



Matemáticas y sus aplicaciones 7

Fernando Macías Romero

Editor



Los capítulos del presente libro son co-creados por los autores al poner su atención e intención en participar en la Second International Conference on Mathematics and its Applications (2CIMA). El 2CIMA emana de la fortuna de contar con el mejor comité organizador que ha puesto la Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas de la BUAP, he aquí los resultados que promueve la riqueza matemática, trabajos tenaces que lograron sobreponerse a los inexorables jueces y fueron autorizados después de un arbitraje riguroso.

BUAP

FCFM

CSMA

60
AÑOS DE
AUTONOMÍA
UNIVERSITARIA

textos
Científicos

Matemáticas y sus aplicaciones 7

*Dedicado al Profesor Agustín Contreras Carreto
por su 60 aniversario*



Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias Físico Matemáticas

Fernando Macías Romero
Editor

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

José Alfonso Esparza Ortiz

Rector

René Valdiviezo Sandoval

Secretario General

Ygnacio Martínez Laguna

Vicerrector de Investigación y Estudios de Posgrado

Flavio Marcelino Guzmán Sánchez

E. D. Vicerrectoría de Extensión y Difusión de la Cultura

Ana María Dolores Huerta Jaramillo

Directora de Fomento Editorial

Martha Alicia Palomino Ovando

Directora de la Facultad de Ciencias Físico Matemáticas

Primera edición, 2016

ISBN: 978-607-525-135-6

© Benemérita Universidad Autónoma de Puebla

Dirección de Fomento Editorial

2 Norte 1404, C.P. 72000

Puebla, Pue.

Teléfono y fax: 01 222 246 8559

Impreso y hecho en México

Printed and made in Mexico

Matemáticas y sus aplicaciones 7

Selección bajo arbitraje riguroso de algunos trabajos presentados en el
Second International Conference on Mathematics and its Applications
(2CIMA), FCFM, BUAP.

Editor

Fernando Macías Romero

Comité científico internacional

Francisco Cala Rodríguez (UAC, CHL), Jesús Ildefonso Díaz (UCM, ESP), Antonio Díaz Ramos (UMA, ESP), Ermínia de Lourdes Campello Fanti (UNESP, BRA), Gudelía Figueroa Preciado (UNISON), Miguel Ángel García Ariza (BUAP), Alejandro García Diego (UNAM), David Herrera Carrasco (BUAP), Miguel Antonio Jiménez Pozo (BUAP), Judy Kennedy (LU, USA), Alejandro Illanes (UNAM), Isabel Llatas Salvador (USB, VEN), Fernando Macías Romero (BUAP), Edgar Martínez Moro (UVA, ESP), Juan Pablo Pinasco (UBA, ARG), Gerardo Raggi Cárdenas (UNAM), José Gregorio Rodríguez Nieto (UNC, CO), Fernando Tohmé (UNS, ARG), Camillo Trapani (UNIPA, ITA), Teresa de Jesús Valerio López (UAQ), Francisco Venegas Martínez (IPN), David Villa Hernández (BUAP), Federico Weinschelbaum (UdeSA, ARG).

Contenido

Presentación	1
Homenaje académico al Profesor Agustín Contreras Carreto por su 60 aniversario	
<i>Breve semblanza.</i> David Herrera Carrasco, Fernando Macías Romero	5
<i>Semblanza a Agustín.</i> José Juan Angoa Amador	9
<i>Al maestro Agustín.</i> María de Jesús López Toriz	11
<i>Homenaje al trabajo académico del Dr. Agustín Contreras Carreto en ocasión de su 60 aniversario.</i> Ángel Tamariz Mascarúa	13
<i>Al Profe Contrí.</i> María del Rocío Macías Prado	15
<i>Generosidad.</i> La rayita	17
<i>Caricaturas 1996.</i> Agustín Contreras Carreto	19
Álgebra	
Capítulo 1. El Anillo de Burnside	31
César Cejudo Castilla, Juan Manuel Ramírez Contreras, David Villa Hernández	
Capítulo 2. El lema de Hensel y el levantamiento de Hensel	53
Ángel Raúl García Ramírez, Carlos Alberto López Andrade	
Capítulo 3. Una variación del contenido fuerte entre formaciones localmente definidas	83
Ismael Gutiérrez García, Anselmo Torresblanca Badillo	

Análisis

Capítulo 4. Classification of the continuous translations in spaces of weighted integrable real functions	109
<i>Álvaro Hernández Cervantes, Miguel Antonio Jiménez Pozo</i>	
Capítulo 5. Aproximación tipo Korovkin y Sistemas de Chebyshev	123
<i>José Margarito Hernández Morales, José Luis Carrasco Pacheco</i>	
Filosofía, Divulgación e Historia de las Matemáticas	
Capítulo 6. El descubrimiento de la inconmensurabilidad por Hipaso de Metaponto	155
<i>Agustín Contreras Carreto, Elizabeth de Gante Coronel, María del Rocío Macías Prado</i>	
Probabilidad y Estadística	
Capítulo 7. Un portafolio de inversión óptimo bajo un enfoque de Conditional Value at Risk (CVaR)	177
<i>Ambrosio Ortiz Ramírez, Abigail Rodríguez Nava, Luis Alberto Sánchez Zacateco</i>	
Topología	
Capítulo 8. Funciones de Whitney admisibles	199
<i>José Gerardo Ahuatzi Reyes, David Herrera Carrasco, Fernando Macías Romero</i>	
Capítulo 9. Tercer producto simétrico único	231
<i>Vianey Córdova Salazar, David Herrera Carrasco, Fernando Macías Romero</i>	
Capítulo 10. Caracterización de los hiperespacios de un continuo localmente conexo	249
<i>Lázaro Flores De Jesús, David Herrera Carrasco, Fernando Macías Romero</i>	
Índice de autores	269

Presentación

Tenemos la fortuna de hacer un recorrido hacia un proceso creativo sin precedentes. Ha llegado el momento de compartir esta sabiduría, de invitar a todo el mundo a embarcarse en el navío que nos conduce hacia la Fuente de todo lo creado. Esta es la razón por la cual editamos el libro que tienen en sus manos. La felicidad que propone este libro por su divulgación, investigación e intercambio de ideas se debe a la generosidad de muchísimos matemáticos que participaron en el denominado *Second International Conference on Mathematics and its Applications* (2CIMA, 2015), un esfuerzo profesional consolidado que ha permitido la participación de grandes personajes de diversas universidades, nacionales y extranjeras, tanto en el desarrollo del 2CIMA como en su memoria escrita, que es el presente libro. La base ha sido un comité organizador especializado, entusiasta y vigoroso emanado de la Academia de Matemáticas de la Facultad de Ciencias Físico Matemáticas de la BUAP. Este producto no es ni siquiera setemesino, es normal, de por lo menos nueve meses de trabajo constante. Por el amor a la matemática es que ha nacido este ejemplar que nos brinda la sabiduría necesaria para mostrarles parte de nuestros quehaceres cotidianos.

Los capítulos de este libro están agrupados por secciones de acuerdo al área temática en el 2CIMA. Dichos capítulos fueron sometidos a arbitraje riguroso.

Agradecemos, con toda el alma, a todos los árbitros su amabilidad, gentileza, dedicación y trabajo científico. Un agradecimiento especial a José Gerardo Ahuatzí Reyes por su apoyo en la edición de esta obra. Gracias por dejar huella.

Fernando Macías Romero
Editor

**Homenaje académico al Profesor Agustín
Contreras Carreto por su 60 aniversario**

Breve semblanza

Agustín Contreras Carreto, nacido en el D.F. el 5 de noviembre de 1956, es un prestigioso topólogo, geómetra, profesor, investigador, sumo escritor, sumo amanuense y sumo divulgador de la matemática, el mejor conferenciante, imitador, caricaturista, dr. magneto, músico, poeta «y loco». Directo, curioso, sociable, cariñoso, jovial; toca el saxofón, piano, acordeón; le gusta la alegría, las palanquetas, las galletas. Sus frutos: Manuel, Carlos, Héctor y Antonio. Busca la felicidad en el ser humano, cree en la bondad de las personas. Detesta las convocatorias BUAP. Es fan de contemplar el atardecer, de celebrar los cumpleaños de la gente, detallista, lector empedernido, le encanta comer y beber. Es un chico que dice que lo mejor de las fiestas se encuentra en la cocina. Buscador de tesoros descabellados. Su película: atrapados sin salida. Su actor: Jack Nicholson. Le emociona el falsete, el aroma, el sabor. Simpatiza con el Chi kung y Tai Chi. Estudió su licenciatura (1975-1979), maestría (1978-1980) y doctorado (1997-2003), todo en Matemáticas, en la Facultad de Ciencias de la UNAM. Fue profesor en la UNAM (1978 - febrero de 1980); profesor en la Universidad de Guanajuato (agosto de 1980 - julio de 1981). Afortunadamente llegó como profesor a la entonces Escuela de Ciencias Físico Matemáticas (ECFM) de la Universidad Autónoma de Puebla el 16 de agosto de 1981 y desde entonces a la fecha es Profesor de Tiempo Completo. Ha dirigido 17 tesis de licenciatura, 4 de maestría y una de doctorado. Es uno de los profesores de la FCFM más populares de todos los tiempos por su pasión y sabiduría. Sus alumnos afirman haber aprendido tanto de él. Sus mejores admiradoras: Hortensia, Pillimba, Delia y Graciela siguen sonriendo gustosas, observando todos sus logros desde el cielo.

Publicaciones

Libros

1. Agustín Contreras Carreto (otros autores: Juan Angoa Amador, Manuel Ibarra Contreras, Raúl Lineras Gracia, Armando Martínez García), «Matemáticas Elementales» Textos Científicos. Dirección de Fomento Editorial, BUAP. Primera ed., 2004. 2da ed., 2008. 3a. reimpresión de la 2da ed. 2015, 1000 ejemplares, ISBN: 978 607 7541 24 0, 277 páginas.
2. Agustín Contreras Carreto (otros autores: Juan Angoa Amador, Jaime Arroyo García, David Herrera Carrasco Manuel Ibarra Contreras, Raúl Lineras Gracia, Fernando Macías Romero, Armando Martínez García, Celestino Soriano

Soriano, Fernando Velázquez Castillo), «Cálculo Diferencial en una Variable» Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, marzo de 2005, tiraje de 1000 ejemplares, ISBN: 968 863 817 X, 261 páginas.

3. Agustín Contreras Carreto (otros autores: J. Juan Angoa Amador, Jaime Arroyo García, David Herrera Carrasco, Raúl Linares Gracia, Héctor Sánchez Morgado, Celestino Soriano Soriano, Fernando Velázquez Castillo), «Álgebra I», Benemérita Universidad Autónoma de Puebla, ISBN: 978 968 9282 412, 200 páginas.
4. Agustín Contreras Carreto (otros autores: J. Juan Angoa Amador, Manuel Ibarra, María de Jesús López Toriz), «Introducción a las estructuras algebraicas», Benemérita Universidad Autónoma de Puebla, primera edición, 2007, ISBN: 9789689182917, 97 páginas.
5. Agustín Contreras Carreto (otros autores: J. Juan Angoa Amador, Raúl Linares Gracia, María de Jesús López Toriz, Armando Martínez García), «Cálculo integral», Fomento Editorial de la Benemérita Universidad Autónoma de Puebla; primera edición, 2015, ISBN: 978 607 487 831 8, 299 páginas.

Capítulos de libros

1. Agustín Contreras, «Los principios de continuidad en los Elementos de Euclides» (Capítulo 6), Encuentro de enseñanza e historia de las matemáticas, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, págs. 71-77. Primera Edición. ISBN: 978 968 9391 13 5. Editores Juan Angoa, Manuel Ibarra. Tiraje de 500 ejemplares. 88 págs. Publicado en marzo de 2008.
2. Juan Angoa, Agustín Contreras, Manuel Ibarra, Ángel Tamariz, «Una introducción a los espacios E -regulares y a la C_p -teoría. Una presentación categórica» (Capítulo 5), Topología y Sistemas Dinámicos II, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, págs. 111-160. Primera Edición. ISBN: 978 698 595 255 2. Editores J. Juan Angoa Amador, José Arrazola, Raúl Escobedo, Alejandro Illanes, Mauricio Osorio, Julio Poisot, Guillermo Sienna, Ángel Tamariz. Tiraje de 500 ejemplares. 214 págs. Publicado en febrero de 2009.
3. Ángel Tamariz Mascarúa, Agustín Contreras Carreto, Manuel Ibarra Contreras «Algunas generalizaciones de la pseudocompacidad» (Capítulo 3), Topología y sus Aplicaciones 1, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, págs. 53-70. Primera Edición. ISBN:

978-607-487-388-7. Editores Juan Angoa, José Arrazola, Raúl Escobedo. Tiraje de 500 ejemplares. 206 págs. Publicado en febrero de 2012.

4. Juan Angoa, Agustín Contreras, Manuel Ibarra, «Algunas reflexiones sobre la historia de la matemática» (Capítulo 4), Matemáticas y sus aplicaciones 5, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, págs. 103-115, Primera Edición 2015, ISBN: 978-607-487-934-6. Editor Fernando Macías Romero. Tiraje de 500 ejemplares. 224 páginas. Publicado el 19 de octubre de 2015.
5. Agustín Contreras Carreto, Elizabeth de Gante Coronel, María del Rocío Macías Prado, «El descubrimiento de la inconmensurabilidad por Hipaso de Metaponto» (Capítulo 6), Matemáticas y sus aplicaciones 7, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, págs. 155-173, Primera Edición 2016, ISBN: 978-607-525-135-6. Editor Fernando Macías Romero. Tiraje de 500 ejemplares. páginas. Publicado el 15 de diciembre de 2016.

Artículos

1. Agustín Contreras-Carreto, Ángel Tamariz-Mascarúa, «On some generalizations of compactness in spaces $C_p(X, 2)$ and $C_p(X, \mathbb{Z})$ », Boletín de la Sociedad Matemática Mexicana (3) volumen 9 (2003), 291- 308.
2. Fidel Casarrubias-Segura, Agustín Contreras-Carreto, Alejandro Ramírez-Páramo, «Some new topological cardinal inequalities», Topology and its Applications 154 (2007) 1307-13013.

Agradecemos al «Contri» su amistad y aportaciones a la FCFM.

David Herrera Carrasco
Fernando Macías Romero

Semblanza a Agustín

La semblanza a alguien, son pedazos de recuerdos seleccionados desde el presente, uno trata de engañar a los lectores diciendo cosas como fueron antes, pero el presente nos engaña dos veces primero cuando creemos que así pasó y segundo cuando el cómo pensamos ahora, manda imperiosamente el qué recordar.

Hechas estas aclaraciones van mis recuerdos. Un día de una semana de un congreso de Educación realizado en Atlixco, conocí a Agustín, un joven matemático que doblando papel pretendía construir con nuestras manos una elipse, algo inusitado, frente a nosotros la matemática tenía un acercamiento manual, nuestros cuerpos pueden crear una materialidad del concepto, el concepto regresando a la materia. Pasados algunos días, y también la sorpresa de esta manifestación corpórea de la matemática, encontré a Agustín buscando a sus alumnos en la Escuela de Ciencias Físico Matemáticas, la total invisibilidad de sus alumnos nos permitió conocernos mejor, así entre tarareos de obras clásicas empecé a conocer algunos datos de Agustín de viva voz. Prófugo de un doctorado, buscó una vida útil en la provincia como maestro, la razón de un posgrado mostró la insuficiencia existencial para él, ahora me doy cuenta, que la conciencia de este hecho la tenemos muchos pero pocos tenemos el valor de abandonar un proyecto tan valioso socialmente.

Al paso de los años, hemos iniciado conjuntamente muchos proyectos matemáticos y no matemáticos, ambos igual de importantes. Recuerdo todos los libros que hemos escrito y fundamentalmente vivido intensamente, las primeras versiones de Matemáticas Básicas, Álgebra I y Cálculo Diferencial, él fue el amanuense en gran parte, y yo el dictador en otra gran parte. Estas tareas nos permitieron degustar el bolero en todas sus grandes cantantes, el jazz y Ella, el tango y Susana, José Alfredo y sus profundas dolencias a flor de piel, Cri Cri y muchísimas más curiosidades melódicas que su aparato de sonido Fisher no paraba de reproducir mientras trabajábamos. Aquí aprovecho para marcar una gran cualidad de Agustín, es un gran creador de ambientes en donde el trabajo, peligrosamente, se convierte en placentero. Así mientras desarrollábamos estos textos, Agustín se daba a la tarea de regalar un bolero, tocado por él y “cantado por mí”, en los cumpleaños de nuestros amigos (el JA). Así pasamos por Guty, Agustín Lara, Miguel Prado, José Alfredo, María Greever, Consuelito, todos absolutos desconocidos por las nuevas generaciones que no saben como enamorarse, lo cual no niega que se enamoren.

Como él mismo ha reconocido, un proyecto de posgrado sin los cuates para él no tiene sentido, así que los cuates le propusimos reanudar nuestros estudios de posgrado y ahí tenemos a Agustín reiniciando su doctorado, finalmente concluido

con el asesoramiento de ese otro gran cuate: Ángel Tamariz. Y henos aquí, doctores e ignorantes, revalorando nuestra ingenua y amorosa relación primaria con la matemática, la cual Agustín nunca perdió, y que ahora comprendo lo ha nutrido existencialmente toda su vida. Si algunos han perdido esta relación amorosa y convertido a la matemática en sólo la fuente de sus ingresos, lo cual ha situado a Agustín en diletante, para mí ha revalorado su convicción de que la matemática fundamentalmente nos hace felices.

Juan Angoa

Al maestro Agustín

Hablando de música, ¿qué sería de los *boleros* sin el maestro Agustín Lara? . . .

¿Qué sería de la *música clásica* sin Bach o Mozart?, como dice el Maestro Agustín Contreras acerca de sus preferencias: no puedo decidir, ambos son maximales.

¿Qué sería del mundo sin soles? recordando a Graciela Salicrup, por cierto uno de esos soles es su discípulo Agustín Contreras Carreto.

Desde un 16 de agosto de 1981, nuestra Facultad contó en su planta docente con el maestro Agustín Contreras Carreto, quien día a día, como emulando a la propiedad Arquimediana, con sus clases (*en todo el sentido de la palabra*, que hoy en día muchos pasan por alto esta hermosa labor), sus seminarios, sus atenciones a cientos de alumnos y calificando a los mismos, ha colaborado a crear a nuestra Facultad de Ciencias Físico Matemáticas de la BUAP, por cierto sin pedir a cambio la dichosa *constancia*.

Podemos prescindir de ciertas burocracias y burócratas y en verdad ¡no pasa nada!, sin embargo, estoy segura que no podemos prescindir del maestro Contri, no imagino a nuestra facultad con tal vacío.

Como es sabido por muchos de Ustedes la Teoría de los Continuos es un área importante de la matemática en nuestra facultad, por los maestros, por los investigadores que a ella se dedican, por los congresos en que ella se incluye y por los alumnos que por ella se interesan, incluso para el revuelo que la burocracia exige. La llegada de esta *bella rama de la topología*, también se la debemos al maestro Agustín y, por supuesto, a todos los organizadores de las *Jornadas Veraniegas de Topología*, ya que a través de Isabel Puga con su plática “Una introducción a los continuos” impartida un jueves 13 de junio de 1996, al seno de las Jornadas, nacieron *los continuos* en Puebla.

Recuerdo cuando le pedí al profesor Agustín que por favor me eligiera un curso a impartir en la *asamblea de la academia*, pues el día citado tenía que exponer en el seminario de la UNAM, bajo la dirección del profesor Janusz Charatonik. Al día siguiente me dice: *le tengo dos noticias*, la mala es que impartirás “Introducción a las estructuras algebraicas” y la buena es que “no hay textos para esta materia” resultado de esta sorpresa nació el primer libro que tenemos en coautoría, por supuesto no podía faltar su peculiar grito *¡perdón!*

Existen tantas cosas maravillosas que nos ha compartido, y sin constancia de por medio, sólo por el gusto que tiene por la *bella matemática*. Es capaz de realizar mil y una cosa en distantes áreas de la cultura; la música, la literatura, la historia, los que hemos tenido la dicha de ver sus manos atadas a las teclas de un *piano*, al

interpretar un bolero o una pieza de Mozart, nos impresiona escuchar todo lo que puede producir y crear, y no precisamente por estar publicando un *paper*, caray . . . le agradezco su gentileza, virtud que lo caracteriza, gracias maestro Agustín.

La milagrosa niña doctora de Tepeaca
(M. López Toriz)

Homenaje al trabajo académico del Dr. Agustín Contreras Carreto en ocasión de su 60 aniversario

*Parecieran agotarse los enigmas
y de madrugada
hay de nuevo aguamiel en el agave...¹*

Mario Calderón en *Hálito de origen*.

El mes de noviembre de 2016 se nos presenta una gran oportunidad para ofrecer nuestros respetos a uno de los académicos que mayor influencia ha tenido en la formación y en la construcción de la Facultad de Ciencias Físico-Matemáticas de la Benemérita Universidad Autónoma de Puebla. Me refiero a nuestro querido amigo, profesor y colega Agustín Contreras Carreto.

Tuve el gusto de conocer a Agustín en el año de 1996, cuando desarrollamos un proyecto de trabajo conjunto con los Profesores Juan Angoa, Manuel Ibarra y Armando Martínez. Este evento que narraría Agustín como “Una cálida tarde de verano en Huejotzingo, una espumosa y fragante sidra, y cinco matemáticos disfrutando de la bebida y del paisaje. Con estos ingredientes y una voluntad común, la alegre conversación sólo podía enjendrar algo bueno”. Esta exitosa confabulación consistió en un trabajo que iniciamos los cinco para crear un grupo de estudio, docencia, difusión e investigación en topología general, e incluyó mi tutoría para un trabajo doctoral que llevó a cabo nuestro querido homenajeado.

Así se inició mi relación académica y personal con Agustín. Tuve desde aquel entonces la oportunidad de trabajar y de convivir con una personalidad única que se caracteriza por su responsabilidad, generosidad, y su rigor y talento matemático. Agustín es el matemático-poeta. Su gran sensibilidad humana, su vocación por el trabajo colectivo², su preocupación por lo justo, su calidez y determinación por entregar a sus alumnos y colegas su conocimiento y dedicación, construyeron una trayectoria académica que ha marcado a varias generaciones de estudiantes y académicos de la FCFM-BUAP. Sus clases, seminarios, libros, artículos y conferencias de divulgación son una muestra del trabajo en equipo, y de la conjunción de lo riguroso-matemático con la calidez humana que toma diversas formas amalgamadas por el buen humor omnipresente en Agustín, su sensibilidad por la música y la poesía, y su crítica a todo lo que significa injusticia.

¹Con estos versos dio término Agustín a su tesis doctoral.

²“La sopa de piedras”, como diría el Dr. Oscar Falcón, Profesor de la Facultad de Ciencias, UNAM, a quien siempre tendremos en la memoria (?-2002)

¿Quién puede olvidar el espíritu amable y divertido que enmarcó las primeras Jornadas Veraniegas de Topología en 1996? Éstas fueron organizadas por este grupo de los *Cinco de Huejotzingo* que estaban constituidas por una serie de conferencias anunciadas siempre por medio de dibujos memorables que diseñaba Agustín. Estas imágenes reflejaban de manera asombrosa algunas de las características de la personalidad tanto de los conferencistas como de los escuchas. En mi opinión, esas Primeras Jornadas fueron el cimiento del desarrollo académico posterior de la topología en la FCFM.

El trabajo doctoral de Agustín es una conjunción de dos objetivos; por un lado logra construir una síntesis preciosamente concretada de las teorías categóricas de los E -espacios, sus extensiones maximales compactas, y de los espacios de funciones continuas $C_p(X, E)$; y por otro lado, resuelve varios problemas relacionados a la pseudocompacidad de espacios del tipo $C_p(X, 2)$. Menciono aquí sólo uno de sus resultados sobresalientes en esta tesis: *Si X es un espacio cero-dimensional y normal, entonces $C_p(X, 2)$ es σ -compacto si y sólo si X es Eberlein-Grothendieck y el conjunto de puntos no aislados de X es compacto.* Queda aún pendiente la edición de la tesis doctoral de Agustín como libro de texto que servirá mucho, tanto a estudiantes de posgrado como a profesores e investigadores en topología.

Termino aquí mi insuficiente homenaje a mi querido compañero Agustín, quien tanto me ha enseñado como matemático y persona. Y junto a sus alumnos, colegas, y sobre todo junto a su queridísima Hortensia, que siempre estará presente, y a sus amadísimos hijos, me sumo a este muy merecido homenaje. Mil felicidades y muchas gracias, Agustín.

Ángel Tamariz Mascarúa

Al Profe Contri

El profesor Agustín Contreras, mejor conocido como el «Profe Contri» es uno de los mejores profesores de la FCFM. Lo vi unas veces antes de entrar a la facultad y me parecía todo un personaje: su bolsa de mandado, sus pelitos en la orejas y su peculiar forma de ser. Cuando recién entré a la universidad, tuve la fortuna de tomar clases con él en el segundo semestre: la materia de geometrías no euclidianas. En ese momento, me enamoré de la geometría.

El Profe Contri, hace sus clases súper divertidas e interesantes. No sólo se aprenden matemáticas sino también la historia detrás de lo que nos está enseñando, haciendo gala de sus dotes histriónicos al encarnar al mismísimo Hipaso o imitar con maestría inigualable la voz de Gauss, entre otros.

Después de tomar la primera materia con él, ya no pude dejar de tenerlo como profesor: fue mi maestro de Geometría Sintética (no puede uno hacerse licenciado en Matemáticas sin haber tomado esa materia con él, es una joya: actúa la historia del intento de la demostración del quinto postulado, del descubrimiento de la incommensurabilidad de los números racionales, etc.; siempre enfatiza esa parte de la historia que todo matemático debe saber) y en otras siete materias.

Cuando conocí otra faceta del Profe Contri fue en la maestría, donde no sólo fue mi asesor, sino mi amigo y mi «médico homeópata» de cabecera. Esos dos años y medio fueron maravillosos, aprendimos juntos cosas y me presentó el mundo de las categorías con una emoción indescriptible. Las clases eran de todo, empezábamos a estudiar homología y terminábamos hablando de música, de política, de historia, de su vida, de sus maestros (con sus voces originales)...

Estas breves palabras no son suficientes ni para captar la esencia de tan maravilloso personaje ni para agradecerle lo que ha hecho por mí y por la facultad, que le debe su inusual amor por las matemáticas y su enseñanza.

María del Rocío Macías Prado

Generosidad

Cuántas y tantas almas podemos decir: “Agustín Contreras fue mi maestro”. No lo sé. Puedo hablar de una experiencia personal, del cambio que significó en mi vida ser su alumno.

Sus clases son todo un acontecimiento, una verdadera fiesta. Todo se transforma, ya no son obligaciones o deberes, se trata de auténtico goce, sumergirse en aguas termales. ¡El paraíso terrenal!

Con el Maestro llegó la geometría en medio de una gran algarabía, con música y danzas. Después de conocerlo uno queda convencido de que debiera celebrarse “El día del quinto postulado de Euclides”, ¡oh maravilla de la transfiguración! Si lo aceptamos, la geometría es una tierna y dulce doncella, heroína del romanticismo (parece que escuchamos “La muerte y la doncella” de Schubert); si no, se convierte en una exuberante y voluptuosa mujer de la narrativa latinoamericana (oiga usted cómo suena el bongó).

Cuántos jóvenes no habrán encontrado su vocación de matemáticos, de geómetras, de topólogos, de artistas, en sus clases.

Para mí significó salir del medioevo y entrar a la Ilustración. El salón se hizo ancho, ancho extendiéndose por días, semanas, meses . . . , descubrimiento de lo mejor del cultivo de los hombres. Euclides, Lobachevski, Mozart, Beethoven, Britten, Lara, Toñita, Escher, van Gogh, Hendrix, Dylan, José Alfredo, Maupassant, (Mapzapán, le decía “el Juan”), Eco, Uderzo y Goscinnny, Pillimba . . . todos son uno. Con Don Agustín queda demostrado que la enseñanza es una realidad absolutamente humana. ¿Qué educación a distancia o universidad en línea puede lograr esto?

Hay que dejar bien claro que nuestro ilustre amigo no conoce de “productividad”, sino de fecundidad. No existe en él un deseo desenfrenado de acumular conocimientos a como dé lugar, Agustín solo sabe de disfrute en la alegría. Está muy lejos de los que presumen de algo. Él es generoso, para recibir y para dar. Como el príncipe Mishkin siempre está dispuesto, abierto en su sencillez, para ayudar a cualquiera que se acerca.

Hoy la burocracia gusta de calificar a los profesores por sus “puntos”, pero nosotros, los que lo conocemos, sabemos que en él hay una especial capacidad de enamorar a los jóvenes de la matemática. Modelando una mujer de una costilla, como cortés ninfagogo, o amigo, la conduce y se la presenta al novio. El alumno prorrumpe en canto . . .

La matemática, con él, se hace arte sin dejar de ser ciencia.

Pero miro espantado una enorme silueta que se dibuja a través de los vidrios de la

puerta de mi departamento, aun así la abro, camino hacia atrás . . . es él, ejerce sobre mí una poderosa fascinación, es una extraña combinación de emociones, sorpresa, miedo, ¡jalegría! Está dotado de una siniestra virtud por medio de su voz que es una especie de chillido y lamento, y al mismo tiempo como un canto de ternura que te invita a imitarlo. ¡El monstruo dada!

El maestro Contri (de cariño) puede dirigir un curso del grado que sea, desde el básico hasta doctorado, lo mismo que el coro de una parvada de guajolotes tlaxcaltecas o al Cuarteto Guarneri interpretando el Razumovsky número 3.

Bueno, basta. Podría seguir al infinito cual canción para coro de niños como la “Old Abram Brown” de Britten.

Choquemos, pues, nuestros vasos como cosacos empedernidos a la salud de nuestro insigne amigo; brindemos por todos los que hemos tomado sus cursos, por todos los que los tomarán, pero sobre todo por los que tenemos la fortuna incalculable de ser sus amigos.

Muchas gracias Maestro Agustín, le estaré eternamente agradecido.

La Rayita, uno que fue alumno.

Caricaturas 1996

*realizadas por el amanuense Agustín Contreras Carreto
para anunciar las sesiones de las primeras Jornadas de
Topología*

¡NO FALTES A LAS

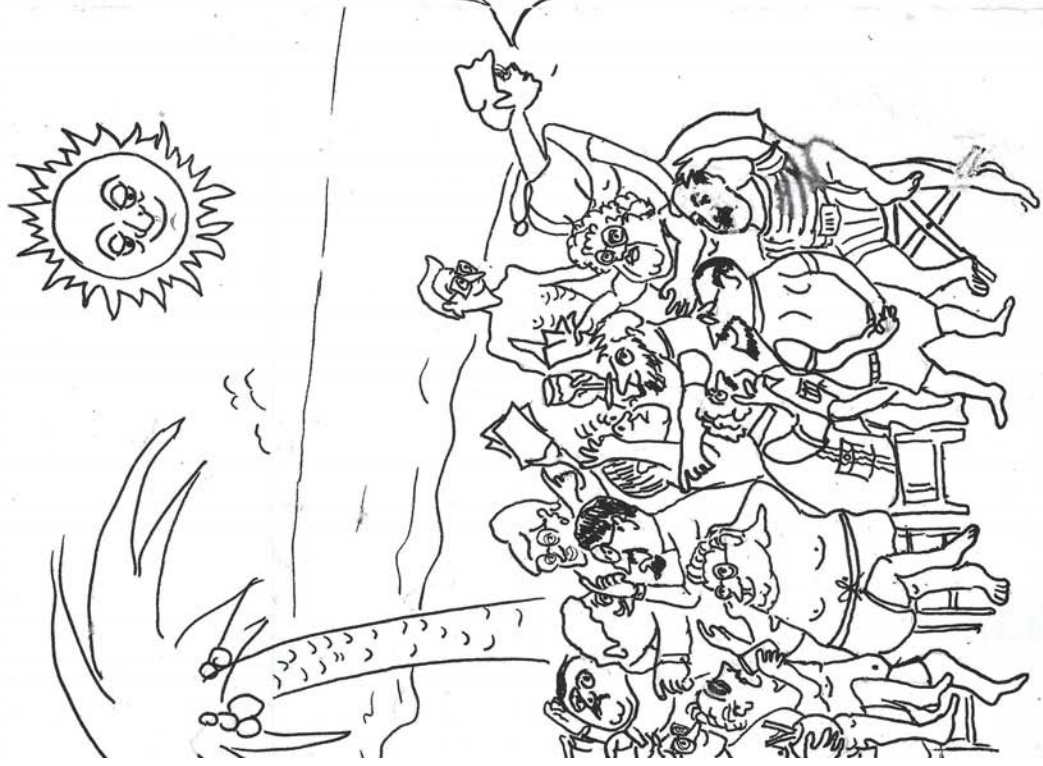
JORNADAS VERANIEGAS DE TOPOLOGÍA!

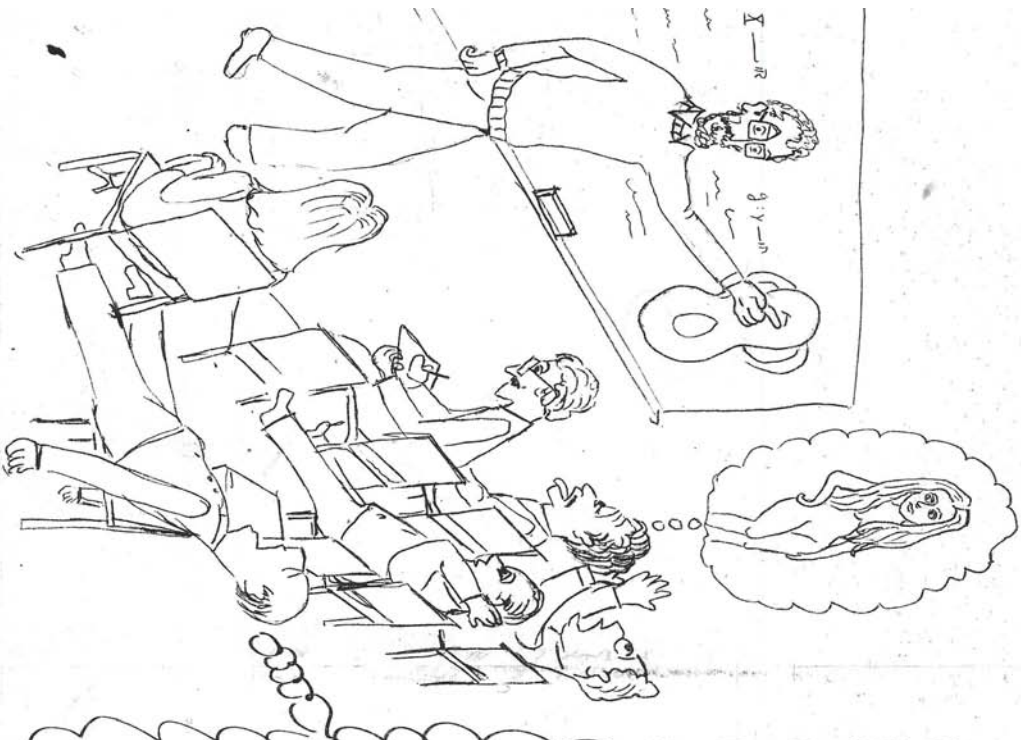
Seminario de espacios $C(X)$ de funciones continuas

Dirigido por Angel Tamariz (U.NAM.)

1ª SESIÓN : JUEVES 30 de Mayo
en el edificio 8 de la F.C.F.M.
a las 12 Horas (sala de conferencias)

Tema: Visión general
(de los espacios $C(X)$)





JORNADAS VERANIEGAS DE TOPOLOGÍA, Seminario de espacios $C(X)$ de funciones continuas

Dirigido por ANGEL TAMARIZ (UNAM.)

2ª SESIÓN: Jueves 6 de Junio (de corpus)
a las 12 Hrs. en la sala de conferencias
del edificio 8 de la F.C.F.M.
¡estará soñado! No faltes

...Y este jueves 13 de Junio, aparte de la 3^a sesión del

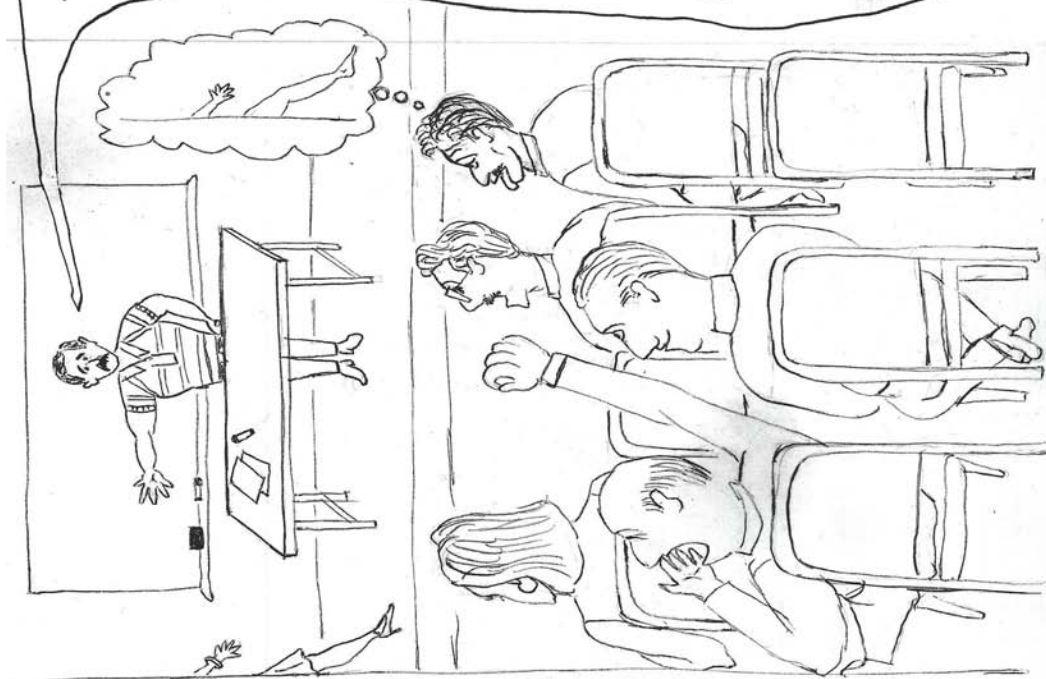
Seminario de espacios C (S)

dirigido por Angel Tamariz Mascarúa y que comenzará a las 12:00 Hrs, tendremos también la presencia de la

Dra. Isabel Puga (U.N.A.M.) quien nos brindará la conferencia:

**"Una introducción a los
continuos"**

a las 4 P.M. en la sala de conferencias del edificio 8 de la F.C.F.M.



Vamos este VIERNES 21 DE JUNIO
a la 4ª sesión del

Seminario de espacios

C(X)

dirigido por ANGEL TAMARIZ (UNAM),
que comenzará a las 12:00 HRS,
pues se presentará también la

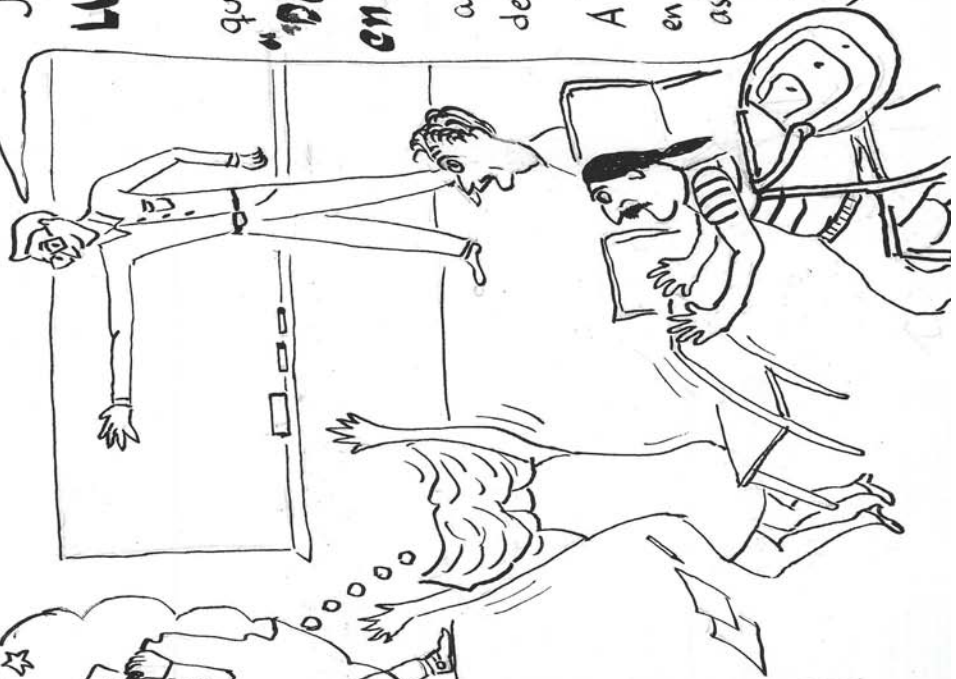
Dra. Silvia de Neymet (UNAM)

para brindarnos la conferencia:

“Teoría de retratos”

a las 4 P.M. en la Sala de
Conferencias del edificio 8 de la F.C.F.M.
(entrada y salida libres)





Toca ahora su turno, este
JUEVES 27 DE JUNIO, a
LUIS MIGUEL VÍLEGAS
 (U.N.A.M.)
 que nos brindará la conferencia
**"Descomposición de espacios
 en subespacios densos ajenos"**
 a las 4 P.M. en la Sala de Conferencia
 del edificio 8 de la F.C.F.M.
 A las 12 horas de ese mismo día y
 en ese mismo lugar, no se olviden de
 asistir a la 5ª sesión del
**Seminario de espacios
 (S)** dirigido por ANGEL TAMARIZ M.

Este Viernes 5 de Julio, a las 12 Hrs.,
asista Ud. a la 6a. Sesión del
**Seminario de espacios $C(X)$
de funciones continuas**
dirigido por ANGEL TAMARIZ MASCARUA
en el salón de conferencias del edificio 8
de la F.C.F.M.

Además, a las 4 P.M. del mismo día y
en el mismo lugar, se presentará el célebre

Dr. Oleg Okunev (U.N.A.M.)
quien dictará la conferencia:

"On Free Topological Groups"



JORNADAS VERANIEGAS DE TOPOLOGÍA

Presentan, este VIÉRNES 12 de Julio :

A las 12:00 Hrs. , la 7ª sesión del
**Seminario de Espacios $C(\Sigma)$ y $C_p(\Sigma)$
de funciones continuas**

dirigido por ANGEL TAMARIZ MASCARÚA (U.N.A.M.)
en el edificio 4 de la F.C.F.M.

Y,

A las 4 P.M. , al tan esperado

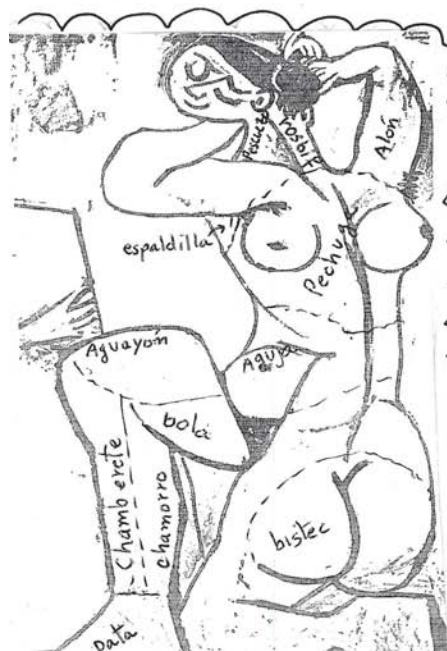
Dr. ADALBERTO GARCÍA-MAYNEZ (U.N.A.M.)

quien dictara' la conferencia

AXIOMAS DE SEPARACIÓN EN TOPOLOGÍA

en el salón de conferencias
del edificio 8 de la F.C.F.M.

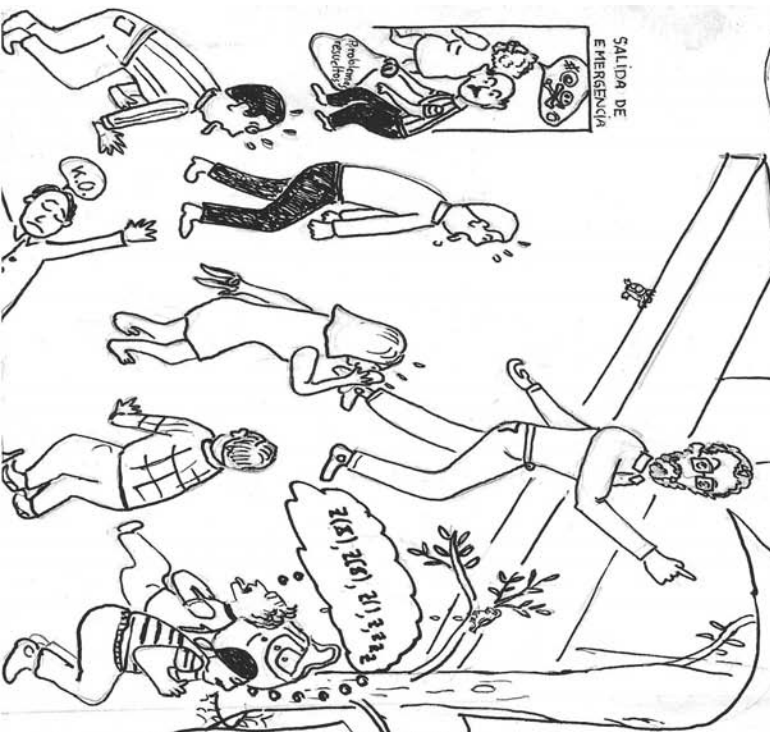
¡NO FALTES!



¿Cómo separar
en espacios
topológicos ?

¡Qué interesante!

Adiós a las **TORNADES**
VERANIEGAS DE
Topoteia



¡ÁNIMO!

No dejes de asistir a
la 8ª. y última sesión del

Seminario de espacios **C(S)**

dirigido por Angel Tamariz M.,
este Viernes 19 de Julio a
las 11:00 hrs. en
algún salón del edificio 4
de la F.C.F.M.

Álgebra

Capítulo 1

El Anillo de Burnside

César Cejudo Castilla, Juan Manuel Ramírez Contreras,
David Villa Hernández
FCFM, BUAP

Resumen

En este trabajo realizaremos una construcción del anillo de Burnside de un grupo finito G . Esta construcción se hará de manera detallada para una mejor comprensión de dicho anillo y con ello, familiarizar al lector con temas relacionados. Cabe mencionar que para una mejor comprensión de este texto se requieren conocimientos básicos de la Teoría de Grupos.

1 G -conjuntos

El anillo de Burnside $B(G)$ es uno de los anillos fundamentales de representación de G . Además, es el objeto universal a considerar en el estudio de la categoría de G -conjuntos finitos. Es un análogo para los G -conjuntos finitos del anillo \mathbb{Z} ; de hecho, el anillo \mathbb{Z} es isomorfo al anillo de Burnside del grupo trivial.

Por otra parte, el estudio de su espectro primo y de sus idempotentes primitivos conduce a varios teoremas de inducción como el de Artin, Brauer y Dress (teoremas de representación). El anillo de Burnside es un invariante del grupo, el cual detecta solubilidad. Andreas Dress, probó que un grupo es soluble si y sólo si los únicos idempotentes de $B(G)$ son los triviales. También probó el teorema de inducción de Dress para $B(G)$, el cual, a partir del hecho de que $B(G)$ actúa en los funtores de Mackey, se tradujo en un teorema de inducción para cada uno de estos funtores, ver [2]. El lector puede encontrar más aplicaciones en [1], [2] y [3].

Definición 1.1. Sean X un conjunto y G un grupo, diremos que X es un G -conjunto si existe una función

$$* : G \times X \longrightarrow X \text{ tal que } (g, x) \longmapsto *((g, x)) = g * x$$

que satisface:

- i) $e * x = x$ para todo $x \in X$, donde e es la identidad en G .

ii) $(gh) * x = g * (h * x)$ para todo $g, h \in G$ y $x \in X$.

Observación 1.2. A la función $*$ usualmente se le conoce como la *acción* de G en X y se dice que G *actúa* en X .

Ejemplo 1.3. Sean G un grupo y X un conjunto, entonces X es un G -conjunto mediante:

$$* : G \times X \longrightarrow X \text{ tal que } (g, x) \longmapsto g * x = x.$$

La acción $*$ es llamada la *acción trivial* de G en X , y cuando consideramos dicha acción decimos que G actúa trivialmente en X .

Ejemplo 1.4. Sea G un grupo y $H \leq G$, entonces G/H (el conjunto de las clases laterales izquierdas de H) es un G -conjunto mediante la multiplicación por la izquierda, es decir:

$$* : G \times G/H \longrightarrow G/H \text{ tal que } (g, g_1H) \longmapsto g * (g_1H) = (gg_1)H$$

para todo $g \in G$ y $g_1H \in G/H$.

Definición 1.5. Sean X y Y conjuntos. Considere los conjuntos $X' = X \times \{1\}$ y $Y' = Y \times \{2\}$ donde $1 \notin Y$ y $2 \notin X$, entonces definimos la unión ajena de X y Y como $X \sqcup Y = X' \cup Y'$. Note que $X' \cap Y' = \emptyset$.

De manera más general, dada $\{X_i\}_{i \in I}$ una familia de conjuntos, la unión ajena de dicha familia es $\bigsqcup_{i \in I} X_i = \bigcup_{i \in I} X'_i$, donde $X'_i = X_i \times \{i\}$.

Ejemplo 1.6. Si $\{X_i\}_{i \in I}$ es una familia de G -conjuntos, entonces $\bigsqcup_{i \in I} X_i$ es un G -conjunto mediante:

$$* : G \times \bigsqcup_{i \in I} X_i \longrightarrow \bigsqcup_{i \in I} X_i \text{ tal que } (g, (x, i)) \longmapsto g * (x, i) = (g *_i x, i)$$

donde $*_i$ es la acción de G en X_i con $x \in X_i$, para algún $i \in I$.

Ejemplo 1.7. Si $\{X_i\}_{i \in I}$ es una familia de G -conjuntos, entonces el producto cartesiano de la familia $\{X_i\}_{i \in I}$ (denotado por $\prod_{i \in I} X_i$) es un G -conjunto mediante:

$$* : G \times \prod_{i \in I} X_i \longrightarrow \prod_{i \in I} X_i \text{ tal que } (g, (x_i)_{i \in I}) \longmapsto g * (x_i)_{i \in I} = (g *_i x_i)_{i \in I}$$

donde $*_i$ es la acción de G en X_i , para cada $i \in I$.

Ejemplo 1.8. Si X y Y son dos G -conjuntos, entonces

$$\text{Hom}(X, Y) = \{f : X \rightarrow Y \mid f \text{ es función}\}$$

es un G -conjunto mediante:

$$* : G \times \text{Hom}(X, Y) \longrightarrow \text{Hom}(X, Y) \text{ tal que } (g, f) \longmapsto g * f$$

donde $(g * f)(x) = g *_{Y} f(g^{-1} *_{X} x)$, para toda $x \in X$, $f \in \text{Hom}(X, Y)$ y $g \in G$. Es claro que $*$ está bien definida, puesto que f es una función y $*_{Y}, *_{X}$ son las acciones de G en Y y X , respectivamente.

Demostremos que $*$ es una acción.

i) Sea $f \in \text{Hom}(X, Y)$, entonces $(e * f)(x) = e *_{Y} f(e *_{X} x) = f(x)$ para toda $x \in X$, por lo tanto $e * f = f$.

ii) Sean $g, g' \in G$. Entonces para $x \in X$ se tiene que:

$$\begin{aligned} ((g'g) * f)(x) &= (g'g) *_{Y} [f(g^{-1}g'^{-1} *_{X} x)] \\ &= g' *_{Y} [g *_{Y} f(g^{-1}g'^{-1} *_{X} x)] \\ &= g' *_{Y} [(g * f)(g'^{-1} *_{X} x)] \\ &= (g' * (g * f))(x). \end{aligned}$$

Por lo tanto $(g'g) * f = g' * (g * f)$.

Observación 1.9. Dada $*$ una acción de G en X se tiene que \sim es una relación de equivalencia en X , donde $y \sim x$ si y sólo si existe $g \in G$ tal que $y = g * x$.

Definición 1.10. De acuerdo a la relación anterior \sim , denotaremos a la clase de equivalencia de $x \in X$ por $\mathcal{O}_G(x)$ y la llamaremos *órbita* de x en G . Observe que $\mathcal{O}_G(x) = \{g * x : g \in G\} \subseteq X$.

Definición 1.11. Sean X, Y dos G -conjuntos con acciones $*_{X}$ y $*_{Y}$, respectivamente. Diremos que $f \in \text{Hom}(X, Y)$ es un morfismo de G -conjuntos si satisface que $f(g *_{X} x) = g *_{Y} f(x)$ para todo $g \in G$ y $x \in X$. Denotamos a la colección de morfismos de G -conjuntos de X a Y por $\text{Hom}_G(X, Y)$.

Ejemplo 1.12. Sean $H \leq K \leq G$, entonces

$$\Pi : G/H \longrightarrow G/K \text{ tal que } aH \longmapsto aK$$

es un morfismo sobreyectivo de G -conjuntos.

- i) Veamos que Π esta bien definida. Sean $aH, a'H \in G/H$ tales que $aH = a'H$, así $(a')^{-1}a \in H \leq K$, en consecuencia $aK = a'K$.
- ii) Veamos que Π es un morfismo de G -conjuntos. Primero, notamos que la acción de G en G/H y en G/K es la definida en el Ejemplo 1.4, así

$$\Pi(g * aH) = \Pi((ga)H) = (ga)K = g * (aK) = g * \Pi(aH).$$

Ejemplo 1.13. Sea X un G -conjunto, la *identidad*

$$Id_X : X \longrightarrow X \text{ tal que } x \longmapsto x$$

es un morfismo de G -conjuntos.

Observación 1.14.

- i) La composición de morfismos de G -conjuntos es nuevamente un morfismo de G -conjuntos.
- ii) El inverso de un isomorfismo de G -conjuntos es también un isomorfismo de G -conjuntos.
- iii) Un morfismo de G -conjuntos manda órbitas en órbitas.

Sean $f \in \text{Hom}_G(X, Y)$, $x \in X$ y $\mathcal{O}_G(x)$ la órbita de x en G . Luego, $f(gx) = g(f(x))$, así $f(\mathcal{O}_G(x)) \subset \mathcal{O}_G(f(x))$. Por otro lado, $gf(x) = f(gx) \in f(\mathcal{O}_G(x))$, por tanto $\mathcal{O}_G(f(x)) \subset f(\mathcal{O}_G(x))$, concluyendo que $\mathcal{O}_G(f(x)) = f(\mathcal{O}_G(x))$.

Definición 1.15. Sean X, Y dos G -conjuntos, diremos que X y Y son isomorfos como G -conjuntos si existe $f \in \text{Hom}_G(X, Y)$ biyectivo. Si X y Y son isomorfos lo denotaremos por $X \cong_G Y$.

Definición 1.16. Sean G un grupo, X un G -conjunto y $x \in X$. Definimos el *estabilizador* de x en G como $\text{stab}_G(x) = \{g \in G : g * x = x\}$.

Observación 1.17. Note que $\text{stab}_G(x)$ es un subgrupo de G .

- i) Sea e la identidad en G , sabemos que $e * x = x$, entonces $e \in \text{stab}_G(x)$.
- ii) Sean $g_1, g_2 \in \text{stab}_G(x)$, sabemos que $(g_1g_2) * x = g_1 * (g_2 * x) = x$, entonces $g_1g_2 \in \text{stab}_G(x)$.
- iii) Sea $g \in \text{stab}_G(x)$, luego $g * x = x$, así $x = g^{-1} * x$, entonces $g^{-1} \in \text{stab}_G(x)$.

De aquí en adelante si no hay lugar a confusión, vamos a abreviar la notación y escribiremos gx en lugar de $g * x$.

Definición 1.18. Llamaremos a un G -conjunto *transitivo* si tiene una única órbita; es decir, si para cada $x, y \in X$ existe $g \in G$ tal que $y = gx$.

Teorema 1.19. Si G es un grupo, X un G -conjunto y $x_0 \in X$, entonces

$$i) \text{ stab}_G(gx_0) = g(\text{stab}_G(x_0))g^{-1}.$$

$$ii) \mathcal{O}_G(x_0) \cong_G G/\text{stab}_G(x_0).$$

iii) Si H es un subgrupo de G , entonces G/H es transitivo.

iv) Si X es un G -conjunto transitivo, entonces $X \cong_G G/H$ para algún $H \leq G$.

Demostración. *i)* Sea $a \in \text{stab}_G(gx_0)$, así $a(gx_0) = gx_0$, de aquí $(ag)x_0 = gx_0$, en consecuencia $g^{-1}agx_0 = x_0$, de aquí que $g^{-1}ag \in \text{stab}_G(x_0)$, con lo que $a \in g(\text{stab}_G(x_0))g^{-1}$, por tanto $\text{stab}_G(gx_0) \subset g(\text{stab}_G(x_0))g^{-1}$. De manera análoga $g(\text{stab}_G(x_0))g^{-1} \subset \text{stab}_G(gx_0)$, por lo tanto

$$\text{stab}_G(gx_0) = g(\text{stab}_G(x_0))g^{-1}.$$

ii) Definimos $f : \mathcal{O}_G(x_0) \rightarrow G/\text{stab}_G(x_0)$ tal que $gx_0 \mapsto g(\text{stab}_G(x_0))$.

Sean $g_1x_0, g_2x_0 \in \mathcal{O}_G(x_0)$. Entonces $g_1x_0 = g_2x_0$ si y sólo si $(g_2^{-1}g_1)x_0 = x_0$, esto último es equivalente a que $g_2^{-1}g_1 \in \text{stab}_G(x_0)$ y esto pasa si y sólo si $g_2(\text{stab}_G(x_0)) = g_1(\text{stab}_G(x_0))$. Por lo tanto, f está bien definida y es inyectiva.

Ahora, sea $g(\text{stab}_G(x_0)) \in G/\text{stab}_G(x_0)$, notemos que $gx_0 \in \mathcal{O}_G(x_0)$ y es tal que $f(gx_0) = g(\text{stab}_G(x_0))$, así f es suprayectiva.

Sólo falta demostrar que f es morfismo de G -conjuntos, para ello tomemos $g, g' \in G$ entonces

$$f(g'(gx_0)) = f((g'g)x_0) = g'g(\text{stab}_G(x_0)) = g'(g(\text{stab}_G(x_0))) = g'(f(gx_0)).$$

Por lo tanto $\mathcal{O}_G(x_0) \cong_G G/\text{stab}_G(x_0)$.

iii) Como G actúa en G/H mediante la acción del Ejemplo 1.4, tenemos que $G/H = \mathcal{O}_G(H)$. Por lo tanto, G/H es transitivo.

iv) Es inmediato de *ii)*. □

Teorema 1.20. Si X es un G -conjunto finito, entonces existe un conjunto de índices I tal que

$$X \cong_G \bigsqcup_{i \in I} G/H_i \text{ donde } H_i \leq G \text{ para cada } i \in I.$$

Demostración. Sean X un G -conjunto y $\{x_i \mid i \in I\}$ un conjunto de representantes de las órbitas de G en X . Entonces $X = \bigcup_{i \in I} \mathcal{O}_G(x_i)$ y por el Teorema 1.19 sabemos que $\mathcal{O}_G(x_i) \cong_G G/\text{stab}_G(x_i)$ para cada $i \in I$; llamemos f_i al isomorfismo de $\mathcal{O}_G(x_i)$ a $G/\text{stab}_G(x_i)$. Así, definimos

$$f : \bigcup_{i \in I} \mathcal{O}_G(x_i) \longrightarrow \bigsqcup_{i \in I} (G/\text{stab}_G(x_i)) \text{ tal que } x \longmapsto (f_i(x), i) \text{ para } x \in \mathcal{O}_G(x_i)$$

el cual es un isomorfismo de G -conjuntos. Finalmente definimos $H_i = \text{stab}_G(x_i)$, con lo cual queda demostrado el teorema. \square

Definición 1.21. Sea $H \leq G$ un subgrupo, llamaremos *conjugado* de H por $g \in G$ al subgrupo de G definido por $gHg^{-1} = \{ghg^{-1} : h \in H\}$.

Observación 1.22. Sea $X = \{H \subseteq G : H \leq G \text{ es subgrupo}\}$, entonces X es un G -conjunto mediante

$$* : G \times X \longrightarrow X \text{ tal que } (g, H) \longmapsto g * H = gHg^{-1}$$

$$i) \quad e * H = eHe^{-1} = H.$$

$$ii) \quad g_1g_2 * H = (g_1g_2)H(g_1g_2)^{-1} = g_1g_2Hg_2^{-1}g_1^{-1} = g_1 * (g_2Hg_2^{-1}) = g_1 * (g_2 * H).$$

Definición 1.23. Al conjunto de órbitas de X bajo la acción de G (ver Observación 1.22) lo denotaremos $\mathcal{C}(G)$ y lo llamaremos las clases de conjugación de subgrupos de G . Denotamos por $[H] \in \mathcal{C}(G)$ la órbita de $H \in X$ bajo esta acción, es decir: $[H] = \{gHg^{-1} : g \in G\}$. Además, $[K] = [H]$ si y sólo si $K = gHg^{-1}$ para algún $g \in G$, es decir, H y K son subgrupos conjugados.

Lema 1.24. Si $H, K \leq G$ son subgrupos, entonces $G/H \cong_G G/K$ si y sólo si $[H] = [K]$.

