \dots, T-3$ y $m = 1, 2, 3$ con los 3 cambios de las últimas tres observaciones tomadas en pares, por tanto $N(k) = 3(T-3) + 3(3-1)/2 = 3(T - (3+1)/2) = 3(T-2)$

La forma de interpretar la sucesión dicótoma se ilustra en el siguiente ejemplo:

En caso de que $\eta_{i,m} = 1$ para todo i y para todo m es porque $Y_1 < Y_2 < \dots < Y_T$, lo cual significa que la sucesión de los Y_t es estrictamente creciente.

Si el caso es $\eta_{i,m} = 1$ para $i = 1, \dots, T-2$ y $\eta_{T-1,1} = 0$, es porque $Y_1 < Y_2 < \dots < Y_{T-1}$, $Y_i < Y_T$ para $i = 1, \dots, T-2$ y $Y_{T-1} > Y_T$, esto significa que la sucesión de los Y_t es creciente hasta la penúltima observación.

Con este mismo razonamiento se puede intuir que cualquier alteración en el orden estricto de incremento de la sucesión de los Y_t se verá reflejada en la presencia de ceros en la sucesión de los $\eta_{i,m}$. De esta manera cuando el número de unos presentes en la sucesión de los $\eta_{i,m}$ es grande se podrá interpretar que la mayor

parte del tiempo la sucesión de los Y_t permaneció incrementándose; mientras que una sucesión que contiene mayor número de ceros que de unos representa que la sucesión de los Y_t esta decreciendo.

Los $\eta_{i,m}$ forman un arreglo cuyo número de filas es $T - 1$ y el número de elementos en las filas varía con el valor de k . En la Tabla 1 se ilustra el tipo de arreglo que se produce cuando $k = 3$:

Dado que para obtener el número de rachas en una sucesión dicótoma es necesario que ésta esté ordenada, se adoptará aquí el orden en que se presentan los elementos en cada una de las filas como se indica en la Tabla 1. De esta manera la posición t en la sucesión dicótoma del elemento $\eta_{i,m}$, esta dada por la siguiente función de los instantes de observación que se están comparando:

$$t(i, m) := \begin{cases} k(i-1) + m & i \leq T - k \\ k(T-k) + \frac{1}{2}(k+i-T-1)(k-i+T) + m & i > T - k \end{cases} \quad (4)$$

Tabla 1. Arreglo de los elementos $\eta_{i,m}$ cuando $k = 3$ para lectura unidireccional

$\eta_{1,1}$	$\rightarrow \eta_{1,2}$	$\rightarrow \eta_{1,3}$					✓	
	$\rightarrow \eta_{2,1}$	$\rightarrow \eta_{2,2}$	$\rightarrow \eta_{2,3}$				✓	
		$\rightarrow \eta_{3,1}$	$\rightarrow \eta_{3,2}$	$\rightarrow \eta_{3,3}$			✓	
			\ddots	\ddots	\ddots		\vdots	
				\ddots	\ddots	\ddots	\vdots	
					$\rightarrow \eta_{T-3,1}$	$\rightarrow \eta_{T-3,2}$	$\rightarrow \eta_{T-3,3}$	✓
						$\rightarrow \eta_{T-2,1}$	$\rightarrow \eta_{T-2,2}$	✓
							$\rightarrow \eta_{T-1,1}$	

De aquí en adelante el subíndice t se utiliza para indexar sucesiones que relacionan observaciones en instantes de tiempo que dependen de i y de m . Con esta forma de leer los subíndices, el arreglo de los $\eta_{i,m}$ se transforma en la sucesión

$$\{\eta_{1,1}, \eta_{1,2}, \eta_{1,3}, \eta_{2,1}, \eta_{2,2}, \eta_{2,3}, \eta_{3,1}, \dots, \eta_{T-3,3}, \eta_{T-2,1}, \eta_{T-2,2}, \eta_{T-1,1}\}$$

la cual puede escribirse en términos de un solo subíndice de la siguiente forma

$$\{\eta_1, \eta_2, \eta_3, \eta_4, \eta_5, \eta_6, \eta_7, \dots, \eta_{3(T-3)}, \eta_{3(T-3)+1}, \eta_{3(T-3)+2}, \eta_{3(T-2)}\}$$

Para la construcción del estadístico de prueba se define una variable indicadora de permanencia en un estado de incremento o decremento en el instante t :

$$\begin{aligned} {}^{-}I_1 &:= 1 \\ {}^{-}I_t &:= \begin{cases} 1 & \eta_t = \eta_{t-1} \\ 0 & \eta_t \neq \eta_{t-1} \end{cases} \quad \text{para } t = 2, \dots, N(k) \end{aligned} \quad (5)$$

Entendiendo por antirraza el caso en el cual la situación de los η_t se mantiene estable en el incremento o decremento se define el número de antirrachas hasta el t -ésimo elemento de la sucesión $\{\eta_1, \eta_2, \dots, \eta_{N(k)}\}$ ⁷ por:

$${}^{-}r_t := \sum_{n=1}^t {}^{-}I_n \quad \text{para } t = 1, 2, \dots, N(k) \quad (6)$$

Nótese que ${}^{-}r_t - 1$ es el número de veces que la sucesión de los η_t se mantuvo estable en un estado de incremento

⁷El número de antirrachas es una función de las rachas, como se observa en Corzo (1995).

o decremento hasta el tiempo $t(i, m)$, en el arreglo de las comparaciones sucesivas $\Delta_{i,m}$ de los elementos de la sucesión $\{Y_1, Y_2, \dots, Y_T\}$. Por ejemplo en la Tabla 1, el ${}^{-}r_{t(2,3)} - 1 = {}^{-}r_8 - 1$ representa el número de veces que la subsucesión $\eta_{1,1}, \eta_{1,2}, \eta_{1,3}, \eta_{2,1}, \eta_{2,2}, \eta_{2,3}, \eta_{3,1}, \eta_{3,2}$ ha permanecido en unos o en ceros; así si esta subsucesión fuera 1, 1, 0, 1, 0, 0, 0, 1 entonces ${}^{-}r_8 - 1 = 3$ indicando tres permanencias en algún estado, dos marcadas con líneas debajo y la última corresponde a los dos ceros en negrilla. Se define la función que indica con un valor positivo un incremento y con un valor negativo un decremento en un instante dado, mediante la función:

$$\delta_{t(i,m)} := g(i, m) \text{sign}(\Delta_{i,m}) = \delta_t \quad (7)$$

cuyo signo está dado por la función de cambio. Dos posibles formas de la función $g(i, m)$ que son independientes de m se utilizarán más adelante: $g(i, m) = 1$ y $g(i, m) = i$. Finalmente como el número de antirrachas en la sucesión de η_t es una variable aleatoria, se utiliza la función $\varsigma({}^{-}r_{N(k)})$ para normar el estadístico de rachas, que se define como sigue:

$$G_k := \frac{1}{\varsigma({}^{-}r_{N(k)})} \sum_{t=1}^{N(k)} \delta_t {}^{-}r_t \quad (8)$$

Valores extremos positivos de G_k se obtienen cuando hay muchas permanencias (${}^{-}r_t - 1$ grande) y son en

su mayoría incrementos (δ_t positivo), éstos reflejan una tendencia de la sucesión Y_1, \dots, Y_T a permanecer en estado creciente. Análogamente, valores extremos negativos de G_k se obtienen cuando hay muchas permanencias (${}^{-}r_t - 1$ grande) y son en su mayoría decrementos (δ_t negativo), éstos reflejan una tendencia de la sucesión Y_1, \dots, Y_T a permanecer en estado decreciente. En consecuencia, la hipótesis nula se rechazará en favor de la alternativa de tendencia a permanecer en algún estado cuando el estadístico toma valores extremos, tanto positivos como negativos.

El estudio de potencia se hará para varios casos particulares de G_k que se distinguen de dos maneras: por el tipo de función que indica el cambio $g(i, m)$ y por la función que normaliza el valor de la estadística $\varsigma({}^{-}r_{N(k)})$. Para facilitar la interpretación estos casos se distinguen con las letras ${}^{-}C_k$ y γ_k y se muestran en la Tabla 2.

Nótese que en los casos estudiados, los estadísticos de la forma ${}^{-}C_k$ tienen como factor de normalización ${}^{-}r_{N(k)}$ y los de la forma γ_k no están normalizados. Además la función $g(i, m)$ se mantiene constante en ${}^{-}C_k$, mientras que γ_k es una función del primero de los dos instantes de observación comparados en la función de cambio.

Tabla 2. Estadísticos basados en rachas para pruebas de tendencia monótona

Estadístico	$g(i, m)$	$\varsigma({}^{-}r_{N(k)})$	Casos: Características
${}^{-}C_k = \frac{1}{{}^{-}r_{N(k)}} \sum_{t=1}^{N(k)} \delta_t {}^{-}r_t$	1	${}^{-}r_{N(k)}$	${}^{-}C_1 : \begin{cases} k = 1; m = 1; \\ N(k) = T - 1 \end{cases}$ ${}^{-}C_{T-1} : \begin{cases} k = T - i; m = 1, \dots, T - i; \\ N(k) = T(T - 1)/2 \end{cases}$
$\gamma_k = \sum_{i,m} \delta_{i,m} {}^{-}r_{i,m}$	i	1	$\gamma_1 : \begin{cases} k = 1; m = 1; \\ N(k) = T - 1 \end{cases}$ $\gamma_2 : \begin{cases} k = 2; m = 1, \text{mín} \{T - i, 2\}; \\ N(k) = 2T - 3 \end{cases}$ $\gamma_5 : \begin{cases} k = 5; m = 1, \dots, \text{mín} \{T - i, 5\}; \\ N(k) = 5T - 15 \end{cases}$ $\gamma_{T-1} : \begin{cases} k = T - i; m = 1, \dots, T - i; \\ N(k) = T(T - 1)/2 \end{cases}$

2.1 Regiones críticas

Tablas de los valores críticos de ${}^{-}C_1$ y ${}^{-}C_{T-1}$ y un procedimiento para obtener la distribución de los estadísticos, se encuentran en Corzo y Gómez (2000). El método que se utiliza para obtener la distribución de los estadísticos γ_k es el mismo que el empleado para el

cálculo de la distribución de ${}^{-}C_1$ y ${}^{-}C_{T-1}$ dado que la única diferencia entre las dos clases de estadísticos esta en el denominador⁸. Para tamaños de muestra entre 5 y 8 se utilizan los valores críticos exactos de las distribuciones de los seis estadísticos y para tamaños de muestra entre 9 y 50 se emplean las aproximaciones

⁸Detalles sobre el cálculo de las distribuciones y regiones críticas de los estadísticos γ_k se pueden consultar en Gómez (2002).

obtenidas mediante simulación. El mecanismo adoptado para aproximar los percentiles consiste en generar un número adecuado de permutaciones aleatorias de los enteros de 1 a T y posteriormente estimar los percentiles con la probabilidad asociada más cercana por debajo, al nivel de significancia deseado⁹.

3. Estudio de potencia

El estudio de la potencia se lleva a cabo simulando

muestras a partir del modelo $Y_t = \pm a(t) + \varepsilon_t$, en el cual $a^j(t)$ representa una función de tendencia monótona creciente (ocho casos de $a^j(t)$ se muestran en la Tabla 3), y ε_t es la parte aleatoria que se generó a partir de las distribuciones de Laplace, logística y normal. En este artículo sólo se presentan los resultados para $a(t)$ dado que los resultados para $-a(t)$ son muy similares (Gómez (2002)).

Tabla 3. Casos de tendencia monótona creciente utilizados

Caso j	$a^j(t)$	$a'^j(t)$: Rapidez de crecimiento
Caso 1	$a^1(t) = t^{1/4}$	$a'^1(t) = \frac{1}{4}t^{-3/4}$
Caso 2	$a^2(t) = t^{2/5}$	$a'^2(t) = \frac{2}{5}t^{-3/5}$
Caso 3	$a^3(t) = t^{1/2}$	$a'^3(t) = \frac{1}{2}t^{-1/2}$
Caso 4	$a^4(t) = t^{3/5}$	$a'^4(t) = \frac{3}{5}t^{-2/5}$
Caso 5	$a^5(t) = t^{7/10}$	$a'^5(t) = \frac{7}{10}t^{-3/10}$
Caso 6	$a^6(t) = t^{4/5}$	$a'^6(t) = \frac{4}{5}t^{-1/5}$
Caso 7	$a^7(t) = t^{9/10}$	$a'^7(t) = \frac{9}{10}t^{-1/10}$
Caso 8	$a^8(t) = t$	$a'^8(t) = 1$

Aproximación de las potencias (para cada caso de $a^j(t)$ de la Tabla 3)

Paso 1: Generar m muestras de tamaño T de la distribución seleccionada para ε_t , obtener $y_t = \pm a^j(t) + \varepsilon_t$ para $t = 1, \dots, T$, y calcular el número de rechazos de la hipótesis nula $\hat{\pi}_1$.

Paso 2: Generar el doble de muestras de tamaño T que en el paso anterior, obtener $y_t = \pm a^j(t) + \varepsilon_t$ para $t = 1, \dots, T$, calcular el número de rechazos de la hipótesis nula $\hat{\pi}_2$ y $|\hat{\pi}_2 - \hat{\pi}_1|$

Paso 3: Si $|\hat{\pi}_2 - \hat{\pi}_1| > 0.005$ ir a paso 2.

Paso 4: Terminar.

Una aproximación del nivel de significancia se obtiene con el anterior algoritmo para $a^j(t) = 0$

En este estudio, se comparan entre sí las potencias de las pruebas que utilizan los estadísticos propuestos ${}^-C_1$, ${}^-C_{T-1}$, γ_1 , γ_2 , γ_5 y γ_{T-1} , y con las potencias de la prueba de Daniels, que utiliza el estadístico:

$$\rho = \frac{12}{T(T^2 - 1)} \sum_{i=1}^T \left(i - \frac{T+1}{2} \right) \left(R_i - \frac{T+1}{2} \right) \quad (9)$$

donde R_i denota el rango de la i -ésima observación. Como ρ está basado en el coeficiente de correlación de Spearman, su rango de variación está entre -1 y 1; así que la prueba rechaza la hipótesis de aleatoriedad cuando ρ se acerca a sus valores extremos.

Como en los tamaños de muestra menores o iguales que ocho no era posible obtener el mismo nivel de significancia en todas las pruebas, la comparación se realizó sobre las potencias empíricas de las pruebas aleatorizadas. Para tamaños de muestra mayores, no fue necesario hacer la aleatorización de las pruebas porque los niveles de significancia resultaron muy parecidos para todas las pruebas.

3.1. Resultados

El estudio del comportamiento de las potencias de las pruebas se realizó de la siguiente manera:

⁹El número de permutaciones a generar es el tamaño de una muestra aleatoria simple del conjunto de las $T!$ permutaciones de los enteros de 1 a T . La varianza de los estadísticos se aproximó por el método Jackknife. Una descripción completa de la metodología de simulación se encuentra en Gómez (2002).

- **Consistencia**¹⁰: Cambiando el tamaño de sucesión y fijando un caso de tendencia y una distribución. Los resultados de esta parte se presentan en la Tabla resumen 1 en la cual las filas muestran los tamaños de muestra que resaltan las diferencias entre las potencias aproximadas de las pruebas comparadas.
- **Potencia**: Cambiando el caso de tendencia y fijando un tamaño de la sucesión y una distribución. Los resultados de esta parte se presentan en la Tabla resumen 2 que contiene en sus cuatro filas las aproximaciones numéricas de las potencias de las siete pruebas para los cuatro casos de tendencia que más diferencias establecen entre las potencias aproximadas de las pruebas comparadas: $t^{1/4}$, $t^{1/2}$, $t^{4/5}$ y t (casos 1, 3, 6 y 8)¹¹.
- **Sensibilidad**¹²: Cambiando la distribución y fijando un tamaño de sucesión y un caso de tendencia.

Las Tablas Resumen presentan las aproximaciones de las potencias de las pruebas con un nivel de significancia de 5%¹³, simulando 50000 muestras¹⁴ de tamaño T que varía de 5 a 50 (incrementando de 1 en 1), para muestras de distribución de Laplace¹⁵, utilizando la siguiente notación:

T : tamaño de la muestra, aumenta de 5 en 5 desde 5 hasta 50

C1: potencia estimada de la prueba con $^{-}C_1$

C(T-1): potencia estimada de la prueba con $^{-}C_{T-1}$

gama1: potencia estimada de la prueba con γ_1

gama2: potencia estimada de la prueba con γ_2

gama5: potencia estimada de la prueba con γ_5

gama(T-1): potencia estimada de la prueba con γ_{T-1}

Daniels: potencia estimada de la prueba de Daniels (con ρ)

3.1.1. Consistencia

¹⁰Se realizó para muestras hasta de tamaño 50.

¹¹En Gómez (2002) se encuentran la totalidad de resultados enunciados en este artículo de manera que el lector interesado en profundizar en el tema pueda consultar un documento más detallado.

¹²Se refiere a la sensibilidad de las pruebas al grado de apuntamiento de las distribuciones muestreadas.

¹³Los resultados para un nivel de significancia del 1% son muy similares a los presentados aquí y se pueden consultar en Gómez (2002).

¹⁴Una muestra corresponde a una sucesión de variables de un tamaño dado con un comportamiento específico.

¹⁵Los resultados de las potencias con las distribuciones logística y normal son muy similares, razón por la cual no se hacen comentarios sobre ellos y sólo se presentan las tablas correspondientes en el anexo.

En la Tabla resumen 1 se evidencia el crecimiento de las potencias, en algunos casos muy lento y en otros casos bastante rápido, dependiendo de la rapidez de crecimiento de las funciones de tendencia y del número de comparaciones entre pares de observaciones consideradas en el estadístico de prueba. En general, el crecimiento más rápido lo presentan las potencias de la prueba de Daniels, seguido de cerca por las potencias de la prueba con $^{-}C_{T-1}$ y, un poco más despacio, por las potencias de la prueba con γ_{T-1} ; las potencias más bajas son las de las pruebas cuyos estadísticos son γ_5 , γ_2 , $^{-}C_1$ y γ_1 , en su orden.

Caso 1: $a^1(t) = t^{1/4}$. Las potencias de las pruebas basadas en γ_5 , γ_2 , γ_1 y $^{-}C_1$ son muy bajas (no superan el 12%) y muestran un comportamiento casi constante. Las potencias de las pruebas basadas en γ_{T-1} y $^{-}C_{T-1}$ presentan un crecimiento moderado llegando sólo al 41,68% y 71,05% respectivamente, mientras que la prueba de Daniels alcanza el 91,63%.

Caso 3: $a^3(t) = t^{1/2}$. Se refleja con claridad la mayor bondad de las pruebas que emplean todas las comparaciones (basadas en ρ , $^{-}C_{T-1}$ y γ_{T-1}), pues rápidamente obtienen la máxima potencia, aunque cada prueba lo hace a partir de un tamaño de muestra diferente ($T = 28$ para la prueba basada en ρ , $T = 37$ para la prueba basada en $^{-}C_{T-1}$ y $T = 46$ para la prueba basada en γ_{T-1}). Las potencias de las pruebas que utilizan pocas comparaciones (basadas en los estadísticos $^{-}C_1$, γ_1 y γ_2) presentan un crecimiento muy lento (sin llegar al 30%). Las potencias de la prueba con γ_5 se distinguen por acercarse con rapidez a las potencias de las pruebas basadas en ρ , $^{-}C_{T-1}$ y γ_{T-1} ; si bien en ese punto parece ser equidistante a los dos grupos descritos, al menos entre tamaños de sucesión 15 a 35.

Una tipología similar a la del caso 3 se observa en los casos 2 y 4, resaltando los siguientes aspectos:

Tabla resumen 1: Consistencia contra algunas alternativas de las pruebas con distribución de Laplace y $\alpha = 5\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	7.47%	7.08%	6.89%	8.03%	8.03%	9.43%	10.27%
10	6.85%	7.04%	7.50%	9.41%	10.47%	13.43%	20.42%
15	6.34%	6.83%	7.83%	10.50%	13.25%	20.99%	30.32%
20	7.09%	6.48%	7.47%	10.48%	17.11%	26.18%	42.48%
25	7.15%	6.69%	7.73%	11.48%	21.39%	34.77%	54.39%
30	7.81%	6.54%	7.47%	11.50%	26.70%	44.38%	65.29%
35	7.38%	6.29%	8.02%	11.61%	31.78%	51.20%	74.40%
40	7.38%	6.44%	7.89%	12.05%	37.22%	59.04%	81.98%
45	7.16%	6.45%	7.45%	12.18%	41.35%	65.93%	87.53%
50	6.64%	6.27%	8.35%	11.42%	41.68%	71.05%	91.63%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	12.89%	11.99%	11.93%	15.00%	15.00%	19.86%	22.40%
10	15.47%	13.95%	17.33%	28.82%	34.83%	47.77%	65.19%
15	15.35%	13.76%	20.14%	39.78%	59.32%	79.85%	91.48%
20	17.08%	13.43%	21.40%	47.83%	80.95%	94.16%	99.13%
25	17.91%	13.89%	23.54%	56.34%	92.81%	99.07%	99.96%
30	19.18%	14.04%	23.52%	60.53%	98.01%	99.93%	99.99%
35	18.75%	13.84%	26.15%	65.08%	99.53%	99.99%	100.00%
40	18.87%	14.08%	26.11%	69.11%	99.94%	100.00%	100.00%
45	18.47%	14.02%	25.92%	72.61%	99.99%	100.00%	100.00%
50	17.55%	14.20%	28.61%	73.55%	99.95%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	28.57%	27.19%	27.22%	34.47%	34.47%	45.71%	50.95%
10	47.97%	41.84%	56.92%	82.88%	88.72%	95.99%	99.15%
15	58.21%	50.17%	74.07%	97.31%	99.76%	100.00%	100.00%
20	67.32%	56.28%	84.52%	99.60%	100.00%	100.00%	100.00%
25	74.12%	62.32%	90.71%	99.95%	100.00%	100.00%	100.00%
30	79.45%	66.93%	94.02%	99.99%	100.00%	100.00%	100.00%
35	82.63%	70.65%	96.67%	100.00%	100.00%	100.00%	100.00%
40	85.41%	73.62%	97.75%	100.00%	100.00%	100.00%	100.00%
45	87.55%	76.89%	98.64%	100.00%	100.00%	100.00%	100.00%
50	88.60%	79.25%	99.16%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	46.40%	45.03%	45.06%	54.19%	54.19%	68.14%	73.57%
10	80.22%	73.53%	88.68%	98.76%	99.50%	99.93%	100.00%
15	92.19%	87.66%	98.04%	100.00%	100.00%	100.00%	100.00%
20	97.48%	94.48%	99.74%	100.00%	100.00%	100.00%	100.00%
25	99.28%	97.74%	99.98%	100.00%	100.00%	100.00%	100.00%
30	99.78%	99.11%	100.00%	100.00%	100.00%	100.00%	100.00%
35	99.93%	99.62%	100.00%	100.00%	100.00%	100.00%	100.00%
40	99.99%	99.89%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.99%	99.96%	100.00%	100.00%	100.00%	100.00%	100.00%
50	100.00%	99.98%	100.00%	100.00%	100.00%	100.00%	100.00%

Caso 2: $a^2(t) = t^{2/5}$. Las potencias de la prueba con γ_5 empiezan a alejarse de las del grupo que emplea poca información, pero más lentamente que en el caso 3. Las pruebas con estadísticos que utilizan todas las comparaciones empiezan a alcanzar la potencia máxima.

Caso 4: $a^4(t) = t^{3/5}$. Hay mayor claridad en el acercamiento de las potencias de la prueba con γ_5 a las de las pruebas que utilizan todas las comparaciones, que cada vez logran el máximo en un tamaño de muestra más pequeño.

Caso 6: $a^6(t) = t^{4/5}$. Las potencias de las pruebas que emplean pocas comparaciones (basadas en los estadísticos ${}^{-}C_1$, γ_1 y γ_2) aumentan más rápido que en los casos 1 a 4; por ejemplo la prueba menos potente que es la basada en γ_1 alcanza una potencia de 62,3% para un tamaño de muestra relativamente pequeño, $T = 25$, potencia que nunca alcanza en los casos 1 a 4. Hay mayor cercanía de las potencias de las pruebas basadas en estadísticos que emplean todas las comparaciones y la prueba basada en γ_5 , que consigue la potencia máxima en $T = 31$, en tanto la prueba basada en γ_{T-1} la consigue para $T = 19$, la basada en ${}^{-}C_{T-1}$ la consigue para $T = 15$ y la prueba de Daniels la consigue para $T = 15$.

En el caso 5: $a^5(t) = t^{7/10}$, las curvas de las potencias estimadas son muy similares en su forma y en la velocidad de aumento, con la diferencia de que las potencias de las pruebas basadas en estadísticos que emplean pocas comparaciones exhiben un crecimiento más lento que en el caso 6.

Caso 8: $a^8(t) = t$. Con excepción de la prueba basada en γ_1 , todas las pruebas obtienen la potencia máxima, desde tamaños de muestra distintos: 10 para la prueba de Daniels, 12 para la prueba basada en ${}^{-}C_{T-1}$, 13 para la basada en γ_{T-1} , 15 para la basada en γ_5 , 30 para la basada en γ_2 y 46 para la basada en ${}^{-}C_1$. El mayor tamaño de muestra $T = 50$ se requiere para alcanzar la potencia máxima (99,98%) de la prueba basada en γ_1 .

En el caso 7: $a^7(t) = t^{9/10}$, las potencias crecen de manera similar al caso 8 aunque con más lentamente, empezando a verse que las potencias de las pruebas con ${}^{-}C_{T-1}$ y ρ se acercan bastante entre sí. De la misma forma las potencias de las pruebas con γ_5 y γ_{T-1} ; y las potencias de las pruebas que emplean pocas comparaciones se van acercando lentamente a las anteriores.

3.1.2. Potencia

Para tamaños de muestra pequeños hay una gran similitud entre las potencias de todas las pruebas. Para

$T = 5$, las pruebas basadas en estadísticos que usan el mismo número de comparaciones (γ_5 y γ_{T-1}) tienen potencias iguales; las pruebas basadas en estadísticos que usan pocas comparaciones (${}^{-}C_1$, γ_1 y γ_2) traslapan todas sus potencias.

Entre $T = 6$ y $T = 14$ se observan mayores diferencias entre las potencias para todos los casos, potencias que empiezan a acercarse nuevamente para $T = 15$, tamaño de muestra para el cual las pruebas basadas en estadísticos que utilizan más comparaciones alcanzan la potencia máxima en el caso 8. Las pruebas basadas en los dos estadísticos (${}^{-}C_1$ y γ_1) que emplean la comparación de pares sucesivos sólo alcanzan potencias de 87% y el 92%, respectivamente, las cuales alcanzan su máxima potencia para $T = 25$.

4. Conclusiones

Hay cinco puntos importantes para resaltar del estudio:

El comportamiento de las potencias exhibe una similitud entre pruebas que utilizan un número cercano de comparaciones que se deberá reflejar en la obtención de eficiencias relativas cercanas a uno, de acuerdo con la tipología descrita para cada caso de tendencia.

Si bien los estadísticos más sencillos presentan potencias bajas en los casos de tendencias más lentas, su bondad mejora hasta llegar a ser competitivos en los casos 6 a 8 (ver tabla 3), con la ventaja de que su cálculo es más rápido.

Las pruebas basadas en estadísticos que emplean mayor número de comparaciones entre las observaciones son más competitivos con la tradicional prueba de Daniels, manteniendo sus potencias un comportamiento análogo y muy cercano.

Es importante notar que para aumentar la potencia de la prueba basada en G_k se debe sospechar qué tipo de tendencia puede existir en la variable para definir así el mejor estadístico a emplear. En casos de tendencias muy lentas sería necesario considerar toda la información; cuando se sospecha una tendencia moderada se podría pensar en comparar únicamente 5 términos consecutivos y obtener resultados similares con mayor rapidez; si se puede suponer que hay una tendencia lineal, entonces con observar tres términos consecutivos sería suficiente y si la tendencia parece más rápida basta con examinar los pares sucesivos.

Tabla resumen 2: Potencias de las pruebas con distribución de Laplace y $\alpha=5$ para algunos tamaños de sucesión

T=5							
Caso	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
1	7.47%	7.08%	6.89%	8.03%	8.03%	9.43%	10.27%
2	10.19%	9.50%	9.40%	11.53%	11.53%	14.71%	16.39%
3	12.89%	11.99%	11.93%	15.00%	15.00%	19.86%	22.40%
4	16.69%	15.60%	15.58%	19.88%	19.88%	26.80%	30.37%
5	21.87%	20.61%	20.62%	26.32%	26.32%	35.33%	39.82%
6	28.57%	27.19%	27.22%	34.47%	34.47%	45.71%	50.95%
7	36.70%	35.28%	35.32%	43.70%	43.70%	56.98%	62.54%
8	46.40%	45.03%	45.06%	54.19%	54.19%	68.14%	73.57%
T=15							
Caso	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
1	6.34%	6.83%	7.83%	10.50%	13.25%	20.99%	30.32%
2	10.22%	9.80%	12.94%	22.65%	34.24%	53.37%	70.21%
3	15.35%	13.76%	20.14%	39.78%	59.32%	79.85%	91.48%
4	24.37%	20.81%	32.88%	64.37%	83.59%	95.67%	99.09%
5	38.88%	32.73%	52.46%	86.40%	96.75%	99.65%	99.99%
6	58.21%	50.17%	74.07%	97.31%	99.76%	100.00%	100.00%
7	78.18%	70.54%	90.72%	99.77%	99.99%	100.00%	100.00%
8	92.19%	87.66%	98.04%	100.00%	100.00%	100.00%	100.00%
T=25							
Caso	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
1	7.15%	6.69%	7.73%	11.48%	21.39%	34.77%	54.39%
2	11.59%	9.67%	13.72%	29.66%	64.87%	85.97%	96.61%
3	17.91%	13.89%	23.54%	56.34%	92.81%	99.07%	99.96%
4	29.61%	22.50%	42.09%	85.33%	99.69%	99.99%	100.00%
5	49.06%	38.17%	68.43%	98.26%	100.00%	100.00%	100.00%
6	74.12%	62.32%	90.71%	99.95%	100.00%	100.00%	100.00%
7	92.77%	85.81%	98.96%	100.00%	100.00%	100.00%	100.00%
8	99.28%	97.74%	99.98%	100.00%	100.00%	100.00%	100.00%

Finalmente, se encuentra evidencia para pensar que el denominador de ${}^{-}C_k$ incluye información importante al estadístico y mejora la prueba, pero los principales cambios están dados en términos del número de comparaciones entre pares de observaciones que se incluyan en la sucesión dicótoma, es decir del valor de k .¹⁶

Referencias

- [1] **Aiyar, R., Guillier C. & Albers, W.**, *Asymptotic Relative Efficiencies of Rank Tests for Trend Alternatives*. Journal of American Statistician Association **74** (179), 226–231.
- [2] **Corzo, J. A.**, *Una prueba de dispersión basada en secuencias*. Tesis de Maestría. Universidad Nacional de Colombia, 1984.
- [3] **Corzo, J. A.**, *Algunas propiedades distribucionales de estadísticas de rachas*. Reporte interno No. 25. Departamento de Matemáticas y Estadística. Universidad Nacional de Colombia, Santa Fe de Bogotá.
- [4] **Corzo, J. A.**, *Análisis de datos a través de rachas*. Coloquio Distrital de Matemáticas y Estadística, Santa Fe de Bogotá, 1996.
- [5] **Corzo, J. A. & E. Gómez**, *Una prueba de rachas para alternativa de tendencia con muestras pequeñas*. Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales **24** (2000), 417–426.
- [6] **Gómez, E.** *Estimación de la potencia de una prueba de rachas para alternativa de tendencia*. Trabajo de grado. Universidad Nacional de Colombia, 1999.
- [7] **Gómez, E.** *Estudio de eficiencia relativa de una prueba de rachas para alternativa de tendencia*. Tesis de maestría. Universidad Nacional de Colombia, (2002).

¹⁶Este artículo contiene resultados parciales de la tesis de maestría de Gómez (2002).

Apéndice

Tabla 3. Consistencia de las pruebas con distribución logística y $\alpha=5\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	7.02%	6.68%	6.48%	7.45%	7.45%	8.65%	9.22%
10	6.46%	6.75%	7.04%	8.59%	9.37%	11.76%	17.58%
15	5.96%	6.53%	7.34%	9.52%	11.43%	17.93%	25.92%
20	6.63%	6.14%	6.99%	9.26%	14.47%	21.97%	36.45%
25	6.74%	6.40%	7.32%	10.31%	17.56%	28.84%	47.01%
30	7.40%	6.28%	6.98%	10.15%	21.98%	37.14%	57.35%
35	7.04%	6.05%	7.54%	10.33%	26.21%	43.04%	66.90%
40	7.00%	6.20%	7.45%	10.68%	30.62%	50.08%	75.19%
45	6.83%	6.22%	7.02%	10.68%	33.95%	57.01%	81.73%
50	6.30%	6.01%	7.86%	10.01%	34.65%	61.94%	86.98%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	11.49%	10.69%	10.61%	13.31%	13.31%	17.38%	19.22%
10	13.19%	12.17%	14.76%	24.16%	29.62%	41.77%	60.72%
15	12.93%	11.92%	17.00%	32.98%	52.38%	75.08%	90.14%
20	14.69%	11.70%	17.84%	39.98%	75.43%	92.04%	99.08%
25	15.32%	12.03%	19.53%	47.35%	89.84%	98.59%	99.97%
30	16.37%	12.18%	19.64%	51.30%	97.10%	99.89%	100.00%
35	15.88%	11.90%	21.56%	55.29%	99.26%	99.99%	100.00%
40	16.04%	12.25%	21.62%	59.31%	99.88%	100.00%	100.00%
45	15.81%	12.27%	21.45%	62.92%	99.99%	100.00%	100.00%
50	14.89%	12.19%	23.62%	63.53%	99.93%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	24.57%	23.21%	23.23%	30.30%	30.30%	41.12%	46.11%
10	41.17%	35.56%	50.48%	80.30%	87.25%	95.82%	99.46%
15	50.09%	42.45%	67.74%	96.76%	99.76%	100.00%	100.00%
20	58.91%	47.58%	78.88%	99.51%	100.00%	100.00%	100.00%
25	65.29%	52.49%	86.64%	99.94%	100.00%	100.00%	100.00%
30	71.37%	57.13%	90.54%	99.99%	100.00%	100.00%	100.00%
35	74.59%	60.89%	94.23%	100.00%	100.00%	100.00%	100.00%
40	77.64%	63.78%	96.01%	100.00%	100.00%	100.00%	100.00%
45	80.28%	66.80%	97.16%	100.00%	100.00%	100.00%	100.00%
50	81.06%	69.39%	98.14%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	41.72%	40.26%	40.29%	50.31%	50.31%	65.14%	71.19%
10	75.30%	68.01%	86.71%	98.86%	99.59%	99.96%	100.00%
15	89.39%	83.57%	97.68%	100.00%	100.00%	100.00%	100.00%
20	96.17%	91.67%	99.70%	100.00%	100.00%	100.00%	100.00%
25	98.66%	96.10%	99.96%	100.00%	100.00%	100.00%	100.00%
30	99.61%	98.28%	100.00%	100.00%	100.00%	100.00%	100.00%
35	99.85%	99.21%	100.00%	100.00%	100.00%	100.00%	100.00%
40	99.96%	99.71%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.98%	99.85%	100.00%	100.00%	100.00%	100.00%	100.00%
50	100.00%	99.94%	100.00%	100.00%	100.00%	100.00%	100.00%

Tabla 4. Consistencia de las pruebas con distribución normal y $\alpha=5\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	6.87%	6.56%	6.33%	7.35%	7.35%	8.36%	8.82%
10	6.40%	6.75%	6.76%	8.37%	9.00%	11.13%	16.54%
15	6.06%	6.60%	7.15%	9.29%	10.91%	17.03%	23.78%
20	6.82%	6.36%	7.21%	9.29%	13.72%	20.39%	33.45%
25	6.58%	6.21%	7.11%	9.74%	16.21%	26.61%	43.83%
30	7.36%	6.18%	7.05%	9.86%	20.41%	34.24%	53.62%
35	6.92%	6.08%	7.35%	9.75%	24.05%	39.29%	62.49%
40	6.96%	6.18%	7.46%	10.12%	28.39%	46.38%	70.85%
45	6.70%	6.08%	6.99%	10.31%	30.57%	52.58%	77.79%
50	6.15%	5.99%	7.66%	9.64%	31.69%	57.58%	83.77%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	10.95%	10.20%	10.10%	12.68%	12.68%	16.30%	17.90%
10	12.26%	11.32%	13.53%	21.90%	27.02%	38.84%	57.77%
15	12.46%	11.64%	15.89%	30.68%	48.43%	71.76%	89.06%
20	13.88%	11.31%	16.87%	36.58%	71.64%	90.37%	98.94%
25	14.12%	11.40%	17.94%	42.99%	87.60%	98.29%	99.96%
30	15.46%	11.54%	18.33%	47.00%	96.31%	99.85%	100.00%
35	15.01%	11.46%	19.91%	50.95%	99.00%	99.99%	100.00%
40	14.91%	11.34%	19.95%	55.03%	99.85%	100.00%	100.00%
45	14.52%	11.34%	19.53%	58.03%	99.98%	100.00%	100.00%
50	13.72%	11.36%	21.82%	58.56%	99.88%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	22.53%	21.16%	21.17%	28.10%	28.10%	38.66%	43.25%
10	37.45%	31.82%	46.49%	78.39%	86.18%	95.70%	99.67%
15	45.91%	38.48%	63.97%	96.39%	99.80%	100.00%	100.00%
20	54.60%	43.55%	75.32%	99.45%	100.00%	100.00%	100.00%
25	60.61%	47.96%	83.27%	99.92%	100.00%	100.00%	100.00%
30	66.84%	52.46%	88.09%	99.99%	100.00%	100.00%	100.00%
35	70.24%	56.05%	92.29%	100.00%	100.00%	100.00%	100.00%
40	73.43%	58.87%	94.56%	100.00%	100.00%	100.00%	100.00%
45	75.69%	61.86%	95.93%	100.00%	100.00%	100.00%	100.00%
50	76.86%	64.11%	97.48%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	38.42%	36.87%	36.90%	47.54%	47.54%	62.98%	69.37%
10	71.97%	64.17%	85.16%	98.98%	99.66%	99.97%	100.00%
15	87.09%	80.16%	97.43%	99.99%	100.00%	100.00%	100.00%
20	95.05%	89.62%	99.64%	100.00%	100.00%	100.00%	100.00%
25	98.15%	94.74%	99.94%	100.00%	100.00%	100.00%	100.00%
30	99.32%	97.34%	99.99%	100.00%	100.00%	100.00%	100.00%
35	99.76%	98.78%	100.00%	100.00%	100.00%	100.00%	100.00%
40	99.91%	99.45%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.96%	99.74%	100.00%	100.00%	100.00%	100.00%	100.00%
50	99.99%	99.88%	100.00%	100.00%	100.00%	100.00%	100.00%

Tabla 5. Consistencia de las pruebas con distribución de Laplace y $\alpha=1\%$

Caso 1							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	2.11%	2.11%	2.22%	2.22%	2.22%	2.22%	2.23%
10	0.35%	2.81%	2.12%	2.73%	3.03%	2.46%	5.78%
15	1.83%	2.18%	2.17%	3.18%	4.04%	6.14%	10.42%
20	1.70%	2.08%	1.89%	3.06%	5.64%	8.34%	18.02%
25	1.65%	1.99%	2.04%	3.53%	7.54%	13.14%	27.10%
30	1.48%	1.86%	1.88%	3.39%	10.42%	20.24%	37.48%
35	1.64%	1.72%	2.07%	3.28%	13.45%	25.15%	47.70%
40	1.76%	1.77%	1.96%	3.50%	17.16%	32.20%	56.44%
45	1.82%	1.74%	1.80%	3.63%	19.62%	39.85%	66.16%
50	1.81%	1.67%	2.19%	3.12%	19.04%	44.86%	74.37%
Caso 3							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	5.41%	5.41%	5.75%	5.75%	5.75%	5.75%	5.84%
10	1.64%	6.53%	6.27%	12.10%	15.13%	17.12%	34.35%
15	6.31%	5.33%	7.46%	19.45%	33.72%	52.81%	71.72%
20	5.96%	5.46%	7.79%	25.01%	59.07%	78.98%	94.20%
25	5.89%	5.31%	8.70%	32.19%	79.66%	94.79%	99.42%
30	5.56%	5.21%	8.73%	35.78%	92.82%	99.32%	99.95%
35	5.97%	5.01%	10.02%	40.27%	97.84%	99.91%	100.00%
40	6.23%	4.87%	9.93%	44.70%	99.50%	100.00%	100.00%
45	6.44%	4.94%	9.72%	48.91%	99.91%	100.00%	100.00%
50	6.41%	4.74%	11.32%	49.16%	99.51%	100.00%	100.00%
Caso 6							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	17.72%	17.72%	18.60%	18.60%	18.60%	18.60%	18.78%
10	14.65%	26.64%	34.59%	62.46%	70.30%	77.93%	93.16%
15	38.46%	32.43%	52.76%	90.82%	98.27%	99.86%	99.98%
20	45.71%	38.78%	66.52%	98.21%	99.98%	100.00%	100.00%
25	52.29%	43.94%	77.07%	99.75%	100.00%	100.00%	100.00%
30	56.25%	48.32%	83.37%	99.96%	100.00%	100.00%	100.00%
35	63.01%	52.04%	89.53%	99.99%	100.00%	100.00%	100.00%
40	68.01%	55.58%	92.45%	100.00%	100.00%	100.00%	100.00%
45	72.03%	59.10%	94.60%	100.00%	100.00%	100.00%	100.00%
50	75.07%	62.04%	96.73%	100.00%	100.00%	100.00%	100.00%
Caso 8							
T	C1	gama1	gama2	gama5	gama(T-1)	C(T-1)	Daniels
5	33.92%	33.92%	35.23%	35.23%	35.23%	35.23%	35.26%
10	44.54%	58.59%	73.44%	94.11%	96.51%	98.42%	99.82%
15	82.01%	77.11%	93.47%	99.94%	100.00%	100.00%	100.00%
20	92.12%	88.09%	98.83%	100.00%	100.00%	100.00%	100.00%
25	96.90%	94.27%	99.81%	100.00%	100.00%	100.00%	100.00%
30	98.60%	97.23%	99.98%	100.00%	100.00%	100.00%	100.00%
35	99.60%	98.65%	99.99%	100.00%	100.00%	100.00%	100.00%
40	99.88%	99.48%	100.00%	100.00%	100.00%	100.00%	100.00%
45	99.96%	99.77%	100.00%	100.00%	100.00%	100.00%	100.00%
50	99.99%	99.91%	100.00%	100.00%	100.00%	100.00%	100.00%

LEVERRIER-FADEEV ALGORITHM AND CLASSICAL ORTHOGONAL POLYNOMIALS

by

J. Hernández, F. Marcellán & C. Rodríguez ¹

To Professor Jairo Charris Castañeda, as a tribute of our mathematical friendship

Resumen

Hernández, J., F. Marcellán & C. Rodríguez: Leverrier–Fadeev Algorithm and Classical Orthogonal Polynomials. *Rev. Acad. Colomb. Cienc.* **28** (106): 39–47, 2004. ISSN 0370-3908.

Usando propiedades estructurales de los polinomios ortogonales clásicos (Hermite, Laguerre, Jacobi y Bessel), se implementa el algoritmo de Leverrier-Fadeev para obtener el polinomio característico de una matriz cuadrada de elementos complejos.

Palabras clave: Polinomio característico, funciones de transferencia, polinomios ortogonales, funcionales lineales clásicos.

Abstract

Using structural properties of classical orthogonal polynomials (Hermite, Laguerre, Jacobi, and Bessel), an implementation of Leverrier-Fadeev algorithm to obtain the characteristic polynomial of a square matrix with complex entries is presented.

Key words: Characteristic Polynomial, Transfer Functions, Orthogonal Polynomials, Classical Linear Functionals.

1. Introduction

For a given matrix $A \in \mathbb{C}^{n \times n}$ an algorithm attributed to Leverrier, Fadeev, and others, allows the simultaneous determination of the characteristic polynomial of A and the adjoint matrix of $sI - A$, where I denotes

the identity matrix in $\mathbb{C}^{n \times n}$. Indeed, if

$$p(s) = \det(sI - A) = s^n + \sum_{k=0}^{n-1} a_{n-k} s^k$$

denotes the characteristic polynomial of A and

¹Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida de la Universidad 30, 28911, Leganés, Spain. jhbenite@math.uc3m.es, pacomarc@ing.uc3m.es, teijeiro@math.uc3m.es

AMS Classification 2000: Primary 42C05, 33C45.

$$\tilde{A}(s) = \text{Adj}(sI - A) = s^{n-1}I + \sum_{k=0}^{n-2} s^k B_{n-k-1}$$

denotes the adjoint matrix of $sI - A$, and taking into account

$$\tilde{A}(s) = p(s)(sI - A)^{-1},$$

then the coefficients (a_k) and the matrices (B_k) can be generated from

$$\begin{aligned} a_1 &= -\text{tr } A, & B_1 &= A + a_1 I, \\ a_k &= -\frac{1}{k} \text{tr}(AB_{k-1}), & B_k &= a_k I + AB_{k-1}, \end{aligned} \quad (1.1)$$

for $k = 2, \dots, n-1$. Here $\text{tr } A$ denotes the trace of the matrix A .

Notice that (1.1) can be read as follows (See [3])

$$\begin{cases} (sI - A)\tilde{A}(s) = p(s)I, \\ \frac{dp(s)}{ds} = \text{tr } \tilde{A}(s). \end{cases} \quad (1.2)$$

Despite the little value from a numerical point of view, this algorithm is useful for theoretical purposes as well as for the applications in linear control theory. More precisely, $\frac{1}{p(s)}\tilde{A}(s)$ is the transfer function of a continuous time linear system with n inputs and n outputs.

The algorithm takes into account the representation of the characteristic polynomial and the adjoint matrix in terms of the canonical basis $\{s^k\}_{k=0}^n$ in the linear space of polynomials with complex coefficients and degree at most n .

From a computational point of view the accuracy of the algorithm using an orthogonal polynomial system is improved. For some particular cases of orthogonal polynomials S. Barnett [1] gave an implementation of the algorithm. The key idea is the relation (1.2) as well as the expression of the derivative of the polynomial P_k , $k = 1, \dots, n$, in terms of the family $\{P_k\}_{k=0}^n$. The aim of our contribution is to present a general approach for families of classical orthogonal polynomials (Hermite, Laguerre, Jacobi, and Bessel) taking into account a characterization of such families obtained in [4]. Indeed it allows to give an expression of P_k as a linear combination of P'_{k+1} , P'_k , and P'_{k-1} . Thus we can show a very simple implementation of the Leverrier algorithm, where parameters associated with the three-term recurrence relation play the main role.

The structure of the paper is the following. In the section 2 we summarize the basic properties of classical orthogonal polynomials. In the section 3 we present the adapted version of Leverrier algorithm for bases of classical orthogonal polynomials, and we analyze it for each family of classical orthogonal polynomials. In the section 4, some examples are tested.

2. Classical Orthogonal Polynomials

Let u be a linear functional in the linear space \mathbb{P} of polynomials with complex coefficients. If $\langle \cdot, \cdot \rangle$ denotes the duality bracket then $c_n = \langle u, x^n \rangle$ is said to be the moment of order n associated with the linear functional u .

The linear functional u is said to be quasi-definite [2] if the principal submatrices of the Hankel matrix $H = (c_{i+j})_{i,j=0}^{\infty}$ are non-singular. In such a case, there exists a unique sequence of monic polynomials $\{P_n\}_{n=0}^{\infty}$ such that

- (i) $\langle u, x^k P_n \rangle = 0$, $k = 0, 1, \dots, n-1$.
- (ii) $\langle u, x^n P_n \rangle \neq 0$.
- (iii) $\deg P_n = n$.

The sequence $\{P_n\}_{n=0}^{\infty}$ is said to be a sequence of monic orthogonal polynomials (SMOP) with respect to u . It is very well known that $\{P_n\}_{n=0}^{\infty}$ satisfies a three-term recurrence relation

$$xP_n(x) = P_{n+1}(x) + \beta_n P_n(x) + \gamma_n P_{n-1}(x), \quad (2.1)$$

$n = 1, 2, \dots$ with $\gamma_n \neq 0$.

The converse is also true and this result is due to several authors despite the fact is known as Favard's Theorem [2].

If $q(x)$ denotes a polynomial, then a new linear functional $\tilde{u} = q(x)u$ can be introduced as follows

$$\langle \tilde{u}, p(x) \rangle = \langle u, p(x)q(x) \rangle \quad (2.2)$$

for every $p \in \mathbb{P}$.

On the other hand, as for a distribution, the derivative of the linear functional u, Du , is given by $\langle Du, p(x) \rangle = -\langle u, p'(x) \rangle$, $p \in \mathbb{P}$.

Definition 2.1. A linear functional u is said to be classical if there exist polynomials ϕ, ψ , with $\deg \phi \leq 2$ and $\deg \psi = 1$ such that

$$D(\phi u) = \psi u. \quad (2.3)$$

Up to a linear change of variables, four cases appear

- (i) $\phi(x) = 1$. This leads to Hermite linear functional with $\psi(x) = -2x$.
- (ii) $\phi(x) = x$. This leads to Laguerre linear functional with $\psi(x) = -x + \alpha + 1$.
- (iii) $\phi(x) = x^2 - 1$. This yields the Jacobi linear functional with $\psi(x) = -(\alpha + \beta + 2)x + \beta - \alpha$.
- (iv) $\phi(x) = x^2$. This yields the Bessel linear functional with $\psi(x) = (\alpha + 2)x + 2$.

Theorem 2.2.(see [4]) *If $\{P_n\}_{n=0}^\infty$ is the SMOP associated with u , then the following statements are equivalent*

- (i) u is a classical linear functional.
- (ii) $\{Q_n\}_{n=0}^\infty$, with $Q_n = \frac{P'_{n+1}}{n+1}$, is a SMOP.
- (iii) $P_n = Q_n + r_n Q_{n-1} + s_n Q_{n-2}$.
- (iv) $\phi(x)Q_n = a_n P_{n+2} + b_n P_{n+1} + c_n P_n$, with $c_n \neq 0$.

TABLE 1. Coefficients in the three-term recurrence relation (2.1)

	β_n	γ_n
Hermite	0	$\frac{n}{2}$
Laguerre	$2n + \alpha + 1$	$n(n + \alpha)$
Jacobi	$\frac{\beta^2 - \alpha^2}{(2n + \alpha + \beta)(2n + \alpha + \beta + 2)}$	$\frac{4n(n + \alpha)(n + \beta)(n + \alpha + \beta)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta)^2(2n + \alpha + \beta + 1)}$
Bessel	$-\frac{2\alpha}{(2n + \alpha)(2n + \alpha + 2)}$	$-\frac{4n(n + \alpha)}{(2n + \alpha - 1)(2n + \alpha)^2(2n + \alpha + 1)}$

TABLE 2. Coefficients in the relation of the Theorem 2.2 (iii)

	r_n	s_n
Hermite	0	0
Laguerre	n	0
Jacobi	$\frac{2n(\alpha - \beta)}{(2n + \alpha + \beta)(2n + \alpha + \beta + 2)}$	$-\frac{4n(n - 1)(n + \alpha)(n + \beta)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta)^2(2n + \alpha + \beta + 1)}$
Bessel	$\frac{4n}{(2n + \alpha)(2n + \alpha + 2)}$	$\frac{4n(n - 1)}{(2n + \alpha - 1)(2n + \alpha)^2(2n + \alpha + 1)}$

3. Leverrier-Fadeev Algorithm

Let $\{P_n\}_{n=0}^\infty$ be a sequence of monic orthogonal polynomials. If we expand the characteristic polynomial $p(s)$ and the adjoint matrix $\tilde{A}(s)$ of a matrix $A \in \mathbb{C}^{n \times n}$ in terms of the above basis in the linear space of polynomials with complex coefficients, then we get:

$$p(s) = P_n(s) + \sum_{k=0}^{n-1} \hat{a}_{n-k} P_k(s), \quad (3.1)$$

$$\tilde{A}(s) = P_{n-1}(s)I + \sum_{k=0}^{n-2} P_k(s)\hat{B}_{n-k-1}. \quad (3.2)$$

From the first identity in (1.2)

$$(sI - A) \left(P_{n-1}(s)I + \sum_{k=0}^{n-2} P_k(s)\hat{B}_{n-k-1} \right) = P_n(s)I + \sum_{k=0}^{n-1} \hat{a}_{n-k}P_k(s)I. \quad (3.3)$$

Taking into account the three-term recurrence relation (2.1) for the family $\{P_n\}_{n=0}^\infty$, (3.3) becomes

$$P_n(s)I + \sum_{k=0}^{n-1} \hat{a}_{n-k}P_k(s)I = [P_n(s) + \beta_{n-1}P_{n-1} + \gamma_{n-1}P_{n-2}]I - P_{n-1}(s)A + \sum_{k=0}^{n-2} (P_{k+1}(s) + \beta_k P_k(s) + \gamma_k P_{k-1}(s)) \hat{B}_{n-k-1} - \sum_{k=0}^{n-2} P_k(s)A\hat{B}_{n-k-1}.$$

Equating coefficients of P_k in the previous expression we get

$$\begin{aligned} A\hat{B}_0 &= -\hat{a}_1 I + \beta_{n-1}\hat{B}_0 + \hat{B}_1, \\ A\hat{B}_1 &= -\hat{a}_2 I + \gamma_{n-1}\hat{B}_0 + \beta_{n-2}\hat{B}_1 + \hat{B}_2, \\ &\vdots \\ A\hat{B}_{n-k-1} &= -\hat{a}_{n-k} I + \gamma_{k+1}\hat{B}_{n-k-2} + \beta_k\hat{B}_{n-k-1} + \hat{B}_{n-k}, \quad k = 1, 2, \dots, n-3, \\ A\hat{B}_{n-1} &= -\hat{a}_n I + \gamma_1\hat{B}_{n-2} + \beta_0\hat{B}_{n-1}, \end{aligned} \quad (3.4)$$

with $\hat{B}_0 = I$. In a matrix form

$$A \begin{bmatrix} \hat{B}_{n-1} \\ \vdots \\ \hat{B}_0 \end{bmatrix} = M \begin{bmatrix} \hat{B}_{n-1} \\ \vdots \\ \hat{B}_0 \end{bmatrix}$$

where $M = J_n - [0|\hat{a}]$. J_n is the Jacobi matrix of dimension n associated with the SMOP $\{P_n\}_{n=0}^\infty$ i. e.

$$J_n = \begin{bmatrix} \beta_0 & \gamma_1 & 0 & \cdots & 0 \\ 1 & \beta_1 & \gamma_2 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & & & \gamma_{n-1} \\ 0 & \cdots & 0 & 1 & \beta_{n-1} \end{bmatrix}$$

and

$$\hat{a} = \begin{bmatrix} \hat{a}_n \\ \hat{a}_{n-1} \\ \vdots \\ \hat{a}_1 \end{bmatrix}.$$

In the literature, the matrix M is called the comrade matrix of A with respect to the orthogonal system $\{P_n\}_{n=0}^\infty$. His characteristic polynomial is $p(s)$. In particular, we get

$$\text{tr } A = -\sum_{j=0}^{n-1} \beta_j - \hat{a}_1.$$

On the other hand, from the second relation in (1.2) for $n = 2, 3, \dots$ we have

$$P'_n(s) + \sum_{k=0}^{n-1} \hat{a}_{n-k}P'_k(s) = nP_{n-1}(s) + \sum_{k=0}^{n-2} P_k(s)\text{tr } \hat{B}_{n-k-1}. \quad (3.5)$$

If $\{P_n\}_{n=0}^\infty$ is a classical family then, from theorem (iii), we get

$$P_k(s) = \frac{P'_{k+1}(s)}{k+1} + r_k \frac{P'_k(s)}{k} + s_k \frac{P'_{k-1}(s)}{k-1}, \quad k = 2, 3, \dots$$

Thus, substitution in (3.5) yields

$$\begin{aligned} P'_n(s) + \sum_{k=0}^{n-1} \hat{a}_{n-k} P'_k(s) &= P'_n(s) + r_{n-1} \frac{n}{n-1} P'_{n-1}(s) + s_{n-1} \frac{n}{n-2} P'_{n-2}(s) + \\ &+ \sum_{k=2}^{n-2} \left(\frac{P'_{k+1}(s)}{k+1} + r_k \frac{P'_k(s)}{k} + s_k \frac{P'_{k-1}(s)}{k-1} \right) \text{tr } \hat{B}_{n-k-1} + \\ &+ \text{tr } \hat{B}_{n-1} P'_1(s) + \text{tr } \hat{B}_{n-2} \left(\frac{P'_2(s)}{2} + r_1 P'_1(s) \right). \end{aligned}$$

Finally, equating the coefficients of P'_k in both hand sides we get

$$\begin{aligned} (n-1)\hat{a}_1 &= nr_{n-1} + \text{tr } \hat{B}_1, \\ (n-2)\hat{a}_2 &= ns_{n-1} + r_{n-2} \text{tr } \hat{B}_1 + \text{tr } \hat{B}_2, \\ &\vdots \\ k\hat{a}_{n-k} &= s_{k+1} \text{tr } \hat{B}_{n-k-2} + r_k \text{tr } \hat{B}_{n-k-1} + \text{tr } \hat{B}_{n-k}, \quad k = 1, 2, \dots, n-3. \end{aligned} \tag{3.6}$$

Thus, in order to obtain (\hat{a}_k) and (\hat{B}_k) we will proceed as follows.

First Step

$$\hat{a}_1 = n(\beta_{n-1} - r_{n-1}) - \text{tr } A. \tag{3.7}$$

Indeed, taking traces in the first equation of (3.4), and (3.6)

$$\begin{cases} \text{tr } A &= -n\hat{a}_1 + n\beta_{n-1} + \text{tr } \hat{B}_1, \\ (n-1)\hat{a}_1 &= nr_{n-1} + \text{tr } \hat{B}_1, \end{cases}$$

and (3.7) follows.

Second Step

$$\hat{B}_1 = A\hat{B}_0 + \hat{a}_1 I - \beta_{n-1} \hat{B}_0. \tag{3.8}$$

Third Step

$$\begin{aligned} 2\hat{a}_2 &= (\gamma_{n-1} - s_{n-1}) \text{tr } \hat{B}_0 + \\ &(\beta_{n-2} - r_{n-2}) \text{tr } \hat{B}_1 - \text{tr } (A\hat{B}_1). \end{aligned} \tag{3.9}$$

Indeed, from the second equation in (3.4) and (3.6)

$$\begin{cases} \text{tr } (A\hat{B}_1) &= n(\gamma_{n-1} - \hat{a}_2) + \beta_{n-2} \text{tr } \hat{B}_1 + \text{tr } \hat{B}_2, \\ \text{tr } \hat{B}_2 &= (n-2)\hat{a}_2 - ns_{n-1} - r_{n-2} \text{tr } \hat{B}_1, \end{cases}$$

and (3.9) follows.

Fourth Step

$$\hat{B}_2 = A\hat{B}_1 + \hat{a}_2 I - \gamma_{n-1} \hat{B}_0 - \beta_{n-2} \hat{B}_1. \tag{3.10}$$

Thus, for $k = 1, 2, \dots, n-3$,

$$\begin{aligned} (n-k)\hat{a}_{n-k} &= (\beta_k - r_k) \text{tr } \hat{B}_{n-k-1} + \\ &(\gamma_{k+1} - s_{k+1}) \text{tr } \hat{B}_{n-k-2} - \text{tr } (A\hat{B}_{n-k-1}), \end{aligned} \tag{3.11}$$

as well as

$$\hat{B}_{n-k} = A\hat{B}_{n-k-1} + \hat{a}_{n-k} I - \gamma_{k+1} \hat{B}_{n-k-2} - \beta_k \hat{B}_{n-k-1}.$$

These results follow from the expressions in (3.4) and (3.6) for $k = 1, \dots, n-3$.

Finally, taking traces in the last equation of (3.4) we get

$$n\hat{a}_n = \beta_0 \text{tr } \hat{B}_{n-1} + \gamma_1 \text{tr } \hat{B}_{n-2} - \text{tr } (A\hat{B}_{n-1}).$$

As a conclusion we get

Theorem 3.1.

(i) For $k = 0, 1, \dots, n-1$,

$$\begin{aligned} (n-k)\hat{a}_{n-k} &= (\beta_k - r_k) \text{tr } \hat{B}_{n-k-1} + \\ &(\gamma_{k+1} - s_{k+1}) \text{tr } \hat{B}_{n-k-2} - \text{tr } (A\hat{B}_{n-k-1}), \end{aligned} \tag{3.12}$$

with the convention $\hat{B}_{-1} = 0$, $r_0 = 0$, $s_1 = 0$.

(ii) For $k = 1, 2, \dots, n-1$

$$\begin{aligned} \hat{B}_{n-k} &= A\hat{B}_{n-k-1} + \hat{a}_{n-k} I - \\ &\gamma_{k+1} \hat{B}_{n-k-2} - \beta_k \hat{B}_{n-k-1}. \end{aligned} \tag{3.13}$$

The implementation of the algorithm is as follows

DATA: $\{\beta_k\}_{k=0}^{n-1}$, $\{\gamma_k\}_{k=1}^n$, $\{r_k\}_{k=0}^{n-1}$, $\{s_k\}_{k=1}^n$.

Initial Condition: $\hat{B}_{-1} = 0$, $\hat{B}_0 = I$.

1. From \hat{B}_{n-k-2} and \hat{B}_{n-k-1} taking into account (3.12) we get \hat{a}_{n-k} .
2. From (3.13) we get \hat{B}_{n-k} .

END

and, for each family of monic orthogonal polynomials, is given in below

3.1. Hermite Case. According to Theorem 3.1 we get

$$(i) \quad (n-k)\hat{a}_{n-k} = \frac{k+1}{2}\text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right).$$

$$(ii) \quad \hat{B}_{n-k} = A\hat{B}_{n-k-1} + \hat{a}_{n-k}I - \frac{k+1}{2}\hat{B}_{n-k-2}. \quad (3.14)$$

In particular, taking traces in (ii) and using (i) we get

$$\text{tr} \hat{B}_{n-k} = k\hat{a}_{n-k}.$$

This is formula (3.12) in [1].

Furthermore, substituting in (i) we get

$$(n-k)\hat{a}_{n-k} = \frac{(k+1)(k+2)}{2}\hat{a}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right), \quad (3.15)$$

i.e.

$$\text{tr} \left(A\hat{B}_{n-k-1} \right) = \frac{(k+1)(k+2)}{2}\hat{a}_{n-k-2} - (n-k)\hat{a}_{n-k}.$$

3.3. Jacobi Case. According to Theorem 3.1 we get

$$(n-k)\hat{a}_{n-k} = \left(\frac{\beta^2 - \alpha^2}{(2k + \alpha + \beta)(2k + \alpha + \beta + 2)} - \frac{2k(\alpha - \beta)}{(2k + \alpha + \beta)(2k + \alpha + \beta + 2)} \right) \text{tr} \hat{B}_{n-k-1} +$$

$$\left(\frac{4(k+1)(k+1+\alpha)(k+1+\beta)(k+1+\alpha+\beta)}{(2k + \alpha + \beta + 1)(2k + \alpha + \beta + 2)^2(2k + \alpha + \beta + 3)} - \frac{4k(k+1)(k+1+\alpha)(k+1+\beta)}{(2k + \alpha + \beta + 1)(2k + \alpha + \beta + 2)^2(2k + \alpha + \beta + 3)} \right) \text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right) =$$

$$\frac{\beta - \alpha}{2k + \alpha + \beta + 2} \text{tr} \hat{B}_{n-k-1} + \frac{4(k+1)(k+1+\alpha)(k+1+\beta)}{(2k + \alpha + \beta + 2)^2(2k + \alpha + \beta + 3)} \text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right).$$

On the other hand, if $\alpha = \beta$ then we are in the Gegenbauer case. The linear functional is symmetric and thus get

$$(n-k)\hat{a}_{n-k} = \frac{4(k+1)(k+1+\alpha)^2}{(2k+2\alpha+2)^2(2k+2\alpha+3)} \text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right)$$

$$= \frac{k+1}{2k+2\alpha+3} \text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right), \quad (3.18)$$

3.2. Laguerre Case. According to Theorem 3.1 we get

$$(i) \quad (n-k)\hat{a}_{n-k} = [(2k + \alpha + 1) - k] \text{tr} \hat{B}_{n-k-1} + (k+1)(k+\alpha+1)\text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right)$$

i.e.

$$(n-k)\hat{a}_{n-k} = (k + \alpha + 1)\text{tr} \hat{B}_{n-k-1} + (k+1)(k + \alpha + 1)\text{tr} \hat{B}_{n-k-2} - \text{tr} \left(A\hat{B}_{n-k-1} \right)$$

$$(ii) \quad \hat{B}_{n-k} = A\hat{B}_{n-k-1} + \hat{a}_{n-k}I - (k+1)(k + \alpha + 1)\hat{B}_{n-k-2} - (2k + \alpha + 1)\hat{B}_{n-k-1} \quad (3.16)$$

Taking traces in (ii) and using (i) we get

$$\text{tr} \hat{B}_{n-k} = k\hat{a}_{n-k} - k\text{tr} \hat{B}_{n-k-1}.$$

Thus we deduce

$$(n-k)\hat{a}_{n-k} = (k + \alpha + 1)\text{tr} \hat{B}_{n-k-1} + (k + \alpha + 1) \left[(k+1)\hat{a}_{n-k-1} - \text{tr} \hat{B}_{n-k-1} \right] - \text{tr} \left(A\hat{B}_{n-k-1} \right),$$

i.e.

$$(n-k)\hat{a}_{n-k} = (k + \alpha + 1)(k+1)\hat{a}_{n-k-1} - \text{tr} \left(A\hat{B}_{n-k-1} \right). \quad (3.17)$$

Up to a normalization this is the formula (3.23b) in [1], when $\alpha = 0$.

or, equivalently

$$\operatorname{tr} \hat{B}_{n-k} = k\hat{a}_{n-k} + \frac{4k(k+1)(k+1+\alpha)^2}{(2k+2\alpha+1)(2k+2\alpha+2)^2(2k+2\alpha+3)} \operatorname{tr} \hat{B}_{n-k-2},$$

i.e.

$$\operatorname{tr} \hat{B}_{n-k} = k\hat{a}_{n-k} + \frac{k(k+1)}{(2k+2\alpha+1)(2k+2\alpha+3)} \operatorname{tr} \hat{B}_{n-k-2}.$$

Notice that the symmetry of the linear functional yields an important simplification in our algorithm.