Demostración. (\Leftarrow) Sabemos que $[K] = [H]$, entonces existe $g \in G$ tal que $K = gHg^{-1}$, así

$$g^{-1}Kg = H.$$

Luego, definimos $f : G/H \longrightarrow G/K$ tal que $aH \longmapsto ag^{-1}K$. Sean $aH, bH \in G/H$. Entonces

$$\begin{aligned} aH = bH &\Leftrightarrow ag^{-1}Kg = bg^{-1}Kg \\ &\Leftrightarrow ag^{-1}K = bg^{-1}K. \end{aligned}$$

Por lo tanto f está bien definida y es inyectiva.

Ahora, sea $bK \in G/K$ con $b \in G$, notemos que $bgH \in G/H$, luego $f(bgH) = bbg^{-1}K = bK$. De aquí, f es suprayectiva y, por tanto, biyectiva.

Por último, $f(g_1(aH)) = f(g_1aH) = (g_1ag^{-1})K = g_1(ag^{-1})K = g_1f(aH)$, con lo que f es morfismo de G -conjuntos. Por lo tanto, $G/H \cong_G G/K$.

(\Rightarrow) Sea $f : G/H \rightarrow G/K$ un isomorfismo de G -conjuntos tal que $H \mapsto gK$ para algún $g \in G$. Así, para todo $h \in H$, $gK = f(H) = f(hH) = hgK$, así $gK = hgK$, luego $g^{-1}hg \in K$, de aquí $g^{-1}Hg \subset K$.

Tomemos f^{-1} , el cual sabemos que es un morfismo de G -conjuntos y note que para cada k en K , $kg^{-1}H = kf^{-1}(K) = f^{-1}(kK) = f^{-1}(K) = g^{-1}H$, así $gkg^{-1} \in H$ para todo $k \in K$, luego $gKg^{-1} \subset H$, en consecuencia $K \subset g^{-1}Hg$, por tanto $K = g^{-1}Hg$, por lo tanto $[H] = [K]$. \square

Definición 1.25. Sean $H, K \leq G$ subgrupos. Si existe $g \in G$ tal que $gHg^{-1} \subseteq K$, diremos que H es *subconjugado* de K .

Teorema 1.26. Si $H, K \leq G$ son subgrupos, entonces $\text{Hom}_G(G/H, G/K) \neq \emptyset$ si y sólo si H es subconjugado de K .

Demostración. (\Rightarrow) Sea $f \in \text{Hom}_G(G/H, G/K)$, entonces $f(eH) = aK$ para alguna $a \in G$. Para cada $h \in H$ tenemos que $hH = H$, entonces se tiene que $f(hH) = f(eH) = aK$, lo anterior implica que $aK = f(hH) = hf(eH) = haK$. Por lo tanto $(a^{-1})ha \in K$, y así tenemos que $(a^{-1})Ha \subset K$.

(\Leftarrow) Sea $\alpha : G/H \rightarrow G/a^{-1}Ha$ un isomorfismo de G -conjuntos. Por otro lado, existe $a \in G$ tal que $a^{-1}Ha \leq K$, entonces del Ejemplo 1.12 tenemos el morfismo de G -conjuntos $\Pi : G/a^{-1}Ha \rightarrow G/K$. En consecuencia

$$(\Pi \circ \alpha) \in \text{Hom}_G(G/H, G/K). \quad \square$$

Observación 1.27. Si H es subconjugado de K , entonces todo conjugado de H es subconjugado de cualquier conjugado de K .

Observación 1.28. En $\mathcal{C}(G)$ hay una relación de orden parcial. Dados $[H], [K] \in \mathcal{C}(G)$, tenemos que

$$[H] \preceq [K] \text{ si y sólo si } H \text{ es subconjugado de } K.$$

2 Propiedades de la marca

Definición 2.1. Sean G un grupo, $H \leq G$ un subgrupo y X un G -conjunto finito. Definimos X^H como el conjunto de todos los puntos de X que quedan fijos bajo la acción de H ; es decir,

$$X^H = \{x \in X : h * x = x \text{ para todo } h \in H\}.$$

Definimos la *Marca* de H en X como el número de elementos de X^H y la denotamos por

$$\varphi_H(X) = |X^H|.$$

Teorema 2.2. Si $H, K \leq G$ son subgrupos y X, Y son G -conjuntos finitos, entonces

i) $\varphi_H(X \sqcup Y) = \varphi_H(X) + \varphi_H(Y)$.

ii) $\varphi_H(X \times Y) = \varphi_H(X)\varphi_H(Y)$.

iii) Si $[H] = [K]$, entonces $\varphi_H(X) = \varphi_K(X)$ para todo G -conjunto finito X .

Demostración. i)

$$\begin{aligned} (X \sqcup Y)^H &= \{z \in X \sqcup Y : h * z = z \ \forall \ h \in H\} \\ &= \{z : (z \in X \times \{1\} \vee z \in Y \times \{2\}) \wedge h * z = z \ \forall \ h \in H\} \\ &= \{z : (z \in X \times \{1\} \wedge h * z = z) \vee (z \in Y \times \{2\} \wedge h * z = z) \ \forall \ h \in H\} \\ &= \{z \in X \times \{1\} \vee z \in Y \times \{2\} : h * z = z \ \forall \ h \in H\} \\ &= X^H \sqcup Y^H. \end{aligned}$$

Por lo tanto $|(X \sqcup Y)^H| = |X^H \sqcup Y^H| = |X^H| + |Y^H|$.

ii)

$$\begin{aligned} (X \times Y)^H &= \{(x, y) \in X \times Y : h * (x, y) = (x, y) \ \forall \ h \in H\} \\ &= \{(x, y) \in X \times Y : (h * x, h * y) = (x, y) \ \forall \ h \in H\} \\ &= \{(x, y) : x \in X^H \text{ y } y \in Y^H \ \forall \ h \in H\} \\ &= X^H \times Y^H. \end{aligned}$$

Por lo tanto $|(X \times Y)^H| = |X^H \times Y^H| = |X^H| |Y^H|$.

iii) Sea X un G -conjunto. Supongamos que $[H] = [K]$, esto implica que existe $g \in G$ tal que $H = gKg^{-1}$, entonces tenemos que:

$$\begin{aligned} X^H &= \{x \in X : h * x = x \ \forall \ h \in H\} \\ &= \{x \in X : gkg^{-1} * x = x \ \forall \ k \in K\} \\ &= \{x \in X : k(g^{-1} * x) = g^{-1} * x \ \forall \ k \in K\} \\ &= \{x \in X : g^{-1} * x \in X^K\} \\ &= \{x \in X : x \in gX^K\} \\ &= gX^K. \end{aligned}$$

Del párrafo anterior tenemos que $|X^H| = |gX^K|$.

Ahora definimos $\psi : X^K \rightarrow gX^K$ tal que $x \mapsto gx$. Observe que, ψ es una biyección, por lo que $|X^K| = |gX^K|$, y con ello se tiene que $|X^H| = |X^K|$. Por lo tanto, para todo G -conjunto X tenemos que $\varphi_H(X) = \varphi_K(X)$. \square

Lema 2.3. Sean $H, K \leq G$ subgrupos, entonces existe una biyección entre

$$(G/K)^H \text{ y } \text{Hom}_G(G/H, G/K).$$

Demostración. i) Supongamos que H no es subconjugado de K , entonces por el Teorema 1.26

$$\text{Hom}_G(G/H, G/K) = \emptyset.$$

Por otro lado tenemos que:

$$\begin{aligned} (G/K)^H &= \{gK \in G/K : (hg)K = gK \text{ para todo } h \in H\} \\ &= \{gK \in G/K : g^{-1}hgK = K \text{ para todo } h \in H\} \\ &= \{gK \in G/K : g^{-1}Hg \subset K\} \\ &= \emptyset. \end{aligned}$$

ii) Supongamos que H es subconjugado de K . Definimos

$$\gamma : (G/K)^H \rightarrow \text{Hom}_G(G/H, G/K) \text{ tal que } aK \mapsto \gamma_a$$

donde $\gamma_a : G/H \rightarrow G/K$ está definida por $gH \mapsto gaK$.

Veamos que γ_a está bien definida. Sean $g_1H = g_2H$, así $g_1 = g_2h$ para algún $h \in H$. Luego, $\gamma_a(g_1H) = g_1aK = g_2(h(aK)) = g_2(aK) = \gamma_a(g_2H)$.

Ahora veamos que γ_a es un morfismo de G -conjuntos. Sean $g' \in G$ y $gH \in G/H$, entonces $\gamma_a(g'(gH)) = \gamma_a((g'g)H) = g'gaK = g'\gamma_a(gH)$.

Veamos que γ está bien definida. Sean $aK, bK \in (G/K)^H$ tales que $aK = bK$. Entonces $b = ak$ para algún $k \in K$. Luego, $\gamma_b(gH) = gbK = gakK = gaK = \gamma_a(gH)$ para toda $gH \in G/H$, por tanto $\gamma_a = \gamma_b$.

Sea

$$\gamma' : \text{Hom}_G(G/H, G/K) \longrightarrow (G/K)^H \text{ tal que } \alpha \longmapsto \alpha(H).$$

Veamos que γ' está bien definida. Dado $h \in H$, tenemos que $h\alpha(H) = \alpha(hH) = \alpha(H)$, por lo que $\alpha(H) \in (G/K)^H$.

Ahora, veamos que $\gamma \circ \gamma' = \text{Id}_{\text{Hom}_G(G/H, G/K)}$.

$$(\gamma \circ \gamma')(\alpha) = \gamma(\gamma'(\alpha)) = \gamma(\alpha(H)) = \gamma(gK) = \gamma_g, \text{ donde } \alpha(H) = gK.$$

Por otro lado $((\gamma \circ \gamma')(\alpha))(xH) = \gamma_g(xH) = xgK = x\alpha(H) = \alpha(xH)$, para todo $xH \in G/H$.

Por lo que $(\gamma \circ \gamma')(\alpha) = \alpha$, de aquí

$$\gamma \circ \gamma' = \text{Id}_{\text{Hom}_G(G/H, G/K)}.$$

Ahora, sea $aK \in (G/K)^H$, $(\gamma' \circ \gamma)(aK) = \gamma'(\gamma_a) = \gamma_a(H) = aK$.

Por tanto

$$\gamma' \circ \gamma = \text{Id}_{(G/K)^H}.$$

Por lo tanto, γ es biyección. □

En vista del Lema 2.3 y el Teorema 1.26 tenemos el siguiente corolario.

Corolario 2.4. Sean $H, K \leq G$ subgrupos. Entonces $\varphi_H(G/K) \neq 0$ si y sólo si H es subconjugado de K .

Definición 2.5. Dados G un grupo y $H \leq G$ subgrupo. Definimos el normalizador de H en G como:

$$N_G(H) = \{g \in G : gH = Hg\}.$$

Lema 2.6. Sean $H, K \leq G$ subgrupos, entonces

$$\varphi_H(G/K) = \left(\frac{|N_G(K)|}{|K|} \right) \alpha(H, K) = \left(\frac{|N_G(H)|}{|K|} \right) \beta(H, K)$$

donde

$$\alpha(H, K) = |\{E \leq G : [E] = [K] \text{ y } H \subseteq E\}|$$

y

$$\beta(H, K) = |\{E \leq G : [E] = [H] \text{ y } E \subseteq K\}|.$$

Demostración. Primero, llamemos $\mathfrak{A} = \{E \leq G : [E] = [K] \text{ y } H \subseteq E\}$. Luego, sabemos que

$$\begin{aligned} (G/K)^H &= \{aK : haK = aK \text{ para toda } h \in H\} \\ &= \{aK : a^{-1}ha \in K \text{ para toda } h \in H\} \\ &= \{aK : a^{-1}Ha \subseteq K\} \\ &= \{aK : H \subseteq aKa^{-1}\}. \end{aligned}$$

Ahora, tomemos $f : \{aK : H \subseteq aKa^{-1}\} \longrightarrow \mathfrak{A}$ tal que $aK \longmapsto aKa^{-1}$.

Veamos que f está bien definida. Sean $aK = bK$, entonces existe $k \in K$ tal que $a = bk$, por lo tanto $aKa^{-1} = bkKk^{-1}b^{-1} = bKb^{-1}$.

Veamos que f es sobreyectiva. Sea $E \in \mathfrak{A}$, así $[E] = [K]$, es decir, existe $a \in G$ tal que $E = aKa^{-1}$. Sólo falta ver que $aK \in (G/K)^H$. Sea $h \in H \subseteq E$, entonces $h = ak'a^{-1}$ para algún $k' \in K$, por lo que $haK = ak'a^{-1}aK = aK$.

Por lo tanto, $\{aK : H \subseteq aKa^{-1}\} = \bigcup_{E \in \mathfrak{A}} f^{-1}(E)$, donde $f^{-1}(E)$ es la imagen inversa de E bajo f .

Veamos que $f^{-1}(E) = \{abK : bK \in N_G(K)/K\}$.

(\subseteq) Sea $cK \in \{aK : H \subseteq aKa^{-1}\}$ tal que $f(cK) = aKa^{-1}$, de aquí que $cKc^{-1} = aKa^{-1}$, así $a^{-1}cKc^{-1}a = K$, con lo cual $a^{-1}c \in N_G(K)$, en consecuencia $a^{-1}c = b \in N_G(K)$, entonces $c = ab$ tal que $bK \in N_G(K)/K$.

(\supseteq) Sea abK con $bK \in N_G(K)/K$, en consecuencia $f(abK) = abKb^{-1}a^{-1} = aKa^{-1} = E$, entonces $abK \in f^{-1}(E)$.

Observemos que si $bK, b'K \in N_G(K)/K$ entonces $bK = b'K$ si y sólo si $abK = ab'K$ por lo que de la igualdad anterior de conjuntos obtenemos que $|f^{-1}(E)| = |N_G(K)/K|$. Por lo tanto,

$$\varphi_H(G/K) = \left(\frac{|N_G(K)|}{|K|} \right) \alpha(H, K).$$

Por otro lado, llamemos $A = \{a \in G : a^{-1}Ha \subseteq K\}$ y tomemos

$$f_1 : A \longrightarrow \{aK : a^{-1}Ha \subseteq K\} \text{ tal que } a \longmapsto aK.$$

Es claro que f_1 está bien definida y es sobreyectiva.

Ahora, tenemos que la imagen inversa de aK bajo f_1 es:

$$\begin{aligned}
 (f_1)^{-1}(aK) &= \{g \in G : f_1(g) = aK\} \\
 &= \{g \in G : gK = aK\} \\
 &= \{g \in G : a^{-1}g \in K\} \\
 &= \{g \in G : g = ak \text{ para algún } k \in K\} \\
 &= \{ak : k \in K\} \\
 &= aK.
 \end{aligned}$$

Así, $|(f_1)^{-1}(aK)| = |aK| = |K|$, de aquí que $|A| = |K| \varphi_H(G/K)$.

Ahora, llamemos $\mathfrak{B} = \{E \leq G : [E] = [H] \text{ y } E \subseteq K\}$ y veamos que

$$|A| = |N_G(H)| \beta(H, K).$$

Para ello, tomemos

$$f_2 : A \longrightarrow \mathfrak{B} \text{ tal que } a \longmapsto a^{-1}Ha.$$

Notemos que f_2 está bien definida. Luego, sea $E \in \mathfrak{B}$, así $[E] = [H]$ por lo que existe $a \in G$ tal que $E = a^{-1}Ha \subseteq K$, en consecuencia $a \in A$, por tanto f_2 es sobreyectiva.

Por otra parte, la imagen inversa de E bajo f_2 es:

$$\begin{aligned}
 (f_2)^{-1}(E) &= \{g \in G : f_2(g) = a^{-1}Ha\} \\
 &= \{g \in G : g^{-1}Hg = a^{-1}Ha\} \\
 &= \{g \in G : ag^{-1}Hga^{-1} = H\} \\
 &= \{g \in G : ga^{-1} \in N_G(H)\} \\
 &= \{g \in G : ga^{-1} = x \text{ para algún } x \in N_G(H)\} \\
 &= \{xa : x \in N_G(H)\} \\
 &= (N_G(H))a.
 \end{aligned}$$

De lo anterior, $|(f_2)^{-1}(E)| = |(N_G(H))a| = |N_G(H)|$. Por lo tanto,

$$|A| = |N_G(H)| \beta(H, K).$$

Concluyendo que

$$\varphi_H(G/K) = \left(\frac{|N_G(H)|}{|K|} \right) \beta(H, K). \quad \square$$

3 El anillo de Burnside

Dado G un grupo finito, sea \mathbb{A} la clase de todos los G -conjuntos finitos; es decir,

$$\mathbb{A} = \{X : X \text{ es un } G\text{-conjunto finito}\}$$

Observación 3.1. En \mathbb{A} hay una relación de equivalencia \sim . A saber, dados $X, Y \in \mathbb{A}$ se tiene que $X \sim Y$ si y sólo si $X \cong_G Y$.

Definición 3.2. Sea $X \in \mathbb{A}$. Denotaremos por $[X]$ a la clase de equivalencia de $X \in \mathbb{A}$ ($[X] = \{Y \in \mathbb{A} : X \cong_G Y\}$) y la llamaremos la clase de G -isomorfismo de X . Así, definimos

$$B^+(G) = \{[X] : X \in \mathbb{A}\}.$$

Teorema 3.3. Sea G un grupo finito, entonces $(B^+(G), +, \cdot)$ es un semianillo conmutativo con unidad, con las siguientes operaciones binarias:

$$i) + : B^+(G) \times B^+(G) \longrightarrow B^+(G) \text{ tal que } ([X], [Y]) \longmapsto [X] + [Y] = [X \sqcup Y].$$

$$ii) \cdot : B^+(G) \times B^+(G) \longrightarrow B^+(G) \text{ tal que } ([X], [Y]) \longmapsto [X] \cdot [Y] = [X \times Y].$$

Demostración. Veamos que la suma está bien definida:

Sean $[X_1], [X_2], [Y_1]$ y $[Y_2] \in B^+(G)$ tales que $[X_1] = [X_2]$ y $[Y_1] = [Y_2]$. Entonces existen $f_1 : X_1 \rightarrow X_2$ y $f_2 : Y_1 \rightarrow Y_2$ isomorfismos de G -conjuntos los cuales inducen el siguiente isomorfismo de G -conjuntos:

$$f_3 : (X_1 \times \{1\}) \cup (Y_1 \times \{2\}) \rightarrow (X_2 \times \{1\}) \cup (Y_2 \times \{2\}) \\ (z, i) \mapsto (f_i(z), i) \text{ para } i = 1, 2.$$

Y por lo tanto $[X_1 \sqcup Y_1] = [X_2 \sqcup Y_2]$. Concluyendo que la suma no depende de los representantes.

Por otro lado, con la notación anterior tenemos que f_1 y f_2 inducen el siguiente isomorfismo de G -conjuntos:

$$f_4 : X_1 \times Y_1 \rightarrow X_2 \times Y_2 \\ (x, y) \mapsto (f_1(x), f_2(y))$$

Obteniendo que $[X_1 \times Y_1] = [X_2 \times Y_2]$, por lo que el producto está bien definido.

En lo siguiente podemos considerar, sin pérdida de generalidad, que para elementos $[X], [Y] \in B^+(G)$ tenemos que $X \cap Y = \emptyset$; esto debido a que $[X] = [X \times \{1\}]$, $[Y] = [Y \times \{2\}]$, y a que la suma no depende del representante.

Veamos que $B^+(G)$ cumple las propiedades de semianillo conmutativo con unidad. Para ello sean $[X], [Y], [Z] \in B^+(G)$ donde X, Y y Z son ajenos dos a dos. Por lo que las uniones en los siguientes puntos serán ajenas.

i) Asociatividad de la suma.

$$\begin{aligned}
 [X] + ([Y] + [Z]) &= [X] + [Y \cup Z] \\
 &= [X \cup (Y \cup Z)] \\
 &= [(X \cup Y) \cup Z] \\
 &= [X \cup Y] + [Z] \\
 &= ([X] + [Y]) + [Z].
 \end{aligned}$$

ii) Conmutatividad de la suma.

$$\begin{aligned}
 [X] + [Y] &= [X \cup Y] \\
 &= [Y \cup X] \\
 &= [Y] + [X].
 \end{aligned}$$

iii) Notamos que $[\emptyset] \in B^+(G)$, así $[X] + [\emptyset] = [X \cup \emptyset] = [X]$. Por tanto, $[\emptyset]$ es el neutro con la suma.

iv) Asociatividad del producto.

$$\begin{aligned}
 [X] \cdot ([Y] \cdot [Z]) &= [X] \cdot [Y \times Z] \\
 &= [X \times (Y \times Z)] \\
 &= [(X \times Y) \times Z] \\
 &= [X \times Y] \cdot [Z] \\
 &= ([X] \cdot [Y]) \cdot [Z].
 \end{aligned}$$

v) Conmutatividad del producto. Sea

$$l : X \times Y \longrightarrow Y \times X \quad \text{tal que } (x, y) \longmapsto (y, x).$$

Claramente l está bien definida, es biyectiva y es de G -conjuntos. Por lo tanto $[X] \cdot [Y] = [Y] \cdot [X]$.

vi) Distributividad.

$$\begin{aligned}
 [X] \cdot ([Y] + [Z]) &= [X] \cdot [Y \cup Z] \\
 &= [X \times (Y \cup Z)] \\
 &= [(X \times Y) \cup (X \times Z)] \\
 &= [X \times Y] + [X \times Z] \\
 &= [X] \cdot [Y] + [X] \cdot [Z].
 \end{aligned}$$

vii) Notamos que $[G/G] \in B^+(G)$. Tenemos que

$$f : X \times G/G \longrightarrow X \text{ tal que } (x, eG) \longmapsto x$$

es un isomorfismo de G -conjuntos. Por tanto, $[G/G]$ es la unidad en $B^+(G)$.

Por lo tanto, $(B^+(G), +, \cdot)$ es un semianillo conmutativo con unidad. \square

Observación 3.4. Dados $G = \{e\}$ y X un conjunto. Entonces X es un $\{e\}$ -conjunto con acción trivial. Además, dos G -conjuntos con la acción trivial son isomorfos si y sólo si existe una biyección entre ellos. De aquí que

$$\phi : B^+(\{e\}) \longrightarrow \mathbb{N} \text{ tal que } [X] \longmapsto |X|$$

es un isomorfismo de semianillos, con lo cual $\mathbb{N} \cong B^+(\{e\})$. En este capítulo consideramos al cero como un elemento de los números naturales.

Teorema 3.5. Sean $[X], [Y], [Z] \in B^+(G)$ tales que $[X] + [Y] = [Z] + [Y]$, entonces $[X] = [Z]$. Es decir, hay cancelación con la suma en $B^+(G)$.

Demostración. Sean $[X], [Y]$ y $[Z] \in B^+(G)$ tales que $[X] + [Y] = [Z] + [Y]$. Como antes podemos asumir sin pérdida de generalidad que X, Y, Z son ajenos dos a dos, y por consiguiente todas las uniones que aparecen a continuación son ajenas. Sean $X = \bigcup_{i=1}^m O_G(x_i)$, $Y = \bigcup_{j=1}^l O_G(y_j)$ y $Z = \bigcup_{k=1}^n O_G(z_k)$, puesto que $[X \cup Y] = [Z \cup Y]$ hay una correspondencia biyectiva entre las órbitas de $X \cup Y$ y $Z \cup Y$, esto implica que $m + l = n + l$, por lo que $m = n$. Por lo tanto $Z = \bigcup_{i=1}^m O_G(z_i)$.

Demostremos que hay cancelación por inducción sobre $1 \leq l$.

i) Para $l = 1$, existe un isomorfismo de G -conjuntos

$$f : X \cup O_G(y_1) \longrightarrow Z \cup O_G(y_1).$$

Observemos que si $f(O_G(y_1)) = O_G(y_1)$ entonces $f|_X : X \rightarrow Z$ es un isomorfismo de G -conjuntos, por lo que $[X] = [Z]$.

En caso contrario, sin pérdida de generalidad, supongamos que $f(O_G(y_1)) = O_G(z_m)$, entonces $[O_G(y_1)] = [O_G(z_m)]$. Por otro lado, sin pérdida de generalidad, podemos suponer que $f(O_G(x_m)) = O_G(y_1)$, entonces $[O_G(x_m)] = [O_G(y_1)]$. Por lo tanto $[O_G(x_m)] = [O_G(z_m)]$. Sea $f_1 : O_G(x_m) \rightarrow O_G(z_m)$ un isomorfismo de G -conjuntos. Por otro lado observemos que

$$f|_{\bigcup_{i=1}^{m-1} O_G(x_i)} : \bigcup_{i=1}^{m-1} O_G(x_i) \rightarrow \bigcup_{i=1}^{m-1} O_G(z_i)$$

es un isomorfismo de G -conjuntos, por lo que este junto con f_1 inducen un isomorfismo entre $\bigcup_{i=1}^m \mathcal{O}_G(x_i)$ y $\bigcup_{i=1}^m \mathcal{O}_G(z_i)$. Por lo tanto $[X] = [Z]$.

ii) Para $l > 1$ tenemos que $[X] + [Y] = [Z] + [Y]$ entonces

$$([X] + [\bigcup_{j=1}^{l-1} \mathcal{O}_G(y_j)]) + [\mathcal{O}_G(y_l)] = ([Z] + [\bigcup_{j=1}^{l-1} \mathcal{O}_G(y_j)]) + [\mathcal{O}_G(y_l)]$$

por lo que utilizando i) podemos cancelar una órbita, en este caso la órbita $\mathcal{O}_G(y_l)$ obteniendo que $[X] + [\bigcup_{j=1}^{l-1} \mathcal{O}_G(y_j)] = [Z] + [\bigcup_{j=1}^{l-1} \mathcal{O}_G(y_j)]$, finalmente por hipótesis de inducción podemos cancelar $l - 1$ órbitas, concluyendo que $[X] = [Z]$.

□

Observación 3.6. El Teorema 3.5 nos permite definir una relación de equivalencia en $B^+(G) \times B^+(G)$. A saber, dados $([X_1], [Y_1])$ y $([X_2], [Y_2]) \in B^+(G) \times B^+(G)$ tenemos que $([X_1], [Y_1]) \sim ([X_2], [Y_2])$ si y sólo si $[X_1] + [Y_2] = [X_2] + [Y_1]$. Denotaremos a la clase de equivalencia de $([X], [Y]) \in B^+(G) \times B^+(G)$ por $[X] - [Y]$.

Definición 3.7. Definimos el *Anillo de Burnside* $B(G)$ de un grupo finito G como:

$$B(G) = \{[X] - [Y] : ([X], [Y]) \in B^+(G) \times B^+(G)\}$$

dotado de las siguientes operaciones binarias de suma y producto:

$$\begin{aligned} ([X_1] - [Y_1]) + ([X_2] - [Y_2]) &= ([X_1] + [X_2]) - ([Y_1] + [Y_2]) \\ ([X_1] - [Y_1])([X_2] - [Y_2]) &= ([X_1][X_2] + [Y_1][Y_2]) - ([X_1][Y_2] + [X_2][Y_1]) \end{aligned}$$

Observación 3.8. $B(G)$ es el anillo de Grothendieck de $B^+(G)$, donde

- i) El neutro de la suma es $[\emptyset] - [\emptyset] = [X] - [X]$ con $[X] \in B^+(G)$.
- ii) Dado $[X] - [Y] \in B(G)$ su inverso aditivo es $[Y] - [X]$.
- iii) En el producto la unidad es $[G/G] - [\emptyset]$.

Teorema 3.9. *Dado un grupo finito G . Entonces $B(G)$ como grupo abeliano es libre como \mathbb{Z} -módulo, generado por los elementos de la forma $[G/H_i]$, donde H_i recorre un conjunto de representantes de $\mathcal{C}(G)$.*

Demostración. Sea G un grupo finito y X un G -conjunto finito. Por el Teorema 1.20, $X \cong_G \bigsqcup_{i \in I} G/H_i$, donde $H_i \leq G$ para cada $i \in I$. Así,

$$[X] = \left[\bigsqcup_{i \in I} G/H_i \right] = \sum_{i \in I} [G/H_i] \in B^+(G).$$

Ahora, por el Lema 1.24 podemos tener repeticiones de algunos $[G/H_i]$, ya que $[G/H_i] = [G/H_j]$ si y sólo si $[H_i] = [H_j]$. Por lo que denotaremos con $a_{H_i} \in \mathbb{N}$ al número de veces que se repite el elemento $[G/H_i]$. En consecuencia, si $[X] \in B^+(G)$, entonces $[X] = \sum_{[H] \in \mathcal{C}(G)} a_H [G/H]$, con $a_H [G/H] \in \mathbb{N}[G/H] \subseteq B^+(G)$. Por lo que, abusando de la notación tenemos que los elementos de $B(G)$ son de la forma:

$$\sum_{[H] \in \mathcal{C}(G)} a_H [G/H], \text{ con } a_H [G/H] \in \mathbb{Z}[G/H] \subseteq B(G).$$

Ahora, sólo falta ver que esta expresión es única. Supongamos que

$$\sum_{[H] \in \mathcal{C}(G)} a_H [G/H] = 0$$

y que no todos los a_H son cero, así tenemos la siguiente igualdad en $B^+(G)$.

$$\sum_{a_K > 0} a_K [G/K] = \sum_{a_L < 0} (-a_L) [G/L].$$

Llamemos $[Y] = \sum_{a_K > 0} a_K [G/K]$ y $[Z] = \sum_{a_L < 0} (-a_L) [G/L]$. Ahora, si $[Y] = [\emptyset]$, entonces \emptyset tendría una órbita distinta del vacío, lo que no es posible. De aquí que, sin pérdida de generalidad, podemos suponer que $[Y] \neq [\emptyset]$ y $[Z] \neq [\emptyset]$. Entonces si $a_{K_0} > 0$ existe $\mathcal{O}_G(y_0) \subseteq Y$ tal que $\mathcal{O}_G(y_0) \cong_G G/K_0$. Por otro lado, $[Y] = [Z]$, entonces $\mathcal{O}_G(y_0) \cong_G \mathcal{O}_G(z_0)$ para algún $z_0 \in Z$ que a su vez $\mathcal{O}_G(z_0) \cong_G G/L_0$ para algún L_0 tal que $a_{L_0} < 0$, en consecuencia $G/K_0 \cong_G G/L_0$, así $[K_0] = [L_0]$, lo que no es posible. Por tanto,

$$B(G) = \bigoplus_{[H] \in \mathcal{C}(G)} \mathbb{Z}[G/H]. \quad \square$$

Observación 3.10. Hay una inclusión

$$\iota : B^+(G) \longrightarrow B(G) \text{ tal que } [X] \longmapsto [X] - [\emptyset].$$

Observación 3.11. Dados G un grupo finito, R un anillo conmutativo y

$$f : B^+(G) \longrightarrow R$$

un morfismo de semianillos, tenemos que f se extiende de manera única a $B(G)$; es decir, existe

$$F : B(G) \longrightarrow R$$

un morfismo de anillos tal que $F \Big|_{B^+(G)} = f$. Además, F esta dada por:

$$F([X] - [Y]) = f([X]) - f([Y]).$$

Lema 3.12. *Sea*

$$\varphi : B^+(G) \longrightarrow \prod_{[H] \in \mathcal{C}(G)} \mathbb{Z} \text{ dada por } [X] \mapsto \varphi([X]) = (\varphi_H(X))_{[H] \in \mathcal{C}(G)},$$

entonces φ es un morfismo inyectivo de semianillos. De aquí que, φ se extiende de manera única a un morfismo inyectivo de anillos:

$$\tilde{\varphi} : B(G) \longrightarrow \prod_{[H] \in \mathcal{C}(G)} \mathbb{Z}.$$

Demostración. Note que $\prod_{[H] \in \mathcal{C}(G)} \mathbb{Z} = \mathbb{Z}^{|\mathcal{C}(G)|}$ es un anillo conmutativo con suma y producto definimos de manera usual (entrada a entrada).

Veamos que φ está bien definida. Sea $[X] \in B^+(G)$. Entonces $(\varphi_H(X))_{[H] \in \mathcal{C}(G)}$ no depende del representante de $[H]$ por el Teorema 2.2 inciso iii). Sea $[Y] \in B^+(G)$ tal que $[X] = [Y]$, esto implica que existe $f : X \rightarrow Y$ un isomorfismo de G -conjuntos. Sean $[H] \in \mathcal{C}(G)$ y $x \in X^H$ entonces $hx = x$ para toda $h \in H$, esto es equivale a que $hf(x) = f(x)$ para toda $h \in H$ y esto pasa si y sólo si $f(x) \in Y^H$. Por lo que tenemos la siguiente biyección

$$\begin{aligned} X^H &\rightarrow Y^H \\ x &\mapsto f(x) \end{aligned}$$

obteniendo que $\varphi_H(X) = \varphi_H(Y)$. Lo anterior se cumple para cada $[H] \in \mathcal{C}(G)$, de modo que $\varphi([X]) = (\varphi_H(X))_{[H] \in \mathcal{C}(G)} = (\varphi_H(Y))_{[H] \in \mathcal{C}(G)} = \varphi([Y])$. Por lo tanto φ está bien definida.

Recuerde que $1_{B(G)} = [G/G]$ y $1_{\mathbb{Z}^{|\mathcal{C}(G)|}} = \underbrace{(1, \dots, 1)}_{|\mathcal{C}(G)|\text{-veces}}.$

Ahora, para cada $H \leq G$ subgrupo, tenemos que $[H] \preceq [G]$, así $\varphi_H(G/G) \neq 0$, de modo que $\varphi_H(G/G) = 1$. Por tanto, $\varphi([G/G]) = (\varphi_H(G/G))_{[H] \in \mathcal{C}(G)} = \underbrace{(1, \dots, 1)}_{|\mathcal{C}(G)| \text{-veces}}$.

De las propiedades de la marca, es fácil ver que φ preserva la suma y el producto (ver Teorema 2.2 inciso i) y ii)). Por lo tanto φ es un morfismo de semianillos.

Ahora veamos que φ es inyectiva. Recordemos que φ es inyectiva si y sólo si para cada $0 \neq a \in B^+(G)$, se tiene que $\varphi(a) \neq 0$.

Sea $a \in B^+(G)$ tal que $a \neq 0$, así $a = \sum_{[H] \in \mathcal{C}(G)} a_H [G/H]$ con $a_H \in \mathbb{N}$, donde no

todos los a_H son cero. Puesto que G es finito, toda cadena estrictamente creciente $[K_1] \prec [K_2] \prec \dots$ se estaciona. Elegimos $[K']$ máximo con respecto a este orden tal que $a_{K'} \neq 0$. Tenemos que

$$a = \sum_{[H] \in \mathcal{C}(G)} a_H [G/H] = \left[\bigsqcup_{i \in I} G/H_i \right]$$

donde $|I|$ es el número de órbitas y $H_i \leq G$ para cada $i \in I$. Entonces

$$\varphi_{K'} \left(\bigsqcup_{i \in I} G/H_i \right) = \sum_{i \in I} \varphi_{K'}(G/H_i).$$

Ahora, por la buena definición de φ tenemos que si $[G/H_i] = [G/H_j]$ entonces $\varphi_{K'}(G/H_i) = \varphi_{K'}(G/H_j)$ y esto ocurre si y sólo si $[H_i] = [H_j]$, por lo tanto

$$\varphi_{K'} \left(\bigsqcup_{i \in I} G/H_i \right) = \sum_{[H] \in \mathcal{C}(G)} a_H \varphi_{K'}(G/H),$$

ver Teorema 3.9 .

Por otro lado sabemos que $\varphi_{K'}(G/H) \neq 0$ si y sólo si $[K'] \preceq [H]$ y puesto que elegimos $[K']$ máximo con respecto a este orden parcial tal que $a_{K'} \neq 0$ entonces

$$\varphi_{K'} \left(\bigsqcup_{i \in I} G/H_i \right) = a_{K'} \varphi_{K'}(G/K') = a_{K'} |N_G(K')/K'| \neq 0.$$

Así la coordenada $[K']$ -ésima de $\varphi(a)$ es distinta de cero y esto implica $\varphi(a) \neq 0$. Por tanto φ es un morfismo inyectivo de semianillos.

Por último veamos que

$$\tilde{\varphi} : B(G) \longrightarrow \prod_{[H] \in \mathcal{C}(G)} \mathbb{Z}$$

es inyectiva:

Sean $[X_1] - [Y_1], [X_2] - [Y_2] \in B(G)$ tales que

$$\tilde{\varphi}([X_1] - [Y_1]) = \tilde{\varphi}([X_2] - [Y_2])$$

lo cual ocurre si y sólo si

$$\begin{aligned} \varphi([X_1]) - \varphi([Y_1]) &= \varphi([X_2]) - \varphi([Y_2]) \\ \Leftrightarrow \varphi([X_1]) + \varphi([Y_2]) &= \varphi([X_2]) + \varphi([Y_1]) \\ \Leftrightarrow \varphi([X_1] + [Y_2]) &= \varphi([X_2] + [Y_1]) \end{aligned}$$

y puesto que φ es inyectiva entonces

$$[X_1] + [Y_2] = [X_2] + [Y_1]$$

obteniendo que

$$[X_1] - [Y_1] = [X_2] - [Y_2]$$

ver Observación 3.6. Por lo tanto $\tilde{\varphi}$ es un morfismo inyectivo de anillos, ver Observación 3.11. \square

Corolario 3.13. Sean G un grupo finito y X, Y dos G -conjuntos finitos, entonces $[X] = [Y]$ si y sólo si $\varphi_H(X) = \varphi_H(Y)$ para toda $[H] \in \mathcal{C}(G)$.

Demostración. El resultado es inmediato por la buena definición y la inyectividad de φ del lema anterior. \square

Observación 3.14. En el caso de un grupo abeliano G , puesto que todos sus subgrupos son normales, tenemos que $N_G(K) = G$ para todo $K \leq G$ y además si $H \leq G$ entonces

$$\alpha(H, K) = |\{E \leq G \mid [E] = [K] \text{ y } H \subseteq E\}| = \begin{cases} 0 & \text{si } H \not\subseteq K \\ 1 & \text{si } H \subseteq K \end{cases}$$

por lo que la fórmula del Lema 2.6 se reduce a

$$\varphi_H(G/K) = \begin{cases} 0 & \text{si } H \not\subseteq K \\ |G/K| & \text{si } H \subseteq K \end{cases}$$

Ejemplo 3.15. Sea C un grupo cíclico de orden 2^3 . Recordemos que un grupo finito es cíclico si y sólo si para cada divisor d de su orden existe a lo más un subgrupo cíclico de orden d . Entonces tenemos que las clases de conjugación de los subgrupos de C son :

$$\mathcal{C}(C) = \{2^3C; 2^2C; 2C; C\}$$

por lo que la estructura de $B(C)$ como \mathbb{Z} -módulo libre es:

$$B(C) = \mathbb{Z} [C/2^3C] \oplus \mathbb{Z} [C/2^2C] \oplus \mathbb{Z} [C/2C] \oplus \mathbb{Z} [C/C].$$

Por otro lado recordemos que

$$\varphi : B(C) \rightarrow \mathbb{Z}^4$$

$$[X] \mapsto (\varphi_{2^3C}(X), \varphi_{2^2C}(X), \varphi_{2C}(X), \varphi_C(X))$$

$$[C/2^3C] \mapsto (2^3, 0, 0, 0)$$

$$[C/2^2C] \mapsto (2^2, 2^2, 0, 0)$$

$$[C/2C] \mapsto (2, 2, 2, 0)$$

$$[C/C] \mapsto (1, 1, 1, 1)$$

Así $\varphi(B(C)) \subseteq \mathbb{Z}^4$, el cual es un anillo isomorfo a $B(C)$, esta generado por las columnas de la matriz:

$$\begin{pmatrix} 8 & 4 & 2 & 1 \\ 0 & 4 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Agradecimientos

Agradecemos a los árbitros, por sus valiosos comentarios, aportaciones y tiempo dedicado a la lectura minuciosa de este trabajo.

Bibliografía

- [1] A. Dress, *Contributions to the theory of induced representations*, Lecture Notes in Math. Algebraic K-theory, Springer-Verlag, 1973.

- [2] S. Bouc, *Burnside rings*, Handbook of algebra, Volume 2, North-Holland, Amsterdam, 2000, 739-804.
- [3] David Villa Hernández, *Functional Equations for Zeta functions of Burnside rings*, JP Journal of Algebra, Number Theory and Applications Vol. 29 (2013) No. 1, 1-16.
- [4] Alberto G. Raggi Cárdenas and Luis Valero Elizondo, *Normalizing isomorphisms between Burnside rings*, Journal of Algebra 277 (2004) 643–657.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570
cesarcc@fcfm.buap.mx
214570948@alumnos.fcfm.buap.mx
dvilla@fcfm.buap.mx

Capítulo 2

El lema de Hensel y el levantamiento de Hensel

Ángel Raúl García Ramírez, Carlos Alberto López Andrade
FCFM, BUAP

Resumen

El Lema de Hensel es una herramienta utilizada en la teoría de números, geometría y topología algebraica, siendo la factorización de polinomios el objetivo principal de su uso. Desarrollado por Kurt Hensel quien introdujo los números p -ádicos, nos muestra un método para hallar factorizaciones de polinomios mónicos, factorizándolos sobre un campo finito de característica un primo p y después “levantando” dicha factorización sobre un anillo de clases residuales módulo una potencia de ese primo. Recientemente esta herramienta y el levantamiento de Hensel se han empleado en la Teoría de Códigos Algebraicos. En este trabajo se exhibe la demostración del Lema de Hensel para polinomios mónicos sobre el anillo \mathbb{Z}_{p^s} , así como del levantamiento de Hensel.

1 Introducción

En 1904 Kurt Hensel introdujo los números p -ádicos y el lema de Hensel en el artículo *Neue Grundlagen der Arithmetik* (cf. [6]). Cuatro años más tarde, en su libro *Theorie der algebraischen Zahlen*, Hensel mostró una versión más general de su lema. Recientemente, en los años 90 del siglo pasado, esta herramienta ha sido muy útil en el estudio de la caracterización de la estructura algebraica de los códigos cíclicos lineales sobre anillos finitos, por ejemplo; sobre anillos de Galois (cf. [5], [14], [7]) y sobre anillos finitos de cadena (cf. [3], [13], [11]). El levantamiento de Hensel está relacionado con la representación p -ádica de los elementos de un anillo de Galois ó de un anillo finito de cadena y tal representación es imprescindible en la definición de la función de Gray en esta clase de anillos finitos ([4], [1], [2], [17], [9], [10]). Se demostró en [5] a través de Hammons, et. al., que el código de Kerdock es la imagen de Gray de un código cíclico lineal extendido sobre \mathbb{Z}_4 . Ellos usan este hecho para resolver un problema “viejo” dando la explicación de la dualidad formal entre dos códigos no lineales binarios, los famosos códigos de Kerdock y Preparata.

Definiciones y conceptos fundamentales sobre el anillo \mathbb{Z}_{p^s} , propiedades de los polinomios y de la divisibilidad en $\mathbb{Z}_{p^s}[x]$ son dados en la sección 2, la demostración

del Lema de Hensel se encuentra desarrollada en la sección 3 y finalmente la demostración del Levantamiento de Hensel se exhibe en la sección 4. Cabe mencionar que este capítulo de libro está inspirado en la lectura de la obra de Wan ([16]).

2 El anillo de polinomios $\mathbb{Z}_{p^s}[x]$

Sea p cualquier número primo, $s \in \mathbb{N}$ y \mathbb{Z}_{p^s} el anillo de enteros módulo p^s , i.e.,

$$\mathbb{Z}_{p^s} = \frac{\mathbb{Z}}{p^s\mathbb{Z}} = \{\overline{0}, \overline{1}, \dots, \overline{p^s - 1}\}$$

con la suma y multiplicación usual de clases:

$$\begin{aligned} \overline{r_1} + \overline{r_2} &= \overline{r_1 + r_2} \\ \overline{r_1} \overline{r_2} &= \overline{r_1 r_2} \end{aligned} \tag{1}$$

para cada $\overline{r_1}, \overline{r_2} \in \mathbb{Z}_{p^s}$.

Observación 2.1. Por la construcción del anillo cociente \mathbb{Z}_{p^s} tenemos que $\overline{ab} = \overline{c}$ sí y sólo si $ab \equiv c \pmod{p^s}$.

Considérese el conjunto $\mathcal{S} = \{0, 1, \dots, p^s - 1\} \subseteq \mathbb{N} \cup \{0\}$ y la función:

$$\begin{aligned} \phi : \mathbb{Z}_{p^s} &\longrightarrow \mathcal{S} \\ \overline{r} &\mapsto r \end{aligned}$$

Es claro que ϕ es biyectiva, entonces podemos inducir la estructura de anillo en el conjunto \mathcal{S} definiendo las operaciones \oplus y \odot en \mathcal{S} de la manera siguiente:

$$\begin{aligned} s_1 \oplus s_2 &:= \phi(\overline{r_1} + \overline{r_2}) \\ s_1 \odot s_2 &:= \phi(\overline{r_1} \cdot \overline{r_2}) \end{aligned}$$

para cada $s_1, s_2 \in \mathcal{S}$ siempre que $s_1 = \phi(\overline{r_1})$ y $s_2 = \phi(\overline{r_2})$ para algunos $\overline{r_1}$ y $\overline{r_2}$ en \mathbb{Z}_{p^s} .

Dadas estas operaciones, ϕ es un *isomorfismo* y, con esta identificación, usaremos cuando sea conveniente que:

$$\mathbb{Z}_{p^s} = \{0, 1, \dots, p^s - 1\}. \tag{2}$$

Recordar que: (a, b) denota al máximo común divisor de a y b , i.e., $(a, b) = \text{mcd}(a, b)$.

El siguiente lema nos será muy útil más adelante.

Lema 2.2. Sean $a \in \mathbb{N}$ y p un número primo. Si $(a, p) = 1$ y $s \in \mathbb{N}$, entonces $ca \equiv 1 \pmod{p^s}$ para algún $c \in \mathbb{Z}$.

Demostración. Haremos inducción sobre s . Veamos que el resultado es válido para $s = 1$. Como $(a, p) = 1$, entonces existen $c, d \in \mathbb{Z}$ tales que $ca + dp = 1$. Así $ca - 1 = -dp$ es decir, $p \mid (ca - 1)$, por lo tanto,

$$ca \equiv 1 \pmod{p^1}.$$

Supóngase que el resultado se cumple para s y demostremos que se cumple para $s + 1$. Por hipótesis, existen $k_1, k_2, l_1, l_2 \in \mathbb{Z}$ tales que:

$$\begin{aligned} k_1 a + k_2 p &= 1, \\ l_1 a + l_2 p^s &= 1. \end{aligned}$$

Multiplicando las igualdades anteriores tenemos que:

$$\begin{aligned} 1 &= k_1 l_1 a^2 - k_1 l_2 a p^s + k_2 l_1 p a - k_2 l_2 p^{s+1} \\ &= (k_1 l_1 a - k_1 l_2 p^s + k_2 l_1 p) a + (-k_2 l_2) p^{s+1} \\ &= ca + dp^{s+1} \end{aligned}$$

donde $c = k_1 l_1 a - k_1 l_2 p^s + k_2 l_1 p$ y $d = -k_2 l_2$. Entonces $ca - 1 = -dp^{s+1}$, es decir,

$$ca \equiv 1 \pmod{p^{s+1}},$$

lo cual concluye la prueba. \square

Corolario 2.3. En \mathbb{Z}_{p^s} se satisface:

- i) Si $\bar{a} \in \mathbb{Z}_{p^s}$ y $(a, p) = 1$, entonces \bar{a} es una unidad en \mathbb{Z}_{p^s} .
- ii) Para todo $\bar{x} \in \mathbb{Z}_{p^s}$ con $\bar{x} \neq 0$, existen $\bar{a} \in \mathbb{Z}_{p^s}$ e $i \in \mathbb{N} \cup \{0\}$ tales que $(a, p) = 1$ y $\bar{x} = \bar{a} \bar{p}^i$.

Demostración. Sea $\bar{a} \in \mathbb{Z}_{p^s}$ con $(a, p) = 1$, por el Lema 2.2 existe $c \in \mathbb{N}$ tal que $ca \equiv 1 \pmod{p^s}$ entonces $\bar{c} \cdot \bar{a} = \bar{1}$, debido a la Observación 2.1. Por lo tanto, \bar{a} es una unidad de \mathbb{Z}_{p^s} . Ahora bien, considérese $\bar{x} \in \mathbb{Z}_{p^s}$ con $\bar{x} \neq \bar{0}$, usando la identificación anterior se sigue que $x \in \mathbb{N}$. Se tienen dos casos:

- Si $(x, p) = 1$ entonces sea $a = x$, de ahí que, $\bar{x} = \bar{a} \bar{p}^0$, lo cual prueba el resultado.

- Si $(x, p) = p$ entonces $p|x$. Además si q es un primo tal que $q|x$ entonces $q \leq p^s - 1$ pues $x \leq p^s - 1$. Por el teorema fundamental de la aritmética, existen $q_1, q_2, \dots, q_k \in \mathbb{N}$ números primos y $n_1, n_2, \dots, n_k \in \mathbb{N}$ tales que:

$$x = q_1^{n_1} q_2^{n_2} \dots q_k^{n_k}$$

como $p|x$ entonces $p = q_j$ para algún $j \in \{1, 2, \dots, k\}$. Sea

$$i = \text{máx} \{n_j \mid (x, p^{n_j}) = p^{n_j}\}.$$

Como $(x, p) = p$ entonces $n_j \geq 1$ y por consiguiente este conjunto es no vacío, entonces:

$$\begin{aligned} x &= q_1^{n_1} q_2^{n_2} \dots q_{j-1}^{n_{j-1}} p^i q_{j+1}^{n_{j+1}} \dots q_k^{n_k} \\ &= \left(q_1^{n_1} q_2^{n_2} \dots q_{j-1}^{n_{j-1}} q_{j+1}^{n_{j+1}} \dots q_k^{n_k} \right) p^i \end{aligned}$$

donde $a = q_1^{n_1} q_2^{n_2} \dots q_{j-1}^{n_{j-1}} q_{j+1}^{n_{j+1}} \dots q_k^{n_k}$, entonces $x = ap^i$ y, por lo tanto, $\bar{x} = \overline{ap^i}$. \square

Ejemplo 2.4. Sean $p = 3$, $s = 2$ entonces $\mathbb{Z}_{3^2} = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$, si definimos el conjunto $\mathcal{A} := \{a \in \mathbb{Z}_{p^s} : (a, p) = 1\}$; en este caso tenemos que $\mathcal{A} = \{1, 2, 4, 5, 7, 8\}$ luego:

$$\begin{aligned} 1 \odot 1 &= 1 \\ 2 \odot 5 &= 1 \\ 4 \odot 7 &= 1 \\ 8 \odot 8 &= 1 \end{aligned} \tag{3}$$

como se afirma en el Corolario 2.3. Así $\mathbb{Z}_{p^s}^* = \mathcal{A}$, donde $\mathbb{Z}_{p^s}^*$ denota el grupo de unidades del anillo \mathbb{Z}_{p^s} .

El siguiente teorema caracteriza a los ideales en el anillo \mathbb{Z}_{p^s} .

Teorema 2.5. *Los ideales principales $\langle \bar{1} \rangle, \langle \bar{p} \rangle, \dots, \langle \overline{p^{s-1}} \rangle, \langle \bar{0} \rangle$, son todos los ideales de \mathbb{Z}_{p^s} . $\langle \bar{p} \rangle$ es el único ideal maximal de \mathbb{Z}_{p^s} y $\frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle} \simeq \mathbb{F}_p$.*

Demostración. Sea I un ideal de \mathbb{Z}_{p^s} . Suponga que $I \neq \langle \bar{0} \rangle$. Sea m el natural más pequeño que es representante de las clases de I , veamos que $I = \langle \bar{m} \rangle$. Sea $\bar{i} \in I$ entonces $\bar{i} \in \mathbb{Z}_{p^s}$. Por el algoritmo euclidiano de la división, existen $q, r \in \mathbb{Z} : i = qm + r$ con $0 \leq r < m$, entonces $r = i - qm$, de ahí que, $\bar{r} = \bar{i} - \bar{q}\bar{m} \in I$, por consiguiente, $\bar{r} \in I$ y $r < m$ lo cual contradice la minimalidad de m , así $\bar{r} = \bar{0}$, en otras palabras, $\bar{i} = \bar{q}\bar{m}$, entonces $\bar{i} \in \langle \bar{m} \rangle$, en consecuencia, $I \subseteq \langle \bar{m} \rangle$. Es claro que

$\bar{m} \in I$, por consiguiente $\langle \bar{m} \rangle \subseteq I$, entonces $I = \langle \bar{m} \rangle$, como queríamos. Dado que $\bar{m} \in I \subseteq \mathbb{Z}_{p^s}$, por el Corolario 2.3, existe $a \in \mathbb{Z}$, tal que:

$$\bar{m} = \bar{a}p^i \text{ para algún } i \in \{0, \dots, s-1\}. \quad (4)$$

Como $(a, p) = 1$ tenemos que $ca \equiv 1 \pmod{p^s}$ para algún $c \in \mathbb{Z}$, así, $ca - 1 = dp^s$ con $d \in \mathbb{Z}$, entonces $cap^i - p^i = dp^s p^i$, de ahí que, $\overline{cap^i} - \bar{p}^i = \overline{dp^s p^i}$, por lo que

$$\begin{aligned} \bar{0} &= \overline{cap^i} - \bar{p}^i \\ &= \bar{c}\bar{m} - \bar{p}^i \end{aligned}$$

de manera que, $\bar{c}\bar{m} = \bar{p}^i$. Entonces $\bar{p}^i \in \langle \bar{m} \rangle$ luego $\langle \bar{p}^i \rangle \subseteq \langle \bar{m} \rangle$, y por (4), $\langle \bar{m} \rangle \subseteq \langle \bar{p}^i \rangle$, de ahí que, $\langle \bar{m} \rangle = \langle \bar{p}^i \rangle$. Por consiguiente, queda demostrado que todo ideal I de \mathbb{Z}_{p^s} es de la forma $I = \langle \bar{p}^i \rangle$ para algún $i \in \{0, 1, \dots, s-1\}$. Además, como $\bar{p}^i \in \langle \bar{p}^i \rangle$ y $\bar{p} \in \mathbb{Z}_{p^s}$ entonces $\bar{p}^i \cdot \bar{p} = \bar{p}^{i+1} \in \langle \bar{p}^i \rangle$, así, $\langle \bar{p}^{i+1} \rangle \subseteq \langle \bar{p}^i \rangle$, por lo tanto,

$$\langle \bar{0} \rangle \subseteq \langle \bar{p}^{s-1} \rangle \subseteq \dots \subseteq \langle \bar{p}^2 \rangle \subseteq \langle \bar{p} \rangle \subseteq \langle \bar{1} \rangle = \mathbb{Z}_{p^s}. \quad (5)$$

Ahora veamos que $\langle \bar{p} \rangle$ es un ideal maximal. Sea J un ideal de \mathbb{Z}_{p^s} tal que:

$$\langle \bar{p} \rangle \subseteq J \subseteq \mathbb{Z}_{p^s}, \quad (6)$$

luego, existe algún i con $i \in \{0, 1, \dots, s-1\}$ tal que $J = \langle \bar{p}^i \rangle$. Tenemos que: si $i = 1$ entonces $J = \langle \bar{p}^1 \rangle = \langle \bar{p} \rangle$, o bien, si $1 < i \leq s-1$ entonces $J = \langle \bar{p}^i \rangle \subseteq \langle \bar{p} \rangle$ debido a (5), entonces, por (6), $J = \langle \bar{p} \rangle$, o bien, si $i = 0$ entonces $J = \langle \bar{p}^0 \rangle = \langle \bar{1} \rangle = \mathbb{Z}_{p^s}$. En cualquier caso, se concluye que $J = \langle \bar{p} \rangle$ o $J = \mathbb{Z}_{p^s}$. Por lo tanto, $\langle \bar{p} \rangle$ es un ideal maximal, y por (5), es el único ideal maximal de \mathbb{Z}_{p^s} , en consecuencia, $\frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle}$ es un campo.

Recordemos que:

- i) $\frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle} = \{\bar{x} + \langle \bar{p} \rangle : \bar{x} \in \mathbb{Z}_{p^s}\}$
- ii) $\langle \bar{p} \rangle = \{\bar{c}\bar{p} : \bar{c} \in \mathbb{Z}_{p^s}\}$

Sea $y \in \frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle}$ entonces existe $\bar{x} \in \mathbb{Z}_{p^s}$ tal que $y = \bar{x} + \langle \bar{p} \rangle$, dividiendo a x por p tenemos que existen $r, s \in \mathbb{Z}$ tales que $x = sp + r$ con $0 \leq r \leq p-1$ entonces

$$\begin{aligned} y &= \overline{sp+r} + \langle \bar{p} \rangle \\ &= \bar{r} + \bar{s}\bar{p} + \langle \bar{p} \rangle \\ &= \bar{r} + \langle \bar{p} \rangle \end{aligned}$$

es decir:

$$\frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle} = \{\bar{r} + \langle \bar{p} \rangle \mid r \in \{0, 1, \dots, p-1\}\}$$

así, $\frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle}$ es un campo finito con p elementos, por lo tanto, $\frac{\mathbb{Z}_{p^s}}{\langle \bar{p} \rangle} \simeq \mathbb{F}_p$. \square

Lema 2.6. Sean $b \in \mathbb{Z}$ con $b > 1$ y $a \in \mathbb{N}$. Entonces existen $d_0, d_1, \dots, d_{n-1}, t \in \mathbb{Z}$ con $0 \leq d_i \leq b-1$ para cada $i \in \{0, 1, \dots, n-1\}$ tales que:

$$a = d_0 b^0 + d_1 b^1 + \dots + d_{n-1} b^{n-1} + t b^n$$

Demostración. Dividiendo a por b tenemos que, existen $d_0, s_0 \in \mathbb{Z}$ tales que $a = s_0 b + d_0$ con $0 \leq d_0 \leq b-1$. Si dividimos ahora, s_0 por b , obtenemos $d_1, s_1 \in \mathbb{Z}$ tales que $s_0 = s_1 b + d_1$ entonces:

$$\begin{aligned} a &= (s_1 b + d_1) b + d_0 \\ &= s_1 b^2 + d_1 b + d_0 \end{aligned}$$

con $0 \leq d_1 \leq b-1$. Procediendo de esta manera $n-1$ veces, tenemos:

$$a = s_{n-1} b^n + d_{n-1} b^{n-1} + d_{n-2} b^{n-2} + \dots + d_1 b + d_0$$

con $d_i \in \{0, 1, \dots, b-1\}$ para $i \in \{0, 1, \dots, n-1\}$, observese que la sucesión de cocientes satisface:

$$a > s_0 > s_1 > \dots \geq 0$$

debido a que la sucesión s_0, s_1, \dots es una sucesión decreciente de enteros no negativos. Tomando, $t = s_{n-1}$ se concluye que:

$$a = d_0 b^0 + d_1 b^1 + \dots + d_{n-1} b^{n-1} + t b^n$$

como queríamos. \square

Por el Lema 2.6, dado $x \in \{0, 1, \dots, p^s - 1\}$ tenemos que:

$$x = c_0 p^0 + c_1 p^1 + \dots + c_{s-1} p^{s-1} + t p^s$$

con $c_i \in \{0, 1, \dots, p-1\}$ para cada $i \in \{0, 1, \dots, s-1\}$ entonces

$$\begin{aligned} \bar{x} &= \overline{c_0 p^0 + c_1 p^1 + \dots + c_{s-1} p^{s-1} + t p^s} \\ &= \overline{c_0 p^0} + \overline{c_1 p} + \dots + \overline{c_{s-1} p^{s-1}} + \overline{t p^s} \\ &= \overline{c_0 p^0} + \overline{c_1 p} + \dots + \overline{c_{s-1} p^{s-1}}, \end{aligned} \tag{7}$$

puesto que $\overline{p^s} = \bar{0} \in \mathbb{Z}_{p^s}$. Definamos ahora la función:

$$\mu : \mathbb{Z}_{p^s} \longrightarrow \mathbb{F}_p \tag{8}$$

dada por $\mu(\bar{x}) = \mu(\overline{c_0 p^0} + \overline{c_1 p} + \dots + \overline{c_{s-1} p^{s-1}}) = c_0$, para cada $\bar{x} \in \mathbb{Z}_{p^s}$.

Lema 2.7. *La función μ es un homomorfismo de anillos con $\text{Ker } \mu = \langle \bar{p} \rangle$.*

Demostración. Sean $\bar{x} = \sum_{i=0}^{s-1} \bar{c}_i \bar{p}^i, \bar{y} = \sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i \in \mathbb{Z}_{p^s}$, entonces

$$\bar{x} + \bar{y} = \sum_{i=0}^{s-1} (\bar{c}_i + \bar{d}_i) \bar{p}^i = \sum_{i=0}^{s-1} (\overline{c_i + d_i}) \bar{p}^i,$$

de ahí que,

$$\begin{aligned} \mu(\bar{x} + \bar{y}) &= \mu\left(\sum_{i=0}^{s-1} (\overline{c_i + d_i}) \bar{p}^i\right) = c_0 + d_0 \\ &= \mu\left(\sum_{i=0}^{s-1} \bar{c}_i \bar{p}^i\right) + \mu\left(\sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) = \mu(\bar{x}) + \mu(\bar{y}), \end{aligned}$$

esto es, $\mu(\bar{x} + \bar{y}) = \mu(\bar{x}) + \mu(\bar{y})$ para toda $\bar{x}, \bar{y} \in \mathbb{Z}_{p^s}$. Por otro lado,

$$\begin{aligned} \mu(\bar{x}\bar{y}) &= \mu\left(\left(\bar{c}_0 + \bar{c}_1 \bar{p} + \bar{c}_2 \bar{p}^2 + \dots + \bar{c}_{s-1} \bar{p}^{s-1}\right) \sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) \\ &= \mu\left(\bar{c}_0 \left(\sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) + \dots + \bar{c}_{s-1} \left(\sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) \bar{p}^{s-1}\right) \\ &= \mu\left(\bar{c}_0 \bar{d}_0 + \bar{c}_0 \left(\sum_{i=1}^{s-1} \bar{d}_i \bar{p}^i\right) + \dots + \bar{c}_{s-1} \left(\sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) \bar{p}^{s-1}\right) \\ &= \mu\left(\bar{c}_0 \bar{d}_0 + \bar{c}_0 \left(\sum_{i=1}^{s-1} \bar{d}_i \bar{p}^i\right) + \dots + \bar{c}_{s-1} \left(\sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) \bar{p}^{s-1}\right) \\ &= c_0 d_0 = \mu\left(\sum_{i=0}^{s-1} \bar{c}_i \bar{p}^i\right) \mu\left(\sum_{i=0}^{s-1} \bar{d}_i \bar{p}^i\right) = \mu(\bar{x}) \mu(\bar{y}) \end{aligned}$$

esto es, $\mu(\bar{x}\bar{y}) = \mu(\bar{x}) \mu(\bar{y})$ para toda $\bar{x}, \bar{y} \in \mathbb{Z}_{p^s}$. Por lo tanto, μ es un homomorfismo.

Sea $\bar{x} \in \langle \bar{p} \rangle$, entonces existe $\bar{c} \in \mathbb{Z}_{p^s}$ tal que $\bar{x} = \bar{c}\bar{p}$. Entonces $\bar{x} = \bar{0}\bar{p}^0 + \bar{c}\bar{p} + \dots + \bar{0}\bar{p}^{s-1}$, luego $\mu(\bar{x}) = \bar{0}$, por lo tanto $\bar{x} \in \text{Ker } \mu$, de ahí que, $\langle \bar{p} \rangle \subseteq \text{Ker } \mu$. Ahora bien, sea $\bar{x} \in \text{Ker } \mu$ entonces $\mu(\bar{x}) = 0$, luego $\bar{x} = \bar{0} + \bar{c}_1 \bar{p} + \dots + \bar{c}_{s-1} \bar{p}^{s-1}$. Como $\bar{c}_i \bar{p}^i \in \langle \bar{p}^i \rangle \subseteq \langle \bar{p} \rangle$, por (5), para cada $i \in \{1, \dots, s-1\}$, esto es, $\bar{c}_i \bar{p}^i \in \langle \bar{p} \rangle$ para cada $i \in \{1, 2, \dots, s-1\}$, así, $\bar{x} \in \langle \bar{p} \rangle$ entonces $\text{Ker } \mu \subseteq \langle \bar{p} \rangle$. Por lo tanto, $\text{Ker } \mu = \langle \bar{p} \rangle$. \square

$\mathcal{R}[x]$ denota el anillo de polinomios en la indeterminada x con coeficientes en el anillo \mathcal{R} , esto es, el conjunto

$$\mathcal{R}[x] = \left\{ f(x) = \sum_{i=0}^n c_i x^i : c_i \in \mathcal{R}, i \in \{0, \dots, n\} \text{ para algún } n \in \mathbb{N} \right\} \quad (9)$$

es un anillo con las operaciones habituales de suma y multiplicación de polinomios.

Podemos extender el homomorfismo definido en (8) de la manera siguiente: sea la función

$$\bar{\cdot} : \mathbb{Z}_{p^s}[x] \longrightarrow \mathbb{F}_p[x] \quad (10)$$

dada por $\overline{-(f(x))} = \sum_{i=0}^n \mu(c_i) x^i$ para cada $f(x) \in \mathbb{Z}_{p^s}[x]$ y denotamos por $\overline{f(x)}$ la imagen de $f(x)$ bajo “ $\bar{\cdot}$ ”.

Como una consecuencia directa del lema anterior tenemos:

Corolario 2.8. *La función $\bar{\cdot}$ es un homomorfismo de anillos.*

Demostración. Recordemos que si $c_i \in \mathbb{Z}_{p^s}$, por (7) tenemos que:

$$c_i = d_{i0} + d_{i1}p + \dots + d_{i(s-1)}p^{s-1} \text{ con } d_{ik} \in \{0, 1, \dots, p-1\}$$

Sean $f(x), g(x) \in \mathbb{Z}_{p^s}[x]$ entonces $f(x) = \sum_{i=0}^n c_i x^i$ y $g(x) = \sum_{i=0}^m d_i x^i$, con $c_i, d_i \in \mathbb{Z}_{p^s}$ para cada i . Sin pérdida de generalidad, supóngase que $m > n$ entonces $f(x) + g(x) = \sum_{i=0}^m (c_i + d_i) x^i$ con $c_i = 0$ para cada $n < i \leq m$. Entonces

$$\begin{aligned} \overline{f+g}(x) &= \sum_{i=0}^m \mu(c_i + d_i) x^i \\ &= \sum_{i=0}^m (c_{i0} + d_{i0}) x^i \\ &= \sum_{i=0}^m c_{i0} x^i + \sum_{i=0}^m d_{i0} x^i \\ &= \sum_{i=0}^n c_{i0} x^i + \sum_{i=0}^m d_{i0} x^i \\ &= \sum_{i=0}^n \mu(c_i) x^i + \sum_{i=0}^m \mu(d_i) x^i \\ &= \overline{f}(x) + \overline{g}(x). \end{aligned}$$

Sabemos que $f(x)g(x) = \sum_{k=0}^l e_k x^k$ donde $e_k = c_0 d_k + c_1 d_{k-1} + \dots + c_k d_0$ y $0 \leq l \leq m+n$, entonces

$$\begin{aligned}\mu(e_k) &= \mu(c_0 d_k + c_1 d_{k-1} + \dots + c_k d_0) \\ &= \mu(c_0)\mu(d_k) + \mu(c_1)\mu(d_{k-1}) + \dots + \mu(c_k)\mu(d_0).\end{aligned}$$

Por otro lado, tenemos que:

$$\begin{aligned}\bar{f}(x)\bar{g}(x) &= \left(\sum_{i=0}^n \mu(c_i)x^i\right) \left(\sum_{i=0}^m \mu(d_i)x^i\right) \\ &= (\mu(c_0) + \mu(c_1)x + \dots + \mu(c_n)x^n) (\mu(d_0) + \mu(d_1)x + \dots + \mu(d_m)x^m),\end{aligned}$$

pero esto lo podemos expresar como $\sum_{k=0}^l \mu(u_k)x^k$ con

$$\mu(u_k) = (\mu(c_0)\mu(d_k) + \dots + \mu(c_k)\mu(d_0)) = \mu(e_k),$$

de esto se sigue:

$$\bar{f}g(x) = \sum_{k=0}^l \mu(e_k)x^k = \sum_{k=0}^l \mu(u_k)x^k = \bar{f}(x)\bar{g}(x),$$

por lo tanto, “ $-$ ” es un homomorfismo. □

Definición 2.9. Sea $\bar{p} \in \mathbb{Z}_p^s \subseteq \mathbb{Z}_p^s[x]$, el ideal generado por $\bar{p} \in \mathbb{Z}_p^s[x]$ se denota por $\ll \bar{p} \gg$ y se define como:

$$\ll \bar{p} \gg := \{f(x)\bar{p} : f(x) \in \mathbb{Z}_p^s[x]\}.$$

Lema 2.10. $\text{Ker } - = \ll \bar{p} \gg$.

Demostración. Sea $h(x) = \sum_{i=0}^n c_i x^i$. Si $h(x) \in \text{Ker } -$ entonces $\bar{h}(x) = \bar{0} \in \mathbb{Z}_p^s$. Pero $\bar{h}(x) = \sum_{i=0}^n \mu(c_i)x^i$, es decir $\mu(c_i) = 0$ con $i \in \{0, 1, \dots, n\}$, entonces $c_i \in \text{Ker } \mu = \langle \bar{p} \rangle$, por el Lema 2.8. Así, existen $\bar{m}_0, \bar{m}_1, \dots, \bar{m}_n \in \mathbb{Z}_p^s$ tales que $c_i = \bar{m}_i \bar{p}$, de ahí que:

$$h(x) = \sum_{i=0}^n c_i x^i = \sum_{i=0}^n (\bar{m}_i \bar{p}) x^i = \left(\sum_{i=0}^n \bar{m}_i x^i\right) \bar{p} = g(x)\bar{p}$$

con $g(x) = \sum_{i=0}^n \bar{m}_i x^i$, luego $h(x) \in \ll \bar{p} \gg$, por consiguiente, $\text{Ker } - \subseteq \ll \bar{p} \gg$. Ahora bien, si $h(x) \in \ll \bar{p} \gg$ entonces existe $g(x) \in \mathbb{Z}_p^s[x]$ tal que $h(x) = g(x)\bar{p}$,

sea $g(x) = \sum_{i=0}^n a_i x^i$ entonces

$$\begin{aligned} h(x) &= \sum_{i=0}^n a_i \bar{p} x^i \\ &= \sum_{i=0}^n c_i x^i \text{ con } c_i = \overline{a_i \bar{p}} \in \langle \bar{p} \rangle \end{aligned}$$

aplicando “ $-$ ” obtenemos:

$$\bar{h}(x) = \sum_{i=0}^n \mu(c_i) x^i = \sum_{i=0}^n 0 x^i = \bar{0}$$

luego, $h(x) \in \text{Ker } -$, de ahí que, $\langle \bar{p} \rangle \subseteq \text{Ker } -$, por lo tanto, $\text{Ker } - = \langle \bar{p} \rangle$. \square

Algunas veces denotaremos a $f(x)$ en $\mathbb{Z}_{p^s}[x]$ o $\mathbb{F}_p[x]$ simplemente por f .

Teorema 2.11. *El ideal $\langle \bar{p} \rangle$ es un ideal primo de $\mathbb{Z}_{p^s}[x]$, todo ideal primo de $\mathbb{Z}_{p^s}[x]$ contiene a $\langle \bar{p} \rangle$, más aún, si un ideal primo contiene a $\langle \bar{p} \rangle$ propiamente, entonces dicho ideal es maximal.*

Demostración. Sean $f(x), h(x) \in \mathbb{Z}_{p^s}[x]$ tales que $f(x) = \sum_{i=0}^n a_i x^i$ y $h(x) = \sum_{i=0}^m b_i x^i$ con $a_n \neq 0 \neq b_m$, de ahí que, los grados de $f(x)$ y $g(x)$ son n y m respectivamente, lo cual denotamos por $\text{grad}(f(x)) = n$ y $\text{grad}(h(x)) = m$, entonces se tiene que $f(x)h(x) = \sum_{k=0}^l c_k x^k$ donde $c_k = \sum_{s=0}^k a_s b_{k-s}$ y si $f(x)h(x) \neq 0$ entonces $\text{grad}(f(x)h(x)) = l$ con $0 \leq l \leq m+n$. Supóngase, sin pérdida de generalidad, que $m > n$ y además, supóngase que $f(x)h(x) \in \langle \bar{p} \rangle$ entonces

$$\overline{f(x)h(x)} = 0 \tag{11}$$

dado que $\text{Ker } - = \langle \bar{p} \rangle$, pero $\overline{f(x)h(x)} = \sum_{k=0}^l \mu(c_k) x^k$ entonces $\mu(c_k) = 0$ para cada $k \in \{0, 1, \dots, l\}$. Como $\mu(c_k) = \mu(a_0)\mu(b_k) + \dots + \mu(a_k)\mu(b_0) \in \mathbb{F}_p$ para cada $k \in \{0, 1, \dots, l\}$ tenemos que para $k = 0$, $0 = \mu(c_0) = \mu(a_0)\mu(b_0)$, pero \mathbb{F}_p es un dominio entero, así $\mu(a_0) = 0$ o $\mu(b_0) = 0$. Supóngase que $\mu(a_0) \neq 0$, entonces $\mu(b_0) = 0$. Ahora bien, cuando $k = 1$, $0 = \mu(c_1) = \mu(a_1)\mu(b_0) + \mu(a_0)\mu(b_1) = \mu(a_0)\mu(b_1)$ puesto que $\mu(b_0) = 0$ y como $\mu(a_0) \neq 0$ tenemos que $\mu(b_1) = 0$, procediendo de esta manera, establecemos que $\mu(b_k) = 0$ para cada $k \in \{0, 1, \dots, m\}$ de ahí que $\bar{h}(x) = 0$, por consiguiente, $h(x) \in \langle \bar{p} \rangle$. Por otro lado, supóngase que $\mu(b_0) \neq 0$ entonces $\mu(a_0) = 0$ y de manera análoga se tiene que $\bar{f}(x) = 0$, en consecuencia, $f(x) \in \langle \bar{p} \rangle$, de ahí que, si $f(x)h(x) \in \langle \bar{p} \rangle$ entonces $f(x) \in \langle \bar{p} \rangle$ o bien $h(x) \in \langle \bar{p} \rangle$. Por lo tanto, $\langle \bar{p} \rangle$ es un ideal primo de $\mathbb{Z}_{p^s}[x]$. Sea \mathbb{P} un ideal primo de \mathbb{Z}_{p^s} . Como $\bar{p} \cdot \overline{p^{s-1}} = \bar{p}^s = \bar{0} \in \mathbb{P}$, entonces

$\bar{p} \in \mathbb{P}$ o $\overline{p^{s-1}} \in \mathbb{P}$, ya que \mathbb{P} es un ideal primo. Si $\bar{p} \in \mathbb{P}$ entonces $\ll \bar{p} \gg \subseteq \mathbb{P}$, ahora bien, si $p^{s-1} \in \mathbb{P}$, entonces $p^{s-2} \cdot \bar{p} = \overline{p^{s-1}} \in \mathbb{P}$, de ahí que, $\bar{p} \in \mathbb{P}$ o $p^{s-2} \in \mathbb{P}$, una vez más, si $\bar{p} \in \mathbb{P}$ concluimos que $\ll \bar{p} \gg \subseteq \mathbb{P}$, o bien, si $p^{s-2} \in \mathbb{P}$ se procede como antes, y continuando de esta manera se concluye que $\bar{p} \in \mathbb{P}$, por lo tanto, $\ll \bar{p} \gg \subseteq \mathbb{P}$. Ahora probaremos que si \mathbb{P} es un ideal primo y $\ll \bar{p} \gg \subsetneq \mathbb{P}$ entonces \mathbb{P} es maximal. Veamos primero que $-$ es un epimorfismo. Sean $g(x) = \sum_{i=0}^n a_i x^i \in \mathbb{F}_p[x]$ y $f(x) = \sum_{i=0}^n b_i x^i \in \mathbb{Z}_{p^s}[x]$, se estableció en (7) que para cada $i \in \{0, 1, \dots, n\}$, $b_i = c_{i0} + c_{i1}p + c_{i2}p^2 + \dots + c_{i(s-1)}p^{s-1}$ donde $c_{ij} \in \mathbb{F}_p$ para cada $i \in \{0, \dots, n\}$, $j \in \{1, 2, \dots, s-1\}$. Sean los b_i 's de tal forma que $c_{i0} = a_i$ para cada $i \in \{0, 1, \dots, n\}$ entonces

$$\bar{f}(x) = \sum_{i=0}^n \mu(b_i)x^i = \sum_{i=0}^n c_{i0}x^i = \sum_{i=0}^n a_i x^i = g(x)$$

así, $-$ es sobre y por tanto epimorfismo. Si denotamos por $\bar{\mathbb{P}}$ a la imagen de \mathbb{P} bajo $-$ y a $-^{-1}(\bar{\mathbb{P}})$ como la imagen inversa de $\bar{\mathbb{P}}$ bajo $-$ entonces $-^{-1}(\bar{\mathbb{P}}) = \mathbb{P}$ y por un teorema de isomorfismos (ver [16, Theorem 12.7, iii]) la función:

$$\psi : \frac{\mathbb{Z}_{p^s}[x]}{\mathbb{P}} \longrightarrow \frac{\mathbb{F}_p[x]}{\bar{\mathbb{P}}}$$

es un isomorfismo de anillos. Por lo tanto $\frac{\mathbb{Z}_{p^s}[x]}{\mathbb{P}} \simeq \frac{\mathbb{F}_p[x]}{\bar{\mathbb{P}}}$.

Como \mathbb{P} es un ideal primo entonces $\mathbb{P} \neq \mathbb{Z}_{p^s}[x]$. Si $\bar{\mathbb{P}} = \mathbb{F}_p[x]$, por lo anterior, se tiene que: $-^{-1}(\bar{\mathbb{P}}) = -^{-1}(\mathbb{F}_p[x]) = \mathbb{Z}_{p^s}[x]$, la última igualdad debido a que $-$ es sobreyectiva, de ahí que, $\mathbb{P} = \mathbb{Z}_{p^s}[x]$, lo cual es una contradicción, por consiguiente, $\bar{\mathbb{P}} \neq \mathbb{F}_p[x]$. Sean $f_0(x), g_0(x) \in \mathbb{F}_p[x]$ tales que $f_0(x)g_0(x) \in \bar{\mathbb{P}}$ entonces existen $f(x), g(x) \in \mathbb{Z}_{p^s}[x]$ con $\bar{f}(x) = f_0(x)$ y $\bar{g}(x) = g_0(x)$, luego $\bar{f}(x)\bar{g}(x) \in \bar{\mathbb{P}}$, de ahí que, $f(x)g(x) \in \mathbb{P}$, pero \mathbb{P} es un ideal primo por lo que $f(x) \in \mathbb{P}$ o $g(x) \in \mathbb{P}$. Si $f(x) \in \mathbb{P}$ entonces $f_0(x) = \bar{f}(x) \in \bar{\mathbb{P}}$ o bien, si $g(x) \in \mathbb{P}$ entonces $g_0(x) = \bar{g}(x) \in \bar{\mathbb{P}}$, por consiguiente, $\bar{\mathbb{P}}$ es un ideal primo de $\mathbb{F}_p[x]$. Sin embargo, $\bar{\mathbb{P}} \neq \langle 0 \rangle$, debido a que, $\ll \bar{p} \gg \subsetneq \mathbb{P}$, ya que, existe $h(x) \in \mathbb{Z}_{p^s}[x]$ tal que $h(x) \in \mathbb{P}$ pero $h(x) \notin \ll \bar{p} \gg$, entonces $\bar{h}(x) \in \bar{\mathbb{P}}$ y $\bar{h}(x) \neq \bar{0}$, por lo tanto, $\bar{\mathbb{P}} \neq \langle 0 \rangle$, en consecuencia, por [16, Theorem 12.27], $\bar{\mathbb{P}}$ es maximal, así $\frac{\mathbb{F}_p[x]}{\bar{\mathbb{P}}}$ es un campo y como $\frac{\mathbb{Z}_{p^s}[x]}{\mathbb{P}} \simeq \frac{\mathbb{F}_p[x]}{\bar{\mathbb{P}}}$ se concluye que \mathbb{P} es maximal. \square

A continuación estudiaremos otro tipo de ideales, estos son los ideales primarios de $\mathbb{Z}_{p^s}[x]$, para este propósito se enuncia la definición siguiente:

Definición 2.12. Sean \mathcal{R} un anillo, $I \subseteq \mathcal{R}$ un ideal y $r, s \in \mathcal{R}$. Diremos que:

- I es un ideal **primario** si y sólo si $rs \in I$ implica que $r \in I$ o existe un $n \in \mathbb{N}$ tal que $s^n \in I$.

- El **radical de I** se denota por \sqrt{I} y es el conjunto:

$$\sqrt{I} := \{r \in \mathcal{R} \mid \exists n \in \mathbb{N} : r^n \in I\}.$$

Lema 2.13. *Sea \mathcal{R} un anillo con unidad, e I un ideal tal que \sqrt{I} es un ideal primo. Entonces $I \neq \mathcal{R}$.*

Demostración. Es claro $\sqrt{I} \neq \mathcal{R}$, puesto que \sqrt{I} es un ideal primo. Si $I = \mathcal{R}$, se sigue que, $1 \in I$, luego dado $n = 1 \in \mathbb{N}$, se tiene que, $1^n = 1 \in I$, de ahí que, $1 \in \sqrt{I}$ entonces $\sqrt{I} = \mathcal{R}$, lo cual es una contradicción. \square

Teorema 2.14. *Sea Q un ideal de $\mathbb{Z}_{p^s}[x]$.*

- i) *Si Q es un ideal primo de $\mathbb{Z}_{p^s}[x]$ entonces \sqrt{Q} es un ideal primo.*
- ii) $\ll \bar{p} \gg \subseteq \sqrt{Q}$
- iii) *Si \sqrt{Q} es un ideal primo y $\ll \bar{p} \gg \subsetneq \sqrt{Q}$ entonces Q es un ideal primo.*

Demostración. Sea Q un ideal de $\mathbb{Z}_{p^s}[x]$.

- i) Supongamos que Q es primo, entonces $Q \neq \mathbb{Z}_{p^s}[x]$. Sean $f(x), g(x) \in \mathbb{Z}_{p^s}[x]$ tales que $f(x)g(x) \in \sqrt{Q}$, entonces existe un $n \in \mathbb{N}$ tal que $(f(x)g(x))^n \in Q$, i.e., $(f(x))^n \cdot (g(x))^n \in Q$. Como Q es primo, se sigue que $(f(x))^n \in Q$ o existe $m \in \mathbb{N}$ de manera que $((g(x))^{nm}) \in Q$. Si $(f(x))^n \in Q$ entonces $f(x) \in \sqrt{Q}$, o bien, $(g(x))^{nm} \in Q$ entonces $g(x) \in \sqrt{Q}$, pues $nm \in \mathbb{N}$. Por lo tanto, \sqrt{Q} es un ideal primo.
- ii) Como $\bar{p}^s = 0$ y Q es un ideal, entonces $\bar{p}^s \in Q$, pero $(\bar{p})^s = \bar{p}^s \in Q$, es decir., existe $s \in \mathbb{N}$ tal que $((\bar{p})^s \in Q)$, por lo tanto, $\bar{p} \in \sqrt{Q}$ y, en consecuencia, $\ll \bar{p} \gg \subseteq \sqrt{Q}$.
- iii) Suponga que \sqrt{Q} es un ideal primo y además que $\ll \bar{p} \gg \subsetneq \sqrt{Q}$. Por el Teorema 2.11, \sqrt{Q} es maximal y como $\sqrt{Q} \neq \mathbb{Z}_{p^s}[x]$, se sigue del Lema 2.13 que $Q \neq \mathbb{Z}_{p^s}[x]$. Sean $f(x), g(x) \in \mathbb{Z}_{p^s}[x]$ tales que $f(x)g(x) \in Q$. Suponga ahora que para todo $m \in \mathbb{N}$ se cumple que $((g(x))^m \notin Q)$ entonces $g(x) \notin \sqrt{Q}$. Definimos $\mathcal{J}(g(x)) := \{t(x)g(x) + q(x) \mid t(x) \in \mathbb{Z}_{p^s}[x] \text{ y } q(x) \in \sqrt{Q}\}$ y sean $l(x), k(x) \in \mathcal{J}(g(x))$, entonces existen $t_1(x), t_2(x) \in \mathbb{Z}_{p^s}[x]$ y $q_1(x), q_2(x) \in \sqrt{Q}$ tales que $l(x) = t_1(x)g(x) + q_1(x)$ y $k(x) = t_2(x)g(x) + q_2(x)$, de ahí que,

$$\begin{aligned} l(x) - k(x) &= t_1(x)g(x) + q_1(x) - t_2(x)g(x) - q_2(x) \\ &= (t_1(x) - t_2(x))g(x) + (q_1(x) - q_2(x)) \\ &= t(x)g(x) + r(x) \end{aligned}$$

con $t(x) = t_1(x) - t_2(x) \in \mathbb{Z}_{p^s}[x]$ y $r(x) = q_1(x) - q_2(x) \in \sqrt{Q}$ entonces $l(x) - k(x) \in \mathcal{J}(g(x))$, por lo tanto, $(\mathcal{J}(g(x)), +) < (\mathbb{Z}_{p^s}[x], +)$. Ahora, si $h(x) \in \mathcal{J}(g(x))$ y $k(x) \in \mathbb{Z}_{p^s}[x]$, entonces existen $q(x) \in \sqrt{Q}$ y $t(x) \in \mathbb{Z}_{p^s}[x]$ tales que $h(x) = t(x)g(x) + q(x)$, así $k(x)h(x) = k(x)t(x)g(x) + k(x)q(x) = s(x)g(x) + r(x)$ con $s(x) = k(x)t(x) \in \mathbb{Z}_{p^s}[x]$ y $r(x) = k(x)q(x) \in \sqrt{Q}$, por consiguiente, $k(x)h(x) \in \mathcal{J}(g(x))$, en consecuencia, $\mathcal{J}(g(x))$ es un ideal. Sea $q(x) \in \sqrt{Q}$, es claro que $q(x) = t(x)g(x) + q(x)$ con $t(x) = 0$ entonces $q(x) \in \mathcal{J}(g(x))$, así, $\sqrt{Q} \subseteq \mathcal{J}(g(x))$, también tenemos que si $t(x) \in \mathbb{Z}_{p^s}[x]$ con $t(x) \neq 0$ y $t(x) \notin \sqrt{Q}$ entonces $t(x)g(x) + q(x) \notin \sqrt{Q}$, pues de lo contrario, $t(x)g(x) + q(x) - q(x) = t(x)g(x) \in \sqrt{Q}$ y \sqrt{Q} es un ideal primo pero ni $t(x)$ ni $g(x)$ pertenecen a \sqrt{Q} . Así, $\mathcal{J}(g(x)) \neq \sqrt{Q}$ y como \sqrt{Q} es maximal, entonces $\mathcal{J}(g(x)) = \mathbb{Z}_{p^s}[x]$. Por lo anterior, $1 \in \mathcal{J}(g(x))$, entonces existen $t(x) \in \mathbb{Z}_{p^s}[x]$ y $q(x) \in \sqrt{Q}$ tales que $1 = t(x)g(x) + q(x)$ y como $q(x) \in \sqrt{Q}$ entonces existe $n \in \mathbb{N}$ tal que $(q(x))^n \in Q$ entonces

$$\begin{aligned} 1^n &= (t(x)g(x) + q(x))^n \\ &= \sum_{i=0}^n \binom{n}{i} (t(x))^{n-i} (g(x))^{n-i} (q(x))^i \\ &= g(x) \sum_{i=0}^n \binom{n}{i} (t(x))^{n-i} (g(x))^{n-i-1} (q(x))^i \\ &= g(x) \underbrace{\sum_{i=0}^{n-1} \binom{n}{i} (t(x))^{n-i} (g(x))^{n-i-1} (q(x))^i}_{T(x)} + (q(x))^n, \end{aligned}$$

esto es, $1 = T(x)g(x) + (q(x))^n$, multiplicando por $f(x)$ obtenemos que $f(x) = T(x)f(x)g(x) + f(x)(q(x))^n$, luego, como $f(x)g(x) \in Q$, $(q(x))^n \in Q$ y Q es un ideal, se sigue que $f(x) \in Q$. Por lo tanto, Q es un ideal primario. \square

Definición 2.15. Sea $h(x) \in \mathbb{Z}_{p^s}$. Diremos que $h(x)$ es un polinomio **primario** si y sólo si $\langle h(x) \rangle$ es un ideal primario de \mathbb{Z}_{p^s} .

Lema 2.16. Sea $f(x)$ un polinomio de $\mathbb{Z}_{p^s}[x]$, supóngase que $\bar{f}(x) = (g(x))^m$ donde $g(x)$ es un polinomio irreducible en $\mathbb{F}_p[x]$ y $m \in \mathbb{N}$. Entonces $f(x)$ es un polinomio primario.

Demostración. Como $\langle f(x) \rangle$ es un ideal de $\mathbb{Z}_{p^s}[x]$, por ii) del Teorema 2.14, se tiene que $\ll \bar{p} \gg \subseteq \sqrt{\langle f(x) \rangle}$. Dado que $(f(x))^1 \in \langle f(x) \rangle$ entonces $f(x) \in \sqrt{\langle f(x) \rangle}$ y $\bar{f}(x) \neq 0$, entonces $\ll \bar{p} \gg \subsetneq \sqrt{\langle f(x) \rangle}$. Veamos que $\sqrt{\langle f(x) \rangle}$ es un ideal primo. Si $1 \in \langle f(x) \rangle$ entonces existe $h(x) \in \mathbb{Z}_{p^s}[x]$ de tal manera que $1 = h(x)f(x)$ entonces $\bar{1} =$

$\bar{h}(x)\bar{f}(x) = \bar{h}(x)(g(x))^m = g(x)(\bar{h}(x)(g(x))^{m-1})$ pero esto implica que $g(x)$ es una unidad, lo cual es una contradicción ya que $g(x)$ es irreducible, entonces $\bar{1} \notin \langle f(x) \rangle$, se sigue que, $\langle f(x) \rangle \neq \mathbb{Z}_{p^s}[x]$ y así $\sqrt{\langle f(x) \rangle} \neq \mathbb{Z}_{p^s}[x]$. Sean ahora $a(x), b(x) \in \mathbb{Z}_{p^s}[x]$ y suponga que $a(x)b(x) \in \sqrt{\langle f(x) \rangle}$, entonces existe $n \in \mathbb{N}$ tal que $(a(x)b(x))^n \in \langle f(x) \rangle$, es decir, existe $q(x) \in \mathbb{Z}_{p^s}$ tal que $(a(x)b(x))^n = (a(x))^n(b(x))^n = q(x)f(x)$. Aplicando el epimorfismo (10) tenemos que $(\bar{a}(x))^n(\bar{b}(x))^n = \bar{q}(x)\bar{f}(x)$, i.e.,

$$\begin{aligned}\bar{a}(x)^n\bar{b}(x)^n &= \bar{q}(x)\bar{f}(x) \\ &= \bar{q}(x)(g(x))^m\end{aligned}$$

de la última igualdad se tiene que $g(x)$ divide al producto $\bar{a}(x)^n\bar{b}(x)^n$ y, como $g(x)$ es irreducible entonces divide a $\bar{a}(x)$ o bien a $\bar{b}(x)$. Supóngase que $g(x)|\bar{a}(x)$ entonces $(g(x))^m = \bar{f}(x)$ divide a $(\bar{a}(x))^m$, es decir, existe $\bar{c}(x) \in \mathbb{F}_p[x]$ tal que $(\bar{a}(x))^m = \bar{c}(x)\bar{f}(x)$, entonces $(\bar{a}(x))^m - \bar{c}(x)f(x) = \bar{0}$ y como (10) es epimorfismo tenemos que $(a(x))^m - c(x)f(x) = \bar{0}$, de ahí que, $(a(x))^m - c(x)f(x) \in Ker = \ll \bar{p} \gg$. Entonces existe $d(x) \in \mathbb{Z}_{p^s}[x]$ tal que $(a(x))^m - c(x)f(x) = d(x)\bar{p}$, por lo tanto, $(a(x))^m = c(x)f(x) + d(x)\bar{p}$. Como $p^s \geq 2$, elevando la igualdad anterior a p^s obtenemos que $((a(x))^m)^{p^s} = (c(x)f(x) + d(x)\bar{p})^{p^s}$ y, aplicando el teorema del binomio tenemos:

$$\begin{aligned}(a(x))^{mp^s} &= \sum_{i=0}^{p^s} \binom{p^s}{i} (c(x)f(x))^{p^s-i} (d(x)\bar{p})^i \\ &= \underbrace{\left[\sum_{i=0}^{p^s-1} \binom{p^s}{i} (c(x))^{p^s-i} (f(x))^{p^s-i-1} (d(x)\bar{p})^i \right]}_{k(x)} f(x) + (d(x)\bar{p})^{p^s}\end{aligned}$$

es decir, $(a(x))^{mp^s} = k(x)f(x) + (d(x))^{p^s}(\bar{p})^{p^s}$ pero $s < p^s$ entonces existe $t \in \mathbb{N}$ tal que $p^s = s + t$, luego $(a(x))^{mp^s} = k(x)f(x) + \bar{p}^s\bar{p}^t = k(x)f(x)$ puesto que $\bar{p}^s = \bar{0}$, de ahí que, $a(x) \in \sqrt{\langle f(x) \rangle}$. Análogamente si $g(x)$ divide a $\bar{b}(x)$ se concluye que $b(x) \in \sqrt{\langle f(x) \rangle}$. Por lo tanto $\sqrt{\langle f(x) \rangle}$ es un ideal primo. Luego, por la parte *iii* del Teorema 2.14, $\langle f(x) \rangle$ es un ideal primario y, por la Definición 2.15, $f(x)$ es un polinomio primario. \square

3 El Lema de Hensel

Definición 3.1. Sean $f_1, f_2 \in \mathbb{Z}_{p^s}[x]$. Diremos que f_1 y f_2 son **coprimos** en $\mathbb{Z}_{p^s}[x]$ si existen $\lambda_1, \lambda_2 \in \mathbb{Z}_{p^s}[x]$ tales que

$$\lambda_1 f_1 + \lambda_2 f_2 = 1.$$

Cabe mencionar que en la Definición 3.1, $\mathbb{Z}_{p^s}[x]$ puede ser reemplazado por $\mathbb{F}_p[x]$.

Lema 3.2. Sean $f_1, f_2 \in \mathbb{Z}_{p^s}[x]$. Entonces f_1 y f_2 son coprimos en $\mathbb{Z}_{p^s}[x]$ si y sólo si \bar{f}_1, \bar{f}_2 son coprimos en $\mathbb{F}_p[x]$.

Demostración. Sean f_1 y f_2 polinomios coprimos en $\mathbb{Z}_{p^s}[x]$, entonces por la Definición 3.1 existen $\lambda_1, \lambda_2 \in \mathbb{Z}_{p^s}[x]$ tales que $\lambda_1 f_1 + \lambda_2 f_2 = 1$ y, aplicando (10) tenemos que $\overline{\lambda_1 f_1 + \lambda_2 f_2} = \bar{1}$, es decir, $\overline{\lambda_1 f_1} + \overline{\lambda_2 f_2} = \bar{1}$, una vez más, por la Definición 3.1, se concluye que \bar{f}_1 y \bar{f}_2 son coprimos en $\mathbb{F}_p[x]$. Recíprocamente, sean \bar{f}_1 y \bar{f}_2 coprimos en $\mathbb{F}_p[x]$, luego por la Definición 3.1, existen $\mu_1, \mu_2 \in \mathbb{F}_p[x]$ tales que $\mu_1 \bar{f}_1 + \mu_2 \bar{f}_2 = \bar{1}$, pero (10) es un epimorfismo, entonces existen $\lambda_1, \lambda_2 \in \mathbb{Z}_{p^s}[x]$ con $\mu_1 = \bar{\lambda}_1$ y $\mu_2 = \bar{\lambda}_2$ tales que $\overline{\lambda_1 f_1 + \lambda_2 f_2} = \bar{1}$, es decir, $\overline{\lambda_1 f_1 + \lambda_2 f_2 - 1} = \bar{0}$, de la última igualdad, se sigue que, $\lambda_1 f_1 + \lambda_2 f_2 - 1 \in \text{Ker} - = \ll \bar{p} \gg$, entonces existe $k \in \mathbb{Z}_{p^s}[x]$ tal que $\lambda_1 f_1 + \lambda_2 f_2 - 1 = k\bar{p}$, de ahí que, $\lambda_1 f_1 + \lambda_2 f_2 = 1 + k\bar{p}$. Definimos $\sigma = \sum_{i=0}^{s-1} (-k\bar{p})^i$, luego tenemos que, $\sigma(\lambda_1 f_1 + \lambda_2 f_2) = \sigma(1 + k\bar{p})$, es decir,

$$\begin{aligned} (\sigma\lambda_1) f_1 + (\sigma\lambda_2) f_2 &= \sigma + \sigma k\bar{p} \\ &= \sigma + \left(\sum_{i=0}^{s-1} (-k\bar{p})^i \right) k\bar{p} \\ &= \sigma + \left((-k\bar{p})^0 + (-k\bar{p})^1 + \dots + (-k\bar{p})^{s-1} \right) k\bar{p} \\ &= \sigma + \left(1 - k\bar{p} + \dots + (-1)^{s-1} k^{s-1} \bar{p}^{s-1} \right) k\bar{p} \\ &= \sigma + k\bar{p} - k^2 \bar{p}^2 + \dots + (-1)^{s-2} k^{s-1} \bar{p}^{s-1} + (-1)^{s-1} k^s \bar{p}^s \\ &= 1 - k\bar{p} + k^2 \bar{p}^2 + \dots + (-1)^{s-1} k^{s-1} \bar{p}^{s-1} \\ &\quad + k\bar{p} - k^2 \bar{p}^2 + \dots + (-1)^{s-2} k^{s-1} \bar{p}^{s-1} + (-1)^{s-1} k^s \bar{p}^s \\ &= 1 \end{aligned}$$

ya que $\bar{p}^s = 0$, de ahí que, $(\sigma\lambda_1) f_1 + (\sigma\lambda_2) f_2 = 1$, por lo tanto, f_1 y f_2 son coprimos en $\mathbb{Z}_{p^s}[x]$. \square

Lema 3.3. Sea $f \in \mathbb{Z}_{p^s}[x]$ un polinomio mónico, supóngase que $\bar{f} = g_1 g_2 \in \mathbb{F}_p[x]$, donde $g_1, g_2 \in \mathbb{F}_p[x]$ son polinomios mónicos coprimos. Entonces existen $f_1, f_2 \in \mathbb{Z}_{p^s}[x]$ polinomios mónicos coprimos tales que

i) $f = f_1 f_2 \in \mathbb{Z}_{p^s}[x]$,

ii) $\bar{f}_i = g_i$ para $i \in \{1, 2\}$.

Demostración. Sea $f \in \mathbb{Z}_{p^s}[x]$ un polinomio mónico, tal que $\bar{f} = g_1 g_2$, donde $(g_1, g_2) = 1$ en $\mathbb{F}_p[x]$, entonces $\bar{f} - g_1 g_2 = \bar{0}$. Como (10) es un epimorfismo, existen $h_1, h_2 \in \mathbb{Z}_{p^s}[x]$ tales que $\bar{h}_1 = g_1$ y $\bar{h}_2 = g_2$, así $\bar{f} - \bar{h}_1 \bar{h}_2 = \overline{f - h_1 h_2} = \bar{0}$; por el Lema 2.10 tenemos que $f - h_1 h_2 \in \ll \bar{p} \gg$, es decir existe $k \in \mathbb{Z}_{p^s}$ tal que $f - h_1 h_2 = k\bar{p}$. Por consiguiente, $f = h_1 h_2 + k\bar{p}$. Como g_1 y g_2 son coprimos en $\mathbb{F}_p[x]$, se sigue del Lema 3.2 que h_1 y h_2 son coprimos en $\mathbb{Z}_{p^s}[x]$ luego, por la Definición 3.1, existen $\lambda_1 \lambda_2 \in \mathbb{Z}_{p^s}[x]$ tales que $\lambda_1 h_1 + \lambda_2 h_2 = 1$. Denotando por v a $k\bar{p}$ tenemos:

$$f = h_1 h_2 + v. \quad (12)$$

Definimos en $\mathbb{Z}_{p^s}[x]$ los polinomios siguientes:

$$\begin{aligned} h_{11} &= h_1 + \lambda_2 v \\ h_{21} &= h_2 + \lambda_1 v \end{aligned}$$

Como $v \in \ll \bar{p} \gg$ entonces $\bar{v} = \bar{0}$; aplicando (10) a cada polinomio, tenemos que $\overline{h_{11}} = \bar{h}_1 = g_1$ y $\overline{h_{21}} = \bar{h}_2 = g_2$ además, al multiplicar estos polinomios obtenemos:

$$\begin{aligned} h_{11} h_{21} &= h_1 h_2 + \lambda_1 h_1 v + \lambda_2 h_2 v + \lambda_1 \lambda_2 v^2 \\ &= h_1 h_2 + (\lambda_1 h_1 + \lambda_2 h_2) v + \lambda_1 \lambda_2 v^2 \\ &= h_1 h_2 + v + \lambda_1 \lambda_2 v^2, \end{aligned}$$

puesto que $\lambda_1 h_1 + \lambda_2 h_2 = 1$ y, usando (12) en la última igualdad se obtiene que $h_{11} h_{21} = f + \lambda_1 \lambda_2 v^2$, de ahí que, $f - h_{11} h_{21} = -\lambda_1 \lambda_2 v^2$, de manera que:

$$f \equiv h_{11} h_{21} \pmod{v^2}.^1 \quad (13)$$

Como $(g_1, g_2) = 1$, $\bar{h}_{11} = g_1$ y $\bar{h}_{21} = g_2$, se sigue del Lema 3.2 que h_{11} y h_{21} son coprimos, es decir existen $\lambda_{11}, \lambda_{21}$ tales que $\lambda_{11} h_{11} + \lambda_{21} h_{21} = 1$. Por (13) tenemos que $f = h_{11} h_{21} + k_1 v^2$ para algún $k_1 \in \mathbb{Z}_{p^s}[x]$, entonces repitiendo el proceso t veces definiendo en cada paso a los polinomios:

$$\begin{aligned} h_{1t} &= h_{1t-1} + \lambda_{2t-1} k_{t-1} v^{2(t-1)} \\ h_{2t} &= h_{2t-1} + \lambda_{1t-1} k_{t-1} v^{2(t-1)} \end{aligned}$$

donde $\overline{h_{1t}} = g_1$, $\overline{h_{2t}} = g_2$, $\lambda_{1t-1} h_{1t-1} + \lambda_{2t-1} h_{2t-1} = 1$ y $f = h_{1t-1} h_{2t-1} + k_{t-1} v^{2(t-1)}$. Es fácil ver que para $t = s$ al multiplicar los polinomios definidos arriba se tiene:

$$f \equiv h_{1s} h_{2s} \pmod{v^{2s}}.$$

¹Usamos la notación de congruencia modular ya que $v^2 | f - h_{11} h_{21}$.

Observe que $v^{2s} = (k\bar{p})^{2s} = k^{2s}\bar{p}^s\bar{p}^s = \bar{0}$, es decir, $f \equiv h_{1s}h_{2s}$ (mód \bar{p}^s), en otras palabras, $f - h_{1s}h_{2s} = \bar{0}$ y, por lo tanto, $f = h_{1s}h_{2s}$. Finalmente, renombrando $f_1 = h_{1s}$ y $f_2 = h_{2s}$ tenemos que: $f = f_1f_2$, $\bar{f}_1 = g_1$, $\bar{f}_2 = g_2$ y $(f_1, f_2) = 1$, con lo cual queda demostrado el lema. \square

A continuación mostraremos un resultado que nos será muy útil en la generalización del Lema 3.3.

Lema 3.4. *Sea R un anillo euclidiano y sean $p_1, p_2, \dots, p_r \in R$ coprimos por pares. Entonces, si $r \geq 3$ se tiene que $(\prod_{i=1}^{r-1} p_i, p_r) = 1$.*

Demostración. Haremos inducción sobre r . Supongamos que $r = 3$. Sean $p_1, p_2, p_3 \in R$ tales que $(p_i, p_j) = 1$ para $i \neq j$ con $i, j \in \{1, 2, 3\}$ entonces, por la Definición 3.1, tenemos que existen $a_1, a_2, b_1, b_2 \in R$ tales que $a_1p_1 + a_2p_3 = 1$ y $b_1p_2 + b_2p_3 = 1$. Multiplicando estas igualdades tenemos:

$$\begin{aligned} 1 &= (a_1p_1 + a_2p_3)(b_1p_2 + b_2p_3) \\ &= a_1b_1p_1p_2 + a_1b_2p_1p_3 + a_2b_1p_2p_3 + a_2b_2p_3^2 \\ &= (a_1b_1)p_1p_2 + (a_1b_2p_1 + a_2b_1p_2 + a_2b_2p_3)p_3 \\ &= \lambda_1p_1p_2 + \lambda_2p_3 \end{aligned}$$

donde $\lambda_1 = a_1b_1$, $\lambda_2 = a_1b_2p_1 + a_2b_1p_2 + a_2b_2p_3 \in R$, por lo tanto, $(p_1p_2, p_3) = 1$. Supóngase que este resultado se cumple para r , veamos que es válido para $r+1$. Sean $p_1, p_2, \dots, p_r, p_{r+1} \in R$ tales que $(p_i, p_j) = 1$ para $i \neq j$ con $i, j \in \{1, 2, \dots, r+1\}$, en particular tenemos que $(p_r, p_{r+1}) = 1$ y, por la hipótesis inductiva, tenemos que $(\prod_{i=1}^{r-1} p_i, p_{r+1}) = 1$ así, por la Definición 3.1 se tiene que existen $c_1, c_2, d_1, d_2 \in R$ tales que:

$$\begin{aligned} c_1p_r + c_2p_{r+1} &= 1, \\ d_1p_1p_2 \cdots p_{r-1} + d_2p_{r+1} &= 1, \end{aligned}$$

multiplicando las igualdades obtenemos:

$$\begin{aligned} 1 &= c_1d_1 \prod_{i=1}^{r-1} p_i p_r + c_1d_2 p_r p_{r+1} + c_2d_1 \prod_{i=1}^{r-1} p_i p_{r+1} + c_2d_2 p_{r+1}^2 \\ &= (c_1d_1) \prod_{i=1}^r p_i + \left(c_1d_2 p_r + c_2d_1 \prod_{i=1}^{r-1} p_i + c_2d_2 p_{r+1} \right) p_{r+1} \\ &= \mu_1 \prod_{i=1}^r p_i + \mu_2 p_{r+1} \end{aligned}$$

donde $\mu_1 = c_1 d_1$, $\mu_2 = c_1 d_2 p_r + c_2 d_1 \prod_{i=1}^{r-1} p_i + c_2 d_2 p_{r+1} \in R$, de ahí que, $\mu_1 \prod_{i=1}^r p_i + \mu_2 p_{r+1} = 1$, por lo tanto, $(\prod_{i=1}^r p_i, p_{r+1}) = 1$, lo cual queríamos demostrar. \square

Ahora demostraremos uno de los resultados más importantes de este trabajo, el famoso **Lema de Hensel**.

Lema 3.5 (Lema de Hensel). *Sea f un polinomio mónico en $\mathbb{Z}_p[x]$ y supóngase que $\bar{f} = g_1 g_2 \cdots g_r \in \mathbb{F}_p[x]$ donde g_1, g_2, \dots, g_r son polinomios mónicos y coprimos por pares sobre \mathbb{F}_p . Entonces existen polinomios mónicos y coprimos por pares f_1, f_2, \dots, f_r sobre \mathbb{Z}_p tales que:*

$$i) f = f_1 f_2 \cdots f_r \in \mathbb{Z}_p[x]$$

$$ii) \bar{f}_i = g_i \text{ para cada } i \in \{1, 2, \dots, r\}$$

Demostración. Haremos inducción sobre el número de factores, r . Para el caso $r = 2$, ver la demostración del Lema 3.3. Ahora, supóngase que el resultado es válido para $r - 1$ factores y veamos que es válido para r factores. Sea $f \in \mathbb{Z}_p[x]$ un polinomio mónico tal que $\bar{f} = g_1 g_2 \cdots g_r$, donde g_1, g_2, \dots, g_r son polinomios mónicos y coprimos por pares sobre \mathbb{F}_p . Denotemos por $h = g_1 g_2 \cdots g_{r-1}$, es claro que, $h \in \mathbb{F}_p[x]$ y $\bar{f} = h g_r$. Por el Lema 3.4, tenemos que $(h, g_r) = 1$ además, por el Lema 3.3, tenemos que existen $\hat{h}, f_r \in \mathbb{Z}_p[x]$ polinomios mónicos coprimos tales que $f = \hat{h} f_r$, donde $\bar{\hat{h}} = h$ y $\bar{f}_r = g_r$. Pero $\bar{\hat{h}} = h = \prod_{i=1}^{r-1} g_i$, por la hipótesis inductiva existen $f_1, f_2, \dots, f_{r-1} \in \mathbb{Z}_p[x]$ polinomios mónicos coprimos por pares tales que $\bar{\hat{h}} = f_1 f_2 \cdots f_{r-1}$ y $\bar{f}_i = g_i$ con $i \in \{1, 2, \dots, r-1\}$, de ahí que, $f = f_1 f_2 \cdots f_{r-1} f_r \in \mathbb{Z}_p[x]$ y $\bar{f}_i = g_i$ para cada $i \in \{1, 2, \dots, r\}$, con lo cual queda demostrado el lema de Hensel. \square

Concluimos esta sección mencionando que el lector interesado puede consultar el libro de McDonald ([12]) para revisar el Lema de Hensel en su versión general sobre anillos conmutativos finitos locales y los artículos [3] y [13] para ver un par de aplicaciones del Lema de Hensel sobre la clase de anillos conmutativos finitos de cadena.

4 Polinomios Básicos irreducibles y el levantamiento de Hensel

A continuación, se demostrará un teorema de factorización “única” para polinomios con coeficientes en el anillo \mathbb{Z}_p .

Teorema 4.1. *Sea f un polinomio mónico en $\mathbb{Z}_p[x]$ con $\text{grad}(f) \geq 1$. Entonces:*

i) f puede ser factorizado de la manera siguiente:

$$f = f_1 f_2 \cdots f_r$$

donde f_1, f_2, \dots, f_r son polinomios mónicos, primarios y coprimos por pares en $\mathbb{Z}_{p^s}[x]$, más aún, para cada $i \in \{1, 2, \dots, r\}$, \bar{f}_i es una potencia de algún polinomio mónico irreducible en $\mathbb{F}_p[x]$.

ii) Esta factorización es única salvo el orden.

Demostración. Sea $f \in \mathbb{Z}_{p^s}[x]$ entonces $\bar{f} \in \mathbb{F}_p[x]$, como $\mathbb{F}_p[x]$ es un dominio de factorización única existen $g_1, g_2, \dots, g_r \in \mathbb{F}_p[x]$ polinomios mónicos irreducibles coprimos por pares tales que

$$\bar{f} = g_1^{e_1} g_2^{e_2} \cdots g_r^{e_r}, \quad (14)$$

para algunos $e_1, e_2, \dots, e_r \in \mathbb{N}$, por el Lema 3.4 tenemos que si $i \neq j$ entonces $(g_i^{e_i}, g_j^{e_j}) = 1$ para $i, j \in \{1, 2, \dots, r\}$, además, debido al Lema de Hensel, tenemos que existen polinomios mónicos, coprimos por pares $f_1, f_2, \dots, f_r \in \mathbb{Z}_{p^s}[x]$ tales que $f = f_1 f_2 \cdots f_r$ y $\bar{f}_i = g_i^{e_i}$ para cada $i \in \{1, 2, \dots, r\}$, y por el Lema 2.16 se sigue que para cada i , f_i es un polinomio primario en $\mathbb{Z}_{p^s}[x]$, con lo cual queda demostrada la primera afirmación. Ahora bien, supóngase que:

$$f_1 f_2 \cdots f_r = h_1 h_2 \cdots h_t \quad (15)$$

son dos factorizaciones de f como producto de polinomios primarios, mónicos coprimos por pares en $\mathbb{Z}_{p^s}[x]$. Se sigue de (15) que $f_1 f_2 \cdots f_r \in \langle h_i \rangle$ para cada $i \in \{1, 2, \dots, t\}$ y como $\langle h_i \rangle$ es un ideal primario, existen $k_i \in \mathbb{Z}$ con $1 \leq k_i \leq r$ y $n_i \in \mathbb{N}$ tales que $f_{k_i}^{n_i} \in \langle h_i \rangle$. Veamos que k_i es único para cada $i \in \{1, 2, \dots, t\}$. Sean $k'_i \neq k_i$ y $n'_i \in \mathbb{N}$ tales que $f_{k'_i}^{n'_i}, f_{k_i}^{n_i} \in \langle h_i \rangle$, es claro que $(f_{k'_i}, f_{k_i}) = 1$ entonces existen $a, b \in \mathbb{Z}_{p^s}[x]$ tales que $a f_{k_i} + b f_{k'_i} = 1$, como $1^{n_i+n'_i-1} = (a f_{k_i} + b f_{k'_i})^{n_i+n'_i-1}$ tenemos que:

$$1 = \sum_{j=0}^{n_i+n'_i-1} \binom{n_i+n'_i-1}{j} a^j f_{k_i}^j b^{n_i+n'_i-1-j} f_{k'_i}^{n_i+n'_i-1-j}$$

al desarrollar la sumatoria siempre tendremos presentes a los factores $f_{k_i}^{n_i}$ y $f_{k'_i}^{n'_i}$ los cuales son elementos del ideal $\langle h_i \rangle$, entonces $1 \in \langle h_i \rangle$ lo cual es una contradicción pues h_i es primario, por lo tanto $k_i = k'_i$. De manera similar, para cada $j \in \{1, 2, \dots, r\}$ existe un único l_j con $l_j \in \{1, 2, \dots, t\}$ tal que $h_{l_j}^{m_j} \in \langle f_j \rangle$, así,

para cada $k_i \in \{1, 2, \dots, r\}$ tenemos que $h_{l_j}^{m_j} = cf_j$. Si $j = k_i$ entonces $h_{l_{k_i}}^{m_{k_i}} = cf_{k_i}$ así $h_{l_{k_i}}^{m_{k_i}n_i} = c^{n_i}f_{k_i}^{n_i} \in \langle h_i \rangle$, es decir, $h_{l_{k_i}}^{m_{k_i}n_i} \in \langle h_i \rangle$ luego existe un $\lambda \in \mathbb{Z}_{p^s}[x]$ tal que $h_{l_{k_i}}^{m_{k_i}n_i} = \lambda h_i$, aplicando el epimorfismo (10) tenemos que $\overline{h_{l_{k_i}}^{m_{k_i}n_i}} = \overline{\lambda} \overline{h_i}$, en otras palabras, $\overline{h_{l_{k_i}}^{m_{k_i}n_i}} \in \langle \overline{h_i} \rangle$. Así tenemos que $l_{k_i} = i$ para cada $i \in \{1, 2, \dots, r\}$ pues de lo contrario dado l_{k_i} existirían j_0, i_0 distintos tales que $j_0 = l_{k_i} = i_0$ pero $(h_{j_0}, h_{i_0}) = h_{j_0} \neq 1$ en contradicción con el Lema 3.2. Usando lo anterior, definimos las siguiente funciones:

$$\begin{aligned} \{1, 2, \dots, t\} &\longrightarrow \{1, 2, \dots, r\} \\ i &\longmapsto k_i \\ \{1, 2, \dots, r\} &\longrightarrow \{1, 2, \dots, t\} \\ j &\longmapsto l_j \end{aligned}$$

las cuales son inyectivas, de ahí que $r \leq t$ y también $t \leq r$ así, $r = t$ y reenumerando $k_i = i$ para cada $i \in \{1, 2, \dots, r\}$ tenemos que $l_j = k_i = i$ entonces $f_i^{n_i} \in \langle h_i \rangle$ y $h_i^{m_i} \in \langle f_i \rangle$. Si $j \neq 1$ tenemos que $(f_1, f_j) = 1$ entonces $(\overline{f_1}, \overline{f_j}) = 1$, luego $\overline{f_2} \overline{f_3} \cdots \overline{f_r}$ y $\overline{f_1}^{n_1}$ son coprimos y por el Lema 3.4 $(f_2 f_3 \cdots f_r, f_1^{n_1}) = 1$. Como $f_1^{n_1} \in \langle h_1 \rangle$ existe algun $c \in \mathbb{Z}_{p^s}[x]$ tal que $f_1^{n_1} = ch_1$. Veamos que el producto $f_2 f_3 \cdots f_r$ y h_1 son coprimos. Dado que $(f_2, f_3 \cdots f_r, f_1^{n_1}) = 1$ existen $\alpha, \beta \in \mathbb{Z}_{p^s}[x]$ tales que $1 = \alpha f_2 f_3 \cdots f_r + \beta f_1^{n_1} = \alpha f_2 f_3 \cdots f_r + \beta (ch_1)$, sea $\gamma = c\beta$ entonces

$$\alpha f_2 f_3 \cdots f_r + \gamma h_1 = 1, \quad (16)$$

de ahí que, $(f_2 f_3 \cdots f_r, h_1) = 1$. Multiplicando (16) por f_1 se obtiene que $\alpha f_1 f_2 \cdots f_r + \gamma h_1 f_1 = f_1$, es decir, $\alpha h_1 h_2 \cdots h_r + \gamma h_1 f_1 = f_1$, usando (15), entonces $f_1 = h_1 (\alpha h_2 h_3 \cdots h_r + \gamma f_1)$, por lo anterior, se tiene que h_1 divide a f_1 . Procediendo de manera similar con $h_1^{m_1}$, se establece que f_1 divide a h_1 y como $(f_1, h_1) = 1$ se sigue que $f_1 = h_1$, análogamente se concluye que $f_i = h_i$ para toda $i \in \{1, 2, \dots, r\}$. \square

Definición 4.2. Sea $f(x)$ un polinomio mónico de grado $m \geq 1$ en $\mathbb{Z}_{p^s}[x]$. Si $\overline{f}(x) \in \mathbb{F}_p[x]$ es irreducible (o primitivo), diremos que $f(x)$ es un **polinomio mónico básico irreducible** (o mónico básico primitivo) en $\mathbb{Z}_{p^s}[x]$.

Los siguientes lemas, serán de suma importancia en la demostración de los resultados más relevantes de esta sección.

Lema 4.3. Si \mathbb{F}_q es un campo finito con q elementos, entonces todo polinomio irreducible $h(x) \in \mathbb{F}_q[x]$ de grado m divide a $x^{q^m} - x$.

Demostración. Como $h(x)$ es irreducible con $\text{grad}(h) = m$, entonces el conjunto:

$$\mathbb{F} = \frac{\mathbb{F}_q[x]}{\langle h(x) \rangle} := \{[f(x)] = f(x) + \langle h(x) \rangle \mid f(x) \in \mathbb{F}_p[x]\}$$

es un campo finito con q^m elementos y por [8, Lemma 2.3] se tiene que, $[x] \in \mathbb{F}$ implica que $[x]^{q^m} = [x]$, entonces $[0] = [x]^{q^m} - [x] = [x^{q^m} - x]$, es decir, $x^{q^m} - x \in \langle h(x) \rangle$, por lo tanto, $h(x) \mid x^{q^m} - x$. \square

Lema 4.4. *Sea f un polinomio no nulo de grado positivo sobre un campo finito \mathbb{F} . Si $(f, f') = 1$ entonces f no tiene factores múltiples.*

Demostración. Suponga que f tiene al menos un factor múltiple, digamos g con $\text{grad}(g) \geq 1$, es decir, $f = g^2h$ para algún $h \in \mathbb{F}[x]$. Entonces $f' = 2gg'h + h'g^2$ factorizando a g tenemos que $f' = g(2g'h + h'g)$, de manera que, g divide a f y f' . Sea $d = (f, f')$ entonces $g \mid d$, de ahí que, $d \neq 1$. \square

Lema 4.5. *Sean \mathbb{F} un campo con característica p , i.e., $\text{Car}(\mathbb{F}) = p$ y $n \in \mathbb{N}$. Si $p \nmid n$ entonces el polinomio $x^n - 1 \in \mathbb{F}_p[x]$ no tiene raíces múltiples.*

Demostración. Sea r una raíz múltiple de $h(x) = x^n - 1$ en $\mathbb{F}_p[x]$. Es claro que $r \neq 0$ y, por [8, Theorem 1.68], también es raíz de $h'(x) = nx^{n-1}$. Entonces $nr^{n-1} = 0$, esto es, $(nr)r^{n-2} = 0$ pero \mathbb{F}_p es un dominio entero y $r \neq 0$, de ahí que, $nr = 0$, por lo tanto, $p \mid n$. \square

Teorema 4.6. *Para cualquier entero $m \geq 1$ existe un polinomio mónico básico irreducible de grado m sobre \mathbb{Z}_{p^s} el cual divide a $x^{p^m-1} - 1$ en $\mathbb{Z}_{p^s}[x]$.*

Demostración. Sea \mathbb{F}_p un campo finito con p elementos. En [8, Corollary 2.11] se muestra que para $m \geq 1$ existe $f_0(x)$ un polinomio irreducible de grado m sobre \mathbb{F}_p . Si $f_0(x) = a_mx^m + a_{m-1}x^{m-1} + \dots + a_1x + a_0$ entonces $f_1(x) = a_m^{-1}f_0(x)$ es un polinomio mónico irreducible de grado m en $\mathbb{F}_p[x]$. Por el Lema 4.3, $f_1(x)$ divide a $x^{p^m} - x$ en $\mathbb{F}_p[x]$. Como $f_1(x)$ es irreducible y $x^{p^m} - x = x(x^{p^m-1} - 1)$ entonces $f_1(x) \mid x$ ó $f_1(x) \mid x^{p^m-1} - 1$. Supongamos que $m = 1$ y sea $f(x) = x - 1 \in \mathbb{Z}_{p^s}[x]$ tal que $\bar{f}(x) = f_1(x)$ entonces $f_1(x) \mid (x^{p^m-1} - 1)$ en $\mathbb{F}_p[x]$ ya que $\bar{f}(x) = f_1(x) = x - 1$, así que $\bar{f}(x) = x - 1$ en $\mathbb{F}_p[x]$ es un polinomio mónico e irreducible tal que $f(x) \mid (x^{p^m-1} - 1)$ en $\mathbb{Z}_{p^s}[x]$, como deseabamos. Por otro lado, si $m > 1$ entonces $f_1(x)$ no puede dividir a x entonces $f_1(x)$ divide a $x^{p^m-1} - 1$ en $\mathbb{F}_p[x]$. Así, existe $g_1(x) \in \mathbb{F}_p[x]$ tal que $x^{p^m-1} - 1 = f_1(x)g_1(x)$, observe que la derivada de $x^{p^m-1} - 1$ es el polinomio $p^m x^{p^m-2} - x^{p^m-2}$ y como $\text{Car}(\mathbb{F}) = p$ entonces $p^m x^{p^m-2} - x^{p^m-2} = -x^{p^m-2}$. Ahora bien, si consideramos los polinomios $\lambda_1 = -1$ y $\lambda_2 = -x$ se sigue que $\lambda_1(x^{p^m-1} - 1) + \lambda_2(-x^{p^m-2}) = -x^{p^m-1} + 1 + x^{p^m-1} = 1$,

así, por la Definición 3.1 y el Lema 4.4, $x^{p^m-1} - 1$ no tiene factores múltiples. Supongamos que $(f_1(x), g_1(x)) = d(x)$ y $\text{grad}(d) \geq 1$ entonces $f_1(x) = k_1(x)d(x)$ y $g_1(x) = k_2(x)d(x)$ para algunos $k_1(x), k_2(x) \in \mathbb{F}_p[x]$, entonces como $x^{p^m-1} - 1 = f_1(x)g_1(x) = k_1(x)k_2(x)(d(x))^2$, tenemos que, $x^{p^m-1} - 1$ tiene a $d(x)$ como factor múltiple, lo cual es una contradicción, por consiguiente, $(f_1(x), g_1(x)) = d(x)$ con $\text{grad}(d) = 1$ pero $f_1(x)$ es mónico, de ahí que, $(f_1(x), g_1(x)) = 1$, entonces por el Lema de Hensel existen polinomios mónicos y coprimos $f_2(x)$ y $g_2(x)$ en $\mathbb{Z}_{p^s}[x]$ tales que $x^{p^m-1} - 1 = f_2(x)g_2(x)$ en $\mathbb{Z}_{p^s}[x]$ con $\overline{f_2(x)} = f_1(x)$ y $\overline{g_2(x)} = g_1(x)$. Si escogemos $f(x) = f_2(x)$ se satisface que $\overline{f(x)} = f_1(x)$ un polinomio mónico e irreducible en $\mathbb{F}_p[x]$ y por la Definición 4.2 $f(x)$ es un polinomio mónico básico irreducible con $\text{grad}(f) = \text{grad}(f_1) = m$ y que divide a $x^{p^m-1} - 1$ en \mathbb{Z}_{p^s} como queríamos demostrar. \square

Definición 4.7. Sea $g(x)$ un polinomio mónico sobre \mathbb{F}_p . Un polinomio mónico $f(x)$ en $\mathbb{Z}_{p^s}[x]$ con $\overline{f(x)} = g(x)$ es llamado un **Levantamiento de Hensel** para $g(x)$ si y sólo si existe $n \in \mathbb{N}$ tal que si $p \nmid n$ entonces $f(x) \mid (x^n - 1)$ en $\mathbb{Z}_{p^s}[x]$.