Furthermore

$$\begin{aligned} \hat{B}_{n-k} &= A\hat{B}_{n-k-1} + \hat{a}_{n-k}I - \frac{4(k+1)(k+\alpha+1)^2(k+2\alpha+1)}{(2k+2\alpha+1)(2k+2\alpha+2)^2(2k+2\alpha+3)} \hat{B}_{n-k-2} \\ &= A\hat{B}_{n-k-1} + \hat{a}_{n-k}I - \frac{(k+1)(k+2\alpha+1)}{(2k+2\alpha+1)(2k+2\alpha+3)} \hat{B}_{n-k-2}. \end{aligned} \quad (3.19)$$

This is, up to the corresponding normalization, the formula (3.20) in [1] for $\alpha = 0$.

3.4. Bessel Case. According to Theorem 3.1 we get

$$\begin{aligned} (n-k)\hat{a}_{n-k} &= \frac{-2\alpha-4k}{(2k+\alpha)(2k+\alpha+2)} \operatorname{tr} \hat{B}_{n-k-1} - \frac{4(k+1)(2k+\alpha+1)}{(2k+\alpha+1)(2k+\alpha+2)^2(2k+\alpha+3)} \operatorname{tr} \hat{B}_{n-k-2} - \operatorname{tr} (A\hat{B}_{n-k-1}) \\ &= \frac{-2}{2k+\alpha+2} \operatorname{tr} \hat{B}_{n-k-1} - \frac{4(k+1)}{(2k+\alpha+2)^2(2k+\alpha+3)} \operatorname{tr} \hat{B}_{n-k-2} - \operatorname{tr} (A\hat{B}_{n-k-1}), \end{aligned}$$

i.e.

$$(n-k)\hat{a}_{n-k} + \frac{1}{k+1+\frac{\alpha}{2}} \operatorname{tr} \hat{B}_{n-k-1} + \operatorname{tr} (A\hat{B}_{n-k-1}) + \frac{k+1}{\left(k+1+\frac{\alpha}{2}\right)^2(2k+\alpha+3)} \operatorname{tr} \hat{B}_{n-k-2} = 0, \quad (3.20)$$

together with

$$\hat{B}_{n-k} = A\hat{B}_{n-k-1} + \hat{a}_{n-k}I + \frac{4(k+1)(k+\alpha+1)}{(2k+\alpha+1)(2k+\alpha+2)^2(2k+\alpha+3)} \hat{B}_{n-k-2} + \frac{2\alpha}{(2k+\alpha)(2k+\alpha+2)} \hat{B}_{n-k-1}. \quad (3.21)$$

4. Example

Consider

$$A = \begin{bmatrix} 1 & -4 & -1 & -1 \\ 2 & 0 & 5 & -4 \\ -1 & 1 & -2 & 3 \\ -1 & 4 & -1 & 6 \end{bmatrix}$$

which has characteristic polynomial

$$a(s) = s^4 - 5s^3 + 9s^2 - 7s + 2.$$

We apply the algorithm for each basis.

4.1. Hermite Basis. From (3.15), $a_1 = -\operatorname{tr} A = -5$, and from (3.14)

$$B_1 = a_1I + A = \begin{bmatrix} -4 & -4 & -1 & -4 \\ 2 & -5 & 5 & -4 \\ -1 & 1 & -7 & 3 \\ -1 & 4 & -1 & 1 \end{bmatrix}.$$

Using (3.15), we get

$$a_2 = 3 - \frac{1}{2} \operatorname{tr} (AB_1) = 12,$$

and from (3.14)

$$B_2 = a_2I - \frac{3}{2}B_0 + AB_1 = \begin{bmatrix} \frac{7}{2} & -1 & -10 & 5 \\ -9 & -\frac{17}{2} & -33 & 3 \\ 5 & 9 & \frac{55}{2} & -3 \\ 7 & 7 & 22 & \frac{3}{2} \end{bmatrix}.$$

Using again (3.15)

$$a_3 = a_1 - \frac{1}{3} \operatorname{tr} (AB_2) = -\frac{29}{2},$$

and from (3.14)

$$B_3 = a_3 I - B_1 + AB_2 = \frac{1}{2} \begin{bmatrix} -8 & 0 & 15 & -12 \\ 4 & 11 & 49 & -14 \\ -1 & -11 & -39 & 11 \\ -3 & -8 & -33 & 7 \end{bmatrix}.$$

Finally

$$a_4 = \frac{1}{4}a_2 - \frac{1}{4}\text{tr}(AB_3) = \frac{29}{4}.$$

Hence, the characteristic polynomial of A is given by (3.1) as

$$a(s) = H_4(s) - 5H_3(s) + 12H_2(s) - \frac{29}{2}H_1(s) + \frac{29}{4}H_0(s).$$

4.2, Laguerre Basis. We consider the family $\{L_n^0\}_{n=0}^\infty$. From (3.17), $a_1 = 16 - \text{tr} A = 11$, and from (3.16)

$$B_1 = a_1 I - 7B_0 + A = \begin{bmatrix} 5 & -4 & -1 & -4 \\ 2 & 4 & 5 & -4 \\ -1 & 1 & 2 & 3 \\ -1 & 4 & -1 & 10 \end{bmatrix}.$$

Using (3.17), we get

$$a_2 = \frac{9}{2}a_1 - \frac{1}{2}\text{tr}(AB_1) = 36,$$

and from (3.16)

$$\begin{aligned} B_2 &= a_2 I - 9B_0 - 5B_1 + AB_1 \\ &= \begin{bmatrix} 4 & -17 & -14 & -11 \\ -1 & -12 & -13 & -13 \\ 1 & 13 & 16 & 9 \\ 3 & 23 & 18 & 22 \end{bmatrix}. \end{aligned}$$

Using again (3.17)

$$a_3 = \frac{4}{3}a_2 - \frac{1}{3}\text{tr}(AB_2) = 35,$$

and from (3.16)

$$\begin{aligned} B_3 &= a_3 I - 4B_1 - 3B_2 + AB_2 \\ &= \begin{bmatrix} -2 & -7 & -4 & -7 \\ -4 & -6 & -1 & -10 \\ 3 & 5 & 2 & 7 \\ 4 & 9 & 4 & 11 \end{bmatrix}. \end{aligned}$$

Finally

$$a_4 = \frac{1}{4}a_3 - \frac{1}{4}\text{tr}(AB_3) = 7.$$

Hence, the characteristic polynomial of A is given by (3.1) as

$$a(s) = L_4^0(s) + 11L_3^0(s) + 36L_2^0(s) + 35L_1^0(s) + 7L_0^0(s).$$

4.3. Jacobi Basis. We consider the family $P_n = P_n^{(0,0)}$ (Legendre Polynomials). From (3.18), $a_1 = -\text{tr} A = -5$, and from (3.19)

$$B_1 = a_1 I + A = \begin{bmatrix} -4 & -4 & -1 & -4 \\ 2 & -5 & 5 & -4 \\ -1 & 1 & -7 & 3 \\ -1 & 4 & -1 & 1 \end{bmatrix}.$$

Using (3.18), we get

$$a_2 = \frac{6}{7} - \frac{1}{2}\text{tr}(AB_1) = \frac{69}{7},$$

and from (3.19)

$$B_2 = a_2 I - \frac{9}{35}B_0 + AB_1 = \begin{bmatrix} \frac{13}{5} & -1 & -10 & 5 \\ -9 & -\frac{47}{5} & -33 & 3 \\ 5 & 9 & \frac{133}{5} & -3 \\ 7 & 7 & 22 & \frac{3}{5} \end{bmatrix}.$$

Using again (3.18)

$$a_3 = \frac{2}{15}\text{tr} B_1 - \frac{1}{3}\text{tr}(AB_2) = -10,$$

and from (3.19)

$$\begin{aligned} B_3 &= a_3 I - \frac{4}{15}B_1 + AB_2 \\ &= \frac{1}{3} \begin{bmatrix} -10 & 2 & 23 & -16 \\ 5 & 19 & 71 & -19 \\ -1 & -17 & -55 & 15 \\ -4 & -14 & -49 & 10 \end{bmatrix}. \end{aligned}$$

Finally

$$a_4 = \frac{1}{12}\text{tr} B_2 - \frac{1}{4}\text{tr}(AB_3) = \frac{26}{5}.$$

Hence, the characteristic polynomial of A is given by (3.1) as

$$a(s) = P_4(s) - 5P_3(s) + \frac{69}{7}P_2(s) - 10P_1(s) + \frac{26}{5}P_0(s).$$

Now, We consider the family $U_n = P_n^{(\frac{1}{2}, \frac{1}{2})}$ (Chebyshev Polynomials of the second kind). From (3.18), $a_1 = -\text{tr} A = -5$, and from (3.19)

$$B_1 = a_1 I + A = \begin{bmatrix} -4 & -4 & -1 & -4 \\ 2 & -5 & 5 & -4 \\ -1 & 1 & -7 & 3 \\ -1 & 4 & -1 & 1 \end{bmatrix}.$$

Using (3.18), we get

$$a_2 = \frac{3}{4} - \frac{1}{2}\text{tr}(AB_1) = \frac{39}{4},$$

and from (3.19)

$$B_2 = a_2 I - \frac{1}{4} B_0 + AB_1 = \begin{bmatrix} \frac{5}{2} & -1 & -10 & 5 \\ -9 & -\frac{19}{2} & -33 & 3 \\ 5 & 9 & \frac{53}{2} & -3 \\ 7 & 7 & 22 & \frac{1}{2} \end{bmatrix}.$$

Using again (3.18)

$$a_3 = \frac{1}{9} \text{tr } B_1 - \frac{1}{3} \text{tr } (AB_2) = -\frac{19}{2},$$

and from (3.19)

$$\begin{aligned} B_3 &= a_3 I - \frac{1}{4} B_1 + AB_2 \\ &= \frac{1}{4} \begin{bmatrix} -12 & 4 & 31 & -20 \\ 6 & 27 & 93 & -24 \\ -1 & -23 & -71 & 19 \\ -5 & -20 & -65 & 13 \end{bmatrix}. \end{aligned}$$

Finally

$$a_4 = \frac{1}{16} \text{tr } B_2 - \frac{1}{4} \text{tr } (AB_3) = \frac{35}{8}.$$

Hence, the characteristic polynomial of A is given by (3.1) as

$$a(s) = U_4(s) - 5U_3(s) + \frac{39}{4}U_2(s) - \frac{19}{2}U_1(s) + \frac{35}{8}U_0(s).$$

4.4. Bessel Basis. We consider the family $B_n = B_n^0$. From (3.20), $a_1 = -1 - \text{tr } A = -6$, and from (3.21)

$$B_1 = a_1 I + A = \begin{bmatrix} -5 & -4 & -1 & -4 \\ 2 & -6 & 5 & -4 \\ -1 & 1 & -8 & 3 \\ -1 & 4 & -1 & 0 \end{bmatrix}.$$

Using (3.20), we get

$$a_2 = -\frac{2}{21} - \frac{1}{6} \text{tr } B_1 - \frac{1}{2} \text{tr } (AB_1) = \frac{102}{7},$$

and from (3.21)

$$\begin{aligned} B_2 &= a_2 I + \frac{1}{35} B_0 + AB_1 \\ &= \begin{bmatrix} \frac{33}{5} & 3 & -9 & 9 \\ -11 & -\frac{22}{5} & -38 & 7 \\ 6 & 8 & \frac{168}{5} & -6 \\ 8 & 3 & 23 & -\frac{2}{5} \end{bmatrix}. \end{aligned}$$

Using again (3.20)

$$a_3 = -\frac{1}{30} \text{tr } B_1 - \frac{1}{6} \text{tr } B_2 - \frac{1}{3} \text{tr } (AB_2) = -\frac{289}{15},$$

and from (3.21)

$$\begin{aligned} B_3 &= a_3 I + \frac{1}{15} B_1 + AB_2 \\ &= \frac{1}{3} \begin{bmatrix} -21 & 1 & 52 & -35 \\ 34 & 43 & 175 & -32 \\ -17 & -43 & -141 & 27 \\ -26 & -31 & -116 & 10 \end{bmatrix}. \end{aligned}$$

Finally

$$a_4 = -\frac{1}{12} \text{tr } B_2 - \frac{1}{4} \text{tr } B_3 - \frac{1}{4} \text{tr } (AB_3) = \frac{84}{5}.$$

Hence, the characteristic polynomial of A is given by (3.1) as

$$\begin{aligned} a(s) &= B_4(s) - 6B_3(s) + \frac{102}{7}B_2(s) - \\ &\quad \frac{289}{15}B_1(s) + \frac{84}{5}B_0(s). \end{aligned}$$

Acknowledgements

The work of the first author has been supported by Fundación Universidad Carlos III de Madrid. The work of the second author has received the financial support from Dirección General de Investigación (Ministerio de Ciencia y Tecnología) of Spain, grant BFM2003-06335-C03-02.

References

- [1] **S. Barnett**, *Leverrier's algorithm for orthogonal polynomials bases*, Linear Alg. and Appl. **236** (1996), 245–263.
- [2] **T. S. Chihara**, *An introduction to orthogonal polynomials*, Gordon and Breach, New York, 1978.
- [3] **J. S. Frame**, *Matrix functions and applications IV: Matrix functions and constituent matrices*, IEEE Spectrum **1** (1964), 123–131.
- [4] **F. Marcellán, A. Branquinho, & J. C. Petronilho**, *Classical orthogonal polynomials: A functional approach*, Acta Appl. Math. **34** (1994), 283–303.

PRECONDICIONAMIENTO DE MÉTODOS ITERATIVOS*

por

Guillermo Cervantes C.¹ & Carlos Mejía S.²

A la memoria de Jairo Charris C.

Resumen

Cervantes, Guillermo & Carlos Mejía S.: Precondicionamiento de métodos iterativos. Rev. Acad. Colomb. Cienc. **26** (106): 49–55, 2004. ISSN 0370-3908.

Se exponen nociones generales sobre métodos iterativos no estacionarios y se explica la importancia de agregarles precondicionamientos. Se presentan varios tipos de métodos iterativos y varias clases de precondicionadores incluyendo algunos definidos en términos de matrices por bloques. Para estos últimos, se presentan algunos resultados sugeridos en la literatura reciente que se lograron probar en detalle. Los resultados se comprueban con ejemplos ilustrativos.

Palabras clave: Precondicionamiento, métodos iterativos, matrices ralas, sistemas lineales.

Abstract

We consider general material related to nonstationary iterative methods and we explain the importance of implementing a preconditioning strategy.

Key words: Preconditioning, iterative methods, sparse matrices, linear systems.

¹Guillermo Cervantes C., Universidad del Norte, Barranquilla, Colombia. email: gcervant@uninorte.edu.co

²Carlos Mejía S., Universidad Nacional de Colombia, Sede de Medellín, Colombia, email: cemejia@unalmed.edu.co

*Este trabajo fue apoyado por la Dirección Nacional de Investigaciones de la Universidad Nacional a través del Proyecto de Investigación código DI00C1236.

AMS Mathematics Subject Classification 2000. 65F10, 65F50, 65F30, 15A06.

1. Introducción

En este trabajo se muestran maneras de acelerar métodos iterativos para la solución numérica del sistema $Ax = b$, con A matriz no singular de $\mathbb{R}^{n \times n}$ y $b \in \mathbb{R}^n$. La aceleración se consigue por preconditionamiento, el cual se describe en la sección 2. Para que estos métodos sean verdaderamente útiles, es conveniente que la matriz A sea *rala*, es decir, que la mayor parte de sus elementos sean cero. Es muy conveniente también que la matriz sea estructurada, por ejemplo, una matriz en la que los elementos no nulos están situados únicamente en unas cuantas diagonales, llamadas bandas. De esta manera, los métodos trabajan más rápido pues no es necesario almacenar la matriz en memoria.

El trabajo está organizado así: Como lo dijimos antes, la idea de preconditionamiento se explica en la sección 2. Con algún detalle, se consideran preconditionamientos basados en métodos iterativos estacionarios y en factorización LU incompleta. En la sección 3, sobre preconditionamiento para matrices definidas por bloques, se ofrecen resultados relacionados con preconditionamientos recientemente definidos para matrices especiales 2×2 por bloques.

2. Precondicionamiento

La idea del preconditionamiento es la reducción del número de iteraciones requerido para la convergencia, transformando el sistema original $Ax = b$ por un sistema $\tilde{A}x = \tilde{b}$, de tal forma que se satisfagan las siguientes propiedades:

- Resolver $\tilde{A}x = \tilde{b}$ no debe incrementar considerablemente el número de operaciones que se requieren para resolver $Ax = b$.
- $Ax = b$ y $\tilde{A}x = \tilde{b}$ tienen la misma solución, es decir, $\tilde{A}^{-1}\tilde{b} = A^{-1}b$.

La matriz \tilde{A} y el vector \tilde{b} se consiguen por premultiplicación o posmultiplicación por una matriz M , llamada *precondicionador* que debe ser fácilmente invertible. Por medio del preconditionador se transforma el sistema $Ax = b$ en otro equivalente con condiciones espectrales más favorables y se reduce el número de iteraciones requeridas para la convergencia, sin incrementar significativamente la cantidad de cálculos por iteración. Cuando esto se hace, hay que poner en una balanza los costos y los beneficios; es decir, el costo de construir y aplicar un preconditionador versus la ganancia en rapidez de convergencia. Ciertos preconditionadores necesitan

una pequeña fase de construcción pero otros pueden necesitar un trabajo sustancial. En el segundo caso la ganancia puede estar en el uso repetido del mismo preconditionador en múltiples sistemas lineales.

2. 1. Métodos Iterativos Estacionarios. Dado el preconditionador M para el sistema $Ax = b$, obtenemos un sistema equivalente $M^{-1}Ax = M^{-1}b$. Si escribimos $A = M - N$, entonces $Ax = b$ se puede escribir como $Mx = Nx + b$.

Así que dada una aproximación x_{k-1} , se puede obtener una nueva aproximación x_k de la siguiente forma:

$$Mx_k = Nx_{k-1} + b$$

multiplicando por M^{-1} se obtiene

$$x_k = M^{-1}Nx_{k-1} + M^{-1}b$$

pero $M^{-1}N = I - M^{-1}A$, entonces

$$\begin{aligned} x_k &= (I - M^{-1}A)x_{k-1} + M^{-1}b \\ &= x_{k-1} + M^{-1}(b - Ax_{k-1}) \\ &= x_{k-1} + M^{-1}r_{k-1} \end{aligned}$$

Si $M = D = \text{diag}(A)$, la iteración se conoce como método de Jacobi. Se supone que los valores en la diagonal de A son diferentes de cero para que M^{-1} pueda estar definida.

Si la matriz A se escribe de la forma $A = D + L_A + U_A$, donde $D = \text{diag}(A)$, L_A es la parte de A estrictamente triangular inferior y U_A es la parte de A estrictamente triangular superior, podemos definir dos métodos iterativos adicionales:

1. $M = D + L_A$, en cuyo caso la iteración se llama método de Gauss-Seidel.
2. $M = \omega^{-1}D + L_A$, donde ω es un parámetro en el intervalo $(0, 2)$. En este caso la iteración se llama SOR (*successive overrelaxation*).

Notemos que si $\omega = 1$, el método SOR se reduce al método de Gauss-Seidel. Si $\omega > 1$, el método se llama de **sobrerelajación**, y si $\omega < 1$, el método se llama de **relajación**.

Cuando A es real y simétrica o compleja hermitiana y se definen $M_1 = \omega^{-1}D + L_A$ y $M_2 = \omega^{-1}D + U_A$, el par de iteraciones

$$\begin{aligned} M_1x_{k-\frac{1}{2}} &= N_1x_{k-1} + \omega b \text{ con } N_1 = M_1 - A \\ M_2x_k &= N_2x_{k-\frac{1}{2}} + \omega b \text{ con } N_2 = M_2 - A \end{aligned}$$

son denominadas método SSOR o SOR simétrico. En este caso el preconditionador M puede ser escrito como $M = \frac{\omega}{2-\omega} (\omega^{-1}D + L_A) D^{-1} (\omega^{-1}D + U_A)$.

Para matrices *consistentemente ordenadas* y que satisfacen la *Propiedad A*, (ver [7]) se conoce un valor óptimo para el parámetro ω . Se calcula en términos del radio espectral de la matriz de iteración de Jacobi, siempre que dicha matriz tenga valores propios reales y su radio espectral sea menor que 1. Dicho valor óptimo es $\omega_{opt} = \frac{2}{1+\sqrt{1-\beta^2}}$, donde β ($\beta < 1$) es el radio espectral de la matriz de iteración de Jacobi. Un dato importante es que las matrices tridiagonales por bloques y en particular, las matrices tridiagonales, satisfacen las dos condiciones mencionadas arriba.

2. 2. Factorización Incompleta. Una clase importante de preconditionadores se obtiene de la factorización LU , que sabemos trabaja para matrices arbitrarias. Sin embargo, existen algunas variantes que explotan la estructura especial que pueda tener la matriz: simétrica, definida positiva, ralas (sparse), etc. Para el caso de las matrices con un patrón de dispersión o ralas, se tiene un procedimiento denominado factorización incompleta, es decir: Si A es una matriz rala, los factores L y U usualmente no tienen el mismo patrón de dispersión de la matriz A , y como la idea es conservar este patrón, entonces se descartan los elementos diferentes de cero en aquellas posiciones donde la matriz A tenía un cero. Tales posiciones se denominan **fill-elements** (posiciones rellenadas). Así se obtendría una factorización aproximada: $A \approx LU$.

La idea de generar factorizaciones aproximadas ha sido estudiada por muchos investigadores. El primero que lo hizo fue R. Varga en 1960 [9]. Posteriormente J. Meijerink y H. Van der Vorst en 1977 [4] la hicieron popular cuando la usaron para generar preconditionadores para el gradiente conjugado y otros métodos iterativos.

Uno de los más importantes preconditionadores de esta clase es el que se obtiene por la factorización incompleta de Cholesky. Si A es una matriz rala, simétrica y definida positiva, se puede calcular una matriz H que tenga el mismo patrón de dispersión que la triangular inferior de A y que de algún modo esté cercana al factor de Cholesky G de A , así el preconditionador toma la forma $M = HH^T$.

Recordemos que para dos matrices B y C , decir que $B \geq C$ significa que para todos los correspondientes elementos, se cumple $b_{ij} \geq c_{ij}$. La factorización incompleta

vía LU está garantizada si A es invertible y diagonalmente dominante o si A es una M-matriz, es decir, si A es una matriz con elementos no positivos fuera de la diagonal principal, invertible y $A^{-1} \geq 0$ (Ver [8], pag. 210).

El siguiente teorema, probado en [4], establece la existencia y unicidad de la factorización incompleta LU para M-matrices.

Teorema 1. Si $A = (a_{ij})_{n \times n}$ es una M-matriz, entonces para cada subconjunto P de $\{(i,j): j \neq i, i, j=1,2,\dots,n\}$ existe una matriz triangular inferior $L = (l_{ij})_{n \times n}$ con unos en la diagonal y una matriz triangular superior $U = (u_{ij})_{n \times n}$ tal que $A = LU - R$, donde $l_{ij} = 0$ si $(i,j) \in P$, $u_{ij} = 0$ si $(i,j) \in P$ y $r_{ij} = 0$ si $(i,j) \notin P$. Los factores L y U son únicos y la descomposición $A = LU - R$ es tal que LU es no singular, $(LU)^{-1} \geq 0$ y $LU \geq A$.

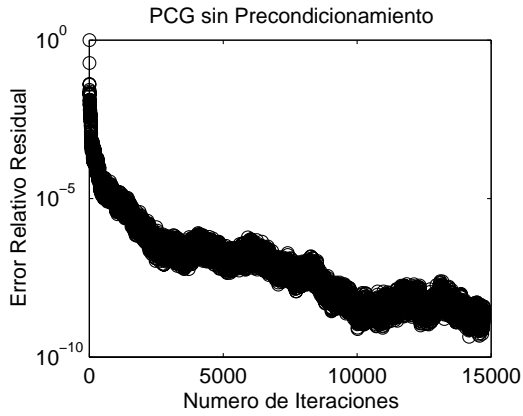
A la luz del teorema anterior, si escogemos $M = LU$, donde L es triangular inferior y U es triangular superior, el sistema $Mx = y$ (**instrucción del algoritmo del gradiente conjugado preconditionado**) se resuelve por la vía usual, es decir, por sustitución regresiva y progresiva. La factorización incompleta que se obtiene descartando todos los **fill-elements** en el proceso LU se conoce como **Factorización Incompleta de Nivel Cero** y se denota **ILU(0)**.

En todos los ejemplos presentados en este trabajo, se utilizan las rutinas que trae MATLAB 6.1 (versión instruccional) para los métodos iterativos *Gradiente Conjugado (CG)* y *Minimum Residual (MINRES)*. El primero se usa para matrices simétricas definidas positivas y el segundo para matrices generales. Además, en los ejemplos se trabaja con un vector de *unos* como solución exacta conocida y con un vector nulo como primera aproximación.

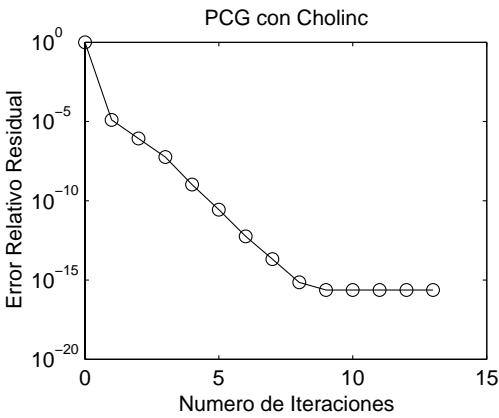
Ejemplo 1. Utilizamos los preconditionamientos Jacobi, SSOR y Factorización Incompleta de Cholesky en un sistema con matriz simétrica definida positiva tomada de la colección *Matrix Market* [6]. La matriz elegida para este ejemplo se llama *bcstk11*. Esta matriz aparece en aplicaciones de ingeniería estructural, tiene 1473 filas y un número de condición con un orden de magnitud de 10^8 .

Los experimentos que hicimos los presentamos en las siguientes figuras y en una tabla de resumen.

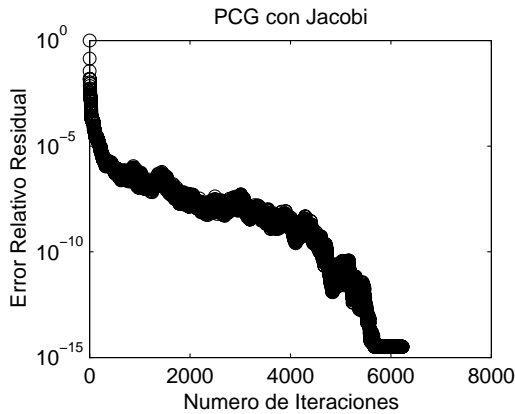
El primer resultado es el correspondiente a ausencia de preconditionamiento.



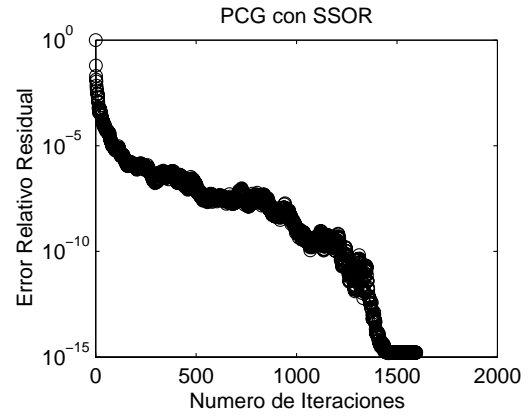
El segundo resultado corresponde a precondicionamiento con factorización incompleta de Cholesky.



En tercer lugar, aparecen los resultados de utilizar el método de Jacobi para precondicionamiento del CG.



Por último, estos son los resultados de aplicar el precondicionamiento SSOR a CG.



En la tabla siguiente resumimos estos experimentos.

Matriz bcsstk11 de Matrix Market			
Método	Iteraciones	Error	Tiempo(sg)
No Precond.	28312	4.4E-12	167
Cholinc(6)	13	8.1E-12	1.13
Jacobi	6236	1.3E-11	38
SSOR $\omega = 1.8$	1695	8.6E-12	511

3. Precondicionadores por Bloques

En las aplicaciones, frecuentemente las matrices deben considerarse por bloques. De hecho, las estrategias anteriores para producir precondicionadores por factorización incompleta y por los métodos iterativos clásicos se pueden adaptar para cuando la matriz A se particiona por bloques.

Otro tipo de precondicionadores por bloques se presenta en casos muy específicos. Por ejemplo, la matriz proveniente de la discretización de la ecuación de Poisson en un cuadrado unitario con n^2 puntos de malla tiene la siguiente estructura

$$\text{Trid por bloques}(-I, T, -I),$$

donde $T = \text{Trid}(-1, 4, -1)$. En este caso, un precondicionador natural es: $M = \text{Diag por bloques}(T)$.

Otra fuente de precondicionadores con estructura de bloque se origina cuando la matriz del sistema tiene una forma denominada **KKT** (Karush, Kuhnn, Tucker). Es decir, matrices *no singulares* que tienen la siguiente estructura:

$$\begin{pmatrix} A & B^T \\ C & 0 \end{pmatrix}, \text{ donde } A \in \mathbb{R}^{n \times n} \text{ y } B, C \in \mathbb{R}^{m \times n}, \quad (1)$$

con $n \geq m$. En muchas aplicaciones A es simétrica y $B = C$, en cuyo caso (1) sería simétrica; en todo caso, si A es o no simétrica, la matriz (1) es generalmente indefinida, es decir, las partes reales de sus valores propios pueden ser positivas o negativas. Matrices con esta estructura se presentan frecuentemente en aplicaciones: problemas de optimización, problemas de fluidos, etc. Murphy, Golub y Wathen muestran en [5] un preconditionador bastante eficiente para sistemas lineales cuya matriz de coeficientes tiene la forma (1). Dicho preconditionador tiene la siguiente estructura cuando la matriz A es invertible:

$$P = \begin{pmatrix} A & 0 \\ 0 & CA^{-1}B^T \end{pmatrix} \quad (2)$$

Este preconditionador fue generalizado posteriormente por Ipsen en [2] para matrices *no singulares* de la forma

$$\begin{pmatrix} A & B^T \\ C & D \end{pmatrix}, \quad (3)$$

donde $A \in \mathbb{R}^{n \times n}$, B y $C \in \mathbb{R}^{m \times n}$ y $D \in \mathbb{R}^{m \times m}$ con $n \geq m$. Dicho preconditionador también se ha utilizado en la solución de sistemas lineales provenientes de la discretización de las ecuaciones linealizadas de Navier-Stokes [3] con muy buenos resultados.

A continuación se detallan varias proposiciones en las cuales al preconditionar sistemas lineales cuyas matrices de coeficientes tengan la forma (1) o (3), usando preconditionadores con la forma (2), se obtienen matrices preconditionadas con a lo más cuatro, tres o dos valores propios distintos. La primera de tales proposiciones aparece demostrada en [5], las restantes aparecen referenciadas pero no probadas en [5] y [2]. Las pruebas detalladas de tales proposiciones pueden consultarse en [1].

El preconditionador (2) es sugerido mediante una descomposición de la matriz (1) usando el complemento de Schur, el cual se define a continuación:

Definición 2. Supongamos que la matriz $A \in \mathbb{R}^{n \times n}$ se puede escribir de la forma $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, donde A_{11} es $r \times r$. Si se supone que A_{11} es no singular, la matriz $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$ se denomina el **Complemento de Schur** de A_{11} en A .

La primera proposición que presentamos afirma que el preconditionador a izquierda (2) genera una matriz preconditionada con polinomio minimal de grado a lo más 4.

Proposición 3. Si la matriz $A = \begin{pmatrix} A & B^T \\ C & 0 \end{pmatrix}$, donde $A \in \mathbb{R}^{n \times n}$ es invertible y $B, C \in \mathbb{R}^{m \times n}$ con $n \geq m$, se preconditiona con (2), entonces la matriz preconditionada $T = P^{-1}A$ satisface $T(T - I)(T^2 - T - I) = 0$.

La segunda proposición afirma algo similar del polinomio minimal de la matriz preconditionada cuando el preconditionamiento se toma a derecha o a izquierda.

Proposición 4. Si $T = AP^{-1}$ o si $T = P_1^{-1}AP_2^{-1}$ con $P_1P_2 = P$ y A como en la proposición anterior, entonces también se satisface que

$$T(T - I)(T^2 - T - I) = 0.$$

La proposición que se presenta a continuación es referenciada sin demostración en [5]. Indica que métodos iterativos basados en subespacios de Krylov, como *min-res*, convergen en a lo más 4 iteraciones, suponiendo que se puede trabajar con *matemática exacta*.

Proposición 5. Para cualquier vector r , el subespacio de Krylov

$$\text{gen}\{r, Tr, T^2r, T^3r, \dots\}$$

con T como en la proposición anterior, es a lo más de dimensión 3 si T es no singular (o 4 si T es singular).

Demostración. Sin pérdida de generalidad supondremos que T es singular. Como

$$Q(t) = t(t - 1) \left(t - \frac{1 + \sqrt{5}}{2} \right) \left(t - \frac{1 - \sqrt{5}}{2} \right) \quad (4)$$

es un polinomio mónico que anula a T , entonces el polinomio minimal de T tiene a lo más grado 4.

Usando el Teorema 1.1.1 de [1], obtenemos que la dimensión de

$$\text{gen}\{r, Tr, T^2r, T^3r, \dots\} \quad (5)$$

es a lo sumo 4. Si T es no singular, el polinomio (4) tendría a lo sumo grado 3 y por tanto el subespacio en (5) tiene a lo más dimensión 3. \square

En la siguiente proposición se presentan variaciones del preconditionador (2), las cuales producen matrices preconditionadas con uno o dos valores propios distintos. Tal proposición es presentada en [5] pero sin demostración.

Proposición 6. Si se escoge $P = \begin{pmatrix} A & B^T \\ 0 & CA^{-1}B^T \end{pmatrix}$ en la proposición 3 el sistema preconditionado que resulta tiene exactamente dos valores propios 1 y -1. Si $CA^{-1}B^T$ es reemplazado por $-CA^{-1}B^T$, entonces la matriz preconditionada tiene a 1 como único valor propio.

El preconditionador trabajado en las proposiciones anteriores puede ser extendido a matrices no singulares de la forma $\mathcal{A} = \begin{pmatrix} A & B^T \\ C & D \end{pmatrix}$, donde la matriz A es invertible. La matriz \mathcal{A} no es necesariamente definida positiva y puede ser compleja, pero aquí trabajamos únicamente el caso real.

La proposición que sigue está propuesta sin demostración en [2] y muestra una extensión del preconditionador (2) desarrollado en la proposición 3.

Proposición 7. Sea $\mathcal{A} = \begin{pmatrix} A & B^T \\ C & D \end{pmatrix}$, donde $A \in \mathbb{R}^{n \times n}$ es invertible, B y $C \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{m \times m}$, $n \geq m$, $C \neq 0$ y $P = \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}$, donde $S = D - CA^{-1}B^T$ es el Complemento de Schur de A . Entonces AP^{-1} y $P^{-1}\mathcal{A}$ tienen como polinomio minimal a $Q(t) = (t-1)^2$.

La siguiente proposición, también de [2], es una generalización de la proposición 4 propuesta arriba.

Proposición 8. Si $P_1 = \begin{pmatrix} I & 0 \\ CA^{-1} & -I \end{pmatrix}$, $P_2 = \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}$ y \mathcal{A} como en la proposición anterior, entonces $P_1^{-1}\mathcal{A}P_2^{-1} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$. Además la matriz $P = P_1P_2 = \begin{pmatrix} A & B^T \\ C & D - 2S \end{pmatrix}$ es tal que $P^{-1}\mathcal{A}$ tiene como polinomio minimal a $Q(t) = (t-1)(t+1)$.

La proposición que se presenta a continuación es una extensión de la proposición 6 y aparece referenciada en [2] pero sin demostración.

Proposición 9. Si $\mathcal{A} = \begin{pmatrix} A & B^T \\ C & D \end{pmatrix}$, con A , B , C y D como en la proposición 7, entonces el preconditionador $P = \begin{pmatrix} A & 0 \\ 0 & -S \end{pmatrix}$, donde S es el complemento de Schur de A , es tal que la matriz preconditionada es $T = AP^{-1} = \begin{pmatrix} I & -B^TS^{-1} \\ CA^{-1} & -DS^{-1} \end{pmatrix}$. Si \mathcal{A} es de la forma KKT, es decir $D = 0$, entonces $T^2 - T = \begin{pmatrix} -B^TS^{-1}CA^{-1} & 0 \\ 0 & I \end{pmatrix}$. Puesto que $(T^2 - T)^2 = T^2 - T$, entonces el polinomio minimal de T es a lo más de grado 4.

Las proposiciones anteriores permiten concluir que para matrices de la forma

$$\mathcal{A} = \begin{pmatrix} A & B^T \\ C & 0 \end{pmatrix},$$

donde $A \in \mathbb{R}^{n \times n}$ y $B, C \in \mathbb{R}^{m \times n}$ con $n \geq m$, que por lo general son indefinidas, los preconditionadores

$$P = \begin{pmatrix} A & 0 \\ 0 & \pm CA^{-1}B^T \end{pmatrix} \text{ y } P = \begin{pmatrix} A & B^T \\ 0 & \pm CA^{-1}B^T \end{pmatrix}$$

logran que el polinomio minimal de $P^{-1}\mathcal{A}$ tenga a lo más 4 valores propios distintos, lo cual implica que, bajo aritmética exacta, métodos Krylov como el **minres** al ser aplicados al sistema lineal preconditionado converja en a lo sumo 4 iteraciones.

Conclusión similar se puede obtener para matrices de la forma

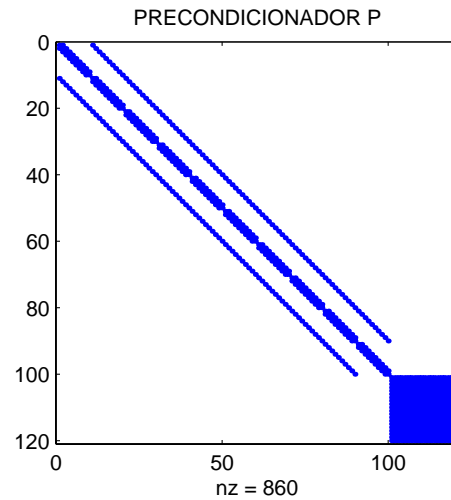
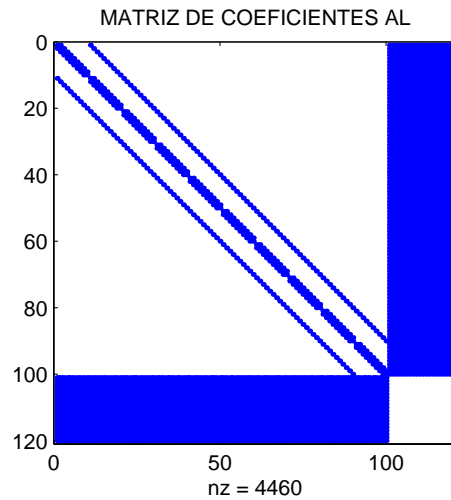
$$\mathcal{A} = \begin{pmatrix} A & B^T \\ C & D \end{pmatrix},$$

donde $A \in \mathbb{R}^{n \times n}$, B y $C \in \mathbb{R}^{m \times n}$ y $D \in \mathbb{R}^{m \times m}$ con $n \geq m$, con preconditionadores

$$P = \begin{pmatrix} A & B^T \\ 0 & \pm S \end{pmatrix}, P = \begin{pmatrix} A & B^T \\ C & D - 2S \end{pmatrix}$$

donde $S = D - CA^{-1}B^T$.

Ejemplo 2. Se diseñaron sistemas lineales con matrices de la forma (1) de tamaños diferentes. En todos los casos, los experimentos se hicieron de manera que la solución exacta sea un vector con todos sus elementos iguales a 1 y siempre se tomó el vector nulo como aproximación inicial. Como criterio de parada se utilizó una tolerancia de $1E-8$.



La tabla siguiente resume los experimentos que se hicieron.

Matriz por Bloques AL, Precondicionador MGW					
Dimensiones		Sin precond.		MGW	
A	AL	# Iter.	Error	# Iter	Error
100	120	130	2.8E-5	2	2.6E-10
400	450	295	8.8E-5	2	1.2E-7
900	1000	326	1.3E-4	2	3.8E-6
2025	4000	2582	1.4E-4	10	2.6E-3
2500	3000	865	1.4E-4	4	2.0E-4
3600	4000	629	1.0E-4	6	1.5E-5

El caso en el que se presentaron 10 iteraciones tuvo final anormal. El mensaje es que dos iterados consecutivos no se distinguen, de acuerdo con la tolerancia dada. Nótese que corresponde a una dimensión pequeña de A , en comparación con la de AL : La dimensión 2025 es menos del 51% de la dimensión 4000.

El otro caso con más de 4 iteraciones es el último, en el que se reportan 6 pero tuvo final normal.

En estos dos casos debe pensarse también que con matrices de tamaño 4000×4000 , no se puede pretender

que los cálculos se hagan exactamente. El inevitable redondeo es un factor que puede explicar el alejamiento moderado de lo que dice la teoría para estos casos.

Referencias

- [1] **G. Cervantes**, *Introducción a los preconditionadores*, Tesis de Maestría, Universidad Nacional de Colombia, Medellín, 2003.
- [2] **I. Ipsen**, *A note on preconditioning nonsymmetric matrices*, SIAM J. Sci. Compt., **23** (3) (2001), 1050–1051.
- [3] **D. Kay, D. Loghin & A. Wathen**, *A preconditioner for steady-state Navier–Stokes equations*, SIAM J. Sci. Compt., **24** (1) (2002), 237–256.
- [4] **J. A. Meijerink & H. A. van der Vorst** *An iterative solution method for linear systems of which the coefficient matrix is a symmetric m -matrix*, Math. Comput. **31** (137) (1977), 148–162.
- [5] **M. Murphy, G. Golub & A. Wathen**, *A note on preconditioning for indefinite linear systems*. SIAM J. Sci. Compt., **21** (6) (2000), 1969–1972.
- [6] **National Institute of Standards and Technology**, *The matrix market*. <http://math.nist.gov/MatrixMarket>.
- [7] **J. M. Ortega**, *Numerical Analysis, a Second Course*. SIAM, 1990.
- [8] **G. W. Stewart**, *Afternotes on Numerical Analysis*. SIAM, 1996.
- [9] **R. Varga**, *Factorization and normalized iteratives methods*. IN **R. Laner** (ed.), *Boundary Problems in Differential Equations*, University of Wisconsin, Madison, 1960, 12–142.

LA NOTACIÓN DIAGRAMÁTICA DE C. S. PEIRCE PARA LOS CONECTIVOS PROPOSICIONALES BINARIOS

por

Arnold Oostra¹

Dedicado al maestro Jairo Charris

Resumen

Arnold Oostra: La notación diagramática de C. S. Peirce para los conectivos proposicionales binarios. Rev. Acad. Colomb. Cienc. **28** (106): 57–70, 2004. ISSN 0370-3908.

Una de las notaciones más geniales para los conectivos proposicionales binarios fue introducida hace un siglo por C. S. Peirce. En esta notación, el símbolo asignado a cada conectivo puede leerse como la tabla de verdad del mismo y simultáneamente como el diagrama de Venn de la operación conjuntista correspondiente. Además de varias propiedades individuales y de múltiples relaciones, las simetrías del sistema completo de los conectivos proposicionales se reflejan de manera visual en los signos propuestos por Peirce.

Palabras clave: C. S. Peirce, conectivo proposicional, diagrama, tautología, simetría, automorfismo.

Abstract

One of the most brilliant notations for the binary propositional connectives was introduced a century ago by C. S. Peirce. In this notation, the symbol assigned to each connective can be read as its truth table and simultaneously as the Venn diagram of the corresponding set operation. Besides several individual properties and numerous relations, the symmetries of the complete system of connectives are visually reflected in the signs proposed by Peirce.

Key words: C. S. Peirce, propositional connective, diagram, tautology, symmetry, automorphism.

¹Profesor del Departamento de Matemáticas y Estadística, Universidad del Tolima. Becario de la Fundación Mazda para el Arte y la Ciencia. email: oostra@bunde.tolinet.com.co

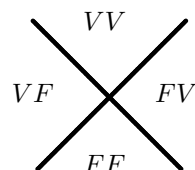
2000 Mathematics Subject Classification: 03A05, 00A30, 03B10.

Es bien sabido que en el cálculo proposicional clásico todos los conectivos pueden obtenerse como combinación de unos pocos, por ejemplo de la negación y la conjunción. Esto quizás explica por qué se acostumbra presentar y asignar símbolos a solo 4 o 5 de los 16 conectivos binarios clásicos y justifica la despreocupación generalizada por estudiar y diseñar notaciones para todos ellos. Siendo símbolos fundamentales en la lógica, un cambio hacia una notación adecuada podría compararse con el cambio de la notación romana para los números enteros a la arábica [8]. Sin embargo, a lo largo del siglo XX fueron propuestas numerosas notaciones para los conectivos binarios [20] que pueden clasificarse, de manera algo burda, en las *alfabéticas* que emplean letras usuales –como la polaca– y las *diagramáticas* que representan cada conectivo mediante un dibujo que procura sintetizar la definición del mismo. Quizás lo única notación que combina de manera armoniosa estas dos cualidades es la propuesta por Shea Zellweger [1, 17, 18, 19, 20, 21].

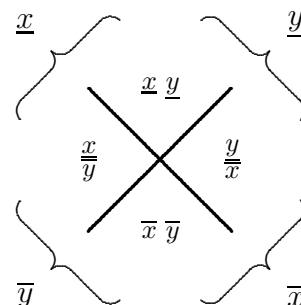
El lógico norteamericano Charles S. Peirce presentó por lo menos dos notaciones diferentes para los conectivos proposicionales binarios. Peirce ha sido destacado como uno de los intelectuales más originales y versátiles de América: entre otras ciencias, realizó aportes significativos a la química, la física, la geodesia, la filología, las matemáticas, la lógica matemática, la filosofía – de estas dos últimas puede decirse que fueron transformadas por las ideas del pensador. Por muy diversas razones Peirce nunca tuvo una posición estable en alguna Universidad, lo cual con seguridad aumentó el volumen de sus manuscritos pero mermó su influencia y el número de sus publicaciones. De hecho, C. S. Peirce ha pasado prácticamente desapercibido y solo en las últimas décadas del siglo XX ha tomado impulso el esfuerzo por estudiar, interpretar y aplicar sus ideas². La notación para los conectivos que Peirce más utilizó y que –por razones explicitadas en este trabajo– puede considerarse una de las mejores notaciones diagramáticas, fue elaborada hacia 1902 pues aparece en manuscritos con fechas de enero de ese año [2].

La idea de Peirce es sencilla en extremo: la tabla de verdad que define un conectivo binario tiene cuatro renglones, cada uno de los cuales puede ser V o F ; el símbolo \times tiene cuatro cuadrantes, cada uno de los

cuales puede dejarse abierto o bien cerrarse uniendo los extremos correspondientes. Se conviene cerrar los cuadrantes correspondientes a F y se adoptan los siguientes rótulos para los cuadrantes.



La idea de encerrar un espacio para indicar negación también se manifiesta en los gráficos existenciales de Peirce, donde un corte –una curva cerrada simple– se interpreta como la negación de su contenido (CP 4.402³, véase también [14, 22]). Por otra parte, un primer argumento para justificar su elección de los cuadrantes lo presenta Peirce en el diagrama siguiente, incluido en CP 4.260. En este dibujo \bar{x} denota la negación de x y \underline{x} su afirmación – Peirce sustenta de manera detallada la notación \bar{x} en CP 4.259 y ella se adopta en lo sucesivo en este documento.



La tabla 1 contiene la lista de los conectivos con sus signos respectivos. Peirce también propuso y empleó variantes cursivas de estos símbolos, sustentando con solidez los cambios introducidos en un pasaje precioso excluido de *Collected Papers* pero editado en una nota al pie en NEM 3.272⁴: en unos casos, se trata de signos más fáciles de escribir; en otros, algún antecesor respetable había introducido un símbolo más apropiado; en el caso del conectivo $=$, “a causa del sistema a veces daremos al signo la forma $x \infty y$ ”.

²Para mayor información véase [4, 5, 6, 7, 13, 15, 16] además de las referencias bibliográficas indicadas en estos documentos.

³Esto es: Parágrafo 402 del volumen 4 del texto *Collected Papers of Charles Sanders Peirce* [10].

⁴Esto es: Página 272 del volumen 3 de *The New Elements of Mathematics* [11].

VV	F	F	F	F	V	V	V	V	F	F	F	F	V	V	V	V
VF	F	F	F	V	F	V	F	F	V	V	F	V	F	V	V	V
FV	F	F	V	F	F	F	V	F	V	F	V	V	V	F	V	V
FF	F	V	F	F	F	F	F	V	F	V	V	V	V	V	F	V
	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠	⊠
	⊠	∧	<	>	∨	△	▽	∞	8	▽	▽	∞	∞	∞	∞	∞
								=								×

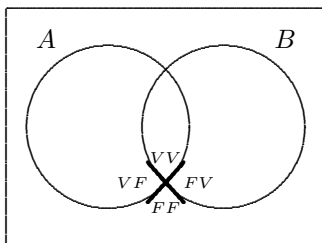
Tabla 1. Notación de Peirce para los conectivos binarios (1902).

En las secciones que siguen se muestra cómo la notación de Peirce refleja con fidelidad muchas propiedades de los conectivos, características que pueden clasificarse en tres *niveles de representación*.

1. Traducción

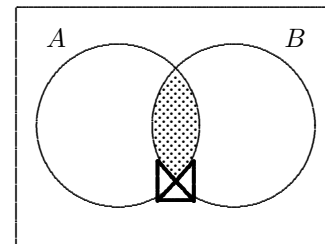
En la notación de Peirce, el signo del conectivo es la tabla de verdad del mismo. Para la traducción solo se requiere una convención sencilla: rotular los cuadrantes y cerrar los F . Esta traducción de signos a tablas de verdad se usa de manera iterada en el estudio de las propiedades de los conectivos (sección 2).

Pero en la notación de Peirce el signo del conectivo también es el diagrama de Venn de la operación conjuntista correspondiente. Si las proposiciones A y B se representan mediante circunferencias secantes, las cuatro regiones resultantes se rotulan –de manera natural– como sigue: la interior a ambas circunferencias es VV ; la interior a A y exterior a B es VF ; la exterior a A e interior a B es FV ; la exterior a ambas es FF . Si los centros de las circunferencias están a la misma altura con el de A a la izquierda y el de B a la derecha, entonces en el punto de corte *inferior* aparece la convención para los signos propuestos por Peirce.



Cabe anotar que Peirce mismo presentó este diagrama como un segundo argumento para justificar su elección de los cuadrantes, en el documento aún inédito *A Proposed Logical Notation* escrito alrededor de 1904.

Con esta convención, la traducción al diagrama de Venn es inmediata: el signo del conectivo se coloca en la intersección *inferior* y entonces las regiones sombreadas corresponden a los cuadrantes *abiertos* del signo. Por ejemplo, el siguiente es el paso del conectivo conjunción a la operación de intersección.



Todo parece indicar que Peirce empleó esta representación diagramática para calcular la tabla de soluciones de la quinta forma tautológica (apartado 2.5).

Las convenciones empleadas en la construcción de los dieciséis signos y en su traducción –a las tablas de verdad y a los diagramas de Venn–, además de muy sencillas, son *universales* en el sentido de que se emplea la misma clave de manera homogénea para representar la totalidad de los conectivos.

2. Propiedades y relaciones de los conectivos

En un artículo publicado en 1880, Christine Ladd-Franklin indicó que “hay una gran ventaja en que cada símbolo tenga una forma tal que su opuesto pueda indicarse volteándolo efectivamente, como el símbolo del Sr. Peirce para la implicación” (citado en [20]). A su vez, en el pasaje editado como nota en NEM 3.272 Peirce escribe: “Debe mencionarse (...) que parte de la propuesta [de la Sra. Franklin] era que cuando la relación significada era simétrica, el signo debería tener una simetría derecha e izquierda”. De esta manera, se busca que algunas propiedades del objeto –por ejemplo, el carácter conmutativo del conectivo– se reflejen en propiedades del signo –la simetría en el eje vertical–. De hecho, en

la notación de Peirce muchas propiedades de los conectivos se manifiestan en los signos y muchas relaciones entre los conectivos se traducen en relaciones entre los signos.

Recuérdese que dos fórmulas proposicionales α, β son *equivalentes* si la fórmula $(\alpha \bowtie \beta)$ es una tautología, esto es, resulta V para cualesquier valores de verdad asignados a las proposiciones atómicas que componen a α y a β .

Siendo O un conectivo cualquiera, ¿cuál es su “opuesto”? En otras palabras, ¿cuándo la fórmula $x O y$ es equivalente a $y \diamond x$, siendo O y \diamond conectivos? Al intercambiar las proposiciones x e y , los valores asignados a VV y FF no cambian, mientras los asignados a VF y FV se intercambian. Así, una respuesta a la segunda pregunta es: cuando el valor asignado por O a VF es el mismo asignado por \diamond a FV , y el valor asignado por O a FV es el mismo asignado por \diamond a VF . En la notación de Peirce esto se expresa como sigue: el cuadrante izquierdo de O coincide con el derecho de \diamond y el derecho de O con el izquierdo de \diamond .

Teorema. *Las expresiones $x O y$, $y \diamond x$ son equivalentes si y solo si \diamond se obtiene de O por reflexión en el eje vertical.*

En particular, un conectivo O es conmutativo –esto es, $x O y$, $y O x$ son equivalentes– si y solo si O es invariante bajo la reflexión en el eje vertical, o lo que es lo mismo, si O “tiene simetría derecha e izquierda”.

De manera similar puede derivarse la siguiente generalización de las conocidas identidades de De Morgan. Aquí *complementación* indica abrir los cuadrantes cerrados y cerrar los abiertos.

Teorema. *Las expresiones $\overline{x O y}$, $\overline{x} \diamond \overline{y}$ son equivalentes si y solo si \diamond se obtiene de O por rotación de 180 grados y complementación.*

Aunque el argumento es un poco más elaborado, también puede probarse de esta manera –diagramática– que \bowtie y \times son los únicos conectivos binarios *completos* en el sentido de que todo conectivo puede expresarse como combinación de cualquiera de ellos [5].

La búsqueda de tautologías. El objetivo primordial perseguido por Peirce al introducir los signos para los conectivos fue la búsqueda sistemática de tautologías, indagación en la cual las características de los signos juegan un papel decisivo. El procedimiento empleado por Peirce al buscar tautologías es el siguiente. Inicialmente

escogía una *forma*, una expresión de lógica proposicional en la cual no solo las proposiciones sino también los conectivos son incógnitas; luego, empleando propiedades de su notación, establecía todas –o muchas de– las *sustituciones* de los conectivos que hicieran una tautología de la forma.

Además de sus variantes, las formas estudiadas por Peirce son las siguientes [2].

$$\begin{aligned} & x O x \\ & x \diamond (x O x) \\ & (x \diamond x) O (x \heartsuit x) \\ & (x \diamond y) O (x \heartsuit y) \\ & (x \spadesuit y) \heartsuit [(y \clubsuit z) \diamond (x O z)] \end{aligned}$$

Aquí $O, \diamond, \heartsuit, \spadesuit$ y \clubsuit son *variables* y en cada forma se buscan los conectivos tales que la expresión es verdadera para *cualquier* proposiciones x, y, z . Nótese que cada forma establece una relación –binaria, ternaria, ... – entre los conectivos.

En el análisis de estas formas Peirce consignó sus soluciones en tres tablas que –aun bajo una mirada muy superficial– presentan una simetría muy notable, simetría debida a los signos empleados. A continuación se presenta de manera sucinta el estudio de las cinco formas.

2.1. Primera forma. La expresión $x O x$ corresponde a un conectivo de aridad 1 luego tiene solo cuatro opciones: constante V (tautología), constante F , x , \overline{x} . Para obtener el primero se requiere que el conectivo O asigne V a las parejas VV, FF –las únicas que pueden intervenir–, lo cual en la notación de Peirce equivale a que los cuadrantes superior e inferior estén abiertos. Así, los conectivos O que hacen de esta expresión una tautología son $\bowtie, \times, \times, \times$. De igual manera se analizan los otros tres valores posibles de $x O x$, forma que induce una *clasificación* de los conectivos binarios mostrada en la tabla 2. Esta es una adaptación de la tabla elaborada por Peirce, la tabla original se presenta en el apéndice (tabla 11).

Los rótulos de las cuatro clases fueron propuestos por Peirce en CP 4.270. Si $O \in \mathcal{N}$ entonces $x O x$ es una tautología; si $O \in \mathcal{E}$, la forma equivale a \overline{x} ; si $O \in \mathcal{W}$ entonces $x O x$ es x ; y si $O \in \mathcal{S}$, es siempre F . La ubicación de las cuatro clases en una gran \times es coherente con la convención para cada signo: en el cuadrante superior siempre V , en el inferior siempre F , etcétera. A su vez,

la asignación de letras a las clases es coherente con la orientación usual de los mapas terrestres. Una versión de la tabla 2 aparece en CP 4.268, pero por desgracia los editores en vez de reproducir los signos introducidos por Peirce los sustituyeron por “un simbolismo más convencional”.

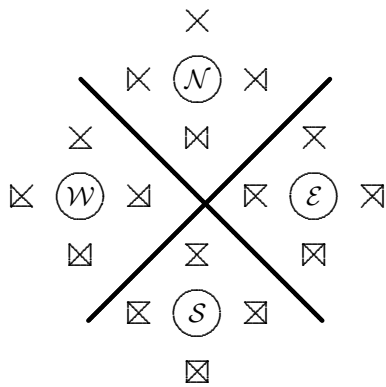


Tabla 2. Clasificación de los conectivos según la forma $x O x$.

2.2. Segunda forma. El análisis de la expresión $x \diamond (x O x)$ es sencilla en la medida en que contiene la forma $x O x$, estudiada de manera exhaustiva en el apartado 2.1. Por ejemplo si $O \in \mathcal{N}$ entonces $x O x$ es una tautología, de manera que la segunda forma se reduce a $x \diamond V$; para que esta forma sea tautología, basta que \diamond asigne V a los valores VV y FV –los únicos posibles–; en la notación de Peirce, esto equivale a que los cuadrantes superior y derecho estén abiertos. De esta manera, las parejas (\diamond, O) que arrojan una tautología de la segunda forma con $O \in \mathcal{N}$ son (\boxtimes, O) , (\boxtimes, O) , (\boxtimes, O) y (\times, O) .

O	\mathcal{N}	\mathcal{E}	\mathcal{W}	\mathcal{S}
\diamond	\boxtimes	\boxtimes	\boxtimes	\boxtimes

Tabla 3. Tautologías de la forma $x \diamond (x O x)$.

De la misma manera se analizan las otras tres elecciones posibles de O . La pequeña tabla 3 contiene los resultados, para su elaboración se adoptó la técnica empleada por Peirce: en cada caso, se escoge el conectivo con más cuadrantes cerrados que haga tautología la expresión estudiada. Por ejemplo, para $O \in \mathcal{N}$ se escoge \boxtimes . Todas las demás soluciones se obtienen abriendo uno por uno los cuadrantes cerrados del conectivo elegido, en todas las formas posibles. En el ejemplo, se

consigna \boxtimes y las soluciones posibles son \boxtimes , \boxtimes , \boxtimes , \times .

2.3. Tercera forma. Como antes, el estudio de la expresión $(x \diamond x) O (x \heartsuit x)$ se basa en la clasificación inducida por la forma $x O x$. Si $\diamond \in \mathcal{N}$ y $\heartsuit \in \mathcal{N}$, la forma se reduce a $V O V$ que es tautología si O asigna V a la pareja VV , es decir, si su cuadrante superior está abierto. De esta manera, las ternas $(\diamond, O, \heartsuit)$ con $\diamond \in \mathcal{N}$, $\heartsuit \in \mathcal{N}$ que arrojan una tautología de esta forma son $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$ y $(\diamond, \times, \heartsuit)$. Como otro ejemplo, si $\diamond \in \mathcal{N}$ y $\heartsuit \in \mathcal{W}$ entonces la forma se reduce a $V O x$ que es tautología si O asigna V a las parejas VV y VF , es decir, si están abiertos sus cuadrantes superior e izquierdo. Las ternas $(\diamond, O, \heartsuit)$ con $\diamond \in \mathcal{N}$, $\heartsuit \in \mathcal{W}$ que arrojan una tautología de esta forma son $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$, $(\diamond, \boxtimes, \heartsuit)$ y $(\diamond, \times, \heartsuit)$.

De esta manera pueden estudiarse los dieciséis casos posibles. Hasta donde se sabe, ni Peirce ni otros investigadores han consignado los resultados para esta forma en una tabla, pero es sencillo hacerlo con la técnica de Peirce indicada en el apartado 2.2: en el primer ejemplo se consigna \boxtimes y las soluciones posibles son \boxtimes , \boxtimes , \boxtimes , \boxtimes , \boxtimes , \boxtimes , \boxtimes , \boxtimes , \times ; en el segundo ejemplo se registra \boxtimes y las soluciones posibles son \boxtimes , \boxtimes , \boxtimes , \times .

La tabla 4 contiene todos los conectivos máximos. Si a la izquierda (conectivo \diamond) se escoge \mathcal{N} y a la derecha (conectivo \heartsuit) se toma \mathcal{N} , en la intersección de las filas correspondientes está \boxtimes ; en la intersección de las filas $\diamond = \mathcal{N}$ y $\heartsuit = \mathcal{W}$ está \boxtimes . Obsérvese la singular simetría especular que tiene esta tabla en los ejes vertical y horizontal, la misma simetría visible en el *Espejo Mágico* de M. C. Escher [3].

2.4. Cuarta forma. La expresión $(x \diamond y) O (x \heartsuit y)$ difiere sutilmente de la forma estudiada en el apartado 2.3. En vez de una variable proposicional ahora aparecen dos –lo cual hace disminuir el número de tautologías posibles– pero aparecen de manera ordenada –lo cual permite considerar simultáneamente las expresiones $x \diamond y$ y $x \heartsuit y$ –. La técnica empleada por Peirce para encontrar todas las tautologías de esta forma consiste en escoger libremente los conectivos \diamond y \heartsuit y, a partir de ellos, encontrar los conectivos O que hagan siempre verdadera la expresión analizada. Por ejemplo, si \diamond es \boxtimes y \heartsuit es \times entonces $x \diamond y$ siempre es F y $x \heartsuit y$ puede tomar los valores V y F , luego se requiere que O asigne V a las combinaciones FV y FF o, lo que es lo mismo, que sus cuadrantes derecho e inferior estén abiertos: en

este caso, las soluciones posibles para O son \boxtimes , \boxtimes , \boxtimes , \boxtimes . Si \diamond es \boxtimes y \heartsuit es \boxtimes entonces $x \diamond y$ toma los valores F, F, F, V –en ese orden– mientras $x \heartsuit y$ toma los valores V, F, F, V ; la forma $(x \diamond y) O (x \heartsuit y)$ se reduce a las cuatro posibilidades $F O V, F O F, F O F, V O V$; para obtener una tautología se requiere que O asigne V a las parejas FV, FF, VV , es decir, que sus cuadrantes superior, derecho e inferior estén abiertos; las soluciones posibles para O son \boxtimes y \boxtimes .

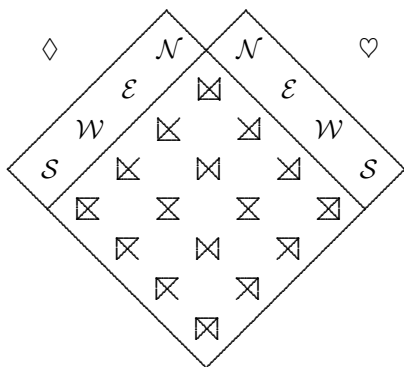


Tabla 4. Tautologías de la forma $(x \diamond x) O (x \heartsuit x)$.