En el Teorema 4.6, tenemos que $\overline{f(x)} = f_1(x)$ y si $p \mid p^m - 1$ entonces existirá un $k \in \mathbb{N}$ tal que $p^m - 1 = kp$ entonces $p^m - kp = 1$, en otras palabras, $(p, p^m) = 1$ lo cual es absurdo, así existe $n = p^m - 1 \in \mathbb{N}$ tal que $p \nmid n$ y $f(x) \mid (x^n - 1)$ en $\mathbb{Z}_{p^s}[x]$. por lo tanto, $f(x)$ es el levantamiento de Hensel de $f_1(x)$.

Sin embargo, no todo polinomio mónico básico irreducible es un levantamiento de Hensel. Por ejemplo, para el polinomio $x + 2 \in \mathbb{Z}_4[x]$ se tiene $\overline{x + 2} = x \in \mathbb{F}_2[x]$. Veamos que $x + 2$ no es un levantamiento de Hensel para x . Sea $n \in \mathbb{N}$, si $2 \nmid n$ entonces existe $k \in \mathbb{N}$ tal que $n = 2k + 1$, como $x^{2k+1} - 1 = (x^{2k} - 2x^{2k-1})(x + 2) - 1$ tenemos que el residuo de la división de $x^{2k+1} - 1$ por $x + 2$ es -1 , esto es, $x^{2k+1} - 1 \equiv -1 \pmod{x + 2}$, por lo tanto, $x + 2 \nmid x^n - 1$, lo cual contradice la Definición 4.7.

Teorema 4.8. Sea $s \in \mathbb{N}$. Un polinomio mónico $g(x) \in \mathbb{F}_p[x]$ tiene un levantamiento de Hensel $f(x) \in \mathbb{Z}_{p^s}$ si y sólo si $g(x)$ no tiene raíces múltiples y $x \nmid g(x)$ en $\mathbb{F}_p[x]$.

Demostración. Supóngase que $f(x)$ es un levantamiento de Hensel para $g(x)$ sobre \mathbb{Z}_{p^s} , entonces $\overline{f(x)} = g(x)$ y existe $n \in \mathbb{N}$ tal que $p \nmid n$ y $f(x) \mid x^n - 1$, así, tenemos que, $x^n - 1 = f(x)h(x)$ para algún $h(x) \in \mathbb{F}_p[x]$ entonces

$$x^n - 1 = \overline{x^n - 1} = \overline{f(x)h(x)} = \overline{f(x)}\overline{h(x)} = g(x)\overline{h(x)}. \quad (17)$$

Como $p \nmid n$, $x^n - 1$ no tiene raíces múltiples en $\mathbb{F}_p[x]$, por el Lema 4.5, y de (17) se sigue que $g(x)$ no tiene raíces múltiples pues de tenerlas entonces serían también raíces de $x^n - 1$, lo cual no puede ocurrir. Más aún, si $x \mid g(x)$ tenemos que

$x|x^n - 1$, por (17), es decir, 0 es raíz de $x^n - 1$ lo cual tampoco puede ocurrir. Por lo tanto, $x \nmid g(x)$. Recíprocamente, supongamos que $g(x)$ no tiene raíces múltiples y que $x \nmid g(x)$ en $\mathbb{F}_p[x]$, entonces $g(0) \neq 0$. Por [8, Lemma 3.1], existe $n \in \mathbb{N}$ tal que $g(x)|x^n - 1$ con $n \leq p^{\text{grad}(g)-1}$. Tenemos que $(n, p) = 1$ o bien $p|n$, entonces existen $m \in \mathbb{N}$ y $e \in \mathbb{N} \cup \{0\}$ con $(m, p) = 1$ tales que $n = mp^e$. Como $p \mid \binom{p^e}{i}$ para $i \in \{1, 2, \dots, p^e - 1\}$ y $\text{Car}(\mathbb{F}_p) = p$ entonces $\binom{p^e}{i} = 0$ para $i \in \{1, 2, \dots, p^e - 1\}$ en \mathbb{F}_p , aplicando el teorema del binomio para $a, b \in \mathbb{F}_p$ se tiene que:

$$\begin{aligned} (a+b)^{p^e} &= \sum_{i=0}^{p^e} \binom{p^e}{i} a^{p^e-i} b^i \\ &= a^{p^e} + \sum_{i=1}^{p^e-1} \binom{p^e}{i} a^{p^e-i} b^i + b^{p^e} \\ &= a^{p^e} + b^{p^e}. \end{aligned}$$

Por lo anterior, $(x^m)^{p^e} = ((x^m - 1) + 1)^{p^e} = (x^m - 1)^{p^e} + 1$, i.e., $(x^m)^{p^e} = (x^m - 1)^{p^e} + 1$, de ahí que, $(x^m - 1)^{p^e} = (x^m)^{p^e} - 1 = x^{mp^e} - 1 = x^n - 1$. Como $g(x)|(x^n - 1)$ entonces $g(x)h(x) = (x^m - 1)^{p^e}$ para algún $h(x) \in \mathbb{F}_p[x]$ y $\text{grad}(g) + \text{grad}(h) = n = mp^e \geq 1$. Si $\text{grad}(g) = 1$ entonces $g(x) = x - 1$, ya que $g(x)$ es mónico y 1 es raíz común de $g(x)$ y $x^n - 1$, además $x - 1|x^m - 1$, por consiguiente, $g(x)|x^m - 1$. Por otro lado, como $g(x)h(x) = (x^m - 1)(x^m - 1)^{p^e-1}$ y $g(x)$ no tiene raíces múltiples, si $\text{grad}(g) = l > 1$ entonces $g(x)|x^m - 1$, ya que de lo contrario, tendría a 1 como raíz múltiple. En ambos casos tenemos que existe $m \in \mathbb{N}$ tal que $p \nmid n$ y $g(x)|x^m - 1$ en $\mathbb{F}_p[x]$, es decir, existe $g_0(x) \in \mathbb{F}_p[x]$ de tal manera que $x^m - 1 = g(x)g_0(x)$ y como $p \nmid m$, $x^m - 1$ no tiene factores múltiples, así $(g(x), g_0(x)) = 1$ en $\mathbb{F}_p[x]$, por el Lema de Hensel, existen $f(x), f_0(x) \in \mathbb{Z}_{p^s}[x]$ polinomios mónicos, coprimos entre sí tales que

$$x^m - 1 = f(x)f_0(x) \quad (18)$$

en $\mathbb{Z}_{p^s}[x]$ con $\bar{f}(x) = g(x)$ y $\bar{f}_0(x) = g_0(x)$. Por (18) se deduce que $f(x)|x^m - 1$ en $\mathbb{Z}_{p^s}[x]$ con $p \nmid m$, por lo tanto, f es el levantamiento de Hensel de $g(x)$. \square

Este teorema y el siguiente lema son fundamentales en la demostración de la unicidad del levantamiento de Hensel.

Lema 4.9. Sean $m, n \in \mathbb{N}$. $x^m - 1|x^n - 1$ si y sólo si $m|n$.

Demostración. Por el algoritmo de la división, existen $s, r \in \mathbb{Z}$ tales que $n = sm + r$ con $0 \leq r < m$. Es fácil verificar, mediante la división larga de $x^n - 1$ por $x^m - 1$, que:

$$x^n - 1 = (x^{(s-1)m+r} + x^{(s-2)m+r} + \dots + x^r)(x^m - 1) + x^r - 1,$$

dado que $n = sm + r$ con $\text{grad}(x^r - 1) < \text{grad}(x^m - 1)$, por consiguiente, $x^n - 1 \equiv x^r - 1 \pmod{x^m - 1}$. Si $m|n$ entonces $r = 0$, luego, $x^r - 1 = 0$ y de lo anterior se sigue que $x^m - 1|x^n - 1$. Por otro lado, si $x^m - 1|x^n - 1$ tenemos que $x^r - 1 = 0$, es decir, $x^r = 1$ entonces $r = 0$, por lo tanto. $m|n$, como queríamos demostrar. \square

Teorema 4.10. *Sea $s \in \mathbb{N}$ y $g(x)$ un polinomio mónico en $\mathbb{F}_p[x]$ sin raíces múltiples tal que $x \nmid g(x)$ en $\mathbb{F}_p[x]$. Entonces $g(x)$ tiene un único levantamiento de Hensel en $\mathbb{Z}_{p^s}[x]$.*

Demostración. Por el Teorema 4.8, $g(x)$ tiene un levantamiento de Hensel en $\mathbb{Z}_{p^s}[x]$, así que sólo probaremos la unicidad. Sean $f^{(1)}(x)$ y $f^{(2)}(x)$ dos levantamientos de Hensel de $g(x)$ en $\mathbb{Z}_{p^s}[x]$ entonces

- i) $f^{(1)}$ y $f^{(2)}$ son polinomios mónicos.
- ii) $\overline{f^{(1)}}(x) = g(x) = \overline{f^{(2)}}(x)$.
- iii) Existen $n_1, n_2 \in \mathbb{N}$ tales que $p \nmid n_1, p \nmid n_2$ con $f^{(1)}(x)|x^{n_1} - 1$ y $f^{(2)}(x)|x^{n_2} - 1$ en $\mathbb{Z}_{p^s}[x]$.

De ii) y iii) se sigue que $g(x)|x^{n_1} - 1$ y $g(x)|x^{n_2} - 1$ en $\mathbb{F}_p[x]$. Así, tenemos dos casos: $n_1 = n_2$ y $n_1 \neq n_2$. Supóngase que $n = n_1 = n_2$, como $\mathbb{F}_p[x]$ es un dominio de factorización única tenemos que existen $h_1(x), h_2(x), \dots, h_r(x)$ polinomios mónicos e irreducibles en $\mathbb{F}_p[x]$ y $e_1, e_2, \dots, e_r \in \mathbb{N}$ tales que $x^n - 1 = h_1^{e_1}(x)h_2^{e_2}(x) \cdots h_r^{e_r}(x)$. Sea $g_i(x) = h_i^{e_i}(x)$ para cada $i \in \{1, 2, \dots, r\}$ entonces

$$x^n - 1 = g_1(x)g_2(x) \cdots g_r(x) \quad (19)$$

en $\mathbb{F}_p[x]$ con $(g_i, g_j) = 1$ para cada $i \neq j$, luego, por el Lema de Hensel, existen $f_1(x), f_2(x), \dots, f_r(x) \in \mathbb{Z}_{p^s}[x]$ tales que $x^n - 1 = f_1(x)f_2(x) \cdots f_r(x)$ en $\mathbb{Z}_{p^s}[x]$ con $f_i(x)$ mónico, $(f_i(x), f_j(x)) = 1$ si $i \neq j$, y $\overline{f_i}(x) = g_i(x) = h_i(x)^{e_i} \in \mathbb{F}_p[x]$ para cada $i \in \{1, 2, \dots, r\}$. Por el Lema 2.16, f_i es un polinomio primario sobre $\mathbb{Z}_{p^s}[x]$ para cada i . Como $g(x)|x^n - 1$, $(g_i(x), g_j(x)) = 1$ si $i \neq j$ y por (19) tenemos que $g(x) = g_1(x)g_2(x) \cdots g_t(x)$ para algún $1 \leq t \leq r$, salvo el orden de $g_1(x), g_2(x), \dots, g_r(x)$ en (19). Finalmente, nombrando a $\overline{f(x)} = \overline{f^{(1)}(x)} = \overline{f^{(2)}(x)} = g(x)$ y $f(x)|x^n - 1$, por el Teorema 4.1, en otras palabras, $f^{(1)}(x) = f^{(2)}(x)$. Por otro lado, si $n_1 \neq n_2$, sea $n = [n_1, n_2]^2$, entonces $n_1|n$ y $n_2|n$, de ahí que, existen $k_1, k_2 \in \mathbb{N}$ tales que $n = k_1 n_1$ y $n = k_2 n_2$. Nuevamente ocurre que $(p, k_1) = 1$ o $k_1 = kp^e$ y $(p, k_2) = 1$ o $k_2 = k'p^{e'}$ con $(p, k) = (p, k') = 1$ y $e, e' \in \mathbb{N}$. Sea $t = kk'n_1n_2$ entonces $p \nmid t$ pero $n_1|t$ y $n_2|t$

²El mínimo común múltiplo de n_1 y n_2

entonces $x^{n_1} - 1 | x^t - 1$ y $x^{n_2} - 1 | x^t - 1$ en $\mathbb{Z}_{p^s}[x]$, por el Lema 4.9, y por el primer caso, debemos concluir que $f^{(1)}(x) = f^{(2)}(x)$. Esto prueba la unicidad del levantamiento de Hensel. \square

Algoritmo 1 Hensel's Step [15, Algorithm 15.10]

Entrada: p un número primo, polinomios $f \in \mathbb{Z}_{p^2}[x]$ y $g, h, s, t \in \mathbb{F}_p[x]$ tales que: $\bar{f} = gh$, g, h mónicos, $\text{grad}(f) = \text{grad}(g) + \text{grad}(h)$, $\text{grad}(t) < \text{grad}(g)$, $\text{grad}(s) < \text{grad}(h)$ y $sg + th = 1$ en $\mathbb{F}_p[x]$.

Salida: Polinomios $g^*, h^*, s^*, t^* \in \mathbb{Z}_{p^2}[x]$ tales que: $f = g^*h^*$ con g^*, h^* mónicos, $\bar{g}^* = g$, $\bar{h}^* = h$, $\bar{s}^* = s$ y $\bar{t}^* = t$.

- 1: Calcule el polinomio $e = f - gh$ en \mathbb{Z}_{p^2} .
 - 2: Hallar polinomios $q, r \in \mathbb{Z}_{p^2}[x]$ tales que $se = qh + r$.
 - 3: Defina los polinomios $g_0 = g(q + 1) + te$, $h_0 = h + r$ y $u = sg_0 + th_0 - 1$.
 - 4: Hallar polinomios $v, w \in \mathbb{Z}_{p^2}[x]$ tales que: $su = vh_0 + w$.
 - 5: Obtener $g^* = g_0$, $h^* = h_0$, $s^* = s - w$ y $t^* = t(1 - u) - vg_0$.
-

En el siguiente ejemplo aplicamos el Algoritmo 1 para factorizar el polinomio dado.

Ejemplo 4.11. Factorizar el polinomio $f = x^3 + 4x + 8 \in \mathbb{Z}_{32}[x]$.

Es claro que $\bar{f} = x^3 + x + 2 \in \mathbb{F}_3[x]$ y además 2 es una raíz de \bar{f} , entonces $x - 2$ divide a \bar{f} en $\mathbb{F}_3[x]$, realizando la división se obtiene que $\bar{f} = (x - 2)(x^2 + 2x + 2)$, de ahí que, $g = x - 2$ y $h = x^2 + 2x + 2$. Aplicando el algoritmo de Euclides se tiene que

$$(2x + 2)(x - 2) + (1)(x^2 + 2x + 2) = 1$$

así $s = 2x + 2$ y $t = 1$. Calculamos en $\mathbb{Z}_9[x]$ el polinomio $e = f - gh = x^3 + 4x + 8 - (x^3 - 2x - 4) = 6x + 3$. Entonces $se = 3x^2 + 6x$ y, al dividir por h se obtiene que

$$se = 3(x^2 + x + 2) - 6x,$$

de ahí, se sigue que $q = 3$ y $r = -6x \in \mathbb{Z}_9$. Luego se calculan los polinomios $g_0 = x - 5$ y $h_0 = x^2 - 4x + 2$ y definimos $u = sg_0 + th_0 - 1 = 3x^2 - 3x$. Nuevamente al dividir su por h_0 tenemos que

$$su = (6x + 6)(x^2 - 4x + 2) + (6x - 3)$$

Finalmente, se concluye que $v = 6x + 6$ y $w = 6x - 3$. Realizando los últimos

calculos se obtiene que

$$\begin{array}{ll} g^* = x - 5 & \overline{g^*} = x - 2 \\ h^* = x^2 - 4x + 2 & \overline{h^*} = x^2 - x + 2 = x^2 + 2x + 2 \\ s^* = -4x + 5 & \overline{s^*} = -x + 2 = 2x + 2 \\ t^* = 4 & \overline{t^*} = 1 \end{array}$$

Además,

$$\begin{aligned} s^*g^* + t^*h^* &= (-4x + 5)(x - 5) + (4)(x^2 - 4x + 2) \\ &= -4x^2 + 7x - 7 + 4x^2 - 7x + 8 \\ &= 1 \end{aligned}$$

como esperabamos.

Algoritmo 2 Método de Graeffe [16, Theorem 13.12]

Entrada: Un polinomio $f_2(x) \in \mathbb{F}_2[x]$ de grado n sin raíces múltiples y tal que $f_2(0) \neq 0$.

Salida: Un polinomio $f(x)$ el levantamiento de Hensel de $f_2(x)$.

- 1: Escriba $f_2(x) = e(x) - d(x)$ donde $e(x)$ sólo contiene términos de f_2 con exponente par y $d(x)$ con los de exponente impar.
- 2: Calcule en $\mathbb{Z}_4[x]$ el polinomio

$$f(x^2) = \begin{cases} + [(e(x))^2 - (d(x))^2] & \text{si } \text{grad}(e) > \text{grad}(d) \\ - [(e(x))^2 - (d(x))^2] & \text{si } \text{grad}(e) < \text{grad}(d) \end{cases}$$

- 3: Cambie x^2 por x en $f(x^2)$.
-

En los siguientes dos ejemplos se emplea el Algoritmo 2 para realizar el levantamiento de Hensel del polinomio dado.

Ejemplo 4.12. Calcular el levantamiento de Hensel para el polinomio $f_2(x) = x^3 + x + 1 \in \mathbb{F}_2[x]$. Obsérvese que $e(x) = 1$ y $d(x) = x^3 + x$ entonces tenemos que $\text{grad}(d) > \text{grad}(e)$, así

$$\begin{aligned} f(x^2) &= - [(1)^2 - (x^3 + x)^2] \\ &= - [-x^6 - 2x^4 - x^2 + 1] \\ &= x^6 + 2x^4 + x^2 - 1. \end{aligned}$$

Luego, haciendo el cambio de x^2 por x en $f(x^2)$ tenemos que $f(x) = x^3 + 2x^2 + x - 1$. Veamos que es el levantamiento de Hensel de f_2 .

$$\overline{f}(x) = \overline{x^3 + 2x^2 + x - 1} = x^3 + x - 1 = x^3 + x + 1 = f_2(x).$$

Además, se tiene tras hacer la división larga que

$$x^3 + 2x^2 + x - 1 \equiv 0 \pmod{x^7 - 1}$$

como queríamos probar.

Ejemplo 4.13. Calcular el levantamiento de Hensel para $g_2(x) = x^4 + x^3 + x^2 + x + 1$. Tenemos pues que $e(x) = x^4 + x^2 + 1$ y $d(x) = x^3 + x$ así $\text{grad}(e) > \text{grad}(d)$ entonces

$$\begin{aligned} g(x^2) &= \left[(x^4 + x^2 + 1)^2 - (x^3 + x)^2 \right] \\ &= \left[x^8 + 2x^6 + x^4 + 2x^4 + 2x^2 + 1 - x^6 - 2x^4 - x^2 \right] \\ &= x^8 + x^6 + x^4 + x^2 + 1 \end{aligned}$$

de ahí que $g(x) = x^4 + x^3 + x^2 + x + 1$. Tenemos que

$$\overline{g}(x) = \overline{x^4 + x^3 + x^2 + x + 1} = x^4 + x^3 + x^2 + x + 1 = g_2(x)$$

Además, se tiene tras hacer la división larga que

$$x^4 + x^3 + x^2 + x + 1 \equiv 0 \pmod{x^5 - 1}$$

por lo tanto, g es el levantamiento de Hensel de g_2 .

Concluimos esta sección invitando al lector interesado a consultar el artículo [13, Definición 3.4]) para ver una aplicación del Levantamiento de Hensel sobre la clase de anillos conmutativos finitos de cadena.

Agradecimientos

Agradecemos al árbitro su amable revisión y valiosos comentarios al trabajo.

Bibliografía

- [1] A. Bonnecaze, P. Solé, and A. R. Calderbank, *Quaternary quadratic residue codes and unimodular lattices*, IEEE Trans. Inform. Theory **41** (1995), 366–377.

- [2] A. Bonnecaze and P. Udaya, *Cyclic codes and self-dual codes over $\mathbb{F}_2 + u\mathbb{F}_2$* , IEEE Trans. Inform. Theory **45** (1999), 1250–1255.
- [3] H. Q. Dinh and S. R. López-Permouth, *Cyclic and negacyclic codes over finite chain rings*, IEEE Trans. Inform. Theory **50** (2004), no. 8, 1728–1744.
- [4] M. Greferath and S. E. Schmidt, *Gray isometries for finite chain rings and a nonlinear ternary $(36, 3^{12}, 15)$ code*, IEEE Trans. Inform. Theory **45** (1999), 2522–2524.
- [5] A. R. Hammons Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, *The \mathbb{Z}_4 -linearity of kerdock, preparata, goethals, and related codes*, IEEE Trans. Inform. Theory **40** (1994), 301–319.
- [6] K. Hensel, *Neue Grundlagen der Arithmetik*, Journal für die reine und angewandte Mathematik, 127 (1904), 51–84. <http://eudml.org/doc/149178>
- [7] P. Kanwar and S. R. López-Permouth, *Cyclic codes over the integers modulo p^m* , Finite Fields and Their Applications **3** (1997), no. 4, 334–352.
- [8] R. Lidl and H. Niederreiter, *Introduction to finite fields and their applications*, (Great Britain: Cambridge University Press, 1994).
- [9] S. Ling and J. T. Blackford, *$\mathbb{Z}_{p^{k+1}}$ -linear codes*, IEEE Trans. Inform. Theory **48** (2002), no. 9, 2592–2605.
- [10] C. A. López-Andrade and H. Tapia-Recillas, *On the linearity and quasi-cyclicity of the gray image of codes over a galois ring*, Groups, Algebras and Applications, vol. CONM/537, AMS, (2011), pp. 255–268.
- [11] C. A. López-Andrade and H. Tapia-Recillas, *On the cyclicity of the gray image of a class of linear cyclic codes over a finite chain ring*, International Journal of Pure and Applied Mathematics, **80** (2012), no. 2, 181–190.
- [12] B. R. McDonald, *Finite Rings with Identity* in Pure and Applied Mathematics, Marcel Dekker, New York, 1974.
- [13] E. Martínez-Moro and I. F. Rúa, *Multivariable Codes over Finite Chain Rings: Serial Codes*, SIAM J. Discrete Math., **20**(4), (2006) 947–959.
- [14] V. Pless and Z. Qian, *Cyclic codes and quadratic residue codes over \mathbb{Z}_4* , IEEE Trans. Inform. Theory **42** (1996), 1594–1600.

- [15] J. von zur Gathen and G. Jurgens, *Modern computer algebra*, (New York: Cambridge University Press, 2003).
- [16] Z. X. Wan, *Lectures on finite fields and Galois rings*, (Beijing: World Scientific Pub. Co. Inc., 2003).
- [17] J. Wolfmann, *Binary images of cyclic codes over \mathbb{Z}_4* , IEEE, Trans. Inform. Theory **47** (2001), 1773–1779.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570

argr_040890@hotmail.com

clopez@fcfm.buap.mx

Capítulo 3

Una variación del contenido fuerte entre formaciones localmente definidas

Ismael Gutiérrez García, Anselmo Torresblanca Badillo

Departamento de Matemáticas, Universidad del Norte

Resumen

En este capítulo presentamos inicialmente una recopilación de resultados sobre clases de grupos finitos solubles, en particular las formaciones, e introducimos la relación de contenido fuerte entre formaciones. Esta define un orden parcial sobre la clase de todas las formaciones de grupos finitos solubles. Finalmente se introduce una variación de esta relación y establecemos condiciones sobre la formación localmente definida \mathfrak{F} , que contiene G -fuertemente a la formación $\mathfrak{F} = \mathfrak{N}\mathfrak{X}$, donde \mathfrak{N} denota la clase de todos los grupos nilpotentes finitos y \mathfrak{X} es una formación local cualquiera.

1 Introducción

El estudio de los grupos finitos solubles ha ocupado un lugar preponderante en el avance de esta parte del álgebra. Uno de los temas más investigado es el de las clases de grupos y esto ha permitido la generalización de resultados importantes como los teoremas de Sylow y la factorización de grupos, entre otros, [1], [4], [6], [11], [13].

En este capítulo nos ocuparemos de presentar algunos resultados conocidos sobre clases de grupos finitos solubles y un resultado reciente sobre la relación de contenido fuerte entre formaciones localmente definidas o saturadas, ver Teorema 6.8. Cabe resaltar que el Teorema 6.8 es un resultado novedoso y se presenta por primera vez a través de este capítulo. La notación usada en este capítulo es la estándar, para ello puede consultarse [1] o [4].

En el universo soluble, los conceptos de formación localmente definida y formación saturada son equivalentes. Este hecho fué demostrado por U. Lubeseder, bajo la orientación de W. Gaschütz, [7]. Posteriormente P. Schmid demostró que la condición de solubilidad en los grupos no era necesaria [12].

Este capítulo está organizado de la siguiente manera: En la sección 2 introduciremos los conceptos de grupos solubles, supersolubles y nilpotentes. En la sección 3 introduciremos los conceptos de clases de grupos, el producto de clases de grupos

y operadores clausura. Mostraremos que el producto de clases de grupo no siempre es asociativo y conmutativo y probaremos además las condiciones necesarias para que el producto de clases de grupos sea asociativo. También en esta sección enumeramos los distintos operadores clausura que utilizaremos con más frecuencia a lo largo del capítulo. En la sección 4 inicialmente daremos los conceptos de formación, formación saturada, subgrupo \mathfrak{X} -máximo, \mathfrak{X} -proyector, y \mathfrak{X} -residuo de un grupo. Al final de esta sección se definen los conceptos de formación producto de clases de grupos, función formación, formación local y definición canónica local. En la sección 5 introduciremos el concepto de contenido fuerte entre formaciones saturadas y mostraremos que este concepto define un orden parcial sobre la clase de las formaciones saturadas. Posteriormente mostraremos algunos resultados que serán cruciales para la sección siguiente. En la sección 6 introduciremos una variación del concepto de contenido fuerte la cual llamaremos G -contenido fuerte entre formaciones locales y la denotaremos \ll_G .

El contenido fuerte entre formaciones localmente definidas fue introducido por E. Cline [2]. Este trabajo fue extendido por P. D'Arcy [3] y posteriormente se introdujo una variación del concepto por I. Gutiérrez en [8] y [9]. Uno de los resultados centrales de este capítulo y que trataremos en la sección 6, consiste en determinar la formación local \mathfrak{H} que contiene G -fuertemente a la formación $\mathfrak{F} = \mathfrak{N}\mathfrak{X}$, donde \mathfrak{N} es la clase de todos los grupos nilpotentes finitos y \mathfrak{X} es cualquier formación local.

2 Grupos solubles, supersolubles y nilpotentes

Definición 2.1. Sean G un grupo, $M \subseteq G$ y $U \leq G$.

- (a) Una cadena o una serie \mathcal{C} de G es cualquier sucesión finita de subgrupos de G , totalmente ordenada por inclusión, cuyo primer elemento es $\langle 1 \rangle$ y cuyo elemento final es G . Esto es

$$(c) \quad \langle 1 \rangle = U_0 \subseteq U_1 \subseteq \dots \subseteq U_n = G.$$

$\langle 1 \rangle$ denota un grupo trivial multiplicativo con un sólo elemento.

- (b) La cadena \mathcal{C} se denomina subnormal si para todo $i \in \{0, \dots, n-1\}$ se verifica que $U_i \trianglelefteq U_{i+1}$. Si además para todo $i \in \{0, \dots, n-1\}$ se verifica que $U_i \trianglelefteq G$, entonces decimos que \mathcal{C} es una cadena **normal** de G y los grupos cocientes U_i/U_{i+1} se denominan **factores** de la serie.
- (c) Si en la serie normal anterior, cada grupo cociente U_i/U_{i+1} está contenido en $Z(G/U_{i+1})$, la serie se llama **central**. Cada factor de la serie se llama central.

- (d) Una cadena normal de G , digamos (\mathcal{C}) , se denomina una **serie principal** de G , si para todo $i \in \{1, \dots, n-1\}$ se verifica que U_{i+1}/U_i es un subgrupo normal mínimo de G/U_i . Los grupos cocientes U_{i+1}/U_i se denominan **factores principales** de G . Se dice que un factor principal H/K de G es **central** en G , si $H/K \leq Z(G/K)$.

Definición 2.2. Sean G un grupo, $\pi \subseteq \mathbb{P}$ y $\pi' = \mathbb{P} \setminus \pi$.

- (a) Diremos que G es un π -grupo si y solo si $\sigma(G) \subseteq \pi$, donde $\sigma(G)$ denota el conjunto de todos los divisores primos de $|G|$.
- (b) Definimos el subgrupo $O_\pi(G)$ de la siguiente manera:

$$O_\pi(G) := \langle N \mid N \trianglelefteq G, N \text{ es un } \pi\text{-grupo} \rangle.$$

- (c) G se llama π -soluble si se verifican:

- (S1) Cada factor principal de G o es un π -grupo o un π' -grupo, y
 (S2) Los π -factores principales de G son abelianos.

- (d) Llamamos a $H \leq G$ un π -subgrupo de Hall de G si todos los divisores primos de $|H|$ pertenecen a π y los divisores primos de $[G : H]$ pertenecen a π' . El conjunto de los π -subgrupos de Hall de G se notará con $\text{Hall}_\pi(G)$.

Definición 2.3. Un grupo finito G es soluble si y solo si G es π -soluble para $\pi = \mathbb{P}$. Es decir, cada factor principal de G es abeliano.

Ejemplo 2.4. (a) Todo grupo abeliano es soluble.

- (b) Los grupos simétricos $\text{Sym}(3)$ y $\text{Sym}(4)$, y los grupos alternantes $\text{Alt}(3)$ y $\text{Alt}(4)$, son solubles. Para verificarlo basta exhibir las series normales

$$\begin{aligned} \langle 1 \rangle &\leq \text{Alt}(3) \leq \text{Sym}(3) \\ \langle 1 \rangle &\leq \langle (12)(34), (13)(24) \rangle \leq \text{Alt}(4) \leq \text{Sym}(4), \end{aligned}$$

cuyos factores son abelianos.

- (c) Para $n \geq 5$ los grupos $\text{Sym}(n)$ y $\text{Alt}(n)$ no son solubles.
- (d) Si $p \in \mathbb{P}$, entonces todo p -grupo finito es soluble.

Teorema 2.5 ([5] (5.1.2)). *El producto de dos subgrupos solubles normales de un grupo es soluble.*

Definición 2.6. Un grupo finito G se denomina supersoluble si admite una serie normal, digamos

$$\langle 1 \rangle = U_0 \subseteq U_1 \subseteq \dots \subseteq U_n = G.$$

donde cada factor U_{i+1}/U_i es un grupo cíclico.

Ejemplo 2.7. (a) Todo grupo abeliano es supersoluble.

(b) Para el 4-grupo de Klein $V_4 = \{1, g, h, k\}$ se tiene la serie normal

$$\langle 1 \rangle \trianglelefteq \langle g \rangle \trianglelefteq V_4$$

con factores cíclicos.

(c) Del Ejemplo 2.4 tenemos que el grupo $\text{Sym}(4)$ es soluble, sin embargo no es supersoluble. En efecto, el grupo $\langle (12)(34), (13)(24) \rangle$ no es cíclico. En forma similar se tiene que el grupo $\text{Alt}(4)$ no es supersoluble.

Teorema 2.8 ([10] (10.5.1), [5] (9.4.5)). *Sea G es un grupo finito, $U \leq G$ y $N \trianglelefteq G$.*

(a) *Si G es supersoluble, entonces U y G/N son supersolubles.*

(b) *Si $G/\Phi(G)$ es supersoluble, entonces G es supersoluble, donde $\Phi(G)$ denota el grupo de Frattini de G .*

Teorema 2.9 ([4] VII (2,2)). *Sea G un grupo finito soluble. Entonces cualquiera de las dos siguientes afirmaciones son equivalentes:*

(a) *G es un grupo supersoluble.*

(b) *Todos los factores principales de G son cíclicos.*

(c) *Todo subgrupo máximo de G tiene índice primo.*

Definición 2.10. Un grupo finito G se denomina nilpotente si admite una serie central

$$(c) \quad \langle 1 \rangle = U_0 \subseteq U_1 \subseteq \dots \subseteq U_n = G.$$

Ejemplo 2.11. (a) Todo grupo abeliano es nilpotente.

(b) Si $p \in \mathbb{P}$, entonces todo p -grupo finito es nilpotente.

(c) Dado que el centro de $\text{Sym}(3)$ es trivial, se tiene que $\text{Sym}(3)$ no es un grupo nilpotente.

(d) El grupo $\text{Sym}(4)$ no es nilpotente.

3 Clases de grupos y operadores clausura

Asumiremos que el lector está familiarizado con la notación y terminología de la teoría de grupos finitos. Asumiremos además que todos los grupos considerados en este capítulo son finitos y solubles.

Definición 3.1. Una clase de grupos \mathfrak{X} es una clase que satisface las siguientes propiedades:

- (a) \mathfrak{X} contiene un grupo de orden 1.
- (b) Si $G \in \mathfrak{X}$, entonces cualquier grupo isomorfo a G pertenece a \mathfrak{X} .

Si G es un grupo, denotaremos por (G) la clase formada por todos los grupos isomorfos a G . Es decir

$$(G) = (H \mid H \text{ es un grupo y } H \cong G).$$

Sean \mathfrak{X} y \mathfrak{D} dos clases de grupos. Si $G \in \mathfrak{X}$ diremos que G es un \mathfrak{X} -grupo y si $\mathfrak{X} \subseteq \mathfrak{D}$ diremos que \mathfrak{X} es una subclase de \mathfrak{D} .

En la siguiente lista se presentan las clases de grupos que más se citarán a lo largo del capítulo.

\emptyset	la clase vacía de grupos;
\mathfrak{A}	la clase de todos los grupos abelianos;
\mathfrak{N}	la clase de todos los grupos nilpotentes;
\mathfrak{U}	la clase de todos los grupos supersolubles;
\mathfrak{S}	la clase de todos los grupos solubles;
\mathfrak{S}_p	la clase de todos los p -grupos, con $p \in \mathbb{P}$;
\mathfrak{S}_π	la clase de todos los π -grupos, con $\pi \subseteq \mathbb{P}$;
\mathfrak{C}	la clase de todos los grupos cíclicos;
\mathfrak{E}	la clase de todos los grupos finitos;
\mathfrak{J}	la clase de todos los grupos simples;

Las siguientes contenciones son inmediatas:

$$\mathfrak{C} \subseteq \mathfrak{A} \subseteq \mathfrak{N} \subseteq \mathfrak{U} \subseteq \mathfrak{S}.$$

Si \mathfrak{X} es una clase de grupos y $\pi \subseteq \mathbb{P}$, denotaremos por \mathfrak{X}_π la clase de todos los grupos en \mathfrak{X} cuyos órdenes involucran solamente primos en π . Si $\pi = \{p\}$, escribiremos \mathfrak{X}_p en lugar de $\mathfrak{X}_{\{p\}}$. Por lo tanto se tiene en particular que $\mathfrak{E}_p = \mathfrak{S}_p$ es la clase de todos los p -grupos, y $\mathfrak{S}_{p'}$ es la clase de los grupos solubles cuyos órdenes involucran solamente primos que pertenecen a $\mathbb{P} \setminus \{p\}$.

Definición 3.2. Si \mathfrak{X} y \mathfrak{D} son dos clases de grupos, definimos el producto $\mathfrak{X}\mathfrak{D}$ de la siguiente manera:

$$G \in \mathfrak{X}\mathfrak{D} \iff \exists N \trianglelefteq G \text{ tal que } N \in \mathfrak{X} \text{ y } G/N \in \mathfrak{D}.$$

Los siguientes dos ejemplos muestran que el producto de clases de grupos no siempre es conmutativo y asociativo respectivamente.

Observación 3.3. Consideremos las clases de grupos \mathfrak{S}_2 y \mathfrak{S}_3 y el grupo $\text{Sym}(3)$. Tenemos que $\text{Sym}(3)$ posee un subgrupo normal de orden 3 pero no posee subgrupos normales de orden 2. Por lo tanto tenemos que $\text{Sym}(3) \in \mathfrak{S}_3\mathfrak{S}_2$ y $\text{Sym}(3) \notin \mathfrak{S}_2\mathfrak{S}_3$. Luego tenemos que $\mathfrak{S}_3\mathfrak{S}_2 \neq \mathfrak{S}_2\mathfrak{S}_3$.

Observación 3.4. Sea $G = \text{Alt}(4)$. Entonces tenemos que el grupo cuatro de Klein

$$V_4 = \{(1), (12)(34), (13)(24), (14)(23)\}$$

es el único subgrupo normal propio de G . Si $H = \langle (12)(34) \rangle$, entonces H es un subgrupo de V_4 , el cual es cíclico y además $|H| = 2$, por lo tanto $H \in \mathfrak{C}$. Además $|V_4/H| = 2$ por lo que $V_4/H \in \mathfrak{C}$. Esto trae como consecuencia que $V_4 \in \mathfrak{C}\mathfrak{C}$. Dado que $|G/V_4| = 3$, se tiene que $G/V_4 \in \mathfrak{C}$, por lo tanto, ya que $V_4 \in \mathfrak{C}\mathfrak{C}$ se verifica que $G \in (\mathfrak{C}\mathfrak{C})\mathfrak{C}$.

Por otra parte, supongamos que $G \in \mathfrak{C}(\mathfrak{C}\mathfrak{C})$, puesto que $\langle 1 \rangle \trianglelefteq G$ es el único subgrupo normal y cíclico de G , por definición $G \in \mathfrak{C}(\mathfrak{C}\mathfrak{C})$ si y sólo si $\langle 1 \rangle \in \mathfrak{C}$ y $G/\langle 1 \rangle \in \mathfrak{C}\mathfrak{C}$, lo cual implica que $G \in \mathfrak{C}\mathfrak{C}$, y similarmente, esto implica que $G \in \mathfrak{C}$, lo cual no es posible. En consecuencia $(\mathfrak{C}\mathfrak{C})\mathfrak{C} \neq \mathfrak{C}(\mathfrak{C}\mathfrak{C})$.

Definición 3.5. (a) Una función de clases c , es una función que envía clases de grupos en clases de grupos. c es llamada operador clausura, si para cualesquiera \mathfrak{X} y \mathfrak{D} clases de grupos, se cumplen las siguientes condiciones:

(OC1) $\mathfrak{X} \subseteq c\mathfrak{X}$ (c es expansiva).

(OC2) $c\mathfrak{X} = c(c\mathfrak{X})$ (c es idempotente).

(OC3) Si $\mathfrak{X} \subseteq \mathfrak{D}$, entonces $c\mathfrak{X} \subseteq c\mathfrak{D}$ (c es monótona).

(b) Una clase de grupos \mathfrak{X} se llama c -cerrada si $\mathfrak{X} = c\mathfrak{X}$. Es claro que si c es un operador clausura, de (OC2) se tiene que $c\mathfrak{D}$ es c -cerrado para cualquier clase de grupo \mathfrak{D} . La clase vacía de grupos \emptyset es c -cerrada para cualquier operador clausura c .

(c) El producto AB de dos funciones de clases se define por composición así

$$AB(\mathfrak{X}) = A(B\mathfrak{X}), \text{ para toda clase de grupos } \mathfrak{X}.$$

Sean G y H dos grupos tales que $G \subseteq H$. En lo sucesivo, $G \trianglelefteq \trianglelefteq H$ indica que G es un subgrupo subnormal de H .

La siguiente lista contiene algunos de los operadores clausura de uso más frecuente. Para una clase de grupos \mathfrak{X} definimos:

$$\begin{aligned} s\mathfrak{X} &= (G \mid G \leq H \text{ para algún } H \in \mathfrak{X}); \\ q\mathfrak{X} &= (G \mid \exists H \in \mathfrak{X} \text{ y un epimorfismo } \varphi : H \rightarrow G); \\ s_n\mathfrak{X} &= (G \mid G \trianglelefteq \trianglelefteq H \text{ para algún } H \in \mathfrak{X}); \\ R_o\mathfrak{X} &= (G \mid \exists N_i \trianglelefteq G, i = 1, \dots, r \text{ con } G/N_i \in \mathfrak{X} \text{ y } \bigcap_{i=1}^r N_i = \langle 1 \rangle); \\ N_o\mathfrak{X} &= (G \mid \exists K_i \trianglelefteq \trianglelefteq G \text{ } i = 1, \dots, r \text{ con } K_i \in \mathfrak{X} \text{ y } G = \langle K_1, \dots, K_r \rangle); \\ \mathfrak{E}_\Phi\mathfrak{X} &= (G \mid \exists N \trianglelefteq G \text{ con } N \leq \Phi(G) \text{ y } G/N \in \mathfrak{X}). \end{aligned}$$

Observación 3.6. Sean \mathfrak{X} una clase de grupos y G un grupo.

- Si la clase \mathfrak{X} es s -cerrada, entonces, siempre que $G \in \mathfrak{X}$ y $N \leq G$, se cumple que $N \in \mathfrak{X}$.
- Si la clase \mathfrak{X} es q -cerrada, entonces, siempre que $G \in \mathfrak{X}$ y $N \trianglelefteq G$, se cumple que $G/N \in \mathfrak{X}$.
- Si la clase \mathfrak{X} es s_n -cerrada, entonces, siempre que $G \in \mathfrak{X}$ y $N \trianglelefteq \trianglelefteq G$, se cumple que $N \in \mathfrak{X}$.
- Si la clase \mathfrak{X} es R_o -cerrada, entonces, siempre que $N_1, N_2 \trianglelefteq G$ con $G/N_i \in \mathfrak{X}$ para $i = 1, 2$ y $N_1 \cap N_2 = \langle 1 \rangle$, se cumple que $G \in \mathfrak{X}$.
- Si la clase \mathfrak{X} es N_o -cerrada, entonces, siempre que $N_1, N_2 \trianglelefteq G$ y $N_1, N_2 \in \mathfrak{X}$ con $G = N_1 N_2$, se cumple que $G \in \mathfrak{X}$.
- Si la clase \mathfrak{X} es \mathfrak{E}_Φ -cerrada, entonces, siempre que $G/\Phi(G) \in \mathfrak{X}$, se cumple que $G \in \mathfrak{X}$.

Lema 3.7 ([4], II, 1.6). *Los operadores definidos anteriormente son operadores clausura.*

Ejemplo 3.8. La clase \mathfrak{S} de los grupos solubles es q -cerrada. Dado que q es expansiva se sigue inmediatamente que $\mathfrak{S} \subseteq q\mathfrak{S}$. Sea $G \in q\mathfrak{S}$, entonces existe $H \in \mathfrak{S}$ y un epimorfismo $\varphi : H \rightarrow G$.

Sea $N = \ker(\varphi) \trianglelefteq H$, como H es soluble se tiene que N y H/N también lo son. Por el primer teorema de isomorfismo para grupos se cumple que $H/N \cong G$. Así que $G \in \mathfrak{S}$. Luego se tiene que $q\mathfrak{S} \subseteq \mathfrak{S}$ y en consecuencia $q\mathfrak{S} = \mathfrak{S}$.

Lema 3.9 ([4], II, 1.10). Si \mathfrak{X} , \mathfrak{D} y \mathfrak{B} son clases de grupos, entonces cada una de las siguientes dos condiciones son suficientes para asegurar que $\mathfrak{X}(\mathfrak{D}\mathfrak{B}) = (\mathfrak{X}\mathfrak{D})\mathfrak{B}$.

$$(a) \quad \mathfrak{X} = \mathfrak{N}_o\mathfrak{X} \quad \text{y} \quad \mathfrak{D} = \mathfrak{Q}\mathfrak{D}.$$

$$(b) \quad \mathfrak{X} = \mathfrak{s}_n\mathfrak{X} \quad \text{y} \quad \mathfrak{D} = \mathfrak{R}_o\mathfrak{D}.$$

Definición 3.10. Sean A y B operadores clausura, diremos que A está contenido en B, que denotaremos con $A \leq B$, si $A\mathfrak{X} \subseteq B\mathfrak{X}$, para toda clase de grupos \mathfrak{X} .

Ejemplo 3.11. Consideremos los operadores clausura \mathfrak{s}_n y \mathfrak{s} . Mostremos que $\mathfrak{s}_n \leq \mathfrak{s}$. Sean \mathfrak{X} una clase de grupos cualquiera y $G \in \mathfrak{s}_n\mathfrak{X}$. Entonces por definición del operador \mathfrak{s}_n se tiene que existe $H \in \mathfrak{X}$ tal que $G \trianglelefteq H$. Luego por la definición de grupo subnormal y del operador clausura \mathfrak{s} se tiene que $G \leq H$, y entonces $G \in \mathfrak{s}\mathfrak{X}$. Dado que G fue arbitrario se tiene que $\mathfrak{s}_n\mathfrak{X} \subseteq \mathfrak{s}\mathfrak{X}$ y en consecuencia $\mathfrak{s}_n \leq \mathfrak{s}$.

4 Formaciones saturadas o localmente definidas

Definición 4.1. Una clase de grupos \mathfrak{F} es llamada formación, si es \mathfrak{Q} -cerrada y \mathfrak{R}_o -cerrada. Si además \mathfrak{F} es \mathfrak{E}_Φ -cerrada, se denomina una formación saturada.

Ejemplos de formaciones son \mathfrak{C} , \mathfrak{A} , \mathfrak{N} , \mathfrak{U} y \mathfrak{S} .

Ejemplo 4.2. La clase \mathfrak{S} es una formación saturada.

Demostración. En el Ejemplo 3.8 se probó que la clase de grupos \mathfrak{S} es \mathfrak{Q} -cerrada. Sea $G \in \mathfrak{R}_0\mathfrak{S}$ arbitrario, entonces existen $N_1, \dots, N_r \trianglelefteq G$ con $\bigcap_{i=1}^r N_i = \langle 1 \rangle$ tales que $G/N_i \in \mathfrak{S}$, y en consecuencia

$$G \cong G/\langle 1 \rangle \cong G/\bigcap_{i=1}^r N_i \in \mathfrak{S}.$$

Es inmediato que $\mathfrak{S} \subseteq \mathfrak{R}_0\mathfrak{S}$, por lo tanto, la clase de grupos \mathfrak{S} es \mathfrak{R}_0 -cerrada. Todo lo anterior demuestra que \mathfrak{S} es una formación.

Mostremos ahora que \mathfrak{S} es \mathfrak{E}_Φ -cerrada. Sea $G \in \mathfrak{E}_\Phi\mathfrak{S}$ arbitrario, entonces existe $N \trianglelefteq G$ con $N \leq \Phi(G)$ tal que $G/N \in \mathfrak{S}$, y dado que $\Phi(G)$ es nilpotente, se cumple que $N \in \mathfrak{S}$. En consecuencia $G \in \mathfrak{S}$. Por otra parte, es claro que $\mathfrak{S} \subseteq \mathfrak{E}_\Phi\mathfrak{S}$. Las dos contenencias anteriores demuestran que $\mathfrak{E}_\Phi\mathfrak{S} \subseteq \mathfrak{S}$. \square

Proposición 4.3. Sea \mathfrak{X} una clase de grupos. Son equivalentes:

$$(1) \quad \mathfrak{R}_0\mathfrak{X} = \mathfrak{X}.$$

(2) Si $N_1, N_2 \trianglelefteq G$ con $N_1 \cap N_2 = 1$ y $G/N_i \in \mathfrak{X}$ para $i = 1, 2$, entonces $G \in \mathfrak{X}$.

Demostración. (1) \Rightarrow (2). Se sigue de la definición de $\mathfrak{R}_0\mathfrak{X}$.

(2) \Rightarrow (1). Supongamos que (2) es válido y que $\mathfrak{X} \neq \mathfrak{R}_0\mathfrak{X}$. Dado que $\mathfrak{X} \subset \mathfrak{R}_0\mathfrak{X}$ se tiene que $\mathfrak{R}_0\mathfrak{X} \setminus \mathfrak{X} \neq \emptyset$.

Sea $A \in \mathfrak{R}_0\mathfrak{X} \setminus \mathfrak{X}$ con orden mínimo. Entonces existen $K_1, \dots, K_r \trianglelefteq A$ con $\bigcap_{i=1}^r K_i = \langle 1 \rangle$ tales que $A/K_i \in \mathfrak{X}$, para $i = 1, \dots, r$. Sin pérdida de generalidad supongamos que para todo subconjunto propio \mathcal{L} de $\{K_1, \dots, K_r\}$ se verifica que $\bigcap_{K \in \mathcal{L}} K \neq \langle 1 \rangle$. Si $r = 1$, entonces se tendría $K_1 = \langle 1 \rangle$ y por lo tanto $A \in \mathfrak{X}$, lo cual es una contradicción. Por lo tanto $r > 1$.

Los subgrupos $N_1 := \bigcap_{i=1}^{r-1} K_i$ y $N_2 := \bigcap_{i=2}^r K_i$ son subgrupos normales no triviales de A tal que

$$(N_1 \cap N_2) \leq \bigcap_{i=1}^r K_i = \langle 1 \rangle.$$

Dado que

$$(A/N_1)/(K_i/N_1) \cong A/K_i \in \mathfrak{X}, \text{ para } i = 1, \dots, r-1,$$

y

$$\bigcap_{i=1}^{r-1} (K_i/N_1) = N_1/N_1,$$

se sigue que $A/N_1 \in \mathfrak{R}_0\mathfrak{X}$ y además $A/N_1 \in \mathfrak{X}$ por la elección de A .

Similarmente se demuestra que $A/N_2 \in \mathfrak{X}$. Entonces por (2) se tiene que $A \in \mathfrak{X}$, lo cual es una contradicción. Esto demuestra que $\mathfrak{R}_0\mathfrak{X} = \mathfrak{X}$. \square

Proposición 4.4. Sea \mathfrak{X} una clase de grupos. Son equivalentes:

(1) $\mathfrak{q}(\mathfrak{X}) = \mathfrak{X}$.

(2) Si $G \in \mathfrak{X}$ y $N \trianglelefteq G$ entonces $G/N \in \mathfrak{X}$.

Demostración. (1) \Rightarrow (2). Sea $G \in \mathfrak{X}$ y $N \trianglelefteq G$. Entonces tomando el epimorfismo canónico $\varphi : G \rightarrow G/N$ se tiene que $G/N \in \mathfrak{q}\mathfrak{X} = \mathfrak{X}$.

(2) \Rightarrow (1). Es inmediato que $\mathfrak{X} \subseteq \mathfrak{q}(\mathfrak{X})$. Sea $H \in \mathfrak{q}(\mathfrak{X})$ cualquiera. Entonces existe $M \in \mathfrak{X}$ y un epimorfismo $\psi : M \rightarrow H$. Si tomemos $S := \ker(\psi) \trianglelefteq M$ se tiene que $M/S \in \mathfrak{X}$.

Por otra parte, por el primer Teorema de isomorfismo para grupos se verifica que $H \cong M/S$. Por lo tanto, $H \in \mathfrak{X}$ y así $\mathfrak{q}(\mathfrak{X}) \subseteq \mathfrak{X}$. \square

Observación 4.5. Teniendo en cuenta las Proposiciones 4.3 y 4.4 se puede definir una formación como una clase de grupos \mathfrak{F} que satisface las siguientes propiedades:

(F1) Si $G \in \mathfrak{F}$ y $N \trianglelefteq G$, entonces $G/N \in \mathfrak{F}$.

(F2) Si $N_1, N_2 \trianglelefteq G$ con $N_1 \cap N_2 = \langle 1 \rangle$ y $G/N_i \in \mathfrak{F}$ para $i = 1, 2$, entonces $G \in \mathfrak{F}$.

Proposición 4.6. *Sea \mathfrak{F} una formación. Entonces, \mathfrak{F} es saturada si y sólo si, siempre que $G/\Phi(G) \in \mathfrak{F}$ se cumple que $G \in \mathfrak{F}$.*

Demostración. Supongamos que \mathfrak{F} es una formación saturada. Sea G un grupo cualquiera tal que $G/\Phi(G) \in \mathfrak{F}$. Entonces, teniendo en cuenta que $\Phi(G) \trianglelefteq G$, $\Phi(G) \trianglelefteq \Phi(G)$ y $G/\Phi(G) \in \mathfrak{F}$, se cumple que $G \in \mathbf{E}_\Phi \mathfrak{F} = \mathfrak{F}$.

Recíprocamente, sea G un grupo y supongamos que de $G/\Phi(G) \in \mathfrak{F}$ se sigue siempre que $G \in \mathfrak{F}$. Dado que $\mathfrak{F} \subseteq \mathbf{E}_\Phi \mathfrak{F}$, sólo restaría probar que $\mathbf{E}_\Phi \mathfrak{F} \subseteq \mathfrak{F}$.

Si $H \in \mathbf{E}_\Phi \mathfrak{F}$, entonces existe $N \trianglelefteq H$ con $N \trianglelefteq \Phi(H)$ tal que $H/N \in \mathfrak{F}$. Como $\Phi(H) \trianglelefteq H$, se tiene que $\Phi(H)/N \trianglelefteq H/N$. Por lo tanto

$$H/\Phi(H) \cong (H/N)/(\Phi(H)/N) \in \mathfrak{F}$$

y así $H \in \mathfrak{F}$. □

Definición 4.7. Sea \mathfrak{X} una clase de grupos. Un subgrupo U de un grupo G es llamado \mathfrak{X} -máximo en G , si se verifican:

(a) $U \in \mathfrak{X}$

(b) Si $U \leq V \leq G$ y $V \in \mathfrak{X}$, entonces $U = V$.

Ejemplo 4.8. Sean $\pi \subseteq \mathbb{P}$, $\mathfrak{X} = \mathfrak{S}_\pi$, la clase de los grupos π -solubles y G un grupo soluble. Sea además $H \in \text{Hall}_\pi G$ cualquiera. Entonces $H \leq G$ y $\sigma(H) \subseteq \pi$. En consecuencia $H \in \mathfrak{X}$.

Supongamos que existe $F \in \mathfrak{X}$ tal que se cumple $H \leq F < G$. Entonces, dado que cualesquiera dos π -subgrupo de Hall de G son conjugados, se tiene que $H \cong F$, es decir $H = F$. Por lo tanto H es \mathfrak{X} -máximo en G .

Entonces, si $\mathfrak{X} = \mathfrak{S}_\pi$, se verifica que los subgrupos \mathfrak{S}_π -máximos de un grupo soluble G son sus π -subgrupos de Hall. Más aún, si H es un π -subgrupo de Hall de G y $K \trianglelefteq G$, entonces HK/K es un π -subgrupo de Hall de G/K .

Definición 4.9. Sea \mathfrak{X} una clase de grupos. Un subgrupo U de un grupo G es llamado un \mathfrak{X} -proyector de G si

$$UK/K \text{ es } \mathfrak{X}\text{-máximo en } G/K \text{ para todo } K \trianglelefteq G.$$

Usaremos el símbolo $\text{Proj}_\mathfrak{X}(G)$ para denotar el conjunto (probablemente vacío) de todos los \mathfrak{X} -proyectores de G .

Ejemplo 4.10. (a) Sea G un grupo finito soluble. Es inmediato que

$$\text{Proj}_{\mathfrak{S}_\pi}(G) = \text{Hall}_\pi(G) \neq \emptyset$$

En particular $\text{Proj}_{\mathfrak{S}_p}(G) = \text{Syl}_p(G)$, donde $\text{Syl}_p(G)$ es el conjunto de los p -subgrupos de Sylow de G .

(b) Sean G un p -grupo no abeliano y \mathfrak{A} la clase de los grupos abelianos finitos. Entonces $\text{Proj}_{\mathfrak{A}}(G) = \emptyset$. Veamos:

Sea G un grupo tal que $|G| = p^a$, con $a \in \mathbb{N}$ y $2 < a$. Supongamos que $U \in \text{Proj}_{\mathfrak{A}}(G)$. Entonces U es abeliano. Si $U = G$, entonces se contradice la hipótesis que G es no abeliano, así que podemos suponer que $U < G$.

Sea $G' = [G, G] \trianglelefteq G$. Como G/G' es abeliano, se tiene entonces que $G/G' \in \mathfrak{A}$ y además UG'/G' es \mathfrak{A} -máximo en G/G' , consecuentemente $UG'/G' \leq G/G'$, y por lo tanto $UG' \leq G$. Dado que U es un \mathfrak{A} -proyector de G , se tiene que U es \mathfrak{A} -máximo en G , así que como $U \leq UG'$ se concluye que $G = UG'$.

Sea $M \triangleleft G$ arbitrario, esto es, M es un subgrupo máximo de G . Entonces dado que G es nilpotente se tiene que $M \trianglelefteq G$, y como cada factor principal H/K de G cumple que $H/K \leq Z(G/K)$, se tiene entonces que G/M es abeliano, luego para cada $x, y \in G$ se satisface que

$$\begin{aligned} x^{-1}My^{-1}M = y^{-1}Mx^{-1}M &\Leftrightarrow x^{-1}y^{-1}M = y^{-1}x^{-1}M \\ &\Leftrightarrow xyx^{-1}y^{-1}M = M \\ &\Leftrightarrow xyx^{-1}y^{-1} \in M, \end{aligned}$$

por lo tanto como M fue arbitrario se tiene que $G' \leq \Phi(G)$.

Por otra parte, existe $S \triangleleft G$ tal que $U \leq S$. Además, se tiene que $G' \leq \Phi(G) \leq S$ por lo cual se cumple que $G = G'U \leq S$, lo cual es una contradicción. Por lo tanto se tiene que $\text{Proj}_{\mathfrak{A}}(G) = \emptyset$.

Teorema 4.11. Sean $\mathfrak{X} = \mathfrak{Q}\mathfrak{X}$ una clase de grupos y G un grupo finito. Si $U \in \text{Proj}_{\mathfrak{X}}(G)$ y $N \trianglelefteq G$, entonces $UN/N \in \text{Proj}_{\mathfrak{X}}(G/N)$.

Demostración. Dado que $U \in \text{Proj}_{\mathfrak{X}}(G)$, se tiene que UM/M es \mathfrak{X} -máximo en G/M para todo $M \trianglelefteq G$. Además es inmediato que $UN/N \in \mathfrak{X}$.

Sea $L/N \trianglelefteq G/N$ cualquiera. Entonces $L \trianglelefteq G$ y así UL/L es \mathfrak{X} -máximo en G/L . Usando un Teorema de isomorfismo tenemos

$$[(UN/N)(L/N)]/(L/N) = (UL/N)/(L/N) \cong UL/L$$

y

$$(G/N)/(L/N) \cong G/L.$$

Por lo tanto UN/N es un \mathfrak{X} -proyector de G/N . □

Definición 4.12. Sea \mathfrak{X} una clase de grupos no vacía y G un grupo. Definimos y denotamos el \mathfrak{X} -residuo de G de la siguiente manera:

$$G^{\mathfrak{X}} := \bigcap \{H \trianglelefteq G \mid G/H \in \mathfrak{X}\}.$$

Consecuentemente $G^{\mathfrak{X}} = \langle 1 \rangle$, si y sólo si $G \in \mathfrak{X}$.

Ejemplo 4.13. El residual abeliano $G^{\mathfrak{A}}$ de un grupo G está dado por su subgrupo derivado, el cual denotaremos por G' . Para demostrarlo, sea $N \trianglelefteq G$ cualquiera tal que $G/N \in \mathfrak{A}$. Entonces

$$\begin{aligned} G/N \text{ es abeliano} &\Leftrightarrow (aN)(bN) = (bN)(aN) \quad \forall a, b \in G \\ &\Leftrightarrow aba^{-1}b^{-1}N = N \quad \forall a, b \in G \\ &\Leftrightarrow [a, b] \in N \quad \forall a, b \in G \\ &\Leftrightarrow G' \leq N \quad \forall a, b \in G. \end{aligned}$$

Dado que N fue arbitrario se sigue que $G' \subseteq G^{\mathfrak{A}}$. Recíprocamente, dado que $G' \trianglelefteq G$ y que G/G' es un grupo abeliano, se sigue que $G^{\mathfrak{A}} \subseteq G'$. Por lo tanto $G^{\mathfrak{A}} = G'$.

Definición 4.14. Sean \mathfrak{X} una clase de grupos y \mathfrak{F} una formación. Definimos

$$\mathfrak{X} \circ \mathfrak{F} := (G : G^{\mathfrak{F}} \in \mathfrak{X}),$$

y llamaremos $\mathfrak{X} \circ \mathfrak{F}$ la formación producto de \mathfrak{X} con \mathfrak{F} .

Ejemplo 4.15. Supongamos que $\mathfrak{X} = (1)$ y $\mathfrak{F} = \mathfrak{S}$. Si $G \in \mathfrak{S}$ entonces $G^{\mathfrak{F}} \in \mathfrak{X}$. Por lo tanto, para cualquier grupo soluble G , se tiene que $G \in \mathfrak{X} \circ \mathfrak{S}$.

Lema 4.16 ([4], II, 2.7). *Sea \mathfrak{F} una formación. Entonces*

$$\bigcap_{p \in \mathbb{P}} \mathfrak{E}_{p'} \mathfrak{S}_p \mathfrak{F} = \mathfrak{N}\mathfrak{F}.$$

Definición 4.17. Una función f que a cada número primo p asigna una formación $f(p)$ se llama función formación. Si f es una función formación, entonces un factor principal H/K de un grupo G se llama f -central si, $G/C_G(H/K) \in f(p)$, para todo número primo p que divide a $|H/K|$. El soporte de la función formación f se denota y define como

$$\text{Supp}(f) := \{p \in \mathbb{P} : f(p) \neq \emptyset\}.$$

Ejemplo 4.18. Sea $f : \mathbb{P} \rightarrow \{\text{Formaciones}\}$ una función definida por

$$f(p) = \mathfrak{S}.$$

Tenemos entonces que f es una función formación. Ahora, dado $G \in \mathfrak{S}$ cualquiera, se tiene que dado un factor principal H/K de G se cumple que

$$G/C_G(H/K) \in \mathfrak{S} = f(p)$$

para todo p que divide a $|H/K|$. Así que cualquier factor principal H/K de G es f -central. Es inmediato que

$$\text{Supp}(f) = \mathbb{P}.$$

Definición 4.19. Dada la función formación f , definiremos y denotaremos por

$$LF(f) := (G \mid \text{todo factor principal de } G \text{ es } f\text{-central}).$$

Diremos que la formación \mathfrak{F} es local o que está definida localmente por f , si $\mathfrak{F} = LF(f)$.

En el siguiente ejemplo presentamos la función formación que define localmente la clase de los grupos nilpotentes.

Ejemplo 4.20. Si definimos $f(p) = (1)$, para todo $p \in \mathbb{P}$, entonces $LF(f) = \mathfrak{N}$.

Demostración. Sea $G \in \mathfrak{N}$. Entonces todo factor principal H/K de G es central. Es decir $H/K \leq Z(G/K)$. Luego, dado cualquier factor principal H/K de G y $p \in \mathbb{P}$ con $p \mid |H/K|$ se tiene

$$\begin{aligned} C_G(H/K) &= \{g \in G \mid h^g K = hK \quad \forall h \in H\} \\ &= \{g \in G \mid hKg^{-1}K = g^{-1}KhK \quad \forall h \in H\} \\ &= \{g \in G \mid hKg^{-1}K = hKg^{-1}K \quad \forall h \in H\} \\ &= \{g \in G \mid hg^{-1}K = hg^{-1}K \quad \forall h \in H\} \\ &= G \end{aligned}$$

Recíprocamente, sea $G \in LF(f)$ y H/K un factor principal de G . Entonces

$$G/C_G(H/K) \in f(p) = (1) \quad \forall p \mid |H/K|.$$

Así que $G = C_G(H/K)$, en consecuencia $H/K \leq Z(G/K)$. Por lo tanto $G \in \mathfrak{N}$. \square

Definición 4.21. Si \mathfrak{F} es una formación local, definimos su característica, denotada $\text{Char}(\mathfrak{F})$, como el conjunto de números primos p para los cuales se verifica que el grupo cíclico de orden p , digamos C_p pertenece a \mathfrak{F} .

Sea p un número primo y $\mathfrak{A}(p-1) := \mathfrak{A} \cap \mathfrak{E}(p-1)$, la formación de todos los grupos abelianos finitos cuyos exponentes dividen a $p-1$. Es un **ejercicio** mostrar que la formación de los grupos finitos supersolubles \mathfrak{A} está definida localmente por la función formación f definida por

$$f(p) = \mathfrak{A}(p-1), \text{ para todo } p \in \mathbb{P}.$$

Ejemplo 4.22. Sea \mathfrak{F} una formación, y sea $f(p) = \mathfrak{F}$ para todo $p \in \mathbb{P}$. Entonces del Lema 4.16 se tiene que

$$LF(f) = \bigcap_{p \in \mathbb{P}} \mathfrak{E}_{p'} \mathfrak{S}_p \mathfrak{F} = \mathfrak{N}\mathfrak{F}.$$

Definición 4.23. Sea \mathfrak{F} una formación.

- Si \mathfrak{F} está definida localmente por f y además para todo $p \in \mathbb{P}$ se cumple que $f(p) \subseteq \mathfrak{F}$ entonces diremos que \mathfrak{F} se define local e inclusiva mediante f .
- Si \mathfrak{S}_p denota la clase de los p -grupos finitos y se cumple que $\mathfrak{S}_p f(p) = f(p)$ para todo primo p , entonces se dice que f es completa y diremos que la definición local de \mathfrak{F} es completa.
- Una función formación F es llamada definición canónica local para \mathfrak{F} si es inclusiva y completa. En éste caso usaremos la notación $\mathfrak{F} = LF(F)$.

Observación 4.24. Cada formación local $\mathfrak{F} = LF(f)$ puede definir una función formación inclusiva y completa g dada por

$$g(p) = \mathfrak{F} \cap \mathfrak{S}_p f(p), \text{ para todo } p \in \mathbb{P},$$

y es claro que como f es una función formación, entonces g también lo es.

Teorema 4.25 ([4], IV (3,7)). *Sea $\mathfrak{F} = LF(f)$, y definamos una función formación*

$$F : \mathbb{P} \longrightarrow \{\text{Clases de grupos}\}$$

como sigue:

$$F(p) = \begin{cases} \mathfrak{Q}(G \in \mathfrak{F} : O_{p'}(G) = \langle 1 \rangle) & \text{para } p \in \text{Char}(\mathfrak{F}) \\ \emptyset & \text{para } p \notin \text{Char}(\mathfrak{F}). \end{cases}$$

Entonces F es la única función formación inclusiva y completa tal que $\mathfrak{F} = LF(F)$.

Teorema 4.26 ([4], IV (4,6)). *(Gaschütz - Lubeseder - Schmid). Una formación de grupos finitos es saturada si y sólo si ésta es local.*

Proposición 4.27. *Sea $\mathfrak{F} = LF(F)$ donde F es una definición canónica local de \mathfrak{F} , y sea f una función formación. Si $\mathfrak{F} = LF(f)$, entonces*

$$F(p) = \mathfrak{S}_p f(p) \cap \mathfrak{F} = \mathfrak{S}_p(f(p) \cap \mathfrak{F}), \text{ para todo } p \in \mathbb{P}.$$

Demostración. Dado que la función formación definida en la Observación 4.24 es inclusiva y completa, se deduce de la unicidad de F en el Teorema 4.25 que $F(p) = \mathfrak{S}_p f(p) \cap \mathfrak{F}$, así que, $F(p) \subseteq \mathfrak{S}_p f(p)$ y $F(p) \subseteq \mathfrak{F}$.

Luego, dado $G \in F(p)$ se tiene que existe $N \trianglelefteq G$ con $N \in \mathfrak{S}_p$ tal que $G/N \in f(p)$, pero $F(p) \subseteq \mathfrak{F}$ y \mathfrak{F} es una formación, así que $G \in \mathfrak{F}$ y en consecuencia $G/N \in \mathfrak{F}$. Por lo tanto, dado $G \in F(p)$ existe $N \trianglelefteq G$ con $N \in \mathfrak{S}_p$ tal que $G/N \in f(p) \cap \mathfrak{F}$, así que $F(p) \subseteq \mathfrak{S}_p(f(p) \cap \mathfrak{F})$. Es claro que $f(p) \cap \mathfrak{F} \subseteq F(p)$ y de éste hecho se deduce que $\mathfrak{S}_p(f(p) \cap \mathfrak{F}) \subseteq \mathfrak{S}_p F(p) = F(p)$. En consecuencia, $F(p) = \mathfrak{S}_p(f(p) \cap \mathfrak{F})$. \square

Teorema 4.28 ([4], IV (3,13)). *Sea $\mathfrak{F} = LF(F)$, y sea \mathfrak{G} una formación no vacía. Asumamos que:*

(i) $\mathfrak{G} = LF(G)$ o

(ii) $\mathfrak{S}_p \mathfrak{G} = \mathfrak{G}$ para todo $p \notin \text{Char}(\mathfrak{F})$.

Entonces $\mathfrak{F} \circ \mathfrak{G} = LF(H)$, donde

$$H(p) = \begin{cases} F(p) \circ \mathfrak{G} & \text{si } p \in \text{Char}(\mathfrak{F}), \\ G(p) & \text{si } p \notin \text{Char}(\mathfrak{F}) \text{ en el caso (i),} \\ \mathfrak{G} & \text{si } p \notin \text{Char}(\mathfrak{F}) \text{ en el caso (ii).} \end{cases}$$

Proposición 4.29 ([4] IV (3,8)(b)). *Sea $\mathfrak{F} = LF(F)$ donde F es una definición canónica local de \mathfrak{F} , y sea f una función formación. Si $f(p) \subseteq \mathfrak{G}$ y $F(p) = \mathfrak{S}_p(f(p) \cap \mathfrak{F})$ para todo $p \in \mathbb{P}$, entonces $\mathfrak{F} = LF(f)$.*

Definición 4.30. Sea \mathfrak{F} una formación saturada y \mathfrak{H} una formación. Definamos la clase $(\mathfrak{F} \downarrow \mathfrak{H})$ como sigue:

$$(\mathfrak{F} \downarrow \mathfrak{H}) := (G \in \mathfrak{G} : \text{Proj}_{\mathfrak{F}}(G) \subseteq \mathfrak{H}).$$

Proposición 4.31. *Sean \mathfrak{F} y \mathfrak{H} dos formaciones. Entonces $(\mathfrak{F} \downarrow \mathfrak{H})$ es una formación.*

Demostración. Sean $G \in (\mathfrak{F} \downarrow \mathfrak{H})$ y $N \trianglelefteq G$ arbitrarios. Sin pérdida de generalidad supongamos que $\text{Proj}_{\mathfrak{F}}(G) \neq \emptyset$. Luego por definición de $(\mathfrak{F} \downarrow \mathfrak{H})$ se tiene que $\text{Proj}_{\mathfrak{F}}(G) \subseteq \mathfrak{H}$. Sea $H \in \text{Proj}_{\mathfrak{F}}(G)$. Entonces tenemos que $H \in \mathfrak{H}$ y del Teorema

4.11 se tiene que HN/N es un \mathfrak{F} -proyector de G/N . Además como \mathfrak{H} es \mathcal{Q} -cerrada se sigue que $HN/N \cong H/(H \cap N) \in \mathfrak{H}$. Por lo tanto, como los \mathfrak{F} -proyector de G/N son conjugados entre sí se tiene que cualquier \mathfrak{F} -proyector de G/N está en \mathfrak{H} , por lo tanto se cumple que $G/N \in (\mathfrak{F} \downarrow \mathfrak{H})$. Luego, como G y N fueron arbitrarios tenemos que $(\mathfrak{F} \downarrow \mathfrak{H})$ es \mathcal{Q} -cerrado. Probemos ahora que $(\mathfrak{F} \downarrow \mathfrak{H})$ es \mathcal{R}_o -cerrado.

Sean $N_1, N_2 \trianglelefteq G \in \mathfrak{S}$ tales que $G/N_i \in (\mathfrak{F} \downarrow \mathfrak{H})$ ($i = 1, 2$) y $N_1 \cap N_2 = \langle 1 \rangle$. Sea $H \in \text{Proj}_{\mathfrak{F}}(G)$ arbitrario, entonces se tiene que

$$H/(H \cap N_i) \cong HN_i/N_i \in \text{Proj}_{\mathfrak{F}}(G/N_i) \quad (i = 1, 2).$$

Por lo tanto como $G/N_i \in (\mathfrak{F} \downarrow \mathfrak{H})$ ($i = 1, 2$) se sigue que $H/(H \cap N_i) \in \mathfrak{H}$ ($i = 1, 2$) y en consecuencia como \mathfrak{H} es una formación se cumple que $H \in \mathcal{R}_o\mathfrak{H} = \mathfrak{H}$. Luego, como H fue arbitrario se tiene que $G \in (\mathfrak{F} \downarrow \mathfrak{H})$. Así hemos probado que $(\mathfrak{F} \downarrow \mathfrak{H})$ es \mathcal{Q} y \mathcal{R}_o -cerrada, y por lo tanto es una formación. \square

Proposición 4.32. Sean \mathfrak{F} y \mathfrak{H} dos formaciones. Entonces $(\mathfrak{H} \downarrow \mathfrak{F})$ satisface la siguiente propiedad:

$$\mathfrak{H} \cap \mathfrak{F} = \mathfrak{H} \cap (\mathfrak{H} \downarrow \mathfrak{F}).$$

Demostración.

$$\begin{aligned} G \in \mathfrak{H} \cap (\mathfrak{H} \downarrow \mathfrak{F}) &\Leftrightarrow G \in \mathfrak{H} \wedge G \in (\mathfrak{H} \downarrow \mathfrak{F}) \\ &\Leftrightarrow G \in \mathfrak{H} \wedge \text{Proj}_{\mathfrak{H}}(G) \subseteq \mathfrak{F} \\ &\Leftrightarrow G \in \mathfrak{H} \wedge G \in \mathfrak{F} \\ &\Leftrightarrow G \in \mathfrak{H} \cap \mathfrak{F}. \end{aligned} \quad \square$$

Teorema 4.33. Sea $\mathfrak{F} = LF(F)$, y para todo $p \in \mathbb{P}$ sea $f^*(p) := (\mathfrak{F} \downarrow F(p))$. Entonces $f^*(p)$ es una definición canónica local completa para \mathfrak{F} .

Demostración. Mostremos inicialmente que $f^*(p)$ es una función formación. Para todo $p \in \mathbb{P}$ se tiene que $(\mathfrak{F} \downarrow F(p))$ es una formación, así que claramente se cumple que $f^*(p)$ es una función formación. Probemos ahora que $f^*(p)$ es completa y que $\mathfrak{F} = LF(f^*(p))$. Sean $p \in \mathbb{P}$ y $G \in \mathfrak{S}_p f^*(p)$ arbitrarios pero fijos. Entonces existe $N \trianglelefteq G$ con $N \in \mathfrak{S}_p$ tal que $G/N \in f^*(p)$. Sea E un \mathfrak{F} -proyector de G arbitrario. Se sigue entonces que $EN/N \cong E/(E \cap N)$ es un \mathfrak{F} -proyector de G/N . Luego se cumple que $EN/N \cong E/(E \cap N) \in F(p)$. Como $E \cap N \leq N$ y N es un p -grupo, se tiene que $E \cap N \in \mathfrak{S}_p$ y además $E \cap N \trianglelefteq E$, así que $E \in \mathfrak{S}_p F(p) = F(p)$. Como E fue un \mathfrak{F} -proyector arbitrario, se concluye que $G \in f^*(p)$. Así que $\mathfrak{S}_p f^*(p) \subseteq f^*(p)$ y dado que $f^*(p) \subseteq \mathfrak{S}_p f^*(p)$ se tiene entonces que $\mathfrak{S}_p f^*(p) = f^*(p)$. Por lo tanto

$f^*(p)$ es una función formación completa. Por otra parte, como F es inclusiva se tiene por la Proposición 4.32 que

$$f^*(p) \cap \mathfrak{F} = (\mathfrak{F} \downarrow F(p)) \cap \mathfrak{F} = \mathfrak{F} \cap (\mathfrak{F} \downarrow F(p)) = \mathfrak{F} \cap F(p) = F(p),$$

de lo cual se sigue que $\mathfrak{S}_p(f^*(p) \cap \mathfrak{F}) = \mathfrak{S}_p F(p) = F(p)$, luego se tiene por la Proposición 4.29 que $\mathfrak{F} = LF(f^*(p))$. Por lo tanto, como p y G fueron arbitrarios se prueba lo deseado. \square

5 La relación de contenido fuerte entre formaciones locales

Definición 5.1. Sean \mathfrak{F} y \mathfrak{H} dos formaciones saturadas. Diremos que \mathfrak{F} está contenido fuertemente en \mathfrak{H} , denotado $\mathfrak{F} \ll \mathfrak{H}$, si, para cada grupo soluble G , todo \mathfrak{F} -proyector de G está contenido (como subgrupo) en un \mathfrak{H} -proyector de G .

Ejemplo 5.2. Sean π y $\tau \subseteq \mathbb{P}$ tales que $\pi \subseteq \tau$, y G un grupo soluble cualquiera. Sea $H \in \text{Proj}_{\mathfrak{S}_\pi}(G)$ arbitrario. Como $\pi \subseteq \tau$, tenemos que existe $H'' \in \text{Hall}_\tau(G)$ tal que $H \leq H''$. Dado que todo τ -subgrupo de Hall de G es un τ -proyector de G , se tiene que $H'' \in \text{Proj}_\tau(G)$. Por lo tanto, como H y G fueron arbitrarios tenemos que $\mathfrak{S}_\pi \ll \mathfrak{S}_\tau$.

Observación 5.3. Se verifica que \ll define un orden parcial sobre la clase \mathcal{S} de las formaciones saturadas. En efecto,

- (a) Para todo $\mathfrak{F} \in \mathcal{S}$ es inmediato que $\mathfrak{F} \ll \mathfrak{F}$. Así que \ll es reflexiva.
- (b) Sean $\mathfrak{F}, \mathfrak{H} \in \mathcal{S}$ tales que $\mathfrak{F} \ll \mathfrak{H}$ y $\mathfrak{H} \ll \mathfrak{F}$.

Sea $H \leq G$, con H un \mathfrak{F} -proyector de G . Entonces existe $M \leq G$ con M un \mathfrak{H} -proyector de G tal que $H \leq M$. Para éste M , existe $N \leq G$ tal que N es un \mathfrak{F} -proyector y $M \leq N$. Sea $K \trianglelefteq G$ arbitrario, entonces tenemos que HK/K y NK/K son \mathfrak{F} -máximos en G/K con $HK/K \leq NK/K$. En consecuencia por la definición de \mathfrak{F} -máximo se tiene que $H = N$ y dado que $H \leq M \leq N$ se concluye que $H = M = N$. Por lo tanto, como $H \in \text{Proj}_{\mathfrak{H}}(G)$ fue arbitrario, tenemos que $\text{Proj}_{\mathfrak{F}}(G) \subseteq \text{Proj}_{\mathfrak{H}}(G)$. La prueba que $\text{Proj}_{\mathfrak{H}}(G) \subseteq \text{Proj}_{\mathfrak{F}}(G)$ es similar. En consecuencia $\text{Proj}_{\mathfrak{F}}(G) = \text{Proj}_{\mathfrak{H}}(G)$. Luego considerando la igualdad anterior se tiene lo siguiente:

$$\begin{aligned} G \in \mathfrak{F} &\Leftrightarrow G \in \text{Proj}_{\mathfrak{F}}(G) \\ &\Leftrightarrow G \in \text{Proj}_{\mathfrak{H}}(G) \\ &\Leftrightarrow G \in \mathfrak{H}. \end{aligned}$$

Con lo cual queda probado que $\mathfrak{F} = \mathfrak{H}$.

- (c) Sean $\mathfrak{F}_1, \mathfrak{F}_2, \mathfrak{F}_3 \in \mathcal{S}$ tales que $\mathfrak{F}_1 \ll \mathfrak{F}_2$ y $\mathfrak{F}_2 \ll \mathfrak{F}_3$. Entonces es inmediato de la transitividad entre subgrupos finitos que $\mathfrak{F}_1 \ll \mathfrak{F}_3$.

Observación 5.4. Para las clases de grupos \mathfrak{N} y \mathfrak{U} se tiene que \mathfrak{N} no está contenida fuertemente en \mathfrak{U} a pesar que $\mathfrak{N} \subset \mathfrak{U}$. Analicemos el siguiente contraejemplo: El grupo $\text{Sym}(4)$ es soluble, ya que posee la serie normal

$$\{1\} \leq V_4 = \langle (12)(34), (13)(24) \rangle \leq \text{Alt}(4) \leq \text{Sym}(4)$$

cuyos factores son abelianos. Se tiene además que $\text{Sym}(4)$ no es supersoluble ya que el factor $V_4/\{1\} \cong V_4$ no es cíclico, en forma similar se tiene que $\text{Alt}(4)$ no es supersoluble. Dado que todo grupo nilpotente es supersoluble, se tiene que $\text{Alt}(4)$ y $\text{Sym}(4)$ no son nilpotentes.

El grupo diedrico de orden 8, denotado con D_4 , es un \mathfrak{N} -proyector de $\text{Sym}(4)$. Si D_4 fuese un \mathfrak{U} -proyector de $\text{Sym}(4)$ se tendría que $|\text{Sym}(4)/D_4| \notin \mathbb{P}$ lo cual es contradictorio. Los \mathfrak{U} -proyectores de $\text{Sym}(4)$ son isomorfos a $\text{Sym}(3)$. Así que no es posible que un \mathfrak{N} -proyector de $\text{Sym}(4)$ esté contenido como subgrupos en algún \mathfrak{U} -proyector de $\text{Sym}(4)$, con lo cual se prueba que la clase de grupos \mathfrak{N} no esta contenida fuertemente en la clase de grupos \mathfrak{U} .

Lema 5.5 ([3], II (2,2)). Sean $\mathfrak{F} = LF(F)$ y $\mathfrak{H} = LF(H^*)$, donde F y H^* son las definiciones canónicas locales de \mathfrak{F} y \mathfrak{H} respectivamente. Entonces $\mathfrak{F} \ll \mathfrak{H}$ si y sólo si para cada $H \in \mathfrak{H}$ un \mathfrak{F} -proyector E de H satisface

$$H^{H^*(p)} \leq E^{F(p)},$$

para cada $p \in \text{Char}(\mathfrak{F})$.

Lema 5.6 ([4], VII (5,3)). Sea $LF(F) = \mathfrak{F} \ll \mathfrak{H} = LF(H^*)$ con F y H^* definiciones canónicas locales de \mathfrak{F} y \mathfrak{H} respectivamente. Sea $\pi \subseteq \mathbb{P}$, si \mathfrak{X} es una formación tal que $F(p) = \mathfrak{S}_p \mathfrak{X}$ para todo $p \in \pi$, entonces existe una formación \mathfrak{D} tal que $H^*(p) = \mathfrak{S}_p \mathfrak{D}$ para todo $p \in \pi$.

Teorema 5.7. Sean \mathfrak{F} y \mathfrak{H} formaciones saturadas con $\mathfrak{F} \ll \mathfrak{H}$. Si $\mathfrak{F} = \mathfrak{N}\mathfrak{X}$ para alguna formación \mathfrak{X} , entonces \mathfrak{H} tiene la forma $\mathfrak{H} = \mathfrak{N}\mathfrak{D}$ para alguna formación \mathfrak{D} .

Demostración. Sean F y H^* las definiciones canónicas locales de \mathfrak{F} y \mathfrak{H} respectivamente, cuyas existencias están garantizadas por los Teoremas 4.25 y 4.26. Sea $\mathfrak{F} = \mathfrak{N}\mathfrak{X}$ para alguna formación \mathfrak{X} . Si definimos la función formación $f(p) = \mathfrak{X}$ para

todo primo p , se tiene por el Ejemplo 4.22 que $LF(f) = \mathfrak{N}\mathfrak{X} = \mathfrak{F}$. Luego por la Proposición 4.27 se cumple que

$$F(p) = \mathfrak{S}_p f(p) \cap \mathfrak{F} = \mathfrak{S}_p \mathfrak{X} \cap \mathfrak{N}\mathfrak{X} = \mathfrak{S}_p(\mathfrak{X} \cap \mathfrak{N}\mathfrak{X}) = \mathfrak{S}_p \mathfrak{X}$$

para todo $p \in \mathbb{P}$. Por el Lema 5.6 existe una formación \mathfrak{D} tal que $H^*(p) = \mathfrak{S}_p \mathfrak{D}$ para todo primo p . Como H^* es una función formación completa se tiene que $H^*(p) = \mathfrak{S}_p H^*(p)$ para todo $p \in \mathbb{P}$, en consecuencia se cumple que $\mathfrak{S}_p \mathfrak{D} = \mathfrak{S}_p H^*(p)$ para todo $p \in \mathbb{P}$, lo cual implica que $H^*(p) = \mathfrak{D}$ para todo primo p . Luego por el Ejemplo 4.22 se tiene que $LF(H^*) = \mathfrak{H} = \mathfrak{N}\mathfrak{D}$. \square

Lema 5.8 ([4], IV (5,20)). *Sea F la definición canónica local de una formación saturada $\mathfrak{F} = LF(F)$, y sea $\pi \subseteq \text{Char}(\mathfrak{F})$ un conjunto de primos. Cualquiera de las siguientes afirmaciones son equivalentes:*

- (a) $\mathfrak{S}_\pi \ll \mathfrak{F}$ (esto es, en cada grupo G los \mathfrak{F} -proyectores tienen índice π');
- (b) $F(q) \subseteq F(p)$ para todo $p \in \pi$ y todo $q \in \mathbb{P}$;
- (c) $F(p) = \mathfrak{F}$ para todo $p \in \pi$;
- (d) $f^*(p) = \mathfrak{S}$ para todo $p \in \pi$;
- (e) $\mathfrak{S}_\pi \mathfrak{F} = \mathfrak{F}$.

6 Una variación del contenido fuerte entre formaciones locales

Definición 6.1. Sean \mathfrak{F} y $\mathfrak{H} = LF(H^*)$ dos formaciones saturadas con H^* una definición canónica local de \mathfrak{H} y sea $\text{Char}(\mathfrak{F}) \subseteq \text{Char}(\mathfrak{H})$. Diremos que \mathfrak{F} está contenido G -fuertemente en \mathfrak{H} , que denotaremos $\mathfrak{F} \ll_G \mathfrak{H}$, si para cada elemento $H \in \mathfrak{H}$ un \mathfrak{F} -proyector E de H satisface que $H^{H^*(p)} \leq E$ para cada $p \in \text{Char}(\mathfrak{F})$.

Observación 6.2. Sean \mathfrak{F} y $\mathfrak{H} = LF(H^*)$ dos formaciones saturadas.

- (a) Si $\mathfrak{F} \ll \mathfrak{H}$ entonces se tiene por el Lema 5.5 que para cada $H \in \mathfrak{H}$ un \mathfrak{F} -proyector E de H satisface $H^{H^*(p)} \leq E^{F(p)} \leq E$ para cada $p \in \text{Char}(\mathfrak{F})$. Por lo tanto se sigue que $\mathfrak{F} \ll_G \mathfrak{H}$.
- (b) Si $\mathfrak{H} \subseteq \mathfrak{F}$, entonces dado $H \in \mathfrak{H}$ se tiene que $H \in \mathfrak{F}$ y por lo tanto $H \in \text{Proj}_{\mathfrak{F}}(H)$. Además, para todo $r \in \text{Char}(\mathfrak{F})$ es claro que $H^{H^*(r)} \leq H$. En consecuencia $\mathfrak{F} \ll_G \mathfrak{H}$. Como cualquier grupo supersoluble es soluble, se tiene que

$\mathfrak{S} \ll_G \mathfrak{U}$, sin embargo \mathfrak{S} no está contenido fuertemente en \mathfrak{U} , ya que por ejemplo, $\text{Sym}(4) \in \mathfrak{S}$ y $\text{Sym}(4) \notin \mathfrak{U}$. Así que $\text{Sym}(4) \in \text{Proj}_{\mathfrak{S}}(\text{Sym}(4))$, y como $\mathfrak{U} = S\mathfrak{U}$, se tiene que no existe un subgrupo $H \in \mathfrak{U}$ con $H \in \text{Proj}_{\mathfrak{U}}(\text{Sym}(4))$ tal que $\text{Sym}(4) \subseteq H$.

Lema 6.3. *Sea \mathfrak{F} una formación saturada con $\mathfrak{F} = s\mathfrak{F}$, esto es, \mathfrak{F} es cerrada bajo subgrupos. Sea $\mathfrak{H} = LF(H^*)$ con H^* la definición canónica local de \mathfrak{H} , y supongamos que*

$$\varrho := \text{Char}(\mathfrak{F}) \subseteq \text{Char}(\mathfrak{H}).$$

Si $\mathfrak{F} \ll_G \mathfrak{H}$, entonces

$$\mathfrak{H} \subseteq \bigcap_{r \in \varrho} \mathfrak{F}H^*(r).$$

Demostración. Sea $H \in \mathfrak{H}$. Entonces un \mathfrak{F} -proyector E de H satisface que $H^{H^*(r)} \leq E$ para todo $r \in \varrho$. Por la definición de \mathfrak{F} -proyector se tiene que $E \in \mathfrak{F} = s\mathfrak{F}$, luego se cumple que $H^{H^*(r)} \in \mathfrak{F}$ para todo $r \in \varrho$. Entonces se tiene que $H \in \mathfrak{F}H^*(r)$ para todo $r \in \varrho$. Con lo cual se prueba que $\mathfrak{H} \subseteq \bigcap_{r \in \varrho} \mathfrak{F}H^*(r)$. \square

Lema 6.4. *Sea $\mathfrak{H} = LF(H^*)$ con $\pi \subseteq \text{Char}(\mathfrak{H})$. $\mathfrak{S}_\pi \ll_G \mathfrak{H}$ si y sólo si $\mathfrak{H} \subseteq \bigcap_{r \in \pi} \mathfrak{S}_\pi H^*(r)$.*

Demostración. Supongamos que $\mathfrak{H} \subseteq \bigcap_{r \in \pi} \mathfrak{S}_\pi H^*(r)$. Sea $H \in \mathfrak{H}$, entonces por hipótesis $H \in \mathfrak{S}_\pi H^*(r)$ para todo $r \in \pi$. En consecuencia para todo $r \in \pi$, existe $N_r \trianglelefteq H$ con $N_r \in \mathfrak{S}_\pi$ tal que $H/N_r \in H^*(r)$. Luego, de la definición de $H^*(r)$ -residuo de H ($H^{H^*(r)}$) se tiene que $H^{H^*(r)} \leq N_r$ para todo $r \in \pi$ y en consecuencia $H^{H^*(r)}$ es un π -grupo para todo $r \in \pi$. Sea $E \in \text{Hall}_\pi(H)$. Entonces E es un π -subgrupo máximo de H , y dado que $H^{H^*(r)} \leq H$ se cumple que $H^{H^*(r)} \leq E$ para todo $r \in \pi$. Por lo tanto $\mathfrak{S}_\pi \ll_G \mathfrak{H}$. El recíproco es inmediato, se sigue del Lema anterior. \square

Observación 6.5. *Sea $\mathfrak{H} = LF(H^*)$ con $\pi \subseteq \text{Char}(\mathfrak{H})$. Definamos $\mathfrak{X} := \bigcap_{r \in \pi} H^*(r)$. Entonces*

$$\bigcap_{r \in \pi} \mathfrak{S}_\pi H^*(r) = \mathfrak{S}_\pi \mathfrak{X}.$$

Veamos:

- (a) Es claro que $\mathfrak{X} \subseteq H^*(r)$ para todo $r \in \pi$. Entonces $\mathfrak{S}_\pi \mathfrak{X} \subseteq \mathfrak{S}_\pi H^*(r)$ para todo $r \in \pi$. De lo cual se sigue que $\mathfrak{S}_\pi \mathfrak{X} \subseteq \bigcap_{r \in \pi} \mathfrak{S}_\pi H^*(r)$.
- (b) Sea $H \in \bigcap_{r \in \pi} \mathfrak{S}_\pi H^*(r)$. Entonces $H^{H^*(r)}$ es un π -subgrupo normal de H , para todo $r \in \pi$, y por la definición de $O_\pi(H)$ se tiene que $H^{H^*(r)} \subseteq O_\pi(H)$ para todo $r \in \pi$. Por lo tanto $H/O_\pi(H) \in H^*(r)$ para todo $r \in \pi$, así que $H/O_\pi(H) \in \mathfrak{X}$. Esto demuestra que $H \in \mathfrak{S}_\pi \mathfrak{X}$.

Teorema 6.6. *Sea \mathfrak{H} una formación saturada y $\pi \subseteq \text{Char}(\mathfrak{H})$. $\mathfrak{S}_\pi \ll_G \mathfrak{H}$ si y sólo si existe una formación \mathfrak{X} tal que $\mathfrak{H} \subseteq \mathfrak{S}_\pi \mathfrak{X}$.*

Demostración. La prueba se sigue del Lema 6.4 y la Observación 6.5. \square

El siguiente teorema prueba que en el caso $\pi = \{p\}$ los conceptos de contenido fuerte y G -contenido fuerte son equivalentes, si p pertenece a la característica de \mathfrak{H} .

Teorema 6.7. *Sea $\mathfrak{H} = LF(H^*)$ una formación saturada con H^* una definición canónica local de \mathfrak{H} y $p \in \text{Char}(\mathfrak{H})$. Entonces dos cualesquiera de las siguientes afirmaciones son equivalentes:*

- (a) $\mathfrak{S}_p \ll \mathfrak{H}$;
- (b) $H^*(q) \subseteq H^*(p)$ para todo $q \in \mathbb{P}$;
- (c) $H^*(p) = \mathfrak{H}$;
- (d) $h^*(p) = \mathfrak{S}$;
- (e) $\mathfrak{S}_p \mathfrak{H} = \mathfrak{H}$;
- (f) $\mathfrak{S}_p \ll_G \mathfrak{H}$.

Demostración.

$$\begin{aligned} \mathfrak{S}_p \ll_G \mathfrak{H} &\Leftrightarrow \mathfrak{H} \subseteq \mathfrak{S}_p H^*(p) && \text{Por Lema 6.4} \\ &\Leftrightarrow \mathfrak{H} \subseteq H^*(p) \\ &\Leftrightarrow \mathfrak{H} = H^*(p) \\ &\Leftrightarrow \mathfrak{S}_p \ll \mathfrak{H} && \text{Por Lema 5.8} \end{aligned}$$

El Lema 5.8 garantiza la equivalencia entre (a), (b), (c), (d) y (e). \square

Teorema 6.8. *Sea \mathfrak{X} una formación y \mathfrak{N} la clase de los grupos nilpotentes. Si \mathfrak{H} es una formación saturada tal que $LF(H^*) = \mathfrak{H} \subseteq \mathfrak{X}$ con H^* la definición canónica local de \mathfrak{H} , entonces $\mathfrak{N}\mathfrak{X} \ll_G \mathfrak{H}$.*

Demostración. Definamos la función formación $f(p) = \mathfrak{X}$, para todo $p \in \mathbb{P}$. Entonces se tiene que $LF(f) = \mathfrak{N}\mathfrak{X}$ y además del teorema de Gaschütz - Lubeseder cumple que $LF(f)$ es una formación saturada. Como $\mathfrak{H} \subseteq \mathfrak{X}$ se sigue de forma inmediata que $\mathfrak{H} \subseteq \mathfrak{N}\mathfrak{X}$. Ahora, dado $H \in \mathfrak{H}$, tenemos que $H \in \mathfrak{N}\mathfrak{X}$ y por lo tanto $H \in \text{Proj}_{\mathfrak{N}\mathfrak{X}}(H)$. Además para todo $r \in \text{Char}(\mathfrak{N}\mathfrak{X})$ se tiene que $H^{H^*(r)} \leq H$. En consecuencia $\mathfrak{N}\mathfrak{X} \ll_G \mathfrak{H}$. \square

Observación 6.9. El recíproco del Teorema anterior en general no se cumple ya que por ejemplo la clase de grupos $\mathfrak{U} \subseteq \mathfrak{N}^2 = \mathfrak{N}\mathfrak{N}$ y en consecuencia $\mathfrak{N}\mathfrak{N} \ll_G \mathfrak{U}$ pero $\mathfrak{U} \not\subseteq \mathfrak{N}$.

Corolario 6.10. $\mathfrak{N}\mathfrak{X} \ll_G (1)$ para cualquier formación \mathfrak{X} .

Demostración. La prueba se sigue en forma inmediata del Teorema anterior teniendo en cuenta que la clase de grupos (1) esta contenida en cualquier formación \mathfrak{X} . □

Bibliografía

- [1] Adolfo Ballester and Luis M. Ezquerro. *Classes of Finite Groups*, Springer, 2006.
- [2] Edward Cline. On an embedding property of generalized Carter subgroups, *Pacific J. Math.* 29, 491-519, 1969
- [3] Patrick D' Arcy. On strong containment of locally defined formations, *Journal of Algebra* 28, 362-373, 1974.
- [4] Klaus Doerk and Trevor Hawkes. *Finite Soluble Groups*. W de G. 1992.
- [5] Derek J. S. Robinson. *A course in the theory of groups*. Second edition, Springer 1996.
- [6] Wolfgang Gaschütz. Zur Theorie der endlichen auflösbaren Gruppen, *Math. Z.* 80, 300-305, 1963.
- [7] Wolfgang Gaschütz and Ursula Lubeseder. Kennzeichnung gesättigter Formationen, *Mathematische Zeitschrift*, Volume: 82, page 198-199, 1963.
- [8] Ismael Gutiérrez. Zur starken Enthaltenseinsrelation für gesättigte Formationen, *Dissertation*, Universität-Mainz, 2002.
- [9] Ismael Gutiérrez. The G-strong containment for locally defined formations of finite soluble groups. *Journal of Algebra*, 2006.
- [10] Hall, Marshall. *Teoria de los grupos*, Trillas, 1979.
- [11] Bertram Huppert. *Endliche Gruppen I. Grundlehren der mathematischen Wissenschaften*. Band 134. Springer- Verlag, 1967.

- [12] Peter Schmid. Every saturated formation is a local formation. *Journal of Algebra* 51, 144 - 148, 1978.
- [13] Guo Wenbin. *The theory of classes of groups*, Science press, Beijing, P.R. China, 1997.

Departamento de Matemáticas y Estadística, Universidad del Norte,
Km 5 vía a Puerto Colombia,
Barranquilla Atlántico, Colombia.

`isgutier@uninorte.edu.co`

`torresbadillo15@hotmail.com`

Análisis

Capítulo 4

Classification of the continuous translations in spaces of weighted integrable real functions

Álvaro Hernández Cervantes¹, Miguel Antonio Jiménez Pozo^{2,*}
¹CIMAT. ²FCFM, BUAP

Abstract

We classify the usual weights $\omega(x)$, $x \geq 1$, for which the positive translations $t \rightsquigarrow f(x+t)$ are well defined and are continuous for every $f \in L^1(\omega(x)dx)$.

Keywords: Bernstein problem, Hölder space, Plessner problem, weighted integration.

1 Introduction

To state the problem we are dealing with, consider a Banach space E of functions f defined on a topological semigroup G such as \mathbb{R}^k , the circle \mathbb{T} , \mathbb{T}^k , the additive semigroup \mathbb{R}_+ of non negative real numbers, or the multiplicative strictly positive real numbers $\mathbb{R}_+^* = \mathbb{R}_+ \setminus \{0\}$. The translation (sometimes also called shift) operator Λ_t , $t \in G$, is defined by

$$\begin{aligned}\Lambda_t : E &\longrightarrow E \\ f &\rightsquigarrow f_t(x) := f(x+t),\end{aligned}$$

whenever $f_t \in E$.

In dealing with many function spaces such as L^p , $1 \leq p < \infty$, \mathcal{C} , Lip , and so on, a fundamental property required for developing the usual mathematic works is the continuity of this translation operator. However it could be not only that $f_t \notin E$, but that $\|f_t - f\|_E \not\rightarrow 0$ even in the case that Λ_t is well defined for every $t \in G$ and a fixed $f \in E$.

Thus, a better reformulation of the problem in hands is to identify the $f \in E$ for which the mapping $t \rightsquigarrow \Lambda_t f$ is well defined and continuous. Or equivalently, to find the largest Λ -invariant subspace $F \subset E$ on which the operator $t \rightsquigarrow \Lambda_t$ is continuous. Of course, not only F is a subspace but also closed in E as well.

*To whom any communication must be addressed

Historically, the first remarkable paper on this subject is the one by Plessner [12] in 1929. He proved that F is the space of absolutely continuous functions when E is the space of 2π -periodic functions of bounded variation that vanish at a fixed point, and the norm of a function is its total variation.

The Plessner paper generated a long series of different researches (Bochner [2], Hardy-Littlewood [8], Shilov [15], Mirkil [11], and others). We shall resume several of them in section 2, to explain the connexion of Plessner problems with Hölder spaces, although the Plessner problem has its own interest in an independent way. Section 3 shows several general results on the continuity of the translation operators to be used in section 4, that presents the answer to the Plessner problem in $L^1(\omega)$ when ω represents different weight functions on \mathbb{R}_+ . To complete the information given in section 2 on the problem of polynomial approximation in Hölder spaces, the reader may see the general papers [3], [4], [5] and [6]. Further it could be wishfully to extend these results on polynomial approximation to Hölder spaces of functions on unbounded intervals where Bernstein weights appear from ([1], [9], [13], [14]). Then results of section 4 need to be applied when ω is a Bernstein weight. We shall summarize this point at the end of the paper.

2 Hölder spaces and the Plessner problem

Many of the results quoted here for 2π -periodic functions can be found in Mirkil [11].

There are different ways of associating Lipschitz spaces to $L^p(\mathbb{T})$, $1 \leq p \leq \infty$. Here we follow (although not always) the tradition of using $p = \infty$, to represent the continuous functions with sup-norm. To this aim the translation given by $f_t(x) = f(x + t)$, $f \in L^p(\mathbb{T})$, $x, t \in \mathbb{R}$ is needed. Then $Lip_1^p(\mathbb{T})$ is conformed with the functions $f \in L^p(\mathbb{T})$, for which

$$\theta_1(f)_p = \sup\{\|f_t - f\|_p/|t|, t \neq 0\} < \infty. \quad (1)$$

In the case $L^p([a, b])$, we have a similar definition with the agreement $f_t(x) = f(x)$, if $x + t \notin [a, b]$. So let X be \mathbb{T} or $[a, b]$.

The functional $\theta_1(\circ)_p$ is a semi-norm. Thus $Lip_1^p(X)$ is a normed space (actually Banach) with the norm

$$\|f\|_{Lip_1^p(X)} = \|f\|_p + \theta_1(f)_p. \quad (2)$$

Now fix $0 < \alpha \leq 1$, a semi-norm $\theta_\alpha(f)_p$ can be defined by (1) with the change of $|t|$ by $|t|^\alpha$. This allows the introduction of the so called *Hölder spaces* $Lip_\alpha^p(X)$, normed by (2) with $\theta_\alpha(f)_p$ instead of $\theta_1(f)_p$. Since definitions of Lipschitz and

Hölder norms coalesce for $\alpha = 1$, both names are used indistinctly to denominate the spaces $Lip_\alpha^p(X)$, for any $\alpha > 0$. Since these Banach spaces $Lip_\alpha^p(X)$ are not separable, one cannot approximate all their functions by polynomials. It happens this is very closely related to the Plessner problem.

In fact, let us take $X = \mathbb{T}$ for simplicity. The translation

$$t \rightsquigarrow \Lambda_t(f)$$

for each f fixed, is well defined in $Lip_\alpha^p(\mathbb{T})$, $0 < \alpha \leq 1$, $1 \leq p \leq \infty$, but it is not continuous.

The theory was developed as follows. For any $0 < \alpha < 1$, $f \in Lip_\alpha^p(\mathbb{T})$, $\delta > 0$, set

$$\theta_\alpha(f, \delta)_p = \sup\left\{\frac{\|ft - f\|_p}{|t|^\alpha} : 0 < |t| \leq \delta\right\}.$$

Of course,

$$\theta_\alpha(f)_p = \sup\{\theta_\alpha(f, \delta)_p : \delta > 0\}.$$

Then for $0 < \alpha < 1$, introduce the linear subspace $lip_\alpha^p(\mathbb{T})$ of $Lip_\alpha^p(\mathbb{T})$ conformed by the functions $f \in Lip_\alpha^p(\mathbb{T})$ such that

$$\lim_{\delta \rightarrow 0} \theta_\alpha(f, \delta)_p = 0.$$

If we extend this definition to $\alpha = 1$, we get that $lip_1^p(\mathbb{T})$, $1 \leq p \leq \infty$, only contains the (a.e) constant functions. So nothing is new in this case and another definition for $\alpha = 1$ needs to be considered (we skip this situation). For the moment being define the linear spaces F_α^p of functions $f \in Lip_\alpha^p(\mathbb{T})$, $1 \leq p \leq \infty$, $0 < \alpha \leq 1$, for which

$$\lim_{t \rightarrow 0} \|ft - f\|_{Lip_\alpha^p(\mathbb{T})} = 0.$$

From Hardy and Littlewood [8], F_1^1 is the space of functions of bounded variation on \mathbb{T} , $F_1^\infty = D_1^\infty(\mathbb{T})$ the space of differentiable functions, and for $1 < p < \infty$, F_1^p is the space of antiderivatives of the functions in $L^p(\mathbb{T})$. In [15], Shilov proved that for any $1 \leq p \leq \infty$, $0 < \alpha \leq 1$, F_α^p is exactly the closure in the norm of $Lip_\alpha^p(\mathbb{T})$ of all trigonometric polynomials. Finally Mirkil [11] closed the circle of ideas by proving that for any $0 < \alpha < 1$ and $1 \leq p \leq \infty$, $F_\alpha^p = lip_\alpha^p(\mathbb{T})$. Thus for $0 < \alpha < 1$, the set of trigonometric polynomials is dense in $lip_\alpha^p(\mathbb{T})$.

Now suppose we are studying similar problems on $X = \mathbb{R}$, or for simplicity $X = \mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Not all the starting points assumed by Mirkil hold in this case.

First, the non constant polynomials are not bounded on \mathbb{R} . Then, traditionally a weight function ω on \mathbb{R}_+ such that $\omega(x) \rightarrow 0$, when $x \rightarrow \infty$, is used to define $\mathcal{C}_B^\omega(\mathbb{R}_+)$ to be the functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ (or \mathbb{C}) continuous, such that

$$\|f\|_\omega = \sup_{x \in \mathbb{R}_+} |f(x)\omega(x)| < \infty.$$

Also define $\mathcal{C}_l^\omega(\mathbb{R}_+)$ as the linear subspace of $f \in \mathcal{C}_B^\omega(\mathbb{R}_+)$, for which $\lim_{x \rightarrow +\infty} f(x)$ exists, and $\mathcal{C}_o^\omega(\mathbb{R}_+)$ the subspace of $\mathcal{C}_l^\omega(\mathbb{R}_+)$, for which the limit is zero.

We can equally consider the function spaces $L^{p,\omega}(\mathbb{R}_+)$, $1 \leq p < \infty$, defined by $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ (or \mathbb{C}) such that

$$\|f\|_{p,\omega} = \left(\int_{\mathbb{R}_+} |f|^p \omega \right)^{\frac{1}{p}} < \infty.$$

Denoting $\mathcal{C}_B^\omega(\mathbb{R}_+) = L^{\infty,\omega}(\mathbb{R}_+)$ we unify the notation. In order to study polynomial approximation in these spaces, one needs that they contain all polynomials \mathbf{p} .

Any weight ω for which polynomials are dense in $L^{p,\omega}(\mathbb{R}_+)$, $1 \leq p \leq \infty$, is denominated a Bernstein weight for the corresponding p ([1], [9], [13], [14]).

If we want to introduce a Hölder subspace in $L^{p,\omega}$, we need first of all to know those $L^{p,\omega}$, for which the translation operator is well defined and is continuous. This situation is different to the ones considered by Mirkil [11]. In fact, when X was finite, the translation was well defined in all the spaces $L^p(X)$, $1 \leq p \leq \infty$. But the general situation is different as we show with the following example.

Suppose for simplicity of calculation that $X = [x \mid x \geq 1] \subset \mathbb{R}_+$ and $\omega(x) = e^{-x^2}$. Of course for any algebraic polynomial \mathbf{p} ,

$$\|\mathbf{p}\omega\|_{L^1} = \int_1^\infty |\mathbf{p}(x)|e^{-x^2} dx < \infty.$$

But taking $f(x) = \frac{e^{x^2}}{x^2}$, while

$$\|f\omega\|_{1,\omega} = \int_1^\infty f(x)\omega(x)dx = \int_1^\infty \frac{dx}{x^2} < \infty,$$

we observe that for $t > 0$,

$$\|f_t\omega\|_{1,\omega} = \int_1^\infty \frac{e^{(x+t)^2}}{(x+t)^2} e^{-x^2} dx \geq \int_1^\infty \frac{e^{2xt}}{(x+t)^2} dx = \infty.$$

Then the translation operator is not well defined in $L^{1,e^{-x^2}}(\mathbb{R}_+)$.

3 Some general results on the continuity of the translation operator in $L^1(\omega)$

Many general results on the properties of the translation operators have been deeply studied and are well known. For instance, in the already quoted Mirkil paper [11] and the Butzer-Berens book [6], among many others. So we shall present here only the general result given by theorem 3.3, that we need for the particular cases treated in section 4.

The following results can be extended to more general topological groups and semi-groups. Let (G, \circ) be one of the classical groups $(\mathbb{R}^k, +)$, $(\mathbb{T}^k, +)$ or (\mathbb{R}_+^*, \cdot) , and μ the corresponding Haar measure. For any fixed $t \in G$, the translation operator

$$\begin{aligned} \Lambda_t : L^1(G) &\longrightarrow L^1(G) \\ f &\rightsquigarrow f_t(x) := f(x \circ t), \end{aligned}$$

is a linear isometry. Moreover, the function

$$\begin{aligned} G &\rightarrow L^1(G) \\ t &\rightsquigarrow f_t \end{aligned} \tag{3}$$

is continuous for any fixed $f \in L^1(G)$.

This result on continuity in (3) is a consequence of the good topological properties of G and the invariant translation property of the Haar measure μ . The new question concerns to know whether it still holds in some form for $L^1(\omega)$, where ω is a real valued weight function on G . Since the question involves continuity properties, it is reasonable to suppose from the very early that ω is a continuous function.

Remember $L^1(\omega)$ is the Banach space of μ -measurable complex valued functions f on G , such that $f\omega \in L^1(G)$, with the norm

$$\|f\|_\omega = \int_G |f(x)|\omega(x)d\mu(x).$$

Following the integration rules we must consider that two functions $f, g \in L^1(\omega)$ are equal (in the same equivalence class) if and only if

$$\mu[f\omega \neq g\omega] = 0. \tag{4}$$

But it may be that there exist μ -measurable functions $f, g : G \rightarrow \mathbb{C}$ that satisfy (4), while $\mu[f \neq g] > 0$. On the other hand, if $\omega(x) = 0$ at some $x_0 \in G$, it is easy to construct $f \in L^1(\omega)$, such that $f_t \notin L^1(\omega)$ for some $t \in G$. Both of these problems disappear as soon as we assume by hypothesis that $\omega(x) > 0$ at any $x \in G$. Of

course it may occur that $\omega(x) \rightarrow 0$ as $x \rightarrow \infty$, when $G = \mathbb{R}^k$ or \mathbb{R}_+^* , for instance, but this is only an apparent problem because $\infty \notin G$.

Organizing ideas, first we shall restrict ourselves to the particular groups $\mathbb{R}^k, \mathbb{T}^k$ and \mathbb{R}_+^* , and strictly positive continuous weight functions ω . In this situation $L^1(\omega)$ is exactly $L^1(\omega d\mu)$, where μ is the Lebesgue measure $dx = d(x_1, x_2, \dots, x_k)$ when $G = \mathbb{R}^k$, $dx/(2\pi)^k = d(x_1, x_2, \dots, x_k)/(2\pi)^k$ when $G = \mathbb{T}^k$, or dx/x when G is the multiplicative group \mathbb{R}_+^* .

As the considered measure ωdu is regular, by Luzin theorem the space $C_K(G)$ of continuous complex valued functions on G with compact support is a dense subspace of $L^1(\omega)$.

Lemma 3.1. *Let $u \in C_K(G)$, then*

$$\begin{aligned} G &\rightarrow L^1(G) \\ t &\rightsquigarrow u_t, \end{aligned}$$

is a well defined and continuous map.

Proof. Any continuous function of compact support is uniformly continuous. □

Lemma 3.2. *Let $u \in C_K(G)$, then*

$$\begin{aligned} G &\rightarrow L^1(\omega) \\ t &\rightsquigarrow u_t, \end{aligned}$$

is a well defined and continuous map.

Proof. Let $t_0 \in G$ and $\epsilon > 0$ be given. Using Lemma 1, we get any open neighborhood V of t_0 whose closure \bar{V} is compact, and such that

$$\sup_{t \in V} \|u_t - u_{t_0}\|_\infty < \epsilon.$$

Let K be the compact support of u , and

$$K' = \{x \circ t : x \in K, t \in \bar{V}\}.$$

Since K' is compact, the bound

$$M = \sup_{x \in K'} |u_t - u_{t_0}|,$$

is finite. Then for any $t \in V$,

$$\int_G |u_t - u_{t_0}| \omega d\mu \leq M \mu(K') \epsilon. \quad \square$$

Theorem 3.3. *Suppose for any $t \in G$ the linear operators*

$$\begin{aligned} \Lambda_t : L^1(\omega) &\rightarrow L^1(\omega) \\ f &\rightsquigarrow f_t \end{aligned}$$

are well defined and bounded. Then for each $f \in L^1(\omega)$, the map

$$\begin{aligned} G &\rightarrow L^1(\omega) \\ t &\rightsquigarrow f_t \end{aligned}$$

is continuous.

Proof. Since $L^1(\omega)$ is a Banach space, the “uniform bound” theorem applies. Then there exists a number $M > 0$ such that

$$\sup_{t \in G} \|\Lambda_t\| \leq M.$$

Now let $f \in L^1(\omega)$ and $\epsilon > 0$ be given. Choose $u \in C_K(G)$ such that

$$\|f - u\|_\omega \leq \epsilon.$$

Fix any $t_0 \in G$. Using Lemma 3.2, get a neighborhood V of t_0 , such that

$$\forall t \in V, \|u_t - u_{t_0}\|_\omega < \epsilon.$$

Then for every $t \in V$, we have

$$\begin{aligned} \|f_t - f_{t_0}\|_\omega &\leq \|f_t - u_t\|_\omega + \|u_t - u_{t_0}\|_\omega + \|u_{t_0} - f_{t_0}\|_\omega \\ &\leq 2M\|f - u\|_\omega + \|u_t - u_{t_0}\|_\omega \\ &\leq (2M + 1)\epsilon. \end{aligned} \quad \square$$

4 On the continuity of the translation operator for particular given weights on \mathbb{R}_+

The general results of section 3 can be easily adapted for the semigroup \mathbb{R}_+ . But the given example of the function $f(x) = \frac{e^{-x^2}}{1+x^2} \in L^1(e^{-x^2})$, where the positive translation f_t is not well defined in this space for any $t > 0$, shows it is worthwhile to classify the main weights ω for which the translation operators are well defined and continuous. This is the goal of this section.

Usually the examination of weighted spaces of functions in \mathbb{R} can be reduced to \mathbb{R}_+ and positive parameters t for the translation. Moreover, since $f_{t_0+t} = (f_{t_0})_t$ and

the analysis of the continuity property can be reduced to $t_0 = 0$, we shall suppose that $t \in [0, 1]$, what is another convenient hypothesis.

Finally, since $0 \in \mathbb{R}_+$, several weights as $\frac{1}{x^p}$, $p > 0$, must be written as $\frac{1}{1+x^p}$. A simplification that facilitates the calculations but does not affect the study on hand is to consider $X = [1, \infty)$, instead of the whole semi-axis \mathbb{R}_+ .

Using the general theorem 3.3 of the last section, to find the weights ω for which the translation is well defined and continuous in $L^1(\omega)$, we only need to prove that Λ_t , is defined and uniformly bounded for $t \in [0, 1]$.

From now on $\omega \in \mathcal{C}(X)$ is a strictly positive weight. Since it is not possible to consider all kind of weights we choose now the main ones.

Theorem 4.1. *Let p a real number and $\omega(x) = x^p$. Then $L^1(\omega)$ is translation invariant and continuous in the parameter t for every fixed $f \in L^1(\omega)$.*

Proof. Let be $0 \leq t \leq 1$ and $f \in L^1(x^p)$ fixed. Then

$$\|f_t\|_\omega = \int_{1+t}^\infty |f(u)|(u-t)^p du \leq \|f\|_{L^1(\omega)},$$

whenever $p > 0$.

If $p = 0$, then $L^1(\omega)$ is the classical Lebesgue space $L^1(X)$.

Now suppose $p < 0$. Set $M > 0$ such that $M \geq (\frac{u-t}{u})^p$, for every $u \geq 1+t$, $0 < t \leq 1$. Then

$$\begin{aligned} \|f_t\|_\omega &= \int_1^\infty |f_t(x)|x^p dx = \int_{1+t}^\infty |f(u)|(u-t)^p du \\ &\leq M \int_{1+t}^\infty |f(u)|u^p du \\ &\leq M \|f\|_{L^1(\omega)}. \end{aligned} \quad \square$$

Of course, the case $p > 0$ is not interesting enough but it was included to give a complete description of the case $\omega(x) = x^p$, p real.

We need the following lemma.

Lemma 4.2. *Let $f \in L^1(\omega)$ and $2 < b < \infty$. If*

$$\int_b^\infty |f_t(x)|\omega(x)dx < \infty.$$

Then $f_t \in L^1(\omega)$.

Proof. We need to prove that

$$I = \int_1^b |f_t(x)|\omega(x)dx < \infty.$$

Since ω is continuous and strictly positive there exists a constant C , depending on b , such that $\omega(u - t) \leq C\omega(u)$. Thus

$$I_t = \int_{1+t}^{b+t} |f(u)|\omega(u - t)du \leq C \int_1^\infty |f(x)|\omega(x)dx = C\|f\|_{L^1(\omega)}. \quad \square$$

Suppose now $\omega(x) = a_0x^m + a_1x^{m-1} + a_2x^{m-2} + \dots + a_m$ is a strictly positive polynomial of real variables. Then there exists $b > 2$ such that for every $x \geq b$, $\omega(x) \leq 2a_0x^m$.

Consequently

$$\int_1^\infty |f(x)|\omega(x)dx \approx cte + 2a_0 \int_b^\infty |f(x)|x^m dx,$$

applying the lemma above $f \in L^1(\omega)$ if and only if $f \in L^1(x^m)$. In particular $f_t \in L^1(\omega)$ whenever f also belongs to this space. A similar argument can be used in the analysis of algebraic rational weights $\omega(x) = \frac{p(x)}{q(x)}$.

From $e^{x-t} = e^x \cdot e^{-t}$ easily follows directly the next result.

Let $\omega(x)$ be the exponential weight e^{-x} . Then $f_t \in L^1(e^{-x})$ for every $f \in L^1(e^{-x})$ and the mapping $t \rightsquigarrow f_t$ is continuous.

Now we wish to consider other exponential weights of type e^{-x^p} .

Theorem 4.3. *Let be $\omega(x) = e^{-x^p}$, $0 < p < 1$.*

If $f \in L^1(\omega)$ and $0 < t \leq 1$, then $f_t \in L^1(\omega)$ and $t \rightsquigarrow f_t$ is continuous.