Como en el apartado anterior, la tabla 5 contiene en cada casilla el conectivo O máximo que arroja una tautología de esta forma. En la intersección de las filas $\diamond = \boxtimes$ y $\heartsuit = \boxtimes$ está \boxtimes ; en la intersección de las filas $\diamond = \boxtimes$ y $\heartsuit = \boxtimes$ está \boxtimes .

Aunque aquí se presenta con otro orden para los conectivos, en esencia la tabla 5 se debe a Peirce –su versión original aparece en el apéndice, tabla 12–. Una versión horizontal (\square) de esta tabla aparece en CP 4.273, con la notación usada por los editores. De la tabla presentada allí, los editores dicen que “es una representación más clara (y precisa) de la tabla dada por Peirce, en razón de una reducción en el número de signos usados”. En el documento [9] se analiza la simetría de las tres tablas de Peirce con la ayuda de la teoría de grupos y, a partir de ese estudio, se refuta técnicamente la afirmación anterior: la razón esencial es que el grupo D_4 de los movimientos rígidos del cuadrado no actúa de manera *natural* sobre los signos empleados por los editores de *Collected Papers*. Allí también se muestra que las tablas 4 y 5 tienen la misma simetría y que ésta se pierde al colocarlas de manera horizontal (\square), de suerte que fue un gran acierto de Peirce el colocar sus tres tablas inclinadas (\diamond).

2.5. Quinta forma. Es evidente que el estudio de la forma $(x \spadesuit y) \heartsuit [(y \clubsuit z) \diamond (x O z)]$ es más complejo que el de las formas anteriores, pues aumenta el número de variables proposicionales y el de conectivos. Para esta forma Peirce renuncia a encontrar *todas* las tautologías, como lo ha logrado en las formas anteriores, y se conforma con *algunas*. Por ejemplo, no considera las tautologías ‘triviales’ que se obtienen cuando el conectivo principal \heartsuit es \boxtimes .

La estrategia empleada por Peirce puede describirse como sigue. En primer lugar, se escogen los conectivos \heartsuit, \diamond de tal manera que la forma auxiliar $P \heartsuit (Q \diamond R)$ sea verdadera excepto para una sola combinación preestablecida de valores para P, Q, R . Luego se determinan los conectivos \spadesuit, \clubsuit, O de tal manera que las tres proposiciones $(x \spadesuit y), (y \clubsuit z), x O z$ no tomen los valores establecidos, con lo cual se garantiza que la expresión completa es una tautología.

Peirce escogió cuatro combinaciones y las rotuló con los conectivos $\boxtimes, \boxtimes, \boxtimes, \boxtimes$. La razón para elegir estos signos es sencilla: en el primer caso los valores excluidos para $P-Q-R$ son $F-F-F$, que es con exactitud el único caso en el cual la implicación $(P \boxtimes Q) \boxtimes R$ es falsa. Con otras palabras, en el primer caso se están buscando aquellos conectivos \heartsuit, \diamond tales que la fórmula $P \heartsuit (Q \diamond R)$ es equivalente a $(P \boxtimes Q) \boxtimes R$. Los demás casos se obtienen al cambiar en la última expresión el conectivo \boxtimes por los otros tres indicados.

Si la combinación escogida es \boxtimes , los conectivos \heartsuit, \diamond pueden determinarse como sigue. Si $F \diamond F$ es F , se requiere que \heartsuit asigne F a la pareja FF y V a todas las demás, luego su único cuadrante cerrado es el inferior y $\heartsuit = \boxtimes$, mientras \diamond tiene cerrado su cuadrante inferior y para garantizar la unicidad se requiere que todos los demás estén abiertos de suerte que también $\diamond = \boxtimes$; en el otro caso, si $F \diamond F$ es V , el conectivo \heartsuit debe asignar F a la pareja FV y V a todas las demás, luego $\heartsuit = \boxtimes$, mientras \diamond tiene abierto su cuadrante inferior y cerrados los demás, $\diamond = \boxtimes$. Las parejas de soluciones posibles para los cuatro casos están consignados en la tabla 6.

pivote	\boxtimes	\boxtimes	\boxtimes	\boxtimes
$\heartsuit \diamond$	$\boxtimes \boxtimes$	$\boxtimes \boxtimes$	$\boxtimes \boxtimes$	$\boxtimes \boxtimes$
	$\boxtimes \boxtimes$	$\boxtimes \boxtimes$	$\boxtimes \boxtimes$	$\boxtimes \boxtimes$

Tabla 6. Conectivos \heartsuit, \diamond para la forma $(x \spadesuit y) \heartsuit [(y \clubsuit z) \diamond (x O z)]$.

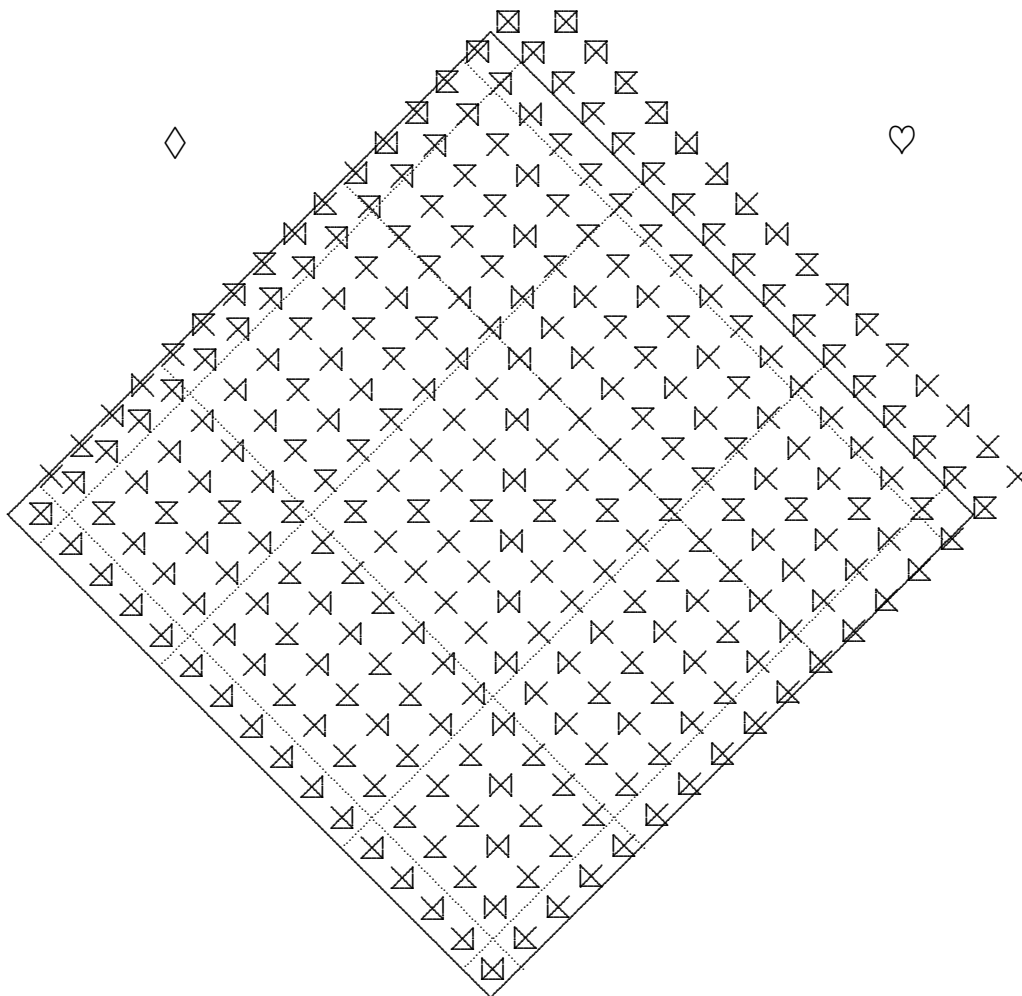


Tabla 5. Tautologías de la forma $(x \diamond y) \text{ O } (x \heartsuit y)$.

El cálculo de los conectivos \spadesuit , \clubsuit , O es más complejo aunque el procedimiento propuesto es el mismo seguido en las formas anteriores: se toman libremente dos conectivos \spadesuit , \clubsuit ; se buscan todos los conectivos O que satisfagan la condición exigida; se consigna el que tiene más cuadrantes cerrados; los demás se obtienen del máximo abriendo sus cuadrantes en todas las maneras posibles. En [2] y en [5] se indican métodos de cálculo del tercer conectivo en términos de los otros dos. Glenn Clark en [2] sostiene que Peirce no dejó indicación alguna sobre este cálculo, pero todo parece indicar que el pasaje CP 4.274 se refiere al mismo –en [5] se desarrollan algunos ejemplos que sostienen esta conjetura–. Los resultados de los 64 cálculos están consignados en la tabla 7. Obsérvese de nuevo la simetría especular en

el eje vertical, presente porque –siguiendo a Peirce– se tomaron órdenes distintos para las filas exteriores de la izquierda (\spadesuit) y de la derecha (\clubsuit). La tabla original de Peirce puede apreciarse en el apéndice (tabla 13). En [9] se demuestra que, con esta notación, solo hay *dos* presentaciones simétricas posibles, ambas con la tabla inclinada (\diamond). Por otra parte, aunque es un hecho de justificación fácil [5] no deja de ser asombroso que la misma tabla arroja las soluciones para *todos los cuatro* casos \boxtimes , \boxtimes , \boxtimes , \boxtimes .

Así, el procedimiento para encontrar tautologías de la quinta forma es el siguiente. En primer lugar se escoge un *pivote* del conjunto $\{\boxtimes, \boxtimes, \boxtimes, \boxtimes\}$. En segunda instancia y con la tabla 6, se toma cualquier pareja de

conectivos (\heartsuit, \diamondsuit) entre las dos asignadas al pivote elegido. Para los conectivos restantes se toman libremente dos conectivos \spadesuit, \clubsuit ; en la esquina superior de la tabla 7 se busca el pivote escogido; en las filas que se cortan en

el pivote se toma \spadesuit a la izquierda y \clubsuit a la derecha; al interior de la tabla, en el cruce de las filas correspondientes, se encuentra el conectivo máximo O que completa la tautología.

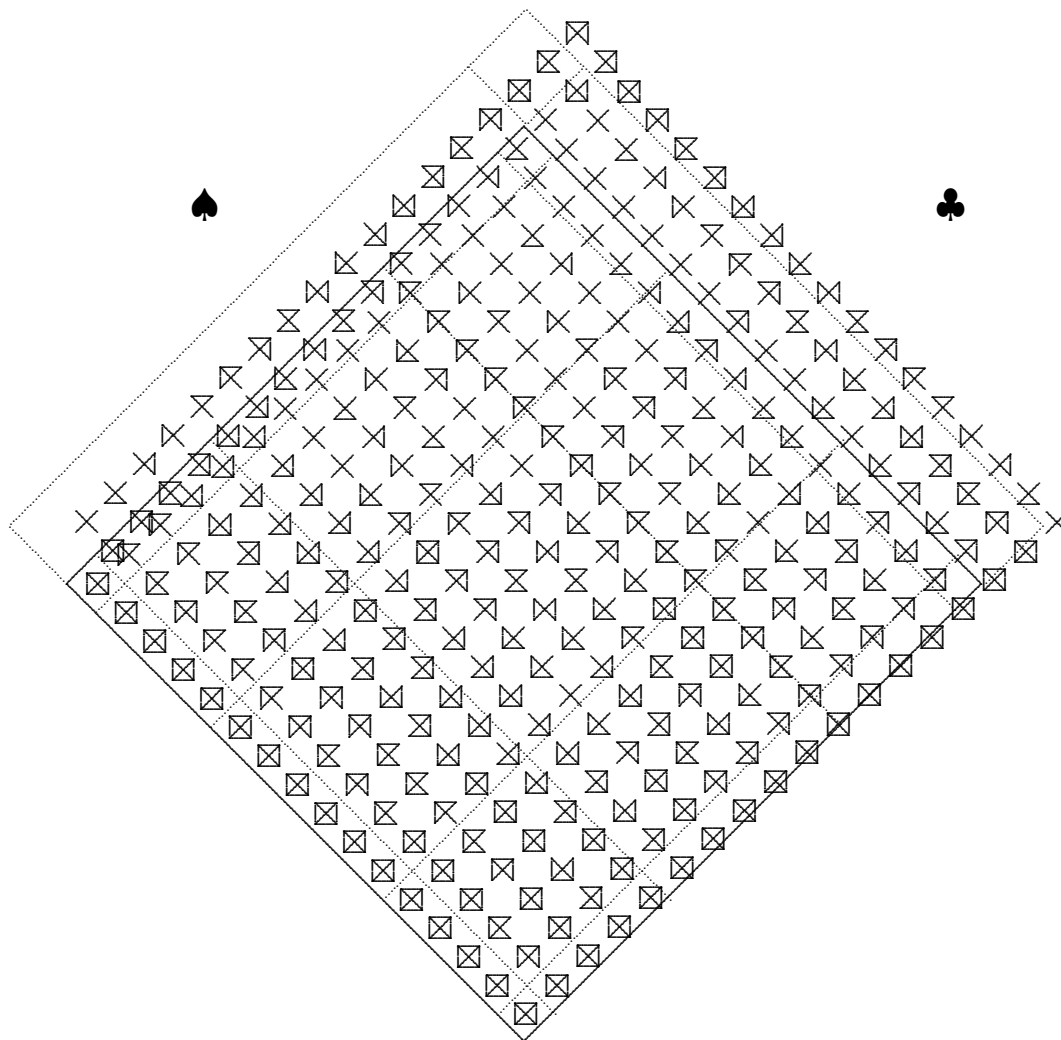


Tabla 7. Conectivos \spadesuit, \clubsuit, O para la forma $(x \spadesuit y) \heartsuit [(y \clubsuit z) \diamondsuit (x O z)]$.

Por ejemplo, se escoge el pivote \boxtimes . De la tabla 6 se elige $(\heartsuit, \diamondsuit) = (\boxtimes, \boxtimes)$.

Se escogen $(\spadesuit, \clubsuit) = (\boxtimes, \boxtimes)$ y la tabla 7 indica – buscando $\spadesuit = \boxtimes$ y $\clubsuit = \boxtimes$ en las filas que se encuentran en \boxtimes arriba – que puede tomarse $O = \boxtimes$ o cualquier conectivo obtenido de él abriendo sus cuadrantes cerrados.

Luego

$$(x \boxtimes y) \boxtimes [(y \boxtimes z) \boxtimes (x \boxtimes z)]$$

es una tautología.

Peirce consideró las siguientes doce variantes de esta forma, que pueden atacarse con el mismo método. La tabla 7 sirve para todas las variantes pero la 6 debe volverse a calcular para cada forma [5].

$$\begin{aligned}
&(x \spadesuit y) \heartsuit [(y \clubsuit z) \diamond (x \circ z)] \\
&(x \spadesuit y) \heartsuit [(x \circ z) \diamond (y \clubsuit z)] \\
&(y \clubsuit z) \heartsuit [(x \spadesuit y) \diamond (x \circ z)] \\
&(y \clubsuit z) \heartsuit [(x \circ z) \diamond (x \spadesuit y)] \\
&(x \circ z) \heartsuit [(x \spadesuit y) \diamond (y \clubsuit z)] \\
&(x \circ z) \heartsuit [(y \clubsuit z) \diamond (x \spadesuit y)] \\
\\
&[(x \spadesuit y) \diamond (y \clubsuit z)] \heartsuit (x \circ z) \\
&[(x \spadesuit y) \diamond (x \circ z)] \heartsuit (y \clubsuit z) \\
&[(y \clubsuit z) \diamond (x \spadesuit y)] \heartsuit (x \circ z) \\
&[(y \clubsuit z) \diamond (x \circ z)] \heartsuit (x \spadesuit y) \\
&[(x \circ z) \diamond (x \spadesuit y)] \heartsuit (y \clubsuit z) \\
&[(x \circ z) \diamond (y \clubsuit z)] \heartsuit (x \spadesuit y)
\end{aligned}$$

A la asombrosa tabla 7 está ligada una historia típica de los documentos de Peirce: en *Collected Papers* no aparece, pues se la confunde con la tabla 5; estuvo perdida durante muchos años en el inmenso legado manuscrito de Peirce; fue reubicada y comprendida apenas en las últimas décadas del siglo XX [2].

Conteo. A partir de las tablas presentadas no es difícil contar las tautologías encontradas por Peirce en las diferentes formas. De la tabla 2 es claro que la primera forma aporta 4 tautologías. En la tabla 3, cada uno de los conectivos consignados representa una familia de 4 soluciones; cada solución da lugar a 4 tautologías –pues O puede escogerse en un conjunto de tantos conectivos–; el total se duplica porque cada tautología de la forma

$x \diamond (x \circ x)$ aporta una de la forma variante $(x \circ x) \heartsuit x$ (¡reflejando el signo \diamond en su eje vertical!). Así, hay $4 \times 4 \times 4 \times 2 = 128$ tautologías de la segunda forma.

Como en la 3, en las tablas 4, 5 y 7 el conectivo que aparece en cierta casilla es el máximo del conjunto de soluciones: si tiene n cuadrantes cerrados, representa una familia de 2^n conectivos. En consecuencia, el número de tautologías aportadas por cada una de estas tablas es la suma $\sum 2^{n_i}$ donde i recorre todas las casillas de la tabla. Para la tercera forma la suma $\sum 2^{n_i}$ es 80, cantidad que debe multiplicarse por 16 –el número de posibles elecciones de la pareja (\diamond, \heartsuit) – lo cual arroja 1280 tautologías. En la tabla 5 se tiene $\sum 2^{n_i} = 680$; aunque Peirce no lo indicó, con esta misma tabla pueden encontrarse otras tantas tautologías de la forma alternativa $(x \diamond y) \circ (y \heartsuit x)$, aumentando la cuenta a 1360. La suma $\sum 2^{n_i}$ de la tabla 7 es 1699, cantidad que debe multiplicarse por la cantidad de pivotes, por la de opciones que da la tabla 6 y finalmente por el número de formas alternativas consideradas. Así, la cantidad de tautologías encontradas de la quinta forma asciende a $1699 \times 4 \times 2 \times 12 = 163104$. En CP 4.271 se indica un total de 24376, que debe ser $24576 = 256 \times 4 \times 2 \times 12$, producto obtenido al considerar solo una tautología por cada casilla de esta tabla en vez de 2^{n_i} .

El gran total de tautologías halladas en las diversas formas estudiadas por Peirce es $4 + 128 + 1280 + 1360 + 163104 = 165876$.

3. Propiedades del sistema

En la sección 2 se mostró cómo muchas propiedades de los conectivos se reflejan en el signo que les corresponde en la notación de Peirce. Además de las propiedades individuales de los conectivos y de las relaciones entre ellos –conmutatividad, tautología de una forma determinada– pueden distinguirse propiedades del sistema completo de los conectivos binarios.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
VV	F	F	F	F	V	V	V	V	F	F	F	F	V	V	V	V
VF	F	F	F	V	F	V	F	F	V	V	F	V	F	V	V	V
FV	F	F	V	F	F	F	V	F	V	F	V	V	V	F	V	V
FF	F	V	F	F	F	F	F	V	F	V	V	V	V	V	F	V

Tabla 8. El sistema de los conectivos binarios.

En la tabla 8 –donde los conectivos están rotulados con los números de 1 a 16– se reconocen varias simetrías, por ejemplo: la reflexión de la tabla en su eje vertical

corresponde a cambiar todas las letras V por F y viceversa. Conforme lo explicitó H. Weyl [12], la simetría de una estructura se estudia mediante el grupo de los automorfismos de la misma –de hecho, la teoría de grupos

es la herramienta precisa que la matemática desarrolló para estudiar la simetría–.

En el caso de los conectivos binarios, no es difícil encontrar *automorfismos lógicos*: la negación –asignar al conectivo O el conectivo \diamond tal que $x \diamond y$ y $\overline{x \text{ O } y}$ son equivalentes–, la negación en los argumentos, la conversión.

Todas las combinaciones de estas funciones básicas arrojan un total de 16 automorfismos. Puede justificarse que no hay más [4, 5], pero no parece nada fácil explicar por qué el número de automorfismos coincide con el de conectivos. La tabla 9 muestra de manera explícita las 16 funciones biyectivas, a la izquierda se indica su forma de cálculo.

$x \text{ O } y$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
$x \text{ O } y$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Reposo
$\overline{y} \text{ O } x$	1	3	5	2	4	10	6	9	8	11	7	13	15	12	14	16	Rotación 90°
$\overline{x} \text{ O } \overline{y}$	1	5	4	3	2	11	10	8	9	7	6	15	14	13	12	16	Rotación 180°
$y \text{ O } \overline{x}$	1	4	2	5	3	7	11	9	8	6	10	14	12	15	13	16	Rotación 270°
$y \text{ O } x$	1	2	4	3	5	7	6	8	9	11	10	12	14	13	15	16	Reflexión EV
$\overline{x} \text{ O } y$	1	4	5	2	3	11	7	9	8	10	6	14	15	12	13	16	Reflexión DA
$\overline{y} \text{ O } \overline{x}$	1	5	3	4	2	10	11	8	9	6	7	15	13	14	12	16	Reflexión EH
$x \text{ O } \overline{y}$	1	3	2	5	4	6	10	9	8	7	11	13	12	15	14	16	Reflexión DD
$\overline{x \text{ O } y}$	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	Complemento
$\overline{\overline{y} \text{ O } x}$	16	14	12	15	13	7	11	8	9	6	10	4	2	5	3	1	Rot 90°+Co
$\overline{\overline{x} \text{ O } \overline{y}}$	16	12	13	14	15	6	7	9	8	10	11	2	3	4	5	1	Rot 180°+Co
$\overline{y \text{ O } \overline{x}}$	16	13	15	12	14	10	6	8	9	11	7	3	5	2	4	1	Rot 270°+Co
$\overline{y \text{ O } x}$	16	15	13	14	12	10	11	9	8	6	7	5	3	4	2	1	Ref EV+Co
$\overline{\overline{x} \text{ O } y}$	16	13	12	15	14	6	10	8	9	7	11	3	2	5	4	1	Ref DA+Co
$\overline{\overline{y} \text{ O } \overline{x}}$	16	12	14	13	15	7	6	9	8	11	10	2	4	3	5	1	Ref EH+Co
$\overline{x \text{ O } \overline{y}}$	16	14	15	12	13	11	7	8	9	10	6	4	5	2	3	1	Ref DD+Co

Tabla 9. Efecto de los automorfismos lógicos.

Estos automorfismos con la composición constituyen un grupo. Se trata de un subgrupo no abeliano de S_{16} isomorfo a $\mathbb{Z}_2 \times D_4$: la última columna de la tabla 9 indica el isomorfismo. Aquí EV indica la reflexión en el eje vertical y EH en el horizontal, mientras DA indica la reflexión en la diagonal que asciende de izquierda a derecha y DD en la que descende de izquierda a derecha –llamada *diagonal principal* en teoría de matrices–. En lenguaje más sintético, los automorfismos corresponden a los movimientos rígidos de un cuadrado que puede cambiar entre dos estados –o dos colores–, cambio que en la tabla se ha descrito como “complemento”.

Todo (sub)grupo de automorfismos de una estructura *actúa* de manera natural sobre la estructura. En este caso, resulta interesante calcular las órbitas, los elementos invariantes y los subgrupos de isotropía de esta acción directamente de la tabla, como se hace con detalle en [4] y [5]. Allí también se calculan subgrupos y cocientes de este grupo.

Cuando se escribe la tabla 9 con la notación de Peirce, el resultado es sorprendente.

En la tabla 10 se observa de inmediato que los automorfismos lógicos corresponden *con toda exactitud* a los movimientos rígidos de los signos: el automorfismo identificado con la rotación corresponde a rotar cada signo 90 grados; cualquier automorfismo caracterizado como una reflexión corresponde a reflejar cada signo en el eje o diagonal del caso; el automorfismo de negación, identificado con el complemento, corresponde a abrir en cada signo los cuadrantes cerrados y cerrar los abiertos. El marco \times con el que se dibujan los signos originales es, sin duda, un cuadrado cuyos lados pueden alternar entre dos estados –cuadrante abierto y cuadrante cerrado– y sus movimientos rígidos corresponden, via un isomorfismo, a los automorfismos del sistema de los conectivos binarios. De esta manera, toda la simetría del sistema de conectivos se refleja en el sistema de signos propuesto por Peirce.

xOy	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	
xOy	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Reposo
$\bar{y}Ox$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Rotación 90°
$\bar{x}O\bar{y}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Rotación 180°
$yO\bar{x}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Rotación 270°
yOx	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Reflexión EV
$\bar{x}Oy$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Reflexión DA
$\bar{y}O\bar{x}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Reflexión EH
$xO\bar{y}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Reflexión DD
xOy	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Complemento
$\bar{y}Ox$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Rot 90°+Co
$\bar{x}O\bar{y}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Rot 180°+Co
$yO\bar{x}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Rot 270°+Co
yOx	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Ref EV+Co
$\bar{x}Oy$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Ref DA+Co
$\bar{y}O\bar{x}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Ref EH+Co
$xO\bar{y}$	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	☒	Ref DD+Co

Tabla 10. Efecto de los automorfismos con la notación de Peirce.

La notación propuesta por Peirce en 1902 para los conectivos proposicionales binarios puede considerarse un prototipo de un sistema de signos diagramático: cada signo *es*, en cierto modo, el objeto representado; las *propiedades* individuales de cada objeto y las *relaciones* entre los objetos se reflejan fielmente en los signos correspondientes; las propiedades colectivas del *universo* de los objetos se manifiestan en el sistema de signos. Este conjunto de signos establece un umbral bastante exigente que deberán superar –o al menos igualar– otras notaciones que se propongan para los conectivos proposicionales binarios.

Bibliografía

[1] Glenn Clark and Shea Zellweger, *Let the mirrors do the thinking*. Mount Union Magazine **93** (1993), 2–5.

[2] Glenn Clark, *New light on Peirce’s iconic notation for the sixteen binary connectives*. In: Nathan Houser, Don D. Roberts and James Van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce*, 304–333. Indiana University Press, Bloomington and Indianapolis, 1997.

[3] M. C. Escher, *Estampas y Dibujos*. Taschen, Köln, 1992.

[4] Mireya García, Jhon Fredy Gómez y Arnold Oostra, *Simetría y Lógica: la notación de Peirce para los 16 conectivos binarios*. Memorias del XII Encuentro de Geometría y sus Aplicaciones, 1–26. Universidad Pedagógica Nacional, Bogotá, 2001.

[5] Mireya García y Jhon Fredy Gómez, *Notación de Peirce para los Conectivos Binarios*. Tesis (Matemáticas). Universidad del Tolima, Ibagué, 2002.

[6] Nathan Houser, *Introduction: Peirce as logician*. In: Nathan Houser, Don D. Roberts and James Van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce*, 1–22. Indiana University Press, Bloomington and Indianapolis, 1997.

[7] Arnold Oostra, *Acercamiento lógico a Peirce*. Boletín de Matemáticas - Nueva Serie **VII** (2000), 60–77.

[8] Arnold Oostra, *Los diagramas de la matemática y la matemática de los diagramas*. Boletín de Matemáticas - Nueva Serie **VIII** (2001), 1–7.

[9] Arnold Oostra, *Simetría en algunas tablas de C. S. Peirce*. Memorias del XIV Encuentro de Geometría y sus Aplicaciones, por aparecer. Universidad Pedagógica Nacional, Bogotá, 2003.

[10] Charles S. Peirce, *Collected Papers of Charles Sanders Peirce*. Charles Hartshorne and Paul Weiss (Eds.), vols. 1–6. Harvard University Press, 1931–1934.

[11] Charles S. Peirce, *The New Elements of Mathematics*. Carolyn Eisele (Ed.), vols. 1–4. Mouton, The Hague, 1976.

[12] Hermann Weyl, *Simetría*. McGraw-Hill, Madrid, 1991.

[13] Fernando Zalamea, *Una jabalina lanzada hacia el futuro: anticipos y aportes de C. S. Peirce a la lógica matemática del siglo XX*. Mathesis **9** (1993), 391–404.

[14] Fernando Zalamea, *Lógica Topológica: una introducción a los Gráficos Existenciales de Peirce*. XIV Coloquio Distrital de Matemáticas y Estadística. Universidad Pedagógica Nacional, Bogotá, 1997.

[15] Fernando Zalamea, *Ariel y Arisbe. Evolución y evaluación del concepto de América Latina en el siglo XX: una visión crítica desde la lógica contemporánea y la arquitectónica pragmática de C. S. Peirce*. Tercer Lugar, Premio de Pensamiento Latinoamericano. Convenio Andrés Bello, Bogotá, 2000.

[16] Fernando Zalamea, *El Continuo Peirceano. Aspectos globales y locales de genericidad, reflexividad y modalidad: una visión del continuo y la arquitectónica pragmática peirceana desde la lógica matemática del siglo XX*. Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, 2001.

[17] Shea Zellweger, *Peirce, iconicity, and the geometry of logic*. In: Myrdene Anderson and Floyd Merrell (Eds.), *On Semiotic Modeling*, 483–507. Mouton de Gruyter, New York, 1991.

- [18] **Shea Zellweger**, *Cognitive frames, cognitive overload, and mind-held diagrams in logic*. In: John Deeley and Terry Pre-witt (Eds.), *Semiotics 1991*, 35–45. University Press of America, New York, 1993.
- [19] **Shea Zellweger**, *On a deep correspondence between sign-creation in logic and symmetry in crystallography*. In: Irmen-gard Rauch and Gerald F. Carr (Eds.), *Semiotics around the World: Synthesis in Diversity*, 821–824. Mouton de Gruyter, New York, 1997.
- [20] **Shea Zellweger**, *Untapped potential in Peirce's iconic notation for the sixteen binary connectives*. In: Nathan Houser, Don D. Roberts and James Van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce*, 334–386. Indiana University Press, Bloomington and Indianapolis, 1997.
- [21] **Shea Zellweger**, *Mathelological semiotics: a lesson in constructing a shape-value notation for elementary logic*. In: Myrdene Anderson, Adalira Sáenz-Ludlow, Shea Zellweger, Victor V. Cifarelli (Eds.), *Educational Perspectives on Mathematics as Semiosis: from thinking to interpreting to knowing*, 285–356. Legas, Ottawa, 2003.
- [22] **J. Jay Zeman**, *The Graphical Logic of C. S. Peirce*. Ph. D. dissertation. University of Chicago, Chicago, 1964.

Apéndice: Las tablas originales de Peirce

A continuación se presentan las versiones originales de las tablas 2, 5 y 7, elaboradas por Charles S. Peirce en 1902. Todas se han copiado de [2].

La tabla 11 es la misma 2 pero en ella Peirce integró la información de la tabla 3, pues los signos colocados en el centro de los cuadrantes corresponden a las soluciones de la segunda forma. En efecto, el signo colocado *a la izquierda* en la circunferencia central de cada cuadrante

es el conectivo \diamond ‘máximo’ que arroja una tautología de la forma $x \diamond (x \text{ O } x)$, tomando el conectivo O de ese cuadrante; el signo colocado *a la derecha* es el conectivo \heartsuit máximo que arroja una tautología de la forma dual $(x \text{ O } x) \heartsuit x$, tomando O de ese cuadrante.

Las tablas 12 y 13 corresponden a las tablas 5 y 7, respectivamente.

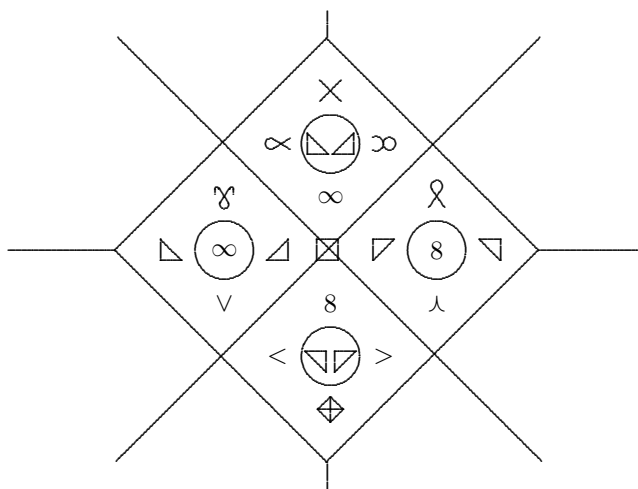


Tabla 11. Primera tabla de Peirce.

The image shows a large diamond-shaped Peirce table. The table is oriented with its vertices at the top and bottom. The grid is composed of various symbols, including diamonds, hearts, and mathematical symbols like infinity and triangles. The symbols are arranged in a complex, repeating pattern across the grid. The table is surrounded by a border of symbols, including diamonds and hearts. The overall structure is a large diamond shape with a grid of symbols inside.

Tabla 12. Segunda tabla de Peirce: tautologías de la forma $(x \diamond y) \text{ O } (x \heartsuit y)$.

The image shows a diamond-shaped Peirce table, which is a 16x16 grid of symbols arranged in a diamond pattern. The symbols are arranged in a regular grid, with each cell containing a unique symbol. The symbols include triangles, diamonds, and crosses, arranged in a regular grid. The table is surrounded by a dotted border. On the left side of the table, there is a spade symbol (♠) and on the right side, there is a club symbol (♣).

Tabla 13. Tercera tabla de Peirce: forma $(x \spadesuit y) \heartsuit [(y \clubsuit z) \diamond (x \circ z)]$.

MARIPOSAS Y 3-VARIEDADES

por

M. H. Hilden*, J. M. Montesinos[†], D. M. Tejada & M. M. Toro[‡]

Dedicado a la memoria del profesor Jairo Charris C.

Resumen

Hilden, M. H., J. M. Montesinos, D. M. Tejada & M. M. Toro: Mariposas y n -variedades. Rev. Acad. Colomb. Cienc. **28** (106): 71–78, 2004. ISSN 0370-3908.

En este artículo mostraremos que para cada diagrama de un nudo 3-coloreable se le puede asociar, en forma canónica, una 3-bola (llamada mariposa) con caras identificadas por reflexiones topológicas, la cual se puede triangular y 4-colorear, de tal manera que la coloración del nudo y de la triangulación resultan compatibles. Como corolario se obtiene que para toda 3-variedad hay al menos una bola asociada a ella con dichas características.

Palabras clave: Nudos, grupo fundamental, 3-variedades, recubrimientos ramificados.

Abstract

In this paper we show that for any 3-colorable knot diagram there is a canonic 3-ball, called butterfly, with faces that are identified by topological reflections. Moreover, we prove that this ball admits a 4-colored triangulation compatible with the 3-coloration of the knot. As a corollary we associate to any 3-manifold at least one of these balls.

Key words: Knots, Fundamental group, 3-manifolds, Branched coverings.

*Department of Mathematics, University of Hawaii, Honolulu, Hawaii 96825, USA. e-mail: mike@math.hawaii.edu

[†]José María Montesinos Amilibia, Departamento de Geometría y Topología, Universidad Complutense de Madrid, Madrid, España. e-mail: montesin@mat.ucm.es

[‡]Escuela de Matemáticas, Universidad Nacional de Colombia, Apartado Aéreo 3840, Medellín, Colombia. e-mail: dtejada@perseus.unalmed.edu.co ; mmtoro@perseus.unalmed.edu.co

AMS 2000 Mathematics Subject Classification: Primary 57M25. Secondary: 57M12.

1. Introducción

Aproximadamente, desde hace un siglo, el problema de la clasificación de 3-variedades está planteado. Prácticamente, desde ese entonces la clasificación de las superficies (2-variedades) es conocida. Específicamente, una superficie (2-variedad) compacta, conexa sin frontera es una esfera o una suma conexa de toros o una suma conexa de planos proyectivos. Este es un teorema sencillo y elegante. Un resultado similar se busca para las 3-variedades.

El teorema de clasificación de superficies se prueba usando métodos de topología combinatoria. Una de las pruebas clásicas se hace tomando polígonos con un número par de lados e identificando los lados por parejas. Se sabe que toda 3-variedad es una bola (o poliedro) con un número par de caras poligonales que se identifican por parejas. Así que uno pensaría que es posible generalizar la prueba de superficies a 3-variedades, pero infortunadamente, al tratar de imitar dicha prueba en dimensión 3, uno de los problemas que se encuentra es que el número de formas de identificar un par de caras fijas puede ser grande, lo cual hace que rápidamente se descarte este método.

Nuestro trabajo pretende reactivar este estudio, concentrándonos en bolas para las cuales identificamos sus caras de una forma precisa y única. Estas bolas las llamamos mariposas ya que la forma de identificar sus caras nos recuerda como se cierran las alas de una mariposa. Dichas bolas en sí mismas no nos producirán cualquier 3-variedad, para ello deberemos recurrir al Teorema de Hilden y Montesinos (véase [Hi], [Mo1]) que relaciona para cada 3-variedad un nudo 3-coloreado.

1.1. Preliminares

1.1.1. Resultados de la teoría de nudos

Recordaremos en esta sección algunos conceptos y teoremas de la teoría de nudos.

Un *nudo* K es una copia homeomorfa de S^1 en $S^3 = \mathbb{R}^3 \cup \{\infty\}$. Una proyección de un nudo sobre un plano, tal que la gráfica tenga sólo puntos simples o dobles y los segmentos que la formen se intersecten transversalmente, es llamada un *diagrama del nudo* (véase [Ro]). Las Figuras 2, 3 y 4 ilustran distintos diagramas.

El *grupo de un nudo* K es el grupo fundamental de $S^3 \setminus K$ y es denotado por $\pi_1(K)$. Para cada diagrama de un nudo existe una presentación canónica de su grupo, la cual es llamada *la presentación de Wirtinger*. Esta

presentación tiene tantos generadores como tramos continuos tenga el diagrama y tantas relaciones como cruces tenga el diagrama, siendo una cualquiera de ellas redundante. Cada una de las relaciones se obtiene teniendo en cuenta la orientación del nudo en dicho cruce.

En las figuras siguientes, donde las letras denotan los generadores asociados a cada tramo continuo, mostramos la relación que aparece con cada estilo de cruce.

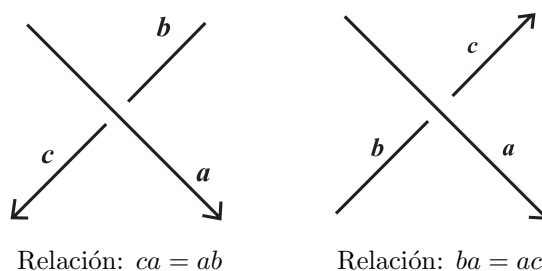


Figura 1.

Ejemplo 1. La presentación

$$\langle a, b, c, d, f, g, h, i, j : da = ac; af = fb; hb = bg; cg = gb; dc = cf; ad = dj; if = fh; gi = if \rangle$$

es una presentación de Wirtinger para el grupo del nudo 9_{45} , que ilustramos en la Figura 2.

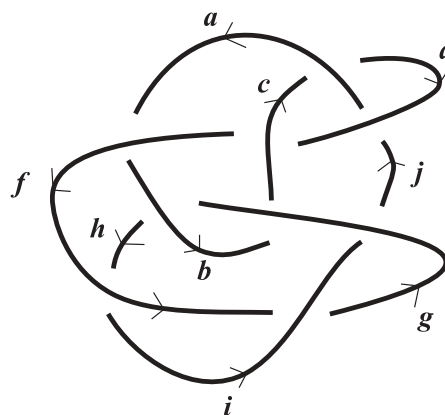


Figura 2.

Definición 2. Decimos que un nudo (o enlace) es *3-coloreable* si admite un diagrama al cual se le puede hacer lo siguiente:

1. Colorear cada arco del diagrama.
2. Usar un total de 3 colores.
3. En cada cruce debe ocurrir una sola de las siguientes posibilidades:
 - (a) Cada arco que está en dicho cruce se colorea con diferente color.

(b) Se usa únicamente un color.

Nota 3. Realmente si un nudo es 3-coloreable, entonces todos sus diagramas lo son. Esto es porque la propiedad de ser 3-coloreable es invariante bajo los movimientos de Reidemeister (véase [FaSt]).

Ejemplo 4. Los siguientes nudos son 3-coloreables.

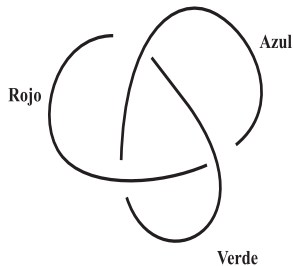


Figura 3a.

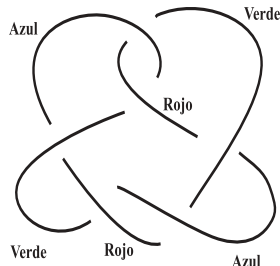


Figura 3b.

Sin embargo, no todo nudo es 3-coloreable, por ejemplo es fácil comprobar que el nudo del ocho de la Figura 4 no lo es.

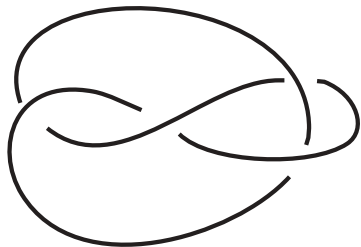


Figura 4.

Sea $\Sigma_3 = \{id, (0\ 1), (1\ 2), (0\ 2), (0\ 1\ 2), (0\ 2\ 1)\}$ el grupo simétrico de orden 3.

Teorema 5. Sea K un nudo 3-coloreable. Entonces su grupo fundamental admite una representación u homomorfismo (no trivial) en Σ_3 .

Demostración. En efecto, sea D_K un diagrama dado del nudo K , el cual está 3-coloreado. Sean $\{\alpha_i\}_{i=1}^n$ el conjunto de generadores de la presentación de Wirtinger para este diagrama. Entonces cada arco tiene asignado un generador y un color. Supongamos que los colores son *verde*, *rojo* y *azul*, definimos $\Phi : \pi_1(K) \rightarrow \Sigma_3$ tal que $\Phi(\text{generador asociado a arco verde}) \mapsto (0\ 1)$, $\Phi(\text{generador asociado a arco rojo}) \mapsto (1\ 2)$ y $\Phi(\text{generador asociado a arco azul}) \mapsto (0\ 2)$. Ahora, ya que las relaciones de Wirtinger se satisfacen en cada cruce y como, además, $(i\ j)(i\ k)(i\ j)^{-1} = (j\ k)$ y $(i\ j)(i\ j)(i\ j)^{-1} = (i\ j)$ para i, j, k diferentes entre sí

en $\{1, 2, 3\}$, entonces se sigue inmediatamente que Φ es una representación u homomorfismo de grupos. \square

Así pues, cada vez que tengamos un diagrama de un nudo 3-coloreado, para cada arco escribiremos la transposición asociada, en vez de escribir el nombre de su color.

Las pruebas de los teoremas que enunciaremos a continuación se pueden consultar en [BuZi], [CrFo], [Fo], [Kaw], [Ka], [Mo2], [Ro], [SeTh1] o en [SeTh2].

Sean K un nudo en S^3 , Σ_n el grupo simétrico de n elementos, $n \in \mathbb{N}$, entonces:

Teorema 6. Para toda representación $\rho : \pi_1(S^3 \setminus K) \rightarrow \Sigma_n$ existe una única n -cubierta ramificada $p(\rho) : M^3 \rightarrow S^3$ asociada a ρ , donde M^3 es una 3-variedad y cuyo conjunto de ramificación es el nudo K (o un subconjunto de K). Además, si ρ es transitiva M^3 es conexa.

Teorema 7. Sea M^3 una 3-variedad. Para toda n -cubierta $p : M^3 \rightarrow S^3$ ramificada sobre un nudo K , existe una única representación $\rho(p) : \pi_1(S^3 \setminus K) \rightarrow \Sigma_n$ asociada a p , salvo conjugación en Σ_n . Además, si M^3 es conexa ρ es transitiva.

El siguiente teorema fue probado, simultánea e independientemente, por M. H. Hilden y J. M. Montesinos.

Teorema 8. Para toda 3-variedad M^3 cerrada y orientable, existe una 3-cubierta ramificada sobre un nudo, la cual es simple.

Demostración. Véase [Hi] y [Mo1]. \square

Recordemos que una cubierta ramificada p sobre un nudo K se llama *simple* si la representación $\rho(p)$ asociada a ella envía los generadores de Wirtinger, de un diagrama de K , a transposiciones. Así pues, el teorema anterior es equivalente a:

Teorema 9. Para toda 3-variedad M^3 cerrada y orientable, existe una 3-cubierta ramificada sobre un nudo 3-coloreable.

El teorema que sigue es debido a Goodman y Onishi (véase [GoOn] y [Te]), si bien no hace parte de la Teoría de Nudos, lo utilizaremos en la prueba de nuestro Teorema Principal.

Teorema 10. Sean B^3 una bola con frontera S^2 , entonces toda triangulación 4-coloreada Δ^2 de S^2 puede ser extendida a una triangulación 4-coloreada Δ de la bola B^3 .

2. Mariposas

En esta sección, inicialmente, damos la definición de mariposa, ilustrándola con algunos ejemplos y, a continuación, veremos que todo diagrama de un nudo tiene una mariposa asociada. En [HMTT1], [HMTT2] y en [Te] se encuentra una definición de mariposa más formal y además, se pueden consultar más detalles de las pruebas que aquí presentamos.

Definición 11. Una *mariposa* es una 3-bola con un número par de caras poligonales (llamadas *alas*) identificadas por parejas. Dos alas se pueden identificar si ellas comparten una arista y la identificación se realiza al hacer una reflexión topológica (no necesariamente isometría) a lo largo de la arista común. La unión de las aristas sobre las cuales se hacen las reflexiones es llamada el *tronco* de la mariposa.

Ejemplo 12. Las 3-bolas siguientes son todas mariposas, las reflexiones están indicadas por las flechas dobles.

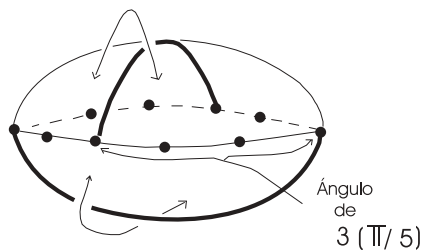


Figura 5a.

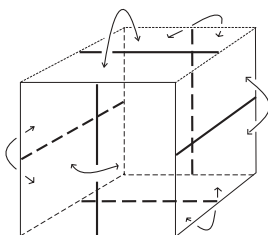


Figura 5b.

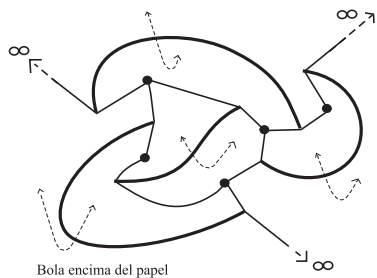


Figura 5c.

Estas tres mariposas tienen todas la particularidad que al hacer la identificación de todas las parejas de alas se obtiene la esfera S^3 (en [HMTT2] se prueba que esto siempre es así) y el tronco se convierte en un nudo inmerso en S^3 . En efecto, la primera mariposa produce el nudo racional $\frac{5}{3}$ que ilustramos a continuación (véase prueba en [HMTT2] y en [Te]).



Figura 6.

La segunda mariposa, debida a Thurston (véase [Thu]) produce los anillos de Borromeo. Y como veremos a continuación, al hacer las identificaciones de las caras de la última mariposa obtenemos el nudo del ocho (Figura 4). Obsérvese que ésta la hemos dibujado en forma diferente, pues allí suponemos que la frontera de la mariposa es el papel y que la 3-bola está sobre el plano del papel.

Definición 13. Si al realizar las identificaciones de las caras de una mariposa el tronco produce un nudo inmerso en S^3 , decimos que *el nudo admite una representación de mariposa*.

Teorema 14. *Todo nudo admite una representación de mariposa.*

Demostración. La prueba se puede consultar en [HMTT1], [Te] o en el artículo “Nudos Combinatorios y Mariposas” de M. M. Toro que aparece en esta misma revista. \square

Dicha prueba es constructiva y permite deducir el siguiente algoritmo:

Algoritmo para construir la mariposa.

1. Partimos de un diagrama del nudo, es decir de una colección S de arcos disjuntos en un plano P (del

papel). (Ilustraremos el algoritmo con el nudo del ocho de la Figura 4.)

2. En el interior de cada región acotada y determinada por el diagrama del nudo marcamos un punto B (que depende de la región). En la región no acotada, dicho punto B es el infinito, pero no lo marcamos.

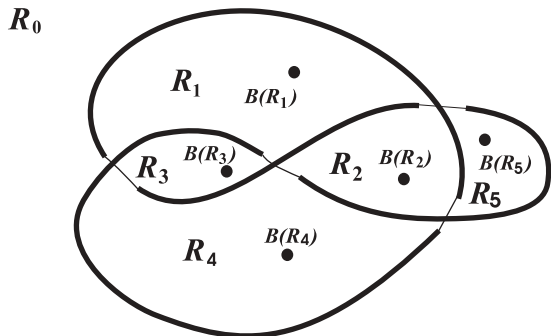


Figura 7.

3. Cada extremo de un arco de S lo unimos con los dos puntos B pertenecientes a las dos regiones adyacentes a dicho arco, por medio de un arco contenido en el plano P del papel. El papel queda así poligonalizado.

4. Las caras adyacentes a cada segmento de S se identifican por reflexión en el segmento, como lo indican las dobles flechas de la Figura 5c.

5. Sobre el plano P del papel suponemos que hay una 3-bola cuya frontera es el plano P del papel y por lo tanto está poligonalizada.

Esta 3-bola es la mariposa buscada, cuya frontera visualizamos en la Figura 5c. En efecto, al hacer las identificaciones en dicha bola obtenemos el cono inmerso en S^3 de la Figura 8, donde se visualiza claramente en nudo del ocho.

Definición 15. La mariposa obtenida con el algoritmo anterior la llamaremos *la mariposa asociada al diagrama del nudo*.

Nota 16. Según dicho algoritmo, por cada arco de la colección S , hay dos alas de la mariposa adyacentes a él, las cuales se identifican por reflexión a lo largo de dicho arco y tienen la forma que se muestra en la Figura 9.

Notación: En cada cara distinguiremos dos clases de vértices: aquellos que son extremos de los arcos del nudo, que denotaremos por A , y los que ya hemos denotado por B , los cuales se identifican todos con el vértice del cono (véase Figura 8), el cual también hemos marcado con B en dicha figura.

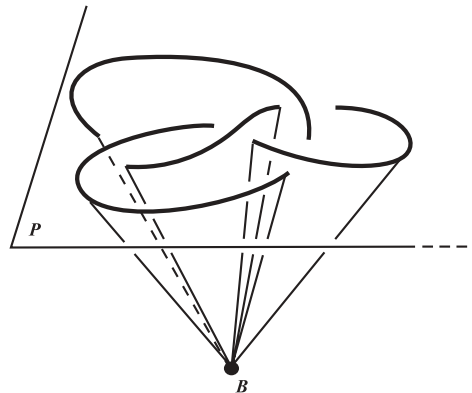


Figura 8.

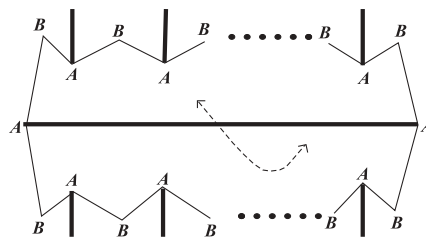


Figura 9.

3. Teorema principal

Primero estableceremos lo que entenderemos por una triangulación en una variedad y en una mariposa.

Definición 17. Sean M una n -variedad, $\Delta = \{T_i\}_{i \in I}$ una colección de n -simplex (véase [Mu]). Si $M = \bigcup_{i \in I} T_i$ y es tal que la intersección entre cada par T_i, T_j ($i \neq j$) es o vacía, o un único $(n - k)$ -simplex ($k < n$), se dice que $\Delta = \{T_i\}_{i \in I}$ es una *triangulación* de M .

Diremos que *una mariposa está triangulada* si tiene una triangulación Δ , tal que al realizar las identificaciones de la frontera, los k -simplex de Δ (contenidos en la frontera) se identifican con k -simplex de Δ (también contenidos en la frontera), para $k = 0, 1, 2$.

Una n -coloración de una triangulación Δ es una función $C : V \rightarrow \{1, 2, \dots, n\}$, donde V es el conjunto de vértices (0-simplex) de Δ , tal que dos vértices de un mismo simplex no tienen el mismo número (o color).

Al comienzo hemos definido lo que es un nudo 3-coloreable. En nuestro Teorema Principal relacionaremos la 3-coloración de un nudo con la 4-coloración de una mariposa asociada a éste. Definimos entonces lo que significa que estas dos coloraciones sean compatibles.

Definición 18. Sea D_K el diagrama de un nudo 3-coloreable, B_K la mariposa asociada a este diagrama. Si B_K está triangulada y 4-coloreada, entenderemos que la 4-coloración de la mariposa B_K es compatible con la 3-coloración del nudo (o que ambas coloraciones son compatibles) si cuando se identifican dos triángulos de la frontera de B_K por la reflexión en un arco cuya transposición asociada es $(i j)$, $i, j \in \{1, 2, 3\}$, entonces los colores de sus vértices realizan dicha transposición. En este caso el índice 4 siempre permanecerá fijo. Así, los vértices de color 4 se identifican sólo con vértices del mismo color.

Ejemplo 19. En la gráfica siguiente tenemos una mariposa asociada al nudo racional $\frac{3}{1}$, llamado trébol, la cual tiene una triangulación 4-coloreada compatible con la 3-coloración del nudo. Por simplicidad en el dibujo no mostramos sino los triángulos de la frontera.

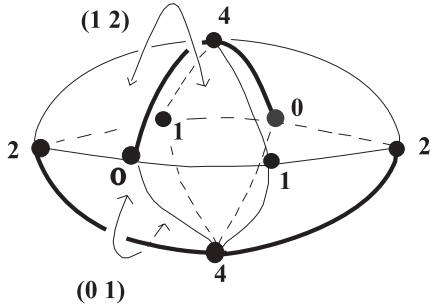


Figura 10.

En el teorema siguiente veremos que siempre es posible encontrar una mariposa de este estilo para cualquier nudo dado.

Teorema 20. [Teorema Principal] *La mariposa asociada al diagrama de un nudo K 3-coloreable es triangulable y 4-coloreable. Además ambas coloraciones son compatibles entre sí.*

Demostración. La prueba de este teorema es constructiva y la exhibiremos dando los pasos específicos de la construcción. Nuestro objetivo es encontrar una triangulación de la frontera de la mariposa que esté bien 4-coloreada y sea compatible con la 3-coloración del nudo K . Una prueba más detallada se encuentra en [HMTT1] y en [Te].

Primero construiremos una triangulación Δ de la frontera de la mariposa asociada al diagrama de K , la cual tendremos que refinar para garantizar que es 4-coloreable.

Paso 1. (Vértices y aristas de Δ .) Los vértices de Δ son todos los puntos de tipo A y B (según el algoritmo), además, un punto por cada arco del nudo, el cual es tomado en el interior de cada arco. Estos últimos puntos los llamaremos de tipo D . Las aristas son las curvas AB y AD comprendidas en las fronteras de las caras de la mariposa y, además, las siguientes curvas disjuntas (excepto en D) DA y DB , que pertenecen a las 2 caras de la mariposa que son adyacentes al arco que contiene a D . Las ilustramos en la figura siguiente:

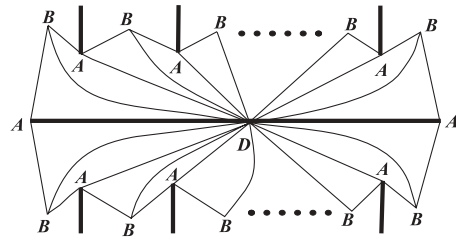


Figura 11.

Paso 2. (Se colorean los vértices de Δ .) A todos los vértices de tipo D se les asignará como color el número 4. Los vértices de tipo A se colorearán según la siguiente regla: si un vértice A pertenece a un arco cuya transposición asociada es $(i j)$, el número asociado será $k \in \{1, 2, 3\}$, $k \neq i$, $k \neq j$.

Los vértices de tipo B se colorean así: comenzamos con uno cualquiera y lo coloreamos con cualquiera de los números 1, 2 ó 3. Una vez coloreado este vértice, los colores de todos los otros quedan predeterminados de la siguiente manera: si R_1 y R_2 son dos regiones cualesquiera que tienen un arco común cuya transposición asociada es $(i j)$, y si se tiene la coloración $k \in \{1, 2, 3\}$ para el vértice $B(R_1)$, entonces el vértice $B(R_2)$ adquiere la coloración de la imagen de k mediante la acción de $(i j)$ que denotaremos por $(i, j)k$. Esta regla para colorear y el hecho de que saliendo de uno de los vértices de tipo B se puede pasar a cualquier otro saltando arcos en forma transversal, ya que K es un nudo, nos implica que es suficiente escoger el color de uno solo de los $B(R_i)$ para que el resto quede coloreado.

Se observa que por la forma en la que hemos coloreado los vértices de Δ y por las relaciones de Wirtinger, la 4-coloración de Δ es compatible con la 3-coloración del nudo.

De esta manera han quedado coloreados todos los vértices de la triangulación Δ , sin embargo, esta coloración puede no ser una buena, ya que puede ocurrir

que queden dos vértices de un mismo triángulo con el mismo color. Si esto ocurriera, dichos vértices deberán ser de tipo A y B , pues los vértices de tipo D son los únicos que tienen el color 4 y ellos no se conectan entre sí por aristas.

Paso 3. (Refinamos Δ .) Pensemos ahora que tenemos un vértice de tipo A (que está en el extremo de un arco cuya transposición asociada es $(i j)$), que tiene el color k , y que se conecta por una arista a un vértice de tipo B que tiene el mismo color k . Es fácil ver que el otro vértice de tipo B que está conectado al del tipo A debe tener también la coloración k , ya que k está fijo con $(i j)$ y, además, si la transposición asociada al arco que sobrepasa en dicho cruce es $(i' j')$ entonces el color de los vértices de tipo A y B del otro lado del cruce debe ser, para ambos, $(i' j')k$, es decir, ellos tendrán también un mismo color.

Con el fin de evitar dicha discrepancia debemos hacer algunas subdivisiones. Tomamos entonces, un punto en el interior de cada una de las aristas en donde se da una discrepancia, de tal manera que al hacer las reflexiones estos puntos resulten identificados (véase Figura 12). Dichos puntos los llamaremos de tipo C y los unimos con los vértices de tipo D de las dos caras adyacentes a la arista a la cual perteneciera el de tipo C .

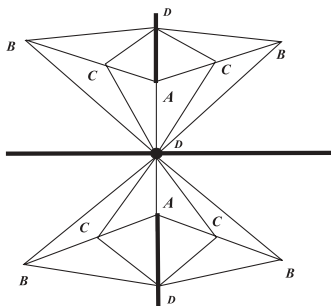


Figura 12.

Paso 4. (Coloreamos los vértices tipo C .) La coloración la hacemos independientemente en todo cruce donde había discrepancia de colores. Para colorear los cuatro puntos de tipo C adyacentes a un cruce, tomamos uno de ellos y le damos el color $i \in \{1, 2, 3\}$, donde i es diferente del color que tengan los extremos de la arista donde está localizado. Inmediatamente, el color de los otros tres puntos de tipo C (del mismo cruce) queda predeterminado, además, dicha coloración no predetermina la coloración de ningún otro punto del tipo C . Ilustremos la situación con la figura siguiente.

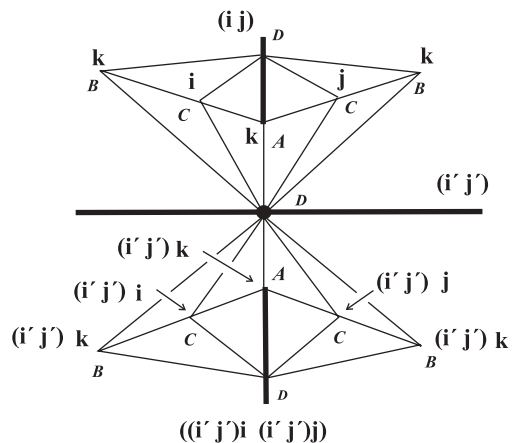


Figura 13.

De nuevo se tiene que, debido a las relaciones de Wirtinger, la coloración de los vértices de tipo C es compatible con la 3-coloración del nudo K .

Paso 5. (Se triangula el interior de la mariposa.) Como ya tenemos la frontera 4-coloreada, utilizamos el Teorema 10, de Goodman y Onishi, para extender esta triangulación y su coloración hacia el interior de la mariposa. \square

Corolario 21. Para toda 3-variedad existe al menos una mariposa triangulada y 4-coloreada asociada a ella.

Demostración. Esto es consecuencia del Teorema 8, de Hilden y Montesinos, y del teorema anterior. \square

Ilustraremos el teorema principal con el siguiente ejemplo.

Ejemplo 22. Tomemos el nudo 3-coloreable 6_1 de la Figura 3b. Usando el algoritmo contruimos la mariposa asociada a este diagrama (Figura 14). Nótese que por claridad en la ilustración hemos puesto el punto del infinito (que es de tipo B) varias veces, sin embargo es el mismo punto.

A continuación aplicamos los Pasos 1 y 2 del Teorema 20 y obtenemos la Figura 15. Luego aplicamos los Pasos 3 y 4 y encontramos la triangulación 4-coloreada de la frontera de la mariposa asociada al nudo 6_1 , la cual es compatible con la 3-coloración del nudo (véase Figura 16). La triangulación en el interior de la bola se obtiene usando el Teorema 10 de Goodman y Onishi.

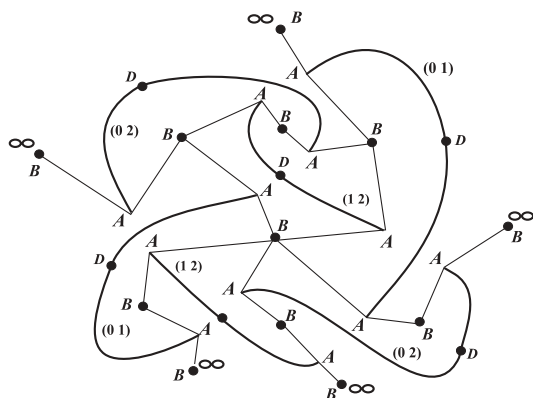


Figura 14.

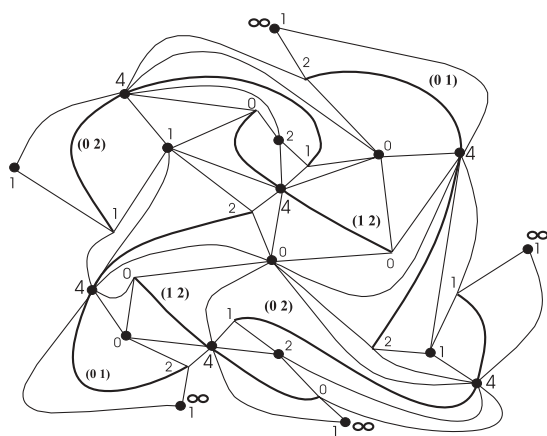


Figura 15.

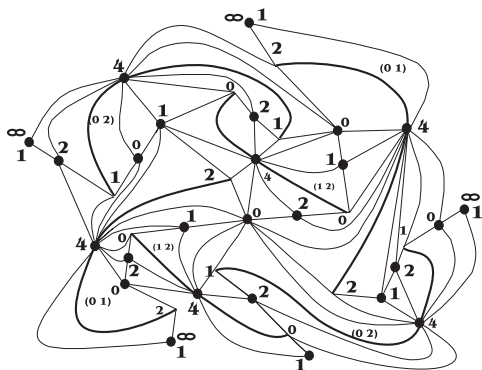


Figura 16.

Finalmente, podemos decir que el estudio de las mariposas abre nuevas perspectivas de trabajo en el estudio de 3-variedades y de nudos. Por ejemplo, se pueden plantear diversas preguntas que relacionen el número de alas de una mariposa asociada a un nudo con, por ejemplo, el número de puentes del nudo, respuestas a este estilo de preguntas permiten definir un invariante

nuevo para nudos (véase [HMTT2]). Por otro lado, las triangulaciones 4-coloreadas están relacionadas con los recientes trabajos de Izmetiev y Joswig ([IzJo]) y permiten probar al menos uno de sus teoremas en forma constructiva y mas simple (véase [HMTT1] y [Te]).

Agradecimientos. Este trabajo se ha realizado en el Grupo de Investigación en Matemáticas de la Universidad Nacional de Colombia Sede Medellín, y ha sido financiado parcialmente por COLCIENCIAS, proyecto 1118-05-13631 y por la DIME, Medellín.