Proof. Suppose

$$\|f\|_\omega = \int_1^\infty |f(x)|\omega(x)dx = \int_1^\infty |f(x)|e^{-x^p} dx < \infty.$$

Fix any $0 < t \leq 1$.

$$\int_1^\infty |f_t(x)|e^{-x^p} dx = \int_1^\infty |f(x+t)|\frac{dx}{e^{x^p}}.$$

Let $g(x)$ be such that $e^{x^p} \cdot g(x) = e^x$. Then

$$g(x) = e^{x \cdot (1-x^{p-1})}.$$

Since $0 < p < 1$, $g(x)$ is an increasing and strictly positive function on $X = [1, \infty)$. Thus the right hand integral above is written as

$$\int_1^\infty |f(x+t)| \frac{dx}{e^{xp}} = \int_1^\infty |f(x+t)|g(x) \frac{dx}{e^x}. \quad (5)$$

Since

$$\int_1^\infty |f(x)|g(x) \frac{dx}{e^x} = \int_1^\infty |f(x)| \frac{dx}{e^{xp}} < \infty,$$

then $f \cdot g \in L^1(e^{-x})$. This implies for $t > 0$, that $(f \cdot g)_t \in L^1(e^{-x})$, where

$$(f \cdot g)_t(x) = f_t(x) \cdot g_t(x) = f(x+t) \cdot g(x+t).$$

From the properties of g ,

$$|f(x+t)g(x+t)| \geq |f(x+t)|g(x) = |f_t(x)|g(x). \quad (6)$$

Using (5) and (6),

$$\begin{aligned} \int_1^\infty |f(x+t)|g(x) \frac{dx}{e^x} &\leq \int_1^\infty |f(x+t)|g(x+t) \frac{dx}{e^x} \\ &= \int_1^\infty |(f \cdot g)_t(x)| \frac{dx}{e^x} < \infty. \end{aligned}$$

Then $f_t \in L^1(\omega)$.

To complete the proof we use the monotone property of g to get that $\|\Lambda_t f\|_\omega$, $0 < t \leq 1$ is uniformly bounded. \square

We already know that translation fails for the weight e^{-x^2} . But we wish to analyse others values of p . Suppose $p > 1$ and set $\omega(x) = e^{-x^p}$ and $f(x) = \frac{e^{xp}}{x^2}$. Obviously $f \in L^1(\omega)$, because

$$\int_1^\infty |f(x)|\omega(x)dx = \int_1^\infty \frac{e^{xp}}{x^2} e^{-x^p} dx = \int_1^\infty \frac{dx}{x^2} < \infty.$$

However, we shall prove that $f_t \notin L^1(\omega)$, for any $t > 0$. To do that we consider the *Taylor Formula* of the function $g(u) = u^p$, $u \geq 1$, $p > 0$;

$$(x+t)^p = x^p + ptx^{p-1} + \frac{p(p-1)}{2} \xi^{p-2} t^2,$$

where $\xi \in (x, x+t)$. Then

$$(x+t)^p > x^p + ptx^{p-1}.$$

From this inequality follows the next one,

$$e^{(x+t)^p} > e^{x^p} \cdot e^{ptx^{p-1}}.$$

Now we can finish the proof of our claim,

$$\begin{aligned} \int_1^\infty |f_t(x)|\omega(x)dx &= \int_1^\infty \frac{e^{(x+t)^p}}{(x+t)^2} \frac{dx}{e^{x^p}} \\ &\geq \int_1^\infty \frac{e^{x^p} \cdot e^{ptx^{p-1}}}{(x+t)^2} \frac{dx}{e^{x^p}} \\ &= \int_1^\infty \frac{e^{ptx^{p-1}}}{(x+t)^2} dx = \infty. \end{aligned}$$

□

Main results of this section are included in the following scheme:

Weights	Exponential parameters	$f \in L^1(w) \Rightarrow f_t \in L^1(w)$	Continuity of the translation
$\omega(x) = x^p$	$p \in \mathbb{R}$	Yes	Yes
$\omega(x) = e^{-x^p}$	$0 < p \leq 1$	Yes	Yes
$\omega(x) = e^{-x^p}$	$1 < p$	No	-----

Final remark. *The asymptotic behaviour of a Bernstein weight in $L^1(\omega(x)dx)$, $x \geq 0$, must be “close” to $\omega(x) = e^{-x^p}$, with $p \geq 1$ (see [9], for instance). But the translation operator in $L^1(e^{-x^p})$ is only well defined for $p \leq 1$. This reduces any possible extension of results in section 2 with exponential weights to the case given by $L^1(e^{-x})$.*

Acknowledgment

This paper have be written under the support of the SEP-VIEP projects of BUAP, Puebla, Mexico.

Bibliography

[1] S. N. Bernstein, *Le probleme de l'approximation des fonctions continues sur tout l'axe reel et l'une de ses applications*, Bull. Math. Soc. France, 52 (1924), 399-410.

- [2] S. Bochner, *Additive set functions on groups*, Ann. Math., 40 (1939), 769-799.
- [3] J. Bustamante, *Problems on approximation in Hölder norms*, Tópicos de Teoría de la Aproximación II, M. A. Jiménez & J. Bustamante, eds, Textos Científicos, Puebla, Mexico., (2007), 9-48.
- [4] J. Bustamante and M. A. Jiménez-Pozo, *Trends in Hölder approximation*, Proceedings of the 5th International Conference on Approximation and Optimization in the Caribbean; M. Lassonde (Ed): Approximation, Optimization and Mathematical Economics; Physica-Verlag, (2001), 81-96.
- [5] J. Bustamante and M. A. Jiménez-Pozo, *Contributions to the study of polynomial approximation in Hölder spaces*, FCFM-BUAP, Matemáticas y sus Aplicaciones 3, Textos Científicos, Fomento Editorial de la Benemérita Universidad Autónoma de Puebla, ISBN: 978 607 487 659 8., Puebla, Mexico., (2013), 5-24.
- [6] P. L. Butzer and H. Berens, *Semi-Groups of Operators and Approximation*, Springer-Verlag, New York/Berlin, (1967).
- [7] N. Dunford, *On a Theorem of Plessner*, Brown University, (1935), 356-358.
- [8] G. H. Hardy and J. E. Littlewood, *Some properties of fractional integrals I*, Math. Zeit., 27 (1927-1928), 565-606.
- [9] D. S. Lubinsky, *A Survey of Weighted Polynomial Approximation with Exponential*, Surveys in Approximation Theory, ISSN 1555-578X. Volume 3, (2007), 1-105.
- [10] K. de Leeuw, *Linear spaces with a compact group of operator*, III. J. Math., 2 (1958), 367-377.
- [11] H. Mirkil, *Continuous Translation of Hölder and Lipschitz Functions*, Can. J. Math., 12 (1960), 674-685.
- [12] A. Plessner, *Eine Kennzeichnung der totalstetig Funktionen*, J. reine angew. Math., 160 (1929), 26-32.
- [13] H. Pollard, *Solution of Bernstein's approximation problem*, Proc. Amer. Math. Soc., 4 (1953), 869-875.
- [14] H. Pollard, *The Bernstein approximation problem*, Proc. Amer. Math. Soc., 6 (1955), 402-411.

- [15] E. Shilov, *Homogeneous rings of functions*, Uspehi Matem. Nauk N.S., 6 (1951), A.M.S. Translation 92.
- [16] A. Zygmund, *Smooth Functions*, Duke Math. J., 12 (1945), 47-76.

Centro de Investigación en Matemáticas, A. C.
Calle Jalisco s/n, Colonia Valenciana,
Guanajuato, Gto. C.P. 36240
alvaro.hernandez@cimat.mx

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570
mjimenez@fcfm.buap.mx

Capítulo 5

Aproximación tipo Korovkin y Sistemas de Chebyshev

José Margarito Hernández Morales, José Luis Carrasco Pacheco
Universidad Tecnológica de la Mixteca

Resumen

Este trabajo gira en torno a la relación entre los espacios de Chebyshev y los subespacios de Korovkin en el espacio de funciones continuas, mencionando y demostrando algunos de los teoremas más importantes de ambas teorías. Se muestra el desarrollo de dos teorías de aproximación diferentes. Señalamos que este es un capítulo de divulgación cuyo principal objetivo es resaltar la relación existente entre las dos teorías mencionadas.

1 Introducción

Algunas veces podemos aproximar funciones complicadas mediante una sucesión de funciones más simples (con las cuales es más fácil de trabajar en la solución de un determinado problema) que dan una solución satisfactoria en ciertas aplicaciones. Hoy en día existe bastante teoría sobre tipos de aproximación de funciones, algunos de estos son: aproximación lineal, aproximación no lineal, aproximación uniforme, aproximación de Chebyshev, aproximación tipo Korovkin, aproximación de Lipschitz, aproximación con peso, aproximación sensible al signo, etc.

El matemático ruso P. L. Chebyshev, en 1854, desarrolló los conceptos que sientan las bases de la Teoría de Aproximación, mediante el problema: dada una función continua f , encontrar un polinomio algebraico p de grado $\leq n$, tal que el máximo de su desviación con respecto a f sobre un intervalo dado, sea más pequeño que el de los otros polinomios con las mismas características. En el problema mencionado, los sistemas de Chebyshev juegan un papel muy importante, ya que tienen una caracterización interesante del polinomio de mejor aproximación. De hecho, un teorema clásico de Haar establece que toda función $f \in C(X)$ tiene un polinomio de mejor aproximación en el subespacio generado por $h_0, \dots, h_n \in C(X)$ y, además éste es único si y sólo si h_0, \dots, h_n es un sistema de Chebyshev de orden $n + 1$, [5].

Por otra parte, en 1953 Korovkin (otro matemático ruso) estableció un teorema que con el tiempo se haría muy célebre. Por su simplicidad y al mismo tiempo su poder han despertado el interés de muchos matemáticos. Se trata de

un criterio que permite decidir cuándo una sucesión de operadores lineales positivos $K_n : C([0, 1]) \rightarrow C([0, 1])$ converge uniformemente al operador identidad $Id : C([0, 1]) \rightarrow C([0, 1])$. Korovkin estableció que basta con verificar la convergencia uniforme para $f \in \{1, x, x^2\}$, este último es llamado un subconjunto de Korovkin. Después se han encontrado otros subconjuntos con las mismas propiedades que $\{1, x, x^2\}$ y se les ha llamado también subconjuntos de Korovkin, mientras que a los espacios generados por estos conjuntos se les llama subespacios de Korovkin, [1].

En tiempos recientes se ha encontrado una relación muy interesante entre los subespacios de Chebyshev (espacios generados por un sistema de Chebyshev) y los subespacios de Korovkin, [1].

2 Sistemas de Chebyshev

En todo el capítulo supondremos que X es un espacio de Banach con escalares reales o complejos, a menos que se indique lo contrario. Sea Y un subespacio vectorial de X . Si $f \in X$, el error de aproximación $E(f)$ de f por elementos de Y se define por

$$E(f) := E(f, Y)_X := \inf_{P \in Y} \|f - P\|.$$

En particular si $\{x_1, \dots, x_n\} \subset X$ se tiene el subespacio generado

$$Y = \text{gen}\{x_1, \dots, x_n\}$$

y, a la combinación lineal $y = \sum_{i=1}^n \lambda_i x_i$ se le llama polinomio de grado n . El grado de aproximación (o error de grado n) E_n de f por polinomios de Y está dado por $E_n(f) = \inf_{P \in Y} \|f - P\|$. Si existe un polinomio $P_0 \in Y$ para el cual $E(f) = \|f - P_0\|$, entonces decimos que P_0 es un elemento de mejor aproximación de f en Y . La función error E es una función que se comporta muy bien, pues es una función continua, más aún es sublineal.

Teorema 2.1. *Si X es un espacio y $E : X \rightarrow \mathbb{R}$ es la función error, entonces E es continua.*

Demostración. Sean $f, g \in X$ fijos y $p \in Y$, entonces por la desigualdad triangular y la definición de ínfimo se sigue que:

$$\begin{aligned} E(f) &\leq \|f - p\| \leq \|f - g\| + \|g - p\| \quad \text{para toda } p \in Y, \text{ luego} \\ E(f) - \|f - g\| &\leq \|g - p\| \quad \text{y aplicando ínfimo sobre toda } p \in Y \text{ se sigue que} \\ E(f) - E(g) &\leq \|f - g\| \quad \text{para cualesquiera } f, g \in X. \end{aligned}$$

Como la última desigualdad es arbitraria indistintamente del orden de f y g se tiene

$$\begin{aligned} E(f) - E(g) &\leq \|f - g\| \\ E(g) - E(f) &\leq \|f - g\| \end{aligned}$$

por tanto

$$|E(f) - E(g)| \leq \|f - g\|. \quad (1)$$

Así dado $\epsilon > 0$, tomemos $0 < \delta < \epsilon$, para obtener:

$$\|f - f_0\| < \delta \text{ implica } |E(f) - E(f_0)| \leq \|f - f_0\| < \delta < \epsilon.$$

Por tanto E es continua. □

Ahora nos centramos en analizar las propiedades de los elementos de mejor aproximación, y como se muestra en [7, pag.59], la existencia de la mejor aproximación está garantizada cuando Y es un subespacio de X de dimensión finita, es decir, el conjunto denotado por

$$B(f) = \{P \in Y \mid E(f) = \|f - P\|\} \neq \emptyset,$$

más aún este conjunto tiene propiedades interesantes como se mostrará más adelante.

Proposición 2.2. *E es una función sublineal.*

Demostración. Probemos que E cumple

i) $E(f + g) \leq E(f) + E(g).$

ii) $E(\alpha f) = |\alpha|E(f).$

Sean $P(f) \in B(f)$ y $P(g) \in B(g)$ (pues $B(f), B(g) \neq \emptyset$, pero no necesariamente únicos), luego

$$E(f + g) \leq \|(f + g) - (P(f) + P(g))\| \leq \|f - P(f)\| + \|g - P(g)\| = E(f) + E(g).$$

Por otro lado, sea $\alpha \in \mathbb{R}$. Si $\alpha \neq 0$, entonces

$$E(\alpha f) = \|\alpha f - \frac{1}{\alpha}P(\alpha f)\| = |\alpha|\|f - \frac{1}{\alpha}P(\alpha f)\| \geq E(f)$$

por tanto $E(\alpha f) \geq |\alpha|E(f)$. Inversamente

$$|\alpha|E(f) = |\alpha|\|f - P(f)\| = \|\alpha f - \alpha P(f)\| \geq E(\alpha f).$$

Las desigualdades anteriores son válidas ya que toda combinación lineal de elementos de mejor aproximación están en Y y para el caso $\alpha = 0$ es obvio, así se tiene la igualdad y por tanto E es sublineal. □

Notemos por *ii*) de la proposición anterior que $E(-f) = E(f)$, es decir, el error E es simétrico con respecto a el inverso aditivo de X .

Proposición 2.3. *Si Y es un subespacio vectorial de X , entonces el conjunto $B(f)$ es cerrado, acotado y convexo.*

Demostración. Se tiene los siguientes dos casos.

a) Si $f \in Y$, entonces $E(f) = 0$, por tanto $B(f) = \{f\}$ el cual es cerrado, acotado y convexo.

b) Sean $f \in X - Y$, $p \in Y$ y $\epsilon > 0$.

Si $d(p, B(f)) = \inf_{q \in B(f)} \|q - p\|$, entonces existe un $q \in B(f)$ tal que

$$\|q - p\| < d(p, B(f)) + \epsilon,$$

dado que $E(f) = \|f - q\|$, luego se sigue que

$$\|f - p\| \leq \|f - q\| + \|q - p\| < E(f) + d(p, B(f)) + \epsilon,$$

es decir, $\|f - p\| < E(f) + d(p, B(f)) + \epsilon$, para todo $\epsilon > 0$. Por tanto se tiene la desigualdad

$$E(f) \leq \|f - p\| \leq E(f) + d(p, B(f)) \quad \text{para toda } p \in Y. \quad (2)$$

Por otro lado, recordemos que $p \in \overline{B(f)}$ si y sólo si $d(p, B(f)) = 0$. Así, afirmamos que $\overline{B(f)} \subset B(f)$. Si $p \in \overline{B(f)}$, entonces $d(p, B(f)) = 0$ y por (2) se sigue que $\|f - p\| = E(f)$, es decir, $p \in B(f)$ y por tanto $B(f)$ es cerrado. Más aún, de

$$E(f) + \|f\| = \|f - p\| + \|f\| \geq \|p\|$$

para todo $p \in B(f)$, se sigue que $B(f)$ es acotado. Finalmente, $B(f)$ es convexo. Sean α y β no negativos con $\alpha + \beta = 1$ y $p, q \in B(f)$, se deduce que,

$$\|\alpha f + \beta f - (\alpha p + \beta q)\| \leq \alpha \|f - p\| + \beta \|f - q\| = \alpha E(f) + \beta E(f)$$

y factorizando se tiene $\|f - (\alpha p + \beta q)\| \leq E(f)$. Además como Y es un subespacio vectorial, entonces $\alpha p + \beta q \in Y$, así $E(f) = \|f - (\alpha p + \beta q)\|$, donde $\alpha + \beta = 1$, es decir, $B(f)$ es convexo. \square

Si Y es de dimensión finita, entonces Y es topológicamente isomorfo a \mathbb{R}^n , es decir, con $\{x_1, \dots, x_n\}$ y $\{e_1, \dots, e_n\}$ bases, respectivamente de Y y de \mathbb{R}^n , la función

$$\phi(y) = \phi \left(\sum_{i=1}^n \lambda_i x_i \right) := \sum_{i=1}^n \lambda_i e_i,$$

es un isomorfismo de espacios vectoriales y un homeomorfismo de espacios topológicos, luego se sigue que $\phi(B(f))$ es cerrado y acotado en \mathbb{R}^n y por tanto compacto. Por otro lado como la compacidad es un invariante topológico se sigue que $B(f)$ es compacto en Y y más aún es compacto en X , pues Y es cerrado en X . Así se tiene el siguiente resultado.

Corolario 2.4. $B(f)$ es compacto y convexo para cada $f \in X$.

Una observación importante es que ϕ preserva la convexidad por ser isomorfismo de espacios vectoriales (note que la convexidad no es invariante topológico).

En muchos casos prácticos es importante saber que existe un único polinomio de mejor aproximación. Decimos que un espacio es estrictamente convexo, si dado $f_1 \neq f_2$ donde $\|f_1\| = \|f_2\| = 1$ y $\alpha_1, \alpha_2 > 0$ con $\alpha_1 + \alpha_2 = 1$, entonces $\|\alpha_1 f_1 + \alpha_2 f_2\| < 1$. Está es una condición suficiente para la unicidad del mejor aproximante y como ejemplo de espacio estrictamente convexo se tiene L_p con $1 < p < \infty$. En estos espacios el polinomio de mejor aproximación es único.

Definición 2.5. Si Y es un subespacio de X tal que para cada $f \in X$ existe un único polinomio de aproximación en Y , entonces se dice que Y es un conjunto de unicidad de X .

Proposición 2.6. Sea Y un subespacio de unicidad de dimensión finita de X . Si $P : X \rightarrow Y$ es el operador de mejor aproximación definido por $f \rightarrow P(f)$, entonces P es continua.

Demostración. Como $P(f)$ es único, entonces el operador P está bien definido. Para la continuidad basta con demostrar que P es acotado. De la desigualdad (1) y haciendo $g = 0$ se sigue que $E(f) \leq \|f\|$ para toda $f \in X$, así

$$\|P(f)\| \leq \|f - P(f)\| + \|f\| = E(f) + \|f\| \leq 2\|f\|,$$

por tanto $\|P\| \leq 2$. □

En algunos espacios especiales existen los subespacios de unicidad, uno de ellos es cuando el espacio es de Hilbert, es decir, cuando el espacio tiene definido un producto interno. En estos espacios una caracterización del mejor aproximante es el siguiente.

Teorema 2.7. Sea $(H, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert y sea H_0 un subespacio de H , un elemento $g \in H_0$ es de mejor aproximación para $f \in H$ si y sólo si cumple la siguiente condición de ortogonalidad.

$$\langle f - g, h \rangle = 0 \quad \text{para toda } h \in H_0. \quad (3)$$

Demostración. Procedemos por contradicción. Supongamos que g es un elemento de mejor aproximación a f y que existe un $h \in H_0$ para el cual no se cumple (3), es decir, $\langle f - g, h \rangle \neq 0$ para algún $h \in H_0$. Sin pérdida de generalidad se puede escoger h tal que $\|h\| = 1$ pues el vector normalizado también es no ortogonal, de hecho todos los vectores sobre la recta generada por h . Además puede escoger a h o $-h$ según sea el caso para suponer $\langle f - g, h \rangle = \delta > 0$. Luego sea el vector $k = g + \delta h \in H$, así se tiene

$$\|f - (g + \delta h)\|^2 = \|f - g\|^2 - \langle f - g, \delta h \rangle - \langle \delta h, f - g \rangle + \|\delta h\|^2 \quad (4)$$

$$= \|f - g\|^2 - \langle f - g, \delta h \rangle - \langle \delta h, f - g \rangle + \delta^2. \quad (5)$$

Como $\langle f - g, h \rangle = \delta$, entonces $\delta \langle f - g, h \rangle = \langle f - g, \delta h \rangle = \delta^2$; de manera semejante se tiene $\langle \delta h, f - g \rangle = \delta^2$, se sigue que

$$\begin{aligned} \|f - k\|^2 &= \|f - (g + \delta h)\|^2 = \|f - g\|^2 - \delta^2 - \delta^2 + \delta^2 \\ &= \|f - g\|^2 - \delta^2 < \|f - g\|^2, \end{aligned}$$

por tanto $\|f - k\| < \|f - g\|$, pero esto no puede ser pues g es de mejor aproximación, entonces $\langle f - g, h \rangle = 0$ para toda $h \in H_0$.

Recíprocamente, como H_0 es un subespacio de H , entonces $H_0 = H_0 + g$. Se definen los conjuntos

$$A = \{\|f - (g + h)\| : h \in H_0\} \quad \text{y} \quad B = \{\|f - k\| : k \in H_0\},$$

respectivamente. Afirmamos que $A = B$; si $\|f - k\| \in A$ y como $k = g + h$ para alguna $h \in H_0$, entonces se sigue que $\|f - k\| = \|f - (g + h)\| \in B$, luego $B \subset A$. La otra contención es clara, pues $g + h \in H_0$, por tanto

$$E(f) = \inf_{k \in H_0} \|f - k\| = \inf_{b \in B} b = \inf_{a \in A} a = \inf_{h \in H_0} \|f - (g + h)\| \geq \|f - g\|.$$

Utilizando la ecuación (4) con $\delta = 1$ y la hipótesis $\langle f - g, h \rangle = 0$, $E(f) = \|f - g\|$. \square

Un producto interior define un funcional lineal real o complejo según sea el producto interior, así la ecuación (3) se puede interpretar como sigue: para el funcional $\lambda(h) := \langle f - g, h \rangle$ se tiene $\lambda(h) = 0$ para toda $h \in H_0$, luego podemos tomar lo anterior como motivación para la siguiente definición en espacios de Banach.

Sea X un espacio y Y un subespacio vectorial de X , decimos que un funcional lineal real o complejo λ es ortogonal a Y si $\lambda(h) = 0$ para toda $h \in Y$ y denotamos por Y^\perp al conjunto de todos los funcionales lineales ortogonales a Y .

Teorema 2.8. *Sea Y un subespacio cerrado de X y $f \in X - Y$. Entonces $f_0 \in Y$ es elemento de mejor aproximación a f en Y si y sólo si existe un funcional lineal acotado λ tal que $\lambda \in Y^\perp$, $\|\lambda\| = 1$ y $\|f - f_0\| = \lambda(f)$. Más aún*

$$E(f) = \sup_{\substack{\lambda \in Y^\perp \\ \|\lambda\|=1}} \lambda(f). \quad (6)$$

Demostración. Sea $f \in X - Y$ y f_0 el mejor aproximante a f en Y , es decir, $\|f - f_0\| = E(f)$. Definimos el funcional lineal $\lambda : H \rightarrow \mathbb{R}$ por

$$\lambda(y + mf) = mE(f), \quad \text{donde } H = Y \oplus \text{gen}\{f\}. \quad (7)$$

Notemos que H es un subespacio vectorial de X y λ está bien definido puesto que para cada $h \in H$ existen únicos $y \in Y$ y $m \in \mathbb{R}$ tales que $h = y + mf$, más aún es lineal ya que si $h = y + mf$, $h' = y' + m'f$ y $\alpha \in \mathbb{R}$, entonces

$$\begin{aligned} \lambda(h + \alpha h') &= \lambda((y + \alpha y') + (m + \alpha m')f) = (m + \alpha m')E(f) \\ &= mE(f) + \alpha m'E(f) = \lambda(y + mf) + \alpha \lambda(y' + m'f) \\ &= \lambda(h) + \alpha \lambda(h'). \end{aligned}$$

Con lo anterior demostraremos las propiedades para el λ del teorema. Primero notemos que $h \in Y$ si y sólo si $m = 0$, entonces $\lambda(h) = 0$ para toda $h \in Y$, es decir, $\lambda \in Y^\perp$. Además se tiene que

$$|\lambda(y + mf)| = |m|E(f) \leq \|mf + y\| \quad \text{para toda } y \in Y \text{ y } m \in \mathbb{R}.$$

Por tanto $\|\lambda\| = 1$ sobre H y

$$\lambda(f) = \lambda(f - y) = E(f) = \|f - f_0\|.$$

Por otro lado notemos que λ es un funcional lineal sobre H el cual está acotado por un funcional sublineal E continuo sobre H (por la proposición (2.2)), más aún λ coincide con E sobre H , por el teorema de Hahn-Banach (ver [1], pag.48) existe una extensión λ a todo X con $\|\lambda\| = 1$ sobre X , además λ mantiene la ortogonalidad sobre Y .

Recíprocamente, como $\lambda(f - y) \leq \|\lambda\| \|f - y\|$ y por hipótesis $\|\lambda\| = 1$, $\lambda \in Y^\perp$ y $\|f - f_0\| = \lambda(f)$, entonces

$$\lambda(f) = \lambda(f) - \lambda(y) = \lambda(f - y) \leq \|f - y\| \quad \text{para todo } y \in Y,$$

así se sigue que

$$\|f - f_0\| = \lambda(f) \leq \inf_{y \in Y} \|f - y\| = E(f) \quad \text{para cada } f \in X. \quad (8)$$

Por tanto $\|f - f_0\| = E(f)$, es decir, f_0 es el mejor aproximante de f en Y . Más aún, de la desigualdad (8), si fijamos f , entonces para toda $\lambda \in Y^\perp$ y $\|\lambda\| = 1$, $E(f)$ es una cota superior, luego

$$\sup_{\substack{\lambda \in Y^\perp \\ \|\lambda\|=1}} \lambda(f) \leq E(f).$$

Finalmente, como f_0 es el mejor aproximante de f sobre Y y por (7) existe un funcional ortogonal sobre Y tal que $\|\lambda\| = 1$, se sigue la desigualdad contraria

$$E(f) = \|f - f_0\| = \lambda(f) \leq \sup_{\substack{\lambda \in Y^\perp \\ \|\lambda\|=1}} \lambda(f). \quad \square$$

Ahora nuestro interés es la caracterización del mejor aproximante en $C(X)$ donde X es compacto Hausdorff, pues estos espacios en la práctica son unos de los más usados por las propiedades que tienen. En el desarrollo de este capítulo se mostrará la fuerza del siguiente teorema ya que al final se caracterizará el polinomio de mejor aproximación de la manera más simple con el teorema de alternancia de Chebyshev.

Teorema 2.9 (Kolmogorov). *Si Y es un subespacio de dimensión finita de $C(X)$, entonces P es un elemento de mejor aproximación de $f \in C(X)$ sobre Y si y sólo si para cada $Q \in Y$ se tiene*

$$\max_{x \in A_0} \Re\{[f(x) - P(x)]\overline{Q(x)}\} \geq 0, \quad (9)$$

donde A_0 denota el conjunto (el cual depende de f y P) de todos los $x \in X$ para los cuales $|f(x) - P(x)| = \|f - P\|$.

Demostración. Procedemos por contradicción, en efecto, supongamos que existe $Q \in Y$ tal que

$$\max_{x \in A_0} \Re\{[f(x) - P(x)]\overline{Q(x)}\} = -2\epsilon < 0,$$

para algún $\epsilon > 0$. Por la continuidad de la parte real \Re ,

$$|\Re\{[f - P]\overline{Q(x)}\} - \Re\{[f - P]\overline{Q(x_0)}\}| < \epsilon$$

para toda $x \in B_\delta(x_0)$, para algún δ y $x_0 \in A_0$. Por lo tanto,

$$\Re\{[f - P]\overline{Q(x)}\} < \epsilon + \Re\{[f - P]\overline{Q(x_0)}\}$$

para toda $x \in B_\delta(x_0)$. Además como

$$\Re\{[f - P]\overline{Q(x_0)}\} \leq \max_{x \in A_0} \Re\{[f(x) - P(x)]\overline{Q(x)}\} = -2\epsilon,$$

entonces

$$\Re\{[f(x) - P(x)]\overline{Q(x)}\} < \epsilon - 2\epsilon = -\epsilon$$

para toda $x \in B_\delta(x_0)$. Tomando ahora el conjunto abierto $G := \bigcup_{x \in A_0} B_\delta(x)$, se tiene que G contiene a A_0 (notemos que $\delta(\epsilon, x)$ y $x \in A_0$), luego

$$\Re\{[f(x) - P(x)]\overline{Q(x)}\} < -\epsilon \text{ para todo } x \in G \supseteq A_0. \quad (10)$$

Por otro lado, supongamos la función $P_1 = P - \lambda Q$ donde $\lambda > 0$ y $M = \max |Q(x)|$, entonces para cada $x \in G$ se sigue que

$$\begin{aligned} |f(x) - P_1(x)|^2 &= |f(x) - P(x) + \lambda Q(x)|^2 \\ &= |f(x) - P(x)|^2 + 2\lambda \Re[(f - P)(x)\overline{Q(x)}] + \lambda^2 |Q(x)|^2 \\ &\leq E(f)^2 - 2\lambda\epsilon + \lambda^2 M^2. \end{aligned}$$

Ahora tomando $\lambda < M^{-2}\epsilon$, se sigue que $\lambda^2 M^2 < \lambda\epsilon$ y entonces se tiene la desigualdad $|f(x) - P_1(x)|^2 < E(f)^2 - \lambda\epsilon < E(f)^2$, es decir,

$$|f(x) - P_1(x)| < E(f) \text{ para toda } x \in G. \quad (11)$$

Por otro lado

$$|f(x) - P(x)| \leq \|f - P\| = E(f) = \inf_{P \in Y} \|f - P\|, \quad x \in X,$$

así se tiene la desigualdad estricta

$$|f(x) - P(x)| < E(f), \quad x \in G^c.$$

Sea $\delta > 0$ tal que $|f(x) - P(x)| < E(f) - \delta$, $x \in G^c$, entonces se tiene que,

$$|f(x) - P_1(x)| \leq |f(x) - P(x)| + \lambda |Q(x)| \leq E(f) - \delta + \lambda M$$

y tomando λ de tal manera que $\lambda < (2M)^{-1}\delta$ se deduce que

$$|f(x) - P_1(x)| < E(f) - \delta + \frac{1}{2}\delta < E(f), \quad x \in G^c$$

De esta última desigualdad y de (11) se sigue

$$|f(x) - P_1(x)| < E(f) \text{ para todo } x \in X \quad (12)$$

Por lo anterior, como $P_1 \in Y$ se obtiene que $\|f - P_1\| < E(f)$, es decir, existe un menor error al que se tiene inicialmente, es una contradicción y por tanto.

$$\max_{x \in A_0} \Re\{[f(x) - P(x)]\overline{Q(x)}\} \geq 0, \quad Q \in Y.$$

Recíprocamente, sea $P_1 \in Y$. Si $P - P_1 := Q$ y $x_0 \in X$, entonces

$$|f(x_0) - P_1(x_0)|^2 = |f(x_0) - P(x_0)|^2 + 2\Re\{[f(x_0) - P(x_0)]\overline{Q(x_0)}\} + |Q(x_0)|^2,$$

donde $\Re\{[f(x_0) - P(x_0)]\overline{Q(x_0)}\} \geq 0$. Luego

$$|f(x_0) - P_1(x_0)| \geq |f(x_0) - P(x_0)| = \|f - P\|,$$

por tanto P_1 no es de mejor aproximación, pues si lo fuera se tendría $|f(x) - P_1(x)| \leq \|f - P_1\| = E(f)$ para todo $x \in X$, así P es el mejor aproximante. \square

En adelante A_0 siempre denotará el conjunto utilizado en la hipótesis del teorema anterior. Como consecuencia del Teorema de Kolmogorov y considerando el espacio de funciones continuas reales $C(X)$ se tiene.

Corolario 2.10. *Si Y es un subespacio de dimensión finita de $C(X)$ y $P \in Y$, entonces P es un elemento de mejor aproximación de $f \in C(X)$ si y sólo si para cada $Q \in Y$ se tiene*

$$\max_{x \in A_0} \{[f(x) - P(x)]Q(x)\} \geq 0.$$

Los teoremas anteriores son caracterizaciones del elemento de mejor aproximación, utilizando funcionales lineales ortogonales. El Teorema de Rivlin-Shapiro que demostraremos utiliza el Teorema de Kolmogorov, así como el Teorema de Carathéodory, cuya prueba se puede ver en ([1], pag.46). Además el Teorema de Hahn-Banach en su versión geométrica que también se utiliza en el desarrollo de la prueba del Teorema Rivlin-Shapiro, se puede ver en ([8], pag.204).

Teorema 2.11. *Sean G y G_0 conjuntos disjuntos convexos en un espacio lineal normado X . Si G_0 es abierto, entonces existe un funcional lineal λ en X y un $\gamma \in \mathbb{R}$ que cumplen*

$$\lambda(g) < \gamma \leq \lambda(f) \text{ para toda } f \in G, \quad g \in G_0.$$

Teorema 2.12 (Carathéodory). *Si $B \subset \mathbb{R}^n$ y $r \leq n$, entonces el casco convexo $\text{cov } B$ consta de todos los elementos de la forma*

$$x = \sum_{i=0}^r p_i x_i, \text{ con } x_i \in B, \quad p_i \geq 0 \text{ y } \sum_{i=0}^r p_i = 1$$

para $r \in \{1, 2, \dots, n\}$.

Teorema 2.13 (Rivlin-Shapiro). *Sea Y un subespacio lineal de $C(X)$ de dimensión n . Una función $P \in Y$ es de mejor aproximación a $f \in C(X)$ si y sólo si existen puntos $x_i \in A_0$, $i = 0, \dots, r$ y números $p_i \geq 0$ tal que*

$$\sum_{i=0}^r p_i = 1 \text{ y } \sum_{i=0}^r p_i [f(x_i) - P(x_i)] Q(x_i) = 0,$$

para todo $Q \in Y$ con $r \leq n$.

Demostración. Note que todo funcional lineal $\lambda : \mathbb{R}^n \rightarrow \mathbb{R}$, está definido en \mathbb{R}^n si y sólo si está definido sólo en la base canónica de \mathbb{R}^n , es decir, λ es funcional lineal en \mathbb{R}^n si y sólo si $\lambda(\mathbf{x}) = \sum_{i=1}^n a_i x_i$, donde $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ y $\mathbf{a} = (\lambda(e_1), \dots, \lambda(e_n))$. Si $\{\phi_1, \dots, \phi_n\}$ es una base para $Y \subset C(X)$, entonces para toda $Q \in Y$, se tiene $Q = \sum_{i=1}^n c_i \phi_i$, con $c_i \in \mathbb{R}$, para $i = 1 \dots, n$.

Por otro lado sea la función continua $\Phi = ((f-P)\phi_1, \dots, (f-P)\phi_n) : X \rightarrow \mathbb{R}^n$ definida por

$$\Phi(x) = ((f-P)(x)\phi_1(x), \dots, (f-P)(x)\phi_n(x)).$$

Como A_0 es un subconjunto cerrado (A_0 es la imagen inversa del escalar $\|f-p\|$ bajo la función continua $|f-p|$) en un espacio X compacto y Hausdorff, entonces A_0 es compacto, luego entonces $\Phi(A_0)$ es un subespacio compacto de \mathbb{R}^n . Más aún, el casco convexo $\text{cov } \Phi(A_0)$ es compacto.

Por otro lado afirmamos:

- i) P es de mejor aproximación a f en Y si y sólo si $\max_{x \in \Phi(A_0)} \lambda(x) \geq 0$ para toda λ funcional lineal.
- ii) Si $\max_{x \in \Phi(A_0)} \lambda(x) \geq 0$ para toda λ funcional lineal, entonces $0 \in \text{cov } \Phi(A_0)$.

Probemos i), supongamos que P es de mejor aproximación a f y sea λ un funcional lineal. Luego

$$(\lambda\Phi)(x) = \lambda(\Phi(x)) = \sum_{i=0}^n \lambda(e_i)(f-P)(x)\phi_i(x) = (f-P)(x) \left(\sum_{i=0}^n \lambda(e_i)\phi_i(x) \right),$$

es decir, $\lambda\Phi(x) = (f-P)(x)Q(x)$, donde

$$Q(x) = \sum_{i=0}^n \lambda(e_i)\phi_i(x) \in Y.$$

Por el Teorema de Kolmogorov se sigue que

$$0 \leq \max_{x \in A_0} (f-P)(x)Q(x) = \max_{x \in A_0} \lambda\Phi(x) = \max_{x \in \Phi(A_0)} \lambda(x). \quad (13)$$

Recíprocamente, supongamos que $0 \leq \max_{x \in \Phi(A_0)} \lambda(x)$. Sea

$$Q(x) = \sum_{i=0}^n c_i \phi_i(x) \in Y$$

y el funcional lineal definido por

$$\lambda(\mathbf{x}) = \sum_{i=0}^n c_i x_i,$$

por (13) y el Teorema de Kolmogorov se tiene el resultado. Para la prueba de ii), procedemos por contradicción. Supongamos que $0 \notin \text{cov } \Phi(A_0)$, como $\text{cov } \Phi(A_0)$ es compacto, este es cerrado en \mathbb{R}^n , luego sea la bola abierta $B_\delta(0)$ en el complemento de $\text{cov } \Phi(A_0)$, ahora si hacemos $G := \text{cov } \Phi(A_0)$ y $G_0 = B_\delta(0)$, entonces se cumple la hipótesis del teorema (2.11) y por tanto existe un funcional lineal λ en \mathbb{R}^n y un $\gamma \in \mathbb{R}$ que cumplen

$$\lambda(g) < \gamma \leq \lambda(f) \text{ para toda } f \in G, g \in G_0.$$

En particular, $0 = \lambda(0) < \gamma \leq \lambda(f)$ para toda $f \in \Phi(A_0)$, es decir, $-\lambda(f) < 0$ para toda $f \in \Phi(A_0) \subseteq \text{cov } \Phi(A_0)$, luego $\max_{f \in \Phi(A_0)} -\lambda(f) \leq 0$, lo cual contradice la hipótesis.

Finalmente, por i), ii) y aplicando el Teorema de Carathéodory, se sigue que $\sum_{i=0}^r p_i y_i = 0$, donde $y_i = \Phi(x_i)$ para algún $x_i \in A_0$ y los $p_i \geq 0$ son tal que $\sum_{i=0}^r p_i = 1$, con $r \leq n$. Ahora sea

$$Q = \sum_{j=0}^r c_j \phi_j \in Y$$

y el correspondiente funcional lineal λ , donde $\lambda(e_j) = c_j$, $j = 1, \dots, n$, entonces

$$\begin{aligned} 0 &= \lambda(0) = \sum_{i=0}^r p_i \lambda(y_i) = \sum_{i=0}^r p_i \left(\sum_{j=0}^n \lambda(e_j)(f - P)(x_i) \phi_j(x_i) \right) \\ &= \sum_{i=0}^r p_i \left(\sum_{j=0}^n (f - P)(x_i) c_j \phi_j(x_i) \right) = \sum_{i=0}^r p_i (f - P)(x_i) \left(\sum_{j=0}^n c_j \phi_j(x_i) \right) \\ &= \sum_{i=0}^r p_i (f - P)(x_i) Q(x_i), \end{aligned}$$

por tanto, existen puntos $x_i \in A_0$, $i = 0, \dots, r$ y números $p_i \geq 0$ tal que

$$\sum_{i=0}^r p_i = 1 \text{ y } \sum_{i=0}^r p_i [f(x_i) - p(x_i)]Q(x_i) = 0$$

para todo $Q \in Y$ con $r \leq n$.

Para el recíproco del teorema, supongamos que

$$\sum_{i=0}^r p_i [f(x_i) - p(x_i)]Q(x_i) = 0.$$

Así $[f(x_i) - p(x_i)]Q(x_i) \geq 0$, para algún $x_i \in A_0$ con $0 \leq i \leq r$, pues $p_i \geq 0$, y por tanto

$$\max_{x \in A_0} [f(x) - p(x)]Q(x) \geq [f(x_i) - p(x_i)]Q(x_i) \geq 0,$$

para todo $Q \in Y$. Finalmente, por el Teorema de Kolmogorov se sigue el resultado. \square

Corolario 2.14. Si P es de mejor aproximación a f en X , entonces P es de mejor aproximación a f en algún subconjunto $\{x_0, \dots, x_r\} \subseteq A_0$.

Como motivación del Teorema Rivlin-Shapiro surge la siguiente definición que es importante porque permite caracterizar de otra forma a los elementos de mejor aproximación en términos de funcionales lineales ortogonales como se verá más adelante.

Definición 2.15. Sea Y un subespacio lineal de $C(X)$ de dimensión finita n , λ^* se llama funcional lineal de aproximación de f y P en Y , si dados $\{x_0, \dots, x_r\} \subseteq A_0$ y $r \leq n$, se tiene

$$\lambda^*(g) = \sum_{i=0}^r \alpha_i g(x_i),$$

donde $\alpha_i = p_i \text{ sign}[f(x_i) - P(x_i)]$, $p_i \geq 0$, $i = 0, \dots, r$, $\sum_{i=0}^r p_i = 1$.

Corolario 2.16. Sea Y un subespacio lineal de $C(X)$ de dimensión finita n . Si $f \notin Y$, entonces P es de mejor aproximación a f en Y si y sólo si existe un funcional lineal de aproximación de f y P , para algunos puntos $\{x_0, \dots, x_r\} \subseteq A_0$ y $r \leq n$, el cual es ortogonal en Y .

Demostración. Supongamos que P es de mejor aproximación a f en Y . Por el Teorema de Rivlin-Shapiro, existen puntos $x_i \in A_0$, $i = 0, \dots, r$ y números $p_i \geq 0$ tal que

$$\sum_{i=0}^r p_i = 1 \text{ y } \sum_{i=0}^r p_i [f(x_i) - P(x_i)]Q(x_i) = 0,$$

para todo $Q \in Y$ con $r \leq n$, luego se define el funcional lineal ortogonal de f y P en Y por

$$\lambda^*(g) = \sum_{i=0}^r \alpha_i g(x_i)$$

donde $\alpha_i = p_i \operatorname{sign}[f(x_i) - P(x_i)]$, $p_i \geq 0$ y $\sum_{i=0}^r p_i = 1$. Como

$$\begin{aligned} \|f - P\| \lambda^*(Q) &= \lambda^*(\|f - P\| Q) \\ &= \sum_{i=0}^r \alpha_i \|f - P\| Q(x_i) \\ &= \sum_{i=0}^r p_i [f(x_i) - P(x_i)] Q(x_i) = 0, \end{aligned}$$

se sigue que $\lambda^*(Q) = 0$ para todo $Q \in Y$, es decir, λ^* es ortogonal en Y . Recíprocamente, por el teorema (2.8), basta probar que el funcional lineal ortogonal λ^* cumple $\|\lambda^*\| = 1$ y $\|f - P\| = \lambda^*(f)$. Como

$$|\lambda^*(g)| = \left| \sum_{i=0}^r \alpha_i g(x_i) \right| \leq \sum_{i=0}^r p_i |g(x_i)| \leq \sum_{i=0}^r p_i \|g\| = \|g\|,$$

luego $\|\lambda^*\| = 1$. Por otro lado, como λ^* es ortogonal en Y se sigue que

$$\begin{aligned} \lambda^*(f) &= \lambda^*(f - P) = \sum_{i=0}^r p_i \operatorname{sign}[f(x_i) - P(x_i)] [f(x_i) - P(x_i)] \\ &= \sum_{i=0}^r p_i (\operatorname{sign}[f(x_i) - P(x_i)])^2 \|f - P\| = \|f - P\|. \quad \square \end{aligned}$$

3 Sistemas de Haar

En la sección anterior todos los teoremas son caracterizaciones del mejor aproximante, finalizando con el teorema y corolario que caracteriza de la manera más sencilla tal polinomio. Ahora necesitamos garantizar la unicidad del polinomio de mejor aproximación. Para ello definimos los subespacios de Haar y en particular los subespacios de Chebyshev en espacios de Hausdorff.

Definición 3.1. Sea X un espacio Hausdorff con al menos $n + 1$ puntos y sean h_0, h_1, \dots, h_n funciones reales continuas. Decimos que el conjunto de funciones $\Phi := \{h_0, h_1, \dots, h_n\}$ es un sistema de Haar de orden $n + 1$ en X , si cada combinación lineal $\lambda_0 h_0 + \lambda_1 h_1 + \dots + \lambda_n h_n$ con coeficientes $\lambda_0, \lambda_1, \dots, \lambda_n \in \mathbb{R}$ no

simultáneamente todos ceros, no tienen más de n ceros distintos en X . El subespacio generado $H = \text{gen}\{h_0, \dots, h_n\}$ de X por el sistema de Haar de orden $n + 1$ en X es llamado subespacio de Haar de orden $n + 1$ en $\mathcal{C}(X)$. En particular si X es compacto Hausdorff decimos que Φ es un Sistema de Chebyshev, respectivamente H es llamado subespacio de Chebyshev.

De la definición y el hecho que X tiene cardinalidad $n + 1$, si Φ es un sistema de Haar, entonces Φ forma un conjunto linealmente independiente en $C(X)$. El conjunto $\{1, x, \dots, x^n\}$ en $[a, b]$ es un sistema de Chebyshev.

Proposición 3.2. Sean X un espacio Hausdorff de cardinalidad $n+1$ y $h_0, h_1 \dots h_n \in \mathcal{C}(X)$. $\Phi = \{h_0, h_1 \dots h_n\}$ forma un sistema de Haar de orden $n + 1$ en $\mathcal{C}(X)$ si y sólo si $\det(h_i(x_j)) \neq 0$ para cualesquiera puntos distintos $\{x_0, \dots, x_n\} \in X$.

Demostración. Sean x_0, \dots, x_n puntos distintos de X . Evaluando h_i en x_j para $j = 0, \dots, n$ y $i = 0, \dots, n$ se tiene el sistema de ecuaciones lineal homogéneo de orden $n + 1$:

$$\begin{aligned} \lambda_0 h_0(x_0) + \lambda_1 h_1(x_0) + \dots + \lambda_n h_n(x_0) &= 0 \\ \lambda_0 h_0(x_1) + \lambda_1 h_1(x_1) + \dots + \lambda_n h_n(x_1) &= 0 \\ &\vdots \\ \lambda_0 h_0(x_n) + \lambda_1 h_1(x_n) + \dots + \lambda_n h_n(x_n) &= 0 \end{aligned} \tag{14}$$

que tiene como única solución a $\lambda_0 = \lambda_1 = \dots = \lambda_n = 0$, pues de lo contrario contradice que Φ es un sistema de Haar, ya que la combinación lineal $\sum_{i=0}^n \lambda_i h_i = 0$ tendría $n + 1$ ceros, así $\det(h_i(x_j)) \neq 0$. Recíprocamente, si existe una combinación lineal $\sum_{i=0}^n \lambda_i h_i = 0$ con coeficientes $\lambda_0, \lambda_1, \dots, \lambda_n \in \mathbb{R}$ no simultáneamente todos ceros, tal que tiene al menos $n + 1$ ceros, entonces el sistema de ecuaciones lineales homogéneo (14) de orden $n + 1$ tiene una solución no trivial $(\lambda_0, \dots, \lambda_n)$, por tanto $\det(h_i(x_j)) = 0$ contradiciendo la hipótesis, así Φ es un sistema de Haar de orden $n + 1$. \square

El teorema anterior es importante porque nos permite encontrar una única función en el subespacio de Haar, cuya gráfica pase por algunos puntos determinados, es decir, se tiene el siguiente corolario.

Corolario 3.3. Sean $x_0, \dots, x_n \in X$ y $\alpha_0, \dots, \alpha_n \in \mathbb{R}$. Si Φ forma un sistema de Haar en X , entonces existe una única combinación lineal $h = \sum_{i=0}^n \lambda_i h_i$ tal que $h(x_i) = \alpha_i$ con $i = 1, \dots, n$.

Demostración. Por el teorema anterior se tiene $\det(h_i(x_j)) \neq 0$ y por tanto el sistema

$$\begin{aligned} \lambda_0 h_0(x_0) + \lambda_1 h_1(x_0) + \dots + \lambda_n h_n(x_0) &= \alpha_1 \\ \lambda_0 h_0(x_1) + \lambda_1 h_1(x_1) + \dots + \lambda_n h_n(x_1) &= \alpha_2 \\ &\vdots \\ \lambda_0 h_0(x_n) + \lambda_1 h_1(x_n) + \dots + \lambda_n h_n(x_n) &= \alpha_n. \end{aligned}$$

tiene solución única, $\lambda_0, \dots, \lambda_n$, así se tiene que $h = \sum_{i=0}^n \lambda_i h_i$ tal que $h(x_i) = \alpha_i$. □

De esta forma decimos que hay una interpolación h de los $\{h_0, \dots, h_n\}$ con $h(x_i) = \alpha_i$ para $i = 0, \dots, n$.

Proposición 3.4. *Si Φ es un sistema de Chebyshev de orden $n + 1$ en $[a, b]$ y $x_1, \dots, x_n \in (a, b)$, entonces la función definida por*

$$D(x, x_1, \dots, x_n) := \begin{vmatrix} h_0(x) & h_0(x_1) & \dots & h_0(x_n) \\ h_1(x) & h_1(x_1) & \dots & h_1(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x) & h_n(x_1) & \dots & h_n(x_n) \end{vmatrix} \tag{15}$$

tiene exactamente por ceros los puntos x_i , para $i = 1, \dots, n$. Más aún, en cada x_i cambia de signo $D(x, x_1, \dots, x_n)$.

Demostración. La primera parte es clara. Por las propiedades de los determinantes se sigue que $D(x_i, x_1, \dots, x_n) = 0$ para $i = 1, \dots, n$. Más aún, como $D(x, x_1, \dots, x_n)$ es una combinación lineal del sistema de Chebyshev Φ , entonces $D(x, x_1, \dots, x_n)$ tiene a lo más n ceros, y por tanto tiene exactamente n ceros. Sea $0 < \delta$ lo suficientemente pequeño tal que $(x_1 - \delta, x_1 + \delta) \subset [a, b]$. Definimos la función continua $g_\delta : [0, \delta] \rightarrow \mathbb{R}$ por

$$g_\delta(t) := \begin{vmatrix} h_0(x_1 - \delta + t) & h_0(x_1 + t) & \dots & h_0(x_n) \\ h_1(x_1 - \delta + t) & h_1(x_1 + t) & \dots & h_1(x_n) \\ \vdots & \vdots & & \vdots \\ h_n(x_1 - \delta + t) & h_n(x_1 + t) & \dots & h_n(x_n) \end{vmatrix}$$

Afirmamos que $g_\delta(t) \neq 0$ para toda $t \in [0, \delta]$. En efecto, supongamos que existe $t_0 \in [0, \delta]$ tal que $g_\delta(t_0) = 0$, entonces

$$\lim_{\delta \rightarrow t_0} g_\delta(t_0) = 0 = -D(x_1 + t_0, x_1, \dots, x_n)$$

contradiendo la primera parte de la proposición, así $g_\delta(t) \neq 0$ para toda $t \in [0, \delta]$. Finalmente observemos que

$$g(0) = D(x_1 - \delta, x_1, \dots, x_n) \text{ y } g(\delta) = -D(x_1 + \delta, x_1, \dots, x_n),$$

luego $D(x_1 - \delta, x_1, \dots, x_n)$ y $D(x_1 + \delta, x_1, \dots, x_n)$ tienen signos opuestos, pues $g \neq 0$ en todo el intervalo $[0, \delta]$, de esta misma forma se prueba que también cambia de signo $D(x, x_1, \dots, x_n)$, para x_2, \dots, x_n . □

Se tiene un resultado más general de la proposición anterior, de hecho la función $D(x, x_1, \dots, x_n)$ genera a todas las funciones en un subespacio de Chebyshev que cambian de signo exactamente en n puntos de un intervalo cerrado $[a, b]$.

Proposición 3.5. Sean Φ un sistema de Chebyshev de orden $n + 1$ en $[a, b]$ y H el subespacio de Chebyshev de orden $n + 1$. Si $h \in H$ tiene exactamente n ceros distintos x_1, \dots, x_n en (a, b) , entonces h necesariamente cambia de signo en cada x_i , para $i = 1, \dots, n$.

Demostración. Si $h \in H$ y x_1, \dots, x_n son los ceros de h , entonces $h = \sum_{i=0}^n \lambda_i h_i$, pues H es un subespacio diferente del trivial y por tanto existe algún $\lambda_i \neq 0$. Sin pérdida de generalidad supongamos que $\lambda_0 \neq 0$, así usando la propiedad de los determinantes y multiplicando a los λ_i diferentes de cero con el renglón i -ésimo correspondiente y sumando al primer renglón en la función $D(x, x_1, \dots, x_n)$, se sigue que

$$D(x, x_1, \dots, x_n) = \frac{1}{\lambda_0} \begin{vmatrix} \sum_{i=0}^n \lambda_i h_i & 0 & \dots & 0 \\ h_1(x) & h_1(x_1) & \dots & h_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ h_n(x) & h_n(x_1) & \dots & h_n(x_n) \end{vmatrix} = \frac{D_1}{\lambda_0} h(x),$$

donde

$$D_1 = \begin{vmatrix} h_1(x_1) & h_1(x_2) & \dots & h_1(x_n) \\ h_2(x_1) & h_2(x_2) & \dots & h_2(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ h_n(x_1) & h_n(x_2) & \dots & h_n(x_n) \end{vmatrix}$$

Entonces h tiene los mismos ceros que $D(x, x_1, \dots, x_n)$, luego por la proposición (3.4), h cambia de signo en cada x_i si y sólo si lo hace $D(x, x_1, \dots, x_n)$ en cada x_i . □

Lema 3.6. Sean X un espacio Hausdorff que tiene al menos $n + 2$ puntos, Φ un sistema de Haar de orden $n+1$ y P un polinomio de mejor aproximación a la función continua f , entonces el subconjunto A_0 para el cual $|f(x) - P(x)| = \|f - P\| = E(f)$ contiene al menos $n + 2$ puntos.

Demostración. Supongamos que $A_0 = \{x_0, \dots, x_s\}$ con $s \leq n + 1$, entonces existe un punto $x \in X$, tal que

$$0 \leq |f(x) - P(x)| < \|f - P\| = E(f),$$

luego entonces $0 < E(f)$. Por otro lado, por el sistema de ecuaciones lineales de $(s + 1) \times (n + 1)$

$$\begin{aligned} \lambda_0 \phi_0(x_0) + \lambda_1 \phi_1(x_0) + \dots + \lambda_n \phi_n(x_0) &= -[f(x_0) - P(x_0)] \\ \lambda_0 \phi_0(x_1) + \lambda_1 \phi_1(x_1) + \dots + \lambda_n \phi_n(x_1) &= -[f(x_1) - P(x_1)] \\ &\vdots \\ \lambda_0 \phi_0(x_s) + \lambda_1 \phi_1(x_s) + \dots + \lambda_n \phi_n(x_s) &= -[f(x_s) - P(x_s)] \end{aligned}$$

se tiene un polinomio $Q = \sum_{i=0}^n \lambda_i \phi_i$ (no necesariamente único) tal que

$$Q(x_k) = -[f(x_k) - P(x_k)]$$

con $k = 0, \dots, s$. Luego aplicando el Teorema de Kolmogorov se tiene

$$\begin{aligned} \max_{x \in A_0} [f(x) - P(x)]Q(x) &= \max_{0 \leq k \leq s} [f(x_k) - P(x_k)]Q(x_k) = \max_{0 \leq k \leq s} -[f(x_k) - P(x_k)]^2 \\ &= -E^2(f) < 0, \end{aligned}$$

pero esto no puede ser ya que contradice la caracterización del polinomio de mejor aproximación. \square

De forma inmediata se prueba la unicidad del polinomio de mejor aproximación usando el teorema anterior.

Teorema 3.7. Si Φ es un sistema de Haar de orden $n + 1$ en X , entonces toda función continua f tiene un único polinomio de mejor aproximación P .

Finalmente, si en X está definido un sistema de Haar se garantiza la unicidad del mejor aproximante. El teorema principal en esta sección es el teorema de Alternancia de Chebyshev que se pueden consultar en ([7], pag.74).

Teorema 3.8 (de Alternancia de Chebyshev). *Sean Φ un sistema de Haar de orden n en $[a, b]$ y $A \subseteq [a, b]$ de al menos $n + 1$ puntos. Si $f \in \mathcal{C}(A)$, entonces P es el polinomio de mejor aproximación a f en A si y sólo si existen $n + 1$ puntos $x_0 < \dots < x_n$ en A_0 y un número $\epsilon = 1$ o -1 , para el cual $f(x_i) - P(x_i) = d(-1)^i \epsilon$ con $i = 0, \dots, n$ y donde $d = \|f - P\|$.*

Notemos que el teorema de alternancia de Chebyshev sigue siendo un teorema de caracterización del polinomio de mejor aproximación y para calcular tal polinomio se puede utilizar el algoritmo de Remez.

4 Medida de Radon

Las medidas de Radon son utilizadas en muchas áreas del análisis, como es en la teoría de la probabilidad, la teoría del potencial o la teoría de la representación integral. En esta sección supondremos, al igual que antes, que X es un espacio localmente compacto y Hausdorff, $\mathcal{C}(X)$ es el espacio normado de las funciones continuas de X en \mathbb{R} con la norma suprema, $\mathcal{K}(X)$ el espacio de las funciones continuas y de soporte compacto ($f \in \mathcal{K}(X)$, si $f \in \mathcal{C}(X)$ y f se anula en el complemento de un subconjunto compacto de X) y finalmente $\mathcal{C}_0(X)$ es el espacio de las funciones continuas que se anulan en el infinito. Por las propiedades del soporte compacto ($\text{sop}(f + g) \subseteq \text{sop}(f) \cup \text{sop}(g)$, $\text{sop}(fg) \subseteq \text{sop}(f) \cap \text{sop}(g)$) se tiene que $\mathcal{K}(X)$ es un subespacio vectorial de $\mathcal{C}(X)$, más aún $\mathcal{K}(X) \subset \mathcal{C}_0(X)$ y $\mathcal{K}(X)$ es denso en $\mathcal{C}_0(X)$. Por otro lado, si X es compacto, entonces $\mathcal{K}(X) = \mathcal{C}_0(X)$. Si el espacio topológico (X, τ) no es compacto, entonces podemos tomar la compactificación por un punto X^* y tomar los espacios correspondientes mencionados arriba. Una motivación importante para definir una medida de Radon es la proposición siguiente.

Proposición 4.1. *Si μ es un funcional lineal positivo en $\mathcal{K}(X)$, entonces para cada $K \subset X$ existe una constante M_K tal que $|\mu(f)| \leq M_K \|f\|$ para todo $f \in \mathcal{K}(X)$ con $\text{sop}(f) \subset K$.*

Demostración. Por el lema de Urysohn, podemos encontrar una función continua $g : X \rightarrow [0, 1]$ tal que $g(x) = 1$ para todo $x \in K$. Si $f \in \mathcal{K}(X)$ con $\text{sop}(f) \subset K$, entonces $|f(x)| \leq \|f\|g(x)$ para todo $x \in X$. Así se tiene

$$-\|f\|g(x) \leq f(x) \leq \|f\|g(x)$$

para toda $x \in X$. Luego $\|f\|g - f \geq 0$ y $\|f\|g + f \geq 0$, y por ser μ lineal y positivo se sigue

$$-\|f\|\mu(g) \leq \mu(f) \leq \|f\|\mu(g),$$

es decir, $|\mu(f)| \leq M_K \|f\|$, donde $M_K = \mu(g)$. □

La definición siguiente resulta ser una generalización natural del resultado anterior cuando estamos tratando funcionales lineales positivos en espacios localmente compactos Hausdorff, pues la constante positiva M_K siempre existe para cada compacto K , utilizando el lema de Urysohn.

Definición 4.2. Un funcional lineal $\mu : \mathcal{K}(X) \rightarrow \mathbb{R}$ es llamado una medida de Radon si para cada K compacto, existe una constante $M_K > 0$ tal que $|\mu(f)| \leq M_K \|f\|$ para toda $f \in \mathcal{K}(X)$ con $\text{sop}(f) \subset K$.

Unas referencia donde se prueban algunas propiedades importantes sobre medidas de Radon son ([6], pag.151 y ([3]). El espacio de todas las medidas de Radon se denota por $\mu(X)$. En nuestro caso nos centramos en el conjunto de las medidas de Radon acotadas $\mu_b(X)$, aquellas que son continuas con respecto a la norma suprema definida por

$$\|\mu\| = \sup_{\|f\| \leq 1} |\mu(f)|.$$

Además, una medida de Radon μ es positivo si $\mu(f) \geq 0$ para toda $f \geq 0$ con $f \in \mathcal{K}(X)$, o equivalentemente si $f \leq g$, entonces $\mu(f) \leq \mu(g)$.

Observemos que el conjunto de las medidas de Radon positivas forma un cono lineal y es denotado por $\mu^+(X)$, así se define una norma en $\mu^+(X)$ por restricción de la norma suprema definida en $\mathcal{C}(X)$. Al conjunto de medidas de Radon positivas y acotadas lo denotamos por $\mu_b^+(X)$. Una medida de Radon acotada es una contracción si $\|\mu\| \leq 1$ y $\mu_b^+(X)$ es el conjunto de medidas de Radon positivas con norma 1, llamadas probabilidades.

El ejemplo más simple de medida de Radon en X son las medidas de Dirac. Dado un $x \in X$ la medida de Dirac en x es la medida de Radon ϵ_x definida por

$$\epsilon_x(f) = f(x) \text{ para cada } f \in \mathcal{K}(X). \quad (16)$$

Notemos que $|\epsilon_x(f)| = |f(x)| \leq \|f\|$ para todo $f \in \mathcal{K}(X)$, luego $\|\epsilon_x\| = 1$ y como consecuencia inmediata tenemos que $\epsilon_x \in \mu_b^+(X)$. Más aún como ϵ_x es positivo, si K es compacto, entonces $|\epsilon_x(f)| \leq \epsilon_x(g) \|f\|$ para todo $f \in \mathcal{K}(X)$, con $\text{sop}(f) \subseteq K$ y $\epsilon_x(g) = g(x)$ por la proposición (4.1). Una combinación lineal de medidas de Dirac es llamada medida discreta en X , así una medida discreta es descrita como una medida de Radon acotada en X de la forma

$$\mu = \sum_{i=1}^n \lambda_i \epsilon_{x_i} \text{ donde } n \geq 1, x_1 \dots x_n \in X \text{ y } \lambda_1, \dots, \lambda_n \in \mathbb{R}. \quad (17)$$

Se observa que

$$|\mu(f)| \leq \|\mu\| = \left\| \sum_{i=1}^n \lambda_i \epsilon_{x_i} \right\| = \sum_{i=1}^n |\lambda_i| \|f(x_i)\| \leq \|f\| \sum_{i=1}^n |\lambda_i|,$$

luego afirmamos que μ es positiva si y sólo si $\lambda_i \geq 0$ para $i = 0, \dots, n$, más aún se tiene $\|\mu\| = \sum_{i=1}^n |\lambda_i|$.

Una forma de construir medidas de Radon a partir de una de ellas es la siguiente. Sea $\mu \in \mu(X)$ y consideremos $g \in \mathcal{C}(X)$, entonces el funcional lineal $\nu : \mathcal{K}(X) \rightarrow \mathbb{R}$ definido por

$$\nu(f)(x) := \mu(f \cdot g) \text{ para cada } f \in \mathcal{K}(X) \quad (18)$$

es una medida de Radon en X .

Se puede ver por medio de la densidad de $\mathcal{K}(X)$ en $\mathcal{C}(X)$ que si $\mu \in \mu_b(X)$ (respectivamente $\mu \in \mu_b^+(X)$), entonces μ se puede extender (de manera única) a un funcional continuo (respectivamente positivo) $\tilde{\mu}$ sobre $\mathcal{C}_0(X)$.

Si X es compacto, entonces toda medida de Radon es acotada, es decir, $\mu(X) = \mu_b(X)$, y por tanto también se sigue que $\mu^+(X) = \mu_b^+(X)$. Existen otras propiedades de las medidas de Radon que operacionalmente son importantes como es la desigualdad de Cauchy-Schwarz y que se pueden generalizar sobre sublátices vectoriales de funcionales lineales arbitrarias, es decir, si $\mu \in \mu^+$, entonces se sigue que $(\mu(|fg|))^2 \leq \mu(f^2)\mu(g^2)$ para toda $f, g \in \mathcal{K}(X)$.

Definición 4.3. Sea $\mu \in \mu(X)$, decimos que la medida de Radon μ es cero en un subconjunto $U \subset X$, si $\mu(f) = 0$ para toda $f \in \mathcal{K}(X)$ con $Sop(f) \in U$. Denotamos por $\mathcal{U}(\mu)$ al conjunto $\{U \subset X \mid \mu \text{ es cero en } U \text{ con } U \text{ abierto en } X\}$, así el Soporte de μ es definido por el subconjunto $Sop(\mu) = X - \bigcup_{U \in \mathcal{U}(\mu)} U$.

Notemos de la definición que $Sop(\mu)$ es un subconjunto cerrado, más aun, es el complemento del mayor subconjunto abierto en X para el cual μ es cero.

Observación 4.4.

- i) Si $x \in Sop(\mu)$, entonces para cada vecindad V de x existe una $f \in \mathcal{K}(X)$ tal que $Sop(f) \subset V$ y $\mu(f) \neq 0$, inversamente $x \notin Sop(\mu)$ si existe una vecindad abierta de V de x en la cual μ es cero.
- ii) $Sop(\mu) = \emptyset$ si y sólo si $\mu = 0$.

El siguiente teorema muestra algunas propiedades importantes que se utilizan en lo sucesivo (la prueba se puede ver en ([1], pag.23)).

Teorema 4.5. Si X es un espacio localmente compacto Hausdorff y $\mu \in \mu(X)$, entonces

- i) $Sop(\mu + \nu) \subset Sop(\mu) \cup Sop(\nu)$ para toda μ, ν . La igualdad se da si μ, ν son positivos.

- ii) Si $f \in \mathcal{K}(X)$ y $f = 0$ en el $Sop(\mu)$, entonces $\mu(f) = 0$. Si $\mu \in \mu_b(X)$, entonces también es válido siempre que $f \in \mathcal{C}_0(X)$.
- iii) Si $f, g \in \mathcal{K}(X)$ y $f = g$ sobre $Sop(\mu)$, entonces $\mu(f) = \mu(g)$.
- iv) Si $\mu \in \mu^+(X)$ y $f \in \mathcal{K}(X)$, entonces $\mu(f) \geq 0$ si $f \geq 0$ sobre $Sop(\mu)$.
- v) Si $\mu \in \mu^+(X)$, $f \in \mathcal{K}(X)$, $f \geq 0$ y si $\mu(f) = 0$, entonces $f = 0$ sobre $Sop(\mu)$.
- vi) Para cada $g \in \mathcal{C}(X)$, $Sop(g, \mu) = \overline{\{x \in Sop(\mu) | g(x) \neq 0\}} \subset Sop(g) \cap Sop(\mu)$.
- vii) Si x_1, \dots, x_n son puntos distintos de X para $n \geq 1$, entonces $Sop(\mu) = \{x_1, \dots, x_n\}$ si y sólo si $\mu = \sum_{i=1}^n \lambda_i \epsilon_{x_i}$ para algunos $\lambda_1, \dots, \lambda_n \in \mathbb{R} - \{0\}$.
- viii) Si $Sop(\mu)$ es compacto, entonces μ es acotada.

Sea I un conjunto dirigido, se define una red de medidas de Radon por $(\mu_i)_{i \in I}^{\leq}$. Decimos que $(\mu_i)_{i \in I}^{\leq}$ converge con el orden \leq , si para cada $f \in \mathcal{K}(X)$ se tiene que

$$(\mu_i(f))_{i \in I}^{\leq} \rightarrow \mu(f),$$

es decir, dado $U_\mu(f)$ abierto, existe un $i_0 \in I$ tal que $\mu_i(f) \in U_\mu(f)$ para todo $i \geq i_0$ y en este sentido decimos que μ_i converge a μ vagamente y lo denotamos por

$$\lim_{i \in I} \mu_i(f) = \mu(f)$$

(vea ([1], pag.46)). Dada una medida de Radon $\mu \in \mu_b^+$, se define para un $H \subset \mathcal{C}_0(X)$ el conjunto

$$D_+(H, \mu) = \{f \in \mathcal{C}_0(X) \mid \text{para cada red } (\mu_i)_{i \in I}^{\leq} \in \mu_b^+ \text{ que cumplen}$$

$$\lim_{i \in I} \mu_i(f) = \mu(f) \text{ siempre que } \sup_{i \in I} \|\mu_i\| < \infty \text{ y}$$

$$\lim_{i \in I} \mu_i(h) = \mu(h) \text{ para toda } h \in H\}.$$

Es claro que $H \subset D_+(H, \mu) \subset \mathcal{C}_0(H)$. Para nosotros es de interés saber cuando $D_+(H, \mu) = \mathcal{C}_0(H)$, en tal caso H es llamado un D_+ -subconjunto de Korovkin con respecto a una medida de Radon positiva y acotada. Por lo anterior se desea conocer la hipótesis sobre H que asegure la igualdad. Un subconjunto de $\mathcal{C}_0(X)$ que está muy relacionada con $D_+(H, \mu)$ es el definido por

$$U_+(H, \mu) = \{f \in \mathcal{C}_0(X) \mid v(f) = \mu(f) \text{ para cada } v \in \mu_b^+ \text{ y}$$

$$\text{se cumple que } v(h) = \mu(h) \text{ para toda } h \in H\}.$$

El siguiente teorema es importante porque caracteriza a los D_+ -subconjunto de Korovkin.

Teorema 4.6. *Sea H un subespacio de $\mathcal{C}_0(X)$, entonces $D_+(H, \mu) = \mathcal{U}_+(H, \mu)$ siempre que $\mu \in \mu_b^+(X)$.*

Demostración. Es claro que $D_+(H, \mu) \subset \mathcal{U}_+(H, \mu)$, ya que si $f \in D_+(H, \mu)$ y $v \in \mu_b^+$ satisface $v|_H = \mu|_H$, se define la red $(\mu_i)_{i \in I} \stackrel{\leq}{\subset}$ con $\mu_i = v$ y es claro que $\sup_{i \in I} \|\mu_i\| = \|v\|$ y $\lim_{i \in I} \mu_i(h) = v(h) = \mu(h)$ para toda $h \in H$ lo cual asegura que

$$\lim_{i \in I} \mu_i(f) = \mu(f) = v(f),$$

entonces $f \in \mathcal{U}(H, \mu)$. Inversamente, supongamos que $f \in \mathcal{U}_+(H, \mu)$ y $(\mu_i)_{i \in I} \stackrel{\leq}{\subset}$ es una red de medidas de Radon positivas y acotadas que cumplen que $\sup_{i \in I} \|\mu_i\| < +\infty$ y $\lim_{i \in I} \mu_i(h) = \mu(h)$ para todo $h \in H$ con $\lim_{i \in I} \mu_i(f) \neq \mu(f)$. Existe $\epsilon > 0$ tal que para toda $i \in I$ existe $\alpha(i) > i$ con la propiedad $|\mu_i - \mu(f)| \geq \epsilon$. Por otro lado, la familia $A = \{\mu_{\alpha(i)}\}$ es fuertemente acotada, es decir, para todo K compacto, existe $M_K \geq 0$ tal que $|\mu_i(f)| \leq M_K \|f\|$ para toda $\mu_{\alpha(i)}$ y toda $f \leq k$. Como es fuertemente acotada, entonces A es relativamente compacto, por tanto en A , existe una subfamilia $(\mu_r)_{r \in I}$ con un filtro \mathcal{F} más fino que el filtro de las secciones finales del conjunto de índices I tal que

$$\lim_{r \in I^{\mathcal{F}}} \mu(h) = \lim_{i \in I} \mu_i(h) = \mu(h)$$

para toda $h \in H$. Por tanto se sigue que $\mu = v$ sobre $\mathcal{U}_+(H, \mu)$ y en particular se tiene que $\mu(f) = v(f)$ lo cual es una contradicción ya que $|\mu_{\alpha(i)}(f) - \mu(f)| \geq \epsilon$ es equivalente a $|\mu_{\alpha(i)}(f) - v(f)| \geq \epsilon$. □

Definición 4.7. Un subespacio H de $\mathcal{C}_0(X)$ es llamado cofinal si para cada $f \in \mathcal{C}_0(X)$, existe $h \in H$ con $f \leq h$. Notemos que es equivalente a la existencia de un $h_0 \in H$ tal que $0 < h_0(x)$ para todo $x \in X$.

Teorema 4.8. *Si X es compacto y H es cofinal en $\mathcal{C}(X)$, entonces para cada $f \in \mathcal{C}(X)$ los siguientes son equivalentes.*

- i) $f \in \mathcal{U}_+(H, \mu)$.
- ii) $\sup_{h \in H, h \leq f} \mu(h) = \mu(f) = \inf_{k \in H, k \geq f} \mu(k)$.
- iii) Para cada $\epsilon > 0$ existe un $h, k \in H$ con $h \leq f \leq k$ y $\mu(k) - \mu(h) < \epsilon$.

Demostración. Observemos que ii) y iii) son equivalentes por definición de ínfimo y supremo. Sea $\epsilon/2 > 0$, entonces existen $h, k \in H$, con $h \leq f \leq k$ tal que $\mu(f) < \epsilon/2 + \mu(h)$ y $\mu(k) < \mu(f) + \epsilon/2$ de donde se sigue que $\mu(k) - \mu(h) < \epsilon$.

Demostremos que *i*) es equivalente a *ii*). Definimos $p : \mathcal{C}(X) \longrightarrow \mathbb{R}$ por

$$p(g) = \inf_{\substack{k \in H \\ k \geq g}} \mu(k).$$

Notemos que p cumple $p(g + f) \leq p(g) + p(f)$ y $p(\alpha g) = \alpha p(g)$ para todo $f, g \in \mathcal{C}(X)$ y $\alpha \geq 0$, es decir, p es sublineal. Aplicando el Teorema de Hahn-Banach, existe un funcional lineal v sobre $\mathcal{C}(X)$ tal que $v \leq p$ y coincide con p en f , $v(f) = p(f)$. Más aún, v es positivo. Si $g \in \mathcal{C}(X)$ con $0 \leq g$, entonces $v(-g) \leq p(-g) \leq \mu(0) = 0$, ya que $-g \leq 0$ y por ser μ y v funcionales lineal, luego $v \in \mu^+(X)$. Más aún, $v = \mu$ sobre H puesto que $p = \mu$ sobre H . Así

$$\mu(f) = v(f) = p(f) = \inf_{\substack{k \in H \\ k \geq f}} \mu(k),$$

por *i*). Para completar esta parte de la demostración, se tiene que $-f \in U_+(H, \mu)$, por lo que $\mu(-f) = p(-f)$, o bien,

$$\mu(f) = -p(-f) = \sup_{\substack{h \in H \\ h \leq f}} \mu(h).$$

Para ver que *ii*) implica *i*), si $v \in \mu^+(X)$ y $v = \mu$ sobre H , entonces de las desigualdades

$$\sup_{\substack{h \in H \\ h \leq f}} v(h) \leq v(f) \leq \inf_{\substack{k \in H \\ f \leq k}} v(k)$$

y por hipótesis se sigue que $v(f) = \mu(f)$. □

El siguiente teorema también caracteriza a los D_+ -subconjuntos de Korovkin con respecto a una medida de Radon positiva y acotada.

Teorema 4.9. *Sea X compacto y H subespacio de $\mathcal{C}(X)$. Los siguientes enunciados son equivalentes.*

- i)* H es un D_+ -subespacio para $\mu \in \mu_b^+(X)$.
- ii)* H es cofinal en $\mathcal{C}(X)$ y $\sup_{\substack{h \in H \\ h \leq f}} \mu(h) = \mu(f) = \inf_{\substack{k \in H \\ k \geq f}} \mu(k)$ para cada $f \in \mathcal{C}(X)$.
- iii)* H es cofinal en $\mathcal{C}(X)$ y para toda $f \in \mathcal{C}(X)$ y $\epsilon > 0$ existen funciones $h, k \in H$ tales que $h \leq f \leq k$ y $\mu(k) - \mu(h) < \epsilon$.

5 Conjuntos y cerradura de Korovkin

La aproximación tipo Korovkin se desarrolla en medio de los conceptos y los resultados obtenidos en la convergencia de las redes de operadores lineales positivos. La caracterización obtenida en relación con la convergencia al operador identidad, tiene una subsecuente estimulación a la investigación sobre otra clase de operadores lineales. Esta sección solo menciona los conceptos básicos y algunos resultados para entender la relación entre los espacios de Chebyshev y los subespacios de Korovkin en el espacio de funciones continuas.

Definición 5.1. Un subconjunto H de $\mathcal{C}_0(X)$ es llamado subconjunto de Korovkin con respecto a un operador lineal positivo T , si se satisface la siguiente propiedad.

Si para toda red arbitraria $(L_i)_{i \in I}^{\leq}$ de operadores lineales positivos de $\mathcal{C}_0(X)$ en $\mathcal{C}_0(Y)$ tal que $\sup_{i \in I} \|L_i\| < \infty$ y $\lim_{i \in I} L_i(h) = T(h)$ para toda $h \in H$, entonces $\lim_{i \in I} L_i(f) = T(f)$ para toda $f \in \mathcal{C}_0(X)$.

Por notación decimos que H es K_+ – subconjunto de Korovkin para T . (19)

Observemos que la equicontinuidad de la red de operadores lineales positivos está descrita por la condición

$$\sup_{i \in I} \|L_i\| < \infty.$$

Si T es un operador lineal positivo tal que $\|T\| < 1$, entonces se llama contracción y si H satisface la propiedad (19) de la definición (5.1), sólo para redes $(L_i)_{i \in I}^{\leq}$ de operadores lineales positivos que son contracciones, entonces H es llamado subconjunto de Korovkin para T con respecto a la contracción lineal positiva y lo denotamos por K_+^1 – subconjunto de Korovkin para T . Notemos que H es un subconjunto de Korovkin de $\mathcal{C}_0(X)$ si y sólo si el subespacio lineal generado $\mathcal{L}(H)$ por H , es un conjunto de Korovkin, así podemos trabajar con la definición (5.1), indistintamente si es conjunto o subespacio. En adelante H es K_+ – subespacio de Korovkin para T . Análogamente, también se define el K_+^1 – subespacio para T .

Notemos que si el espacio X es compacto y $1 \in H$, entonces la red automáticamente es equicontinua bajo la definición (19), pues como X es compacto, entonces $\mathcal{C}_0(X) = \mathcal{C}(X)$ y por tanto la convergencia en $\mathcal{C}_0(X)$ está determinado por las bolas de la norma suprema, luego si $1 \in H$, entonces $\lim_{i \in I} L_i(h) = T(h)$, para toda $h \in H$ y en particular $\lim_{i \in I} L_i(1) = T(1)$, así $L_i \in B_\epsilon(T(1))$ para toda $i_0 \leq i$, es decir, $\|L_i(1) - T(1)\| < \epsilon$ para toda $i_0 \leq i$, por tanto se sigue que $\|L_i\| < \epsilon + \|T(1)\|$, además como X es compacto se tiene $\|L_i(1)\| = \|L_i\|$ para todo $i \in I$, por tanto $\sup_{i \in I} \|L_i\| < \infty$, es decir, si X es compacto y $1 \in H$, entonces no se necesita pedir $\sup_{i \in I} \|L_i\| < \infty$ de la definición (5.1). Los subespacios que introducimos a continuación son importantes en la caracterización de los subespacios de Korovkin.

Definición 5.2. Se le llama clausura de Korovkin de H con respecto al operador lineal positivo T o K_+ - *clausura* de H para T , al subconjunto

$$K_+(H, T) = \{f \in \mathcal{C}_0(X) \mid \lim_{i \in I} L_i(f) = T(f) \text{ para cada red } (L_i)_{i \in I}^{\leq} \text{ de operadores lineales de } \mathcal{C}_0(X) \text{ en } \mathcal{C}_0(Y) \text{ que satisface la condición } \sup_{i \in I} \|L_i\| < \infty \text{ y } \lim_{i \in I} L_i(h) = T(h) \text{ para todo } h \in H\}.$$

Notemos que al igual que los conjuntos de Korovkin, la K_+ - *clausura* con respecto a T de un conjunto H de $\mathcal{C}_0(X)$ coincide con el K_+ - *clausura* para T del subespacio generado por H . Uno de los resultados importantes que podemos mencionar es que $K_+(H, T) = \mathcal{C}_0(X)$ siempre y cuando H sea K_+ - *subespacio* de Korovkin para T . Análogamente se define el subespacio K_+^1 - *clausura* de H para la contracción T con la definición (5.2) tomando sólo operadores lineales positivos que son contracciones. El primer resultado importante que caracteriza tanto la K_+ - *clausura* y a la K_+^1 - *clausura* de un conjunto con respecto a un operador lineal positivo $T : \mathcal{C}_0(X) \rightarrow \mathcal{C}_0(Y)$ consiste en demostrar la relación existente entre estos subespacios y la convergencia de medidas de Radon considerando para cada $y \in Y$, $\mu_y^T : \mathcal{C}_0(X) \rightarrow \mathbb{R}$ definido por $\mu_y^T(f) = Tf(y)$ para cada $f \in \mathcal{C}_0(X)$. Así $\mu_y^T = \epsilon_y \circ T$. Más aún, si H es un subespacio de $\mathcal{C}_0(X)$, entonces

$$K_+(H, T) = \bigcap_{y \in Y} U_+(H, \mu_y^T) = \bigcap_{y \in Y} D_+(H, \mu_y^T),$$

esto muestra la relación que existe entre los subconjunto de Korovkin con respecto a un operador lineal positivo T y los D_+ - subconjuntos de Korovkin con respecto a una medida de Radon positiva y acotada, donde primero se analiza para el caso de contracciones y después se puede generalizar para operadores lineales positivos.

Definición 5.3. Sean X y Y dos espacios localmente compactos Hausdorff y $T : \mathcal{C}_0(X) \rightarrow \mathcal{C}_0(Y)$ un operador lineal positivo. Decimos que T es finitamente definido de orden n , si existen n funciones $\phi_1, \dots, \phi_n : Y \rightarrow X$ y n funciones reales $\psi_1, \dots, \psi_n : Y \rightarrow \mathbb{R}$ tales que

$$T(f) = \sum_{i=1}^n \psi_i \cdot (f \circ \phi_i) \text{ para cada } f \in \mathcal{C}_0(X). \tag{20}$$

Se asume que el lado derecho de la ecuación (20) de la definición anterior pertenece a $\mathcal{C}_0(Y)$, bajo cualquiera de las siguientes hipótesis

- a) Si Y es compacto, entonces suponer que $\phi_1, \dots, \phi_n, \psi_1, \dots, \psi_n$ son continuas.

- b) Si Y no es compacto, entonces suponer que $\psi_1, \dots, \psi_n \in \mathcal{C}_0(Y)$ y ϕ_1, \dots, ϕ_n son continuas.

Así en ambos casos T está bien definido.

En adelante nosotros denotamos por $\mathcal{F}_n(X, Y)$ el conjunto de todos los operadores finitamente definidos de orden n de $\mathcal{C}_0(X)$ en $\mathcal{C}_0(Y)$ y por $\mathcal{F}_n^1(X, Y)$ el conjunto de todos los operadores finitamente definidos de orden n de $\mathcal{C}_0(X)$ en $\mathcal{C}_0(Y)$ para el cual admite una representación de la forma de la ecuación (20) y $\sum_{i=1}^n \psi_i = 1$. Observemos que $\mathcal{F}_n(X, Y) \subset \mathcal{F}_{n+1}(X, Y)$ y $\mathcal{F}_n^1(X, Y) \subset \mathcal{F}_{n+1}^1(X, Y)$, para cada n , más aún, un operador finitamente definido $T : \mathcal{C}_0(X) \rightarrow \mathcal{C}_0(Y)$ de orden 1 es de la forma $T(f) = \psi(f \circ \phi)$ para cada $f \in \mathcal{C}_0(X)$ para algún $\phi : Y \rightarrow X$ y una función real positiva $\psi : Y \rightarrow \mathbb{R}$.

Los operadores finitamente definidos están relacionados con las medidas de Radon positivas discretas, es decir, si $T : \mathcal{C}_0(X) \rightarrow \mathcal{C}_0(Y)$ es un operador finitamente definido de orden n con una representación (20), entonces tenemos

$$\mu_y^T(f) = Tf(y) = \sum_{i=1}^n \psi_i(y) \cdot f(\phi_i(y)) = \sum_{i=1}^n \psi_i(y) \cdot \epsilon_{\phi_i(y)}(f)$$

para cada $y \in Y$ y toda $f \in \mathcal{C}_0(X)$, luego entonces la medida

$$\mu_y^T = \sum_{i=1}^n \psi_i(y) \cdot \epsilon_{\phi_i}$$

es una medida de Radon positiva y discreta para cada $y \in Y$.

Proposición 5.4. Sean $T : \mathcal{C}_0(X) \rightarrow \mathcal{C}_0(Y)$ un operador lineal positivo y H un subespacio de dimensión n en $\mathcal{C}_0(X)$. Supongamos que X no es compacto y que H separa puntos de X . Si para cada $x \in X$, existe un $h \in H$ tal que $h(x) \neq 0$, entonces existen $n + 1$ funciones $\phi_1, \dots, \phi_{n+1} : Y \rightarrow X$ y $n + 1$ funciones reales positivas $\psi_1, \dots, \psi_{n+1} : Y \rightarrow \mathbb{R}$ tales que

$$T(h) = \sum_{i=1}^{n+1} \psi_i \cdot (h \circ \phi_i) \text{ para cada } h \in H.$$

En consecuencia, si H es un K_+ -subespacio para T (o un K_+^1 -subespacio, si X es compacto y $\|T\| \leq 1$), entonces T es finitamente definido de orden $n + 1$. Finalmente, si X y Y son ambos compactos y $T(1) = 1$ y si H es un K_+^1 -subespacio para T , entonces $T \in \mathcal{F}_{n+1}^1(X, Y)$.

Definición 5.5. Sea un espacio localmente compacto Hausdorff. Un subconjunto H de $\mathcal{C}_0(X)$ se dice K_+ -subconjunto de orden n en $\mathcal{C}_0(X)$, si es un K_+ -subconjunto de T para cada espacio Y localmente compacto Hausdorff y cada operador finitamente definido $T \in \mathcal{F}_n(X, Y)$. Si además H es un subespacio de $\mathcal{C}_0(X)$, entonces se dice K_+ -subespacio de orden n en $\mathcal{C}_0(X)$.

Notemos que cada K_+ -subespacio de orden n en $\mathcal{C}_0(X)$ es un K_+ -subespacio de orden p con $1 \leq p \leq n$. Finalmente, cuando X es compacto se tiene que $\mathcal{C}(X) = \mathcal{C}_0(X)$ y tenemos el resultado siguiente que relaciona la teoría de Korovkin con la teoría de Chebyshev.

Teorema 5.6. Sea X el intervalo $[a, b] \subset \mathbb{R}$. Si H es un subespacio de $\mathcal{C}(X)$ de dimensión $n + 1$, entonces se tiene las siguientes afirmaciones.

- i) Si n es par, entonces H es un subespacio de Chebyshev de orden $n + 1$ en $\mathcal{C}(X)$ si y sólo si H es un K_+ -subespacio de orden $\frac{n}{2}$ en $\mathcal{C}(X)$.
- ii) Si n es impar y H es un subespacio de Chebyshev de orden $n + 1$ en $\mathcal{C}(X)$, entonces H es un K_+ -subespacio de orden $\frac{n-1}{2}$ en $\mathcal{C}(X)$.

Demostración. Sean los dos casos

$$r_n = \begin{cases} \frac{n}{2} & \text{si } n \text{ es par} \\ \frac{n-1}{2} & \text{si } n \text{ es impar} \end{cases}$$

entonces para cada conjunto de puntos diferentes $\{x_1, x_2, \dots, x_{r_n}\}$ de X , siempre se tiene que $\sum_{i=1}^{r_n} \omega(x_i) \leq n$, donde

$$\omega(x) = \begin{cases} 2 & \text{si } x \in (a, b), \\ 1 & \text{si } x = a, b. \end{cases}$$

Luego (vea el teorema 2.3.5 de [1], pag.104 y el teorema 3.4.7, ii) de [1], pag.184), para cada $\mu \in C_+(x_1, x_2, \dots, x_{r_n}) = \{\mu \in \mu_b^+(X) \mid \text{existen } \alpha_1 \dots \alpha_n \in \mathbb{R}_+ \text{ tales que } \mu = \sum_{i=1}^n \alpha_i \epsilon_{x_i}\}$. Se tiene que H es un D_+ -subespacio si y sólo si H es un K_+ -subespacio de orden r_n en $\mathcal{C}(X)$. Ahora probemos el inverso del inciso a). Sea H el subespacio generado por las funciones linealmente independiente $\{h_0, \dots, h_n\}$. Si H no es un subespacio de Chebyshev de orden n en $\mathcal{C}(X)$, entonces existen $n + 1$ puntos distintos $\{x_0, x_1, \dots, x_n\}$ de X y $n + 1$ números $\alpha_0, \dots, \alpha_n$ no todos ceros tales que $\sum_{i=0}^n \alpha_i h_j(x_i) = 0$ para cada $j = 0, \dots, n$. Ahora definimos el conjunto $J = \{\text{todos los } i \text{ tal que } \alpha_i > 0\}$ y observando que

$$\sum_{i=1}^n \alpha_i h_j(x_i) = \sum_{i \in J} \alpha_i \epsilon_{x_i}(h_j) + \sum_{i \notin J} \alpha_i \epsilon_{x_i}(h_j) = 0,$$

notemos que podemos suponer que la cardinalidad de J es menor que $\frac{n}{2}$. Si $J = \emptyset$, entonces $\mu = 0$, luego definimos la medida de Radon positiva como $\mu = \sum_{i \in J} \alpha_i \epsilon_{x_i}$ y $v = -\sum_{i \notin J} \alpha_i \epsilon_{x_i}$. Por tanto, como $\mu(h_j) = v(h_j)$ para $j = 0, \dots, n$, entonces $\mu = v$ sobre H y por el teorema 3.4.7, ii), (vea [1], pag.184) se sigue que $\mu = v$ sobre $\mathcal{C}(X)$ y por tanto es una contradicción, pues todos los puntos $\{x_0, x_1, \dots, x_n\}$ son distintos. \square

Los siguientes ejemplos ilustran los sistemas de Chebyshev.

Ejemplo 5.7.

- i) $1, x, \dots, x^n$ en $[a, b]$.
- ii) $1, \cos x, \dots, \cos nx$ en $[0, \pi]$.
- iii) $1, \sin x, \dots, \sin nx$ en $(0, \pi)$.
- iv) $e^{a_i x}$ con $a_1 < \dots < a_n$ en $[a, b]$.
- iv) x^{a_i} con $a_1 < \dots < a_n$ en $[a, b]$.

Así por ejemplo si n es par, el subespacio de Chebyshev generado por el sistema $\{1, x, \dots, x^n\}$ en $[a, b]$ es de orden $n+1$ si y sólo si es un K_+ - subespacio de orden $\frac{n}{2}$. Si n es impar, entonces el subespacio de Chebyshev de orden $n+1$ implica que es un K_+ - subespacio de orden $\frac{n-1}{2}$.

Concluimos este escrito, puntualizando la importancia que presenta el último resultado. Probar que un conjunto de funciones es un sistema de Chebyshev es relativamente fácil utilizando la teoría de Chebyshev, específicamente con la proposición (3.2) y, siendo el conjunto generado por este sistema un conjunto de Korovkin se pueden formar, mediante la teoría de aproximación tipo Korovkin, procesos de aproximación adecuados, como por ejemplo por los operadores de Bernstein para aproximación por polinomios algebraicos o por medio de núcleos, como los de Fejer o los de Jackson para el caso de aproximación trigonométrica, que nos proporcionen aproximaciones de las funciones que se requiera aproximar.

Agradecimientos

Agradecemos a la Universidad Tecnológica de la Mixteca el apoyo para llevar a cabo este trabajo, así como a la Benemérita Universidad Autónoma de Puebla por permitirnos presentarlo en el congreso CIMA.

Bibliografía

- [1] Altomare F. and Campiti M. *Korovkin-type Approximation Theory and its Applications*, Walter de Gruyter Studies in Mathematics 17 (1994).
- [2] Altomare F. *Korovkin-type Theorems and Approximation by Positive Linear Operators*. Surveys in Approximation Theory, Volume 6 (2010). pp. 92-164.
- [3] Bernd Anger, Claude Portenier. *Radon Integrals Theory*, Volume 6 (2010). Progress in Mathematics 103,(1992).
- [4] Bauer H. *Theorems of Korovkin type for adapted spaces* (French summary), Ann. Inst. Fourier (Grenoble) 23 (1973), fasc. 4, 245-260. MR 50 # 10643. SC 5.2.d, 5.5.d, 5.6.d.
- [5] Cheney E. W. *Introduction to Approximation Theory*, McGraw-Hill, USA (1966).
- [6] Gustave Choquet. *Lectures on Analysis*, W.A. Benjamin, INC.Advanced Book Program, volumen 1(1969).
- [7] DeVore R. A. and G. G. Lorentz. *Constructive Approximation*, Springer-Verlag, New York, (1993).
- [8] Royden, H.L. *Real analysis*, 2nd MacMillan, New York, (1968).(Russian).

Instituto de Física y Matemáticas,
Universidad Tecnológica de la Mixteca
Km. 2.5, carretera a Acatlima,
Huajuapán, Oaxaca. C.P. 69000

jmh@mixteco.utm.mx
pacheco@mixteco.utm.mx

Filosofía, Divulgación e Historia de las Matemáticas

Capítulo 6

El descubrimiento de la inconmensurabilidad por Hipaso de Metaponto

Agustín Contreras Carreto, Elizabeth de Gante Coronel,
María del Rocío Macías Prado
FCFM, BUAP

Resumen

En este capítulo hablaremos de la primera gran crisis en la historia de la matemática y, de hecho, de la ciencia o, más apropiadamente, de la filosofía. Se trata del descubrimiento de la existencia de segmentos inconmensurables. El origen exacto de tal descubrimiento está rodeado de espesa bruma y ha generado una activa controversia entre los diferentes puntos de vista de los historiadores de la matemática acerca de esta cuestión. Aquí ofreceremos sólo uno de dichos puntos de vista y trataremos de explicar someramente por qué el descubrimiento de la inconmensurabilidad representó una gran crisis en la Escuela Pitagórica.

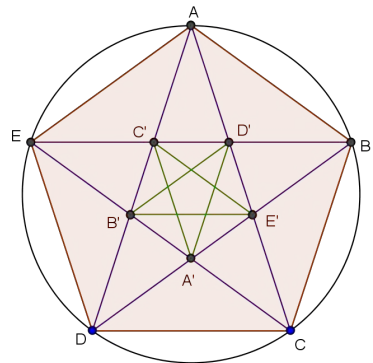
1 Introducción

En una mañana del año 480 antes de Cristo los gritos de un niño rompieron la tranquilidad de Crotona. Anunciaban que el gran Hipaso de Metaponto había muerto en un naufragio ocurrido en la noche anterior. Era muy conocido en esa ciudad por su gran inventiva: hacía música con placas de metal de diferentes grosores y de distintos diámetros, y construía poliedros regulares como nunca antes se habían visto. Pocos sabían que precisamente su genial cerebro lo había convertido en un peligroso subversivo y que ese naufragio no fue casualidad: durante varias semanas antes había estado cabizbajo y no quería asistir a las reuniones de La Hermandad. Los otros miembros de la secta sospechaban que el filósofo de Metaponto seguramente había descubierto algo muy importante, pero sólo recurriendo a los extraordinarios encantos de una irresistible y bella dama, quien embriagó y sedujo al sabio, fue que obtuvieron el gran secreto, una verdadera blasfemia contra el inolvidable Maestro Pitágoras: Hipaso demostró que existían por lo menos dos segmentos que no son conmensurables: la diagonal y el lado de un mismo pentágono regular.

2 Consideraciones preliminares acerca de la conmensurabilidad y de la inconmensurabilidad

La introducción a este capítulo es un tanto ficticia, un ardid para interesar al lector acerca de la figura de Hipaso de Metaponto (en griego clásico: Ἱππασος Μεταποντῖνος) pero no está tan alejada de las investigaciones históricas sobre este personaje, del cual se conoce realmente muy poco. Morris Kline menciona en [6] que, según una vieja leyenda, los pitagóricos viajaban por mar y que uno de ellos, Hipaso de Metaponto, fue arrojado por la borda por haber encontrado un elemento del universo que contradecía la doctrina pitagórica de que todos los fenómenos del universo se pueden reducir a números enteros. Por otro lado, Howard Eves señala en [3] que, de acuerdo con una versión, Hipaso fue desterrado de la comunidad pitagórica y fue erigida una tumba con su nombre como si él hubiera muerto. Kurt von Fritz, en su interesante artículo [8], ofrece plausibles argumentos para afirmar que el descubrimiento de la inconmensurabilidad fue hecho por Hipaso de Metaponto alrededor del año 450 a. C., y no en el cuadrado, sino en el pentágono regular. En otras palabras, von Fritz piensa que los primeros dos segmentos inconmensurables descubiertos en la historia de la Humanidad, no fueron la diagonal y el lado de un cuadrado, sino la diagonal y el lado de un pentágono regular, lo cual equivale, como veremos posteriormente, a que el primer número irracional no fue $\sqrt{2}$, sino el número áureo $\phi = (1 + \sqrt{5})/2$. Según Von Fritz el pentágono regular era, para los pitagóricos, una figura más misteriosa que el cuadrado.

Las diagonales de un cuadrado se cortan en un punto, pero si se intersecan todas las diagonales del pentágono, se formará un pentágono más pequeño que también será regular, y si se trazan ahora las diagonales de éste, se obtendrá un pentágono regular más pequeño aún, y así sucesivamente, siguiendo esta construcción en un proceso de apariencia interminable o que involucra el infinito. En una especie de operación inversa, si comenzamos con un pentágono regular cualquiera y prolongamos sus lados lo suficiente para que se intersequen formando una estrella, al unir las puntas de dicha estrella se obtendrá un pentágono regular más grande, cuyas diagonales son los lados de la estrella. Todo esto lo conocían los pitagóricos desde tiempos de Pitágoras, es decir, desde el siglo VI a. C. y les fascinaba tanto que la hermosa estrella pentagonal fue un símbolo distintivo de la Hermandad Pitagórica. Aún ahora se conoce a esa figura como la estrella pitagórica y sigue siendo



un símbolo famoso entre los esotéricos. Es allí, dice von Fritz, donde debe haber un misterio y no en el sencillo cuadrado, pero se necesitó de una mente audaz para atreverse a develarlo y formular la pregunta clave: bueno y... ¿cuánto mide la diagonal de un pentágono regular usando como unidad su lado?

Para tratar de aclarar los términos usados en esta pregunta, comencemos con una importante definición:

Definición 2.1. Se dice que **dos segmentos** \overline{AB} y \overline{CD} **son conmensurables**, si existe un segmento \overline{EF} que cabe un número entero de veces en \overline{AB} y un número entero de veces en \overline{CD} , es decir, si las longitudes de los segmentos \overline{AB} , \overline{CD} y \overline{EF} son a , b y c , respectivamente, entonces existen números naturales m y n tales que $a = mc$ y $b = nc$. **Los segmentos** \overline{AB} y \overline{CD} **son inconmensurables**, si no son conmensurables.

No es difícil para nosotros, en estos tiempos en que tenemos claro el concepto de número real y sus propiedades, percatarse de que, si dos segmentos \overline{AB} y \overline{CD} son conmensurables, entonces la razón o el cociente de sus longitudes es un número racional. En efecto, si las longitudes de los segmentos \overline{AB} , \overline{CD} y \overline{EF} son a , b y c , respectivamente, y existen números naturales m y n tales que $a = mc$ y $b = nc$, entonces $a/b = mc/nc = m/n$. También se cumple la proposición recíproca: si la razón entre las longitudes de dos segmentos es un número racional, entonces ellos son conmensurables. Para demostrar esta afirmación, supongamos que \overline{AB} y \overline{CD} son segmentos de longitudes a y b , respectivamente, y que a/b es un número racional. Entonces existen números naturales m y n tales que $a/b = m/n$. Sea \overline{EF} un segmento cuya longitud c es la n -ésima parte de b . Entonces $b = nc$ y, por lo tanto, $a = (a/b)b = (m/n)(nc) = mc$, es decir, el segmento \overline{EF} cabe m veces en el segmento \overline{AB} y, por la definición del segmento \overline{EF} , cabe n veces en el segmento \overline{CD} . Por consiguiente, los segmentos \overline{AB} y \overline{CD} son conmensurables.

Resumamos estas reflexiones en un solo enunciado:

Proposición 2.1. *Dos segmentos son conmensurables si y sólo si la razón entre sus longitudes es un número racional.*

Para medir la longitud de un segmento cualquiera necesitamos una unidad de medida. Actualmente podemos usar, según nos convenga, el centímetro, el metro, la pulgada o la yarda, pero a mediados del siglo V a. C. no se habían inventado estas maravillosas unidades. Sin embargo, desde mucho antes de ese siglo, se usaba **el método de sustracción mutua** para medir la longitud de un segmento cualquiera, usando como unidad de medida otro segmento cualquiera. Los griegos lo llamaban **antiféresis**, palabra construida a partir del verbo ἀνθυφαίρεω, que emplea Euclides

en la proposición 1 del libro VII de los *Elementos*, y que está compuesta a su vez por $\upsilon\phi\alpha\iota\rho\epsilon\omega$, que significa “sustraer gradualmente”, y por el prefijo *anti*, usado aquí para designar la idea de dos cosas que se corresponden. La antiféresis es entonces una sustracción recíproca o mutua, como la llamamos antes.

Supongamos que tenemos dos segmentos cualesquiera \overline{AB} y \overline{CD} de longitudes a y b , respectivamente, y que $a > b$. La antiféresis trata de buscar una “unidad de medida común” entre los dos segmentos dados, es decir, un segmento de longitud u que quepa un número entero de veces en a y un número entero de veces en b . La idea fundamental del procedimiento es que, si existe un segmento de longitud u que mida simultáneamente a a y b , mide igualmente a un segmento de longitud $a - b$ (y por tanto a $a - 2b$, $a - 3b$, etc.), y viceversa: si u mide a la vez a b y $a - b$, entonces mide también a a . La primera etapa de la antiféresis consiste en recortar de la longitud a , tantas veces como sea posible, la longitud b , colocando el segmento \overline{CD} sobre el segmento \overline{AB} , una y otra vez hasta que ya no quepa completo en \overline{AB} . Así podemos saber cuántas veces enteras cabe \overline{CD} en el segmento \overline{AB} ; digamos que cabe c_1 veces, donde c_1 es un número natural positivo. Si \overline{CD} cabe exactamente c_1 veces en \overline{AB} , entonces ya podemos expresar la longitud de \overline{AB} en términos de la de \overline{CD} : $a = bc_1$. No sólo hemos medido la longitud de \overline{AB} en función de la de \overline{CD} , sino que hemos hallado un segmento, en este caso \overline{CD} , que cabe un número entero de veces (a saber, c_1) en \overline{AB} y un número entero de veces (una vez) en \overline{CD} . En otras palabras, hemos visto que \overline{AB} y \overline{CD} son conmensurables.

Si \overline{CD} no cabe exactamente c_1 veces en \overline{AB} , entonces, al colocar c_1 veces \overline{CD} sobre el segmento \overline{AB} , es decir, cuando se ha retirado c_1 veces la longitud b a la longitud a , queda un pequeño segmento residual \overline{EF} , de longitud r_1 . En este caso tenemos:

$$a = bc_1 + r_1,$$

donde $0 < r_1 < b$.

Ahora debemos trabajar con los segmentos de longitudes b y r_1 , porque aún estamos buscando nuestra unidad de longitud u . La segunda etapa de la antiféresis es intercambiar el papel de los dos segmentos \overline{CD} y \overline{EF} : puesto que \overline{EF} es el más pequeño, la longitud de éste es la que se va a retirar de b , tantas veces como sea posible: supongamos que \overline{EF} cabe c_2 veces en el segmento \overline{CD} , donde c_2 es un número natural positivo. Si \overline{EF} cabe exactamente c_2 veces en \overline{CD} , entonces ya podemos expresar las longitudes de \overline{CD} y \overline{AB} en términos de la de \overline{EF} :

$$b = r_1c_2 \quad \text{y}$$

$$a = bc_1 + r_1 = r_1c_2c_1 + r_1 = r_1(c_2c_1 + 1).$$

Llamando m_1 a $c_2c_1 + 1$ y n_1 a c_2 , obtenemos $a = m_1r_1$ y $b = n_1r_1$ y con ello, que

$$\frac{a}{b} = \frac{m_1r_1}{n_1r_1} = \frac{m_1}{n_1}.$$

Entonces, $a = (m_1/n_1)b$. Así pues, hemos logrado expresar la longitud de \overline{AB} en función de la de \overline{CD} y hemos hallado un segmento, en este caso \overline{EF} , que cabe un número entero de veces (a saber, m_1) en \overline{AB} y un número entero de veces (a saber, n_1) en \overline{CD} . En otras palabras, hemos visto que \overline{AB} y \overline{CD} son conmensurables.

Si \overline{EF} no cabe exactamente c_2 veces en \overline{CD} , entonces, al colocar c_2 veces \overline{EF} sobre el segmento \overline{CD} , queda un pequeño segmento residual \overline{GH} , de longitud r_2 . En este caso tenemos:

$$b = r_1c_2 + r_2,$$

donde $0 < r_2 < r_1$.

Repitiendo el procedimiento, trataremos de ver cuántas veces enteras cabe \overline{GH} en el segmento \overline{EF} : supongamos que cabe c_3 veces, donde c_3 es un número natural positivo. Si \overline{GH} cabe exactamente c_3 veces en \overline{EF} , entonces ya podemos expresar las longitudes de \overline{EF} , \overline{CD} y \overline{AB} en términos de la de \overline{GH} :

$$\begin{aligned} r_1 &= r_2c_3, \\ b &= r_1c_2 + r_2 = r_2c_3c_2 + r_2 = r_2(c_3c_2 + 1) \quad \text{y} \\ a &= bc_1 + r_1 = r_2(c_3c_2 + 1) + r_2c_3 = r_2(c_3c_2 + 1 + c_3). \end{aligned}$$

Llamando m_2 a $c_3c_2 + 1 + c_3$ y n_2 a $c_3c_2 + 1$, obtenemos $a = m_2r_2$ y $b = n_2r_2$ y con ello, que

$$\frac{a}{b} = \frac{m_2r_2}{n_2r_2} = \frac{m_2}{n_2}.$$

Entonces, $a = (m_2/n_2)b$. Así pues, hemos expresado la longitud de \overline{AB} en función de la de \overline{CD} , hallando un segmento, en este caso \overline{GH} , que cabe un número entero de veces (a saber, m_2) en \overline{AB} y un número entero de veces (a saber, n_2) en \overline{CD} . Por consiguiente, \overline{AB} y \overline{CD} son conmensurables.

Si \overline{GH} no cabe exactamente c_3 veces en \overline{EF} , entonces, al colocar c_3 veces \overline{GH} sobre el segmento \overline{EF} , queda un pequeño segmento residual \overline{IJ} , de longitud r_3 . En este caso tenemos:

$$r_1 = r_2c_3 + r_3,$$

donde $0 < r_3 < r_2$. Repetiríamos así el procedimiento anterior para ahora ver cuántas veces enteras cabe \overline{IJ} en el segmento \overline{GH} : supongamos que cabe c_4 veces, donde c_4 es un número natural positivo. Si \overline{IJ} cabe exactamente c_4 veces en \overline{GH} ,

entonces ya podemos expresar las longitudes de \overline{GH} , \overline{EF} , \overline{CD} y \overline{AB} en términos de la de \overline{IJ} :

$$\begin{aligned} r_2 &= r_3c_4, \\ r_1 &= r_2c_3 + r_3 = r_3c_4c_3 + r_3 = r_3(c_4c_3 + 1), \\ b &= r_1c_2 + r_2 = r_3(c_4c_3 + 1)c_2 + r_3c_4 = r_3n_3, \end{aligned}$$

donde

$$\begin{aligned} n_3 &= c_4c_3 + 1 + c_4 \quad \text{y} \\ a &= bc_1 + r_1 = r_3n_3c_1 + r_3(c_4c_3 + 1) = m_3r_3, \end{aligned}$$

donde $m_3 = c_4c_3 + 1 + n_3c_1$. Obtenemos $a = m_3r_3$ y $b = n_3r_3$ y con ello, que

$$\frac{a}{b} = \frac{m_3r_3}{n_3r_3} = \frac{m_3}{n_3}.$$

Entonces, $a = (m_3/n_3)b$. Así, la longitud de \overline{AB} se ha expresado en función de la de \overline{CD} (más precisamente, como el producto de b por un número racional) y tenemos en este caso que \overline{GH} cabe un número entero de veces (a saber, m_3) en \overline{AB} y un número entero de veces (a saber, n_3) en \overline{CD} ; en otras palabras, que \overline{AB} y \overline{CD} son conmensurables.

Si \overline{GH} no cabe un número entero de veces en \overline{EF} , seguiríamos el procedimiento. Inductivamente, si ya logramos hallar k segmentos de longitudes $r_1, r_2, r_3, \dots, r_k$, respectivamente, tales que:

$$\begin{aligned} a &= bc_1 + r_1, \\ b &= r_1c_2 + r_2, \end{aligned}$$

y, para cada $j \in \{3, 4, \dots, k\}$,

$$r_{k-2} = r_{k-1}c_k + r_k,$$

donde

$$0 < r_k < r_{k-1} < \dots < r_1 < b < a$$

entonces el último segmento, el de longitud r_k , tendría dos posibilidades:

- (i) que quepa exactamente un número entero de veces, digamos c_{k+1} , en el segmento de longitud r_{k-1} , o
- (ii) que quepa c_{k+1} veces en el segmento de longitud r_{k-1} , pero quedando un residuo de longitud r_{k+1} , mayor que cero pero menor que r_k .

Definición 2.2. Diremos que el método de antiféresis para determinar cuánto mide \overline{AB} en términos del (o usando como unidad el) segmento \overline{CD} es finito, si \overline{CD} cabe exactamente un número entero de veces en \overline{AB} o existen un número natural k , y k segmentos de longitudes $r_1, r_2, r_3, \dots, r_k$, respectivamente, tales que:

$$\begin{aligned} a &= bc_1 + r_1, \\ b &= r_1c_2 + r_2, \end{aligned}$$

y, para cada $j \in \{3, 4, \dots, k\}$,

$$r_{k-2} = r_{k-1}c_k + r_k,$$

donde

$$0 < r_k < r_{k-1} < \dots < r_1 < b < a$$

y ocurre (i), es decir, tales que finalmente el último residuo distinto de cero, el segmento de longitud r_k , ya cabe un número entero de veces en el segmento de longitud r_{k-1} .

El lector puede ya estar convencido, siguiendo el razonamiento empleado en los párrafos anteriores de que si el método de antiféresis para determinar cuánto mide \overline{AB} usando como unidad el segmento \overline{CD} es finito, se podrán expresar todas las longitudes, comenzando por r_{k-1} y terminando en a , como múltiplos enteros de r_k , obteniéndose así la conmensurabilidad de los segmentos \overline{AB} y \overline{CD} y la representación de la longitud a como un número racional por la longitud b . Así pues, teniendo en cuenta la Proposición 2.1, podemos afirmar que:

Proposición 2.2. *Dados dos segmentos \overline{AB} y \overline{CD} , de longitudes a y b , respectivamente, las siguientes proposiciones son equivalentes:*

- (a) \overline{AB} y \overline{CD} son conmensurables.
- (b) El método de antiféresis para determinar cuánto mide \overline{AB} en términos de \overline{CD} es finito.
- (c) Existe un número racional q tal que $a = qb$

3 Algoritmo de Euclides vs. antiféresis

Dados dos segmento \overline{AB} y \overline{CD} , un segmento que cupiera un número entero de veces en \overline{AB} y un número entero de veces en \overline{CD} sería una especie de divisor común de

ambos segmentos. El lector seguramente recuerda que, si a y b son números enteros, **un divisor común de a y b** es un número entero c que es divisor de a y que también es divisor de b . Esto significa, precisamente, que c cabe un número entero de veces en a y un número entero de veces en b , es decir, que existen números enteros m y n tales que $a = mc$ y $b = nc$. El **máximo común divisor de a y b** es simplemente el mayor de los divisores comunes de ambos números. Suele denotarse por $m. c. d.(a, b)$ y es un número natural, porque 1 siempre es un divisor común de a y de b . Por ejemplo, los divisores de 18 son $\pm 1, \pm 2, \pm 3, \pm 6, \pm 9$ y ± 18 . Los divisores de 60 son $\pm 1, \pm 2, \pm 3, \pm 4, \pm 5, \pm 6, \pm 10, \pm 12, \pm 15, \pm 30$ y ± 60 . Los divisores comunes son $\pm 1, \pm 2, \pm 3$ y ± 6 . El mayor de todos estos divisores comunes es 6. Entonces $m. c. d.(18, 60) = 6$.

Una manera muy conocida de obtener el m.c.d. de dos números naturales a y b , consiste en desarrollar cada uno de ellos como un producto de potencias de sus divisores primos y formar un nuevo número multiplicando los números divisores primos comunes, elevando cada uno de ellos al menor exponente con el que aparezcan en el desarrollo de a o en el de b . En nuestro ejemplo, $18 = 2 \times 3^2$ y $60 = 2^2 \times 3 \times 5$. Los primos 2 y 3 aparecen en ambos desarrollos, pero 3 aparece elevado al cuadrado en el desarrollo de 18 y elevado a la 1 en el de 60. Escogemos el de menor exponente: $3^1 = 3$; en el caso de 2, escogemos también $2^1 = 2$. Entonces $m. c. d.(18, 60) = 2 \times 3 = 6$

Otro eficaz modo de obtener el m.c.d. de dos números, por ejemplo, 18 y 60, es así: primero veríamos cuántas veces cabe 18 en 60. Esto equivale a hacer una división con residuo, como las que hacíamos en primaria. Al hacerlo, obtenemos un cociente (que en este caso es 3) y un residuo (que es 6) que debe ser menor que el divisor (que es 18): Entonces 18 cabe 3 veces completas (enteras) en 60 y sobran 6 (unidades). Podemos expresar esto así $60 = 3 \times 18 + 6$. Ahora hay que ver cuántas veces cabe 6 en 18. De nuevo haríamos la división de 18 entre 6, pero ésta va a tener residuo cero, porque 6 es un divisor de 18: $18 = 3 \times 6$. Claro que también 6 es divisor de 60 ($60 = 10 \times 6$), pero lo interesante es que, despejando 6 en la expresión $60 = 3 \times 18 + 6$, obtenemos $6 = 60 - 3 \times 18$, igualdad en la que es fácil ver que si un número natural d es divisor común de 18 y de 60, entonces también es divisor de 6 y por lo tanto es menor o igual a 6. Así 6 es el mayor divisor común de 18 y 6, el $m. c. d.(18, 60)$. Este procedimiento es un ejemplo de aplicación de un algoritmo conocido con el nombre de **Algoritmo de Euclides**. El lector iniciado en matemáticas y que haya estudiado algo de la teoría de los números enteros, seguramente ya se percató de la gran similitud que existe entre el método de sustracción mutua para hallar un segmento que quepa un número entero de veces en un segmento dado \overline{AB} y un número entero de veces en otro segmento dado \overline{CD} , con el Algoritmo de Euclides. Este algoritmo se llama así no porque

Euclides lo hubiera descubierto por primera vez (era conocido por Pitágoras en el siglo VI a. C.), sino porque aparece en su obra, los *Elementos*, como la Proposición 2 del libro VII, bien enunciado y demostrado quizá por primera vez. Nosotros aquí lo recordamos en una versión actualizada:

Teorema 3.1 (Algoritmo de Euclides (E. VII. 2)). *Dados dos números naturales a y b , con $a > b$, existen números enteros no negativos k , c_1, \dots, c_{k+1} y r_1, \dots, r_k , tales que:*

$$a = bc_1 + r_1, \quad 0 \leq r_1 < b \quad (1)$$

$$b = r_1c_2 + r_2, \quad 0 \leq r_2 < r_1 \quad (2)$$

y, para cada $j \in \{3, 4, \dots, k\}$,

$$r_{k-2} = r_{k-1}c_k + r_k, \quad 0 \leq r_k < r_{k-1} \quad (3)$$

$$r_{k-1} = r_k c_{k+1}. \quad (4)$$

Entonces r_k , es decir, el último residuo distinto de cero, es el máximo común divisor de a y b .

Para demostrar este algoritmo se aplica repetidamente el **Algoritmo de la división** que, para el caso de números naturales, no es otra cosa que la célebre “división con residuo” que nos enseñaron en la primaria, como vimos en nuestro ejemplo (el de hallar el m. c. d.(18, 60)):

Teorema 3.2 (Algoritmo de la división para números naturales). *Si a y b son números naturales, entonces existen números enteros no negativos c (cociente) y r (residuo), tales que:*

$$a = bc + r, \quad 0 \leq r < b \quad (5)$$

Como primer paso en la demostración del Algoritmo de Euclides, aplicamos el Algoritmo de la división a los números a y b , obteniéndose la igualdad (1).

Si $r_1 = 0$, entonces b divide a a y, por supuesto, también a b , y es entonces el máximo común divisor de a y de b , y aquí termina el proceso. Si $r_1 \neq 0$, entonces es un número natural y, aplicando el Algoritmo de la división a los números b y r_1 , se obtiene la igualdad (2).

Si $r_2 = 0$, entonces r_1 divide a b y también a a , ya que

$$a = bc_1 + r_1 = r_1c_2 + r_1 = r_1(c_2 + 1)$$

y se puede demostrar muy fácilmente, despejando r_1 en la igualdad (1) que si un número natural d divide a a y a b , entonces r_1 divide a d . En otras palabras, r_1 es el máximo común divisor de a y b , y aquí se termina el proceso. Si $r_2 \neq 0$, entonces es un número natural y, aplicando el Algoritmo de la división a los números r_1 y r_2 , se obtiene la igualdad (3) para $k = 3$, es decir,

$$r_1 = r_2c_3 + r_3, \quad \text{con } 0 \leq r_3 < r_2.$$

Si $r_3 = 0$, entonces r_2 divide a r_1 y también a b y a a , ya que

$$\begin{aligned} b &= r_1c_2 + r_2 = r_2c_3 + r_2 = r_2(c_3 + 1) \quad \text{y} \\ a &= bc_1 + r_1 = r_2(c_3 + 1)c_1 + r_2c_3 = r_2(c_3c_1 + c_1 + c_3), \end{aligned}$$

y se puede demostrar también que r_2 es el máximo común divisor de a y b , y aquí se termina el proceso. Si $r_3 \neq 0$, entonces es un número natural y, aplicando el Algoritmo de la división a los números r_2 y r_3 , se obtiene la igualdad (3) para $k = 4$, es decir,

$$r_2 = r_3c_4 + r_4, \quad \text{con } 0 \leq r_4 < r_3.$$

Y así sucesivamente. Sin embargo, los residuos $r_1, r_2, r_3, r_4, \dots$, van formando una sucesión estrictamente decreciente de números enteros no negativos, la cual no puede ser infinita, es decir, en algún momento habrá un primer residuo que sea cero. Más formalmente,

$$\text{existe } k \in \mathbb{N} \text{ tal que } r_k \neq 0, \text{ pero } r_{k+1} = 0.$$

Entonces, el algoritmo debe terminar en un número finito de pasos, y se puede demostrar, usando las igualdades (1) a (4) del enunciado del Algoritmo de Euclides, que ese último residuo distinto de cero es el máximo común divisor de a y b . En símbolos,

$$r_k = \text{m. c. d.}(a, b).$$

4 La demostración de Hipaso

Para lograr expresar la longitud de un segmento \overline{AB} usando la longitud de otro segmento \overline{CD} como unidad, tanto los egipcios, como los babilonios y los griegos de antes del año 450 a. C., usaban la antiféresis como un proceso finito, exactamente como sucede en el Algoritmo de Euclides, porque en la práctica, los segmentitos que van quedando, como residuos sucesivos del proceso, tienen longitudes (r_1, r_2, \dots, r_k)

cada vez más pequeñas hasta que en algún momento ya no les interesaban (su interés es puramente práctico) o los burdos instrumentos de medición de esos tiempos ya no las detectaban. Antes del año mencionado o, para ser menos aventurados, antes de que comenzara la segunda mitad del siglo V a. C., se pensaba que tal método era siempre finito, lo cual equivale, como vimos en la sección anterior, a que cualesquiera dos segmentos son conmensurables. Ahora bien, los egipcios y los babilonios, aunque habían desarrollado una impresionante cantidad de procedimientos prácticos, muchos de ellos muy sofisticados, para resolver problemas matemáticos de muy diversa índole, tenían principalmente un interés práctico y jamás nació entre ellos la posibilidad de pensar de que la antiféresis pudiera no ser un proceso finito o de que pudieran existir segmentos que no fueran conmensurables. Pero en aquella noche del año 450 a. C. estaba **Hipaso de Metaponto** mirando la hermosa estrella pentagonal que servía de símbolo de La Hermandad Pitagórica y se le ocurrió la osadía de medir la diagonal de un pentágono regular usando como unidad su lado pero ahora el interés de este filósofo no era práctico, sino teórico, y el instrumento de medición que emplearía era el más fino que se puede hallar sobre la faz de la Tierra: la mente humana.

Hipaso puso en la mesa todos los ingredientes necesarios para llevar a cabo su medición teórica: los pitagóricos de entonces ya conocían (porque provenían quizá desde los tiempos de Tales de Mileto) todos los teoremas acerca de congruencia y semejanza de triángulos (de hecho, las demostraciones que poseían de todos los teoremas de semejanza de triángulos se basaban en que cualesquiera dos segmentos eran, para ellos, conmensurables) y polígonos regulares. También que un triángulo es isósceles si y solamente si los ángulos que se oponen a los lados iguales, son también iguales y que los ángulos opuestos por el vértice son iguales entre sí. Sabían que la suma de los ángulos interiores de todo triángulo es dos veces un ángulo recto. Podían deducir de aquí, dividiendo el pentágono en tres triángulos por medio de diagonales, que la suma de los ángulos interiores de un pentágono regular es seis veces un ángulo recto y que cada uno de ellos mide entonces $6/5$ de ángulo recto. Cabe mencionar que la unidad de medida de ángulos era el ángulo recto, ya que, por supuesto, no existían las medidas en grados ni en radianes ni nada por el estilo. Denotaremos esta medida con el símbolo \perp . También denotaremos a la longitud de un segmento \overline{AB} como AB , y a la medida de un ángulo ABC por $\angle ABC$.