Referencias

- [BuZi] **G. Burde and H. Zieschang**, *Knots*, Walter de Gruyter, New York, NY (1985).
- [CrFo] **R. Crowell and R. Fox**, *Introduction to knot theory*, Springer Verlag, New York, NY (1963).
- [FaSt] **D. Farmer and Th. Stanford**, *Knots and Surfaces*, Mathematical World, **6**, American Mathematical Society (1995).
- [Fo] **R. Fox**, *A quick trip through knot theory*, *Topology of 3-Manifolds and related topics*, Prentice-Hall, p.p.120-167 (1962).
- [GoOn] **J. Goodman and H. Onishi**, *Even Triangulations of S^3 and the coloring of graphs*. *Trans. Amer. Mat. Soc.* **246** (1978), 501–510 .
- [Hi] **M. H. Hilden**, *3-fold branched coverings of S^3* , *Amer. J. of Math.* **98** (1974), 989–997.
- [HMTT1] **M. H. Hilden, J. M. Montesinos, D. M. Tejada and M. M. Toro**, *Knots, Butterflies and 3-manifolds*, preimpreso (2003).
- [HMTT2] **M. H. Hilden, J. M. Montesinos, D. M. Tejada and M. M. Toro**, *Butterflies*, preimpreso (2003).
- [IzJo] **I. Izmetiev and M. Joswig**, *Branched Coverings defined by Triangulations*, por aparecer en *Adv. Geometry*, arXiv:math.GT/0108202v2 (2001).
- [Ka] **L. Kauffman**, *On Knots*, University Press, Princeton (1987).
- [Kaw] **A. Kawachi**, *A survey of Knot Theory*, Birkhauser, Basel, Switzerland (1996).
- [Mo1] **J. M. Montesinos**, *3-manifolds as 3-fold branched covers of S^3* , *Quart. J. Math.* **27** (1976), 85–94.
- [Mo2] **J. M. Montesinos**, *Sobre la conjetura de Poincaré y los recubridores ramificados sobre un nudo*, Tesis Doctoral. Univ. Complutense de Madrid (1972).
- [Mu] **J. R. Munkres**, *Elements of Algebraic Topology*, Addison-Wesley Publishing Company, Inc. New York, N.Y. (1984).
- [Ro] **D. Rolfsen**, *Knots and links*, Publish or Perish, Princeton (1985).
- [SeTh1] **H. Seifert and W. Threlfall**, *Old and new results on Knots*, *Canad. J. Math.*, **2** (1950), 1–15 .
- [SeTh2] **H. Seifert and W. Threlfall**, *A textbook of Topology*, Academic Press., New York-London (1980).
- [Te] **D. M. Tejada**, *Variedades, triangulaciones y representaciones*, preimpreso (2003).
- [Thu] **W. Thurston**, *Three-Dimensional Geometry and Topology*, Preprint (1990).

NUDOS COMBINATORIOS Y MARIPOSAS*

por

Margarita María Toro Villegas ¹

Dedicado a la memoria del profesor Jairo Charris C.

Resumen

Toro Villegas, Margarita María: Nudos combinatorios y mariposas. Rev. Acad. Colomb. Cienc. **28** (106): 79–86, 2004. ISSN 0370-3908.

En este artículo mostraremos una aplicación del concepto de nudo combinatorio para obtener la n -mariposa asociada a un diagrama de n cruces de un nudo K . Se describirán los algoritmos implementados en el programa *Mathematica*.

Palabras clave: Nudos, enlaces, combinatoria, 3-variedades.

Abstract

In this paper we use the concept of combinatory knot in order to construct a n -butterfly associated to a knot diagram with n crosses. We describe an algorithm and its implementation in the software *Mathematica*.

Key words: Knots, links, combinatory, 3-manifold.

1. Introducción

Este artículo tiene un carácter teórico-computacional. Nuestro resultado central, *todo nudo admite una presentación como una n -mariposa*, tiene una prueba algorítmica. Presentamos dicha prueba y desarrollamos e implementamos el algoritmo, utilizando el programa *Mathematica*. Para la implementación de dicho algoritmo, se utiliza el concepto de nudo combinatorio, desarrollado como herramienta computacional, ver [10] y

[11]. Se muestran mediante un ejemplo concreto, el nudo 4_1 o nudo del ocho, los resultados que se obtienen con nuestro programa. Este trabajo está estrechamente relacionado con el artículo *Mariposas y 3-Variedades*, ver [4], que aparece en este número de la Revista de la Academia Colombiana de Ciencias.

El artículo está organizado de la siguiente forma: en la primera sección definimos el concepto de nudo combinatorio. Como el trabajo está pensado para un público

¹Universidad Nacional de Colombia-Sede Medellín. e-mail: mmtoro@perseus.unalmed.edu.co

* Proyecto financiado parcialmente por COLCIENCIAS, proyecto 1118-05-13631, y por DINAIN, proyecto DI00C1236
AMS 2000 Mathematics Subject Classification: Primary 57M25, Secondary: 57M15, 05C99.

no especializado en los conceptos de la teoría de nudos, en esta sección damos algunas definiciones básicas de la teoría de nudos. En la segunda sección mostramos unos ejemplos motivadores del concepto de n -mariposa. En la tercera sección damos una definición intuitiva del concepto de n -mariposa, (para un desarrollo formal ver [4], [5]) y probamos el resultado teórico central. En la siguiente sección describimos el algoritmo que nos permite construir una n -mariposa a partir de un nudo, y desarrollamos su implementación en *Mathematica*. Se muestran unos cómputos concretos, para el nudo 4_1 . En la última sección se hacen comentarios y se establecen algunas preguntas interesantes relacionadas con el tema, que muestran la dirección actual de nuestra investigación.

2. Nudos Combinatorios

En esta sección definimos el concepto de nudo combinatorio, como una extensión del concepto de nudo. Para ello recordaremos algunos conceptos básicos; recomendamos el libro de Adams, [1], para una introducción general e intuitiva a la teoría de nudos. Para un tratamiento formal, recomendamos [2], [6] y [7].

Definición 1. $K \subset \mathbb{R}^3$ es un nudo si existe un homeomorfismo $\phi : S^1 \hookrightarrow K \subset \mathbb{R}^3$.

Un enlace es la imagen de una unión disjunta de copias de S^1 . En este artículo trataremos especialmente nudos, pero toda la teoría se puede aplicar a enlaces. Esto se hará en detalle en un trabajo posterior.

Definición 2. Dos nudos K y K' son equivalentes si existe un homeomorfismo $h : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ tal que $h(K) = K'$

Como una herramienta de trabajo para estudiar nudos se acostumbra utilizar la *proyección regular* del nudo en un plano, es decir, una proyección sobre un plano de tal forma que los segmentos que forman la proyección se intersecten transversalmente y la gráfica sólo tenga puntos dobles. Los siguientes son ejemplos de proyecciones de nudos.

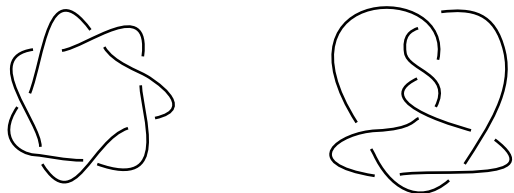


Figura 1

Dibujamos la proyección de tal manera, que en los puntos dobles, que llamaremos *cruces*, se indique, mediante una línea cortada, cuál segmento va por encima y cuál va por debajo. A los segmentos continuos, sin interrupción, se les acostumbra llamar *arcos*.

Los nudos se clasifican como “mansos” y “salvajes”. Los nudos mansos, o poligonales, son aquellos que admiten una proyección con un número finito de puntos doble. Nosotros estudiaremos sólo nudos mansos o poligonales. Es posible considerar nudos con o sin orientación, es decir, podemos asignarle al nudo un sentido en el cual recorrerlo. En ese caso, decimos que el nudo es orientado.

Al utilizar proyecciones de nudos, que se llaman también diagramas de nudos, cambiamos un poco el problema, y necesitamos garantizar que el estudio de los diagramas de nudos en efecto permite estudiar los nudos mismos. Afortunadamente tenemos el siguiente teorema, para la prueba ver [2].

Teorema 3. *Cualquier nudo poligonal K es equivalente, bajo una rotación arbitrariamente pequeña, a un nudo poligonal que admite una proyección regular.*

Estudiar los nudos según su proyección presenta una desventaja, que es fuente a su vez de muchas preguntas de la teoría: *Si dos nudos tienen la misma proyección, ellos son equivalentes, pero el recíproco no es cierto.* Por ejemplo, los dos diagramas siguientes corresponden al mismo nudo, que se acostumbra llamar *trivial* o *suelto*.



Figura 2.

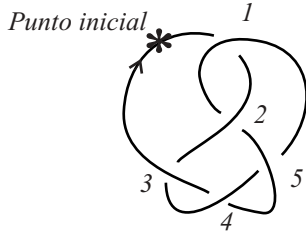
Además, Reidemeister resolvió el problema de la falta de unicidad al introducir los movimientos, conocidos hoy con su nombre, y probar el teorema: *Dos diagramas representan nudos equivalentes si y sólo si uno de ellos se puede deformar en el otro mediante una sucesión finita de movimientos de Reidemeister*, (ver [2]). Claro que éste es un resultado de carácter teórico, y en la práctica, dados dos diagramas no existe un algoritmo efectivo, que siempre permita decidir si los nudos que representan son o no equivalentes.

Este es uno de los problemas centrales de la teoría de nudos, que, aunque ha tenido muchos progresos en los últimos años, aún no está completamente resuelto.

A pesar de la dificultad que se presenta por la falta de unicidad al representar un nudo mediante un diagrama, históricamente los diagramas se han usado extensivamente para estudiar nudos.

Nosotros queremos desarrollar una codificación para los diagramas, que nos permita poder manipularlos en el computador y efectuar los cálculos y operaciones que están definidos para los diagramas de nudos.

Nuestra codificación consiste en tomar un punto inicial en el diagrama del nudo orientado y enumerar los cruces del nudo con enteros positivos $\{a_1, a_2, \dots, a_n\}$, típicamente usamos $\{1, 2, \dots, n\}$. Luego recorremos el nudo y escribimos una lista con los números de los cruces en el orden de recorrido, positivo si el cruce se recorre por encima y negativo si el cruce se recorre por debajo. Se termina al llegar de nuevo al punto inicial. De esta forma cada cruce se recorre dos veces y su correspondiente índice se escribe una vez con signo positivo y otra con signo negativo.



Cruces = $\{-1, 2, -3, 4, -5, 1, -2, 5, -4, 3\}$

Figura 3.

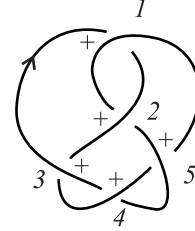
Como queremos que nuestra codificación determine orientación, adicionalmente a cada cruce se le asigna un signo, que está determinado por su orientación, siguiendo la regla de la mano derecha, de la forma indicada en la figura:



Figura 4.

La información sobre los signos de los cruces se almacena en una segunda lista, en la que cada posición representa el número del cruce. En la posición i -ésima se escribe un 1 si el cruce i -ésimo es positivo y un -1

si el cruce es negativo. Así, el nudo anterior tiene como codificación:



$\{-1, 2, -3, 4, -5, 1, -2, 5, -4, 3\}, \{1, 1, 1, 1, 1\}$

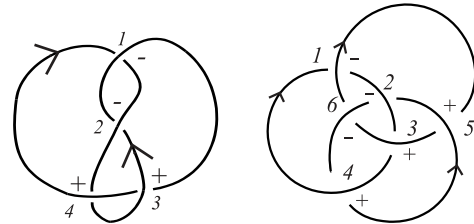
Figura 5.

Definición 4. La codificación de un diagrama de un nudo N es la asignación de una lista de listas $\{\{i_1, i_2, \dots, i_{2n}\}, \{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}\}$ tal que, como conjuntos,

$$\begin{aligned} \{i_1, i_2, \dots, i_{2n}\} &= \{a_1, a_2, \dots, a_n\} \cup \\ &\quad \{-a_1, -a_2, \dots, -a_n\}, \\ a_i \text{ un entero}, \epsilon_j &\in \{1, -1\}, \end{aligned}$$

A este tipo de listas las llamamos *nudos combinatorios*. Estos nudos combinatorios, aparte de permitirnos hacer muchos cálculos de operaciones e invariante de nudos, se han convertido en un objeto de estudio en sí mismos y son la fuente de preguntas muy interesantes, ver [10].

Los siguientes ejemplos muestran el nudo del ocho, o N_1 , y los anillos de Borromeo. Este último es un enlace, y lo que hacemos con nuestra codificación es que la lista de cruces a su vez está dividida en sub-listas, una por cada una de las componentes del enlace.



$N_1 =$ Nudo del ocho $N_2 =$ Anillos de Borromeo

Figura 6

Los respectivos nudos combinatorios son:

$$\begin{aligned} N_1 &= \{\{-1, 2, -4, 3, -2, 1, -3, 4\}, \{-1, -1, 1, 1\}\}. \\ N_2 &= \{\{\{-1, 2, -3, 4\}, \{1, -5, 3, -6\}, \{5, -2, 6, -4\}\}, \\ &\quad \{-1, -1, 1, 1, 1, -1\}\} \end{aligned}$$

La información referente a la orientación de cada cruce es altamente redundante, pero esta redundancia ayuda como forma de autocontrol. Note que como permitimos que el conjunto de índices para los cruces sea un subconjunto cualquiera S de \mathbb{N} , hay que tener cuidado con la sucesión que representa la orientación de los cruces, que tiene un sentido posicional: la orientación del cruce i se escribe en el puesto i . Por lo tanto, la lista de la orientación debe tener longitud igual a $Max\{S\}$, y contendrá información acerca de un conjunto de cruces ficticios. Esto no constituye ningún obstáculo en nuestro trabajo, al contrario, la posibilidad de trabajar

con cruces ficticios nos da una gran versatilidad para implementar conceptos de la teoría de nudos, como los movimientos de Reidemeister, ver [10].

Un aspecto fundamental que debemos considerar es que, dado un nudo, la codificación depende del punto inicial, de la forma de enumeración y, muy especialmente, del diagrama del nudo. Por ejemplo, el nudo $N1$ de la Figura 4, se puede codificar como $\{\{3, -2, 1, -3, 4, -1, 2, -4\}, \{-1, -1, 1, 1\}\}$. El problema con el punto de origen se resuelve definiendo clases de equivalencia en el conjunto de todos los nudos combinatorios, de la siguiente forma

$$\begin{aligned} \{\{i_1, i_2, \dots, i_{2n}\}, \{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}\} &\approx \{\{i_2, \dots, i_{2n}, i_1\}, \{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}\} \\ &\approx \dots \\ &\approx \{\{i_{2n}, \dots, i_1, i_2\}, \{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}\}. \end{aligned}$$

Esta equivalencia corresponde a la idea de lista circular, que es muy natural en *Mathematica*.

El problema de la forma de enumeración de los cruces, se resuelve mediante la relación de equivalencia en el conjunto de las codificaciones, determinada por:

$$\{\{i_1, i_2, \dots, i_{2n}\}, \{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}\} \approx \{\{j_1, j_2, \dots, j_{2n}\}, \{\delta_1, \delta_2, \dots, \delta_s\}\} \tag{1}$$

si y sólo si existe una biyección $\sigma : \{i_1, i_2, \dots, i_{2n}\} \rightarrow \{j_1, j_2, \dots, j_{2n}\}$ tal que, para todo $k, 1 \leq k \leq 2n$

$$\begin{aligned} \sigma(-i_k) &= -\sigma(i_k) \\ \epsilon_{|i_k|} &= \delta_{|\sigma(i_k)|}. \end{aligned} \tag{2}$$

Es decir, que exista una reenumeración de índices.

Tomando el conjunto de nudos combinatorios y partiéndolo por las relaciones de equivalencia definidas, hemos resuelto el problema de la selección de índices y del punto de origen. De ahora en adelante usamos el término nudo combinatorio para referirnos a su clase de equivalencia bajo estas relaciones.

La dependencia del diagrama específico del nudo es un problema que depende de la teoría de nudos, y no de la forma de codificación, y es el mismo con el que se encontraron los matemáticos a principios del siglo XX al querer estudiar los nudos vía su proyección en un plano. Resolvemos este problema utilizando una adaptación de los movimientos de Reidemeister.

2.1. Ejemplos Motivadores de Mariposas

Considere el enlace $N2$ de la Figura 4, que se conoce en la literatura con el nombre de anillos de Borromeo. Thurston probó que se puede obtener a partir de una poligonalización de la superficie de una bola, identificando

caras, mediante reflexiones a lo largo de aristas. En este caso tomamos un cubo y cada una de las caras las dividimos en dos, quedándonos así 12 caras. Hacemos identificación de las caras según lo indican las flechas,

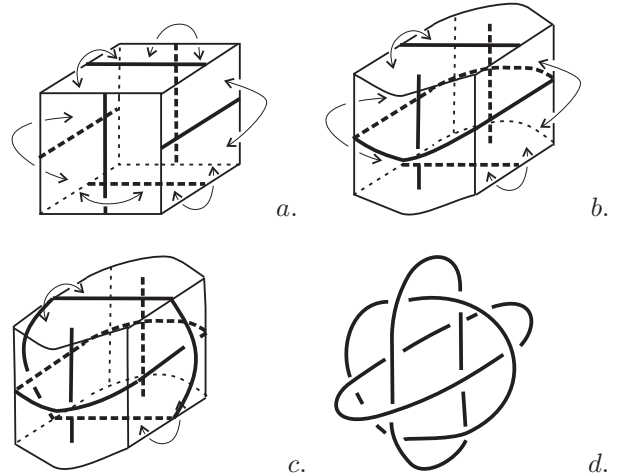


Figura 7

De la misma forma, todo nudo racional se puede obtener mediante un proceso similar, identificando caras, mediante reflexión a lo largo de aristas comunes, de la poligonalización de la superficie de una bola, esta vez mediante 4 hexágonos.

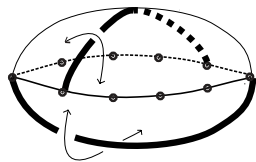


Figura 8

Si nombramos los vértices de los polígonos, tenemos un algoritmo que nos permite, mediante el uso de nudos combinatorios, producir los nudos o enlaces a partir de la poligonalización de la 3-bola y la identificación de las caras. Este algoritmo será presentado en un trabajo posterior, por ahora, nos interesa dedicarnos a la existencia de n -mariposas.

2.2. Resultado Central

Sin entrar en detalles formales, diremos que una n -mariposa es una 3-bola cuya frontera S está poligonalizada en $2n$ polígonos, que se van a identificar por parejas, mediante reflexiones a lo largo de la arista en común. Los polígonos de la mariposa los llamamos *alas*, la unión de las aristas a lo largo de las cuales se hace la identificación la llamamos *tronco*.

En este sentido, los ejemplos anteriores dan lugar a los resultados:

- a. *Todo nudo o enlace racional admite una representación como una 2-mariposa.*
- b. *Los anillos de Borromeo también admiten una representación como una 6-mariposa.*

El resultado teórico central de éste artículo es probar que todo nudo o enlace admite una presentación como una n -mariposa, para algún $n \in \mathbb{N}$.

La prueba es constructiva, y provee un algoritmo para construir la mariposa a partir de un diagrama D del nudo. Por ahora, trabajamos sólo con nudos, pero el trabajo se puede extender a enlaces, y probablemente a grafos.

Teorema 5. *Todo nudo admite una representación de n -mariposa, para algún $n \in \mathbb{N}$.*

Demostración: Sea K un nudo o enlace en $S^3 = \mathbb{R}^3 \cup \{\infty\}$, de tal manera que su proyección regular esté

totalmente contenida en un plano P unido con un número finito de pequeñas bolas, que son vecindades de sus cruces. Tomamos ahora un punto B que no pertenezca a la unión del plano P con las bolas alrededor de los cruces, y formamos el cono del nudo con vértice B .

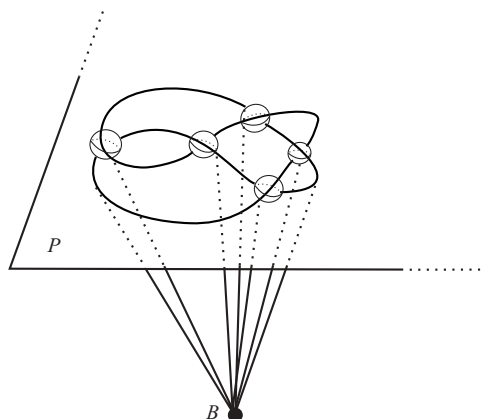


Figura 9

Al cortar a lo largo del cono, se obtiene precisamente una bola con caras, que al ser identificadas por reflexiones a lo largo de los arcos provenientes del diagrama del nudo, define la mariposa.

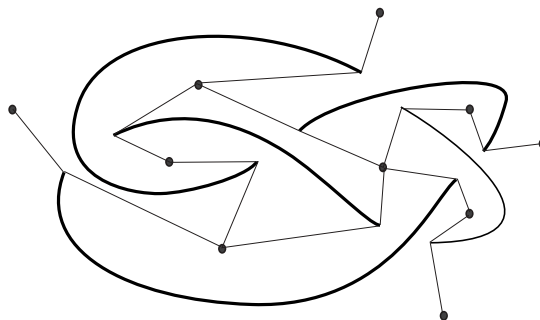


Figura 10

Nota: Esta prueba puede aplicarse también a enlaces, pero hay que hacer restricciones cuando la proyección del enlace no es conexa.

Corolario 6. *Un nudo no trivial K con n -cruces admite una representación de n -mariposa.*

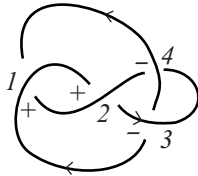
La prueba de este resultado es constructiva, y permite dar un algoritmo para construir la mariposa a partir de un diagrama D_K del nudo K orientado. El número de mariposa coincide con el número de cruces del diagrama.

Este algoritmo es el que queremos implementar en *Mathematica*. Lo interesante es que al implementar el algoritmo, se van aclarando algunos puntos teóricos y surgen preguntas nuevas.

3. Implementación del algoritmo

Simplemente reproducimos la prueba del teorema anterior, en forma algorítmica. Para cada uno de los pasos del algoritmo describimos cómo se hizo la implementación. No mostramos códigos, pero sí los resultados obtenidos por los programas para el nudo del ocho, con el que se ilustra el algoritmo.

Paso 1. Tomamos un diagrama D_K del nudo K , que es una colección de arcos disjuntos orientados. Para implementarlo en el computador, tomamos el nudo combinatorio.



$$\text{ocho} = \{\{1, -2, 3, -4, 2, -1, 4, -3\}, \{1, 1, -1, -1\}\}$$

Figura 11

Note que el nudo que estamos usando para los ejemplos, es la imagen espejo del que usamos en la Figura 5.

Paso 2. Tomamos las regiones R_i , determinadas por la proyección del nudo en el plano, elegimos un punto en cada región, que denotamos $B[i]$. En la región no acotada, tomamos el punto en el infinito.

Para determinar las regiones del plano, la información dada por el nudo combinatorio no es suficiente y, por lo tanto, debemos hacer algunos ajustes en la codificación.

Cambio de la codificación: Reemplazamos el cruce i por $c[i]$, e introducimos dos nuevos vértices: $a[i,-1]$ y $a[i,1]$, que son puntos correspondientes al cruce i cuando se recorre por debajo, con la convención que $a[i,-1]$ se recorre antes que $a[i,1]$. Al identificar las alas, estos puntos quedan identificados. Al crear la nueva lista, incorporamos la información de los signos de los cruces en el signo que le asignamos a $c[i]$. No se crea confusión, ya que la información de cuándo se pasa por encima y cuándo por debajo ya está consignada en los vértices de tipo a .

Creamos una nueva lista en la cual efectuamos los reemplazos

$$i \rightarrow -c[i]*\text{signocruce}(i)$$

$$-i \rightarrow (a[i,-1],c[i]*\text{signocruce}(i),a[i,1])$$

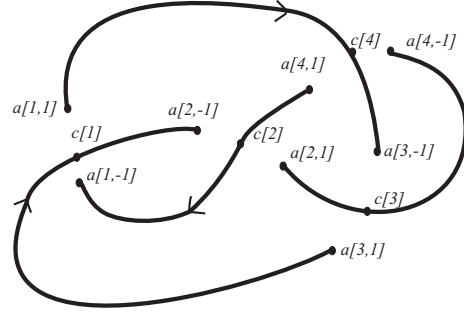


Figura 12

El programa que hace ésto ya está implementado para trabajar con enlaces. El resultado en el caso del nudo del ocho es:

$$\text{cambiar notacion[ocho]} = \{-c[1], a[2, -1], c[2], a[2, 1], c[3], a[4, -1], -c[4], a[4, 1], -c[2], a[1, -1], c[1], a[1, 1], c[4], a[3, -1], -c[3], a[3, 1]\}$$

Para producir las regiones en las que la proyección del nudo divide el plano, construimos un complejo simplicial, en el cual los vértices son todos los vértices de la codificación.

Los lados son las parejas de vértices consecutivos, recordando que la lista es circular. En nuestro caso tenemos:

$$\text{lados} = \{\{-c[1], a[2, -1]\}, \{a[2, -1], c[2]\}, \{c[2], a[2, 1]\}, \{a[2, 1], c[3]\}, \{c[3], a[4, -1]\}, \{a[4, -1], -c[4]\}, \{-c[4], a[4, 1]\}, \{a[4, 1], -c[2]\}, \{-c[2], a[1, -1]\}, \{a[1, -1], c[1]\}, \{c[1], a[1, 1]\}, \{a[1, 1], c[4]\}, \{c[4], a[3, -1]\}, \{a[3, -1], -c[3]\}, \{-c[3], a[3, 1]\}, \{a[3, 1], -c[1]\}\}$$

Las caras del complejo simplicial se construyen de la siguiente forma: Empezamos a recorrer los lados, en sentido positivo, si se llega a un vértice de tipo $a[i,\pm 1]$ continuamos con el vértice que sigue en la codificación, pero si llegamos a un vértice de tipo $c[i]$ “viramos a la izquierda”, como se muestra en la figura 13.

Repetimos este proceso hasta haber recorrido todos los lados 2 veces, una en sentido positivo y otra en sentido negativo. De esta forma hemos construido una partición del plano en $n+2$ regiones.

Implementamos este proceso en un programa, llamado **producir regiones mar**, que toma el nudo combinatorio y produce las regiones.

El programa **producirregionesmar** de hecho está diseñado para trabajar con cualquier nudo (o enlace) combinatorio, y calcula una poligonalización de una superficie orientable de genero mínimo, sobre la cual se puede proyectar el nudo combinatorio.

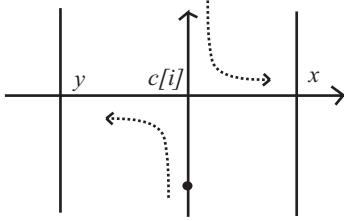


Figura 13

En nuestro caso tenemos

$$\begin{aligned} \text{regiones} = & \{ \{a[2, -1], c[2], a[4, 1], c[4], a[1, 1], c[1]\}, \\ & \{a[1, -1], c[2], a[2, -1], c[1]\}, \\ & \{c[1], a[3, 1], c[3], a[2, 1], c[2], a[1, -1]\}, \\ & \{c[4], a[4, -1], c[3], a[3, 1], c[1], a[1, 1]\}, \\ & \{c[3], a[3, -1], c[4], a[4, 1], c[2], a[2, 1]\}, \\ & \{c[3], a[4, -1], c[4], a[3, -1]\} \end{aligned}$$

Nótese que son 6 regiones.

Por cada una de estas regiones tomamos un vértice $b[i]$. No es relevante saber cual corresponde a la región no acotada. Para simplicidad del diagrama, repetimos este vértice.

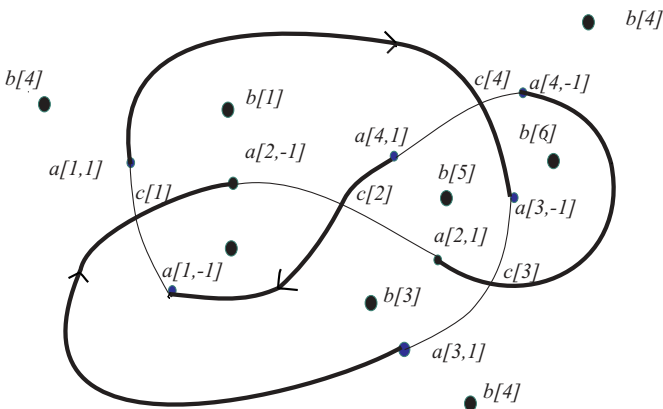


Figura 14

Paso 3. Los extremos de cada uno de los arcos del diagrama, que llamamos vértices de tipo A , los unimos con los puntos $B(i)$, de las dos regiones adyacentes al arco. Así el plano está poligonalizado. Este proceso lo implementamos en dos procesos independientes:

a. Producir los arcos, que serán el Tronco de la mariposa. Se usa el programa **prodarcos**, que a partir de un nudo combinatorio produce los arcos.

Los arcos se determinan por sus extremos, pero para poder construir los polígonos, debemos tener en cuenta sobre cuales cruces pasa cada arco. En este caso no importa el signo de los vértices de tipo $c[i]$.

$$\text{arcos} = \{ \{a[2, 1], c[3], a[4, -1]\}, \{a[4, 1], c[2], a[1, -1]\}, \{a[1, 1], c[4], a[3, -1]\}, \{a[3, 1], c[1], a[2, -1]\} \}$$

b. Construir las alas. Se usa un programa, que requiere como parámetros de entrada las regiones y los arcos, escritos con la convención indicada.

Eliminamos de las regiones los vértices de la forma $c[i]$, y se construye la poligonalización de la superficie uniendo, para cada región j , el punto $b[j]$ con los puntos extremos de los lados que forman la región. Estos son los puntos extremos de los arcos, que son del tipo $a[i, \pm 1]$. De esta forma se construyen los lados de los polígonos. El tronco está formado por los arcos, al eliminar los vértices de la forma $c[i]$.

En nuestro caso tenemos la 4-mariposa dada por:

mariposa:

$$\begin{aligned} & \{ \{a[2, 1], a[4, -1]\}, \{b[3], a[2, 1]\}, \\ & \{b[3], a[3, 1]\}, \{b[4], a[3, 1]\}, \{b[4], a[4, -1]\} \}, \\ & \{ \{a[2, 1], a[4, -1]\}, \{b[5], a[2, 1]\}, \\ & \{b[5], a[3, -1]\}, \{b[6], a[3, -1]\}, \{b[6], a[4, -1]\} \}, \\ & \{ \{a[4, 1], a[1, -1]\}, \{b[1], a[4, 1]\}, \{b[1], a[2, -1]\}, \\ & \{b[2], a[2, -1]\}, \{b[2], a[1, -1]\} \}, \\ & \{ \{a[4, 1], a[1, -1]\}, \{b[5], a[4, 1]\}, \{b[5], a[2, 1]\}, \\ & \{b[3], a[2, 1]\}, \{b[3], a[1, -1]\} \}, \\ & \{ \{a[1, 1], a[3, -1]\}, \{b[1], a[1, 1]\}, \{b[1], a[4, 1]\}, \\ & \{b[5], a[4, 1]\}, \{b[5], a[3, -1]\} \}, \\ & \{ \{a[1, 1], a[3, -1]\}, \{b[4], a[1, 1]\}, \{b[4], a[4, -1]\}, \\ & \{b[6], a[4, -1]\}, \{b[6], a[3, -1]\} \}, \\ & \{ \{a[3, 1], a[2, -1]\}, \{b[3], a[3, 1]\}, \{b[3], a[1, -1]\}, \\ & \{b[2], a[1, -1]\}, \{b[2], a[2, -1]\} \}, \\ & \{ \{a[3, 1], a[2, -1]\}, \{b[4], a[3, 1]\}, \{b[4], a[1, 1]\}, \\ & \{b[1], a[1, 1]\}, \{b[1], a[2, -1]\} \} \end{aligned}$$

$$\text{tronco} = \{ \{a[2, 1], a[4, -1]\}, \{a[4, 1], a[1, -1]\}, \{a[1, 1], a[3, -1]\}, \{a[3, 1], a[2, -1]\} \}$$

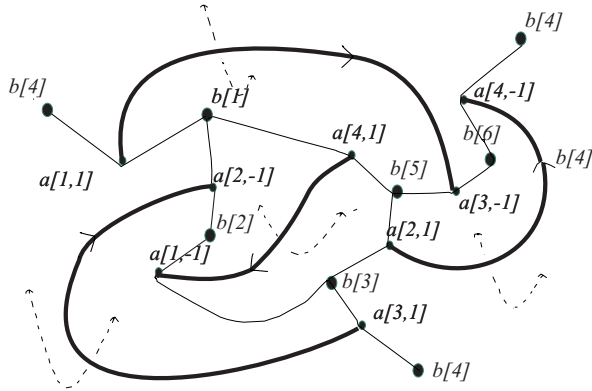


Figura 15

Paso 4. Sobre el plano P suponemos que hay una 3-bola cuya frontera es el plano poligonizado.

Paso 5. Las caras adyacentes a cada arco α se identifican por una reflexión a lo largo de α , para cada arco α que se encuentre en el tronco. Para construir la mariposa, se identifican las alas simétricas a lo largo del arco común.

Fin del algoritmo

4. Comentarios

Este trabajo presenta muchas ideas nuevas que aún se están desarrollando. Se presentan los conceptos de una forma intuitiva, pero todos pueden ser formalizados, ver [5]. Se tienen una gran cantidad de preguntas sobre las cuales estamos trabajando. Por ejemplo:

¿Cómo reducir el número de alas de una n -mariposa?

¿Qué variaciones hay al extender los resultados a enlaces y a grafos?

Vimos en los ejemplos que los vértices de la n -mariposa tienen distinto comportamiento. ¿Cómo clasificar los vértices?

Se sabe que todo diagrama de un nudo da lugar a un nudo combinatorio, pero existen nudos combinatorios que no provienen de diagramas. ¿Para este tipo de nudos combinatorios, qué sentido tiene aplicar los algoritmos descritos en este trabajo?

Referencias

- [1] **Adams, C. C.** *The Knot Book: An elementary introduction to the Mathematical Theory of Knots*. W. H. Freeman and Company-New York, 2001
- [2] **Burde, G. and Zieschang, H.** *Knots*, Walter de Gruyter, Berlin-New York, 1985.
- [3] **Cairns, G. and Elton, D.** *The Planarity Problem for Signed Gauss Words*, Journal of Knot Theory and its Ramifications, **Vol. 2**, No. 4 (1993) 359-367.
- [4] **Hilden, M., Montesinos, J., Tejada, D. y Toro, M.**, *Mariposas y 3-Variedades*, Festival Académico, dedicado a la memoria del Profesor J. Charris, Universidad Nacional de Colombia, Bogotá, agosto, 2003.
- [5] **Hilden, M., Montesinos, J., Tejada, D. y Toro, M.**, *Butterflies*, Preprint, 2003
- [6] **Murasugi, K.** *Knot Theory and its Applications*, Birkhäuser, Boston-Basel-Berlin, 1996.
- [7] **Rolfsen, D.** *Knot and Links*, Publish or perish, Inc., 1976.
- [8] **Tejada, D.** *Variedades, Triangulaciones y Representaciones*, Preprint, 2003.
- [9] **Thurston, W.** *Three Dimensional Geometry and Topology*, Princeton, NJ, 1997.
- [10] **Toro, M.** *Programación en Mathematica con aplicaciones a la Teoría de Nudos*, Preprint, 2003.
- [11] **Toro, M.** *Clasificación de nudos: Una aproximación computacional*. Preprint, 2002.

UNA LÓGICA MODAL PARA LA GEOMETRÍA PLANA DE LOBACHEVSKI

por

Alfonso Ríder Moyano¹ & Rafael María Rubio Ruiz²

Dedicado a la memoria del profesor Jairo Charris C.

Resumen

Ríder Moyano, Alfonso & Rubio Ruiz, Rafael María : Una lógica modal para la geometría plana de Lobachevski. Rev. Acad. Colomb. Cienc. **28** (106): 87–94, 2004. ISSN 0370-3908.

Habitualmente, las geometrías de incidencia están basadas en estructuras bisurtidas formadas por puntos y rectas, y conectadas por una relación entre ambas clases. En lo que sigue, introducimos una estructura monosurtida, que llamamos marco de Lobachevski, la cual resulta adecuada, para construir una base semántica que permita su consideración en el lenguaje modal. Construiremos así un sistema axiomático para dicho lenguaje, que estará determinado por la estructura creada, es decir probaremos su corrección y completitud.

Palabras clave: Lógica modal, geometría de Lobachevski.

Abstract

Usually, the incidence geometries are based in two-sorted structures formed of points, lines connected each other for a relationship. Furthermore, we introduce a one-sorted structure that we call Lobachevski frame which results appropriate to build a semantic base that allows its consideration in modal language. We build an axiomatic system for that language will be determined by the structure created.

Key words: Modal Logic, Lobachevski Geometry

¹Departamento de Matemáticas, Universidad de Córdoba, Córdoba, España, email: ma1rimoa@uco.es

²Departamento de Matemáticas, Universidad de Córdoba, Córdoba, España, email: ma1rurur@uco.es

AMS 2000 Mathematics Subject Classification: AMS Subject Class. (2000): 03B45, 51A15.

1. Introducción

El desarrollo de sistemas lógicos deductivos, susceptibles de ser utilizados como herramientas para la ingeniería del conocimiento, se ha convertido actualmente en una de las líneas más prolíferas de investigación en lógica matemática. En particular, los sistemas basados en lenguajes modales, resultan especialmente adecuados para las ciencias de la computación. Bajo estos enfoques, han sido ampliamente estudiados sistemas que modelizan las relaciones temporales en sus diversas interpretaciones y aspectos de uso. En los últimos años, empiezan a ser también numerosos los autores que se han interesado en el estudio de las propiedades y relaciones geométricas, bajo el prisma de dichos lenguajes modales, si bien es claro, que el estudio de las propiedades y relaciones espaciales resulta ser considerablemente más complejo que el caso del tiempo. Esto es debido, al hecho, de que aquellas estructuras suelen estar compuestas por diversas clases de objetos (multisurtidas): puntos, líneas, planos, etc., estando además ligados tales objetos mediante relaciones, a su vez variadas: colinealidad, paralelismo, ortogonalidad, incidencia, etc.

En este trabajo, siguiendo la estela marcada por distintos autores (véase [3],[17]), construimos una estructura monosurtida (es decir, con una única clase de objetos), que resulta ser adecuada para expresarla en un lenguaje modal, y capaz de capturar de forma equivalente, los modelos de la geometría plana de Lobachevski.

2. Axiomática del plano de Lobachevski

Definición 1. Una terna (P, L, in) , donde P y L son dos conjuntos e $in \subset A \times B$ una relación binaria entre ellos, diremos que es un plano de incidencia si se verifican los siguientes axiomas:

- (IE0) $P \cap L = \emptyset$
- (IE1) $(\forall x, y \in P)(\exists z \in L)(x \text{ in } z \wedge y \text{ in } z)$
- (IE2) $(\forall x, y \in P)(\forall z, t, w \in L)((x \text{ in } z \wedge y \text{ in } z \wedge x \text{ in } t \wedge y \text{ in } t) \rightarrow (x = y \vee z = t))$
- (IE3) $(\forall z \in L)(\exists x, y \in P)(x \neq y \wedge x \text{ in } z \wedge y \text{ in } z)$
- (IE4) $(\forall x \in P)(\exists z, t \in L)(z \neq t \wedge x \text{ in } z \wedge x \text{ in } t)$

Si $x \text{ in } z$ diremos que el punto x es incidente con la recta z , o de forma coloquial, que la recta z pasa por el punto x .

La geometría de incidencia es la estructura geométrica mínima, susceptible de ser ampliada con nuevos axiomas. Una posible ampliación es la inserción de

axiomas de paralelismo. Así definiremos la relación “ser paralela” del siguiente modo:

Sean $z, t \in L$,

$$z \parallel t \text{ si y solo si } (\forall x \in P)(x \text{ in } z \wedge x \text{ in } t \rightarrow z = t)$$

Con esta nueva relación, se pueden construir esencialmente, dos tipos diferenciados de geometrías planas en el siguiente sentido:

Dados $x \in P$ y $z \in L$, tales que x no incida en z :

- (P1) Existe una única recta t tal que x incide en t y z es paralela a t (geometría afín).
- (P2) Existen dos rectas distintas t y l paralelas a z e incidentes con x (geometría de Lobachevski).

Definición 2. Una estructura $S = (P, L, in, \parallel)$, siendo P y L dos conjuntos, e $in, \parallel \subset P \times L$ dos relaciones entre ambos, diremos que es un plano de Lobachevski si (P, L, in) es un plano de incidencia, y además se verifica que:

$$(II) (\forall x \in P)(\forall z \in L)(\exists t, l \in L)((t \neq l) \wedge (z \parallel t) \wedge (z \parallel l) \wedge (x \text{ in } t) \wedge (x \text{ in } l) \vee (x \text{ in } z))$$

Definición 3. Sean

$$S = (P, L, in, \parallel) \quad \text{y} \quad S' = (P', L', in', \parallel')$$

dos planos de Lobachevski. Una aplicación

$$f : P \cup L \mapsto P' \cup L'$$

se dirá que un homomorfismo de S en S' si:

- (i) Para todo $x \in P$ y todo $z \in L$ se cumple que $f(x) \in P'$ y $f(z) \in L'$
- (ii) Para todo $x \in P$ y todo $z \in L$, si $x \text{ in } z$ entonces $f(x) \text{ in}' f(z)$
- (iii) Para toda pareja $z, z' \in L$, si $z \parallel z'$ entonces $f(z) \parallel f(z')$

En lo que sigue, denotaremos por Σ_i la categoría de los planos de Lobachevski (PL), donde sus objetos vienen dados por los planos de Lobachevski, y cuyos morfismos se acaban de definir arriba.

3. Marco de Lobachevski

Construiremos ahora una nueva categoría compuesta de estructuras monosurtidas y la cual resultará ser equivalente a la anterior.

Sea $S = (P, L, in, \parallel)$ un plano de Lobachevski y consideremos el subconjunto W del producto cartesiano $P \times L$, dado por $W = \{(x, z) \mid x \in P, z \in L, x \in z\}$. Sobre W , definiremos las siguientes relaciones binarias:

1. (\equiv_1) Diremos que $(x_1, z_1) \equiv_1 (x_2, z_2)$ si y solo si $x_1 = x_2$.
2. (\equiv_2) Analogamente se dirá que $(x_1, z_1) \equiv_2 (x_2, z_2)$ si y solo si $z_1 = z_2$.
3. (\parallel) Por último, abusando de nuestra notación, diremos que $(x_1, z_1) \parallel (x_2, z_2)$ si y solo si $z_1 \parallel z_2$.

Desde un punto de vista intuitivo, los elementos de W pueden considerarse como puntos o rectas dependiendo del contexto en que aparezcan, es decir $x \equiv_1 y$ significará que considerados como puntos son iguales y de manera análoga al leer $x \equiv_2 y$ interpretaremos que x e y son la misma recta.

Definición 4. La estructura $W(S) = (W, \equiv_1, \equiv_2, \parallel)$, se denominará marco geométrico asociado al plano de Lobachevski $S = (P, L, in, \parallel)$, o sencillamente marco geométrico sobre S .

Sobre $W(S)$ daremos una nueva relación “(on)” a partir de \equiv_1 y \equiv_2 de la siguiente forma:

$$(x_1, z_1) \text{ on } (x_2, z_2) \text{ si y solo si } x_1 \in z_2$$

Tanto la relación *on* como su inversa on^{-1} puede definirse como composición de \equiv_1 y \equiv_2 de manera alternada.

Lema 1. Sea S un plano de incidencia y $W = W(S)$ el marco geométrico asociado a S . Entonces para todo $x, y \in W$ se tiene:

- (i) $x \text{ on } y$ si y solo si $(\exists z \in W)(x \equiv_1 z \wedge z \equiv_2 y)$.
- (ii) $x \text{ on}^{-1} y$ si y solo si $(\exists z \in W)(x \equiv_2 z \wedge z \equiv_1 y)$.

Demostración. Basta utilizar la definición de *on* y on^{-1} . \square

Lema 2. Sea S un plano de Lobachevski y $W = W(S)$ el marco geométrico sobre S . Entonces:

(ML0) \equiv_1 y \equiv_2 son relaciones de equivalencia tales que

$$(\forall x, y \in W)(x \equiv_1 y \wedge x \equiv_2 y \rightarrow x = y)$$

(ML0*) $z \parallel t \leftrightarrow (\forall x \in W)(x \text{ on } z \wedge x \text{ on } t \rightarrow z \equiv_2 t)$

(ML1) $(\forall x, y \in W)(\exists z \in W)(x \text{ on } z \wedge y \text{ on } z)$

(ML2) $(\forall x, y, z, t \in W)(x \text{ on } z \wedge y \text{ on } z \wedge x \text{ on } t \wedge y \text{ on } t) \rightarrow$

$$(x = y \vee z = t)$$

(ML3) $(\forall z \in W)(\exists x, y \in W)(x \not\equiv_1 y \wedge x \text{ on } z \wedge y \text{ on } z)$

(ML4) $(\forall x \in W)(\exists z, t \in W)(z \not\equiv_2 t \wedge x \text{ on } z \wedge x \text{ on } t)$

(ML5) $(\forall x, z \in W)(\exists t, l \in W)(t \not\equiv_2 l \wedge z \parallel t \wedge z \parallel l \wedge x \text{ on } t \wedge x \text{ on } l)$

Demostración. Es de comprobación directa, a partir de las definiciones. \square

Definición 5. LLamaremos marco de Lobachevski a cualquier estructura de la forma $W_F = (W, \equiv_1, \equiv_2, \parallel)$, donde W es un conjunto no vacío y $\{\equiv_1, \equiv_2, \parallel\}$ relaciones binarias en W_F y para la cual además son válidos los axiomas (ML0), ..., (ML5) del lema anterior.

Después de esta definición los marcos geométricos sobre un plano de Lobachevski serán, en particular, marcos de Lobachevski. En lo que sigue, denotaremos mediante Φ_i a la clase formada por los marcos de Lobachevski (ML).

Haciendo un abuso de notación, designaremos también mediante Φ_i , a la categoría cuyos objetos son los marcos de Lobachevski y cuyos morfismos vienen dados por las aplicaciones

$$f : W_F = (W, \equiv_1, \equiv_2, \parallel) \mapsto W'_F = (W', \equiv'_1, \equiv'_2, \parallel')$$

tales que, para todo $x, y \in W$, verifican que:

- (i) si $x \equiv_i y$, entonces $f(x) \equiv'_i f(y)$, $i = 1, 2$
- (ii) si $z \parallel z'$, entonces $f(z) \parallel f(z')$.

4. Equivalencia entre las categorías plano y marco de Lobachevski

Procediendo de forma inversa a como hicimos en el apartado anterior, para cada marco de Lobachevski $W_F = (W, \equiv_1, \equiv_2, \parallel)$, construiremos un plano de Lobachevski asociado a él que denominaremos plano de Lobachevski sobre W_F y denotaremos por $S(W_F)$.

A tal efecto, procedemos con las siguientes definiciones. Consideremos $W_F = (W, \equiv_1, \equiv_2)$ un marco de Lobachevski y dado $x \in W$, denotaremos por $|x|_i, i = 1, 2$ la clase de equivalencia de x determinada por la relación $\equiv_i, i = 1, 2$. Definamos ahora los siguientes conjuntos y relaciones:

$$P_S(W_F) = P = W/\equiv_1 = \{|x|_1 : x \in W\}$$

$$L_S(W_F) = L = W/\equiv_2 = \{|z|_2 : z \in W\}$$

$$in_S(W_F) = in = \{(|x|_1, |z|_2) : x \in W, z \in W, x \text{ on } z\}$$

$$\parallel_S(W_F) = \parallel = \{(|z|_2, |z'|_2) : z \in W, z' \in W, z \parallel z'\}$$

Si x, x' representantes de $|x|_1$, y z, z' representantes de $|z|_2$ tales que $x \text{ on } z$, utilizando el lema 1 aseguramos la existencia de $x_1 \in W_F$ cumpliendo

$$x' \equiv_1 x \equiv_1 x_1 \equiv_2 z \equiv_2 z'$$

y al ser \equiv_1 y \equiv_2 relaciones de equivalencia, concluimos que x' on z' . Lo cual prueba que la relación *in* está bien definida.

Lema 3. *Sea W_F un marco de Lobachevski, entonces $S(W_F)$ es un plano de Lobachevski.*

Demostración. Se deduce de forma directa, de la propia construcción. \square

Teorema 1. *Para cada marco de Lobachevski W_F , el marco W'_F , construido como $W(S(W_F))$ es isomorfo a W_F .*

Demostración. Basta comprobar que la aplicación $f : W_F \mapsto W'_F$ dada como

$$f(x) = (|x|_1, |x|_2) \quad \forall x \in W$$

está bien definida y es un isomorfismo. \square

Teorema 2. *Para cada plano de Lobachevski S , el plano S' , construido como $S(W(S))$ es isomorfo a S .*

Demostración. Definamos la aplicación $f : S \mapsto S'$, que a cada elemento $x \in P$ y $z \in L$ le asigna

$$f(x) = |(x, z_1)|_1$$

$$f(z) = |(x_1, z)|_2$$

donde la existencia de $z_1 \in L$ con $x \in z_1$ nos la garantiza el axioma (IE4) y la de $x_1 \in P$ verificando $x_1 \in z$, el axioma (IE3).

No Es difícil comprobar que la aplicación f es un isomorfismo. \square

Siendo Σ_i y Φ_i las categorías de los planos y marcos de Lobachevski respectivamente, las construcciones hechas anteriormente, nos permiten definir sendas aplicaciones $W : \Sigma_i \rightarrow \Phi_i$ y $S : \Phi_i \rightarrow \Sigma_i$ entre los objetos de ambas categorías.

Sea ahora f , un homomorfismo entre dos planos de S y S' , y consideremos la aplicación

$$W(f) : W(S) \mapsto W(S')$$

que viene dada del siguiente modo:

Si $(x, z) \in W(S)$, entonces $W(f)((x, z)) = (f(x), f(z))$

Con esta definición es claro que el digrama que sigue será conmutativo, convirtiendo a W en un funtor entre

las categorías Σ_i y Φ_i .

$$\begin{array}{ccc} S & \xrightarrow{f} & S' \\ W \downarrow & & \downarrow W \\ W(S) & \xrightarrow{W(f)} & W(S') \end{array}$$

Analogamente, si $f : W \mapsto W'$ es un homomorfismo de marcos de Lobachevski, la aplicación

$$S(f) : S(W_F) \mapsto S(W'_F)$$

definida mediante

$$S(f)(|x|_1) = |f(x)|_1, S(f)(|z|_2) = |f(z)|_2$$

hace conmutativo el siguiente diagrama y convierte a S en un funtor entre las categorías Φ_i y Σ_i .

$$\begin{array}{ccc} W_F & \xrightarrow{f} & W'_F \\ S \downarrow & & \downarrow S \\ S(W_F) & \xrightarrow{S(f)} & S(W'_F) \end{array}$$

Para concluir esta sección, diremos que en estas condiciones, un bien conocido teorema de la teoría de categorías (véase [11]), nos permite asegurar que Σ_i y Φ_i conforman categorías equivalentes.

5. Lenguaje

Consideremos un lenguaje multimodal con operadores monarios

$\{[\equiv_1], [\equiv_2], [\neq_1], [\neq_2], [|||]\}$, un conjunto infinito numerable de literales proposicionales $\{p_1, p_2, \dots\}$, junto con las constantes \perp, \top y las conectivas usuales $\{\neg, \vee, \wedge, \rightarrow, \leftrightarrow\}$.

El conjunto de *fbf* se define de forma usual. Denotaremos con letras latinas mayúsculas A, B, C, \dots a las fórmulas multimodales de este lenguaje. Además, para simplificar la escritura definimos los siguientes operadores modales secundarios:

- $[\neq]A = [\neq_1]A \wedge [\neq_2]A$
- $[U]A = A \wedge [\neq]A$
- $OA = A \wedge [\neq]\neg A$
- $[on]A = [\equiv_1][\equiv_2]A$
- $[on^{-1}]A = [\equiv_2][\equiv_1]A$

- $\langle R \rangle A = \neg[R]\neg A$, donde
 $[R] \in \{\equiv_1, \equiv_2, \neq_1, \neq_2, \|\|\}$

6. Semántica

La semántica que en principio consideramos, para la interpretación del anterior lenguaje será la de Kripke sobre los marcos de la forma $W_F = \{W, \equiv_1, \equiv_2, \neq_1, \neq_2, \|\|\}$, donde W es no vacío y $\equiv_1, \equiv_2, \neq_1, \neq_2$ y $\|\|$ son relaciones binarias arbitrarias sobre W .

Si V es una valoración, de forma inductiva diremos que la fórmula A es verdadera en un estado x de W si:

$$x \models_V p \text{ si y solo si } x \in V(p)$$

$$x \models_V \neg A \text{ si y solo si } x \not\models_V A$$

$$x \models_V A \wedge B \text{ si y solo si } x \models_V A \text{ y } x \models_V B$$

$x \models_V [R]A$ si y solo si $(\forall y \in W)(xRy \rightarrow y \models_V A)$,
donde

$$R \in \{\equiv_1, \equiv_2, \neq_1, \neq_2, \|\|\}$$

Recordemos que si $M = (W_F, V)$ es un modelo basado sobre el marco W_F y ϕ es una fórmula tal que $x \models_V \phi$ para todo $x \in W$, se dirá que ϕ es verdadera en M y escribiremos $M \models \phi$. Cuando ϕ sea verdadera en todo modelo basado sobre un marco W_F , se dirá que ϕ es válida en el marco W_F , lo cual se denotará $F \models \phi$, diciendo en tal caso, que F es un marco de ϕ .

Si ϕ es válida en todo marco F perteneciente a una clase de marcos \mathbf{C} , se dirá que ϕ es \mathbf{C} -válida y lo denotaremos $\models_{\mathbf{C}} \phi$. Diremos, en tal caso, que \mathbf{C} es una clase de marcos de ϕ .

El conjunto de todas las fórmulas válidas en una cierta clase de marcos \mathbf{C} se denomina, la lógica de la clase \mathbf{C} . En nuestro caso a la lógica de la clase de los marcos de Lobachevski la denotaremos por LML. Por otra parte, recordemos que si Σ y Σ' son dos clases de marcos, decimos que Σ es modalmente definible en Σ' por una fórmula A , si para cada marco $F \in \Sigma'$, A es válida en F si y solo si $F \in \Sigma$.

Definición 6. Un marco $W_F = (W, \equiv_1, \equiv_2, \neq_1, \neq_2, \|\|)$ diremos que es \neq -estándar si satisface:

$$(\forall x, y \in W)(x \neq y \leftrightarrow x \neq_1 y \vee x \neq_2 y)$$

Recordemos que no en todos los marcos las modalidades secundarias $[\neq]$ y $[U]$, tienen la interpretación semántica habitual (o estándar), pero en el caso de los marcos \neq -estándar sí.

En nuestro caso, los marcos de Lobachevski son \neq -estándar.

Definición 7. Un marco $W_F = (W, \equiv_1, \equiv_2, \neq_1, \neq_2, \|\|)$ lo llamaremos *cuasi-estándar* si es \neq -estándar y la relación \neq_i es complemento de la relación \equiv_i , $i = 1, 2$.

Lema 4. La clase de marcos *cuasi-estándar* es modalmente definible en la clase de marcos \neq -estándar mediante las siguientes fórmulas:

$$\begin{aligned} [\text{MGEI}^*] \quad & \langle U \rangle p \rightarrow \langle \equiv_i \rangle p \vee \langle \neq_i \rangle p, \quad i = 1, 2 \\ [\text{MGEI}^{**}] \quad & \langle \equiv_i \rangle [\neq] p \rightarrow [\neq] p, \quad i = 1, 2 \end{aligned}$$

Demostración. Consideremos un marco $W_F = \{W, \equiv_1, \equiv_2, \neq_1, \neq_2, \|\|\} \neq$ -estándar, entonces W_F es *cuasi-estándar* si y solo si se cumple:

1. $\equiv_i \cup \neq_i = W \times W$
2. $\equiv_i \cap \neq_i = \emptyset$

Ahora basta con comprobar, que los axiomas MGEI* y MGEI** son válidos si y solo si (1) y (2) se cumplen en W_F . \square

Lema 5. La clase de los marcos de Lobachevski es modalmente definible en la clase de marcos *cuasi-estándar* mediante las siguientes fórmulas:

$$\begin{aligned} (\text{LMGL0}) \quad & [\equiv_i] p \rightarrow p, \quad \langle \equiv_i \rangle [\equiv_i] p \rightarrow p, \\ & [\equiv_i] p \rightarrow [\equiv_i] [\equiv_i] p, \quad i = 1, 2 \\ (\text{LMGL0}^*) \quad & \langle U \rangle p \rightarrow \langle \|\| \rangle p \vee \langle on^{-1} \rangle \langle on \rangle p, \quad \langle \equiv_2 \rangle p \rightarrow \langle \|\| \rangle p, \\ & [\equiv_2] q \wedge \langle on^{-1} \rangle \langle on \rangle ([\neq] q \wedge \neg q) \rightarrow [\|\|] p \\ (\text{MGL1}) \quad & \langle U \rangle p \rightarrow \langle on \rangle \langle on^{-1} \rangle p \\ (\text{LMGL2}) \quad & \langle on \rangle (p \wedge \langle on^{-1} \rangle ([\neq] q \wedge r)) \rightarrow \\ & (\langle on \rangle (\langle \equiv_2 \rangle p \vee \langle on^{-1} \rangle q) \vee \langle \equiv_1 \rangle r) \\ (\text{LMGL}) \quad & p \rightarrow \langle on^{-1} \rangle \langle \neq_1 \rangle \langle on \rangle p \\ (\text{LMGL4}) \quad & p \rightarrow \langle on \rangle \langle \neq_2 \rangle \langle on^{-1} \rangle p \\ [\text{LMGL5}] \quad & \langle U \rangle p \wedge q \rightarrow \langle on \rangle (\langle \|\| \rangle p \wedge \langle \neq_2 \rangle (\langle \|\| \rangle p \wedge \langle on^{-1} \rangle q)) \vee \\ & \langle U \rangle (\langle on \rangle p \wedge q) \end{aligned}$$

Demostración. Las fórmulas (LMGL0) se corresponden con las propiedades reflexiva, simétrica y transitiva de \equiv_i , $i = 1, 2$.

Las de (LMGL0*), se corresponden respectivamente con las condiciones de primer orden:

1. $x \|\| y \vee (\exists z \in W)(x \text{ on}^{-1} z \wedge z \text{ on } y)$
2. $x \equiv_2 y \rightarrow x \|\| y$
3. $(\exists z \in W)(x \text{ on}^{-1} z \wedge z \text{ on } y) \wedge x \|\| y \rightarrow x \equiv_2 y$

cuya conjunción es equivalente a la definición dada de paralelismo.

El resto de expresiones modales (LMGL1), ..., (LMGL5), se corresponden con los axiomas de primer orden (ML1)...(ML5) respectivamente.

Desarrollaremos la prueba de correspondencia, para el caso del axioma que caracteriza a esta geometría, (LMGL5). Observemos para ello, que la fórmula en cuestión, resulta ser una fórmula de Sahlqvist y tenemos por tanto, asegurada la existencia de una fórmula

de primer orden que se corresponde con ella. Usando la traslación estándar, obtendremos como primer paso, la fórmula de segundo orden correspondiente a (LMGL5).

$$\begin{aligned}
& ST_x(\langle U \rangle p \wedge q \rightarrow \langle on \rangle(\langle \parallel \rangle p \wedge \langle \neq_2 \rangle(\langle \parallel \rangle p \wedge \langle on^{-1} \rangle q)) \vee \langle U \rangle(\langle on \rangle p \wedge q)) \\
& \quad \Downarrow \\
& ST_x(\langle U \rangle p \wedge q) \rightarrow ST_x(\langle on \rangle(\langle \parallel \rangle p \wedge \langle \neq_2 \rangle(\langle \parallel \rangle p \wedge \langle on^{-1} \rangle q)) \vee \langle U \rangle(\langle on \rangle p \wedge q)) \\
& \quad \Downarrow \\
& (\forall P, Q)((\exists z)(Pz \wedge Qx) \rightarrow (\exists l)(x \text{ on } l) \wedge (\exists z_2)(l \parallel z_2 \wedge Pz_2) \wedge \\
& \quad (\exists t)(l \neq_2 t) \wedge (\exists z_4)(t \parallel z_4 \wedge Pz_4) \wedge (\exists x_1)(t \text{ on}^{-1} x_1 \wedge Qx_1) \vee \\
& \quad (\exists x_2, z_5)(x_2 \text{ on } z_5 \wedge Pz_5 \wedge Qx_2)) \\
& \quad \Downarrow \\
& (\forall P, Q)(\forall z)(Pz \wedge Qx \rightarrow (\exists l)(x \text{ on } l) \wedge (\exists z_2)(l \parallel z_2 \wedge Pz_2) \wedge \\
& \quad (\exists t)(l \neq_2 t) \wedge (\exists z_4)(t \parallel z_4 \wedge Pz_4) \wedge (\exists x_1)(t \text{ on}^{-1} x_1 \wedge Qx_1) \vee \\
& \quad (\exists x_2, z_5)(x_2 \text{ on } z_5 \wedge Pz_5 \wedge Qx_2))
\end{aligned}$$

Elijamos las mínimas asignaciones de P y Q que hacen el antecedente verdadero, $\sigma(P) \equiv \lambda u.(z = u)$, $\sigma(Q) \equiv \lambda u.(x = u)$; en consecuencia obtenemos

$$\begin{aligned}
& (\forall z)(z = z \wedge x = z \rightarrow (\exists l)(x \text{ on } l) \wedge (\exists z_2)(l \parallel z_2) \wedge (z = z_2) \wedge \\
& (\exists t)(l \neq_2 t) \wedge (\exists z_4)(t \parallel z_4) \wedge (z = z_4) \wedge (\exists x_1)(t \text{ on}^{-1} x_1) \wedge (x = x_1) \vee \\
& (\exists x_2, z_5)(x_2 \text{ on } z_5) \wedge (z = z_5) \wedge (x = x_2)) \\
& \quad \Downarrow \\
& (\forall z)(\exists l)((x \text{ on } l) \wedge (l \parallel z) \wedge (\exists t)(l \neq_2 t) \wedge (t \parallel z) \wedge (t \text{ on}^{-1} x) \vee (x \text{ on } z)) \\
& \quad \Downarrow \\
& (\forall z)(\exists l, t)((x \text{ on } l) \wedge (l \parallel z) \wedge (l \neq_2 t) \wedge (t \parallel z) \wedge (x \text{ on } t) \vee (x \text{ on } z))
\end{aligned}$$

De esta forma se llega a la correspondencia local de nuestra fórmula y basta generalizar para obtener el resultado:

$$(\forall x, z)(\exists l, t)((x \text{ on } l) \wedge (l \parallel z) \wedge (l \neq_2 t) \wedge (t \parallel z) \wedge (x \text{ on } t) \vee (x \text{ on } z))$$

7. Axiomática

Proponemos la siguiente axiomatización para LML:

Axiomas

[I] Todas las tautologías de la lógica de proposiciones

[II] $[R](p \rightarrow q) \rightarrow [R]p \rightarrow [R]q$, $R \in \{\equiv_1, \equiv_2, \neq_1, \neq_2, \parallel\}$, $\langle \neq_i \rangle[\neq_i]p \rightarrow p$, $i = 1, 2$, $\langle \parallel \rangle[\parallel]p \rightarrow p$.

$[\neq_1][\neq]p \rightarrow p$

$[\neq_2]p \wedge [\neq]p \rightarrow [\neq][\neq]p$

$[(R \subseteq U)][U]p \rightarrow [R]p$, $R \in \{\equiv_1, \equiv_2\}$

[LMGL*] $\langle U \rangle p \rightarrow \langle \equiv_i \rangle p \vee \langle \neq_i \rangle p$, $i = 1, 2$

[LMGL**] $\langle \equiv_i \rangle[\neq]p \rightarrow [\neq_i]p$, $i = 1, 2$

- [LMGL0] $[\equiv_i]p \rightarrow p, \langle \equiv_i \rangle [\equiv_i]p \rightarrow p, [\equiv_i]p \rightarrow [\equiv_i][\equiv_i]p, i = 1, 2$
 (LMGL0*) $\langle U \rangle p \rightarrow \langle \parallel \rangle p \vee \langle on^{-1} \rangle \langle on \rangle p, \langle \equiv_2 \rangle p \rightarrow \langle \parallel \rangle p, [\equiv_2]q \wedge \langle on^{-1} \rangle \langle on \rangle ([\neq]p \wedge \neg q) \rightarrow [\parallel]p$
 (LMGL1) $\langle U \rangle p \rightarrow \langle on \rangle \langle on^{-1} \rangle p$
 (LMGL2) $\langle on \rangle (p \wedge \langle on^{-1} \rangle ([\neq]q \wedge r)) \rightarrow ([on](\langle \equiv_2 \rangle p \vee [on^{-1}]q) \vee \langle \equiv_1 \rangle r)$
 (LMGL3) $p \rightarrow \langle on^{-1} \rangle \langle \neq_1 \rangle \langle on \rangle p$
 (LMGL4) $p \rightarrow \langle on \rangle \langle \neq_2 \rangle \langle on^{-1} \rangle p$
 (LMGL5) $\langle U \rangle p \wedge q \rightarrow \langle on \rangle (\langle \parallel \rangle p \wedge \langle \neq_2 \rangle (\langle \parallel \rangle p \wedge \langle on^{-1} \rangle q)) \vee \langle U \rangle (\langle on \rangle p \wedge q)$

Reglas de inferencia:

- (MP) modus ponens: si $\vdash A$ y $\vdash A \rightarrow B$ entonces $\vdash B$,
 (RN) regla de necesidad: si $\vdash A$ entonces $\vdash [R]A$, $R \in \{\equiv_1, \equiv_2, \neq_1, \neq_2, \parallel\}$,
 (IR) regla de irreflexividad: si p no aparece en A y $\vdash Op \rightarrow A$ entonces $\vdash A$,
 (RS) regla de sustitución de letras proposicionales.

8. Corrección y Completitud

El sistema axiomático LML será correcto respecto a la clase de los marcos de Lobachevski, si por una parte todos sus axiomas son válidos en dicha clase y por otro lado, sus reglas de inferencia son también admisibles.

La primera cuestión la hemos asegurado en la propia construcción. En cuanto a la segunda, solo restará hacer la comprobación en el caso de la regla de irreflexividad, ya que las restantes son las usuales en cualquier lógica normal y son válidas en cualquier clase de marcos.

Así pues, supongamos que $Op \rightarrow A$ es válida en nuestra clase de marcos y que por contra A no lo sea. Existirá por tanto una valoración V y un punto w en alguno de nuestros marcos, tal que $w \not\models_V A$. Consideremos una nueva valoración V' definida como $V'(p) = V(p)$ si $p_i \neq p$ y $V'(p) = \{w\}$. De esta forma $w \models_{V'} Op$ y como $Op \rightarrow A$ es una fórmula válida tendremos que $w \models_{V'} A$, pero al no intervenir el literal p en A , concluimos que $w \models_V A$, lo que nos lleva a una contradicción.

Por otra parte, para asegurar la completitud del sistema, haremos uso de un resultado general denominado teoremas SD, y debido a Yde Venema (véase [20]). En el trabajo reseñado, el autor da una prueba general de completitud para lógicas multimodales conteniendo el operador \neq , la regla de irreflexividad (IR) para $[\neq]$ y satisfaciendo ciertas condiciones:

- Por cada operador $[\$]$, el lenguaje debe contener el operador inverso $[\$^{-1}]$, relacionado con $[\$]$ mediante los siguientes axiomas. $\langle \$ \rangle [\$^{-1}]A \rightarrow A$ y

$\langle \$^{-1} \rangle [\$] \rightarrow A$. Obviamente la lógica LML satisface esta condición; para cada operador $\$$, el inverso es él mismo.

- La lógica (normal) debe contener los mínimos axiomas que hagan a los operadores tensionados junto con los axiomas (\neq_1) , (\neq_2) y $(R \subseteq U)$. El resto de axiomas adicionales deben ser fórmulas de Sahlqvist.

En estas condiciones el teorema SD asegura la completitud de la lógica, en la clase de marcos modalmente definible por los axiomas adicionales. En nuestro caso LML se encuentra en las condiciones del teorema, por lo que es completa respecto a los marcos de Lobachevski.

REFERENCIAS

- [1] **Balbani P.** *The modal multilogic of geometry*, J. of Applied Nonclassical Logics, **8**(1998),259–281.
- [2] **Balbani P., V. Dugat, L. Fariñas del Cerro & A. López.** *Éléments de Géométrie Mécanique*. Hermès, 1994.
- [3] **Ph. Balbani, L. Fariñas del Cerro, T. Tinchev & D. Vakarelov.** *Modal Logic for incidence geometries*. Journal of Logic and Computation, **7** (1997), 59–78.
- [4] **Blumental L.** *A modern view of Geometry*. Dover.
- [5] **Coxeter H.** *Fundamentos de Geometría*. Limusa.
- [6] **Gabbay D.** *An irreflexivity lemma with applications to axiomatizations of conditions on tense frames*. In *Aspects of Philosophical Logic*, U. Mönnich, ed. pp. 67–89. Reidel, 1981.
- [7] **Gabbay D. & I. Hodkinson.** *An axiomatization of the temporal logic with until and since over the real numbers*. Journal of Logic and Computation, **1** (1990), 229–259.

- [8] **Gadbay D., I. Hodkinson & M. Reynolds.** *Temporal Logic: Mathematical Foundations and Computational Aspects*, Vol. I. Oxford University Press, 1995
- [9] **Hilbert D.** *Les Fondements de la Géométrie*. Dunod, Paris, 1971.
- [10] **Hughes G. & M. Cresswell.** *A companion to Modal Logic*. Methuen, London, 1984.
- [11] **Jacobson N.** *Basic Algebra II*. W. H. Freeman and Company, New York, 1980.
- [12] **Marx M.** *Complexity of modal logics of relations*. Technical Report ML-97-02, Institute for logic, Language and Computation, University of Amsterdam, 1997.
- [13] **Marx M. & Y. Venema.** *Multidimensional Modal Logic*. Kluwer Academic Publishers, Dordrecht, 1997.
- [14] **Rijke M.** *The modal logic of inequality*. *Journal of Symbolic Logic*, **57** (1992), 566–584.
- [15] **Sahlqvist H.** *Completeness and correspondence in the first and second order semantics for modal logic*. In *Proceeding of the Third Scandinavian Logic Symposium*, Uppsala, Sweden, 1973, s. Kanger, ed. pp: 110–118. North-Holland. Amsterdam, 1975.
- [16] **Patrick Blackburn, Mararten de Rijke & Y de Venema.** *Modal Logic*. Cambridge University Press, United Kingdom, 2001.
- [17] **Vakarelov D.** *A modal theory of arrow*. In *Proceedings of the logic in AI. European Workshop JELIA' 92*. Berlin, Germany, September 1992. D. Pearce and G. Wagner. eds. pp. 1-24. Vol. 633 of Lectures Notes in Artificial Intelligence. Springer-Verlag, Berlin 1992.
- [18] **Venema Y.** *Modal derivation rules*. ITLI prepublication series for mathematical logic and foundations ML-01-07, University of Amsterdam. Amsterdam, 1991.
- [19] **Venema Y.** *Many dimensional Modal Logic*. Doctoral dissertation. University of Amsterdam, 1992.
- [20] **Venema Y.** *Derivation rules as anti-axioms in modal logic*. *Journal of symbolic Logic*, **58** (1993), 1003–1034.
- [21] **Wright G.** *A modal logic of space*. In *The Philosophy of Nicholas Rescher*, F. Sosa, ed. Reidel, Dordrecht, 1979.

p -ADIC OSCILLATORY INTEGRALS AND NEWTON POLYHEDRA

by

W. A. Zuniga-Galindo ¹

Dedicated to the Memory of Professor Jairo Charris

Resumen

Zuniga-Galindo, W. A. : p -adic oscillatory integrals and Newton Polyhedra. Rev. Acad. Colomb. Cienc. **28** (106): 95–99, 2004. ISSN 0370-3908.

En este artículo damos una estimativa asintótica para integrales oscilantes p -ádicas que dependen de dos parámetros. Estas integrales son soluciones de ecuacionesseudodiferenciales p -ádicas del tipo Schrödinger.

Palabras clave: Cuerpos p -ádicos, integrales oscilatorias, funciones zeta locales de Igusa, poliedros de Newton, ecuacionesseudodiferenciales.

Abstract

In this paper we give an asymptotic estimate for p -adic oscillatory integrals depending of two parameters. These integrals are solutions of Schrödinger-type pseudo-differential equations.

Key words: p -adic fields, oscillatory integrals, Igusa local zeta function, Newton polyhedra, pseudo-differential equations.

1. Introduction

Let K be p -adic field, i.e. a finite extension of \mathbb{Q}_p , R_K the ring of integers of K , P_K the maximal ideal of

R_K , and $\overline{K} = R_K/P_K$ the residue field of K . The cardinality of the residue field of K is denoted by q , thus $\overline{K} = \mathbb{F}_q$. For $z \in K$, $v(z) \in \mathbb{Z} \cup \{+\infty\}$ denotes the

¹Department of Mathematics and Computer Science, Barry University, 11300 N. E. Second Avenue, Miami Shores, Florida 33161, USA.
email: wzuniga@mail.barry.edu
AMS Subject Classification: Primary 46S10, 11S40.

valuation of z , $|z|_K = q^{-v(z)}$, and $ac\ z = z\pi^{-v(z)}$ where π is a fixed uniformizing parameter for R_K .

Let Ψ denote a standard additive character of K , thus, for $z \in K$, $\Psi(z) = \exp(2\pi i \text{Tr}_{K/\mathbb{Q}_p}(z))$, where Tr denotes the trace.

Let $\phi(\xi) \in R_K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$, be a non-constant polynomial, $t \in K$, with $v(t) < 0$, $x = (x_1, \dots, x_n) \in K^n$, with $v(x_i) < 0$, $i = 1, \dots, n$, and

$$m = \min \{v(x_1), \dots, v(x_n), v(t)\}.$$

We put $\langle x, \xi \rangle = \sum_i x_i \xi_i$, for $x, \xi \in K^n$. To these data we associate the following parametric exponential sum

$$J(x, t, \phi, K) = J(x, t) = q^{-mn} \sum_{\xi \bmod \pi^m} \Psi(t\phi(\xi) + \langle x, \xi \rangle). \quad (1.1)$$

The exponential sum $J(x, t)$ can be expressed as an integral of the form

$$J(x, t) = \int_{R_K^n} \Psi(t\phi(\xi) + \langle x, \xi \rangle) |d\xi|, \quad (1.2)$$

where $|d\xi|$ is the Haar measure of K^n normalized so that the volume of R_K^n is 1. A more general type of oscillatory integrals is

$$I_h(x, t, \phi, K) = I(x, t) = \int_{K^n} \Psi(t\phi(\xi) + \langle x, \xi \rangle) h(\xi) |d\xi|, \quad (1.3)$$

where h is a Bruhat-Schwartz function, i.e. a locally constant function with compact support. Integrals of the form $\int_K \Psi(a\xi + b\xi) h(\xi) |d\xi|$ are called Gaussian ones. These integrals have explicitly calculated in several cases, and they appear in certain p -adic quantum models [12, Chap. 1, Sect. V, and Chap. 3].

The integrals $I(x, t)$ are the non-archimedean counterpart of

$$u(x, t) = \int_{\mathbb{R}^n} \exp(2\pi i(t\phi(\xi) + \langle x, \xi \rangle)) \widehat{f}(\xi) |d\xi|, \quad (1.4)$$

where \widehat{f} is the Fourier transform of f . Consider the Schrödinger-type equation

$$\frac{\partial u}{\partial t} = i\phi(D)u, \quad u(x, 0) = f(x), \quad (1.5)$$

here $\phi(D)$ is a pseudo-differential operator having symbol $\phi(\xi)$. Then function $u(x, t)$ is a solution for the initial value problem (1.5) (see e.g. [10, Chap. VII, VIII]).

As a consequence of the previous considerations it is natural to ask if integrals (1.2) and (1.3) satisfy some differential equation. At this point, it is important to mention that there are deep connections between differential equations and exponential sums over finite fields [5], [7].

Integrals (1.2) and (1.3) satisfy pseudo-differential equations of Schrödinger-type. Let $\mathcal{S}(K^n)$ denote the \mathbb{C} -vector space of Schwartz-Bruhat functions over K^n . The dual space $\mathcal{S}'(K^n)$ is the space of distributions over K^n . A pseudo-differential operator of Schrödinger-type $A(\partial)$, with symbol $|\tau - \phi(\xi)|_K$, is an operator of the form

$$A(\partial) : \mathcal{S}(K^n) \rightarrow \mathcal{S}(K^n)$$

$$\Phi \rightarrow \mathcal{F}_{(\xi, \tau) \rightarrow (x, t)}^{-1} (|\tau - \phi(\xi)|_K \mathcal{F}_{(y, \varrho) \rightarrow (\tau, \xi)}(\Phi)), \quad (1.6)$$

where

$$\mathcal{F} : \mathcal{S}(K^n) \rightarrow \mathcal{S}(K^n)$$

$$\Phi \rightarrow \int_{K^n} \Psi(-\langle x, y \rangle) \Phi(x) dx \quad (1.7)$$

is the Fourier transform. The operator $A(\partial)$ has self-adjoint extension with dense domain in $L^2(K^n)$.

The initial value problem

$$A(\partial)z = 0, \quad z(x, 0) = h(x) \in \mathcal{S}(K^n) \quad (1.8)$$

is the non-archimedean counterpart of (1.5). By passing to the Fourier transform in (1.8), we get

$$|\tau - \phi(\xi)|_K \mathcal{F}_{(y, \varrho) \rightarrow (\tau, \xi)}(z) = 0, \quad (1.9)$$

from where it follows that any distribution of the form

$$z(x, t) = (\mathcal{F}^{-1}g)(x, t),$$

with $g(\xi, \tau)$ a distribution with support in

$$\{(\xi, \tau) \in K^{n+1} \mid |\tau - \phi(\xi)|_K = 0\},$$

is a solution of $A(\partial)z = 0$. In the case when

$$g(\xi, \tau) = \widehat{h}(\xi) \delta(\tau - \phi(\xi)),$$

where δ is the Dirac distribution, and \widehat{h} is the Fourier transform of h , the distribution $z(x, t)$ takes the form (1.3). Finally, since $I(x, 0) = h(x)$, it holds that $I(x, t)$ is a solution for the initial value problem (1.8). In particular the exponential sums $J(x, t)$ satisfy (1.8), when $h(x, 0)$ is equal to the characteristic function of R_K^n .

The theory of non-archimedean pseudo-differential operators is emerging motivated for its potential use in p -adic physics [12], [8].

The main result of this paper (cf. Theorem 3.1) gives an asymptotic estimation of $|I(x, t)|$ for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, in the case in which the singular locus of $t\phi(\xi) + \langle x, \xi \rangle$ is a subset of $\{(t, x, \xi) \mid x = \xi = 0\}$, and ϕ is a generic polynomial with an algebraically isolated singularity at the origin. The proof of the main result depends on the description of the poles of the Igusa zeta function for non-degenerate polynomials [3], [4], [13], [14], and a theorem of Igusa that establishes a connection between the poles of local zeta functions and the asymptotic expansions of certain p -adic oscillatory integrals [6, Theorem 8.4.2 (3)].

2. Exponential sums and Newton polyhedra

We set $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$. Let $f(\xi) = \sum_l a_l \xi^l \in K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$ be a polynomial in n variables satisfying $f(0) = 0$. The Newton polyhedron $\Gamma(f)$ of f is defined as the convex hull in \mathbb{R}_+^n of the set

$$\bigcup_{m \in \{l \in \mathbb{N}^n \mid a_l \neq 0\}} (m + \mathbb{R}_+^n).$$

We denote by $\langle \cdot, \cdot \rangle$ the usual inner product of \mathbb{R}^n , and identify \mathbb{R}^n with its dual by means of it. We set

$$\langle a_\gamma, x \rangle = m(a_\gamma),$$

for the equation of the supporting hyperplane of a facet γ (i.e. a face of codimension 1 of $\Gamma(f)$) with perpendicular vector $a_\gamma = (a_1, \dots, a_n) \in \mathbb{N}^n \setminus \{0\}$, and $\sigma(a_\gamma) = \sum_i a_i$.

Definition 2.1. A polynomial $f(\xi) = \sum_i a_i \xi^i \in K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$, is called *globally non-degenerate with respect to its Newton polyhedron* $\Gamma(f)$, if it satisfies the following two properties:

- (1) the origin of K^n is a singular point of $f(\xi)$;
- (2) for every face $\gamma \subset \Gamma(f)$ (including $\Gamma(f)$ itself), the polynomial

$$f_\gamma(\xi) = \sum_{i \in \gamma} a_i \xi^i$$

has the property that there is no $\xi \in (K \setminus \{0\})^n$ such that

$$f_\gamma(\xi) = \frac{\partial f_\gamma}{\partial \xi_1}(\xi) = \dots = \frac{\partial f_\gamma}{\partial \xi_n}(\xi) = 0.$$

For a polynomial $f(\xi) \in K[\xi]$ globally non-degenerate with respect to its Newton polyhedron $\Gamma(f)$, we

set

$$\beta(f) = \max_{\tau_j} \left\{ -\frac{\sigma(a_j)}{m(a_j)} \right\},$$

where τ_j runs through all facets of $\Gamma(f)$ satisfying $m(a_j) \neq 0$. We note that

$$T_0 = (-\beta(f)^{-1}, \dots, -\beta(f)^{-1}) \in \mathbb{Q}^n$$

is the intersection point of the boundary of the Newton polyhedron $\Gamma(f)$ with the diagonal $\{(t, \dots, t) \mid t \in \mathbb{R}\} \subset \mathbb{R}^n$.

We put

$$E(z, f, K) = E(z, f) = \int_{R_K^n} \Psi(zf(\xi)) |d\xi|,$$

with $z \in K$. Igusa showed that the asymptotic behavior of $E(z, f, K)$, when $|z|_K \rightarrow \infty$, is controlled by the largest pole of the meromorphic continuation of the local zeta function

$$Z(s, f, \chi) = \int_{R_K^n} \chi(ac f(\xi)) |f(\xi)|_K^s d\xi, \quad \operatorname{Re}(s) > 0, \quad (2.1)$$

associated to f , and a multiplicative character χ of R_K^\times . More precisely, if γ_f the maximum of the real parts of the poles of $Z(s, f, \chi)$, and $\gamma > -1$, then

$$|E(z, f)| \leq C(K) |z|_K^{\gamma_f + \epsilon}, \quad \text{for } |t|_K \rightarrow \infty, \quad (2.2)$$

where $C(K)$ is a constant, and $\epsilon > 0$ (see e.g. [2, Corollary 1.4.5], or [6, Theorem 8.4.2 (3)]).

Theorem 2.2. Let K be a non-archimedean local field, and let $f(\xi) \in R_K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$, be a globally non-degenerate polynomial with respect to its Newton polyhedron $\Gamma(f)$. If $\beta(f) > -1$, then

$$|E(z, f)| \leq C(f, K) |z|_K^{\beta(f) + \epsilon}, \quad (2.3)$$

for $|z|_K \rightarrow \infty$, and any $\epsilon > 0$, here $C(f, K)$ is constant depending on f and K .

The theorem follows from (2.2) by showing the following two facts. First, the poles of $Z(s, f, \chi)$ have the form

$$s = -\frac{\sigma(a_\gamma)}{m(a_\gamma)} + \frac{2\pi i}{\log q} \frac{k}{m(a_\gamma)}, \quad k \in \mathbb{Z},$$

for some facet γ of $\Gamma(f)$ with perpendicular a_γ , and $m(a_\gamma) \neq 0$, or

$$s = -1 + \frac{2\pi i}{\log q} k, \quad k \in \mathbb{Z}.$$

Second, the maximum of the real parts of the poles of $Z(s, f, \chi_{\text{triv}})$ is $\beta(f)$, when $\beta(f) > -1$ (cf. [14,

Theorems A, B]). The description of the largest pole of $Z(s, f, \chi_{\text{triv}})$ when f is non-degenerate with respect to its Newton polyhedron $\Gamma(f)$ and $\beta(f) > -1$ follows from observations made by Varchenko in [11] and was originally noted in the p -adic case in [9] (although it is misstated there as $\alpha(f) \neq -1$). The case $\beta(f) = -1$ is treated in [4]. The case of $\beta(f) < -1$ is more difficult and is established in [4] with some additional conditions on τ_0 by using a difficult result on exponential sums. The author proved the case $\beta(f) \geq -1$ for polynomials with coefficients in a non-archimedean local field of arbitrary characteristic [14].

We put $\|x\|_K = \max_i \{|x_i|_K\}$, for $x \in K^n$.

Proposition 2.3. *Let $f(\xi) \in R_K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$, be a non-constant polynomial without singularities on K^n . Then*

$$|E(z, f)| \leq C(f, K) |z|_K^{-1+\epsilon}$$

for $|z|_K \rightarrow \infty$, and any $\epsilon > 0$, here $C(f, K)$ is a positive constant depending only on f and K .

Proof. The stationary phase formula implies that $E(z, f) = 0$ for $|z|_K \rightarrow \infty$. The proof is a slightly variation of the proof of Lemma (2.4) in [14]. \square

Proposition 2.4. *Let $\phi(\xi) \in R_K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$, be a globally non-degenerate polynomial with respect $\Gamma(\phi)$, with $\beta(\phi) > -1$. If the singular locus of the polynomial $t\phi(\xi) + \langle x, \xi \rangle$ is contained in $\{(t, x, \xi) \mid x = 0\}$, then*

$$|J(x, t)| \leq C(\phi, K) [\min(\|x\|_K, |t|_K)]^{\beta(\phi)+\epsilon},$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and $\epsilon > 0$, here $C(\phi, K)$ a positive constant depending only on ϕ and K .

Proof. For $x \neq 0$, and $t \neq 0$, $t\phi(\xi) + \langle x, \xi \rangle$ does not have singular points on K^n . If $\frac{\|x\|_K}{|t|_K} \leq 1$, Proposition 2.3 implies that

$$|J(x, t)| \leq C_0(K) |t|_K^{-1+\epsilon},$$

for $|t|_K \rightarrow \infty$, and $\epsilon > 0$. If $\frac{|t|_K}{\|x\|_K} < 1$, Proposition 2.3 implies that

$$|J(x, t)| \leq C_1(K) \|x\|_K^{-1+\epsilon},$$

for $\|x\|_K \rightarrow \infty$, and $\epsilon > 0$. Therefore if $x \neq 0$, and $t \neq 0$,

$$|J(x, t)| \leq C_3(K) \min(\|x\|_K, |t|_K)^{-1+\epsilon}, \quad (2.4)$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and $\epsilon > 0$.

For $x = 0$, and $t \neq 0$, Theorem 2.2 implies that

$$|J(x, t)| \leq C_4(\phi, K) |t|_K^{\beta(\phi)+\epsilon}, \quad (2.5)$$

for $|t|_K \rightarrow \infty$, and $\epsilon > 0$. For $x \neq 0$, and $t = 0$,

$$J(x, t) = 0, \quad (2.6)$$

for $|t|_K \geq 1$. Since $\beta(f) > -1$, estimates (2.4), (2.5), and (2.6) imply that

$$|J(x, t)| \leq C(\phi, K) [\min(\|x\|_K, |t|_K)]^{\beta(\phi)+\epsilon},$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and $\epsilon > 0$. \square

The proof of the following Proposition is similar to the previous one.

Proposition 2.5. *Let $\phi(\xi) \in R_K[\xi]$, $\xi = (\xi_1, \dots, \xi_n)$, be polynomial such that $t\phi(\xi) + \langle x, \xi \rangle$ has no singular points on K^n , for any $t \in K$, and $x \in K^n$, then*

$$|J(x, t)| \leq C(K) [\min(\|x\|_K, |t|_K)]^{-1+\epsilon},$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and $\epsilon > 0$, here $C(K)$ a positive constant.

3. Main Result

We shall say that the origin of K^n is an *algebraically isolated singularity* of $\phi(\xi) \in K[\xi_1, \dots, \xi_n]$, if the origin is the only solution of the system

$$\phi(\xi) = \frac{\partial \phi}{\partial \xi_1}(\xi) = \dots = \frac{\partial \phi}{\partial \xi_n}(\xi) = 0. \quad (3.1)$$

Theorem 3.1. *Let $\phi(\xi) \in R_K[\xi_1, \dots, \xi_n]$ be a non-constant polynomial with an algebraically isolated singularity at the origin, such that $\phi(\xi)$ is globally non-generate with respect $\Gamma(\phi)$, and $\beta(\phi) > -1$. If the singular locus of the polynomial $t\phi(\xi) + \langle x, \xi \rangle$ is contained in $\{(t, x, \xi) \mid x = 0\}$, then*

$$|I(x, t)| \leq C(\phi, K) [\min(\|x\|_K, |t|_K)]^{\beta(\phi)+\epsilon}, \quad (3.2)$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and any $\epsilon > 0$, here $C(\phi, K)$ a positive constant depending only on ϕ and K .

Proof. Let $\cup_i (z_i + \pi^{e_0} R_K^n)$ be a finite covering of the support of h such that $h|_{z_i + \pi^{e_0} R_K^n} = h(z_i)$. Then

$$I(x, t) = \sum_i c_i \int_{R_K^n} \Psi(t\phi(z_i + \pi^{e_0}\xi) + \langle \pi^{e_0}\xi, x \rangle) |d\xi|, \quad (3.3)$$

where $c_i = q^{-e_0 n} h(z_i) \Psi(\langle z_i, x \rangle)$. We put $\phi(z_i + \pi^{e_0}\xi) = \phi(z_i) + \pi^{e_0} \phi_i^*(\xi)$, with $\phi_i^*(\xi) \in K[\xi]$. Without loss of generality we may assume that $\pi^{e_0} \phi_i^*(\xi) \in R_K[\xi]$. With this notation, (3.3) can be rewritten as

$$I(x, t) = \sum_i c_i \Psi(t\phi(z_i)) J(x, t, \pi^{e_0} \phi_i^*(\xi)). \quad (3.4)$$

If z_i is the origin then

$$|J(x, t, \pi^{e_0} \phi_i^*(\xi))| \leq C_1(\phi, K) [\min(\|x\|_K, |t|_K)]^{\beta(\phi)+\epsilon}, \quad (3.5)$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and any $\epsilon > 0$, with $C_1(\phi, K)$ a positive constant depending only on h and K (cf. Proposition 2.4).

If z_i is not the origin, then $\pi^{e_0} \phi_i^*(\xi)$ does not have singularities on K^n . Then Proposition 2.5 implies that

$$|J(x, t, \pi^{e_0} \phi_i^*(\xi))|_{\mathbb{C}} \leq C_2(K) [\min(\|x\|_K, |t|_K)]^{-1+\epsilon}, \quad (3.6)$$

for $\min(\|x\|_K, |t|_K) \rightarrow \infty$, and any $\epsilon > 0$, with $C_2(K)$ a positive constant. The result follows from (3.5) and (3.6) by using the fact that $\beta(\phi) > -1$.

3.1. Remarks

1. The main result is valid for non-archimedean local fields of positive characteristic (cf. [14, Theorems A, B, and Corollary 6.1]).
2. Recently R. Cluckers showed that

$$\left| \int_{R_K^n} \Psi(\langle y, f(\xi) \rangle) |d\xi| \right|_C \leq C \|y\|_K^\alpha, \quad \alpha < 0,$$

for $\|y\|_K \rightarrow \infty$, when $f(\xi) = (f_1(\xi), \dots, f_n(\xi))$ is a dominant polynomial map. This result does not provide any information about α [1, Chap. VI].

References

- [1] **Cluckers R.**, *Cell decomposition and p -adic integration*, Ph.D. Thesis 2002, available at <http://front.math.ucdavis.edu/math.LO/0301023>.
- [2] **Denef J.**, *Report on Igusa's local zeta function*, Seminaire Bourbaki 1990/1991 (730–744) in Asterisque 201-203 (1991), 359–386.
- [3] **Denef J.**, *Poles of p -adic complex powers and Newton polyhedra*, Nieuw archief voor wiskunde, **13** (1995), 289–295.
- [4] **Denef J., Hoornaert Kathleen**, *Newton polyhedra and Igusa local zeta function*, To appear in Journal of Number Theory.
- [5] **Dwork B.**, *Bessel functions as p -adic functions of argument*, Duke Math. J. **41** (1974), 711–738.
- [6] **Igusa Jun-Ichi**, *An introduction to the theory of local zeta functions*, AMS/IP Studies in Advanced Mathematics, v. 14, 2000.
- [7] **Katz Nicholas M.**, *Exponential sums and differential equations*, Princeton University Press, 1990.
- [8] **Kochubei A. N.**, *Pseudodifferential equations and stochastics over non-archimedean fields*, Marcel Dekker, 2001.
- [9] **Lichtin B. & Meuser D.**, *Poles of a local zeta function and Newton polygons*, Compos. Math. **55** (1985), 313–332.
- [10] **Stein Elias M.**, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, 1993.
- [11] **Varchenko A.**, *Newton polyhedra and estimation of oscillating integrals*, Funct. Anal. Appl. **10** (1976), 175–196.
- [12] **Vladimirov V. S., Volovich I. V., & Zelenov E. I.**, *p -adic Analysis and Mathematical Physics*, World Scientific, Singapore, 1994.
- [13] **Zuniga-Galindo W. A.**, *Igusa's local zeta functions of semiquasihomogeneous polynomials*, Trans. Amer. Math. Soc. **353**, (2001), 3193–3207.
- [14] **Zuniga-Galindo W. A.**, *Local zeta functions and Newton polyhedra*, Nagoya Math. J., **172** (2003), 31–58.

OVERVIEW ON MODELS IN HOMOTOPICAL ALGEBRA

by

Roberto Ruiz S.¹

To the memory of Professor Jairo Charris Castañeda

Resumen

Ruiz S., Roberto : Overview on models in homotopical algebra. Rev. Acad. Colomb. Cienc. **28** (106): 100–121, 2004. ISSN 0370-3908.

Un functor covariante $\Delta \rightarrow \mathcal{A}$ se dice un *objeto modelo de \mathcal{A}* . Los objetos modelo producen en \mathcal{A} un tema de estudio muy parecido a la topología algebraica cuando \mathcal{A} es la categoría de los espacios topológicos. En este trabajo se describen los escenarios en los cuales se desarrollan estos conceptos y las principales resultados desarrollados por el autor sobre objetos modelos.

Palabras clave: Categoría modelo, objetos simpliciales, cosimpliciales, homotopía, levantamiento.

Abstract

A covariant functor $\Delta \rightarrow \mathcal{A}$ is called a *model object of \mathcal{A}* . Model objects produce in \mathcal{A} a subject matter very much as algebraic topology when \mathcal{A} is the category of topological spaces. Here we describe the settings on which such concepts are developed and describe the main features developed by the author about model objects.

Key words: Model category, Simplicial objects, Cosimplicial, Homotopy, Lifting.

¹Departamento de Matemáticas, Universidad del Valle, Cali, Colombia. Email: robruizs@yahoo.com
AMS Classification 2000: 18G55

1. Model, closed model, and pre-model categories

By a model category (Daniel Quillen [QD67]) we mean a category \mathcal{A} together with three distinguished classes of morphisms F (fibrations), C (cofibrations) and WE (weak equivalences) such that:

- M.0. *Axiom of admissibility:* \mathcal{A} is closed under finite projective and inductive limits.
- M.1. *Lifting axiom:* Given a solid arrow diagram in \mathcal{A} , where i is a cofibrations and p is a fibration and where either i or p is a weak equivalence, then the dotted arrow exists.

$$\begin{array}{ccc} X & \longrightarrow & Y \\ i \downarrow & \nearrow & \downarrow p \\ Z & \longrightarrow & K \end{array}$$

- M.2. *Factorization axiom:* Any morphism f in \mathcal{A} can be factored $f = pi$ where i is a cofibration and weak equivalence and p is a fibration. Also $f = pi$ where i is a cofibration and p is a fibration and a weak equivalence.
- M.3. F is closed under composition, base change and any isomorphism is a fibration. C is closed under composition, cobase change and any isomorphism is a cofibration.
- M.4. The base extension of a morphism which is a fibration and weak equivalence is a weak equivalence. The cobase extension of a map which a cofibration and weak equivalence is a weak equivalence.
- M.5. *Triangular axiom:* If in a commutative diagram

$$\begin{array}{ccc} X & \longrightarrow & Y \\ & \searrow & \nearrow \\ & Z & \end{array}$$

two of the morphisms are weak equivalences so is the third.

If (\mathcal{A}, F, C, WE) is a model category, then there exists an associated category, called the *homotopy category* of (\mathcal{A}, F, C, WE) denoted simply by $H_\circ\mathcal{A}$ which is the localization of \mathcal{A} with respect to WE . $H_\circ\mathcal{A}$ is characterized by the existence of a functor $r : \mathcal{A} \rightarrow H_\circ\mathcal{A}$ and the following universal property of $(r, H_\circ\mathcal{A})$: for every $f \in WE$, $r(f)$ is an isomorphism. If there exists another pair (F, \mathcal{B}) , $F : \mathcal{A} \rightarrow \mathcal{B}$ such that for every $f \in WE$, $F(f)$ is an isomorphism then, there exists a unique functor $\rho : H_\circ\mathcal{A} \rightarrow \mathcal{B}$ such that $\rho r = F$.

Note that in general, if $r(f)$ is an isomorphism, f need not be a weak equivalence. There is however, a special kind of model category in which $r(f)$ is an isomorphism if, and only if f is a weak equivalence. They are called *closed model categories*. Before we define them, we give some notation: In a model category a morphism which is both a fibration and a weak equivalence is called a trivial fibration. TF denotes the class of such morphisms. A morphism which is both a cofibration and a weak equivalence is called a trivial cofibration, TC denotes the class of such maps. We call the classes F , C , TC , TF and WE the *classes of structural maps* of the model category.

From our point of view, the main feature of closed model categories is that, with the exception of WE , the classes of structural maps are characterized by lifting properties: PM2 in next definition.

On the other hand one can also have “almost” model categories which fail to be model because WE fails to behave well. They are “pre model categories”.

By a *pre-model category*, (Roberto Ruiz, [RR76]) we mean a category \mathcal{A} together with five classes of maps F , C , TF , TC and WE such that:

- P.M.1. $TF \subseteq F$ and $WE = TF \circ TC$ (i.e. a map is a weak equivalence if and only if it factors as a trivial cofibration followed by a trivial fibration).
- P.M.2. F , C , TF , TC admit the following characterization by liftings.
- i f is a fibration if and only if f has right lifting property with respect to TC .
 - ii f is a trivial fibration if and only if f has right lifting property with respect to C .
 - iii If f has left lifting property with respect to TF , then f is a cofibration.
 - iv If f has left lifting property with respect to F , then f is trivial cofibration.
- P.M.3 Any map f admits two factorizations:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ h \searrow & & \nearrow i \\ & K & \end{array} \quad \text{and} \quad \begin{array}{ccc} X & \xrightarrow{f} & Y \\ h' \searrow & & \nearrow i' \\ & K' & \end{array}$$

where h is a cofibrations and i is a trivial fibration and h' is a trivial cofibration and i' is a fibration.

Remarks:

- i and ii implies that iii and iv in P.M.2 become equivalences.

- ii Any isomorphism belong to each one of the classes F, C, TF, TC (P.M.2) and hence to WE .
- iii F, C, TF, TC are closed under composition and retracts.
- iv $TC \subseteq C$, and furthermore $TF = F \cap WE$ and $TC = C \cap WE$.
- v Let

$$\begin{array}{ccc} X & \longrightarrow & Y \\ \bar{f} \downarrow & & \downarrow f \\ Z & \longrightarrow & K \end{array}$$

be a Cartesian square in \mathcal{A} . If f is a fibration (resp. trivial fibration) so is \bar{f} . Let

$$\begin{array}{ccc} X & \longrightarrow & Y \\ f \downarrow & & \downarrow \bar{f} \\ Z & \longrightarrow & K \end{array}$$

be a cocartesian square in \mathcal{A} . If f is a cofibration (resp. a trivial cofibration) so is \bar{f} .

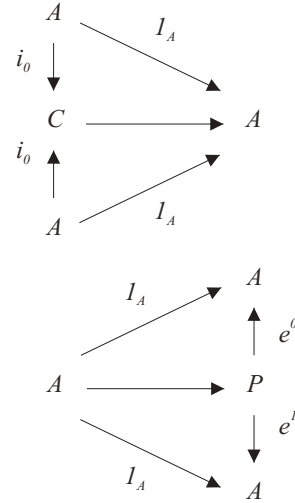
- vi $F \cap C \cap WE =$ isomorphisms of $\bar{\mathcal{A}}$.

Thus a pre model category have good behavior in liftings (as good as in closed model categories) but bad behaved WE : it only goes up to containing all isomorphisms and factorization $WE = TF \circ TC$.

The ‘‘closure’’ of a model category is define as follows [RR77]: given a model category (\mathcal{A}, F, C, WE) there exists a unique pre model category $(\bar{\mathcal{A}}, \bar{F}, \bar{C}, \bar{TF}, \bar{TC}, \bar{WE})$ such that $F \subseteq \bar{F}, C \subseteq \bar{C}, TF \subseteq \bar{TF}, TC \subseteq \bar{TC}$ and $WE \subseteq \bar{WE}$. This pre model category measures the extent in which a model category is a closed model category. It is proved that when the original model category is closed, it coincides with its closure. The closure can also be viewed as ‘‘the theory of liftings’’ of the model category. The uniqueness of the closure is implied by the following fact, meaningful in itself. If \mathfrak{A} is a class of morphisms in \mathcal{A} and we denote by $[\mathfrak{A}]$ the class of all retracts of members of \mathfrak{A} , then one has that $\bar{F} = [F], \bar{C} = [C], \bar{TF} = [TF], \bar{TC} = [TC]$.

Hence a model category is closed if and only if the classes F, C, TF, TC (or equivalently F, C, WE) are closed under retracts. In pre model, model and closed model categories a ‘‘cylinder’’ object (resp. ‘‘path’’ object) of A is a diagram as the next first (resp. second)

one.



C is the ‘‘actual’’ cylinder object of A , i_j ($j=0,1$) are trivial cofibrations (which are usually inclusions) and P is the actual ‘‘path object’’ of A and C^j are trivial fibrations (which usually are evaluation functions). Then one has for $f, g : A \rightarrow B$ that f is ‘‘left homotopic’’ to g if there exists a cylinder of A and $h : C \rightarrow B$ such that $h \circ i_0 = f$ and $h \circ i_1 = g$. The dual procedure provides ‘‘right homotopy’’. If the map $\phi \rightarrow A$ (ϕ initial object) belongs to C (A is a ‘‘co fibrant object’’) then left homotopy is an equivalence relation on $Hom(A, B)$ and left homotopy implies right homotopy [QD67]. Since dual assertions hold, right and left homotopy coincide and are equivalence relations on $Hom(A, B)$ when A is co fibrant and B is fibrant.

Elsewhere homotopy is done through homotopy systems: Let \mathcal{A} be a category. A homotopy system (Kan, [KD55,58]) $Z = (I, J_0, J_1, q)$ consists of the following:

- i A ‘‘cylinder’’ covariant functor $I : \mathcal{A} \rightarrow \mathcal{A}$.
- ii Three natural transformations $J_0 : 1_{\mathcal{A}} \rightarrow I, J_1 : 1_{\mathcal{A}} \rightarrow I, q = I \rightarrow 1_{\mathcal{A}}$ such that $qJ_0 = qJ_1 = 1$.

Homotopy is then given as follows: Let $f, g : X \rightarrow Y$ be morphisms and \mathcal{A} . We say that f is homotopic to g , denoted $f \simeq g$, if there exists a morphism $\rho : I(X) \rightarrow Y$ in \mathcal{A} such that $\rho \circ J_0(X) = f$ and $\rho \circ J_1(X) = g$.

The homotopy relation as defined is not in general an equivalence relation. But it is reflexive and compatible with composition: if $f, g : X \rightarrow Y, h : Y \rightarrow Z, k : K \rightarrow X$ then $f \simeq g$, implies that $hf \simeq hg$ and $fk \simeq gk$.

If \sim denotes the equivalence relation generated by \simeq , then one has the following definition: Let $f : X \rightarrow Y$ a morphism in \mathcal{A} . We say that f is a homotopy equivalence if $[f] \in \mathcal{A}(X, Y)/\sim$ is an isomorphism in $\text{Mor } \mathcal{A}/\sim$.

Fibrations and cofibrations for a homotopy system are given in Kamps [KK69]. We will provide now examples of homotopy systems. The most common ones are given in categories with final object different from the initial one, if it exists. We will denote by $*$ the final object and by \emptyset the initial one.

If \mathcal{A} is a category, then say that \mathcal{A} is *pointed* if $* \cong \emptyset$. Otherwise is *unpointed*. Let \mathcal{A} be unpointed. For an object A of \mathcal{A} the morphisms $* \rightarrow A$ are called the *points* of A and $\mathcal{A}(*, A)$ is called the *underlying set* of A . In unpointed categories there are plenty of homotopy systems. In fact, Let \mathcal{A} be an unpointed category closed for finite products. Let X be an object of \mathcal{A} and x_0, x_1 (if that order) points of X . Then there exists a natural isomorphism $\iota : 1_A \rightarrow 1_A \times *$, and the bimap $I : \mathcal{A} \rightarrow \mathcal{A}$, $A \mapsto A \times X$, $f \mapsto f \times 1_X$ is a covariant functor. Further, $d_i(A) : A \xrightarrow{\iota_A} A \times * \xrightarrow{1_A \times X_i} A \times X$, for $i = 0, 1$ and $\pi_1 : A \times X \rightarrow A$ are natural transformations on $A \in \text{Obj } \mathcal{A}$ and $\pi_1 \circ d_i(A) = 1_A$. Thus (I, d_0, d_1, s) is a homotopy system on \mathcal{A} .

Normal homotopies in the categories of topological spaces Top and simplicial sets $\Delta^\circ S$ are of this kind, the first one induced by $\{I = [0, 1], 0, 1\}$ and the second by $\{\Delta[1], \epsilon^0, \epsilon^1\}$ where $\epsilon^i : [0] \rightarrow [1]$, $0 \mapsto i$ induce simplicial points (still denoted by ϵ^0, ϵ^1) on the second. We assume the reader familiar with them.

Let $Z = (I, J_0, J_1, q)$ be a homotopy system in \mathcal{A} . We say that a map $F : E \rightarrow B$ is a *fibration* (or a Z fibration) if f has the right lifting property with respect to J_0 i.e. to the class of maps $J_0(X) : X \rightarrow I(X)$, $X \in \text{Obj } \mathcal{A}$. A map $i : A \rightarrow X$ is called a (Z) cofibration if for any commutative diagram.

$$\begin{array}{ccc} A & \xrightarrow{J_0(A)} & I(A) \\ \downarrow i & & \downarrow \\ X & \xrightarrow{f} & Y \end{array}$$

there exists a homotopy $\overline{\varphi} : I(X) \rightarrow Y$ such that $\overline{\varphi}I(i) = g$ and $\overline{\varphi}J_0(X) = f$.

One has the following properties for fibrations and cofibrations in a homotopy system [KK69]: Isomorphisms are fibrations and cofibrations. Projections are fibrations. Fibrations are closed under composition, base extension, and retracts. Cofibrations are closed under composition, co-base extension, and retracts.

2. Simplicial Systems, category, functor

We change a little Quillen's version of simplicial categories [QD67] to the notion of *simplicial systems* in a given category. The reason is that, as we will see, there may be more than one way in which a category is a simplicial category. For us then a "simplicial category" will be a pair formed by a category and a "simplicial system".

Let \mathcal{A} be a category. By a *simplicial system* in \mathcal{A} we mean a functor $\underline{Hom}_{\mathcal{A}} : \mathcal{A} \times \mathcal{A} \rightarrow \Delta^\circ S$ such that the following conditions hold:

- S.1. For any objects X, Y, Z of \mathcal{A} there exists a "composition" (simplicial) map:

$$\underline{Hom}_{\mathcal{A}}(X, Y) \times \underline{Hom}_{\mathcal{A}}(Y, Z) \rightarrow \underline{Hom}_{\mathcal{A}}(X, Z)$$

level wise denoted by $(f, g) \mapsto g \circ f$, which is associative in the sense that, if $f \in \underline{Hom}_{\mathcal{A}}(X, Y)_n$, $g \in \underline{Hom}_{\mathcal{A}}(Y, Z)_n$, and $h \in \underline{Hom}_{\mathcal{A}}(Z, K)_n$, then $(h \circ f) \circ g = h \circ (f \circ g)$.

- S.2. There exists a natural isomorphism

$$\lambda : \mathcal{A}(,) \rightarrow [\underline{Hom}_{\mathcal{A}}(,)]_0$$

denoted by $\mathcal{A}(X, Y) \rightarrow (\underline{Hom}_{\mathcal{A}}(X, Y))_0$; $u \mapsto \tilde{u}$ such that if $u \in \mathcal{A}(X, Y)$, $f \in \underline{Hom}_{\mathcal{A}}(Y, Z)_n$ and $g \in \underline{Hom}_{\mathcal{A}}(W, X)_n$ then

$$s_0^n(u) = \underline{Hom}_{\mathcal{A}}(u, Z)_n(f)$$

and

$$s_0^n(u) \circ g = \underline{Hom}_{\mathcal{A}}(W, u)_n(g)$$

where (abusing notation) s_0^n denotes the composite of the (in general different) functions

$$\underline{Hom}_{\mathcal{A}}(X, Y)_p \xrightarrow{s_0} \underline{Hom}_{\mathcal{A}}(X, Y)_{p+1}$$

for $p = 0, 1, \dots, n-1$

By a *simplicial category* we mean a pair $(\mathcal{A}, \underline{Hom}_{\mathcal{A}})$ where \mathcal{A} is a category and $\underline{Hom}_{\mathcal{A}}$ is a simplicial system on \mathcal{A} . As for functors among them:

Let $(\mathcal{A}, \underline{Hom}_{\mathcal{A}})$ and $(\mathcal{B}, \underline{Hom}_{\mathcal{B}})$ be two simplicial categories. By a *simplicial functor*

$$F : (\mathcal{A}, \underline{Hom}_{\mathcal{A}}) \rightarrow (\mathcal{B}, \underline{Hom}_{\mathcal{B}})$$

we mean a functor $F : \mathcal{A} \rightarrow \mathcal{B}$ together with maps

$$\underline{Hom}_{\mathcal{A}}(X, Y) \rightarrow \underline{Hom}_{\mathcal{B}}(F(X), F(Y)); f \mapsto F(f)$$

such that $F(f \circ g) = F(f) \circ F(g)$ and $F(\tilde{u}) = \tilde{F}(u)$. We say that F is *strictly simplicial* if the maps

$$\underline{Hom}_{\mathcal{A}}(X, Y) \rightarrow \underline{Hom}_{\mathcal{B}}(F(X), F(Y))$$

define a natural transformation (denoted again by F).

$$F : \underline{Hom}_{\mathcal{A}} \rightarrow \underline{Hom}_{\mathcal{B}}(F \times F)$$

Since the simplicial functors that we will use are always strictly simplicial we will talk simply of “simplicial functors” and refer to the “strict” part only when specially necessary.

Simplicial categories have cylinders and path objects: Let $(\mathcal{A}, \underline{Hom}_{\mathcal{A}})$ be a simplicial category. Let X be an object of \mathcal{A} and K a simplicial set.

- i By a cylinder object associated to (X, K) we mean a pair $(X \otimes K, \alpha)$ where $X \otimes K$ is an object of \mathcal{A} and

$$\alpha : K \rightarrow \underline{Hom}_{\mathcal{A}}(X, X \otimes K)$$

is a simplicial map such that for each Y in \mathcal{A} the simplicial map

$$\varphi : \underline{Hom}_{\mathcal{A}}(X \otimes K, Y) \rightarrow (\underline{Hom}_{\mathcal{A}}(X, Y))^K$$

next defined is an isomorphism: φ_n has domain $\underline{Hom}_{\mathcal{A}}(X \otimes K, Y)_n$, codomain $\Delta^{\circ} S(K \times \Delta[n], \underline{Hom}_{\mathcal{A}}(X, Y))$ and $\rho \mapsto (\mathbf{o}) \circ (\alpha \times 1) \circ (1 \times \underline{\rho})$. More explicitly the image of ρ is the composition of the maps: $1 \times \underline{\rho}$ with domain $K \times \Delta[n]$, and codomain $K \times \underline{Hom}_{\mathcal{A}}(X \otimes K, Y)$; $\alpha \times 1$ has domain $K \times \underline{Hom}_{\mathcal{A}}(X \otimes K, Y)$, and codomain $\underline{Hom}_{\mathcal{A}}(X, X \otimes K) \times \underline{Hom}_{\mathcal{A}}(X \otimes K, Y)$; \mathbf{o} with domain $\underline{Hom}_{\mathcal{A}}(X, X \otimes K) \times \underline{Hom}_{\mathcal{A}}(X \otimes K, Y)$, and codomain $\underline{Hom}_{\mathcal{A}}(X, Y)$ where

$$\underline{\rho} : \Delta[n] \rightarrow \underline{Hom}_{\mathcal{A}}(X \otimes K, Y)$$

is the simplicial map associated to ρ , namely the unique simplicial map such that $\underline{\rho}(l_{[n]}) = \rho$.

- ii By a path object associated to (X, K) we mean a pair (X^K, β) where X^K is an object of \mathcal{A} and $\beta : K \rightarrow \underline{Hom}_{\mathcal{A}}(X^K, X)$ is a simplicial function such that the induced map

$$\psi : \underline{Hom}_{\mathcal{A}}(Y, X^K) \rightarrow \underline{Hom}_{\mathcal{A}}(Y, X)^K$$

described below is an isomorphism: ψ_n with domain $\underline{Hom}_{\mathcal{A}}(Y, X^K)_n$, and codomain

$$(K \times \Delta[n], \underline{Hom}_{\mathcal{A}}(Y, X)),$$

where

$$\rho \mapsto (\mathbf{o}) \circ (Pr_2, Pr_1) \circ (1 \times \underline{\rho})$$

i.e. the image of ρ is the composition of the maps: $1 \times \underline{\rho}$ with domain $K \times \Delta[n]$, and codomain $K \times \underline{Hom}_{\mathcal{A}}(Y, X^K)$; (Pr_1, Pr_2) with domain $K \times \underline{Hom}_{\mathcal{A}}(Y, X^K)$, and codomain $\underline{Hom}_{\mathcal{A}}(Y, X^K) \times \underline{Hom}_{\mathcal{A}}(X^K, X)$; \mathbf{o} with domain $\underline{Hom}_{\mathcal{A}}(Y, X^K) \times \underline{Hom}_{\mathcal{A}}(X^K, X)$, and codomain $\underline{Hom}_{\mathcal{A}}(Y, X)$

Let $(\mathcal{A}, \underline{Hom}_{\mathcal{A}})$ be simplicial category. Let X be an object of \mathcal{A} and K, L simplicial sets. There are canonical isomorphisms $X \otimes (K \times L) \cong (X \otimes K) \otimes L$ and $(X^K)^L \cong X^{K \otimes L}$ when all of the objects involved are defined.

Recall that the homotopy relation on simplicial sets in general is not an equivalence relation. We denote by \simeq the homotopy relation in $\Delta^{\circ} S$ and by \sim the equivalence relation induced by \simeq . Recall further that the functional simplicial set associated to spaces X and Y (where the idea of simplicial categories was taken from) is the simplicial set $Hom(X, Y)$ whose n -th level is given as follows: $\underline{\Delta}$ denotes the cosimplicial space with $\underline{\Delta}^n = \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} \mid x_i \geq 0, \sum x_i = 1\}$ the cofaces $\delta^i : \underline{\Delta}^n \rightarrow \underline{\Delta}^{n+1}$ is the function that adds 0 in the i -th coordinate and $\rho^i : \underline{\Delta}^n \rightarrow \underline{\Delta}^{n-1}$, $x_i + x_{i+1}$ in the i -th coordinate. $\underline{\Delta}$ will be called “the standard model” in Top . Now, back to $Hom(X, Y)$ we take

$$\underline{Hom}(X, Y)_n = Top(\underline{\Delta}^n \times X, Y)$$

faces and degeneracies induced by those of $\underline{\Delta}$. It is clear that

$$\underline{Hom}(X, Y)_0 \cong Top(X, Y)$$

and is well known that two maps $f, g : X \rightarrow Y$ are homotopic in Top if as members of $\underline{Hom}(X, Y)_0$ they are homotopic. The generalization of this situation to simplicial categories is as follows: Let $(\mathcal{A}, \underline{Hom}_{\mathcal{A}})$ be a simplicial category, $f, g : X \rightarrow Y$ be morphisms in \mathcal{A} . We say that f is strictly homotopic to g , and denote it $f \simeq g$, if their images $\bar{f}, \bar{g} \in \underline{Hom}_{\mathcal{A}}(X, Y)_0$ are homotopic i.e. if $\bar{f} \simeq \bar{g}$. We say that f is homotopic to g , denoted $f \sim g$ if $\bar{f} \sim \bar{g}$ in $\underline{Hom}_{\mathcal{A}}(X, Y)$. We denote $\pi_0(\underline{Hom}_{\mathcal{A}}(X, Y))$ by $\pi_0(X, Y)$. The category $\pi_0 \mathcal{A}$ is defined as having as objects those of \mathcal{A} and for each pair X, Y of $Obj \mathcal{A} = (Obj \pi_0 \mathcal{A})$.

$$\pi_0 \mathcal{A}(X, Y) = \pi_0[\underline{Hom}_{\mathcal{A}}(X, Y)]$$

with composition induced by the one in \mathcal{A} .

The existence of path and cylinder objects provides a nice representation of $\underline{Hom}_{\mathcal{A}}(X, Y)$. In fact if

$(\mathcal{A}, \underline{Hom}_{\mathcal{A}})$ is a simplicial category and \mathcal{B} is a subcategory of $\Delta^{\circ}S$, then we say that \mathcal{A} admits a cylinder through \mathcal{B} if there exists a functor $\mathcal{A} \times \mathcal{B} \rightarrow \mathcal{A}$ such that for each $X \in \mathcal{A}$ and $K \in \mathcal{B}$ the image of (X, K) is a cylinder object of X in \mathcal{A} , say

$$\begin{array}{ccc} X & & \\ d^1 \downarrow & \searrow I_X & \\ X \otimes K & \xrightarrow{s} & X \\ d^0 \uparrow & \nearrow I_X & \\ X & & \end{array}$$

We say that \mathcal{A} admits paths through \mathcal{B} if there exists a functor $\mathcal{A} \times \mathcal{B}^{\circ} \rightarrow \mathcal{A}$ such that for each $X \in \mathcal{A}$ and $K \in \mathcal{B}$ the image of (X, F) is a path object of X in \mathcal{A} say

$$\begin{array}{ccc} & & X \\ & \nearrow I & \\ X & \xrightarrow{s} & X^K \\ & \searrow & \\ & & X \\ & & \downarrow D_1 \\ & & X \end{array}$$

Now we consider $\Delta : \Delta \rightarrow \Delta^{\circ}S$, the standard models of $\Delta^{\circ}S$. On it $\Delta_n = \Delta[n]$ is the simplicial set with $\Delta[n]_m$ the set of increasing functions $[n] \rightarrow [m]$ where $[p] = \{0, 1, \dots, p\}$. $d_i : \Delta[n]_m \rightarrow \Delta[n]_{m-1}$ maps $\alpha : [m] \rightarrow [n]$ in $\alpha \circ \delta_i$ where $\delta_i : [m-1] \rightarrow [m]$ is the 1-1 and increasing function which misses i in $[m]$. Further $\rho_j : \Delta[n]_m \rightarrow \Delta[n]_{m+1}$ maps α in $\alpha \circ \rho_j$ where $\rho_j : [m+1] \rightarrow [m]$ is the onto increasing function which repeats j in $[m]$. $\Delta[n]$ is thus a simplicial set. The $\Delta[n]$'s form a cosimplicial object of $\Delta^{\circ}S$ whose n -th level is $\Delta[n]$, $d_p^i : \Delta[n]_p \rightarrow \Delta[n+1]_p$ sends α in $\delta_i \circ \alpha$ and $s_p^j : \Delta[n]_p \rightarrow \Delta[n-1]_p$ sends α in $\rho_j \circ \alpha$. One thus have that $\Delta[n]$, $d^i : \Delta[n] \rightarrow \Delta[n+1]$ and $s^j : \Delta[n] \rightarrow \Delta[n-1]$ conform a cosimplicial object of $\Delta^{\circ}S$, "the model of the simplicial $\Delta[n]$'s".

We denoted by Δ again the image $\Delta(\Delta)$ which is a subcategory of $\Delta^{\circ}S$.

Suppose that \mathcal{A} admits cylinders through Δ . Then for each $X \in \mathcal{A}$ there exist a composition functor $\Delta \rightarrow \mathcal{A} \times \Delta \rightarrow \mathcal{A}$ with $[n] \mapsto X \otimes \Delta[n]$ and

$$(w : [n] \rightarrow [m]) \mapsto (1_X \otimes w^X : X \otimes \Delta[n] \rightarrow X \otimes \Delta[m])$$

This composition is a cosimplicial object of \mathcal{A} . Similarly if \mathcal{A} admits paths through Δ , then for each $X \in \mathcal{A}$ one

has a simplicial object of \mathcal{A} , $\Delta \rightarrow \mathcal{A} \times \Delta \rightarrow \mathcal{A}$ with $[n] \mapsto (X, \Delta[n]) \mapsto X^{\Delta[n]}$ and

$$(w : [n] \rightarrow [m]) \mapsto (1_X, w^*) \mapsto (X^{w^*} : X^{\Delta[m]} \rightarrow X^{\Delta[n]}).$$

One uses these two object to prove the following.

Suppose \mathcal{A} admits cylinders through Δ . Then for each pair X, Y of objects of \mathcal{A} , $\underline{Hom}_{\mathcal{A}}(X, Y)$ is (up to isomorphism) the simplicial set whose n -th level is $\underline{Hom}_{\mathcal{A}}(X, Y)_n = \mathcal{A}(X \otimes \Delta[n], Y)$. Suppose \mathcal{A} admits paths through Δ . Then for each X, Y objects of \mathcal{A} , $\underline{Hom}_{\mathcal{A}}(X, Y)$ is (up to isomorphism) the simplicial set whose n -th level is

$$\underline{Hom}_{\mathcal{A}}(X, Y)_n = \mathcal{A}(X, Y^{\Delta[n]})$$

Therefore in case \mathcal{A} admits cylinders through Δ , homotopy in \mathcal{A} is a left homotopy. For $f, g : X \rightarrow Y$ morphisms of \mathcal{A} , $f \sim g$ if and only if there exists a morphism $H : X \otimes \Delta[1] \rightarrow Y$ such that $H \circ d^1 = f$ and $H \circ d^0 = g$.

Similarly when \mathcal{A} admits paths through Δ , homotopy in \mathcal{A} is a (right) homotopy: $f \sim g$ if and only if there exists a morphism $T : X \rightarrow Y^{\Delta[1]}$ in \mathcal{A} such that $d_1 \circ T = f$ and $d_0 \circ Tg$.

Of course when \mathcal{A} admits path and cylinders through Δ , homotopy in \mathcal{A} is given by the (then) equivalent ways above.

We present now in some detail the most typical example of a "model object" at work which at long last produces a model category. Then we introduce its generalization to "model object" (or simply "model") and show that they produce pre model categories. In order to deal with the difference (from model to pre model categories) we work the missing parts (homotopy) by means of the homotopy system of its simplicial system.

3 Kan Fibrations and Trivial Fibrations

Let X be a simplicial set. Let $n \in N^*$. By a (simple) n -box we mean a subset $\{x_0, \dots, \hat{x}_k, \dots, x_n\}$ of X_n , (where \hat{x}_k means that there is no element indexed on k), such that for each pair $i, j = 0, 1, 2, \dots, n+1$, if $i < j$ then $d_i x_j = d_{j-1} x_i$. By a trivial n -box of X we mean a subset $\{x_0, \dots, x_{n+1}\}$ of X such that for each pair $i, j = 0, 1, \dots, n+1$ with $i < j$ then $d_i x_i = d_{j+1} x_i$.

The equalities $d_i x_j = d_{j-1} x_i$ for simple and trivial boxes will be referred to as the *compatibility relations* of the x_i 's. They are fulfilled emptyly for $n = 0$. So

0-boxes of X are subsets of X_0 with one element, and trivial 0-boxes, pair (x_0, x_1) , x_0 not necessarily different from x_1 .

Let $S = \{x_0, \dots, \hat{x}_k, \dots, x_n\}$ (respectively, $S' = \{x_0, \dots, x_{n+1}\}$). We say that x is a filler of S (resp. of S') if for each $i \neq k$, $d_i(x) = x_i$ (resp. for each i , $d_i(x) = x_i$).

If $f : X \rightarrow Y$ is a simplicial function and $S = \{x_0, \dots, \hat{x}_k, \dots, x_{n+1}\}$ is a box in X (resp. $S' = \{x_0, \dots, x_{n+1}\}$ is a trivial n -box in X) then $f(S) = \{f(x_0), \dots, f(x_{n+1})\}$ is an n -box in Y , (resp. $f(S')$ is a trivial n -box in Y).

Let $f : X \rightarrow Y$ be a simplicial map. f is said to be a *Kan fibration* if given any n -box S such that the image box $f(S)$ admits a filler $y \in Y_{n+1}$, then there exists a filler $x \in x_{n+1}$ of S , such that $f_{n+1}(x) = y$. f is said to be a trivial fibration if f_0 is onto and if given a trivial n -box S in X such that $f(S)$ admits a filler $y \in Y_{n+1}$, then there exists a filler $x \in x_{n+1}$ of S such that $f(x) = y$.

A simplicial set X is called a *Kan complex* (also is said to be (Kan) *fibrant*, or to have *Kan extension condition*) if the simplicial function $X \rightarrow *$ (where $*$ is any simplicial point) is a Kan fibration.

Let X be a simplicial set and $x \in X_n$. Recall that there exists a unique simplicial map $\tilde{x} : \Delta[n] \rightarrow X$ such that, $\tilde{x}(1_{[n]}) = x$. It is clear that if $f : X \rightarrow Y$ is a simplicial map, then $f \circ \tilde{x} = \tilde{f}(x)$. One also have that $\tilde{x} \circ d^i = d^i(\tilde{x})$ and $\tilde{x} \circ s^j = s^j(\tilde{x})$.

We call a set $\{a_0, \dots, \hat{a}_k, \dots, a_j\} = S$ with $a_j : \Delta[n] \rightarrow X$ a *functional n -box* in X if $a_j \circ d^i = d^{j-1} \circ a_i$, for $i < j$, $i, j \neq k$. When a_k is not omitted we call the set a *trivial functional n -box* in X . If $X \xrightarrow{f} Y$, then we denote by $f(S) = \{f \circ a_0, \dots, f \circ \hat{a}_k, \dots, f \circ a_{n+1}\}$ the functional n -box *image* of S .

We call $\Delta[n+1] \xrightarrow{a} X$ a filler of S if $a \circ d^i = a_j$ for $i \neq k$. Similarly for a trivial functional n -box.

Fibrations and trivial fibrations can be given by means of functional boxes: f is a Kan fibration if and only if for each functional n -box S whose image $f(S)$ admits a filler b , there exists a filler a of S such that $f \circ a = b$. f is a trivial fibration if and only if f_0 is onto and for each functional trivial n -box S whose image $f(S)$ admits a filler b there exists a filler a of S

such that $f \circ a = b$. One can produce trivial boxes from standard ones. A simple but tedious proof can be supplied for the following: if $S = \{x_0, \dots, \hat{x}_k, \dots, x_{n+1}\}$ is an n -box ($n \geq 1$) in a simplicial set X , then the set $\{\underline{x}_0, \dots, \underline{x}_n\} \subseteq X_{n-1}$ defined by

$$\underline{x}_i = \begin{cases} d_{k-1}(x_i) & \text{if } i < k \\ d_k(x_{i+1}) & \text{if } i \geq k \end{cases}$$

is a trivial $n-1$ box in X , which we denote by $d_k(S)$.

If $S = \{x_0, \dots, \hat{x}_k, \dots, x_{n+1}\}$ is a trivial n -box, such that $d_k(S)$ admits a filler x , then

$$\{x_0, \dots, x_{k+1}, \dots, x_{n+1}\}$$

is a trivial n -box.

It can be seen that any trivial fibration is a Kan fibration. Then one have a pre model category as follows: F is the class of Kan fibrations, TF is the class of trivial fibrations, C (resp. TC) is the class of maps with left lifting property for TF (resp. F) and WE (weak equivalences) is $TF \circ TC$. Quillen [QD67] proves that in fact F, C, WE is a closed model structure in $\Delta^\circ S$, that is $(\Delta^\circ S, F, C, WE)$ is a closed model category.

The use of maps of the kind $\Delta[n] \rightarrow X$ on the theorems allows us to generalize the concepts of fibrations, trivial and Kan complexes changing maps $\Delta[n] \rightarrow X$ to maps $Y^n \rightarrow X$ where $Y : \Delta \rightarrow \Delta^\circ S$ is a covariant functor. The version by boxes of the kind $\{x_0, \dots, \hat{x}_k, \dots, x_{n+1}\}$ and $\{x_0, \dots, x_{n+1}\}$ will be useful in order to relate the generalization through Y with the standard theory.

Now we provide the generalization of the concept of fibrations to get fibrations and trivial fibrations associated to a functor $Y : \Delta \rightarrow \mathcal{A}$.

We reestablish in detail the general definitions so as to point out how the standard theory is in fact a particular case of the theory induced by Y . Before we complete the pre model category associated to Y we give the counterpart (for Y) of the theorems in the previous paragraph of fibrations and trivial fibrations (when $\underline{\Delta}$ was used).

Next we show how these two concepts, Y fibrations and Y trivial fibrations, can be characterized (as well as some of their properties) by the use of the singular functor $S_Y : \mathcal{A} \rightarrow \Delta^\circ S$ associated to Y . Finally we complete the pre model category associated to Y .

4. Y fibrations and Y Trivial Fibrations

The generalizations of the concepts of boxes, trivial boxes and fillers can be given as follows.

Definition:

- i Let X be an object of \mathcal{A} . By a *simple Y^n box* (or a *Y^n box*) in X we mean a family of maps $\{a_0, \dots, \hat{a}_k, \dots, a_{n+1}\}$, $0 \leq k \leq n+1$, where $a_i : Y^n \rightarrow X$ is such that if $i < j$, $i, j \neq k$, then $a_j \circ d^i = a_i \circ d^{j-1}$, $d^i, d^{j-1} : Y^{n-1} \rightarrow Y^n$.
- ii By a *trivial Y^n box* for $n \geq 1$ we mean a family $\{a_0, \dots, a_{n+1}\}$ of maps $a_i : Y^n \rightarrow X$ such that $a_j \circ d^i = a_i \circ d^{j-1}$ if $i < j$, then $a_j \circ d^i = a_i \circ d^{j-1}$. A trivial Y^0 box is a family $\{a_0, a_1\}$, $a_i : Y^0 \rightarrow X$.
- iii Let $S = \{a_0, \dots, \hat{a}_k, \dots, a_{n+1}\}$ be a Y^n box (resp. $S' = \{a'_0, \dots, a'_{n+1}\}$ be a trivial Y^n box). By a *filer* of S (resp. S') we mean a map $a : Y^{n+1} \rightarrow X$ such that for each $i \neq k$ (resp. for each i) the following diagram commutes

$$\begin{array}{ccc}
 & Y^{n+1} & \xrightarrow{a} & X \\
 & \uparrow d^i & \nearrow a_i & \\
 & Y & &
 \end{array}$$

If $f : X \rightarrow Y$ is a simplicial map and if $S = \{a_0, \dots, \hat{a}_k, \dots, a_{n+1}\}$ is a Y^n box in X then it is clear that $\{fa_0, \dots, f\hat{a}_k, \dots, fa_{n+1}\}$ is a Y^n box in Y . Similarly the Y^n box (resp. trivial Y^n box) of the fa_i 's will be called the image of S by f and will be denoted in both the simple and trivial case by $f(S)$.

We require Y to have the following property : Let K_{n+1} be any set formed with at least $n-1$ of the maps $d^i : Y^n \rightarrow Y^{n+1}$ in \mathcal{A} . Then for any K_{n+1} there exists an object B on \mathcal{A} and a map $j : B \rightarrow Y^{n+1}$ such that each equation $j \circ X = d^i$ has solution (denoted $d_i|$) and further if B' and $j' : B' \rightarrow Y^{n+1}$ admit solutions to $j' \circ X = d^i$ then there exists a unique $H : B \rightarrow B'$ such that $j' \circ H = j$.

Thus j is unique up to isomorphism. When in K_{n+1} the map d^k is missing we denote j by $i : Y[n+1, k] \rightarrow Y^{n+1}$ and when $K_{n+1} = \{d^0, d^1, \dots, d^n\}$, j is denoted by $i : \delta Y^{n+1} \rightarrow Y^{n+1}$.

When $d^i : Y^n \rightarrow Y^{n+1}$ have underlying functions (and Y^n underlying set) then

$$Y[n+1, k] = \bigcup_{i \neq k} d^i(Y^n) \text{ and } \delta Y^{n+1} = \bigcup_{i=0}^n d^i(Y^n)$$

Of course in general there exist a unique

$$i_k : Y[n+1, k] \rightarrow \delta Y^{n+1}$$

such that (abusing the notation) $i \circ i_k = i$.

- i Let $f : X \rightarrow K$ be a map in \mathcal{A} . We say that f is a *Y fibration* if given any Y^n box S such that $f(S)$ admits a filler b , then there exists a filler a of S such that $fa = b$.
- ii f is said to be a *Y trivial fibration* if for any trivial Y^n box whose image $f(S)$ admits a filler b , there exists a filler a of S such that $fa = b$. Furthermore f has "the Y^0 lifting property" i.e. for any $b : Y^0 \rightarrow K$ there exists $a : Y^0 \rightarrow X$ such that $fa = b$.
- iii X is *Y Kan complex* (resp. *F trivial complex*) if $X \rightarrow *$ is a Y fibration (resp. Y trivial fibration).

Consider a family $\{a_0, \dots, \hat{a}_k, \dots, a_{n+1}\} = S$, $a_i : Y^n \rightarrow X$ ($n \geq 1$). Then S is a Y^n box (resp. a trivial Y^n box) iff there exists a map (filler)

$$Y[n+1, k] \xrightarrow{a} X \text{ (resp. } a : \delta Y^{n+1} \rightarrow X)$$

such that for each $i \neq k$ then $a \circ d^i| = a_i$. The family $\{a_0, \dots, a_{n+1}\}$ is a trivial Y^n box if there exists a map $a : \delta Y^n \rightarrow X$ such that $a \circ d^i = a_i$ for each i .

The effect on Y fibrations and Y trivial fibrations is the following:

Let $f : X \rightarrow K$ be a map on \mathcal{A} . Then f is a Y fibrations (resp. Y trivial fibration) iff f has right lifting property with respect to inclusions of the kind $Y[n, k] \xrightarrow{i} Y^n$, $n \geq 1$ and $0 \leq k \leq n$. (resp. $\delta Y^n \rightarrow Y^n$, $n \geq 0$).

Note that X is a Y Kan complex iff every Y^n box admits a filler. Furthermore X is an F trivial complex iff every trivial Y^n box admits a filler.

Recall that given $Y : \Delta \rightarrow \mathcal{A}$, there exists associated to it the singular functor $S_Y : \mathcal{A} \rightarrow \Delta^\circ S$, given by $(S_Y(X))_n = \mathcal{A}(Y^n, X)$ which gets faces and degeneracies from the ones of Y by composition, and if $f : X \rightarrow K$ then $S_Y(f)$ is the simplicial functions whose n -th level is given by $(S_Y(f))_n(a) = fa$. Note that $S_Y(X)$ is a simplicial set and that the elements of $(S_Y(X))_n$ are the kind of maps $Y^n \rightarrow X$ we have used to define Y^n boxes and trivial Y^n boxes.

Similarly fillers on Y^n boxes, trivial or otherwise, are fillers in the corresponding simplicial set $S_Y(X)$. It is rather simple to verify also, that $f : X \rightarrow K$ has the Y^0

lifting property if and only if $(S_Y(f))_0$ is onto. One gets then the following results.

Let $f : X \rightarrow K$ be a \mathcal{A} map: f is a Y fibration if and only if $S_Y(f)$ is a Kan fibration. f is a Y trivial fibration if and only if $S_Y(f)$ is a trivial fibration. Finally, any Y trivial fibration is a Y fibration.

5. Completing the pre model category associated to Y

Now we complete the Y structure on \mathcal{A} . To Kan fibration and trivial fibration one adds the following definition:

- i A map $f : X \rightarrow K$ is called a Y trivial cofibration if it has left lifting property with respect to the class of Y fibrations.
- ii f is called a Y cofibration if it has left lifting property with respect to the class of Y trivial fibration.
- iii f is called a Y weak equivalence if it factors as $j \circ h$ where h is a Y trivial cofibration and j is a Y trivial fibration.

From these classes of maps we are primarily concerned with fibrations, cofibrations and weak equivalences.

In what follows we will give conditions under which this five classes of maps form a pre model structure in \mathcal{A} . Thus far we have:

- i The following classes of maps are closed under composition and contain all of the isomorphism of \mathcal{A} : Y -fibrations, Y trivial fibrations, Y cofibrations, Y trivial cofibrations.
- ii The base extension of a map which is a Y fibration (resp. Y trivial fibration) is again a Y fibration (resp. a Y trivial fibration) The co base extension of a map which is a Y cofibration (resp. a Y trivial cofibration) is again a Y cofibration (resp. a Y trivial cofibration).
- iii The classes of Y fibration, Y trivial fibration, Y cofibration, and Y trivial cofibrations are closed under retracts.
- iv Suppose a map $X \xrightarrow{f} Y$ factors as $k \circ h$ with h a Y trivial cofibration and k a Y fibration (resp. with h a Y cofibration and k a Y trivial fibration). If f has right lifting property for Y trivial cofibrations, then f is a Y fibration. (resp. if f has right lifting property for Y cofibrations, then f is a Y trivial fibrations).

If the factorization axiom holds for the Y structure then: f is a Y fibration if and only if it has right lifting property for Y trivial cofibrations. Also f is a Y trivial fibration if and only if it has right lifting property with respect to Y cofibration.

As we have mentioned in a model category the prefix “trivial” has special meaning which for the Y structure is still valid. In fact f is a Y trivial fibration if and only if f is a Y fibration and a Y weak equivalence. Also f is a Y trivial cofibration if and only if is a Y cofibration and Y weak equivalence.

Thus so far the following axioms for model and closed model hold for the classes of structural maps associated to Y ; for model categories: $M.0, M.1, M.3, M.4$. For closed model categories: $C.M.1, C.M.4$ and partially $C.M.3$: $Y - F, Y - TF, Y - C$ are closed under retracts.

Now we want to give conditions on Y so that the factorization axiom hold. From the remark of proposition 2.24 it will follow as well that the classes of structural maps associated to Y form a pre model category.

6. Smallness and the factorization axiom

The concept of smallness that we use here is actually the sequential one used by Quillen [QD67] as well as the procedure to build up factorization of maps whenever there exists a family $\{A_i \rightarrow B_i\}$ with the A_i small. A more refined version can be found in RR94.

Let A be an object of \mathcal{A} . We say that A is *small* if

$$\text{Hom}(A, \varinjlim Z^m) = \varinjlim \text{Hom}(A, Z^m)$$

for any family $\{Z^m\}$, $m \in \mathbb{N}$. More precisely, A is small if for any sequence $Z^0 \rightarrow Z^1 \rightarrow Z^2 \rightarrow \dots$ of maps and for any $\alpha : A \rightarrow \varinjlim Z^n$, there exists $m \in \mathbb{N}$ and a map $\alpha : A \rightarrow Z^m$ such that $j \circ \alpha = \alpha$.

Let $Y : \Delta \rightarrow \mathcal{A}$ be a covariant functor. We say that Y is small if for each $n \in \mathbb{N}^*$, Y^n is small.

The cosimplicial character of Y implies smallness of some subobjects of the Y^n 's whenever Y is small. In fact let $I \subseteq [n] = \{0, 1, \dots, n\}$. Then $Y[n, I]$ is small. In particular $Y[n, k]$ and δY^n are small.

Notice also that if A is small and B is retract of A in the usual sense i.e. there are maps $B \xrightarrow{\rho} A$ and $A \xrightarrow{\eta} B$ such that $\eta\rho = 1_B$, then B is also small.

Here are some examples of small Y 's.

- i Δ is obviously small i.e. for each $n \in M$, $\Delta[n]$ is small. Hence so are $\Delta[n, k]$ and $\delta\Delta[n]$.
- ii If Y is small then for any $n \in N$, the object obtained by dropping the first n levels and last n cofaces and co-degeneracies, $RC^n(Y)$, is small. Hence in particular $RC^{(n)}(\Delta)$ is small for any $n \in N$.
- iii If Y is small then the functor $R_Y : \Delta^\circ S \rightarrow \Delta^\circ S$, left adjoint of $S_Y : \Delta^\circ S \rightarrow \Delta^\circ S$, send small objects into small object. That follows since S_Y commutes with sequential direct limits. Even more: Y is small if and if S_Y commutes with sequential direct limits.

We establish now the general situation implied by smallness about factorization of arrows.

Proposition: Let $A = \{A_i \rightarrow B_i \mid i \in I\}$ be a family of maps in \mathcal{A} . If for each $i \in I$, A_i is small, then any map $f : X \rightarrow Y$ factors as $f = k \circ h$ where h and k have the following properties.

- i h has the left lifting property for the class of maps with right lifting property for A .
- ii k has right lifting property for A .

Thus if $Y : \Delta \rightarrow \mathcal{A}$ is small then any map $f : X \rightarrow K$ factors in two ways: $f = k \circ h$ with $h \in Y - C$ and $k \in Y - TF$ and also as $f = k' \circ h'$ with $h' \in Y - TC$ and $k' \in Y - F$. That is to say, if Y is small, then factorizations axiom holds for the Y -structure. Also f is a Y fibration if and only if f has right lifting property for $Y - TC$ and f is a Y trivial fibration if and only if f has right lifting property for $Y - C$.

Hence we also have the following: when Y is small, \mathcal{A} together with the classes (or structural maps) $Y - F$, $Y - TF$, $Y - C$, $Y - TC$ and $Y - WE$ form a pre model category.

7. The Homotopy System Associated to Y

Although in Kan [KD55,561] the development of homotopy groups associated to homotopy system is done through cubical complexes, we have found easier to use simplicial sets and their standard homotopy groups in order to associate homotopy groups to a homotopy. In fact, as we will see, Y homotopy as we define it will become, for a suitable Y , a “simplicial homotopy”, namely the homotopy associated to the Y simplicial system (a concept studied in next paragraph) in which the machinery of standard homotopy of simplicial sets is available. Here, however, we develop Y homotopy independently of the Y simplicial system.

In this and the next paragraph we assume that Y is a “pointed” model of \mathcal{A} , that is to say Y^0 is the final object of \mathcal{A} .

Note that the following defines a homotopy system in \mathcal{A} :

- i $I : \mathcal{A} \rightarrow \mathcal{A}; X \mapsto X \times Y^1; \alpha \mapsto \alpha \times 1_{Y^1}$.
- ii $J_i : 1_{\mathcal{A}} \rightarrow I; X \mapsto J_i(X) : X \cong X \times Y^0 \xrightarrow{1 \times d^i} X \times Y^1$.
- iii $g : I \rightarrow 1_{\mathcal{A}}; X \mapsto g(X) : X \times Y^1 \xrightarrow{1 \times S^0} X \times Y^0 \cong X$.

We call it the Y homotopy system in \mathcal{A} and the corresponding homotopy is the Y homotopy. If $f, g : X \rightarrow K \in \mathcal{A}$, and f is homotopic to g through this homotopy we write $f \stackrel{Y}{\sim} g$, $f \sim g(Y)$ or when the use of Y^1 is to be emphasized we write $f \stackrel{Y^1}{\sim} g$.

Therefore $f \stackrel{Y}{\sim} g$ if and only if here exists $H : X \times Y^1 \rightarrow K$ such that $H \circ d^0 = f$ and $H \circ d^1 = g$.

As for the effect of $S_Y : \mathcal{A} \rightarrow \Delta^\circ S$ itself on Y homotopy, if $f \stackrel{Y}{\sim} g$, then $S_Y(f) \sim S_Y(g)$, where \sim denotes the standard homotopy in $\Delta^\circ S$. However we will present a structure which permits the use of simplicial homotopy in full power in favor and the one in \mathcal{A} induced and Y , including homotopy groups, exact homotopy sequences, etc, as done by Quillen [QD67].

8. The Y Simplicial Structure of \mathcal{A}

The Y simplicial structure is, roughly speaking, the generalization of functional complexes in $\Delta^\circ S$, when in $(X^K)_n = \Delta^\circ S(K \times \Delta[n], X)$ the models $\Delta[n]$ are substituted by Y^n , where $Y : \Delta \rightarrow \Delta^\circ S$.

The formal definition of simplicial system in categories, and the concept of simplicial category was given in the introduction. Here study first the machinery available in the simplicial system induced by special functors $Y : \Delta \rightarrow \mathcal{A}$, then the relations with the Y structure and Y homotopy. In this paragraph we consider a pointed Y which realizes products on standard models. Therefore its realization $R_Y : \Delta^\circ S \rightarrow \mathcal{A}$ commutes with finite products. On the other hand the following assignments define a functor $Hom_Y : \mathcal{A} \times \mathcal{A} \rightarrow \Delta^\circ S$:

- i For $X, K \in \mathcal{A}$, $Hom_Y(X, K) = S_{Y \times \overline{X}}(K)$.
- ii For $f : X^1 \rightarrow X$ and $g : K \rightarrow K^1$, $Hom_Y(f, g)$ is the map given level wise by

$$Hom_Y(f, g)_n : Hom_Y(X, K)_n \rightarrow Hom_Y(X^1, K^1)_n$$

which maps $\alpha \mapsto g \circ \alpha \circ (1_{Y^n} \times f)$.

Notice that $Hom_Y(X, K)$ is a generalization of K^X . In fact $Hom_Y(X, K)_n = S_{Y \times \bar{X}}(K)_n = \Delta^\circ S(Y^n \times X, K)$.

We show next that Hom_Y admits an associated simplicial composition, as required in a simplicial systems.

- i The following defines a simplicial functor for any simplicial sets X, K, Z :

$$Hom_Y(X, K)_n \times Hom_Y(X, Z)_n \rightarrow Hom_Y(X, ZK)_n$$

$$(\alpha, \beta) \mapsto \beta \circ \alpha = \beta \circ (\alpha \times 1_{Y^n}) \circ (1 \times \Delta(Y^n))$$

Or more graphically $\beta \circ \alpha$ is the following composition

$$X \times Y^n \xrightarrow{1 \times \Delta[n]} X \times Y^n \xrightarrow{\alpha \times 1_{Y^n}} K \times Y^n \xrightarrow{\beta} Z$$

- ii The composition of part i is associative in the sense that for X, K, Z, T and $f \in Hom_Y(X, K)_n$, $g \in Hom_Y(Z, K)_n$ and $h \in Hom_Y(Z, T)_n$ then $(h \circ f) \circ g = h \circ (f \circ g)$.

To complete the simplicial system we are required to show that $Hom_Y(X, K)$ is a simplicial set built up on $\Delta^\circ S(X, K)$ or to say better whose 0-level is $\Delta^\circ S(X, K)$ which behave appropriately with the simplicial composition.

- i The functions $\lambda(X, K)$ given by $\mathcal{A}(X, K) \rightarrow Hom_Y(X, K)_0$; $u \mapsto \tilde{u} (= X \times Y^0 \xrightarrow{\cong} X \xrightarrow{u} K)$ define a natural isomorphism.

- ii Let $f \in Hom_Y(K, Z)_n$, $u \in \Delta^\circ S(X, K)$ and let s_0^n denote the composition
- $$Hom_Y(X, K)_0 \xrightarrow{s_0} Hom_Y(X, K)_1 \xrightarrow{s_0} \dots \rightarrow Hom_Y(X, K)_n$$
- If $g \in Hom_Y(W, X)_n$ and $u \in \Delta^\circ(X, K)$ then

$$Hom_Y(W, u)_n(g) = s_0^n(u) \circ g$$

Part i is clear. For part ii notice that $s_0^n(\tilde{u}) = u_0((S^0)^n \times 1_X)$. Furthermore if $f : Y^n \times K \rightarrow Z$ one has a composition

$$Y^n \times X \xrightarrow{\Delta(Y^n) \times 1} Y^n \times Y^n \times X \xrightarrow{1 \times u_0((s_0^n)^n \times 1_X)} Y^n \times K \xrightarrow{f} Z$$

On the other hand recall that

$$Hom_Y(u, Z)_n : \mathcal{A}(Y^n \times K, Z) \rightarrow \mathcal{A}(Y^n \times X, Z)$$

maps $f \mapsto (Y^n \times X \xrightarrow{1 \times u} Y^n \times K \xrightarrow{f} Z)$.

We want to prove that the composition given above and the image of f by the last function coincide. But (level wise) the image of (Y^n, X) by the composition above is

$$\begin{aligned} f[1_{Y^n} \times \tilde{u}_0((S^1)_n \times 1_X)](\Delta(y^n) \times 1)(y^n, x) \\ = f[1_{Y^n} \times \tilde{u}_0((S^0)^n \times 1_X)](y^n, y^n, x) \\ = f(y^n, \tilde{u}((S^0)^n(y^n)), x) \\ = f(y^n, \tilde{u}(*, x)) \\ = f(y^n, u(x)) \end{aligned}$$

Part ii is proved similarly.

Remark:

- i Hom_Y with the simplicial composition given and the natural isomorphism

$$\Delta^\circ S(X, K) \rightarrow Hom_Y(X, K_0)$$

form a simplicial system which we refer to as the *simplicial system associated to Y or the Y simplicial system in \mathcal{A}* .

- ii So far the only condition used has been that Y is pointed. For the existence of cylinder and path objects as well as for the equivalence of right and left homotopy, we will need that Y realizes products on standard models.

In order to develop path objects X^K and cylinders objects $X \otimes K$ for the models in \mathcal{A} , and since one of the properties desired is the equality

$$Hom_Y(X \otimes_Y K, Y) \cong (Hom_Y(X, Y))^K$$

(see in the introduction: Simplicial system), we notice that

$$Hom_Y(X, Y)^K = S_{\Delta \times \bar{K}} S_{Y \times \bar{X}}(Y)$$

Therefore it would be helpful to have a relation between composition of singular functors and singular functors of composite functors. That we do next.

Let $Y, Z : \Delta \rightarrow \mathcal{A}$ be cosimplicial objects of \mathcal{A} . One has the composition $\Delta \xrightarrow{Z} \Delta^\circ S \xrightarrow{R_Y} \mathcal{A}$, which is again a model in \mathcal{A} . Then: *There exists a natural isomorphism* $\eta : S_Y \circ S_Z \rightarrow S_{R_Y \circ Z}$.

In fact, one can see that the isomorphism of adjointness of the pair $\Delta^\circ S \xrightarrow{R_Y} \mathcal{A} \xrightarrow{S_Y} \Delta^\circ S$ applied level wise to Z gives an isomorphism

$$\eta_n : \Delta^\circ S((R_Y(Z^n), X) \leftrightarrow \Delta^\circ S(Z^n, S_Y(X)))$$

with inverse, say ρ_n . It follows from naturality of the singular and realization functors that η is a simplicial map and is natural on X .

9. Existence of Cylinder and Path Objects

Let $X \in \mathcal{A}$ and $K \in \Delta^\circ S$. There exists a pair $(X \otimes_Y K, \alpha)$ with $X \otimes_Y K$ in \mathcal{A} and $\alpha_K : K \rightarrow$

$Hom_Y(X \otimes_Y K)$, which induces a natural isomorphism,

$$\alpha_K^* : Hom_Y(X \otimes_Y K, Y) \cong [Hom_Y(X, Y)]^K$$

Remark:

The isomorphism

$$Hom_Y(X \otimes_Y K, Y) \cong [Hom_Y(X, Y)]^K$$

has not been proved to be the one used by Quillen [QD67] within the framework of the simplicial structure of a simplicial category, but it works remarkably well and we saw no reason to insist on Quillen's isomorphism. The same remark is valid in the case of path objects, next.

Let $X \in \mathcal{A}$ and $K \in \Delta^\circ S$. There exists a pair $(X^K(\text{rel.}Y), \beta)$ where $X^K(\text{rel.}Y)$ is an object of \mathcal{A} and

$$\beta : K \rightarrow Hom_Y(X^K(\text{rel.}Y), X)$$

is a simplicial map, which induces a natural isomorphism

$$\beta_K^* : Hom_Y(Y, X^K(\text{rel.}Y)) \rightarrow [Hom_Y(Y, X)]^K$$

The following easy to check formulas are useful in the study of relations between the Y simplicial system, Y homotopy, and Y pre model structures.

Proposition:

i \otimes_Y is "associative" in the following sense:

$$X \otimes_Y (K \otimes L) = (X \otimes_Y K) \otimes_Y L$$

ii $(X^K(\text{rel.}Y))^L(\text{rel.}Y) \cong X^{K \times L}(\text{rel.}Y)$.

iii $S_Y (Hom_Y(X, K)) \cong S_T(K)$ where $T = (R_Y \circ Y) \times \overline{X}$.

iv For each n , $S_Y(Hom_Y(Y, X))_n = \Delta^\circ S(Y, X^N)$ where $N = R_Y(Y^n)$.

v $S_Y(\overline{X^{R_Y(K)}}) \cong (S_Y(X))^K \cong S_L(X)$ where $L = Y \times \overline{R_Y(K)}$.

10. Relative Subdivisions

In this parts we dealt with the relation among homotopies induces by models. Since there is little difference when using Δ or other category δ we use the later most of the time.

Kan [KD55,561,562] and other authors [FR68] have given characterizations of functors $\Delta^\circ S \rightarrow \Delta^\circ S$ called "subdivision functors" which are distinguished, among other things because the diagram below commutes up

to homotopy equivalence.

$$\begin{array}{ccc} \Delta^\circ S & \xrightarrow{H} & \Delta^\circ S \\ & \searrow \Downarrow & \downarrow \Downarrow \\ & & Top \end{array}$$

$|| : \Delta^\circ S \rightarrow Top$ denotes Milnor's geometric realization [MJ57].

Here we extend the concept of subdivision and give techniques to build subdivision functors in a more general context. First we work with the standard scheme category Δ and then, in a further generalization, any category δ is used as scheme. Thus instead of simplicial objects we work with general pre sheaves.

For the first part, instead of the geometric realization functor we use the realization of a model $M : \Delta \rightarrow Top$, and instead of normal homotopy we use that of a homotopy system γ . If $Sd : \Delta^\circ S \rightarrow \Delta^\circ S$ denotes the desired subdivision functor, $R_M : \Delta^\circ S \rightarrow Top$ the realization induced by M , and " \sim " the homotopy equivalence induced by γ , then one would like to have that $R_M(Sd(X)) \sim R_M(X)$ for each simplicial set X .

If \mathcal{A} and \mathcal{B} are categories, and a functor $R : \mathcal{A} \rightarrow \mathcal{B}$ admits right adjoint $S : \mathcal{B} \rightarrow \mathcal{A}$, we say that (R, S) is an adjoint pair. As for notation, $\Delta\mathcal{B}$ denotes the category of the cosimplicial objects of \mathcal{B} and $\Delta^\circ\mathcal{B}$ the simplicial ones. Also, if X is a simplicial object of \mathcal{B} we denote $X(n) = X_n$, and if $w : [n] \rightarrow [n]$ is an arrow of Δ , then $X(w) = w_*$. For the cosimplicial case the notation will be respectively X^n and w^* . Finally, if A is an object of \mathcal{B} then the constant functor of value A will be denoted by \dot{A} when is considered as a cosimplicial object of \mathcal{B} and \underline{A} when considered as simplicial object of \mathcal{B} .

11. Singular Functors and Realizations

According to the theory of adjoint pairs, a functor $Y : \Delta \rightarrow \mathcal{B}$ defines a functor $S_Y : \mathcal{B} \rightarrow \Delta^\circ S$ given on the objects by $(S_Y(A))_n = \mathcal{B}(Y^n, A)$ and if $w : [n] \rightarrow [n]$ then $(S_Y(A))(w)$ is given by composition with $Y(w)$. Now, if f is a morphism of \mathcal{B} then $S_Y(f)$ is given level wise by composition with f .

When \mathcal{B} is a co complete category S_Y admits a left adjoint denoted $R_Y : \Delta^\circ S \rightarrow \mathcal{B}$. The categories S (sets), $\Delta^\circ S$ and Top are co complete and therefore singular functors admit realizations. We now characterize them.

Given a model $Y : \Delta \rightarrow S$ (respectively $Z : \Delta \rightarrow Top$) then its induced realization is given by

$$R_Y(X) = \left(\coprod_n X_n \times Y^n \right) / t ,$$

respectively,

$$R_Z(X) = \left(\coprod_n X_n \times Z^n \right) / t ,$$

where \coprod denotes the union (respectively the topological sum), $X_n \times Y^n$ the set theoretical product (respectively $X_n \times Z^n$ denotes the topological product taking X^n as a discrete space), and the quotient is the set theoretical (resp. topological) quotient for the equivalence relation t that for each $w : [n] \rightarrow [n]$, x in X^n , and y in Y^n (respectively z in Z^n) identifies the couples $(w_*(x), y)$ and (x, w^*, y) (respectively $(w_*(x), z)$ and (x, w^*, z)). Also, if $f : X \rightarrow K$, then $R_Y(f)([x, y]) = [f_n(x), y]$ (respectively $R_Z(f)([x, z]) = [f_n(x), z]$), where $[x, y]$ denotes the class of equivalence of the couple (x, y) .

One proves that R_Y (respectively R_Z) is a covariant functor and that is left adjoint of S_Y (respectively S_Z).

In the case of $Y : \Delta \rightarrow \Delta^\circ S$ we denote, for each Y^n in $\Delta^\circ S$, $Y^n(m) = Y_m^n$. Thus in Y_m^n , n varies contra variantly and m co variantly. Is easy to see that, for each m , Y_m defines a set theoretical model $Y_m : \Delta \rightarrow S$ given by $Y_m[n] = Y_m^n$ and for each $w : [n] \rightarrow [m]$, $Y_m(w) = (Y(w))_m$.

For each $Y : \Delta \rightarrow \Delta^\circ S$ an adjoint functor of S_Y , say R_Y , is given on the objects by $(R_Y(X))_m = R_{Y_m}(X)$, and if $r : [n] \rightarrow [m]$ then $(R_Y(X))(r) = R_{Y_r}(X)$, where $Y_r : Y_n \rightarrow Y_m$ is the cosimplicial function $(Y_r)^p = Y^p(r)$ and $R_{Y_r} : R_{Y_n} \rightarrow R_{Y_m}$ is the natural transformation induced by Y_r on the set theoretical realization functors. On the morphisms, if $f : X \rightarrow K$ then $(R_Y(f))_m = R_{Y_m}(f)$.

The proof is easy but long. A known fact, worthwhile to clarify is the following: given two adjoint pairs $\mathcal{A} \xrightarrow{R} \mathcal{B} \xrightarrow{S} \mathcal{A}$ and $\mathcal{A} \xrightarrow{R'} \mathcal{B} \xrightarrow{S'} \mathcal{A}$, there exists a one to one and onto function $G : Trans(S, S') \rightarrow Trans(R', R)$, where $Trans$ denotes the class of natural transformations of the first functor in the second. On the other hand if $Y, Y' : \Delta \rightarrow \mathcal{B}$ are two models of \mathcal{B} and $a : Y \rightarrow Y'$ is a cosimplicial morphism then it induces a natural transformation $S_a : S_{Y'} \rightarrow S_Y$, given on an object A of \mathcal{B} by $(S_a : (A))_n(T) = T \circ a^n$ for each T in $(S_{Y'}(A))_n$. Here we have denoted $G(a) = R_a$. In previous results this mechanism was responsible for the equality $R_{Y_r} = G(S_{Y_r})$. Notice that in the set theoretical and

topological cases R_a is given by $R_a(X)[x, y] = [x, a(y)]$, while in the simplicial one is given by $(R_a : (X))_p = R_{a_p}(X)$.

The following result will be very useful in what follows:

Let \mathcal{B} be a co complete category, $Y : \Delta \rightarrow \mathcal{B}$ a model on \mathcal{B} and $\mathcal{B} \xrightarrow{I} \mathcal{B} \xrightarrow{J} \mathcal{B}$ an adjoint pair (notice that $\Delta \xrightarrow{Y} \mathcal{B} \xrightarrow{I} \mathcal{B}$ is a model on \mathcal{B}). Then there exists a natural isomorphism $P : R_{I \circ J} \cong I \circ R_Y$.

The proof follows from the fact that

$$\Delta^\circ S \xrightarrow{I \circ R_Y} \mathcal{B} \xrightarrow{S_Y \circ J} \Delta^\circ S$$

is an adjoint pair and $S_{I \circ Y}$ is naturally isomorphic to $S_Y \circ J$. In particular when \mathcal{B} is S , $\Delta^\circ S$ or Top , and $I : \mathcal{B} \rightarrow \mathcal{B}$ denotes the functor ‘‘product by A ’’, for a A fixed in \mathcal{B} , then I admits an adjoint when \mathcal{B} is S and $\Delta^\circ S$, but when \mathcal{B} is Top it happens only for some A . In such a case, for a model $Y : \Delta \rightarrow \mathcal{B}$ one will have that $\Delta \xrightarrow{Y} \mathcal{B} \xrightarrow{I} \mathcal{B}$ is precisely $Y \times \ddot{A}$. The proposition affirms then that

$$R_Y \times \ddot{A}(X) \xrightarrow{P} R_Y(X) \times A$$

naturally.

12. Homotopies on Models

A homotopy system on a category has an extension to a system on the category of its models. Indeed if $\gamma = (I, d^0, d^1, s)$ is a homotopy system in a category \mathcal{B} then γ induces on $\Delta \mathcal{B}$ the system $\Delta \gamma = (\Delta I, \Delta d^0, \Delta d^1, \Delta s)$ as follows: $\Delta I : \Delta \mathcal{B} \rightarrow \Delta \mathcal{B}$ is the normal extension of I to $\Delta \mathcal{B}$: if Y it belongs to $\Delta \mathcal{B}$ then $\Delta I(Y)^n = I(Y^n)$ and $\Delta I(Y)(w) = I(w)$. The natural transformations are given by the equalities $(\Delta d^i(Y))^n = d^i(Y^n)$ and $(\Delta s(Y))^n = s(Y^n)$, $i = 0, 1$.

Given two models Y, Z over \mathcal{B} and two cosimplicial morphisms $F, G : Y \rightarrow Z$ then it is clear that $F \overset{\Delta r}{\simeq} G$ if, and only if, for each n there exists h^n (γ -homotopy) of F^n in G^n such that for each $w : [n] \rightarrow [m]$, $Z(w)h^n = h^m I(Y(w))$.

Let us see now that homotopies and homotopy equivalences on models induce homotopies and homotopy equivalences on the realizations. We first advance some observations.

Given a homotopy system γ on \mathcal{B} , then to the natural transformation $d^i : 1_{\mathcal{B}} \rightarrow I$ there corresponds, in a bi univocal form, a natural transformation $D^i : J \rightarrow 1_{\mathcal{B}}$, when (I, J) is an adjoint pair. Similarly, if we denote

$d_{R_Y}^i : R_Y \rightarrow I \circ R_Y$ the transformation that in X has as morphism $d_{R_Y}^i : R_Y(X) \rightarrow I(R_Y(X))$, then the corresponding transformation on the singular functors is given for each A in \mathcal{B} by $S_Y(D^i(A)) : S_Y(J(A)) \rightarrow S_Y(A)$ and we denote it by $S_Y D^i$. Other two corresponding natural transformations (natural isomorphisms) that will use are $p' : S_{I \circ Y} \rightarrow S_Y \circ J$ with $p : I \circ R_Y \rightarrow R_Y \circ I$. Finally, there exists a cosimplicial morphism $Y \rightarrow Y \circ I$ that in the level n is $d_{Y_n}^i$. We will denote it by d_Y^i . We have then the corresponding natural transformations $R_{d_Y^i} : R_Y \rightarrow R_{I \circ Y}$ and $S_{d_Y^i} : S_{I \circ Y} \rightarrow S_Y$.

The following fact will be of great help in what follows: For $i = 0, 1$ the following diagram of natural transformations commutes:

$$\begin{array}{ccc}
 & R_Y & \\
 d_{R_Y}^i \swarrow & & \searrow R_{d_Y^i} \\
 I \circ R_Y & \xrightarrow{p} & R_{I \circ Y}
 \end{array}$$

Let \mathcal{B} be a category and γ a homotopy system whose cylinder I admits right adjoint J . Let Y and Z be models on \mathcal{B} . Then:

- i For each pair of simplicial morphisms $F, G : Y \rightarrow Z$, if $F \stackrel{\Delta^r}{\sim} G$ then for each simplicial set X one has that $R_F(X) \stackrel{\sim}{\sim} R_G(X)$, naturally on X . In particular,
- ii If Y is $\Delta\gamma$ -homotopically equivalent to Z , then for each simplicial set X , $R_Y(X)$ is γ -homotopically equivalent to $R_Z(X)$, naturally on X .

This result assures that if the cylinder I of a homotopy system γ on \mathcal{B} has right adjoint J , then the homotopy system induced by γ on the models, which we have denoted $\Delta\gamma$, has a direct effect on the realizations defined by the models. Indeed the proposition assures that if the models are homotopically equivalent then their realizations produce homotopically equivalent images for the original system on \mathcal{B} . In particular if we consider on Top the normal homotopy and if we take as Y the cosimplicial space of the topological simplexes Δ^n then R_Y is the geometric realization of Milnor. If Z is another model on Top , homotopically equivalent to the first one, then for each simplicial set X one has that $|X| \simeq R_Z(X)$, where \simeq denotes the normal homotopy equivalence on Top .

Notice that although each Δ^n is homotopically equivalent to a point not all realization of simplicial sets are

so. This happens because the homotopy equivalence in question on the models is not simplicial and therefore is not an equivalence of the extended homotopy. Naturally, if a model is level wise homotopically equivalent to a point, for a given homotopy, so that the given model and the cosimplicial point are homotopically equivalent, then the relation homotopy induced by the model is trivial.

In categories with final object, an object is homotopically trivial if it is homotopically equivalent to the final object. If the concept of sub object also exists and these are preserved by morphisms, then one has that if a model is homotopically trivial then the final model object is a sub object of the given model. Said otherwise: if the final model is not a sub object of the model then the last one cannot be homotopically trivial.

In the case of S , $\Delta^\circ S$ and Top this means that if a model doesn't have cosimplicial points then it is not homotopically trivial.

This advantage, together with the one of having Eilenberg–Zilber representations make of the models without cosimplicial points specially capable tools for good homotopy theories (see for example [RCRR81]).

Up to now we have studied conditions so that two models $Y, Z : \Delta \rightarrow \mathcal{B}$ have homotopically equivalent realizations for a homotopy system γ on \mathcal{B} . Now we consider functors $\Delta^\circ S \rightarrow \Delta^\circ S$ capable of preserving the homotopy type of the realizations.

Let $Y : \Delta \rightarrow \Delta^\circ S$ be a model of $\Delta^\circ S$ and $M : \Delta \rightarrow \mathcal{B}$ a model of \mathcal{B} . We consider a homotopy system $\gamma = (I, d^0, d^1, s)$ on \mathcal{B} for which I has a right adjoint J . In general the diagram

$$\begin{array}{ccc}
 & \xrightarrow{Y} & {}^\circ S \\
 M & \searrow & \downarrow R_M \\
 & & \mathcal{B}
 \end{array}$$

is not commutative, but it is when Y is the model of the simplicial simplexes $\Delta[n]$. However, when the diagram commutes up to a $\Delta\gamma$ -homotopy equivalence, then $R_M(S_Y(X))$ and $R_M(X)$ are γ -homotopically equivalent on \mathcal{B} . We develop this point next.

We will call $R_Y : \Delta^\circ S \rightarrow \Delta^\circ S$ a *subdivision* (or a subdivision of the identity) relative to the pair (γ, M) if $R_M \circ Y \stackrel{\Delta^r}{\sim} M$. When Δ denotes the model of the

$\Delta[n]$ simplexes, R_{Δ} is the identity of $\Delta^{\circ}S$ up to isomorphism. Then what we are doing is to compare at the homotopy level R_Y with $1_{\Delta^{\circ}S}$, a notion that allows us to assimilate $R_Y(X)$ to a “subdivision” of X . For example when γ is the homotopy system induced by $M^0 \xrightarrow{d^i} M^1 \xrightarrow{s} M^0$ ($i = 0, 1$), by means of the product, we can accept that if $Y \xrightarrow{\Delta^r} \Delta$ then the model Y is a “subdivision” of Δ . It will be seen that this implies that $R_M(X) \xrightarrow{\gamma} R_M(R_Y(X))$ which means that, at level of realizations (via R_M), X and $R_Y(X)$ have the same γ -homotopy type. For consistency we assimilate $R_Y(X)$ as a subdivision of X , achieved upon the subdivision Y of Δ limited only by homotopy. Then one goes further using any homotopy γ , and imposing conditions on Y so that the previous results still hold. The conditions on Y and Δ can hold, up to weak homotopy equivalence. That is the reason for the name of relative subdivisions.

Theorem: Let γ be a homotopy system on \mathcal{B} whose cylinder admits a right adjoint. Let M be a model on \mathcal{B} and Y a model on $\Delta^{\circ}S$. If R_Y is a subdivision relative to (γ, M) then there exists a natural transformation $a : R_M \rightarrow R_M \circ R_Y$ such that for each X in $\Delta^{\circ}S$, $a_X : R_M(X) \rightarrow R_M(R_Y(X))$ is a γ -homotopy equivalence.

For the proof, let us remember that, since $R_M \circ Y \xrightarrow{\Delta^r} M$, there exists a transformation $e : R_{R_M \circ Y} \rightarrow R_M$ such that for each X , e_X is a γ -homotopy equivalence. The theorem will be proved if we exhibit a natural isomorphism $R_M \circ R_Y \rightarrow R_{R_M \circ Y}$, or equivalently one natural isomorphism $S_{R_M \circ Y} \rightarrow S_Y \circ S_M$, but it is obtained by means of the following chain of natural isomorphisms $(S_{R_M \circ Y}(A))_n = \mathcal{B}((R_M \circ Y)^n, A) = \mathcal{B}(R_M(Y^n), A) = \Delta^{\circ}S(Y^n, S_M(A)) = (S_Y(S_M(A)))_n$.

A pair (Y, M) , where Y is a model of $\Delta^{\circ}S$ and M is a model of \mathcal{B} , is said to be a *singular pair* if $R_M \circ Y \cong M$.

It is clear that if (Y, M) is a singular pair then there exists a natural isomorphism $R_M \circ R_Y \rightarrow R_M$. Therefore Y is a subdivision relative to (γ, M) for any homotopy system γ on \mathcal{B} . The converse is also true since “ \cong ” is the homotopy relation of particular homotopy systems on any category. For example that is the case for any homotopy system in which $d^0 = d^1$.

Up to now we have defined subdivisions of the identity relative to pair (γ, M) . Now we consider subdivisions of any models.

Let γ be a homotopy system on \mathcal{B} whose cylinder admits a right adjoint and let M be a model of \mathcal{B} . If Y

and Z are models of $\Delta^{\circ}S$ we say that Y is a *subdivision of Z relative to (γ, M)* , denoted $Y \xrightarrow{(\gamma, M)} Z$, if $R_M \circ Y \xrightarrow{\Delta^r} R_M \circ Z$. Notice that if the homotopy of γ is transitive then the relation $\xrightarrow{(\gamma, M)}$ is an equivalence relation on the class of the models of $\Delta^{\circ}S$.

The relation $\xrightarrow{(\gamma, M)}$ compares models through its realizations in a weak way, via the homotopy equivalence induced by γ on the models of \mathcal{B} . Also, it compares their realizations (via R_M) by means of the homotopy equivalence of γ in \mathcal{B} . For if Y, Z, γ , and M are as above and if $Y \xrightarrow{(\gamma, M)} Z$ then there exists a natural transformation $a : R_M \circ R_Y \rightarrow R_M \circ R_Z$ such that for each X in $\Delta^{\circ}S$ the morphism $a_X : R_M(X) \rightarrow R_M(R_Z(M))$ is a γ -homotopy equivalence.

13. Subdivisions on Pre Sheaves

It is known that a covariant functor $Y : \delta \rightarrow \mathcal{A}$, where δ and \mathcal{A} are any categories, induces a covariant functor $S_Y : \mathcal{A} \rightarrow \delta^{\circ}S$. If \mathcal{A} is co complete then S_Y has a left adjoint functor, denoted here by $R_Y : \delta^{\circ}S \rightarrow \mathcal{A}$. Generalizing the terminology of [MJ57] and [RR76] we call

- i The functor $Y : \delta \rightarrow \mathcal{A}$ a δ -*model of \mathcal{A}* , or a *model of \mathcal{A}* when δ is clear.
- ii S_Y the *singular functor* and R_Y the *realization functor* of Y .

So far we have studied the repercussions on the functors R_Y and R_Z produced by the existence of a homotopy equivalence $Y \simeq Z$, induced in $\delta\mathcal{A}$ by a preset homotopy of \mathcal{A} when δ is the category Δ . The homotopy induced in $\Delta\mathcal{A}$ imposes conditions of naturality that occasionally can be very restrictive. One can see however that if one considers naturality on a restricted number of arrows of Δ , then the theory still works. This is equivalent to change the category Δ restricting arrows.

For generalization we take any unrestricted category δ and work as follows: first we give the extension of a homotopy system of \mathcal{A} to $\delta\mathcal{A}$ ($\delta\mathcal{A}$ the category of covariant functors $\delta \rightarrow \mathcal{A}$ and natural transformations as morphisms) and we show that if Y and Z are models of \mathcal{A} , homotopically equivalent for that extension, then for each pre sheaf X of $\delta^{\circ}S$, it happens that in \mathcal{A} , $R_Y(X)$ is homotopically equivalent to $R_Z(X)$, naturally on X . Further we fix a realization $R_M : \delta^{\circ}S \rightarrow \mathcal{A}$ and generalize the case of [KD55,561,562] to give conditions on two functors $F, G : \delta^{\circ}S \rightarrow \delta^{\circ}S$ so that the realizations

of $F(X)$ and $G(X)$ for R_M are naturally homotopically equivalent. Those said conditions define a relationship on the functors $\delta^\circ S \rightarrow \delta^\circ S$ that will be called of *relative subdivision*, extending the terminology of [KD57]. In general a standard homotopy in $\delta^\circ S$ doesn't exist. So we fix a homotopy system there, another in \mathcal{A} , and we give conditions relating these two systems by means of the functor $R_Y : \delta^\circ S \rightarrow \mathcal{A}$ induced by $Y : \delta \rightarrow \mathcal{A}$ and we say that “ Y carries the system of $\delta^\circ S$ into the one of \mathcal{A} ”.

If both Y and Z carry the system of $\delta^\circ S$ into the one of \mathcal{A} and if Y is homotopically equivalent to Z (for the homotopy of $\delta^\circ S$ extended from \mathcal{A}) then for each A of \mathcal{A} , $S_Y(A)$ is homotopically equivalent to $S_Z(A)$ in $\delta^\circ S$. When $\delta = \Delta$ and homotopy is the normal one in $\Delta^\circ S$ and $Y \simeq Z$ then the homologies induced by Y and Z in \mathcal{A} are isomorphic.

With normal changes the proofs for Δ can be adapted to the new generalization.

In what follows \mathcal{A} denotes a co complete category. For two models $Y, Z : \delta \rightarrow \mathcal{A}$ and a morphism of models $f : Y \rightarrow Z$ (natural transformation) we denote by $S_f : S_Z \rightarrow S_Y$ the transformation induced by f on singular functors.

We dealt now with the construction of a natural transformation $R_f : R_Y \rightarrow R_Z$.

Recall by [RCRR812] that if we denote by $\Phi_Y : 1 \rightarrow S_Y \circ R_Y$ and $\Psi_Y : R_Y \circ S_Y \rightarrow 1$ adjointness transformations for (R_Y, S_Y) , by Ψ_Z, Φ_Z , those of (R_Z, S_Z) and by

$$\begin{array}{ccc} \mathcal{A}(R_Y(X), A) & \longleftrightarrow & \delta^\circ S(X, S_Y(A)) \\ \alpha & \rightarrow & \alpha^Y \\ \beta_Y & \leftarrow & \beta \end{array}$$

(respectively $\alpha \rightarrow \alpha^Z, \beta \rightarrow \beta_Z$) the adjointness isomorphisms, then

$$R_f(X) = [S_f(R_Z(X)) \circ \Phi_Z(X)]_Y : R_Y(X) \rightarrow R_Z(X)$$

The extension of homotopy systems on \mathcal{A} to systems on $\delta\mathcal{A}$ is the following one: If $\eta = (I, d^0, d^1, s)$ is a homotopy system on \mathcal{A} then $\delta\eta = (\delta I, \delta d^0, \delta d^1, \delta s)$ denotes the system in $\delta\mathcal{A}$ with cylinder $\delta I(Y) = I \circ Y$ and $\delta I(\lambda)_x = I(\lambda_x)$ for each morphism $\lambda : Y \rightarrow Z$ in $\delta\mathcal{A}$, where $\delta d^i(Y) : Y \rightarrow \delta I(Y)$ is given by $\delta d^i(Y)_x = d^i(Y(x))$ and $\delta s(Y)$ by $\delta s(Y)_x = s(Y(x))$, $x \in \delta$.

The δn homotopy is a natural extension of the homotopy η because if Y, Z are objects of $\delta\mathcal{A}$ and $F, G : Y \rightarrow Z$ morphisms, then $F \overset{\delta\eta}{\sim} G$ (F is homotopic to G via $\delta\eta$) if, and only if, there exists a family

$h_X : I(Y(X)) \rightarrow Z(X)$ of η -homotopies $F_X \sim \rightarrow G_X$ such that if $w : X \rightarrow X'$ is a morphism of δ , then $Z(w)_0 h_X = h_{X'} \circ I(Y(x))$.

Since the homotopy relation of a homotopy system is not generally an equivalence relation we fix and keep an order that we exemplify for the system $\delta\eta$: If Z, Y are two models of \mathcal{A} (on δ) then Z is $\delta\eta$ homotopically equivalent to Y if there exist morphisms $F : Z \rightarrow Y$ and $G : Y \rightarrow Z$ such that $G \circ F \overset{\delta\eta}{\sim} 1_Z$ and $F \circ G \overset{\delta\eta}{\sim} 1_Y$.

If the cylinder of $\eta = (I, d^0, d^1, s)$ admits right adjoint $J : \mathcal{A} \rightarrow \mathcal{A}$ then J is completed in a right homotopy system denoted $\eta^* = (J, D^0, D^1, S)$, for which $f \overset{\eta}{\sim} g$ if and only if $f \overset{\eta^*}{\sim} g$. In fact if we denote the adjointness transformations by $\Phi : 1_{\mathcal{A}} \rightarrow J \circ I$ and $\Psi : I \circ J \rightarrow 1_{\mathcal{A}}$. For $\alpha : I(A) \rightarrow B$, $\alpha^* = J(\alpha) \circ \Phi_A$ and for $\beta : A \rightarrow J(B)$, $\beta_* = \Psi_B \circ I(\beta)$, then to the transformations d^i ($i = 0, 1$) there correspond the transformations $D_A^i : J \rightarrow 1_{\mathcal{A}}$ given for each A in \mathcal{A} by $D_A^i = \Psi_A \circ d^i J(A)$ or $\bar{d}^i = D^i$ with the notation of [RCRR811].

For a model $Y : \delta \rightarrow \mathcal{A}$, one has another model $\delta \xrightarrow{Y} \mathcal{A} \xrightarrow{I} \mathcal{A}$ on \mathcal{A} whose realization functors $R_{I \circ Y}$ and singular $S_{I \circ Y}$ are related with R_Y and S_Y in the following way.

The adjoint pairs $(I \circ R_Y, S_Y \circ J)$ and $(R_{I \circ Y}, S_{I \circ Y})$ are equivalent. That is to say that there exists a natural isomorphism $\bar{\rho}_Y : S_{I \circ Y} \rightarrow S_Y \circ J$ (and, equivalently, a natural isomorphism $\rho_Y : I \circ R_Y \rightarrow R_{I \circ Y}$).

In fact $(\bar{\rho}_Y(A))_X$ is the isomorphism of \mathcal{A} given by the following chain of natural isomorphisms: $[S_{I \circ Y}(A)]_X = \mathcal{A}(I \circ Y_{(X)}, A) = \mathcal{A}(I(Y_{(X)}), A) \stackrel{*}{=} \mathcal{A}(I \circ Y_{(X)}, A) = S_Y(J(A))_X$, where ρ_Y is the isomorphism induced by $\bar{\rho}_Y$ at realizations.

The effect on the realizations, of homotopy among models, is the following one:

Let $Y, Z : \delta \rightarrow \mathcal{A}$ be two models and $F, G : Y \rightarrow Z$ morphisms of \mathcal{A} . If $F \overset{\delta\eta}{\sim} G$ then there exists a natural transformation $\mathcal{H} : I \circ R_Y \rightarrow R_Z$ such that for each X in $\delta^\circ S$, \mathcal{H}_X defines a homotopy $R_F(X) \overset{\mathcal{H}_X}{\sim} R_G(X)$.

That is so because a $\delta\eta$ homotopy of F into G , say H produces a commutative diagram, where of course $d^i R_Y$ ($i = 0, 1$) is the transformation given by $d^i R_Y(X) = d^i(R_Y(X))$. One takes $\mathcal{H} = R_H \circ \rho_Y$. As a consequence one also has that if Y and Z are models of \mathcal{A} and Y is $\delta\eta$ homotopically equivalent to Z then there exists a natural transformation $\lambda : R_Y \rightarrow R_Z$ such that for

each X in $\delta^\circ S$, $\lambda_X : R_Y(X) \rightarrow R_Z(X)$ is a η homotopy equivalence.

$$\begin{array}{ccccc}
 & & R_Y & & \\
 & \swarrow d^0 R_Y & \downarrow R_{st(\eta)} & \searrow R_F & \\
 I \circ R_Y & \xrightarrow{\gamma} & R_{st(\eta)} & \xrightarrow{R_H} & R_Z \\
 & \swarrow d^1 R_Y & \uparrow R_{st'(\eta)} & \searrow R_G & \\
 & & R_Y & &
 \end{array}$$

When $\mathcal{A} = Top$ (or Kelly) and $\delta = \Delta$ for a model $Y : \Delta \rightarrow Top$ the realizations $R_Y(X)$ will be called $Y - CW$ complexes. They are thought of as complexes with (in general non Euclidean) cells Y^n . The last result establishes a sufficient condition so that the $Y - CW$ complex $R_Y(X)$ is homotopically equivalent to the CW complexes $|X|$ ($|$ Milnor's realization of [MJ57]).

In the topological case $Y - CW$ complexes are a particular case of cellular complexes with non Euclidean cells [RCRR812] studied by the author, professor Carlos Ruiz at National University of Colombia, and Joaquin Luna [RCLJ82]. However, for any co complete category \mathcal{A} and any model $Y : \delta \rightarrow \mathcal{A}$ on \mathcal{A} the concept of Y -complex exist. In fact the sub category of Y -complexes of \mathcal{A} is precisely $R_Y(\delta^\circ Conj)$, $R_Y : \delta^\circ Conj \rightarrow \mathcal{A}$.

Moving to the case of relative subdivisions among functors $\delta^\circ S \rightarrow \delta^\circ S$ we think of an object of $\delta^\circ S$ as a set theoretical skeleton that serves as a pattern to patch the pieces provided by $Y(X)$ of \mathcal{A} to obtain an object of \mathcal{A} . Up to now we have “modified” a model Y to obtain another Z , in such a way that for each skeleton X the obtained objects, one via patching objects $Y(X)$ and the other via patching objects $Z(X)$, were homotopically equivalent. Now we are interested in how to modify the skeleton X into another X' so that for $Y : \delta \rightarrow \mathcal{A}$ fixed, the objects obtained of them, patching $Y(X)$ and $Y(X')$, are homotopically equivalent in \mathcal{A} .

We consider the process $X \mapsto X'$ as a functor $\delta^\circ S \rightarrow \delta^\circ S$ which is the result of a “modification” of the (identity) functor $X \mapsto X$. Accepting that, it is clear that a more general situation arises considering two functors $F, G : \delta^\circ S \rightarrow \delta^\circ S$ for which we want to decide in which sense $G(X)$ is a “modification” of $F(X)$, in such a way that $G(X)$ and $F(X)$ are homotopically equivalent.

We use the way suggested by the precedent theory namely, we consider the realization functors $\delta^\circ S \xrightarrow{F} \delta^\circ S \xrightarrow{R_Y} \mathcal{A}$ and $\delta^\circ S \xrightarrow{G} \delta^\circ S \xrightarrow{R_Y} \mathcal{A}$, provided F and G each admits right adjoint. In such a case F and G are “modelable”. In other words they are realizations induced by models and we can concentrate on the models that define them. We will use the term “subdivision” rather than “modification” following the case exposed by Kan [KD57].

Let us fix a model $M : \delta \rightarrow \mathcal{A}$ and a system η in \mathcal{A} with right adjoint η^* . Let $Y, Z : \delta \rightarrow \delta^\circ S$ be models of $\delta^\circ S$. We will say that Z is a subdivision of Y relative to the couple (η, M) if $R_M \circ Y$ is $\delta\eta$ homotopically equivalent to $R_M \circ Z$. We denote it by $Y \xrightarrow{(\eta, M)} Z$.

If the homotopy η is transitive then $\xrightarrow{(\eta, M)}$ is an equivalence relation on the models of $\delta^\circ S$. With no conditions on η one has the wanted result: If $Y \xrightarrow{(\eta, M)} Z$, then there exists a natural transformation $L : R_M \circ R_Y \rightarrow R_M \circ R_Z$ such that for each object of $\delta^\circ S$, $L_X : R_M(R_Y(X)) \rightarrow R_M(R_Z(X))$ is a η homotopy equivalence.

When F and G admit right adjoint it will be said that G is a subdivision of F relative to (η, M) if $F \circ \delta \xrightarrow{(\eta, M)} G \circ \delta$ ($\delta : \delta \rightarrow \delta^\circ S$ the natural inclusion). Since $F \circ \delta$ (respectively $G \circ \delta$) is the model defining F (respectively G), what we have is that if G is a subdivision of F relative to (η, M) then for each X in $\delta^\circ S$, $R_M(F(X))$ is homotopically equivalent to $R_M(G(X))$ naturally on X . If $G \xrightarrow{(\eta, M)} 1_{\delta^\circ S}$ we will simply say that G is a “subdivision functor module (η, M) ”.

Notice that so far we have obviated the use of homotopy systems in $\delta^\circ S$. However there exists an intermediate step giving a homotopy system in $\delta^\circ S$ that is carried by M into η (see below). If Y is homotopically equivalent to Z for the extension of the homotopy from $\delta^\circ S$ then $Y \xrightarrow{(\eta, M)} Z$. In such a case, a sufficient condition for $R_Z(\delta^\circ S \rightarrow \delta^\circ S)$ to be a subdivision of R_Y relative to (η, M) is that $Y \simeq Z(mod \delta\eta)$.

14. Isomorphic Homologies

Returning to the case $\Delta \rightarrow \mathcal{A}$, the singular functor $S_Y : \mathcal{A} \rightarrow \Delta^\circ S$ gives place in an obvious way to a homology on \mathcal{A} induced by Y . With the same homotopy concept among models, $Y, Z : \Delta \rightarrow \mathcal{A}$ one can compare S_Y with S_Z using the standard homotopy of $\Delta^\circ S$. The process is also generalizable as we will see later. In this

paragraph we consider the effect produced by homotopy equivalencies of models on the respective singular functors.

We consider in \mathcal{A} a homotopy system $\eta = (I, d^0, d^1, s)$, with right adjoint $\eta^* = (J, D^0, D^1, \mathcal{S})$. There are then adjoint pairs $\mathcal{A} \xrightarrow{I} \mathcal{A} \xrightarrow{J} \mathcal{A}$ and $\mathcal{A} \xrightarrow{1_A} \mathcal{A} \xrightarrow{1_A} \mathcal{A}$. Once fixed adjointness transformations there exist isomorphisms $Trans(1_A, I) \rightarrow Trans(J, 1_A)$ and $Trans(I, 1_A) \rightarrow Trans(1_A, J)$. The first one sends d^i in D^i ($i = 0, 1$) and the second sends s in \mathcal{S} . If one also has an adjoint pair $\mathcal{A} \xrightarrow{R} \mathcal{B} \xrightarrow{S} \mathcal{A}$, with fixed adjointness transformations then there are two adjoint pairs in which we are interested namely $\mathcal{A} \xrightarrow{R} \mathcal{B} \xrightarrow{S} \mathcal{A}$ and $\mathcal{A} \xrightarrow{RI} \mathcal{B} \xrightarrow{JS} \mathcal{A}$, together with the following natural transformations and notations:

- $R(d_A^i) : R(A) \rightarrow R(I(A))$, $A \in \mathcal{A}$, $i = 0, 1$, denoted Rd^i .
- $R(s_A) : R(I(A)) \rightarrow R(A)$, $A \in \mathcal{A}$, denoted Rs .
- $D_{S(B)}^i : J(S(B)) \rightarrow S(B)$, $B \in \mathcal{B}$, $i = 0, 1$, denoted D_S^i .
- $S_{S(B)} : S(B) \rightarrow J(S(B))$, $B \in \mathcal{B}$, denoted S_S .

One can verify that if the adjointness transformations of (RI, JS) are those obtained from the pairs (I, J) and (R, S) , the isomorphism $Trans(R, RI) \rightarrow Trans(JS, S)$ then sends Rd^i in D_S^i ($i = 0, 1$) and $Trans(RI, R) \rightarrow Trans(S, JS)$ sends R_S in S_S . Similarly if $\beta \xrightarrow{H} \mathcal{A} \xrightarrow{T} \mathcal{B}$ is an adjoint pair then also is $\mathcal{B} \xrightarrow{H} \mathcal{A} \xrightarrow{I} \mathcal{A} \xrightarrow{J} \mathcal{A} \xrightarrow{T} \mathcal{B}$ and we have transformations and notations as follows:

- $d_{H(B)}^i : I(H(B)) \rightarrow H(B)$, $B \in \mathcal{B}$, $i = 0, 1$ denoted d_H^i .
- $\mathcal{S}_{H(B)} : H(B) \rightarrow I(H(B))$, $B \in \mathcal{B}$, denoted \mathcal{S}_H .
- $T(D_A^i) : T(J(A)) \rightarrow T(A)$, $A \in \mathcal{A}$, $i = 0, 1$ denoted TD^i .
- $T(\mathcal{S}_A) : T(A) \rightarrow T(J(A))$, $A \in \mathcal{A}$, denoted $T\mathcal{S}$.

with the obvious correspondences among them. Using the notation of [RCRR811] for two adjoint pairs $\mathcal{A} \xrightarrow{F_1} \mathcal{B} \xrightarrow{G_1} \mathcal{A}$ ($i = 1, 2$) that assigns $r \rightarrow \bar{r}$ by the isomorphism $Trans(F_1, F_2) \rightarrow Trans(G_1, G_2)$ one has the following group of formulas:

- i $\overline{Rd^i} = D_S^i$.
- ii $\overline{R_S} = S_S$.
- iii $\overline{d_H^i} = TD^i$.
- iv $\overline{\mathcal{S}_H} = T\mathcal{S}$.

In our case we are using models $Y : \delta \rightarrow \mathcal{A}$ each one with a adjoint pair $\delta^\circ S \xrightarrow{R_Y} \mathcal{A} \xrightarrow{S_Y} \delta^\circ S$, a homotopy system $\eta = (I, d^0, d^1, s)$ in $\delta^\circ S$ with right adjoint

$\eta^* = (\mathbb{J}, \mathbb{D}^0, \mathbb{D}^1, \mathbb{S})$ and the system η in \mathcal{A} with adjoint η^* .

We will say that Y realizes η into η (or equivalently η^* realizes into η^*) if a natural transformation $\epsilon : R_Y I \rightarrow IR_Y$ exists such that $\epsilon \circ (R_Y d^i) = d_{R_Y}^i$ for $i = 0, 1$.

As has been seen in [RCRR811], if $\mathcal{I} : \delta \rightarrow \delta^\circ S$ denotes the inclusion functor, then the previous definition can be given equivalently, restricting IR_Y and $R_Y I$ to $\mathcal{I}(\delta)$. That is to say demanding that for each x of δ a morphism $\epsilon_{Y(x)} : I(Y(x)) \rightarrow R_Y(I(Y(x)))$ exists such that $\epsilon_{Y(x)} \circ (R_Y d^i)_{Y(x)} = (d_{R_Y}^i)_{Y(x)}$ naturally on x .

When one uses $\Delta^\circ S \xrightarrow{R_Y} \mathcal{A} \xrightarrow{S_Y} \Delta^\circ S$ with the normal system in $\Delta^\circ S$, Y^0 final object of \mathcal{A} and η the system induced by product with $Y^0 \xrightarrow{d^i} Y^1$ ($i = 0, 1$) one has a natural transformation, $\epsilon_K : R_Y(K \times \Delta[1]) \rightarrow R_Y(K) \times R_Y(\Delta[1]) \simeq R_Y(K) \times Y^1$, the first part of which is induced by the projections and the second by the natural isomorphism $R_Y(\Delta[n]) \simeq Y^n$ connecting the normal cylinder in $\Delta^\circ S$ with the one of $R_Y(K)$ in \mathcal{A} . One also has the following commutative diagram for the inclusion $d_K^i : K \rightarrow K \times \Delta[1]$.

$$\begin{array}{ccc}
 & R_Y(K) & \\
 R_Y(d_K^i) \swarrow & & \searrow d_{R_Y(K)} \\
 R_Y(K \times \Delta[1]) & \xrightarrow{\quad \kappa \quad} & R_Y(K) \times Y^i
 \end{array}$$

It shows that R_Y also connects the inclusions of objects into their cylinders. This is the situation we just generalized. In [RCRR811] conditions were given so that ϵ becomes an isomorphism, case in which R_Y respects cylinders together with the inclusions and therefore transmits homotopies. Our condition here is then weaker than the one used normally for transmission of homotopies.

Let us suppose now that $F, G : Y \rightarrow Z$ are morphisms between models $Y, Z : \delta \rightarrow \mathcal{A}$. For each A , object of \mathcal{A} , one has two morphisms $S_F(A), S_G(A) : S_Z(A) \rightarrow S_Y(A)$. Let us see the implication derived on them by the fact that F and G are homotopic via $\delta\eta$.

Let us suppose that Y realizes η into η . Let $F, G : Y \rightarrow Z$ be two morphisms. If F is $\delta\eta$ -homotopic to G , then a natural transformation $S_Z(A) \rightarrow \mathbb{J}(S_Y(A))$ exists that is a right (or η^* -) homotopy of $S_F(A)$ into $S_G(A)$, for each A .

The proof is long and will be omitted.

Let $Y, Z : \Delta \rightarrow \mathcal{A}$ be two models of \mathcal{A} which realize η into η . If Y is $\delta\eta$ homotopically equivalent to Z then a natural transformation $\lambda : S_Y \rightarrow S_Z$ exists such that λ_A is a η (equivalently η^*) homotopy equivalence.

15. Examples

Let $Y : \Delta \rightarrow \mathcal{A}$ be a cosimplicial model of \mathcal{A} where \mathcal{A} is co complete, closed for finite products and with final object $*$. Let us suppose that $Y^0 = *$. Then one has in \mathcal{A} a homotopy system $\eta = (I, d^0, d^1, s)$ where

$$\begin{aligned} I(A) &= A \times Y^1 \\ d_A^i &: A \simeq A \times Y^0 \xrightarrow{1 \times d^i} A \times Y^1 (d^i : Y^0 \rightarrow Y^1) \\ s_A &: A \times Y^1 \xrightarrow{1 \times s^0} A \times Y^0 \cong A. \end{aligned}$$

Let us suppose now that $(-) \times Y_1$ has right adjoint. It is usually denoted $(-)^{Y^1}$. If η denotes the normal system in $\Delta^\circ S$ then Y realizes η in η . If a priori Y is $\Delta\eta$ equivalent to Z then a cosimplicial morphism exists $F : Z \rightarrow Y$, equivalence of $\Delta\eta$ homotopy, for which $d^i \circ F^0 = F^1 \circ d^i$.

This in turn implies at long last the following commutative diagram

$$\begin{array}{ccc} R_Z(X) & & \\ \downarrow R_Z(d_X^i) & \searrow d_{R_Z(X)}^i & \\ R_Z(X \times \Delta[1]) & \xrightarrow{(1_{R_Z(X)} \times F^1) \circ Pr} & R_Z(X) \times Y^1 \end{array}$$

Where Pr denotes the map $(R_Z(\pi_1), iso \circ R_Z(\pi_2))$. Therefore one has the following:

Let \mathcal{A} be a co complete category, closed for finite products, with final object $*$, and let $Y : \Delta \rightarrow \mathcal{A}$ be such that $Y^0 = *$ and $(-) \times Y^1$ has right adjoint. Let η be the system induced by $* = Y^0 \xrightarrow{d^i} Y^1 \xrightarrow{s^0} Y^0$ ($i = 0, 1$), and $Z : \Delta \rightarrow \mathcal{A}$ any model of \mathcal{A} . If Y is $\Delta\eta$ homotopically equivalent to Z then a natural transformation $\lambda : S_Y \rightarrow S_Z$ exists such that for each A of \mathcal{A} , $\lambda_A : S_Y(A) \rightarrow S_Z(A)$ is a homotopy equivalence in $\Delta^\circ S$.

Now we consider examples of homotopically equivalent models. There is a class of models in Top for which the homotopy extended from Top preserve details from the original that can be helpful. As an illustration let Y be a model of Top . We say that Y is convex if for each

n, Y^n is a convex subspace of a vector topological space (on \mathbb{R}) and for each $w : [n] \rightarrow [m]$, $w^* : Y^n \rightarrow Y^m$ is a lineal function. Then

- i If Y is a convex model, Z is any model of Top and $F, G : Z \rightarrow Y$ are any two cosimplicial continuous functions, then F is homotopic to G by the homotopy of ΔTop , extension of the normal homotopy of Top .
- ii If Y and Z are convex models of Top admitting cosimplicial functions $Y \rightarrow Z$ and $Z \rightarrow Y$ then Y is homotopically equivalent to Z for the homotopy of ΔTop extension of that of Top .

In what follows the homotopy of ΔTop (respectively of $\Delta Kelly$) extension of the normal homotopy of Top (respectively of $Kelly$) will be called the **normal homotopy of ΔTop** (respectively of $\Delta Kelly$).

We have that if Y is a convex model of Top (respectively $Kelly$) with cosimplicial points, then Y is null-homotopic, that is to say, homotopically equivalent (for the normal homotopy) to a cosimplicial point.

The converse is true for any not necessarily convex space. In fact, if Y has cosimplicial points and Z doesn't have, then Y and Z are not homotopically equivalent.

For the model of the complexes Δ^n of Top (or $Kelly$) its realization is the geometric realization [MJ57]. It is clear that $|K|$ is not in general equivalent (homotopically) to the discrete space $|_ \circ K$, which happens to be the realization $R_Y(K)$ for Y a cosimplicial point, even though each Δ^n is homotopically null.

Our definition of homotopy among models is somehow taking in consideration these facts. For convex spaces of the category $Kelly$ the relationship is very precise as we will see. But for the general case the existence of a natural transformation $\lambda : R_Y \rightarrow R_Z$ such that for each X in $\Delta^\circ S$, λ_X is a η homotopy equivalence doesn't seem to imply that $Y \xrightarrow{\delta\eta} Z$, unless λ_X^{-1} form a natural transformation at least for $X = \Delta[n]$, $n = 0, 1, \dots$.

Recall that for $Y : \Delta \rightarrow \mathcal{A}$, $RC(Y)$ ("The right cut of Y ", see [RR76]) is the cosimplicial object of \mathcal{A} defined this way: $[RC(Y)]^n = Y^{n+1}$, the $d^i : RC(Y)^n \rightarrow RC(Y)^{n+1}$ are the same $d^i : Y^{n+1} \rightarrow Y^{n+2}$ but only for $i = 0, 1, \dots, n+1$. Similarly for the $s^j : RC(Y)^n \rightarrow RC(Y)^{n-1}$ are the same $s^j : Y^{n+1} \rightarrow Y^n$ but only for $j = 0, \dots, n-1$. It is easy to see that $RC : \Delta\mathcal{A} \rightarrow \Delta\mathcal{A}$ is a covariant functor and that $d : 1 \rightarrow RC$ given for

$[d(Y)]^n = d^{n+1} : Y^n \rightarrow Y^{n+1}$ is a natural transformation. If $Y : \Delta \rightarrow Top$ (respectively Kelly) is convex then so is $RC(Y)$. Therefore if $Y : \Delta \rightarrow Top$ is a convex model and there exists a continuous cosimplicial function $RC(Y) \rightarrow Y$ then Y and $RC(Y)$ are homotopically equivalent for the normal homotopy in ΔTop .

Taking the case of Δ one has that:

$$\begin{aligned} d^i : RC(\Delta)^n &\rightarrow RC(\Delta)^{n+1} \\ (x_0, \dots, x_n) &\mapsto (x_0, \dots, x_{i-1}, x_i, \dots, x_{n+1}) \end{aligned}$$

for $i = 0, \dots, n+1$

$$\begin{aligned} s^j : RC(\Delta)^n &\rightarrow RC(\Delta)^{n+1} \\ (x_0, \dots, x_{n+1}) &\mapsto (x_0, \dots, x_j + x_{j+1}, \dots, x_{n+1}) \end{aligned}$$

for $j = 0, \dots, n+1$

$RC(\Delta)$ is not homotopically equivalent to Δ because the first one has the cosimplicial point $P^n = (0, \dots, 0, 1) \in \mathbb{R}^{n+2}$.

That example shows that, in general, do not exist simplicial morphisms $RC(Y) \rightarrow (Y)$ and also that $RC(\Delta)$ is homotopically null.

Now $RC(\Delta) - P$ is a cosimplicial space and $d : \Delta \rightarrow RC(\Delta) - P$ since $P^n \notin d^n(\Delta^n)$. Further $RC(\Delta) - P$ is convex since $(0, \dots, 0, 1)$ is a vertex of Δ^{n+1} . Finally the $f^{n-1} : (\Delta^n - P^{n-1}) \rightarrow \Delta^{n-1}$ that maps

$$(x_0, \dots, x_n) \mapsto \left(\sum_{i=0}^{n-1} x_i \right)^{-1} (x_0, \dots, x_{n-1})$$

is a cosimplicial continuous function $(RC(\Delta) - P) \rightarrow (\Delta - P^n)$.

Therefore one can also “extend” Δ to a model W with $W^n = R^{n+1}$ co faces and co degeneracies defined identically as those of Δ , thus Δ forming a cosimplicial subspace of W . But W also has a cosimplicial point, namely 0. Therefore it is homotopically null and cannot exist continuous cosimplicial functions $W \rightarrow \Delta$. Contrary to the previous case $W - 0$ is not a cosimplicial space. But if we consider $\mathcal{H}^n = \{(x_0, \dots, x_n) \mid \sum_{i=0}^n x_i = 0\}$

then $W - \mathcal{H}$ is a cosimplicial space and $0 \subseteq \mathcal{H} \subseteq W$. Now the function $W - \mathcal{H} \rightarrow \Delta$, $(x_0, \dots, x_n) \mapsto \left(\sum_{i=0}^n x_i \right)^{-1} (x_0, \dots, x_n)$ is continuous cosimplicial. However we cannot apply results for convex models since $W - \mathcal{H}$ is not convex (it is not connected). In this case it is clear that $W - \mathcal{H}$ is not homotopically equivalent to Δ since for each $n \neq 0$, $(W - \mathcal{H})^n$ is not homotopically null while Δ^n is, and clearly from the definition one has

that a necessary condition for $Y \xrightarrow{\delta\eta} Z$ is that for each $x \in \delta$, $Y(x) \xrightarrow{\eta} Z(x)$.

The model \mathcal{H} will be called by obvious reasons the **model of the hyperplane of W** . We notice that $\mathcal{H}^0 = \{0\}$ and \mathcal{H}^1 is the straight line at 135° that goes through the origin. Also $d^0 = d^1 : \mathcal{H}^0 \rightarrow \mathcal{H}^1$, which implies the interesting fact that this model’s homotopy is the equality and of course, homotopy equivalencies are the homeomorphisms.

If we consider $\mathcal{H}(k)^n = \{(x_0, \dots, x_n) \in R^{n+1} \mid \sum x_i = k\}$ with co faces and co degeneracies like the ones of W , then for each n , $\mathcal{H}(k)^n$ is homeomorphic to \mathcal{H}^n . However $\mathcal{H}(k)$ doesn’t have cosimplicial points if $k \neq 0$. This can be seen from the fact that if $\mathcal{H}(k)^0$ has cosimplicial points then has only one, since $\mathcal{H}(k)^0$ has one element, which must belong to the cosimplicial point. But $d^0(k) = (0, k)$ and $d^1(k) = (0, k)$. As $(0, k) = (k, 0) \leftrightarrow k = 0$, then $\mathcal{H}(k)$ doesn’t have cosimplicial points if $k \neq 0$, although $R_{\mathcal{H}}$ and $R_{\mathcal{H}(k)}$ are level wise homeomorphic. Since $\partial\Delta[1]$ is realized by $R_{\mathcal{H}}$ into $\{(0, 0)\} \subseteq \{(x, y) \in R^2 \mid x + y = 0\}$ and by $R_{\mathcal{H}(k)}$ into the subspace $\{(0, k), (k, 0)\}$ of $\{(x, y) \in R^2 \mid x + y = k\}$ then it is clear that the realizations are not homotopically equivalent. Apart from this $\mathcal{H}(k) \simeq \mathcal{H}(l)$ if $k \neq 0 \neq l$.

The question on whether two models can have homotopically equivalent realizations without being themselves homotopically equivalent is partially clarified for convex models of the category Kelly with next result which improves the late version: Let Y, Z be two convex models of Kelly. The following statements are equivalent:

- i Y is homotopically equivalent to Z for the normal homotopy of ΔKelly .
- ii There exist natural transformations $\lambda : R_Y \rightarrow R_Z$ and $\rho : R_Z \rightarrow R_Y$.

We notice that $\Delta, RC(\Delta) - P, W$ and $\mathcal{H}(k)$ are models of the category Kelly. Since in this category and in that of the simplicial sets the functor $(-)\times A$ has right adjoint $(-)^A$ for each A , then in particular the cylinder functors $(-)\times\Delta[1]$ in $\Delta^\circ S$ and $(-)\times I$ have right adjoint. It is also known that Kelly is a co complete category. At the level of realizations one thus has that

- i There exists a natural transformation $\lambda : R_{RC(\Delta)} \rightarrow R_P$ such that for each simplicial set X , $\lambda_X : R_{RC(\Delta)}(X) \rightarrow R_P(X)$ is a homotopy equivalence.

- ii There exists a natural transformation $\eta : R_W \rightarrow R_P(X)$ such that for each simplicial set X , $\eta_X : R_W(X) \rightarrow R_P(X)$ is a homotopy equivalence.
- iii There exists a natural transformation $\rho : R_{RC(\Delta)} \rightarrow R_\Delta = | \cdot |$ that for each simplicial set X , $\rho_X : R_{RC(\Delta)}(X) \rightarrow |X|$ is a homotopy equivalence.

Because of parts i and ii for each simplicial set the spaces $\pi_0(X)$ (discreet), $R_{RC(\Delta)}(X)$ and $R_W(X)$ are homotopically equivalent, naturally on X .

As for the singular functors since the homotopy of Δ in Kelly is the normal homotopy then by using relations already given for singular functors one has that: *there exists a transformation $\epsilon : S_{RC(\Delta)} \rightarrow S_\Delta = Sing$ such that for each space A of Kelly, $\epsilon_A : S_{RC(\Delta)}(A) \rightarrow Sing A$ is a homotopy equivalence in $\Delta^\circ S$.*

Under the context of the theory here developed parts i and iii above are not extendable to singular functors since *cosimplicial points do not carry the homotopy of $\Delta^\circ S$ into that of Kelly*. In fact suppose the opposite and consider the diagrams ($i = 0, 1$).

$$\begin{array}{ccc}
 R_P(X) & & \\
 \downarrow R_P(d_X^i) & \searrow d_{R_P(X)}^i & \\
 R_P(X \times \Delta[1]) & \xrightarrow{R_P} & R_P(X) \times I
 \end{array}$$

When X is a simplicial point $R_P(X)$ and $R_P(X \times \Delta[1])$ have a single point. Therefore $\epsilon_X R_P(d_X^0) = \epsilon_X R_P(d_X^1)$, or equivalently $(*, 0) = (*, 1), * \in R_P(X)$.

In the same way W and $R_C(A)$ are homotopically equivalent but the theory cannot be applied at level of singular functors. Indeed by a similar argument to the one just given it follows that W and $RC(\Delta)$ do not carry the homotopy of $\Delta^\circ S$ in that of Kelly.

References

- [AM95] **Marc Aubry**, *Homotopy Theory and Models*, DMV Seminar B. 24, Birkhäuser Verlag, 1965.
- [DFDJ] **F. Bauer & J. Dugundji**, *Categorical Homotopy and Fibrations*, Trans. Am. Math. Soc. **140** (1969), 239-256.
- [CJ90] **J. Cintura**, *Closed structures on reflective subcategories of the category of topological spaces*, Theory and Appl of Categories **37** (1990), 237-247.
- [DC56] **C. H. Dowker**, *Homotopy extension theorems*, Proc. London Math. Soc. **6** (3) (1956), 100-116.
- [FR] **R. Fritsch**, *On subdivisions of semisimplicial sets*, Proc. of the International Symposium of Topology and Applications. Herceg-Novi (Yugoslavia), 25-31.8, 1968, 156-163.
- [GPZM67] **P. Gabriel & M. Zisman**, *Calculus on Fractions and Homotopy Theory*, Springer-Verlag, Berlin, 1967.
- [HK91] **K. A. Hardie & K. H. Kamps**, *The coherent homotopy category over a fixed space is a category of fractions*, Topology and its Applications **40** (1991), 265-274.
- [HU59] **S. T. Hu**, *Homotopy Theory*, Academic Press, New York, 1959.
- [JI84] **I. James**, *General Topology and Homotopy Theory*, Springer-Verlag, Heidelberg, 1984.
- [KK] **K. H. Kamps**, *Fibrations and cofibrations in categories with homotopy systems*. Proc. of the International Symposium of Topology and Applications. Herceg-Novi (Yugoslavia), 25-31.8, 1968, 211-218.
- [KK73] **K. H. Kamps**, *Bemerkungen zum Homotopiesatz für semisimpliziale Faserungen*, Bulletin de L'Académie Polonaise des Sciences **21** (4)1973.
- [KK97] **K. H. Kamps**, *Abstract Homotopy and simple homotopy theory*, World Scientific, Singapore, 1997.
- [KD58] **Daniel Kan**, *Adjoint functors*, Trans. Amer. Math. Soc. **87** (1958), 294-329.
- [KD55] **Daniel Kan**, *Abstract Homotopy I*, Proc. National Acad. Sciences **41** (1955), 1092-1096.
- [KD561] **Daniel Kan**, *Abstract Homotopy II*, Proc. National Acad. Sciences **42** (1956), 225-228.
- [KD562] **Daniel Kan**, *Abstract Homotopy III*, Proc. National Acad. Sciences **42** (1956), 419-421.
- [KD57] **Daniel Kan**, *On CSS complexes*, Amer. J. Math. **79** (1957), 449-476.
- [MJ90] **Jean P. Mayer**, *Cosimplicial homotopies*, Proc. Amer. Math. Soc. **108** (1990), 9-17.
- [MP67] **Peter May**, *Simplicial Objects in Algebraic Topology*, D. van Nostrand, Princeton, 1967.
- [MJ57] **J. W. Milnor**, *The geometric realization of a semi-simplicial complex*, Ann. Math. **65** (1957), 357-362.
- [OG89] **Guillermo Ortiz**, *Modelos de homotopía*, Tesis de Maestría, Universidad del Valle, Colombia, 1989.
- [QD67] **Daniel Quillen**, *Homotopical Algebra*, Lectures Notes 43, Springer-Verlag, Berlin, 1967.
- [RCLJ82] **Carlos Ruiz & Joaquín Luna**, *Normalidad en realizaciones generalizadas ($R_Y(X)$)*, Collectanea Mathematica **33** (1982), 289-297.
- [RCRR71] **Carlos Ruiz & Roberto Ruiz-Salguero**, *La condición de Kan y la noción de grupoide*, Rev. Colombiana Mat. **5** (1971), 59-82.
- [RCRR73] **Carlos Ruiz & Roberto Ruiz-Salguero**, *Kan fibrations which are homomorphisms of simplicial groups*, Rev. Colombiana Mat. **7** (1973), 23-43.
- [RCRR76] **Carlos Ruiz & Roberto Ruiz-Salguero**, *Any equivalence relation over a category is a simplicial homotopy*, Rev. Colombiana Mat. **10** (1976), 151-160.

- [RCRR78] **Carlos Ruiz & Roberto Ruiz-Salguero**, *Remarks about the Eilenberg-Zielber type decomposition in cosimplicial sets*, Rev. Colombiana Mat. **12** (1978), 61–82.
- [RCRR811] **Carlos Ruiz & Roberto Ruiz-Salguero**, *Conditions over a realization functor to commute with finite products*, Rev. Colombiana Mat. **15** (1971), 113–146.
- [RCRR812] **Carlos Ruiz & Roberto Ruiz-Salguero**, *Characterization of the set theoretical geometric realization in the non Euclidean case*, Proc. Amer. Math. Soc. **81** (1981), 321–324.
- [RR] **Roberto Ruiz-Salguero & Guillermo Ortiz**, *Levantamientos y homotopía abstracta*, Lecturas Matemáticas **9** (1-3)(1988), 69–100.
- [RR75] **Roberto Ruiz-Salguero**, *Changes of Models in Algebraic Topology*, Ph. D. Thesis, Temple University, 1975.
- [RR77] **Roberto Ruiz-Salguero**, *The closure of a model category*, Rev. Colombiana Mat. **11** (1977), 19–50.
- [RR79] **Roberto Ruiz-Salguero**, *Subdivisiones relativas*, Rev. Colombiana Mat. **13** (1979), 92–95.
- [RR80] **Roberto Ruiz-Salguero**, *Subdivisiones relativas sobre prehaces y homologías simpliciales isomorfas*, Rev. Colombiana Mat. **14** (1980), 17–42.
- [RR94] **Roberto Ruiz-Salguero**, *Liftings in categorical homotopy*, J. Pure Appl. Algebra **96** (1994), 157–172.
- [VR73] **R. M. Vogt**, *Homotopy limits and colimits*, Math. Z. **134** (1973), 11–52.
- [WH49] **J. H. Whitehead**, *Combinatorial homotopy*, Bull. Amer. Math. Soc. **55** (1949), 231–245.

GEOMETRÍA DE VARIEDADES BANDERA

por

Marlio Paredes y Sofía Pinzón¹

Dedicado a la memoria del profesor Don Jairo Charris

Resumen

Paredes, Marlio & Sofía Pinzón: Geometría de variedades bandera. Rev. Acad. Colomb. Cienc. **28** (106): 123–134, 2004. ISSN 0370-3908.

En este trabajo hacemos un recuento histórico de resultados sobre la geometría de las variedades bandera. Principalmente abordamos la relación existente entre métodos combinatorios y propiedades geométricas.

Palabras clave: Variedades bandera, geometría hermitica, sistemas de raíces, grupos de Lie semisimples.

Abstract

In this work we make a historical recount of results on the geometry of flag manifolds. Mainly we study the relation between combinatorial methods and geometric properties.

Key words: Flag manifolds, Hermitian geometry, root systems, semi-simple Lie groups.

1. Introducción

Una variedad bandera es un espacio homogéneo reductivo $G/C(S)$, donde G es un grupo de Lie complejo semisimple y $C(S)$ es el centralizador de un toro S (no necesariamente maximal en G). Cuando S es un toro maximal decimos que la variedad bandera es maximal, y la denotamos por \mathbb{F} .

Nuestro objeto inicial de estudio son las variedades bandera hermiticas, es decir, una variedad bandera dotada de una estructura cuasicompleja J .

Presentamos aquí algunas nociones básicas que nos permitan introducir, en el lenguaje de álgebras de Lie

¹Escuela de Matemáticas, Universidad Industrial de Santander, Apartado Aereo 678, Bucaramanga, Santander, Colombia. Email: mparedes@uis.edu.co y Email: spinzon@uis.edu.co Este trabajo ha sido financiado parcialmente por Colciencias, Contrato No. RC-240-2001.

AMS Classification 2000: 14M15, 53C55, 22F30, 05C20.

semisimples complejas de dimensión finita, los resultados existentes sobre la geometría de variedades bandera.

Hacemos un recuento histórico de los principales trabajos realizados alrededor del estudio de las estructuras cuasicomplejas y su relación con la combinatoria de los torneos (grafos dirigidos completos). Presentamos las condiciones establecidas en este contexto para que la variedad bandera maximal clásica admita métricas (1,2)-simpléticas (ver, por ejemplo [20], [25], [7],[8],[9], [32]).

Mostramos la generalización conseguida por Pinzón en [27] de algunos de los resultados ya conocidos de la geometría hermitica, para el caso de f -variedades bandera; es decir, una variedad bandera dotada de un endomorfismo \mathcal{F} que satisface $\mathcal{F}^3 + \mathcal{F} = 0$.

Mostramos además que en el caso de la variedad bandera maximal clásica, es decir, $\mathbb{F}(n) = U(n)/(U(1) \times \dots \times U(1))$, donde $U(n)$ es el grupo unitario de matrices, existe una correspondencia biunívoca entre una f -estructura y un digrafo (grafo dirigido no necesariamente completo). Basados en esta representación extendemos la noción de torneo localmente transitivo, dada en [5], a digrafos no completos, y mostramos que en este caso esta condición extendida es necesaria y suficiente para que la variedad bandera maximal clásica sea (1,2)-simplética.

Con esto comenzamos el estudio orientado hacia encontrar una caracterización de las propiedades geométricas de este tipo de variedades, inmersa dentro de la teoría hermitica generalizada.

2. Nociones básicas

Definición 2.1. Un grupo de Lie G es un grupo cuya estructura subyacente tiene una estructura de variedad diferenciable, de tal forma que la aplicación producto

$$p : (g, h) \in G \times G \rightarrow gh \in G$$

es diferenciable.

Ejemplo 2.2.

1. $GL(n, \mathbb{R})$, el grupo de las transformaciones lineales invertibles de \mathbb{R}^n , o en otras palabras el grupo de las matrices invertibles $n \times n$. Este grupo es un subespacio abierto del espacio vectorial $M_n(\mathbb{R})$, y por tanto es una variedad diferenciable. El producto es el producto usual de matrices.

2. Cualquier espacio vectorial de dimension finita con la operación $+$ es un grupo de Lie.

3. El grupo G de las matrices triangulares superiores

$n \times n$ con entradas diagonales iguales a 1:

$$G = \left\{ \begin{pmatrix} 1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \right\},$$

es un grupo de Lie.

4. Grupo unitario de matrices $U(n) = \{X \in GL(n, \mathbb{C}) : X^{-1} = \bar{X}^t\}$, donde X^t es la traspuesta de X .

5. Grupo especial lineal $Sl(n, \mathbb{C}) = \{X \in GL(n, \mathbb{C}) : \det(X) = 1\}$.

6. Grupo ortogonal complejo $O(n, \mathbb{C}) = \{X \in GL(n, \mathbb{C}) : X^{-1} = X^t\}$

Definición 2.3. Una álgebra de Lie consiste de un espacio vectorial \mathfrak{g} dotado de un producto corchete o conmutador $[\cdot, \cdot]$, el cual es bilineal, antisimétrico y satisface la identidad de Jacobi, esto es, para cualesquiera $X, Y, Z \in \mathfrak{g}$ tenemos

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0.$$

Definición 2.4. Sea \mathfrak{g} una álgebra de Lie. Una subálgebra de Lie es un subespacio vectorial \mathfrak{h} de \mathfrak{g} que es cerrado por el corchete; esto es, $[X, Y] \in \mathfrak{h}$ si $X, Y \in \mathfrak{h}$.

Ejemplo 2.5. $\mathfrak{gl}(n, \mathbb{C})$, el espacio de todas las transformaciones lineales de un espacio vectorial de dimensión n en \mathbb{C} , o sea el álgebra de las matrices de tamaño $n \times n$.

El corchete está dado por

$$[X, Y] = XY - YX, \quad X, Y \text{ matrices.}$$

Algunas subálgebras de Lie de $\mathfrak{gl}(n, \mathbb{C})$ son:

1. $\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : X + \bar{X}^t = 0\}$.
2. $\mathfrak{so}(n, \mathbb{C}) = \{X \in \mathfrak{sl}(n, \mathbb{C}) : X + X^t = 0\}$.
3. $\mathfrak{sl}(n, \mathbb{C}) = \{X \in \mathfrak{gl}(n, \mathbb{C}) : \text{tr} X = 0\}$.
4. $\mathfrak{sp}(n, \mathbb{C}) = \{X \in \mathfrak{sl}(2n, \mathbb{C}) : XJ + JX^t = 0\}$, donde J , escrito en bloques de tamaño $n \times n$, es:

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

Definición 2.6. Una representación de un grupo de Lie G en un espacio vectorial V es una acción ρ de G en V tal que todas las transformaciones $\rho(g)$ son aplicaciones lineales de V .

Definición 2.7. Sean \mathfrak{g} una álgebra de Lie, V un espacio vectorial y $\mathfrak{gl}(V)$ la álgebra de Lie de las transformaciones lineales de V . Una aplicación lineal $\rho : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ tal que $\rho([X, Y]) = [\rho(X), \rho(Y)]$, es decir, ρ es un homomorfismo entre álgebras de Lie, es llamada una *representación* de \mathfrak{g} en V .

Existe una representación natural de un grupo de Lie G en su álgebra de Lie. Esta representación es construida de la siguiente forma: un elemento $g \in G$ define el automorfismo interno $C_g(x) = gxg^{-1}$. Es claro que $C_g(1) = 1$, por tanto la derivada en la identidad, $d(C_g)_1$, es una aplicación lineal de \mathfrak{g} en \mathfrak{g} .

Definición 2.8. La representación adjunta $Ad : G \rightarrow Gl(\mathfrak{g})$ de G en su álgebra de Lie \mathfrak{g} es definida por $Ad(g) = d(C_g)_1$.

La representación infinitesimal de Ad es la representación adjunta de \mathfrak{g} .

Definición 2.9. Sea \mathfrak{g} una álgebra de Lie. Su representación adjunta, es la aplicación $ad : \mathfrak{g} \rightarrow gl(\mathfrak{g})$ definida por

$$ad(Y)(X) = [X, Y].$$

La identidad de Jacobi garantiza que la aplicación ad es un homomorfismo de álgebras de Lie, donde el corchete de $gl(\mathfrak{g})$ es dado por el conmutador.

Definición 2.10. Dada una representación ρ de dimensión finita de la álgebra de Lie \mathfrak{g} , definimos en \mathfrak{g} una forma bilineal simétrica llamada forma traza, que está dada por

$$\beta_\rho(X, Y) = tr(\rho(X)\rho(Y)).$$

En el caso de que la representación sea la *adjunta*, esta forma es llamada *Cartan–Killing*, y será denotada de manera más simple por $\langle \cdot, \cdot \rangle$.

Definición 2.11. Una álgebra de Lie \mathfrak{g} es llamada semisimple si, y solamente si, su forma Cartan–Killing es no degenerada.

3. Variedad bandera clásica

Una variedad bandera es un espacio homogéneo reductivo $G/C(S)$, donde G es un grupo de Lie complejo semisimple y $C(S)$ es el centralizador de un toro S (no necesariamente maximal en G). Cuando S es un toro maximal decimos que la variedad bandera es maximal, y la denotamos por \mathbb{F} . Por ejemplo, en el caso clásico, G es el grupo especial unitario $SU(n)$, y $C(S)$ debe ser conjugado a un subgrupo de la forma $SU(n_1) \times \cdots \times SU(n_k)$, con n_1, \dots, n_k enteros positivos tales que $n_1 + \cdots + n_k = n$. Si $m_i = n_1 + \cdots + n_i$, el cociente $SU(n)/(SU(n_1) \times \cdots \times SU(n_k))$ puede ser identificado con el conjunto $F(m_1, \dots, m_k)$ de “banderas parciales” $\{0\} = E_0 \subset E_{m_1} \subset \cdots \subset E_{m_{k-1}} \subset E_{m_k} = \mathbb{C}^n$, donde E_i es un subespacio de \mathbb{C}^n de dimensión i . El caso $n_r = 1$, para todo $1 \leq r \leq k$, será denotado por $\mathbb{F}(n)$, y

esta puede ser identificada con el conjunto de “banderas totales” $\{0\} = E_0 \subset E_1 \subset \cdots \subset E_{n-1} \subset E_n = \mathbb{C}^n$.

4. Variedad bandera general

Sean \mathfrak{g} una álgebra de Lie semisimple sobre \mathbb{C} y $\mathfrak{h} \subseteq \mathfrak{g}$ una subálgebra de Cartan de \mathfrak{g} , es decir, una subálgebra abeliana maximal tal que para cada $H \in \mathfrak{h}$, $ad(H)(\mathfrak{g}) \subset \mathfrak{g}$ es semisimple. Sea α un funcional lineal sobre el espacio vectorial complejo \mathfrak{h} , y denótese por \mathfrak{g}_α el subespacio lineal de \mathfrak{g} dado por

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} : [H, X] = \alpha(H)X, \text{ para todo } H \in \mathfrak{h}\}.$$

Notése que para $\alpha = 0$, $\mathfrak{g}_\alpha = \mathfrak{h}$. Si $\alpha \neq 0$ y $\mathfrak{g}_\alpha \neq \{0\}$, el funcional lineal α es llamado una raíz (de \mathfrak{g} con respecto a \mathfrak{h}). En tal caso \mathfrak{g}_α es llamado un *subespacio raíz*. Denotamos por Π el conjunto de raíces no nulas del par $(\mathfrak{g}, \mathfrak{h})$ y por B la forma Cartan–Killing en $\mathfrak{g} \times \mathfrak{g}$, es decir, $B(X, Y) = \langle X, Y \rangle = tr(adXadY)$, para $X, Y \in \mathfrak{g}$. Dado que \mathfrak{g} es semisimple, B es no degenerada sobre $\mathfrak{g} \times \mathfrak{g}$, la restricción de B a $\mathfrak{h} \times \mathfrak{h}$ es también no degenerada, y así para cada $\alpha \in \Pi$ existe un único $H_\alpha \in \mathfrak{h}$ tal que $B(H, H_\alpha) = \langle H, H_\alpha \rangle = \alpha(H)$ para todo $H \in \mathfrak{h}$. Tómese $(\alpha, \beta) = B(H_\alpha, H_\beta)$ para toda $\alpha, \beta \in \mathfrak{h}^*$; se sigue entonces que (\cdot, \cdot) es una forma bilineal simétrica no degenerada sobre \mathfrak{h}^* .

Teorema 4.1. [35] Sean \mathfrak{g} una álgebra de Lie semisimple compleja y \mathfrak{h} una subálgebra de Cartan de \mathfrak{g} .

1. La álgebra de Lie \mathfrak{g} admite una descomposición en espacios de raíces:

$$\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha \in \Pi} \mathfrak{g}_\alpha.$$

2. Los espacios raíz \mathfrak{g}_α , $\alpha \in \Pi$, son de dimensión compleja uno.
3. Si α y β son dos raíces cualesquiera (incluyendo 0) y $\beta \neq -\alpha$, entonces \mathfrak{g}_α y \mathfrak{g}_β son ortogonales con relación a B .
4. Si α es una raíz no nula, $\Pi \cap \mathbb{Z}\{\alpha\} = \{\alpha, -\alpha\}$.
5. Para cada $\alpha \in \Pi$ existe un vector $X_\alpha \in \mathfrak{g}_\alpha$ de forma tal que para todo $\alpha, \beta \in \Pi$ tenemos:
 - (a) $[X_\alpha, X_{-\alpha}] = H_\alpha$, $[H, X_\alpha] = \alpha(H)X_\alpha$ (para todo $H \in \mathfrak{h}$);
 - (b) $[X_\alpha, X_\beta] = 0$ si $\alpha + \beta \neq 0$ y $\alpha + \beta \notin \Pi$;
 - (c) $\langle X_\alpha, X_\beta \rangle = 1$ si $\alpha + \beta = 0$ y $\langle X_\alpha, X_\beta \rangle = 0$ en los otros casos.

(d) $[X_\alpha, X_\beta] = m_{\alpha,\beta}X_{\alpha+\beta}$, si $\alpha + \beta \in \Pi$ con $m_{\alpha,\beta} \in \mathbb{R}$, y

$$\begin{aligned} m_{-\alpha,-\beta} &= m_{\alpha,\beta} \\ m_{-\alpha,\alpha+\beta} &= m_{\alpha+\beta,-\beta} = m_{-\beta,-\alpha}. \end{aligned} \quad (1)$$

Los elementos $\{X_\alpha : \alpha \in \Pi\}$ que satisfacen el numeral 5 en el anterior teorema serán llamados una *base de Weyl o Cartan-Weyl* de \mathfrak{g} módulo \mathfrak{h} .

Las variedades bandera poseen una descripción más general en términos de estos sistemas de raíces en álgebras de Lie, de la siguiente forma:

Sea \mathfrak{g} una álgebra de Lie semisimple compleja y \mathfrak{h} una subálgebra de Cartan de \mathfrak{g} ; denótese por Π el conjunto de raíces del par $(\mathfrak{g}, \mathfrak{h})$. Tomamos una base de Weyl de \mathfrak{g} fija. Sea $\Pi^+ \subset \Pi$ una escogencia de raíces positivas. Denótese por Σ el sistema simple de raíces correspondiente. Tome Θ un subconjunto de raíces de Σ y $\langle \Theta \rangle$ el conjunto de raíces generado por Θ . Sobre \mathfrak{g} tenemos la siguiente descomposición:

$$\begin{aligned} \mathfrak{g} &= \mathfrak{h} \oplus \sum_{\alpha \in \langle \Theta \rangle^+} \mathfrak{g}_\alpha \\ &\oplus \sum_{\alpha \in \langle \Theta \rangle^+} \mathfrak{g}_{-\alpha} \oplus \sum_{\beta \in \Pi^+ \setminus \langle \Theta \rangle^+} \mathfrak{g}_\beta \oplus \sum_{\beta \in \Pi^+ \setminus \langle \Theta \rangle^+} \mathfrak{g}_{-\beta}, \end{aligned} \quad (2)$$

donde \mathfrak{g}_α , $\alpha \in \Pi$, es el espacio raíz complejo correspondiente a α .

Sea ahora \mathfrak{p}_Θ la subálgebra parabólica de \mathfrak{g} determinada por Θ . Entonces,

$$\mathfrak{p}_\Theta = \mathfrak{h} \oplus \sum_{\alpha \in \langle \Theta \rangle^+} \mathfrak{g}_\alpha \oplus \sum_{\alpha \in \langle \Theta \rangle^+} \mathfrak{g}_{-\alpha} \oplus \sum_{\beta \in \Pi^+ \setminus \langle \Theta \rangle^+} \mathfrak{g}_\beta. \quad (3)$$

Así, la ecuación (2) puede ser reescrita como

$$\mathfrak{g} = \mathfrak{p}_\Theta \oplus \sum_{\beta \in \Pi^+ \setminus \langle \Theta \rangle^+} \mathfrak{g}_{-\beta}. \quad (4)$$

La variedad bandera general \mathbb{F}_Θ asociada al par $\{\mathfrak{g}, \Theta\}$ corresponde al espacio homogéneo $\mathbb{F}_\Theta = G/P_\Theta$, donde G es el grupo de Lie complejo cuya álgebra de Lie es \mathfrak{g} y P_Θ es el normalizador de \mathfrak{p}_Θ en G .

Denotamos por \mathfrak{u} la forma real compacta de \mathfrak{g} , es decir, la forma de Cartan–Killing restringida a \mathfrak{u} es negativa definida. Sea $U \subset G$ el subgrupo conexo asociado a \mathfrak{u} . Sea $K_\Theta = P_\Theta \cap U$, el cual por construcción, es el centralizador de un toro. U actúa transitivamente sobre \mathbb{F}_Θ , y así podemos escribir $\mathbb{F}_\Theta = U/K_\Theta$. Si $\Theta = \emptyset$, $\mathbb{F}_\Theta = \mathbb{F}$ corresponde a la variedad bandera maximal, y en el caso contrario corresponde a una variedad bandera parcial.

De lo anterior, tenemos 9 diferentes variedades bandera generales, asociadas a las diferentes álgebras de Lie semisimples, finitas, existentes:

- Álgebras clásicas:
 1. $A_l, l \geq 1$; corresponde a $\mathfrak{sl}(l+1)$.
 2. $B_l, l \geq 2$; corresponde al álgebra de matrices antisimétricas de dimensión impar $\mathfrak{so}(2l+1)$.
 3. $C_l, l \geq 3$; corresponde al álgebra simpléctica $\mathfrak{sp}(l)$.
 4. $D_l, l \geq 4$; corresponde al álgebra de matrices antisimétricas de dimensión par $\mathfrak{so}(2l)$.
- Álgebras excepcionales:
 1. G_2 es la álgebra construida sobre el espacio vectorial $\mathfrak{sl}(3) \oplus V \oplus V^*$, donde $V = \mathbb{K}^3$ con \mathbb{K} un campo cualquiera.
 2. E_6, E_7, E_8 . La construcción de las álgebras E_6, E_7 puede ser realizada a partir de la construcción de E_8 , que es el álgebra construida sobre el espacio vectorial $\mathfrak{sl}(9) \oplus V \oplus V^*$, donde $V = \Lambda^3 \mathbb{K}^9$ con \mathbb{K} un campo cualquiera.

5. Métricas invariantes

Denótese por b_0 el origen de \mathbb{F}_Θ , visto como espacio homogéneo de U . Una métrica riemanniana U -invariante ds_Λ^2 en \mathbb{F}_Θ es completamente determinada por sus valores en el origen, esto es, por un producto interno (\cdot, \cdot) en $T_{b_0}(\mathbb{F}_\Theta)$.

Cualquier producto interno en $T_{b_0}(\mathbb{F}_\Theta)$, U -invariante tiene la forma $(X, Y)_\Lambda = -\langle \Lambda \circ X, Y \rangle$, con $\Lambda : T_{b_0}(\mathbb{F}_\Theta) \rightarrow T_{b_0}(\mathbb{F}_\Theta)$ positivamente definida con relación a la forma Cartan–Killing, y \circ es el producto de Hadamard o producto de matrices entrada por entrada. La invarianza de $(\cdot, \cdot)_\Lambda$ equivale a afirmar que, en la base de Weyl,

$$\Lambda(X_\alpha) = \lambda_\alpha X_\alpha, \quad (5)$$

con $\lambda_\alpha = \lambda_{-\alpha} > 0$.

6. El espacio tangente de una variedad bandera

Considérese la variedad bandera general $\mathbb{F}_\Theta = G/P_\Theta$. Denótese por U la forma real compacta de G correspondiente a \mathfrak{u} . Sea $K_\Theta = P_\Theta \cap U$. K_Θ , por construcción, el centralizador de un toro. Sea la subálgebra real $\mathfrak{t}_\Theta = \mathfrak{u} \cap \mathfrak{p}_\Theta$; denotaremos por $\mathfrak{t}_\Theta^{\mathbb{C}}$ su complejificación. Podemos escribir

$$\mathfrak{t}_\Theta^{\mathbb{C}} = \mathfrak{h} \oplus \sum_{\alpha \in \langle \Theta \rangle^+} \mathfrak{g}_\alpha \oplus \sum_{\alpha \in \langle \Theta \rangle^+} \mathfrak{g}_{-\alpha}. \quad (6)$$

Sea \mathfrak{u} una forma real compacta de \mathfrak{g} . Asumimos \mathfrak{u} como el subespacio real generado por $i\mathfrak{h}_{\mathbb{R}}$, y A_{α}, S_{α} , con $\alpha \in \Pi \setminus \Theta$, donde $A_{\alpha} = X_{\alpha} - X_{-\alpha}$ y $S_{\alpha} = i(X_{\alpha} + X_{-\alpha})$. La variedad bandera general $\mathbb{F}_{\Theta} = U/K_{\Theta}$ es un espacio homogéneo reductivo. De hecho, sea $\mathfrak{u}_{\beta} = \mathfrak{u} \cap (\mathfrak{g}_{\beta} \oplus \mathfrak{g}_{-\beta})$, $\beta \in \Pi \setminus \langle \Theta \rangle$, y $\mathfrak{q}_{\Theta} = \sum_{\beta \in \Pi \setminus \langle \Theta \rangle} \mathfrak{u}_{\beta}$.

Por tanto,

- (i) $\mathfrak{u} = \mathfrak{t}_{\Theta} \oplus \mathfrak{q}_{\Theta}$, $\mathfrak{t}_{\Theta} \cap \mathfrak{q}_{\Theta} = \emptyset$;
- (ii) $Ad(K_{\Theta})\mathfrak{q}_{\Theta} \subset \mathfrak{q}_{\Theta}$, lo que implica $[\mathfrak{t}_{\Theta}, \mathfrak{q}_{\Theta}] \subset \mathfrak{q}_{\Theta}$.

Identificamos $\mathfrak{q}_{\Theta} = T_{b_0}(\mathbb{F}_{\Theta})$. Esta identificación es dada por $X \in \mathfrak{q}_{\Theta} \rightarrow X_{b_0} \in T_{b_0}(\mathbb{F}_{\Theta})$, es decir, por evaluación de $X \in \mathfrak{q}_{\Theta}$ en b_0 como un campo vectorial sobre $T_{b_0}(\mathbb{F}_{\Theta})$ (véase [17], pág. 191).

El espacio tangente a \mathbb{F}_{Θ} en b_0 se identifica naturalmente con el subespacio $\mathfrak{q}_{\Theta} = \mathfrak{u} \ominus \mathfrak{t} = \sum_{\beta \in \Pi \setminus \langle \Theta \rangle} \mathfrak{u}_{\beta}$, generado por $A_{\alpha}, S_{\alpha}, \alpha \in \Pi \setminus \langle \Theta \rangle$. Análogamente, el espacio tangente complejificado de \mathbb{F}_{Θ} es identificado con $\mathfrak{q}^{\mathbb{C}} = \mathfrak{g} \ominus \mathfrak{h} = \bigoplus_{\alpha \in \Pi \setminus \langle \Theta \rangle} \mathfrak{g}_{\alpha}$. Por el item (ii) inmediatamente anterior, la acción adjunta de K_{Θ} deja \mathfrak{q}_{Θ} invariante y lo descompone en componentes irreducibles que son invariantes por la acción adjunta de K_{Θ} (véase [34]).

7. f -Estructuras invariantes

K. Yano en 1961 introdujo la noción de f -estructura para cualquier variedad riemanniana de la siguiente forma.

Definición 7.1. Sea M una variedad riemanniana n -dimensional. Un campo tensorial \mathcal{F} en TM de tipo (1,1) tal que $\mathcal{F}^3 + \mathcal{F} = 0$, es llamado una f -estructura sobre TM .

Una f -estructura definida en una variedad bandera general \mathbb{F}_{Θ} será denotada por \mathcal{F}^{Θ} . Introducimos ahora la noción de “ f -estructura U -invariante” en \mathbb{F}_{Θ} .

Definición 7.2. Una f -estructura \mathcal{F}^{Θ} es llamada U -invariante si para cada $x \in \mathbb{F}_{\Theta}$ el endomorfismo $\mathcal{F}_x^{\Theta} : T_x\mathbb{F}_{\Theta} \rightarrow T_x\mathbb{F}_{\Theta}$ satisface $du_x \circ \mathcal{F}_x^{\Theta} = \mathcal{F}_{u_x}^{\Theta} \circ du_x$, para todo $u \in U$.

Claramente, \mathcal{F}^{Θ} es diagonalizable en $\mathfrak{q}_{\Theta}^{\mathbb{C}}$ con valores propios $i, 0, -i$.

La U -invarianza de \mathcal{F}^{Θ} garantiza que $\mathcal{F}^{\Theta}(\mathfrak{g}_{\alpha}) \subseteq \mathfrak{g}_{\alpha}$ para todo $\alpha \in \Pi \setminus \langle \Theta \rangle$, con igualdad cuando \mathcal{F}^{Θ} es una estructura cuasicompleja invariante (véase [34]).

Así, \mathcal{F}^{Θ} es determinado de manera única por los valores $\varepsilon_{\alpha}^{\Theta} \in \{0, \pm 1\}$, $\alpha \in \Pi \setminus \langle \Theta \rangle$, definidos por $\mathcal{F}^{\Theta}(X_{\alpha}) = i\varepsilon_{\alpha}^{\Theta}X_{\alpha}$. Estos valores satisfacen $\varepsilon_{-\alpha} = -\varepsilon_{\alpha}$, por tanto \mathcal{F}^{Θ} es definida por sus valores en $\Pi^+ \setminus \langle \Theta \rangle^+$.

Es claro que una estructura cuasicompleja J , es decir, una estructura que satisface $J^2 = -1$, es una f -estructura, por lo cual consideramos un problema interesante extender los resultados obtenidos, hasta el momento en [20], [25], [8], [9], para variedades bandera maximales clásicas cuasihermíticas invariantes (1,2)-simpléticas en su relación con grafos dirigidos completos (torneos) y en [32] para variedades bandera maximales asociadas a cualesquiera de las álgebras semisimples finitas mencionadas anteriormente.

Con relación a la variedad bandera maximal clásica $\mathbb{F}(n) = U(n)/T$, y a estructuras cuasicomplejas, en 1999 X. Mo and C. J. C. Negreiros mostraron en [20] que una condición necesaria para que esta variedad fuera (1,2)-simplética consistía en analizar el torneo asociado a la estructura cuasicompleja y garantizar que no contenía un 4-subtorneo conado, o sea un 4-subtorneo isomorfo a uno de los dos 4-torneos que aparecen en la Figura 1. Si el torneo no contiene este tipo de subtorneo, es llamado libre de cono.

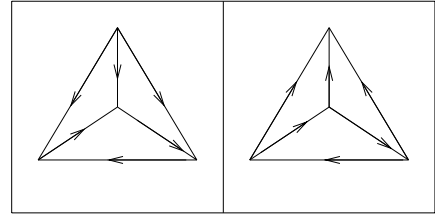


Figura 1: 4-Torneos conados

En 2000 M. Paredes [25] conjetura la suficiencia de la condición libre de cono, basado en los estudios realizados sobre $\mathbb{F}(3), \mathbb{F}(4), \mathbb{F}(5), \mathbb{F}(6), \mathbb{F}(7)$. Con base en esto, Cohen, Negreiros, San Martin [8] completan la demostración del teorema:

Teorema 7.3. $(\mathbb{F}(n), J)$ admite una métrica invariante (1,2)-simplética Λ si, y solamente si, el torneo asociado T_J es libre de cono.

Decimos que la estructura cuasicompleja J admite métrica (1,2)-simplética o que J es (1,2)-admisibile.

En 2002 L. A. B. San Martin and C. J. C. Negreiros, [32] utilizando la combinatoria geométrica de los sistemas de raíces y sus grupos de Weyl, obtienen, sobre una variedad bandera maximal asociada a cualquiera de las álgebras de Lie semisimples finitas ya mencionadas,

una caracterización de las estructuras cusicomplejas invariantes que admiten métricas que convierten la variedad en (1,2)-simpléctica. La metodología utilizada consistió en partir desde el correspondiente grupo de Weyl afín y el conjunto de alcobas (ver [33]), determinados por el sistema de raíces asociado. San Martín y Negreiros definieron una nueva clase de estructuras cusicomplejas llamadas de tipo afín, es decir, a cada alcoba asociaron una determinada estructura cusicompleja, mostrando que, sobre una variedad bandera maximal general, una estructura cusicompleja invariante es (1,2)-admisibles si y solamente si es de tipo afín. En este mismo artículo los autores observaron que la clase de las estructuras cusicomplejas invariantes (1,2)-admisibles es la principal entre las 16 clases dadas por Gray y Hervella [12], y muestran que, en el caso de la variedad bandera maximal general, las 16 clases se reducen a cuatro clases.

En el caso de la variedad bandera maximal para una álgebra de Lie semisimple compleja general no se puede hacer uso de la combinatoria dada por los grafos, pero sí se puede utilizar una combinatoria más general, basada en sistema de raíces. En [9] Cohen, Negreiros y San Martín dieron una interpretación, en términos de sistemas de raíces, de la condición *libre de cono*, y muestran la necesidad de esta condición para que una variedad bandera maximal general, sea (1,2)-simpléctica. Además mostraron que esta condición es también suficiente en el caso de las banderas asociadas a álgebras de Lie semisimples que no contienen componentes de tipo $B_l, l \leq 3, G_2$ o F_4 ; esta afirmación fue corregida en [27], mostrando que la condición libre de cono, dada en términos de raíces, es suficiente para cualquier bandera maximal asociada a una álgebra de Lie semisimple finita.

En 2003 Silva en [34] estudió las variedades bandera generales parciales y obtuvo condiciones geométricas para la clase de las estructuras cusicomplejas hermíticas que convierten la variedad bandera parcial en (1,2)-simpléctica. En este caso se muestra que las dieciséis 16 clases dadas por Gray y Hervella en [12] se reducen a cinco 5 clases.

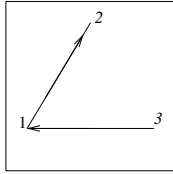


Figura 2: Grafo del Ejemplo 1.1.

Otra motivación para intentar extender los resultados ya mencionados fue la extensión, por parte de Rawnsley [29], de la noción de variedad cuasihermítica (1,2)-simpléctica a f -variedad (1,2)-simpléctica, y por esto la relación entre este tipo de variedades y la existencia de aplicaciones armónicas mediante aplicaciones holomorfas (véase [2]).

8. f -variedad (1,2)-simpléctica y aplicaciones armónicas

Considérese ahora (N, h, \mathcal{F}) una f -variedad riemanniana. Recordemos que (véase por ejemplo [17]) para $X, Y \in TN$, la derivada covariante de \mathcal{F} está dada por:

$$(\nabla \mathcal{F})(X, Y) = \nabla_X \mathcal{F}(Y) - \nabla_Y \mathcal{F}(X) - \mathcal{F}([X, Y]). \quad (7)$$

Definición 8.1. Definimos $(\nabla \mathcal{F})^{(1,1)}(X, Y)$ como $\nabla \mathcal{F}(X, Y)$ si $X \in TN^+$ y $Y \in TN^-$. (N, h, \mathcal{F}) es llamada f -(1,2)-simpléctica si $(\nabla \mathcal{F})^{(1,1)} = 0$.

Observamos que cuando \mathcal{F} es una estructura cusicompleja, la forma de Kähler $\sigma(X, Y) = h(X, \mathcal{F}(Y))$ es antisimétrica; Rawnsley y Salamon ([29], [30]) mostraron que $(\nabla \mathcal{F})^{(1,1)} = 0$ precisamente cuando (N, h, \mathcal{F}) es (1,2)-simpléctica, es decir, $d\sigma^{1,2} = 0$.

Sea (M, g, J) una variedad riemanniana con J una estructura cusicompleja y forma de Kähler ω . El siguiente teorema debido a Black [2], el cual tiene origen en el artículo [18], es una de las fuertes motivaciones de nuestro trabajo.

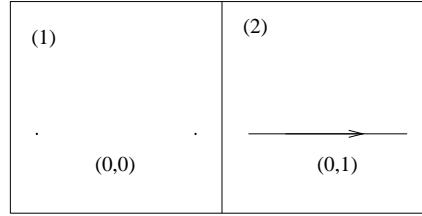
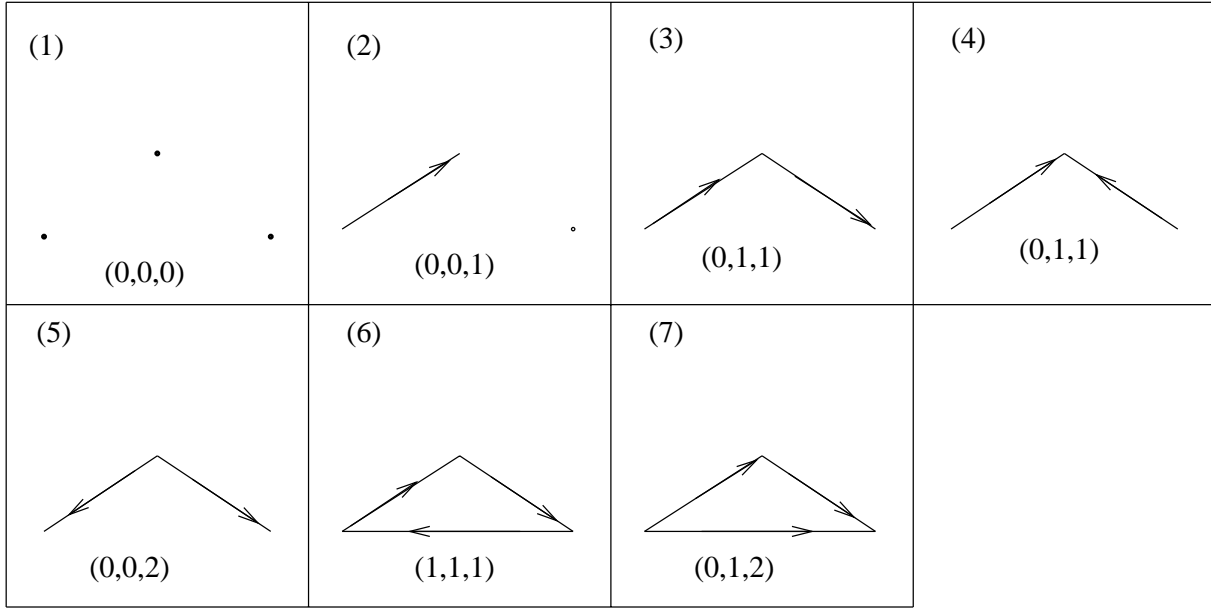
Teorema 8.2. Sea $\phi : (M, g, \mathcal{F}) \rightarrow (N, h, \mathcal{F})$ una aplicación tal que

1. ϕ es f -holomorfa, es decir, satisface $d\phi \circ J = \mathcal{F} \circ d\phi$,
2. M es co-simpléctica, es decir, $d^*(\omega) = 0$,
3. $(\nabla \mathcal{F})^{(1,1)} = 0$.

Entonces ϕ es armónica.

Este teorema en el estudio de aplicaciones armónicas en f -variedades bandera induce un análisis intensivo de las propiedades de las f -estructuras que admiten métricas (1, 2)-simpléticas, o sea, las f -estructuras (1,2)-admisibles.

Presentamos a seguir algunos de los resultados que obtuvimos al estudiar las f -variedades bandera maximales clásicas.

Figura 3: Clases de isomorfismos para $n = 2$.Figura 4: Clases de isomorfismos para $n = 3$.

9. El caso $(\mathbb{F}(n), \Lambda, \mathcal{F})$ (1,2)-simpléctica

Recordemos que $\mathbb{F}(n)$ es la variedad bandera maximal clásica, o sea, ella está asociada al álgebra de Lie, $\mathfrak{sl}(n, \mathbb{C})$, es decir el caso de la bandera maximal asociada a las álgebras de tipo A_n .

En este caso el conjunto formado por $E_{jk}, j \neq k$ y $E_{jj} - E_{kk}, j < k$, es una base de Weyl para $\mathfrak{sl}(n, \mathbb{C})$, \mathfrak{h} es la subálgebra de las matrices diagonales, $\Theta = \emptyset$, así \mathfrak{q} es generada por las matrices $A_{jk} = E_{jk} - E_{kj}$ y $S_{jk} = i(E_{jk} + E_{kj})$. Donde

$$E_{jk} = \begin{pmatrix} & & & & k & & \\ & 0 & 0 & \cdots & 0 & \cdots & 0 \\ & 0 & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ j & 0 & 0 & \cdots & 1 & \cdots & 0 \\ & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ & 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

Para $\mathbb{F}(n)$ y \mathcal{F} una estructura cuasicompleja, es posible establecer una correspondencia entre \mathcal{F} y un grafo orientado completo (torneo) (véase [20], [25], [8]).

Mo y Negreiros probaron en [21] que existen $3^{\binom{n}{2}}$ f -estructuras $U(n)$ -invariantes sobre una variedad bandera compleja, y que ellas están en correspondencia 1:1 con las matrices antisimétricas $\epsilon(\mathcal{F}) = \{\epsilon_{ij}\}$, donde ϵ_{ij} toma valores en el conjunto $\{-1, 0, 1\}$, lo que sugiere una correspondencia 1:1 entre f -estructuras invariantes y grafos orientados (digrafos) no necesariamente completos. La matriz $\epsilon(\mathcal{F})$ puede ser identificada con la matriz de incidencia de algún grafo orientado finito no completo $\mathcal{G} = (V, E)$, con $V = \{1, \dots, n\}$ de la siguiente forma: si $\mathcal{F}(E_{jk}) = i\epsilon_{jk}E_{jk}$, entonces $\mathcal{G}(\mathcal{F})$ es tal que para $j < k$,

$$\begin{aligned} j \rightarrow k &\iff \epsilon_{jk} = 1, \\ j \leftarrow k &\iff \epsilon_{jk} = -1 \\ j \not\leftrightarrow k &\iff \epsilon_{jk} = 0. \end{aligned}$$

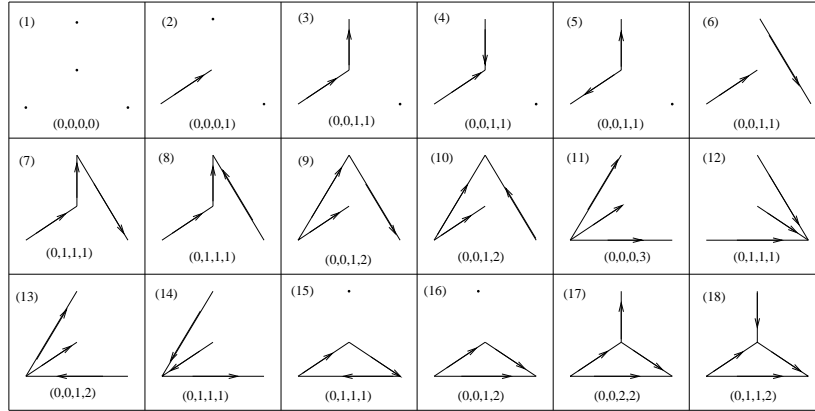


Figura 5: Clases de isomorfismos de digrafos con 4 vértices.

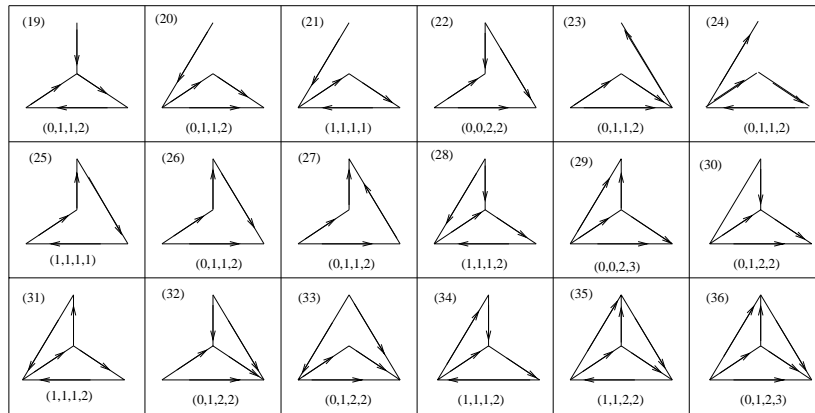


Figura 6: Clases de isomorfismos de digrafos con 4 vértices.

Una construcción similar es válida para las métricas $\Lambda = \{\lambda_{jk}\}$; identificamos cada λ_{jk} con un peso positivo sobre el arco jk de E .

Ejemplo 9.1. Consideremos

$$\mathbb{F}(3) = U(3)/(U(1) \times U(1) \times U(1)) = U(3)/T.$$

En este caso,

$$\mathfrak{q} = T(\mathbb{F}(3))_{(b_0)} = \left\{ \begin{pmatrix} 0 & a & b \\ -\bar{a} & 0 & c \\ -\bar{b} & -\bar{c} & 0 \end{pmatrix} : a, b, c, \in \mathbb{C} \right\}.$$

Tomemos la siguiente f -estructura sobre $\mathbb{F}(3)$:

$$\begin{pmatrix} 0 & a & b \\ -\bar{a} & 0 & c \\ -\bar{b} & -\bar{c} & 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 & ia & -ib \\ i\bar{a} & 0 & 0 \\ -i\bar{b} & 0 & 0 \end{pmatrix}.$$

El digrafo de la Figura 2 es el digrafo asociado a esta f -estructura.

Definición 9.2. Sean \mathcal{G}_1 un digrafo con n vértices $\{1, \dots, n\}$ y \mathcal{G}_2 un digrafo con m vértices $\{1, \dots, m\}$.

Um homomorfismo entre \mathcal{G}_1 y \mathcal{G}_2 es una aplicación $\phi : \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ tal que

$$s \xrightarrow{\mathcal{G}_1} t \implies \phi(s) \xrightarrow{\mathcal{G}_2} \phi(t) \quad \text{ó} \quad \phi(s) = \phi(t),$$

y

$$s \not\xrightarrow{\mathcal{G}_1} t \implies \phi(s) \not\xrightarrow{\mathcal{G}_2} \phi(t).$$

Cuando ϕ es bijectiva decimos que \mathcal{G}_1 y \mathcal{G}_2 son isomorfos.

Para $n = 2$ existen 2 clases de isomorfismos; para $n = 3$ aparecen 7 clases de isomorfismos, y para $n = 4$, 42 clases de isomorfismos (veáanse las Figuras 3, 4, 5-7).

Definición 9.3. Sea $\Gamma(N)$ el espacio de los campos vectoriales sobre una variedad diferenciable N .

- (i) Definimos $(\nabla\mathcal{F})^{(1,1)} : \Gamma^+(N) \times \Gamma^-(N) \rightarrow \Gamma(N)$.
- (ii) (N, h, \mathcal{F}) es llamada $(1, 2)$ -simpléctica si $(\nabla\mathcal{F})^{(1,1)} = 0$.

Nuestra intención es caracterizar las f -estructuras invariantes que admiten métricas (1,2)-simplécticas. En otras palabras, deseamos caracterizar los digrafos $\mathcal{G} = (V, E)$ que admiten pesos positivos $\Lambda = (\lambda_{jk})$, los cuales satisfacen las condiciones (8), (9) y (10). Para esto analizamos en $\mathbb{F}(2), \mathbb{F}(3), \mathbb{F}(4)$ las f -estructuras que admiten métricas (1,2)-simplécticas.

Las Figuras 3, 4, 5-7 muestran las clases de isomorfismos de f -estructuras invariantes en $\mathbb{F}(n)$ para $n = 2, 3, 4$, respectivamente. Para $n = 2, 3$ todas las clases son (1,2)-admisibles.

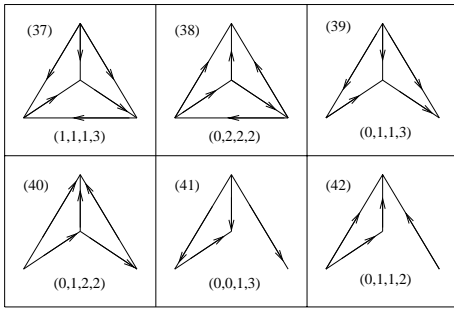


Figura 7: Clases de isomorfismos de digrafos con 4 vértices.

En $\mathbb{F}(4)$ existen 42 clases de isomorfismos de f -estructuras invariantes (ver Figuras , , y 7). Las 36 que aparecen en las figuras 5 y 6 son (1,2)-admisibles.

De acuerdo con Black [2], las condiciones para que $(\mathbb{F}(n), \Lambda, \mathcal{F})$ sea (1,2)-simpléctica son determinadas por

las siguientes reglas:

Si $k \rightarrow j, k \rightarrow l, j \not\leftrightarrow l$, entonces $\lambda_{jk} = \lambda_{kl}$; (8)

Si $j \rightarrow k, l \rightarrow k, j \not\leftrightarrow l$, entonces $\lambda_{jk} = \lambda_{kl}$; (9)

Si $k \rightarrow j, j \rightarrow l, k \rightarrow l$, entonces $\lambda_{kl} = \lambda_{jk} + \lambda_{jl}$.(10)

Estas reglas se resumen en la Figura 8.

A partir del comportamiento, en los casos mencionados, de las f -estructuras que admiten métrica (1,2)-simpléctica, vimos la necesidad de las siguientes definiciones:

Definición 9.4. Sea $\mathcal{G} := (V, E)$ un digrafo no completo.

1. Dados los subdigrafos de \mathcal{G} determinados por los subconjuntos

$$\mathcal{G}_p(v) = \{w \in V : vw \in E\},$$

$$\mathcal{G}_g(v) = \{w \in V : wv \in E\}.$$

Decimos que v es un *ganador* (respectivamente *perdedor*) en \mathcal{G} si $\mathcal{G}_p(v)$ (respectivamente $\mathcal{G}_g(v)$) es igual a $V \setminus \{v\}$. (Veáse Figura 9).

2. Decimos que \mathcal{G} es trivial si $|E'| = 0$.
3. \mathcal{G} es transitivo si la relación “ \rightarrow ” es transitiva (o sea, para $i, j, k \in V'$, $i \rightarrow j \rightarrow k$ implica $i \rightarrow k$);
4. \mathcal{G} es relativamente conexo si para todo $i, j, k \in V$, $i \rightarrow j$ implica $i \leftrightarrow k$ o $j \leftrightarrow k$.
5. \mathcal{G} es llamado localmente transitivo si cada uno de los subdigrafos $\mathcal{G}_p(v)$ y $\mathcal{G}_g(v)$ es transitivo y relativamente conexo.

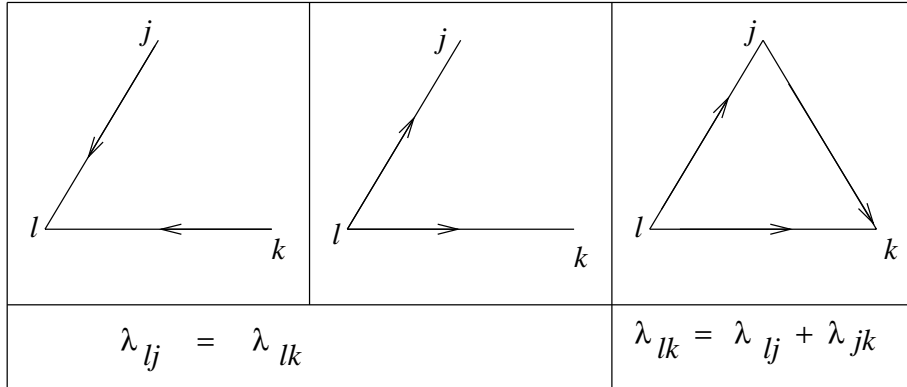


Figura 8: Condiciones de la métrica.

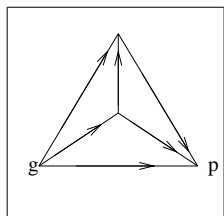


Figura 9: 4-torneo canónico.

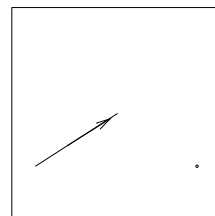


Figura 10: Digrafo no relativamente conexo.

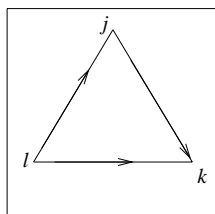


Figura 11: Triángulo transitivo.

Observación 9.5. De la Definición tenemos:

1. Transitividad local significa que el digrafo $\mathcal{G}_p(v), \mathcal{G}_g(v)$ omite determinados 4-subdigrafos, es decir, aquellos que poseen un subconjunto no vacío de flechas de un 3-ciclo (ver Figura 7).
2. En el caso no completo, la definición es menos fuerte que la definición de torneo localmente transitivo dado que esta definición permite que alguno de estos subdigrafos, por ejemplo $\mathcal{G}_g(v)$, pueda ser trivial, mientras que $\mathcal{G}_p(v)$ puede ser transitivo y relativamente conexo.
3. Si $\max\{|\mathcal{G}_g(v)|, |\mathcal{G}_p(v)|\} \leq 2$, para todo $v \in V$, entonces \mathcal{G} es localmente transitivo.

En $\mathbb{F}(3)$ se puede ver fácilmente que las 7 clases de isomorfismos de f -estructuras invariantes son todas localmente transitivas, verificando que los digrafos asociados a estas f -estructuras (ver Figura) son localmente transitivos.

En $\mathbb{F}(4)$, de las 42 clases de isomorfismos de f -estructuras invariantes apenas las 36 que aparecen en las Figuras 5 y 6 son localmente transitivas.

De acuerdo con la Figura 7, un 4-digrafo no localmente transitivo tiene un ganador o un perdedor, pero no ambos. Los arcos del 3-digrafo obtenido quitando el ganador/perdedor forman un subconjunto no vacío de los arcos de un 3-ciclo (ver Figura 7).

Lema 9.6. \mathcal{G} es localmente transitivo si, y solamente si, cada 4-subdigrafo de \mathcal{G} es localmente transitivo.

Proposición 9.7. En $\mathbb{F}(4)$ una f -estructura es localmente transitiva si y solamente si admite métrica (1,2)-simpléctica.

Definición 9.8.

- (i) Un triángulo transitivo es un digrafo completo transitivo $\mathcal{G}_t = (V_t, E_t)$ con $|V_t| = 3$. (Véase la Figura 11).
- (ii) Un digrafo $\mathcal{G} = (V, E)$ será llamado completamente no transitivo si \mathcal{G} no contiene triángulos transitivos; a la f -estructura asociada se la llamará de la misma forma. (Véase la Figura 12).

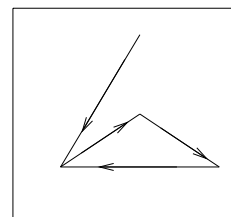


Figura 12: Digrafo completamente no transitivo.

Proposición 9.9. digrafo completamente no transitivo es localmente transitivo.

Demostración. En un digrafo completamente no transitivo, $\mathcal{G} = (V, E)$ y los conjuntos $\mathcal{G}_g(v)$ e $\mathcal{G}_p(v)$ son triviales en el sentido de la Definición . Al mismo tiempo, un digrafo completamente no transitivo admite métricas (1,2)-simplécticas, es decir, pesos positivos $\{\lambda_e > 0, e \in E\}$, los cuales satisfacen las identidades (8)-(10).

De hecho, debido a la ausencia de triángulos transitivos, el sistema de ecuaciones (8)-(10) no posee identidades del tipo (10); luego la métrica Cartan–Killing ($\lambda \equiv 1$), la cual satisface automáticamente (8-9), es (1,2)-simpléctica (en general, \mathcal{G} admite métricas (1,2)-simplécticas no-constantes junto con la métrica Cartan–Killing).

Observamos que la métrica Cartan–Killing sobre un digrafo \mathcal{G} es (1,2)-simpléctica si y solamente si \mathcal{G}' es completamente no transitivo. En particular, si \mathcal{G} es completo, la métrica Cartan–Killing es (1,2)-simpléctica si, y solamente si, $|V| \leq 3$, como fue observado en [7].

Teorema 9.10. *La f -estructura asociada al digrafo $\mathcal{G} = (V, E)$ admite métricas (1,2)-simplécticas si, y solamente si, \mathcal{G} es localmente transitivo.*

En la construcción de las métricas (1,2)-simplécticas sobre un digrafo localmente transitivo \mathcal{G} nuestra metodología fue la de reducir \mathcal{G} a un digrafo completamente no transitivo asociado \mathcal{G}' , con el mismo conjunto de vértices de \mathcal{G} .

Para finalizar, aplicando el Teorema 8.2 y la Proposición 9.9, tenemos:

Teorema 9.11. *Sea (M, g) una variedad riemanniana de dimensión par, con estructura cuasicompleja J , forma de Kähler Ω y $\phi : (M, g, J) \rightarrow (\mathbb{F}(n), \mathcal{F})$ una aplicación f -holomorfa, es decir, $(d\phi \circ J = \mathcal{F} \circ d\phi)$ con \mathcal{F} completamente no transitivo. Supóngase además M es co-simpléctica, es decir $d^*\Omega = 0$. Entonces ϕ es armónica con respecto a la métrica Cartan–Killing.*

Bibliografía

- [1] J. F. ADAMS, *Lectures on Lie Groups*, New York, 1969.
- [2] M. BLACK, *Harmonic Maps into Homogeneous Spaces*, Pitman Research Notes, Math series 255. Longman, Harlow, 1991.
- [3] A. BOREL, *Kählerian coset spaces of semi-simple Lie groups*, Proc. Nat. Acad. of Sci. **40**:1147-1151, 1954.
- [4] A. BOREL AND F. HIRZEBRUCH, *Characteristic classes and homogeneous spaces I*, Amer. J. Math. **80**:458-538, 1958.
- [5] A. E. BROUWER, *The enumeration of locally transitive tournaments*. Afdeling Zuivere Wiskunde [Department of Pure Mathematics], **138**. Mathematisch Centrum, Amsterdam, 1980.
- [6] F. E. BURSTALL AND S. SALAMON, *Tournaments, flags and harmonic maps*, Math Ann. **277**:249-265, 1987.
- [7] N. COHEN, C. J. C. NEGREIROS AND L. A. B. SAN MARTIN, *Characterization of (1,2)-symplectic metrics on flag manifolds*, Contemporary Math. **288**:300-304, 2002.
- [8] N. COHEN, C. J. C. NEGREIROS AND L. A. B. SAN MARTIN, *(1,2)-Symplectic metrics, flag manifolds and tournaments*, Bull. of London Math. Soc. **34**:641-649, 2002.
- [9] N. COHEN, C. J. C. NEGREIROS AND L. A. B. SAN MARTIN, *A rank-three condition for invariant (1,2)-symplectic almost Hermitian structures on flag manifolds*, Bull. Braz. Math. Soc. New Series **33**: 49-73, 2002.
- [10] J. EELLS AND S. SALAMON, *Twistorial constructions of harmonic maps of surfaces into four-manifolds*, Ann. Scuola Norm. Sup. Pisa (4) **12**:589-640, 1985.
- [11] J. EELLS AND L. LEMAIRE, *Another Report on Harmonic Maps*, Bull. of London Math. Soc. **20**:385-524, 1988.
- [12] A. GRAY AND L. M. HERVELLA, *The sixteen classes of almost hermitian manifolds and their linear invariants*, Ann. Mat. Pura Appl **123**:35-58, 1980.
- [13] S. HELGASON, *Differential Geometry, Lie Groups and Symmetric Spaces*, Acad. Press, 1978.
- [14] J. E. HUMPHREYS, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, 1980.
- [15] S. ISHIHARA AND K. YANO, *On Integrability Conditions of a structure f satisfying $\mathcal{F}^3 + \mathcal{F} = 0$* , Quart. J. Math. Oxford (2), **15**:217-222, 1964.
- [16] V. KAC, *Infinite dimensional Lie algebras*. Cambridge University Press, 1985.
- [17] S. KOBAYASHI AND K. NOMIZU, *Foundations of Differential Geometry*, Vol. 1, Interscience Publishers, 1963.
- [18] A. LICHTNEROWICZ, *Applications harmoniques et variétés Kähleriennes*, Symp. Math. **3** (Bologna), 341-402, 1970.
- [19] E. C. L. SANTOS, *Estruturas Quase-Hermitianas Invariantes e Ideais Abelianos*, Disertación de Maestría, Universidade Estadual de Campinas, 2003.
- [20] X. MO AND C. J. C. NEGREIROS, *Tournaments and Geometry of Full Flag Manifolds*, Proceedings of the XI Brazilian Topology Meeting, Rio Claro, Brazil, World Scientific, 1999.
- [21] X. MO AND C. J. C. NEGREIROS, *Horizontal f -estructuras, ϵ -matrices and Equiharmonic moving flags*, Relatório de Pesquisa, Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas, 1999.
- [22] X. MO AND C. J. C. NEGREIROS, *(1,2)-Symplectic structure on flag manifolds*, Tôhoku Math. J. **52**:271-282, 2000.
- [23] W. MOON, *Topics on Tournaments*, Holt, Reinhart and Winston, 1968.
- [24] C. J. C. NEGREIROS, *Some remarks about harmonic maps into flag manifolds*, Indiana Univ. Math. J. **37**:617-636, 1988.
- [25] M. PAREDES, *Aspectos da geometria complexa das variedades bandeira*, Tesis doctoral, Universidade Estadual de Campinas, Brasil, 2000.
- [26] M. PAREDES, *Families of (1,2)-symplectic metrics on flag manifolds*, Internat. J. Math. Math. Sci. **29**:651-664, 2002.
- [27] S. PINZÓN, *Variiedades Bandeira, f -Estruturas e Métricas (1,2)-Simplécticas*, Tesis doctoral, Universidade Estadual de Campinas, Brasil, 2003.
- [28] A. PRESSLEY AND G. SEGAL, *Loop Groups*, Clarendon Press, Oxford, 1986.
- [29] J. H. RAWNSLEY, *f -Structures, f -Twistor Spaces and Harmonic Maps*, Lec. Notes in Math. **1164**, 1986.
- [30] S. SALAMON, *Harmonic and holomorphic maps*, Lec. Notes in Math. **1164**, Springer 1986.
- [31] L. A. B. SAN MARTIN *Algebras de Lie*, Campinas-S.P., Editora da Unicamp, 1999.
- [32] L. A. B. SAN MARTIN AND C. J. C. NEGREIROS, *Invariant almost Hermitian structures on flag manifolds*, Adv. Math. **178**: 277-310, 2003.

- [33] J. Y. SHI, *Alcoves corresponding to an affine Weyl group*, J. London Math. Soc. **35**:42-45, 1987.
- [34] R. C. DE JESUS SILVA, *Estruturas Quase Hermitianas Invariantes em Espaços Homogêneos de Grupos Semi-simples*, Tesis doctoral, Universidade Estadual de Campinas, Brasil, 2003.
- [35] G. WARNER, *Harmonic Analysis on Semi-Simple Lie Groups, I.*, Springer-Verlag, 1972.
- [36] J. A. WOLF AND A. GRAY, *Homogeneous spaces defined by Lie group automorphisms, I.*, J. Diff. Geom. **2**:77-114, 1968.
- [37] J. A. WOLF AND A. GRAY, *Homogeneous spaces defined by Lie group automorphisms, II.*, J. Diff. Geom. **2**:115-159, 1968.
- [38] K. YANO, *On a structure defined by a tensor field of type (1,1) satisfying $\mathcal{F}^3 + \mathcal{F} = 0$* , Tensor **14**:99-109, 1963.

CONTRIBUCIÓN AL ESTUDIO DE LAS ECUACIONES DIFERENCIALES PARCIALES DE TIPO ELÍPTICO*

por

Jorge Cossio[†]

Resumen

Jorge Cossio: Contribución al estudio de las ecuaciones diferenciales parciales de tipo elíptico. Rev. Acad. Colomb. Cienc. **28** (106): 135–145, 2004. ISSN 0370-3908.

En este artículo se presentan los resultados más importantes de mi trabajo de investigación en el estudio de la existencia y de las propiedades de las soluciones de ecuaciones diferenciales parciales no lineales de la forma

$$\begin{cases} \Delta u + \lambda f(u) = 0 & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (1)$$

donde $\lambda \in \mathbf{R}$, Ω es un dominio acotado en \mathbf{R}^N con frontera suave, $\Delta = \sum_{i=0}^n \frac{\partial}{\partial x_i^2}$ es el operador de Laplace y $f : \mathbf{R} \rightarrow \mathbf{R}$ es una función no lineal. Se presentan teoremas obtenidos utilizando teoría de bifurcación, métodos variacionales y un principio de minimax desarrollado por el autor en colaboración con A. Castro y J. M. Neuberger ([Cas-Cos-Nu1], 1997). Además, se incluyen algunos algoritmos para construir y visualizar soluciones a problemas no lineales del tipo (1) y una serie de preguntas abiertas.

Palabras clave: Ecuaciones elípticas semilineales, teoría de bifurcación, métodos variacionales, construcción de soluciones.

Abstract

In this paper I present the most important results of my research studying the solutions of nonlinear partial differential equations of the type

*2000 Mathematics Subject Classification: 35B30, 35B32, 35J20, 35J25, 35J60, 35J65, 47H12, 47H15, 65D07, 65N99. Este trabajo ha sido apoyado parcialmente por Colciencias, Proyecto Código 1118-05-11412.

[†]Escuela de Matemáticas, Universidad Nacional de Colombia, Apartado Aéreo 3840, Medellín, Colombia. email: jcossio@unalmed.edu.co

$$\begin{cases} \Delta u + \lambda f(u) = 0 & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases} \quad (1)$$

where $\lambda \in \mathbf{R}$, Ω is a smooth bounded domain in \mathbf{R}^N , $\Delta = \sum_{i=0}^n \frac{\partial}{\partial x_i^2}$ is the Laplacian operator, and $f : \mathbf{R} \rightarrow \mathbf{R}$ is a nonlinear function. Our theorems were obtained by using bifurcation theory, variational methods, and a minmax principle developed by the author in collaboration with A. Castro and J. M. Neuberger ([Cas-Cos-Nu1], 1997). We also present some theorems related with algorithms for approximating solutions to nonlinear problems of type (1), and some open questions.

Key words: Semilinear elliptic equations, bifurcation theory, variational methods, construction of solutions.

1. Introducción

El interés de la comunidad científica en el análisis no lineal ha ido creciendo significativamente en las últimas décadas. Prácticamente todos los centros de investigación en Matemáticas en el mundo dedican un esfuerzo sustancial al entendimiento de los problemas no lineales.

El desarrollo del análisis no lineal ha permitido no solamente el tratamiento unificado de varios problemas clásicos en Física y en Matemáticas sino también el surgimiento de nuevas teorías no lineales de gran importancia por sí mismas y por sus aplicaciones.

Una de las herramientas más importantes usadas en análisis no lineal es la *teoría de bifurcación*. Bifurcación significa cambios en la estructura del conjunto de soluciones de una ecuación funcional cuando los parámetros que intervienen en la ecuación varían.

Antes de discutir algunos aspectos matemáticos de la teoría de bifurcación, describiremos uno de los ejemplos físicos más familiares de bifurcación: el Problema de Taylor de un fluido en rotación.

En éste consideramos un fluido viscoso incompresible que yace entre un par de cilindros concéntricos verticales que rotan. Suponemos que el cilindro exterior está en reposo y que el cilindro interior rota a una velocidad angular w . El movimiento del fluido se describe por las ecuaciones de Navier-Stokes y éstas tienen una solución explícita. El flujo así obtenido se denomina el flujo de Couette, que existe para todos los valores de w . Si w es pequeña, las partículas se mueven en orbitas circulares y la velocidad depende de la distancia al cilindro interior. Si se excede un cierto valor de w , nuevos movimientos se superponen sobre el flujo de Couette y aparecen los denominados vórtices de Taylor. Para este valor de w ocurre bifurcación.

Para más detalles, el lector interesado puede consultar [Ke-An], 1969.

La teoría de bifurcación, que será discutida en la Sección 2, estudia la existencia de soluciones locales y globales para ecuaciones no lineales de la forma

$$F(\lambda, u_\lambda) = 0, \quad (2)$$

donde $\lambda \in \mathbf{R}$, F es una función no lineal y u_λ pertenece a un cierto espacio de funciones. Los pilares fundamentales de la teoría de bifurcación aparecen inicialmente en los trabajos de Lyapunov y Schmidt, alrededor de 1908, basados en los problemas propuestos por Poincaré en 1885 en conexión con astrofísica. Los aportes de Krasnoselskii y su escuela, alrededor de 1950, permiten la consolidación de la teoría. Pero han sido los trabajos de M. Crandall y P. Rabinowitz ([Cr-Ra], 1971), en lo concerniente a la existencia de ramas locales de soluciones, y de P. Rabinowitz ([Ra1], 1971), en lo relacionado con la existencia de ramas globales de soluciones, los que han permitido su gran desarrollo. La teoría de bifurcación ha sido muy útil en el estudio de una amplia variedad de problemas no lineales tanto en ecuaciones diferenciales, entre los que destacamos varios problemas clásicos como el problema de Bénard, el problema de las ondas de agua y el problema de los tres cuerpos restringido, como en Física y en Ingeniería, en problemas de elasticidad, convección térmica y fluidos en rotación ([Ra2], 1985, y [Am-Pro], 1993).

En relación con la teoría de bifurcación, el autor y A. Castro encontraron, en un trabajo publicado en 1993 ([Cas-Cos1]), una condición suficiente que garantiza la existencia de ramas locales de soluciones para ecuaciones del tipo (1) y mostraron aplicaciones a la existencia de soluciones para sistemas de ecuaciones diferenciales y a la existencia de soluciones periódicas de ecuaciones elípticas no lineales de segundo orden. Posteriormente, el autor, en un trabajo publicado en la Revista de la Academia Colombiana de Ciencias Exactas Físicas y Naturales ([Cos], 1995), en el cual utiliza de manera esencial los resultados de la teoría de bifurcación debidos a Crandall y Rabinowitz ([Cr-Ra], 1971, y [Ra1], 1971),

demostró la existencia de múltiples soluciones para el problema no lineal

$$\begin{cases} \Delta u + f(u) = 0 & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (3)$$

con ciertas condiciones sobre la no linealidad f .

También, usando los resultados de Crandall y Rabinowitz, el autor y A. Castro, en un artículo publicado en 1993 en la Revista Colombiana de Matemáticas ([Cas-Cos2]), demostraron la existencia de múltiples soluciones radiales para el problema (3) en el caso en que el dominio Ω es simétrico.

Antes de describir otras herramientas importantes usadas en el estudio de problemas no lineales, quisiera mencionar que el problema de Dirichlet no lineal (3) modela una gran variedad de problemas que aparecen naturalmente en diferentes áreas de la física, la astrofísica, la ingeniería, la biología y la geometría. Menciono a continuación algunos modelos en astrofísica en los que aparecen distintos tipos de no linealidades. Por ejemplo, en la ecuación de Lane-Emden, que aparece en el estudio de problemas estelares ([Ch], 1939), la no linealidad es del tipo potencia,

$$\begin{cases} \Delta u + u^p = 0 & u > 0 \text{ en } \Omega, p > 1, \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (4)$$

donde u^p es proporcional a la densidad de la estrella gaseosa y el dominio Ω es $B_R(0)$, la bola abierta en \mathbf{R}^3 de radio R y centro en el origen. Henon en 1973 ([He]), estudiando las estructuras estelares en rotación, propuso la siguiente no linealidad, que es una variante del problema (4)

$$\begin{cases} \Delta u + |x|^l u^p = 0 & u > 0 \text{ en } \Omega, p > 1, l > 0, \\ u = 0 & \text{en } \partial\Omega. \end{cases} \quad (5)$$

Para la escogencia de esta no linealidad, Henon comentó: “Esta escogencia, aunque arbitraria, tiene la ventaja de ser simple y conveniente”. Posteriormente, Lieb y Yau en 1987 ([Li-Ya]), estudiando la teoría de colapsos estelares de Chandrasekhar, demostraron que la ecuación para el problema de las estrellas enanas blancas, sin efecto general relativístico, es equivalente a la ecuación

$$\Delta u + 4\pi(2u + u^2)^{\frac{3}{2}} = 0 \quad \text{en } B_R(0). \quad (6)$$

En los problemas (4), (5) y (6) aparece la pregunta de si existe o no solución u para la ecuación semilineal en estudio; y, de existir solución, si ésta es única o si existen múltiples soluciones.

En la Sección 3 de este trabajo presentamos otra herramienta muy importante usada en análisis no lineal:

los *métodos variacionales*. Las ideas topológicas iniciales de estos métodos fueron vislumbradas por Poincaré y Birkhoff a finales del siglo XIX y fueron desarrolladas posteriormente, en los años 20 y 30 del siglo pasado, por Morse ([Mors1], 1925, y [Mors2], 1934) y Ljusternik y Schnirelman ([Lj-Sc], 1934). Con esta técnica las soluciones de los problemas no lineales se consiguen como los puntos críticos de un funcional asociado con la ecuación en estudio. En el caso en que el funcional sea acotado inferior o superiormente es razonable tratar de demostrar que el funcional alcanza su mínimo o su máximo, pero si el funcional asociado no es acotado ni superior ni inferiormente se deben buscar puntos de silla o puntos de tipo minimax. Un método específico para encontrar puntos críticos de tipo minimax, que será presentado en la Sección 3, es el bien conocido *Teorema del Paso de la Montaña*, publicado por A. Ambrosetti y P. Rabinowitz en 1973 ([Am-Ra]), del cual se han conseguido extensiones y variaciones en años recientes que han demostrado ser muy útiles en las aplicaciones a ecuaciones diferenciales.

Otro resultado importante para encontrar puntos críticos, que se presentará en la Sección 3, es el *método de reducción de Lyapunov-Schmidt*, el cual tuvo sus orígenes en las investigaciones de los profesores Lazer, Landesman y Meyers ([Lan-Laz-Me], 1975) y Castro y Lazer ([Cas-Laz2], 1979) y que permite reducir el estudio de los puntos críticos del funcional asociado, el cual está generalmente definido en un espacio de dimensión infinita, al estudio de los puntos críticos de un funcional definido en un subespacio, generalmente, de dimensión finita. Además de los dos métodos mencionados, se presentará en la Sección 3 un *principio de minimax* desarrollado por el autor en colaboración con los investigadores A. Castro y J. M. Neuberger ([Cas-Cos-Nu1], 1997), que permite obtener *soluciones que cambian de signo* para problemas superlineales.

Utilizando el Teorema del Paso de la Montaña, el método de reducción de Lyapunov-Schmidt y el principio de minimax mencionado se presentarán al final de la Sección 3 distintas aplicaciones al estudio de la solución de ecuaciones diferenciales no lineales del tipo (3), las cuales han sido obtenidas por el autor en colaboración con otros investigadores en el área ([Cas-Cos3], 1994, y [Cas-Cos-Nu3], 1998).

En la Sección 4 enfocaremos nuestra atención en la presentación de algunos *algoritmos para construir y visualizar soluciones de problemas no lineales* del tipo (3). Estos algoritmos son importantes ya que los

teoremas que garantizan la existencia de soluciones de ecuaciones diferenciales no lineales son, en general, no constructivos y, además, ejemplos explícitos que apoyen o nieguen conjeturas y que sugieran direcciones para estudios futuros son siempre bienvenidos. Inicialmente se discute un método diseñado por el autor y el profesor H. Arango para construir soluciones radiales al problema (3) en el caso en el cual el dominio es una bola en \mathbf{R}^N y f es una función continua lineal por tramos ([Ar-Cos1], 1996, y [Ar-Cos2], 2000). El método está basado en un spline de funciones de Bessel. Concluiremos la sección con la presentación de un nuevo algoritmo para construir y visualizar las soluciones al problema (3) que se consiguen vía el método de reducción de Lyapunov-Schmidt ([Cos-Le-Nu], 2001).

Las ideas de investigación desarrolladas por el autor en colaboración con otros investigadores en el área, que serán presentadas en las distintas secciones de este trabajo, nos conducen a formular en la Sección 5 una serie de **preguntas abiertas**, que esperamos contribuyan a una mejor comprensión de las soluciones de problemas de Dirichlet no lineales.

2. Teoría de bifurcación y aplicaciones

El Teorema de Punto Fijo de Banach establece que en un espacio métrico completo (X, d) , una función $f : X \rightarrow X$ para la cual existe una constante $k \in (0, 1)$ tal que

$$d(f(x), f(y)) \leq kd(x, y), \quad \forall x, y \in X,$$

tiene un único punto fijo; es decir, existe un único punto $x_0 \in X$ tal que $f(x_0) = x_0$.

Una consecuencia importante de este resultado es el Teorema de la Función Inversa que establece que si una función F envía una vecindad U de $u_0 \in X$ en Y , donde X y Y son espacios de Banach, y $F \in C^1(U)$ con $L = F'(u_0)$ uno a uno y sobreyectiva entonces la ecuación

$$F(u) = f$$

tiene una solución única en una vecindad de u_0 , para toda f en una cierta vecindad de $f_0 = F(u_0)$.

Originado en los trabajos de A. M. Lyapunov y E. Schmidt a comienzos del siglo XX y relacionado con los problemas propuestos por H. Poincaré en 1885 en conexión con astrofísica, se inició el programa de extender los resultados anteriores cuando $F'(u_0)$ no es invertible. Este programa dio lugar a lo que hoy conocemos como *teoría de bifurcación*.

La situación típica en bifurcación está relacionada con una familia de un parámetro $F(\lambda, u)$, con $\lambda \in \mathbf{R}$ y $F \in C(\mathbf{R} \times X, Y)$, donde X y Y son espacios de Banach. Se supone que

$$F(\lambda, 0) = 0, \quad \forall \lambda \in \mathbf{R}. \quad (7)$$

Se dice que un punto $(\lambda, 0)$ es un punto de bifurcación de $F(\lambda, u) = 0$ si toda vecindad de él contiene soluciones no triviales de la ecuación.

Bajo hipótesis simples, M. Crandall y P. Rabinowitz demostraron el siguiente teorema que garantiza la existencia de una *rama local* de soluciones no triviales de la ecuación

$$F(\lambda, u) = 0. \quad (8)$$

Teorema 1. ([Cr-Ra], 1971) Sean X, Y espacios de Banach, Λ un conjunto abierto en \mathbf{R} y $F \in C^m(\Lambda \times X, Y)$, $m \geq 2$. Suponga que

$$\begin{aligned} F(\lambda, u) &= Bu - \lambda u + N(\lambda, u) \\ N(\lambda, 0) &= 0, \quad D_u N(\lambda, 0) = 0, \end{aligned} \quad (9)$$

donde $B : X \rightarrow Y$ es un operador lineal y acotado. Si λ_0 es un valor propio simple de B con vector propio $v_0 \neq 0$, entonces $(\lambda_0, 0)$ es un punto de bifurcación de $F(\lambda, u) = 0$. Además, existen funciones de clase C^{m-1}

$$\begin{aligned} \lambda^*(s) &= \lambda_0 + O(|s|) \\ u^*(s) &= sv_0 + O(s^2) \end{aligned} \quad (10)$$

para s cerca de cero, tales que

$$F(\lambda^*(s), u^*(s)) = 0. \quad (11)$$

Todos los ceros de F cerca de $(\lambda_0, 0)$ son las soluciones triviales $u = 0$ o están dados por (10).

El resultado más importante relacionado con la existencia de *ramas globales* de soluciones, es decir cuando λ varía sobre \mathbf{R} , fue obtenido por P. Rabinowitz aplicando la teoría de grado de Leray-Schauder. Este resultado, que se presenta a continuación, establece que bajo ciertas condiciones de compacidad cada rama de soluciones se extiende al infinito en $\mathbf{R} \times X$ o se va a otro punto de bifurcación.

Teorema 2. ([Ra2], 1971) Suponga que $F(\lambda, u) = u - G(\lambda, u)$ con $(\lambda, u) \in \mathbf{R} \times X$, X un espacio de Banach y G compacto. Si $G(\lambda, u) = \lambda Lu + H(\lambda, u)$, donde L es lineal y compacto, $H(\lambda, u) = o(\|u\|)$ cuando $u \rightarrow 0$ y $\mu \in r(L) := \{\mu \in \mathbf{R}/\mu^{-1} \in \sigma(L) \text{ -el espectro de } L\}$ es de multiplicidad impar entonces $(\mu, 0)$ es un punto de bifurcación. Además S , la clausura del conjunto de

ceros no triviales de F , tiene una componente de soluciones C que contiene a $(\mu, 0)$ y no es acotada en $\mathbf{R} \times X$ o contiene $(\hat{\mu}, 0)$, donde $\mu \neq \hat{\mu} \in r(L)$.

En colaboración con A. Castro ([Cas-Cos1], 1993), dimos una condición suficiente para que un punto (λ, u) fuera un punto de bifurcación local de la ecuación (8), cuando ésta toma la forma

$$F(\lambda, u) := Lu + \lambda N(u) = 0, \quad (12)$$

donde L es un operador lineal definido en un espacio de Hilbert H , L tiene inversa compacta y $N : H \rightarrow H$ es de clase C^2 con $(L^{-1}N)'(0) = L^{-1}$ y $N(0) = 0$.

Sea $-\Lambda \neq 0$ un valor propio de L tal que

$$H = \text{Ker}(L + \Lambda I) \oplus \text{Rango}(L + \Lambda I) =: X \oplus Y.$$

Sean $K \subset X$ un cono cerrado tal que $K \neq X$ y P la proyección sobre X . Se demostró el siguiente teorema.

Teorema 3. ([Cas-Cos1], 1993) *Si para algún $r_1 > 0$*

$$P(N - I)(\{u \in H : |u| \leq r_1\}) \subset K \quad (13)$$

entonces $(\Lambda, 0)$ es un punto de bifurcación de la ecuación (12). Además, para $r > 0$ suficientemente pequeño y $K \cap (-K) = \{0\}$ la ecuación (12) tiene dos soluciones diferentes (α, u) , (β, v) con $\|Pu\| = \|Pv\| = r$, $|\alpha - \Lambda| < r$ and $|\beta - \Lambda| < r$.

A diferencia de los resultados clásicos, nuestro resultado no depende ni de la multiplicidad del valor propio Λ ni de la estructura variacional de la ecuación donde el teorema se aplica. Este resultado tiene aplicaciones a la existencia de soluciones para sistemas de ecuaciones diferenciales y a la existencia de soluciones periódicas de ecuaciones elípticas no lineales de segundo orden.

Sea $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots$ la sucesión de valores propios de $-\Delta$ con condición de Dirichlet cero en la frontera.

En un trabajo publicado en la Revista de la Academia Colombiana de Ciencias Exactas Físicas y Naturales, en el cual se utilizan de manera esencial los resultados de M. Crandall y P. Rabinowitz (veáanse los Teoremas 1 y 2), estudié el problema (3) y demostré que tiene al menos cuatro soluciones cuando el rango de la derivada de la no linealidad incluye al menos los dos primeros valores propios.

Teorema 4. ([Cos], 1995) *Si $0 < f'(0) < \lambda_1$, $\lambda_2 < f'(\infty)$, λ_2 es un valor propio de multiplicidad impar y $f'(\infty) < \lambda_j$ (donde λ_j es el siguiente valor propio distinto de λ_2) entonces el problema (3) tiene al menos cuatro soluciones.*

A diferencia del Teorema A de [Cas-Cos3], publicado en 1994, este resultado no depende del acotamiento global de la derivada de la función no lineal y se demuestra utilizando solamente técnicas de teoría de bifurcación.

También, haciendo uso del teorema de bifurcación global de P. Rabinowitz, en un trabajo en colaboración con A. Castro publicado en la Revista Colombiana de Matemáticas, demostramos que el problema (3) tiene $4j - 1$ soluciones radiales (i.e. $u(x) = u(\|x\|)$) cuando Ω es una bola, la no linealidad tiene un cero positivo y el rango de la derivada de la no linealidad incluye al menos los primeros j valores propios.

Teorema 5. ([Cas-Cos2], 1993) *Si f tiene un cero positivo y $f'(0) = f'(\infty) > \lambda_j$ entonces (3) tiene al menos $4j - 1$ soluciones radiales.*

En ese mismo artículo se presenta una demostración más simple del siguiente resultado obtenido anteriormente por M. Esteban.

Teorema 6. ([Es], 1985) *Si $0 < f'(0) < \lambda_{j+1}$ y $\lambda_{j+k} < f'(\infty)$ entonces (3) tiene al menos $2k + 1$ soluciones radiales.*

Los Teoremas 5 y 6 se demuestran obteniendo una descripción de la gráfica del conjunto de soluciones del problema

$$\begin{cases} \Delta u + \lambda f(u) = 0 & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega, \end{cases} \quad (14)$$

donde $\lambda \in \mathbf{R}$.

3. Principios de minimax y aplicaciones

Otra herramienta muy importante en análisis no lineal la constituyó la introducción de métodos topológicos en el estudio de problemas variacionales; es decir, problemas en los cuales las soluciones se consiguen como puntos críticos de un funcional.

Las ideas topológicas iniciales de estos métodos fueron vislumbradas por Poincaré y Birkhoff a finales del siglo XIX y fueron desarrolladas posteriormente, en los años 20 y 30 del siglo pasado, por Morse ([Mors1], 1925, y [Mors2], 1934) y Ljusternik y Schnirelman ([Lj-Sc], 1934). En particular, Ljusternik y Schnirelman, en el caso de problemas variacionales en variedades de dimensión finita, dieron una cota inferior para el número de puntos críticos en términos de un invariante topológico llamado la categoría de Ljusternik y Schnirelman. La teoría de Morse para funciones Φ no degeneradas dio una clasificación más fina de los puntos críticos en términos

de las formas cuadráticas asociadas con la segunda derivada $\Phi''(u_0)$. En la década del 60 del siglo pasado, las ideas de la teoría de Morse fueron puestas en el contexto de la Geometría Diferencial en variedades de dimensión infinita por R. Palais y S. Smale ([Pa-Sm], 1964) quienes reemplazaron la hipótesis de dimensión finita de la teoría original por una hipótesis de compacidad sobre el funcional.

Como se mencionó anteriormente, con los métodos variacionales las soluciones de los problemas no lineales se consiguen como los puntos críticos de un funcional asociado con el problema en estudio. En el caso en que el funcional sea acotado inferior o superiormente es razonable tratar de demostrar que el funcional alcanza su mínimo o su máximo. Para funciones convexas un resultado clásico en esta dirección es el siguiente.

Teorema 7. Sean X un espacio de Banach reflexivo y Φ una función semicontinua inferiormente definida en X . Si Φ es convexa, acotada inferiormente y $\Phi(x) \rightarrow \infty$ cuando $x \rightarrow \infty$ entonces Φ alcanza su mínimo.

Si Φ no es convexa entonces Φ no necesariamente alcanza su mínimo. Sin embargo el siguiente resultado de I. Ekeland demuestra la existencia de puntos que son casi puntos de mínimo.

Teorema 8. ([Ek], 1974) Sean X un espacio métrico completo y $\Phi : X \rightarrow \mathbf{R} \cup \{+\infty\}$ una función semicontinua inferiormente y acotada inferiormente. Entonces dado $\epsilon > 0$ existe $u_\epsilon \in X$ tal que

$$\Phi(u_\epsilon) \leq \inf_X \Phi + \epsilon, \quad (15)$$

y

$$\Phi(u_\epsilon) < \Phi(u) + d(u, u_\epsilon), \quad \forall u \in X, u \neq u_\epsilon. \quad (16)$$

Una condición de compacidad para funciones C^1 , que es utilizada en la demostración de la existencia de puntos críticos es la llamada condición de Palais-Smale (P-S), que establece que si una sucesión $(u_n) \in X$ es tal que $|\Phi(u_n)| < M$ y $\Phi'(u_n) \rightarrow 0$ entonces (u_n) tiene una sub-sucesión convergente. Tales funciones siempre alcanzan su ínfimo.

Teorema 9. Si Φ es una función de clase C^1 definida en un espacio de Banach, satisface (P-S) y es acotada inferiormente entonces Φ alcanza su mínimo.

Si el funcional asociado con el problema no lineal no es acotado ni superior ni inferiormente se deben buscar

puntos de silla, los cuales se encuentran usando argumentos de tipo minimax. En éstos se comienza considerando una cierta clase A de conjuntos Σ en un cierto espacio topológico, luego se construye

$$\max_{u \in \Sigma} \Phi(u)$$

para cualquier $\Sigma \in A$, y finalmente se forma

$$c = \inf_{\Sigma \in A} \max_{u \in \Sigma} \Phi(u)$$

y se trata de demostrar que el número c es un punto crítico de Φ .

Un método específico muy útil para encontrar puntos de tipo minimax, que extiende las ideas de Poincaré y Birkhoff, es el conocido *Teorema del Paso de la Montaña*, publicado por A. Ambrosetti y P. Rabinowitz y que se presenta a continuación.

Teorema 10. ([Am-Ra], 1973) Si Φ es una función de clase C^1 definida en un espacio de Banach X , satisface (P-S) y la siguiente condición geométrica

$$\begin{cases} \Phi(0) = 0, \Phi(x) \geq \alpha > 0, \forall x \in X \text{ con } \|x\| = R & \text{y} \\ \Phi(x_0) \leq 0, \text{ para algún } x_0 \in X \text{ con } \|x_0\| > R, \end{cases} \quad (17)$$

entonces el siguiente número es un valor crítico de Φ ,

$$c = \min_{g \in P} \max_{0 \leq t \leq 1} \Phi(g(t)) \geq \alpha,$$

donde $P = \{g \in C([0, 1], X) / g(0) = 0, g(1) = x_0\}$. Es decir, existe $u \in X$ tal que

$$\Phi'(u) = 0 \quad \text{y} \quad \Phi(u) = c.$$

En años recientes se han conseguido extensiones y variaciones de este teorema que han demostrado ser muy útiles en las aplicaciones a ecuaciones diferenciales.

Otra técnica muy importante en análisis no lineal es el llamado *método de reducción de Lyapunov-Schmidt*, el cual permite reducir el estudio de los puntos críticos de un funcional Φ definido en un espacio de Hilbert H de dimensión infinita al estudio de los puntos críticos de un funcional $\hat{\Phi}$ definido en un subespacio cerrado de H , generalmente, de dimensión finita. El método de reducción, que presentamos a continuación, tuvo sus orígenes en las investigaciones de Lazer, Landesman y Meyers ([Lan-Laz-Me], 1975) y Castro y Lazer ([Cas-Laz2], 1979).

Teorema 11. ([Cas1], 1981) Sea H un espacio de Hilbert real y sean X y Y subespacios cerrados de H tales que $H = X \oplus Y$. Sea $\Phi : H \rightarrow \mathbf{R}$ un funcional de clase C^1 . Si existen $m > 0$ y $\alpha > 1$ tales que

$$\langle \nabla \Phi(x + y) - \nabla \Phi(x + y_1), y - y_1 \rangle \geq m \|y - y_1\|^\alpha,$$

$\forall x \in X, \forall y, y_1 \in Y$, entonces

(i) Existe una función continua $\psi : X \rightarrow Y$ tal que

$$\Phi(x + \psi(x)) = \min_{y \in Y} \Phi(x + y).$$

Además, $\psi(x)$ es el único elemento de Y tal que

$$\langle \nabla \Phi(x + \psi(x)), y \rangle = 0, \quad \forall y \in Y.$$

(ii) La función $\hat{\Phi} : X \rightarrow \mathbf{R}$ definida por $\hat{\Phi}(x) = \Phi(x + \psi(x))$ es de clase C^1 y

$$\langle \nabla \hat{\Phi}(x), x_1 \rangle = \langle \nabla \Phi(x + \psi(x)), x_1 \rangle, \quad \forall x, x_1 \in X.$$

(iii) $x \in X$ es un punto crítico de $\hat{\Phi}$ si y sólo si $x + \psi(x)$ es un punto crítico de Φ .

Utilizando el Teorema del Paso de la Montaña, el método de reducción de Lyapunov-Schmidt y la teoría de grado de Leray-Schauder, en un trabajo en colaboración con A. Castro, demostramos que el problema (3) tiene al menos cuatro soluciones no triviales, dos de las cuales son de un signo y las otras dos cambian de signo, cuando el rango de la derivada de la no linealidad incluye al menos los dos primeros valores propios.

Teorema 12. ([Cas-Cos3], 1994) Si $f'(0) < \lambda_1$, $f'(\infty) \in (\lambda_k, \lambda_{k+1})$ con $k \geq 2$, y $f'(t) \leq \gamma < \lambda_{k+1}$, entonces (3) tiene al menos cuatro soluciones no triviales. Además, uno de los siguientes casos ocurre:

a) k es par y (3) tiene dos soluciones que cambian de signo.

b) k es par y (3) tiene cinco soluciones, tres de las cuales tienen el mismo signo.

c) k es impar y (3) tiene dos soluciones que cambian de signo.

d) k es impar y (3) tiene tres soluciones del mismo signo.

En [Cos-Ve], con la colaboración de C. Vélez, sin considerar la hipótesis del acotamiento global de la derivada de la no linealidad del Teorema 12 y con la restricción k par, demostramos que el problema (3) tiene por lo menos tres soluciones no triviales. La prueba utiliza argumentos del tipo de paso de la montaña y teoría de grado.

Teorema 13. ([Cos-Ve], 2003) Si $f'(0) < \lambda_1$, $f'(\infty) \in (\lambda_k, \lambda_{k+1})$ con k un entero par, $k \geq 2$, entonces (3) tiene por lo menos tres soluciones no triviales, una es positiva, otra es negativa y la tercera cambia de signo.

En colaboración con S. Herrón, usando el Teorema del Paso de la Montaña, el Teorema de Punto de Silla de P. Rabinowitz ([Ra3], 1986) y argumentos del índice de Morse desarrollados por A. Lazer y S. Solimini ([Laz-Sol], 1988) demostramos la existencia de por lo menos tres soluciones para el problema (3) sin la restricción sobre k del Teorema 13.

Teorema 14. ([Cos-He], 2003) Si $f'(0) < \lambda_1$, $f'(\infty) \in (\lambda_k, \lambda_{k+1})$ con $k \geq 2$, y todos los puntos críticos del funcional J (véase su definición en (*), abajo) son no degenerados entonces (3) tiene por lo menos tres soluciones no triviales: $u_1 > 0$ en Ω , $u_2 < 0$ en Ω y u_3 . Las soluciones de un signo tienen índice de Morse menor o igual a 1 y u_3 tiene índice de Morse mayor o igual a 2.

En un artículo publicado en 1997, escrito en colaboración con los profesores J. M. Neuberger y A. Castro, demostramos **un principio de minimax** que nos permitió establecer condiciones suficientes que garantizan la existencia de soluciones que *cambian de signo exactamente una vez* para problemas de Dirichlet del tipo (3), cuando f es superlineal; es decir,

$$f'(\infty) := \lim_{s \rightarrow \infty} \frac{f(s)}{s} = \infty$$

y subcrítica; es decir,

$$|f(s)| \leq c_1 |s|^p + c_2, \quad \text{con } 1 \leq p \leq \frac{2N}{N-2}.$$

Teorema 15. ([Cas-Cos-Nu1], 1997) Si $f'(0) < \lambda_1$ entonces (3) tiene al menos tres soluciones no triviales: $\omega_1 > 0$ en Ω , $\omega_2 < 0$ en Ω y ω_3 . La función ω_3 cambia de signo exactamente una vez en Ω , es decir, $(\omega_3)^{-1}(\mathbf{R} - \{0\})$ tiene exactamente dos componentes conexas. Si son no degeneradas, las soluciones de un signo tienen índice de Morse 1 y la solución que cambia de signo tiene índice de Morse 2. Además,

$$J(\omega_3) \geq J(\omega_1) + J(\omega_2),$$

donde

$$J(u) = \int_{\Omega} \left(\frac{1}{2} |\nabla u|^2 - F(u) \right) dx, \quad \forall u \in H_0^1(\Omega) \quad (*)$$

y

$$F(s) = \int_0^s f(t) dt.$$

Hasta donde sabemos, este resultado es el primero en establecer la existencia de una solución del problema (3) que *cambia de signo exactamente una vez*. Desarrollos como los obtenidos por Z. Q. Wang ([Wa], 1991) no implican la existencia de soluciones que cambien de signo,

mucho menos que cambien de signo exactamente una vez. También, nuestras pruebas mejoran la información acerca del valor del funcional J en los puntos críticos y del índice de Morse de las soluciones.

La solución que cambia de signo se obtiene minimizando el funcional J en el subconjunto

$$S_1 = \{u \neq 0 : \langle \nabla J(u), u \rangle = 0, u_+, u_- \neq 0, \langle \nabla J(u_+), u_+ \rangle = 0\}. \quad (18)$$

Usando el principio de minimax mencionado arriba para problemas superlineales y los resultados de teoría de grado de A. Castro y J. Cossio ([Cas-Cos3], 1994), demostramos la existencia de soluciones que cambian de signo para problemas de Dirichlet asintóticamente lineales; es decir, cuando la no linealidad satisface

$$f'(\infty) := \lim_{s \rightarrow \infty} \frac{f(s)}{s} \in \mathbf{R}.$$

El primer resultado en esta dirección es el siguiente:

Teorema 16. ([Cas-Cos-Nu3], 1998) *Si $tf''(t) > 0$ para $t \neq 0$ y $f'(-\infty), f'(+\infty) \in (\lambda_2, \infty)$, entonces (3) tiene al menos cuatro soluciones. Una de las soluciones cambia de signo exactamente una vez y si es aislada su grado local es $+1$.*

Este último teorema incluye el caso en el cual (3) presenta saltos no lineales, i.e., el intervalo $(f'(-\infty), f'(+\infty)) \cup (f'(+\infty), f'(-\infty))$ contiene un valor propio λ_k . A su vez el Teorema nos permite extender los resultados del Teorema 12. En efecto, tiene lugar el siguiente teorema.

Teorema 17. ([Cas-Cos-Nu3], 1998) *Si $tf''(t) > 0$ para $t \neq 0$ y $f'(-\infty), f'(+\infty) \in (\lambda_k, \lambda_{k+1})$ para $k \geq 2$, entonces (3) tiene por lo menos cinco soluciones, dos de las cuales cambian de signo. Además, una de estas dos soluciones cambia de signo exactamente una vez.*

Además, demostramos que en el Teorema no necesariamente se consiguen más de dos soluciones que cambien de signo.

Teorema 18. ([Cas-Cos-Nu3], 1998) *Si $k = 2$ en el Teorema entonces (3) tiene exactamente dos soluciones que cambian de signo; ambas cambian de signo una sola vez.*

En [Cas-Cos-Nu2], publicado en 1997, con la colaboración de A. Castro y J. M. Neuberger, demostramos la existencia de soluciones no radiales que cambian de signo cuando Ω es una bola en \mathbf{R}^N y f es asintóticamente lineal. Sea $\lambda_1^r < \lambda_2^r < \dots < \lambda_k^r < \dots$ la sucesión de valores propios de $-\Delta$ actuando en funciones radiales de

$H_0^1(\Omega)$. Recordamos que $\lambda_1 = \lambda_1^r$ y $\lambda_2 < \lambda_2^r$. Se probó el siguiente teorema.

Teorema 19. ([Cas-Cos-Nu2], 1997) *Si $tf''(t) > 0$ para $t \neq 0$, $f'(\infty) \in (\lambda_k, \lambda_{k+1})$ con $k \geq 2$, $f'(t) \leq \gamma < \lambda_{k+1}$ para todo $t \in \mathbf{R}$ y $\lambda_1 < \lambda_k < \lambda_{k+1} \leq \lambda_2^r$, entonces el problema (3) tiene al menos dos soluciones que son no radiales y cambian de signo. Además, una de estas dos soluciones cambia de signo exactamente una vez.*

4. Construcción y visualización de soluciones para problemas elípticos semilineales

Otro tópico de investigación en el que he trabajado es el de construcción y visualización de soluciones para problemas elípticos semilineales. Este trabajo es importante ya que en su gran mayoría los teoremas de existencia de soluciones para problemas semilineales son no constructivos; además siempre serán bienvenidos ejemplos explícitos que apoyen o nieguen conjeturas y que sugieran direcciones para estudios futuros.

En dos trabajos escritos en colaboración con el profesor H. Arango, el primero publicado en la Revista Colombiana de Matemáticas ([Ar-Cos1], 1996) y el segundo en el Electronic Journal of Differential Equations ([Ar-Cos2], 2000), presentamos métodos para construir explícitamente soluciones con simetría radial para el problema (3) cuando el dominio Ω es una bola en \mathbf{R}^N y la función f es continua y lineal por tramos. El método de construcción está basado en un spline de funciones de Bessel y en el “shooting method”. Estos trabajos han sido inspirados por un trabajo previo de los profesores E. Deumens y H. Warchall ([De-War], 1988), en el que los autores estudian una ecuación de onda no lineal en el espacio \mathbf{R}^{n+1} .

Consideremos la siguiente función continua lineal por tramos $f : \mathbf{R} \rightarrow \mathbf{R}$ definida por

$$f(t) = \begin{cases} \Lambda^2 t - (\Lambda^2 - \beta^2) & \text{si } t > 1 \\ \beta^2 t & \text{si } -1 \leq t \leq 1 \\ \Lambda^2 t + (\Lambda^2 - \beta^2) & \text{si } t < -1, \end{cases} \quad (19)$$

donde Λ y β son constantes.

Sea $0 < \lambda_1 < \lambda_2 < \dots < \lambda_k < \dots$ la sucesión de valores propios $-\Delta$ actuando en funciones radiales de H_0^1 . Demostramos el siguiente resultado.

Teorema 20. ([Ar-Cos1], 1996) *Si $\beta^2 \leq \lambda_k$ y $\Lambda^2 > \lambda_p$ ($k \leq p$) entonces existen $(p - k + 1)$ pares*

de soluciones radiales de (3), u_i y v_i ($k \leq i \leq p$), tales que para cada i , $u_i(0) < 0 < v_i(0)$ y u_i, v_i tienen exactamente $(i - 1)$ nodos en $(0, \pi)$.

También se construyeron explícitamente soluciones radiales para la ecuación (3), cuando la no linealidad $f : \mathbf{R} \rightarrow \mathbf{R}$ está definida por

$$f(t) = \begin{cases} \alpha^2 t & \text{si } t \leq \frac{\beta}{2} \\ -\alpha^2 t + \alpha^2 \beta & \text{si } \frac{\beta}{2} \leq t \leq \beta \\ \alpha^2 t - \alpha^2 \beta & \text{si } t \geq \beta \end{cases} \quad (20)$$

donde α y β son constantes.

Teorema 21. ([Ar-Cos2], 2000) Si $\lambda_j < \alpha^2 < \lambda_{j+1}$ y $\beta > 0$ entonces para cada $0 \leq i \leq j - 1$ existen soluciones radiales únicas v_i y u_i para el problema (3) con i nodos en $(0, \pi)$, tales que $v_i(0) > \beta > 0$ y $0 < u_i(0) < \beta$.

Presentamos una gráfica de la solución obtenida con el programa Mathematica, siguiendo el método de construcción empleado en el artículo (véase la Figura 1).

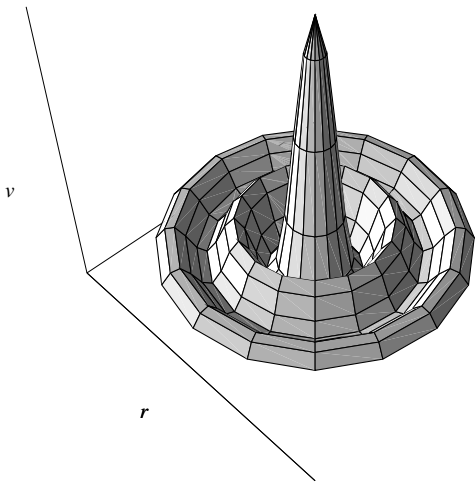


Figura 1. Solución radial $v(r)$ en 3 dimensiones con $\alpha = 5.1$, $\beta = 2.0$ y 4 nodos.

También, en una investigación en colaboración con los profesores J. M. Neuberger y S. Lee ([Cos-Le-Nu], 2001) logramos construir un nuevo algoritmo para aproximar soluciones que se consiguen vía el método de reducción de Lyapunov-Schmidt para problemas de Dirichlet sublineales en dominios acotados.

5. Perspectivas

Como se puede observar de los resultados presentados en las Secciones 2, 3 y 4 de este trabajo, el autor, en colaboración con Alfonso Castro y otros investigadores en el área, ha introducido nuevas ideas en el estudio de la existencia, las propiedades cualitativas y la construcción y visualización de soluciones para el problema elíptico semilineal

$$\begin{cases} \Delta u + f(u) = 0 & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega. \end{cases} \quad (21)$$

Estas ideas nos conducen a formular una serie de preguntas abiertas, que serán especificadas en esta sección, y que esperamos contribuyan a una mejor comprensión de las soluciones de problemas de Dirichlet no lineales.

Perspectiva I. Motivados por el principio de Minimax presentado en el Teorema 15 y por los resultados obtenidos en los Teoremas , y sería interesante estudiar la *existencia de soluciones no radiales que cambian de signo exactamente una vez para el problema (21)*. Recientemente, A. Castro y H. Aduén (véase [Ad-Cas], 2003) demostraron la existencia de infinitas soluciones no radiales para problemas de Dirichlet superlineales.

Perspectiva II. Motivados por los resultados obtenidos en los Teoremas 12, y nos parece importante estudiar la *existencia de soluciones que cambian de signo en dominios generales*. Conjeturamos que es posible debilitar la hipótesis de convexidad del Teorema y conseguir resultados similares.

Perspectiva III. Como el Principio de Minimax de Castro-Cossio-Neuberger (véase Teorema 15) también vale cuando f se reemplaza por λf , sería interesante *demostrar que las soluciones correspondientes u_λ están en la misma curva de bifurcación*. Conjeturamos que ellas se bifurcan del segundo valor propio y que existe un conjunto conexo de soluciones que cambian de signo exactamente una vez y conecta $(0, \infty)$ con $(\lambda_2, 0)$. A medida que vayan apareciendo más soluciones al problema (21) es importante estudiar las correspondientes curvas de bifurcación.

Perspectiva IV. Motivados por el caso unidimensional, el caso radial y los resultados del Teorema 15, sería importante *encontrar condiciones suficientes que permitan demostrar que para cada entero positivo k el problema (21) tiene una solución con k regiones nodales*.

Perspectiva V. Mientras el caso unidimensional, el caso radial y el caso en el cual la no linealidad f es impar sugieren que el problema (21) tiene $2k + 1$ soluciones cuando $f'(\infty) \in (\lambda_k, \lambda_{k+1})$, el contraejemplo de Dancer

(véase [Da], 1976) nos indica que esto no siempre es posible. Sería interesante *demostrar que cuando $k \rightarrow \infty$ el número de soluciones de (21) tiende a ∞ .*

Perspectiva VI. Motivados por los resultados obtenidos por el autor, en colaboración con H. Arango y J. M. Neuberger, en el estudio del problema (21) en dominios simétricos (véase [Ar-Cos1], 1996; [Ar-Cos2], 2000; y [Cos-Le-Nu], 2001) es importante *estudiar la construcción explícita de soluciones, para el problema mencionado, en dominios no simétricos.*

Referencias

- [Ad-Cas] **H. Aduén and A. Castro**, *Infinitely Many Nonradial Solutions to a Superlinear Dirichlet Problem*, Proc. Amer. Math. Soc. **131** (2003), 835–843.
- [Am-Pro] **A. Ambrosetti and G. Prodi**, *A Primer of Nonlinear Analysis*, Cambridge Studies in Advanced Mathematics 34, Cambridge University Press, Cambridge 1993.
- [Am-Ra] **A. Ambrosetti and P. Rabinowitz**, *Dual variational methods in critical point theory*, J. Funct. Anal. **14** (1973), 349–381.
- [Ar-Cos1] **H. Arango y J. Cossio**, *Construcción de Soluciones Radialmente Simétricas para un Problema Elíptico Semilineal*, Rev. Colombiana Mat. **30** (1996), 77–92.
- [Ar-Cos2] **H. Arango and J. Cossio**, *Explicit Construction, Uniqueness, and Bifurcation curves of Solutions for a Nonlinear Dirichlet Problem in a Ball*, Electronic Journal of Differential Equations, Conf. 05, 2000, 1–12.
- [Cas1] **A. Castro**, *Métodos de Reducción via Minimax*, Primer Simposio Colombiano de Análisis Funcional, Medellín, Colombia, (1981).
- [Cas-Cos1] **A. Castro and J. Cossio**, *A Bifurcation Theorem and Applications*, Dynamic Systems and Applications **2** (1993), 221–226.
- [Cas-Cos2] **A. Castro and J. Cossio**, *Multiple Radial Solutions for a Semilinear Dirichlet Problem in a ball*, Rev. Colombiana Mat. **27** (1993), 15–24.
- [Cas-Cos3] **A. Castro and J. Cossio**, *Multiple Solutions for a Nonlinear Dirichlet Problem*, SIAM J. Math. Anal. **25** (1994), 1554–1561.
- [Cas-Cos-Nu1] **A. Castro, J. Cossio and J. M. Neuberger**, *A Sign-Changing Solution for a Superlinear Dirichlet Problem*, Rocky Mountain J.M. **27** (1997), 1041–1053.
- [Cas-Cos-Nu2] **A. Castro, J. Cossio and J. M. Neuberger**, *On Multiple Solutions of a Nonlinear Dirichlet Problem*, Nonlinear Analysis, Theory, Methods & Applications, **30** (1997), 3657–3662.
- [Cas-Cos-Nu3] **A. Castro, J. Cossio and J. M. Neuberger**, *A Minimax Principle, Index of the Critical Point, and Existence of Sign-changing Solutions to Elliptic Boundary Value Problems*, Electronic Journal of Differential Equations **1998** (1998), 1–18.
- [Cas-Laz1] **A. Castro and A. C. Lazer**, *Applications of a Max-min Principle*, Rev. Colombiana Mat. **10** (1976), . 141–149.
- [Cas-Laz2] **A. Castro and A. C. Lazer**, *Critical Point Theory and the Number of Solutions of a Nonlinear Dirichlet Problem*, Ann. Mat. Pura Appl. **70** (1979), 113–137.
- [Cos] **J. Cossio**, *Múltiples Soluciones para un Problema Elíptico Semilineal*. En *Memorias de la III Escuela de Verano en geometría diferencial, ecuaciones diferenciales parciales y análisis numérico*. Academia Colombiana de Ciencias Exactas Físicas y Naturales, Colección memorias No. 7, 1995, 53–59.
- [Cos-Le-Nu] **J. Cossio, S. Lee, and J. M. Neuberger**, *A Reduction Algorithm for Sublinear Dirichlet Problems*, Nonlinear Analysis, **47** (2001), 3379–3390.
- [Cos-He] **J. Cossio and S. Herrón**, *Nontrivial Solutions for a Semilinear Dirichlet Problem with Nonlinearity Crossing Multiples Eigenvalues*, Submitted for publication in the Journal of Dynamics and Differential Equations.
- [Cos-Ve] **J. Cossio y C. Vélez**, *Soluciones no Triviales para un Problema de Dirichlet Asintóticamente Lineal*, Aceptado para publicación en la Revista Colombiana de Matemáticas, 2003.
- [Ch] **S. Chandrasekhar**, *An Introduction to the Study of Stellar Structure*, University of Chicago Press, 1939.
- [Cr-Ra] **M. Crandall and P. Rabinowitz**, *Bifurcation from Simple Eigenvalue*, J. Funct. Anal. **8** (1971), 321–340.
- [Da] **E. Dancer**, *Counterexamples to Some Conjectures on the Number of Solutions of Nonlinear Equations*, Math. Ann. **16** (1976), 1361–1376.
- [De-War] **E. Deumens and H. Warchall**, *Explicit Construction of all Spherical Symmetric Solitary Waves for a Nonlinear Wave Equation in Multiple Dimensions*, Nonlinear Analysis, Theory, Methods and Applications **12** (1988), 419–447.
- [Ek] **I. Ekeland**, *On the Variational Principle*, J. Math. Anal. Appl. **47** (1974), 324–353.
- [Es] **M. Esteban**, *Multiple Solutions of Semilinear Elliptic Problems in a Ball*, J. Differential Equations **57** (1985), 112–137.
- [He] **M. Henon**, *Numerical Experiments on the Stability of spherical Stellar Systems*, Astro. Astrophys. **24** (1973), 229–238.
- [Li-Ya] **E. Lieb and H. T. Yau**, *The Chandrasekhar Theory of Stellar Collapse as the Limit of Quantum Mechanics*, Commun. Math. Phys. **112** (1987), 147–174.
- [Ke-An] **J. Keller and S. Antman**, *Bifurcation Theory and Nonlinear Eigenvalue problems*, W. A. Benjamin, Inc., New York, 1969, 395–409.
- [Lan-Laz-Me] **E. M. Landesman, A. C. Lazer, and D. Myers**, *On Saddle Point Problems in the Calculus*

- of Variations, the Ritz Algorithm, and Monotone Convergence*, J. Math. Anal. Appl. **53** (1975), 594–614.
- [Laz-Sol] **A. C. Lazer and S. Solimini**, *Nontrivial Solutions of Operator Equations and Morse Indices of Critical Points of Min-Max Type*, Nonlinear Analysis TMA **12** (1988), 761–775.
- [Lj-Sc] **L. Ljusternik and L. Schnirelman**, *Méthodes Topologiques dans les Problèmes Variationnels*, Hermann, Paris, 1934.
- [Mors1] **M. Morse**, *Relations Between the Critical Points of a Real Function of n Independent Variables*, Trans. Amer. Math. Soc. **27** (1925), 345–396.
- [Mors2] **M. Morse**, *The Calculus of Variations in the Large*, Amer. Math. Soc. Colloq. Publ., 18, 1934.
- [Pa-Sm] **R. Palais and S. Smale**, *A Generalized Morse Theory*, Bull. Amer. Math. Soc. **70** (1964), 165–171.
- [Ra1] **P. Rabinowitz**, *Some Global Results for Nonlinear Eigenvalue Problems*, J. Funct. Anal. **7** (1971), 487–513.
- [Ra2] **P. Rabinowitz**, *Topological Methods in Bifurcation Theory*, Séminaire de Mathématiques Supérieures, Séminaire Scientifique OTAN, Les Presses de l'Université de Montreal, Montreal, 1985.
- [Ra3] **P. Rabinowitz**, *Minimax Methods in Critical Point Theory with Applications to Differential Equations*, Regional Conference Series in Mathematics, 65, Providence, R. I., AMS (1986).
- [Wa] **Z. Q. Wang**, *On a Superlinear Elliptic Equation*, Ann. Inst. H. Poincaré. Analyse Non Linéaire **8** (1991), 43–57.



**Jairo Charris Castañeda
(1939-2004)**