Ahora sí comenzamos, dijo Hipaso, teniendo ante sí una dibujo como el del pentágono de la figura 1: como el pentágono $ABCDE$ es regular, todos sus lados son iguales,

$$AB = BC = CD = DE = DA,$$

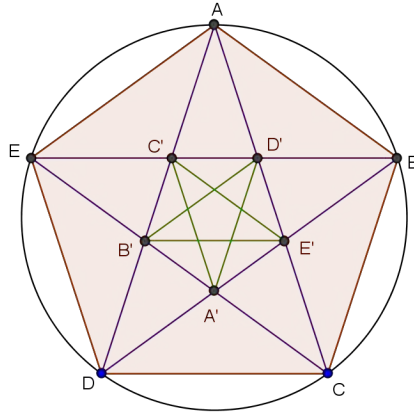


Figura 1: Pentágono

y todos sus ángulos también,

$$\angle ABC = \angle BCD = \angle CDE = \angle DEA = \angle EAB.$$

Así, los triángulos $\triangle ABC$, $\triangle BCD$, $\triangle CDE$, $\triangle DEA$ y $\triangle EAB$, son congruentes entre sí, lo que implica que todas las diagonales del pentágono son iguales entre sí (y con ello, por el criterio lado-lado-lado de congruencia, que los triángulos $\triangle AEB$, $\triangle BAC$, $\triangle CBD$, $\triangle DCE$ y $\triangle EDA$ son congruentes) y que los ángulos $\angle AEB$, $\angle ABE$, $\angle BAC$, $\angle BCA$, $\angle CBD$, $\angle CDB$, $\angle DCE$, $\angle DEC$, $\angle EDA$ y $\angle EAD$ también. Esto nos lleva a que todos los triángulos $\triangle ABD'$, $\triangle BCE'$, $\triangle CDA'$, $\triangle DEB'$ y $\triangle EAC'$, son isósceles y congruentes entre sí (criterio ángulo-lado-ángulo), y con ello, a que

$$AC' = AD' = BD' = BE' = CE' = CA' = DA' = DB' = EB' = EC'$$

y a que

$$\angle AD'B = \angle BE'C = \angle CA'D = \angle DB'E = \angle EC'A.$$

Entonces los triángulos $\triangle AC'D'$, $\triangle BD'E'$, $\triangle CE'A'$, $\triangle DA'B'$ y $\triangle EB'C'$ son congruentes entre sí (criterio lado-lado-lado) y el pentágono $A'B'C'D'E'$ es un pentágono regular (sus lados son iguales en longitud y sus ángulos interiores miden lo mismo) y, por lo tanto, es semejante al primer pentágono $ABCDE$. En particular, el ángulo $\angle B'C'D'$ mide $6/5\angle$. Esta también es la medida del ángulo $\angle AC'E$ (opuestos por el vértice) y, como el triángulo $\triangle AC'E$ es isósceles, entonces

$$2\angle AEC' + 6/5\angle = 2\angle AEC' + \angle AC'E = 2\angle.$$

Despejando, queda que

$$\angle AEC' = (1/2)(2 - (6/5))_{\perp} = 2/5_{\perp}.$$

Por otro lado,

$$\angle EC'B' = 2_{\perp} - (6/5)_{\perp} = (2 - (6/5))_{\perp} = 4/5_{\perp}.$$

Como $\triangle EC'B'$ es isósceles, se deduce fácilmente que $\angle B'EC' = 2/5_{\perp}$. Así,

$$\angle AEB' = \angle AEC' + \angle C'EB' = 4/5_{\perp} = \angle AB'E,$$

lo que implica que $\triangle AEB'$, y $\triangle BCD'$, $\triangle BCA'$, etc., son isósceles. Entonces $AE = AB'$ (y $DE = DC'$, etc.). Como el pentágono $A'B'C'D'E'$ es semejante al pentágono $ABCDE$,

$$\angle B'E'A' = \angle EDA = 2/5_{\perp} = \angle B'DE',$$

de lo que se deduce que $\triangle DB'E'$ es isósceles. En particular, $DB' = B'E'$.

En suma, usando un lenguaje moderno para resumir el posible argumento de Hipaso en su intento de usar la antiféresis, denotemos con d_1 y l_1 a la diagonal y al lado del pentágono $ABCDE$, respectivamente, y con d_2 y l_2 a la diagonal y al lado del pentágono $A'B'C'D'E'$, también respectivamente. Podríamos seguir denotando por d_3 y l_3 , d_4 y l_4 , y así sucesivamente, a las diagonales y lados de los pentágonos que van apareciendo al seguir el proceso indefinidamente. Entonces:

$$\begin{aligned} d_1 &= AD = AB' + B'D = AE + B'E' = l_1 + d_2, \\ l_1 &= DC' = DB' + B'C' = B'E' + B'C' = d_2 + l_2. \end{aligned}$$

Así, l_1 cabe una vez en d_1 y sobra un segmento de longitud d_2 , d_2 cabe una vez en l_1 y sobra un segmento de longitud l_2 . El siguiente paso consiste en ver cuántas veces cabe l_2 en d_2 . Pero aquí es donde se aplica el contundente argumento teórico de que, como el segundo pentágono es semejante al primero, las veces que cabe su lado en su diagonal, son las mismas que las veces que cabe el lado del primer pentágono en su diagonal, porque todos los argumentos que nos llevaron a probar la primera de estas igualdades, serán los mismos, *mutatis mutandis*, que para demostrar las siguientes igualdades:

$$d_2 = l_2 + d_3,$$

$$l_2 = d_2 + l_3,$$

y también éstas

$$d_3 = l_3 + d_4,$$

$$l_3 = d_4 + l_4,$$

y las que seguirían en un proceso que nunca termina, porque los pentagonitos nunca se colapsan a un punto, aunque uno ya no los vea más que con los ojos de la imaginación. Se obtiene entonces una sucesión decreciente de longitudes de segmentitos y que no acaba jamás:

$$d_1 > l_1 > d_2 > l_2 > d_3 > l_3 > d_4 > \dots$$

Entonces, en este caso, la antiféresis no termina nunca. Hipaso comprendía que esto significa que existen segmentos inconmensurables y que este descubrimiento daba al traste con una creencia sólida del mundo pitagórico, en la que habían basado toda una filosofía y las demostraciones de todo un conjunto de teoremas, como los de semejanza. Revelar el descubrimiento sería una declaración hereje, la más terrible de las blasfemias contra el Maestro Pitágoras y su doctrina. No se sabe cómo se conoció este corrosivo secreto que Hipaso no pudo ocultar por mucho tiempo, quizá revelado a una seductora y bella dama, que supo extraérselo con sus encantos, del mismo modo en que Dalila descubrió el misterio de la fuerza sobrehumana de Sansón. Lo que sí seguramente ocurrió, es que tuvo que exponer sus razonamientos en un juicio ante los miembros de La Hermandad. Alguno de ellos pudo haberle cuestionado:

—Bueno, pero tu demostración sólo indica que la antiféresis no funciona para la diagonal y el lado de un pentágono regular. Pero qué tal si por otro método se puede demostrar que dichos segmentos sí son conmensurables.

—No su señoría —contestó Hipaso con toda la honestidad intelectual que lo caracterizaba, sabiendo que su respuesta era su sentencia de muerte— si existiera un segmento u que cupiera un número entero de veces, digamos m_1 en d_1 , y un número entero de veces, n_1 , en l_1 , entonces se tendría que, por las igualdades de arriba,

$$\begin{aligned}d_2 &= d_1 - l_1 = m_1u - n_1u = (m_1 - n_1)u = m_2u, \\l_2 &= l_1 - d_2 = n_1u - m_2u = (n_1 - m_2)u = n_2u, \\d_3 &= d_2 - l_2 = m_2u - n_2u = (m_2 - n_2)u = m_3u,\end{aligned}$$

y, análogamente, $d_4 = m_4u$, $d_5 = m_5u$, $d_6 = m_6u$, \dots , donde $m_1 > m_2 > m_3 > m_4 > \dots$ (porque $d_1 > d_2 > d_3 > d_4 > \dots$), y todas estas m_i son enteros positivos (porque ninguna d_i es igual a cero). Tendríamos entonces algo imposible: una sucesión interminable y estrictamente decreciente de números enteros positivos. Esta contradicción nos lleva a afirmar que no existe tal segmento u y, por consiguiente, que *la diagonal y el lado del pentágono regular, son inconmensurables.*

5 El primer número irracional

Por nuestra Proposición 2.2 sabemos que demostrar que existen segmentos inconmensurables equivale a demostrar que existen números irracionales. Para ver cuál es el número irracional que encontró Hipaso (él no lo planteó así, por supuesto), volvamos a la figura del pentágono. En él se observan dos triángulos semejantes: $\triangle ABC$ y $\triangle AB'E'$. Debido a dicha semejanza, se tiene: $AD/AB' = DC/B'E'$, lo que, con la notación que usamos en la demostración de Hipaso, equivale a $d_1/l_1 = l_1/d_2$. Llamemos λ a esta razón, es decir, $\lambda = d_1/l_1$. Entonces se obtienen las igualdades:

$$d_1 = l_1 + d_2 \quad \text{y} \quad l_1^2 = d_1 d_2,$$

lo que implica que

$$l_1^2 = (l_1 + d_2)d_2 = l_1 d_2 + d_2^2.$$

Dividiendo entre d_2^2 , queda

$$l_1^2/d_2^2 = (l_1/d_2) + 1,$$

o lo que es lo mismo,

$$\lambda^2 = \lambda + 1.$$

Resolviendo esta ecuación de segundo grado y teniendo en cuenta que $\lambda > 0$, se obtiene $\lambda = (1 + \sqrt{5})/2$. Este es el célebre número áureo. Entonces, si aceptamos que los primeros dos segmentos que se supo que eran inconmensurables fueron la diagonal y el lado de un pentágono regular, tendremos que aceptar que el primer número irracional que se descubrió es el número áureo. Hay mucha discusión entre los historiadores acerca de ello. El otro número que compite por la prioridad como número irracional, es la raíz cuadrada de 2. En efecto, en muchos libros de texto que tratan el tema de los números reales, señalan que el primer número irracional descubierto fue $\sqrt{2}$, el número real positivo cuyo cuadrado es 2, y la demostración que dan acerca de la irracionalidad de este número es la clásica demostración por reducción al absurdo (ver, por ejemplo, [7]).

Como dice Benoît Rittaud en [7], «desde el punto de vista pedagógico, una reducción al absurdo como la de la demostración clásica, posee en general un débil poder de convicción. Es necesaria cierta práctica de razonamiento matemático para comprender el sentido exacto y aceptar la validez de una introducción al tema como: ‘voy a demostrarles que $\sqrt{2}$ no se puede escribir en forma de fracción. Supongo que $\sqrt{2} = p/q$, y luego...’. De buenas a primeras se dice una cosa y su contrario en el espacio de dos frases consecutivas. Esta dificultad, inherente a este modo de razonamiento, hace que, según lo que han observado varios autores (como Rapad Szabó), la demostración clásica tenga pocas posibilidades de haber sido aquella gracias a la

cual se descubrió la irracionalidad de la raíz cuadrada de 2 (o la inconmensurabilidad de la diagonal del cuadrado y de su lado), en la medida en que su estructura misma hace de ella un razonamiento que no pueda llevar a cabo sino alguien que ya esté al corriente del resultado final».

La demostración clásica suele atribuírsele a Aristóteles, quien en sus *Primeros Analíticos* escribió: «Suponer que la diagonal es conmensurable implica confundir los números pares y los números los impares». También escribió esta afirmación en otras de sus obras. En un apéndice a los *Elementos* de Euclides aparece la demostración original de Aristóteles, que parece algo más complicada que la mencionada arriba, lo cual es lógico porque en ese tiempo no se manejaban los números reales con las propiedades que aceptamos modernamente. Según el mismo Aristóteles, son los pitagóricos los que demuestran la existencia de segmentos inconmensurables y de hecho, la irracionalidad de $\sqrt{2}$ (es decir, la inconmensurabilidad de un segmento de longitud $\sqrt{2}$ y un segmento de longitud 1).

Los primeros textos que mencionan explícitamente las magnitudes inconmensurables son los *Diálogos* de Platón. En uno de ellos, el «Menón» (escrito hacia el año 380 a. C.), Sócrates hace tomar conciencia a un esclavo del hecho de que la diagonal de un cuadrado es su duplicadora, es decir, que dado un cuadrado cualquiera, podemos duplicar su área construyendo un cuadrado cuyo lado sea igual a la diagonal del cuadrado original, pero para resolver este problema no se necesita saber que la diagonal es inconmensurable con su lado y Platón no habla de ello. El momento del diálogo en el que más se acercan al tema es cuando Sócrates, ante el aprieto del esclavo para hallar el lado del cuadrado de área doble, le dice: «si no quieres dar el número, muéstralo sobre el dibujo». Rittaud señala en [7] que «es necesario ser indulgente para ver en esta sugerencia una alusión a la inexistencia de una relación de números enteros que estén en proporción con el lado y la diagonal». En otro diálogo, «La República», encontramos una breve mención de las «líneas inconmensurables» y, más adelante, una exposición poco clara en la que se trata de «diagonales irracionales» y se deja entrever la inconmensurabilidad de la diagonal y el lado del cuadrado.

Mucho más explícito es otro diálogo de Platón llamado «Teetetos». Este diálogo fue escrito en el año 368/367 a. C., poco tiempo después de la muerte del matemático Teetetos, al término de una batalla en la que fue fatalmente encontrado (ver [8]). En la primera parte del diálogo, el viejo matemático Teodoro de Cirene es representado demostrando a un grupo de hombres jóvenes, entre los cuales se encuentra el joven Teetetos, la irracionalidad de las raíces cuadradas de 3, 5, 6 y hasta 17, por supuesto, sin incluir a las raíces cuadradas de 4 y de 9. El hecho de que Teodoro se haya detenido en 17 ha dado mucho que pensar, pero más todavía la cuestión de por qué no se menciona la irracionalidad de $\sqrt{2}$ y se comienza desde la raíz cuadrada de 3.

Se cree que esta ausencia, lejos de ser el resultado de una ignorancia, muestra de la irracionalidad de $\sqrt{2}$ debía ser demasiado conocida como para que fuera útil citarla. Así pues, el descubrimiento de la inconmensurabilidad tuvo que haberse realizado, no sólo antes de Aristóteles, sino también, antes de Platón. De hecho, la mayoría de los historiadores está de acuerdo en que fue realizado en el seno de la Hermandad Pitagórica. A modo de resumen listamos algunos de los argumentos de von Fritz para hacer plausible su hipótesis de que fue precisamente el pitagórico Hipaso de Metaponto quien lo llevó a cabo y en el pentágono:

El descubrimiento tuvo que haber sido al nivel elemental de las matemáticas griegas de la primera mitad del siglo V a. C. El carácter de la investigación científica tan desarrollado en el siglo V a. C. no sólo hizo posible, sino muy probable que el descubrimiento fuera hecho a pesar de que su conocimiento era mínimo. Aún más significativo es el hecho de que la prueba completa de la raíz cuadrada de dos, como es presentada usualmente y que se atribuye a Aristóteles en el siglo IV a. C., usa los términos conmensurable e inconmensurable, lo cual hace presuponer que el concepto de inconmensurabilidad ya era conocido cuando se hizo la demostración en el cuadrado. Los pitagóricos usaban el pentagrama como insignia, lo cual debió de haber atraído su atención y evocado un deseo de analizar matemáticamente su forma. Es probable que Hipaso para lograr el desarrollo del dodecaedro, hubiera tratado de investigar acerca de números y radios incorporados en el pentagrama y el pentágono regular.

6 Importancia del descubrimiento de la inconmensurabilidad

La idea de que dos magnitudes, y más concretamente, dos segmentos, tienen siempre una parte alcuota común, es decir que son conmensurables, representa una etapa primigenia e inevitable en el desarrollo del pensamiento intuitivo matemático, tanto en el horizonte histórico como en el escolar y, por supuesto, en el ámbito artesanal, por necesidades de la medida siempre aproximada de longitudes. La aparición de las magnitudes inconmensurables marcó una inflexión radical en la evolución histórica de la geometría griega, ya que puso fin al sueño filosófico pitagórico acerca del número como esencia del universo, eliminó de la geometría la posibilidad de medir siempre con exactitud y fue lo que imprimió a la matemática griega una orientación geométrica deductiva plasmada en los *Elementos* de Euclides. El descubrimiento de segmentos inconmensurables marca un hito en la historia de la filosofía, porque no es algo empírico, sino puramente teórico; conduce a un trastorno lógico que estremece los cimientos de la matemática y filosofía griegas, ya que invalida todas las pruebas pitagóricas de los teoremas que hablan de proporciones y la visión fi-

losófica del papel central de los números enteros. Que el lado y la diagonal de un pentágono no tengan medida común implica que, si un número representa a uno de los segmentos, ningún número puede representar al otro. Así pues, es imposible reconocerlos numéricamente juntos. Sin embargo, ante nuestros ojos se muestran conviviendo juntos en un hermoso pentágono. Lo real aquí no es discutible y excede manifiestamente la capacidad numérica. Estas magnitudes geométricas escapan a la numericidad. Para reconstruir el edificio dañado, los griegos desarrollaron una teoría interna referente sólo a las magnitudes geométricas. Establecieron proporciones entre estas magnitudes, pero se negaron a llamarlas números. Como dice V. Gómez Pin en [5], la crisis abierta por el descubrimiento de la inconmensurabilidad, tuvo como consecuencia que la geometría fuera en parte privilegiada en detrimento de la aritmética. Para los griegos, tal descubrimiento, como dice W. Dunham en [1], establecía firmemente la superioridad de la geometría respecto de la aritmética en todas las matemáticas griegas a partir de entonces.

Muchos historiadores se han dedicado a este acontecimiento y existen entre ellos opiniones encontradas, pero nadie puede dudar que considerar el descubrimiento de la inconmensurabilidad como un gran momento en la historia de las matemáticas, como lo hizo Howard Eves en [4], es un gran acierto.

Agradecimientos

Agradecemos infinitamente las acertadas y eruditas sugerencias del árbitro, que mejoraron este trabajo en todos los aspectos, y a la maestra Mónica Macías por su valiosísimo apoyo técnico.

Bibliografía

- [1] W. Dunham, *Viaje a través de los genios*, Pirámide, Madrid, 1992.
- [2] C. H. Edwards, Jr. *The Historical Development of the Calculus*, Springer - Verlag, New York, Heidelberg, Berlin, 1937.
- [3] H. Eves, *An Introduction to the History of Mathematics*, Holt, Rinehart and Winston, New York, Fourth Edition, 1976.
- [4] H. Eves, *Great Moments in Mathematics (before 1650)*, The Mathematical Association of America, 1980.
- [5] V. Gómez Pin, *La tentación pitagórica*, Síntesis, Madrid, segunda edición, 2010.

- [6] M. Kline, *Matemáticas. La pérdida de la certidumbre*, Siglo XXI de España Editores, Madrid, 1985.
- [7] B. Rittaud, *Qué irracional. El fabuloso destino de $\sqrt{2}$* , Consejo Nacional para la Cultura y las Artes, Colección QED, Cd. de México, México, 2009.
- [8] K. von Fritz, *The discovery of incommensurability by Hippasus of Metaponto*, Ann. Math. (2) 46 (1945), 242-264.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570

acontri@fcfm.buap.mx

corogant@live.com

ochiris@gmail.com

Probabilidad y Estadística

Capítulo 7

Un portafolio de inversión óptimo bajo un enfoque de Conditional Value at Risk (CVaR)

Ambrosio Ortiz Ramírez¹, Abigail Rodríguez Nava²,
Luis Alberto Sánchez Zacateco¹

¹Escuela Superior de Economía, Instituto Politécnico Nacional,

²Universidad Autónoma Metropolitana–Unidad Xochimilco

Resumen

En este trabajo se presenta la metodología para construir un portafolio óptimo con una medida de riesgo coherente: Valor en Riesgo Condicional (CVaR) dada por [13]. El problema a resolver consiste en minimizar el CVaR de un portafolio compuesto por n activos. Dada la complejidad de la función objetivo asociada al CVaR se incluye una función auxiliar que permite transformar tal función objetivo en otra que depende de los pesos de los activos del portafolio y de la función auxiliar. Bajo estas condiciones hay un caso particular en el cual la minimización del CVaR es un problema de optimización manejable: cuando la función de densidad de probabilidad conjunta de los rendimientos de los activos del portafolio está representada en un conjunto de escenarios. Esto suele ser el tipo de datos que se encuentran en la práctica: se generan escenarios de pérdidas y ganancias de un portafolio, lo cual permite minimizar el CVaR por el método de simulación histórica. La metodología se aplica a un portafolio compuesto por una muestra de 10 acciones que componen el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores, con una periodicidad que comprende desde el 01/01/2014 al 22/07/2015 con un total de 406 observaciones. Se calcula el VaR por el método de simulación histórica al 95% y al 99% de confianza. Asimismo se determinaron el CVaR mínimo y los pesos que lo conforman con los mismos niveles de confianza. La evidencia empírica muestra que $\min_x \text{CVaR}_\alpha(x) \geq \text{VaR}_\alpha(x)$ para los dos niveles de confianzas elegidos. Este resultado implica que implementar metodologías más robustas para la medición del riesgo de mercado como lo es el CVaR, es importante para un administrador de riesgos, en un intento de reducir el impacto de las variaciones abruptas en los precios de los activos que se negocian en los mercados financieros y que forman los portafolios de inversión.

1 Introducción

La optimización de portafolio es uno de los enfoques más conocidos en la selección de portafolios de inversión. La primera técnica para resolver el problema de selección de portafolio fue desarrollada por Harry Markowitz, 1952 [11]. En el denominado modelo de optimización de portafolio en media-varianza (MV) propuesto por Markowitz, el rendimiento del portafolio se mide por el rendimiento esperado del portafolio y el riesgo asociado se mide por la varianza de los rendimientos de portafolio. Una de las extensiones de optimización de portafolios se presenta en [5], la propuesta consiste en la solución del problema de conformar un portafolio óptimo cuando los activos tienen rendimientos provenientes de distribuciones α -estables.

La varianza como medida de riesgo presenta algunas desventajas. Al controlar (minimizar) la varianza no sólo conduce a una baja desviación del rendimiento esperado en la parte izquierda de la distribución de pérdidas, sino también en el lado derecho, por lo cual puede limitar posibles ganancias [18]. Por lo tanto, algunas medidas de riesgo se han propuesto como alternativa para reemplazar la varianza tales como el Valor en Riesgo (VaR), el cual modela y mide el riesgo en términos de percentiles de la distribución de pérdidas. En lugar de basarse en valores negativos y positivos del rendimiento esperado, considera sólo el lado negativo del rendimiento esperado como riesgo y representa la pérdida máxima pronosticada con un nivel de confianza particular (por ejemplo, 99%) durante un cierto periodo de tiempo (por ejemplo, un día) [7, 13].

El VaR como medida de riesgo es muy popular. Sin embargo, tiene algunas propiedades que limitan su uso [3, 12], como la falta del subaditividad; es decir, VaR de dos portafolios puede ser mayor que la suma de los VaRs de cada portafolio. Además, el VaR es no convexo y suave, además de tener varios mínimos locales, lo cual es un problema puesto que se busca un mínimo global [7, 12]. Así se han desarrollado medidas de riesgo alternativas como el Valor en Riesgo Condicional (CVaR)¹ que se interpreta como el valor de la esperanza condicional de la distribución de pérdida dado que excede el VaR [4]. VaR implica que ¿Cuál es la pérdida máxima que tiene el portafolio?, mientras que CVaR: ¿Cómo se espera el incurrir en pérdidas bajo condiciones extremas?. Experimentos numéricos muestran que el CVaR mínimo ordinariamente conduce a soluciones óptimas cercanas a un VaR mínimo, ya que VaR nunca excede CVaR [13]. CVaR tiene mejores características que el VaR. Es un problema de optimización convexo, y por lo tanto es fácil optimizar [7]. Está demostrado que se pueden utilizar técnicas de programación lineal para optimización del CVaR como medida de riesgo [13, 9]. A su vez en [19] se muestra la aplicación de optimización robusta a un portafolio con el peor caso del CVaR y

¹ CVaR también se conoce como Expected Shortfall (Déficit Esperado).

con información parcial de la distribución de probabilidad, se muestran las ventajas mediante ejercicio de simulación Monte Carlo.

El objetivo central de este trabajo es presentar los conceptos de riesgo, su clasificación en un marco simple con su relación con lo que es una medida de riesgo coherente, asimismo proporcionar las bases teóricas de la metodología de optimización de un portafolio con el CVaR como función objetivo en el contexto de [13].

Este trabajo está organizado como sigue. En la siguiente sección se presenta una introducción al concepto de riesgo, una clasificación de los riesgos y medidas coherentes de riesgo. En la sección 3 se describe la metodología general para calcular el VaR, en específico los pasos para calcular el VaR por el método de simulación histórica, en la sección 4 se plantea el problema de optimización del CVaR, en la sección 5 se aplica la metodología a un portafolio de acciones, por último, se concluye en la sección 6.

2 Riesgo, clasificación y medidas coherentes de riesgo

El término riesgo tiene sus raíces en el latín *risicare* (o *resecare*) que significa atreverse que significa atreverse y en el griego *rizha* (o *riza*) que significa navegar por un acantilado para alcanzar la costa [17]. La definición más común de riesgo es: exposición a la posible pérdida, daño, o alguna otra circunstancia desafortunada. En el siglo XVI ya se tenía una concepción moderna del riesgo, ya que se definía como una contingencia que podía presentarse. Posteriormente, a mediados del siglo XX los conceptos riesgo e incertidumbre se trataron de manera indistinta; sin embargo, actualmente existe una diferencia muy clara entre incertidumbre y riesgo; mientras la incertidumbre es subjetiva y no medible, lo que implica una distribución de frecuencia desconocida, el riesgo es cuantificable y objetivo, con una distribución de frecuencias conocida.

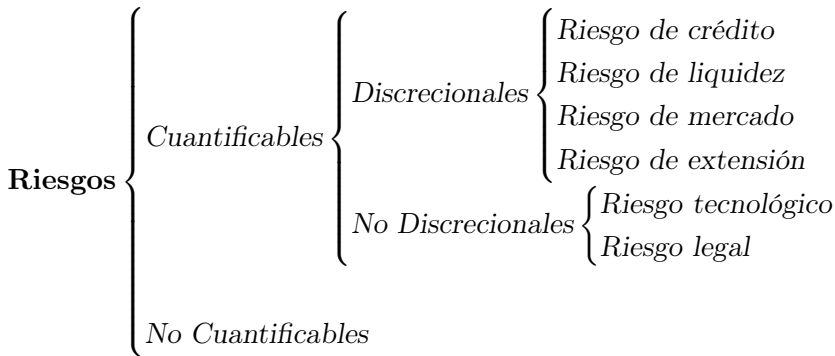
Otro enfoque entre incertidumbre y riesgo es el expuesto en [8], quien define la incertidumbre, riesgo y la medición de éstos como:

- Incertidumbre: Falta de certidumbre, esto es, la existencia de más de una posibilidad. El verdadero resultado (valor, estado) no es conocido.
- Medición de la incertidumbre: Un conjunto de probabilidades asignado a un conjunto de posibilidades. Ejemplo: existe un 80 % de que los alumnos aprueben el curso, 10 % de que reprobren.
- Riesgo: Es un estado de la incertidumbre donde algunas posibilidades pueden ser pérdidas, catástrofes, o algún resultado inesperado .

- **Medición del riesgo:** Un conjunto de posibilidades el cual se mapea aun conjunto de probabilidades y a un conjunto de cuantificación de pérdidas. Ejemplo: existe un 25 % de probabilidad de que el IPyC disminuya a los niveles de octubre 2008 con una pérdida generalizada de 500 billones de pesos.

En este contexto se afirma que se puede tener incertidumbre sin riesgo, pero no riesgo sin incertidumbre [8]. Se puede tener incertidumbre en cuanto al ganador de un partido de fútbol, pero si se apuesta no existe el riesgo de alguna pérdida.

Otra clasificación de riesgos que se encuentra en la práctica de la administración de riesgos es la siguiente (Véase [6]):



1.— **Riesgos cuantificables**, que son aquellos para los cuales es posible conformar bases estadísticas que permitan medir sus pérdidas potenciales, y dentro de estos se encuentran los siguientes:

a.— **Riesgos discrecionales**, que son aquellos resultantes de la toma de una posición de riesgo, Tipos: riesgo de crédito, liquidez y de mercado, de extensión.

i) **Riesgo de crédito o crediticio**. Se define como la pérdida potencial por la falta de pago de un acreditado o contraparte en las operaciones que efectúan los Organismos de Fomento y Entidades de Fomento, incluyendo las garantías reales o personales que les otorguen, así como cualquier otro mecanismo de mitigación utilizado por dichos Organismos de Fomento y Entidades de Fomento.

ii) **Riesgo de liquidez**. Se define como la pérdida potencial por la imposibilidad o dificultad de renovar pasivos o de contratar otros en condiciones normales para los Organismos de Fomento y Entidades de Fomento, por la venta anticipada o forzosa de activos a descuentos

inusuales para hacer frente a sus obligaciones o bien, por el hecho de que una posición no pueda ser oportunamente enajenada, adquirida o cubierta mediante el establecimiento de una posición contraria equivalente.

- iii)* **Riesgo de mercado.** Se define como la pérdida potencial por cambios en los Factores de Riesgo que inciden sobre la valuación o sobre los resultados esperados de las operaciones activas, pasivas u Operaciones Causantes de Pasivo Contingente, tales como tasas de interés, tipos de cambio e índices de precios, entre otros.
- iv)* **Riesgo de extensión.** Se define como la pérdida potencial por la posibilidad de no cubrir la totalidad de los créditos con los pagos establecidos, como consecuencia de la obligación que tienen el Infonavit y Fovissste de eximir al acreditado de su pago alcanzado al vencimiento del plazo del crédito.

b.— **Riesgos no discrecionales**, son aquellos resultantes de la realización de sus actividades, pero que no son producto de la toma de una posición de riesgo, tales como el **riesgo operacional**, que se define como la pérdida potencial por fallas o deficiencias en los controles internos, por errores en el procesamiento y almacenamiento de las operaciones o en la transmisión de información, así como por resoluciones administrativas y judiciales adversas, fraudes o robos, y comprende, entre otros, al riesgo tecnológico y al riesgo legal, en el entendido de que:

- i)* El **riesgo tecnológico** se define como la pérdida potencial por daños, interrupción, alteración o fallas derivadas del uso o dependencia en el hardware, software, sistemas, aplicaciones, redes y cualquier otro canal de distribución de información en la prestación de servicios con los clientes o derechohabientes de los Organismos de Fomento y Entidades de Fomento.
- ii)* El **riesgo legal** se define como la pérdida potencial por el incumplimiento de las disposiciones legales y administrativas aplicables, la emisión de resoluciones administrativas y judiciales desfavorables y la aplicación de sanciones, en relación con las operaciones que los Organismos de Fomento y Entidades de Fomento llevan a cabo.

2.— **Riesgos no cuantificables**, que son aquellos derivados de eventos imprevistos para los cuales no se puede conformar una base estadística que permita medir las pérdidas potenciales.

De esta manera, el riesgo se puede definir como la variación del valor de un portafolio de inversión, de crédito o de seguros con respecto a su valor actual, debido

a variaciones en los factores que los componen. Eso significa que, a diferencia de lo que se considera generalmente, tanto las desviaciones positivas y negativas del valor del portafolio se consideran riesgo, lo que se conoce comúnmente como volatilidad de los factores.

2.1 Medidas coherentes de riesgo

Existen escenarios a los cuales se asocia un riesgo mayor en comparación con otros. Se trata entonces del problema de encontrar una forma de medir y cuantificar el riesgo en cada escenario. La medición del riesgo no es un tema nuevo y ya se han propuesto varias formas para cuantificarlo. Si se parte de la definición de riesgo como la desviación de un resultado esperado entonces ejemplos de medidas de riesgo son la varianza, la desviación estándar, desviación media absoluta, etc. Así, dado un conjunto de observaciones (por ejemplo de rendimientos de algún activo) la varianza representa la media aritmética del cuadrado de las desviaciones respecto a la media, mientras que la desviación estándar es la raíz cuadrada positiva de la varianza. Por su parte, la desviación media absoluta es el promedio aritmético del valor absoluto de las desviaciones respecto a la media. Las tres medidas nos indican cuánto se desvían las observaciones de su media. Surgen entonces dos preguntas: ¿cuál medida de riesgo es mejor?, ¿cuales son las características de una medida de riesgo apropiada?. Para responder a esas preguntas se encuentran en la literatura trabajos que proponen propiedades deseables que debería cumplir una medida de riesgo para ser caracterizada como coherente. Para introducir los axiomas de coherencia en seguida se enuncia una definición formal de medida de riesgo.

Sea un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y sea el horizonte de riesgo T . Sean $L^0(\Omega, \mathcal{F}, \mathbb{P})$ los conjuntos de variables aleatorias en (Ω, \mathcal{F}) , que son casi seguramente finitos. Los riesgos financieros son representados por un conjunto $\mathcal{M} \in L^0(\Omega, \mathcal{F}, \mathbb{P})$ de variables aleatorias, que se interpretan como pérdidas de un portafolio sobre un horizonte de tiempo T . El horizonte de tiempo es libre, es decir sin especificar, aunque en la práctica siempre es finito. Se supone que \mathcal{M} es un cono convexo, es decir, si $L_1 \in \mathcal{M}$ y $L_2 \in \mathcal{M}$ implica que $L_1 + L_2 \in \mathcal{M}$ y que $\lambda L_1 \in \mathcal{M}$ para cada $\lambda > 0$. Las medidas de riesgo son funciones reales $\rho \rightarrow \mathbb{R}$ definidos en los conos de variables aleatorias que satisfacen ciertas propiedades. Se interpreta $\rho(L)$ como el monto de capital que se deberá agregar a una posición con pérdida L . La interpretación de L difiere de aquella propuesta en [3], donde la variable aleatoria $L \in \mathcal{M}$ es interpretada como el valor futuro (en lugar de pérdidas) de una posición que se tiene actualmente, también las tasas de interés son cero para establecer las medidas de riesgo. Una medida de riesgo ρ es coherente de acuerdo a Artzner [3] si satisface cuatro axiomas:

1. **Invarianza ante traslaciones.** Para toda $L \in \mathcal{M}$ y cada $l \in \mathbb{R}$ se cumple que $\rho(L + l) = \rho(L) + l$. Una medida de riesgo invariante ante traslaciones implica que si cuenta con un activo riesgoso y se forma una cartera agregando un activo libre de riesgo, el riesgo de la cartera disminuye en la cantidad del activo libre de riesgo.
2. **Subaditividad.** Para toda $L_1 + L_2 \in \mathcal{M}$ se tiene que $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$. La idea principal de esta propiedad es que “una unión no crea riesgo extra”, esto refleja la idea de que el riesgo se puede reducir diversificando los activos.
3. **Homogeneidad positiva.** Para toda $L \in \mathcal{M}$ y cada $\lambda > 0$ se cumple que $\rho(\lambda L) = \lambda \rho(L)$. En este caso, cuando una riqueza bajo riesgo representada por un activo es multiplicada por un factor positivo, el riesgo y la incertidumbre crecen en la misma proporción. La falta de liquidez de los mercados financieros, resta realismo a esta propiedad porque en este caso no se tendría una relación lineal. Esta propiedad es fácilmente justificada si se supone subaditividad, ya que esta implica que:

$$\rho(nL) = \rho(L + \dots + L) = n\rho(L),$$

dado que no hay diversificación entre las pérdidas de este portafolio es natural que la desigualdad anterior se convierta en igualdad, de donde surge entonces la homogeneidad positiva. Observe que subaditividad y homogeneidad positiva implica que la medida de riesgo ρ sea convexa en \mathcal{M} .

4. **Monotonía.** Para toda $L_1 + L_2 \in \mathcal{M}$ tal que $L_1 \leq L_2$ entonces se cumple casi seguramente que $\rho(L_1) \leq \rho(L_2)$.

En síntesis la importancia de los axiomas anteriores en la medición del riesgo mercado puede explicarse como sigue: la homogeneidad positiva se relaciona con el tema de liquidez; la subaditividad refleja la idea de que la diversificación no crea riesgo adicional; la monotonía regula a las medidas de riesgo de tipo semi-varianza [16], la invarianza ante traslaciones implica que agregar un activo con rendimiento constante a otro con rendimiento aleatorio puede reducir el riesgo.

La varianza, desviación estándar y desviación media absoluta ninguna resulta ser una medida coherente. La varianza sólo cumple subaditividad, mientras que la desviación estándar y la desviación media absoluta cumplen con ser subaditivas y tener homogeneidad positiva pero no con la monotonía ni con ser invariantes ante traslaciones.

En [1] sección 3 se demuestra que CVaR satisface los cuatro axiomas, por consiguiente, califica como una medida de riesgo coherente. De hecho, en [2] se muestra

que cualquier medida de riesgo coherente puede representarse como una combinación convexa de CVaRs con diferentes niveles de confianza. Además, CVaR es una función convexa con respecto a posiciones en un portafolio, lo cual permite la construcción de algoritmos eficientes de optimización. En particular, Rockafellar y Uryasev [13] muestran que el CVaR puede minimizarse utilizando técnicas de programación lineal, lo cual permite ejecutar cálculos a gran escala de manera práctica, eficiente y estable.

3 Modelo de simulación histórica

Dentro de la literatura sobre riesgos financieros se encuentran diversas metodologías para medir el riesgo de mercado de los portafolios de inversión, que tradicionalmente la medida estándar para hacerlo es con el VaR. A continuación se describirá el modelo de Simulación Histórica el cual es un método no paramétrico basado en datos históricos para calcular el VaR.

Si $G(V)$ es la distribución de probabilidad de los cambios futuros del valor de mercado de un portafolio, entonces dado un nivel de confianza α , se puede encontrar la mayor disminución del valor de mercado del portafolio (V) en relación con el valor vigente, tal que la probabilidad de que la reducción del valor del portafolio sea mayor a V es $1 - \alpha$. Esta mayor pérdida del valor del portafolio (V) es el valor en riesgo.

Es decir, para estimar el VaR es necesario conocer la distribución de probabilidad de los cambios futuros del valor de mercado del portafolio y de cada una de sus posiciones durante el periodo de tenencia del portafolio. En general, para estimar la distribución de probabilidad y el VaR se deben ejecutar las siguientes fases [10]:

1. Identificar los factores de riesgo que pueden influir en el valor de mercado del portafolio de inversión.
2. Construir la distribución de probabilidad de los cambios de los factores de riesgo que podrían ocurrir durante el horizonte de inversión².
4. Estimar la distribución de probabilidad de los cambios en el valor de mercado del portafolio, a partir de los efectos de los factores de riesgo de las distribuciones de probabilidad estimadas en la fase anterior³.

² Cabe mencionar que en esta fase el objetivo es predecir un rango de posibles cambios en el factor de riesgo, y no el realizar un pronóstico específico del factor de riesgo.

³ Es pertinente resaltar que en este método el riesgo se modela directamente de la distribución empírica de los rendimientos de activos generados por los factores de riesgo ya identificados previamente.

3. Calcular el VaR de las posiciones individuales en cada activo y de todo el portafolio de inversión.

En función de los supuestos y alcances que se consideran para realizar las fases anteriores, los modelos estadísticos de valor en riesgo se pueden clasificar en: modelos de portafolio o de varianza-covarianza, donde la revaluación del portafolio se realiza mediante aproximaciones analíticas, y los modelos de simulación: histórica y de Monte Carlo, los cuales construyen la distribución de probabilidad a partir de la generación de escenarios y la revaluación del portafolio con cada uno de ellos.

El método de simulación histórica consiste en generar escenarios de los factores de riesgo (tasas de interés, tipo de cambio, precio de acciones, etc.) a partir de la información observada en un determinado número de días. La estimación del VaR consiste en las siguientes etapas:

1. Se genera una serie histórica del factor de riesgo (F).
2. Se construye la serie de rendimientos, es decir, se estiman las variaciones diarias de los factores de riesgo:

$$R_{t-1} = \ln \left(\frac{F_t}{F_{t-1}} \right)$$

3. Se estima una serie alternativa del factor de riesgo. Para ello, al valor actual del factor de riesgo se le agrega el valor de las variaciones calculadas en el paso anterior:

$$F_0 \times \exp \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} F_{01} \\ F_{02} \\ \vdots \\ F_{0n} \end{pmatrix}$$

4. El portafolio se revalúa con cada uno de los valores estimados de los factores de riesgo.
5. Se calculan las pérdidas y ganancias del portafolio, las cuales se obtienen de la diferencia entre cada uno de los valores del portafolio estimados en cada uno de los escenarios, y el valor del portafolio vigente en la fecha de valuación.
6. Se calcula el VaR con base en el nivel de confianza (percentil⁴ o cuantil) elegido.

⁴ Son los valores que dividen a la distribución en 100 partes iguales, cada una de las cuales engloba el 1% de las observaciones. En total habrá 99 percentiles.

4 Planteamiento del problema de optimización del CVaR

Considere activos $S_1, \dots, S_n, n \geq 2$ con rendimientos aleatorios. Suponga que μ_i denota el rendimiento esperado del activo S_i , y también considere x_i como la proporción de participación en el i -ésimo activo. Se representa el rendimiento esperado del portafolio resultante x como sigue:

$$E[x] = \mu_1 x_1 + \dots + \mu_n x_n = \mu^T x. \quad (1)$$

Asimismo, se supone que el conjunto de portafolios factibles es un conjunto no vacío poliédrico representado por $\Omega = \{x | Ax = b, Cx \geq d\}$, donde A es una matriz de $m \times n$, b es un vector de dimensión m , C es una matriz de $p \times n$ y d es un vector de dimensión p . En particular, una de las restricciones al conjunto Ω es $\sum_{i=1}^n x_i = 1$.

Sea $f(x, y)$ la función de pérdida cuando se elige el portafolio x de un conjunto de portafolios factibles, y es la realización de los eventos aleatorios (el vector de los rendimientos de los n activos). Considere las pérdidas de los rendimientos del portafolio, $f(x, y)$, como el negativo del rendimiento del portafolio, que es una función convexa (lineal) de las variables del portafolio x :

$$f(x, y) = -y^T x = -[y_1 x_1 + \dots + y_n x_n]. \quad (2)$$

Se supone que el vector aleatorio y tiene una función de densidad de probabilidad denotada por $p(y)$. Para un vector de decisión fijo x , la función de distribución acumulada de la pérdida asociada con ese vector se calcula como sigue:

$$\beta(x, \gamma) = \int_{f(x, y) \leq \gamma} p(y) dy. \quad (3)$$

Entonces, para un nivel de confianza α dado, se representa el α -VaR asociado al portafolio x como:

$$\text{VaR}_\alpha(x) = \min \{ \gamma \in \mathbb{R} | \beta(x, \gamma) \geq \alpha \}. \quad (4)$$

$\text{VaR}_\alpha(x)$ es el extremo izquierdo del intervalo no vacío de valores γ tal que se cumple que $\beta(x, \gamma) = \alpha$. En otras palabras, VaR_α es un valor mínimo tal que la probabilidad de que la pérdida sea menor o igual a este valor es igual a α . Generalmente en la práctica, el nivel de probabilidad, que es también llamado el nivel de confianza (α) se fija igual a 0.90, 0.95 ó 0.99.

Para variables aleatorias continuas, CVaR_α se define como el valor esperado de las pérdidas que exceden a $\text{VaR}_\alpha(x)$ [16], es decir:

$$\text{CVaR}_\alpha(x) = \frac{1}{(1 - \alpha)} \int_{f(x, y) \geq \text{VaR}_\alpha(x)} f(x, y) p(y) dy \quad (5)$$

que es equivalente a:

$$\text{CVaR}_\alpha(x) = \text{VaR}_\alpha(x) + \text{E} \left[f(x, y) - \text{VaR}_\alpha(x) \mid f(x, y) > \text{VaR}_\alpha(x) \right] \quad (6)$$

es decir, es una esperanza condicional. Observe que

$$\begin{aligned} \text{CVaR}_\alpha(x) &= \frac{1}{(1 - \alpha)} \int_{f(x, y) \geq \text{VaR}_\alpha(x)} f(x, y) p(y) dy \\ &\geq \frac{1}{(1 - \alpha)} \int_{f(x, y) \geq \text{VaR}_\alpha(x)} \text{VaR}_\alpha(x) p(y) dy \\ &= \frac{\text{VaR}_\alpha(x)}{(1 - \alpha)} \int_{f(x, y) \geq \text{VaR}_\alpha(x)} p(y) dy \\ &\geq \text{VaR}_\alpha(x). \end{aligned} \quad (7)$$

i.e. el CVaR de un activo o de un portafolio es siempre al menos tan grande como su VaR. Por consiguiente, activos o portafolios con un CVaR pequeño también tienen VaR pequeño. Además por definición de que el CVaR es una esperanza condicional, la última integral de la función de densidad es igual a 1 y como α está entre (0, 1) entonces el cociente de VaR entre $1 - \alpha$ debe de hacerse mayor y por consiguiente el CVaR es mayor al VaR.

Para una distribución de probabilidad discreta (donde el evento y_j ocurre con probabilidad p_j , para $j = 1, \dots, n$), la definición anterior de CVaR es:

$$\text{CVaR}_\alpha(x) = \frac{1}{(1 - \alpha)} \sum_{j: f(x, y_j) \geq \text{VaR}_\alpha(x)} p_j f(x, y_j). \quad (8)$$

Ejemplo 4.1. Suponga que se dispone función de pérdida $f(x, y)$ para una decisión dada x como $f(x, y) = -y$, con $y = 75 - j$ con probabilidad 1% para $j = 0, \dots, 99$. Si el valor en riesgo con $\alpha = 95\%$ es $\text{VaR}_{95\%}(x) = 20$, ya que la pérdida es de 20 ó más con el 5% de probabilidad. El valor en riesgo condicional según (8) es:

$$\text{CVaR}_{95\%}(x) = \frac{1}{0.05} (20 + 21 + 22 + 23 + 24) \times 1\% = 22.$$

Lo anterior conduce al siguiente resultado.

Teorema 4.2. Siempre se cumple que $\text{CVaR}_\alpha(x) \geq \text{VaR}_\alpha(x)$, lo que significa que CVaR de un portafolio es siempre por lo menos tan grande como su VaR. En consecuencia, portafolios con un CVaR pequeño también tienen un VaR pequeño. Sin embargo, en general minimizar VaR y CVaR no es equivalente.

Demostración. Véase [13]. □

Puesto que la definición de CVaR implica claramente la función VaR, es difícil trabajar y optimizar tal función. En cambio, se considera una función auxiliar más simple [13]:

$$F_\alpha(x, y) = \gamma + \frac{1}{(1 - \alpha)} \int_{f(x,y) \geq \gamma} (f(x, y) - \gamma)p(y)dy, \tag{9}$$

y/o

$$F_\alpha(x, y) = \gamma + \frac{1}{(1 - \alpha)} \int (f(x, y) - \gamma)^+ p(y)dy, \tag{10}$$

donde: $a^+ = \max\{a, 0\}$. Esta función, considerada como una función de γ , tiene las siguientes propiedades importantes que lo hace útil para el cálculo de VaR y CVaR [13]:

1. F_α es una función convexa de γ .
2. VaR es un punto mínimo de F_α .
3. El valor mínimo sobre γ de la función F_α es CVaR.

Como consecuencia de las propiedades anteriores, se deduce inmediatamente que, para minimizar el $\text{CVaR}_\alpha(x)$ sobre x , se tiene que minimizar la función $F_\alpha(x, \gamma)$ respecto a x y γ al mismo tiempo:

$$\min_x \text{CVaR}_\alpha(x) = \min_{x, \gamma} F_\alpha(x, \gamma). \tag{11}$$

Por lo tanto, es posible optimizar el CVaR directamente, sin necesidad de calcular el VaR primero. Puesto que se tiene el supuesto que la función de pérdida $f(x, y)$ es una función convexa (lineal) de las variables portafolio x , $F_\alpha(x, \gamma)$ es también una función convexa (lineal) de x . En este caso, dado que el conjunto de portafolios factible Ω también es convexo, los problemas de optimización en (11) son problemas de optimización convexos, los cuales se pueden resolver mediante técnicas de optimización ya conocidas para este tipo de problemas.

En lugar de usar la función de densidad $p(y)$ de los eventos aleatorios en (10), que a menudo es imposible o indeseable calcular, se puede utilizar o generar un número de escenarios y_i , con $i = 1, \dots, T$. En este caso, considere la siguiente aproximación a la función $F_\alpha(x, \gamma)$:

$$\bar{F}_\alpha(x, \gamma) = \gamma + \frac{1}{(1 - \alpha)T} \sum_{i=1}^T (f(x, y_i) - \gamma)^+. \tag{12}$$

Ahora, en el problema $\min_x \text{CVaR}_\alpha(x)$, reemplazar $F_\alpha(x, \gamma)$ con $\bar{F}_\alpha(x, \gamma)$

$$\min_{x, \gamma} \left[\gamma + \frac{1}{(1 - \alpha)T} \sum_{i=0}^T (f(x, y_i) - \gamma)^+ \right]. \quad (13)$$

Para resolver este problema de optimización, se incluyen variables artificiales z_i para reemplazar al término $(f(x, y) - \gamma)^+$. Para ello, se añaden restricciones $z_i \geq 0$ y $z_i \geq f(x, y_i) - \gamma$ al problema:

$$\begin{aligned} \min_{x, z, \gamma} & \left[\gamma + \frac{1}{(1 - \alpha)T} \sum_{i=1}^T z_i \right] \\ \text{s.a. } & z_i \geq 0, \quad i = 1, \dots, T \\ & z_i \geq f(x, y_i) - \gamma, \quad i = 1, \dots, T \\ & x \in \Omega. \end{aligned} \quad (14)$$

Puesto que $f(x, y)$ es lineal en x , todas las expresiones $z_i \geq f(x, y) - \gamma$ representan restricciones lineales, y por lo tanto el problema es un problema de programación lineal que se puede resolver de manera eficiente mediante simplex o métodos de punto interior. Es importante señalar que el problema depende de la cantidad de escenarios generados y por lo tanto, se deben emplear técnicas de programación lineal de gran escala. Una de las extensiones del problema anterior es el denominado CVaR-promedio [15].

5 Aplicación y análisis de resultados

En esta sección se considera un portafolio formado por 10 acciones al que se calcula el VaR por el método de simulación histórica. La periodicidad abarca desde el 01/01/2014 al 22/07/2015 con un total de 406 observaciones.

Cuadro 1: Resultados de VaR al 95 % de confianza por simulación histórica.

VALOR EN RIESGO POR SIMULACIÓN HISTÓRICA				
Acción	No. Títulos	Valor a 22jul15	VaR(0.95)	% VaR/Valor
AC	1	97	-1.99	2.0 %
ALFAA	1	34	-1.01	3.0 %
ALPEKA	1	24	-0.69	2.9 %
ALSEA	1	52	-1.38	2.7 %
AMXL	1	16	-0.32	2.0 %
ASURB	1	237	-6.20	2.6 %
BIMBOA	1	43	-1.04	2.4 %
BOLSAA	1	30	-0.81	2.7 %
CEMEXCPO	1	14	-0.34	2.4 %
WALMEX	1	37	-0.79	2.1 %
Total	10	583	-9.36	1.6 %

Fuente: elaboración propia.

El Cuadro 1 muestra los resultados del cálculo del VaR al 95 % de confianza para el portafolio y por cada acción. La interpretación del resultado en este cuadro es: el VaR del portafolio es de $-\$9.36$, por lo que se espera que el portafolio perderá $-\$9.36$ o menos en un 95 % de los escenarios, o bien se espera perderá $-\$9.36$ o más en el 5 % de los escenarios. En relación a la afirmación que el VaR no es una medida coherente de riesgo porque no cumple con el axioma de subaditividad, se puede verificar que el VaR del portafolio que es de $-\$9.36$ es mayor que la suma del VaR de cada activo igual $-\$14.56$.

El Cuadro 2 muestra los resultados del cálculo del VaR al 99 % de confianza para el portafolio y por cada acción. La interpretación del resultado en este cuadro es: el VaR del portafolio es de $-\$13.57$, por lo que se espera que el portafolio perderá $-\$13.57$ o menos en un 99 % de los escenarios, o bien se espera perderá $-\$13.57$ o más en el 1 % de los escenarios. Al igual que en el caso anterior se puede verificar que el VaR del portafolio es mayor que la suma del VaR de cada activo igual a $-\$22.33$.

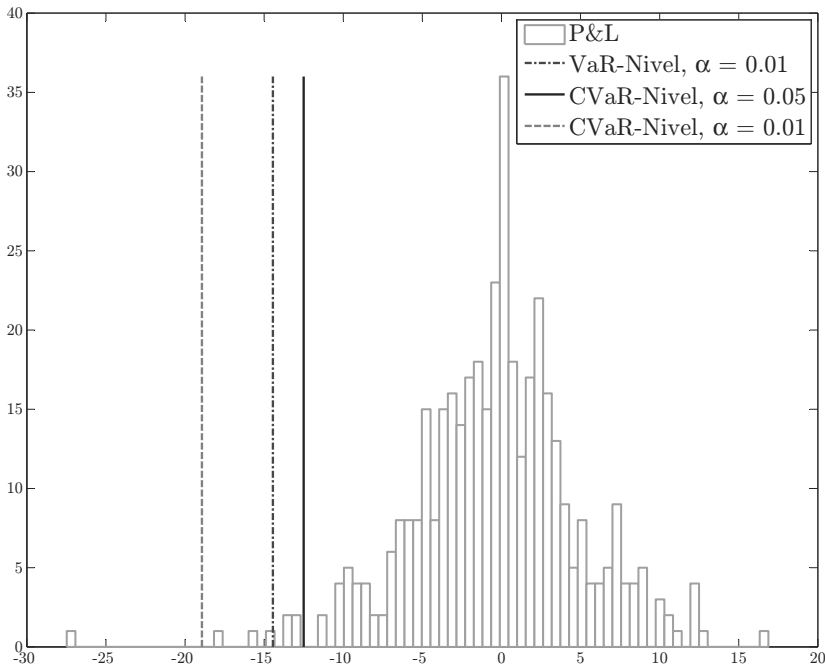
La Gráfica 1 muestra el histograma de las pérdidas y ganancias del portafolio con el VaR al 99 % de confianza y el CVaR al 95 % y al 99 % de confianza. Se observa que $CVaR_{\alpha}(x) \geq VaR_{\alpha}(x)$ con $\alpha = 0.01$.

Cuadro 2: Resultados de VaR al 99% de confianza por simulación histórica.

VALOR EN RIESGO POR SIMULACIÓN HISTÓRICA				
Acción	No. Títulos	Valor a 22jul15	VaR(0.99)	% VaR/Valor
AC	1	97	-2.81	2.9%
ALFAA	1	34	-1.55	4.6%
ALPEKA	1	24	-1.28	5.4%
ALSEA	1	52	-2.16	4.2%
AMXL	1	16	-0.56	3.5%
ASURB	1	237	-9.14	3.9%
BIMBOA	1	43	-1.65	3.9%
BOLSAA	1	30	-1.29	4.3%
CEMEXCPO	1	14	-0.49	3.4%
WALMEX	1	37	-1.39	3.7%
Total	10	583	-13.57	2.3%

Fuente: elaboración propia.

Figura 1: Histograma del vector de pérdidas y ganancias del portafolio. Con $\alpha = 0.01$, $CVaR_\alpha(x) = -18.93$, $VaR_\alpha(x) = -14.45$. (Fuente: elaboración propia).



Cuadro 3: Portafolio óptimo con CVaR al 95 % de confianza.

CVaR=-9.22		
Acción	Pesos óptimos	No. de acciones
AC	0.28	2.8
ALFAA	-0.06	-0.6
ALPEKA	-0.02	-0.2
ALSEA	0.12	1.2
AMXL	0.18	1.8
ASURB	0.13	1.3
BIMBOA	0.15	1.5
BOLSAA	0.01	0.1
CEMEXCPO	0.07	0.7
WALMEX	0.15	0.15

Fuente: elaboración propia.

Cuadro 4: Portafolio óptimo con CVaR al 99 % de confianza.

CVaR=-10.68		
Acción	Pesos óptimos	No. de acciones
AC	0.36	3.6
ALFAA	0.01	0.1
ALPEKA	0.06	0.6
ALSEA	0.15	1.5
AMXL	0.16	1.6
ASURB	0.12	1.2
BIMBOA	0.15	1.5
BOLSAA	-0.01	-0.1
CEMEXCPO	-0.12	-1.2
WALMEX	0.10	1.0

Fuente: elaboración propia.

El Cuadro 3 muestra el CVaR mínimo obtenido y los pesos óptimos al 95 % de confianza. Se observa que de los diez activos seleccionados, dos de ellos que son ALFAA y ALPEKA, la ponderación es negativa, es decir implica que se tomarán posiciones cortas (venta) en ambos activos. Para los demás activos la ponderación es positiva, es decir, se tomarán posiciones largas (compra) en tales activos.

En el Cuadro 4 muestra el CVaR mínimo obtenido y los pesos óptimos al 99 % de confianza. En este caso, se observa que de los diez activos seleccionados, dos de ellos que son BOLSAA y CEMEXCPO, la ponderación es negativa, es decir implica que se tomarán posiciones cortas (venta) en ambos activos. Para los demás activos la ponderación es positiva, es decir, se tomarán posiciones largas (compra) en tales activos. En ambos cuadros se observa que la acción con mayor ponderación es la de AC de Arca Continental, S.A.B. De C.V., lo cual se debe a que fue el activo con la volatilidad anualizada en rendimiento más pequeña (18 %) durante el periodo de tiempo considerado.

A continuación se calcula el VaR del portafolio por simulación histórica con los pesos obtenidos con el CVaR mínimo y se comparan con el CVaR mínimo tanto al 95 % como al 99 % de confianza, los resultados se muestran en el Cuadro 5.

Se observa que $\min_x \text{CVaR}_\alpha(x) \geq \text{VaR}_\alpha(x)$ para los dos niveles de confianza elegidos, además se verifica empíricamente que el CVaR de un activo o de un portafolio es siempre al menos tan grande como su VaR. Por último, de acuerdo con

Cuadro 5: Portafolio óptimo con CVaR al 95 % y 99 % de confianza.

	Min CVaR	$\sum_{i=1}^{10} VaR_i$	VaR	CVaR
95 %	-9.22	-19.91	-12.12	-17.65
99 %	-10.68	-29.90	-18.70	-29.43

Fuente: elaboración propia.

los resultados obtenidos, para el caso de los portafolios óptimos a los dos niveles de confianza, se verifica el VaR es mayor que el CVaR y que la suma del VaR de cada activo.

6 Conclusiones

En el problema de portafolio de Markowitz se plantea y resuelve la relación inversa entre la media y la varianza del rendimiento de portafolio. En su planteamiento original, el problema es encontrar un portafolio de mínima varianza con una restricción de rendimiento esperado. La varianza no es una buena medida del riesgo para distribuciones asimétricas. El Valor en Riesgo presenta varias desventajas, una de ellas es que no es una medida de riesgo coherente a no cumplir con el axioma de subaditividad. Por ello se han propuesto medidas de riesgo alternativas como el Valor en Riesgo condicional que es una medida de riesgo coherente, que está asociada a pérdidas y toma en cuenta solamente la cola izquierda de la distribución de pérdidas y ganancias de un portafolio, y en términos sencillos mide la pérdida promedio en toda la cola con un cierto nivel de probabilidad.

En esta investigación se presentó un método alternativo para la optimización de portafolios en un contexto de minimización del CVaR dada por [13]. En lugar de incluir la varianza o el Valor en Riesgo como restricción de riesgo o como función objetivo en la formación de un portafolio óptimo se utiliza el CVaR. Hay varias ventajas en el uso de CVaR como medida de riesgo y en la optimización de portafolios sobre el VaR. En primer lugar, CVaR puede utilizarse para medir y controlar el “riesgo de cola”. En segundo lugar, CVaR es una medida de riesgo más adecuada en el marco de ser una medida de riesgo coherente. En tercer lugar, CVaR es convexo con respecto a las variables de decisión y es estable, la propiedad de convexidad elimina la posibilidad de un mínimo local que es diferente de un mínimo global, por lo tanto, es más conveniente en el planteamiento y solución del problema de optimización de portafolio sin el supuesto de que los factores de riesgo de mercado se distribuyan como una normal, tal supuesto no se cumple en la mayoría de los mercados financieros, donde la distribución de los factores de riesgo de mercado se

caracterizan por tener colas gordas, asimetría, conglomerados de volatilidad y alta influencia a eventos extremos, entre otros.

Con un portafolio compuesto por una muestra de 10 acciones que componen el Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores, con una periodicidad que comprende desde el 01/01/2014 al 22/07/2015 con un total de 406 observaciones, se obtuvieron en primer lugar el VaR por el método de simulación histórica al 95 % y al 99 % de confianza con una asignación de una unidad de cada acción. Desde el enfoque de medidas coherentes riesgo el análisis de los resultados muestra que el VaR no cumple con el axioma de subaditividad, puesto que el VaR del portafolio es mayor que la suma del VaR de cada activo para los dos niveles de confianza. En segundo lugar se determinaron el CVaR mínimo y los pesos de los activos que lo conforman al 95 % y al 99 % de confianza. La evidencia empírica muestra que se verifica que $\min_x \text{CVaR}_\alpha(x) \geq \text{VaR}_\alpha(x) \geq \text{CVaR}_\alpha(x) \geq \sum_{i=1}^n \text{VaR}_\alpha^i$, con el VaR calculado por el método de simulación histórica.

Finalmente, aunque el VaR es una medida de riesgo estándar en la administración de riesgo de mercado, debido a su sencillez conceptual, la facilidad de cálculo, y su aplicabilidad, debe ser complementada con otras medidas como el CVaR. Cada medida de riesgo tiene sus ventajas y desventajas, por lo cual complementar los resultados del VaR con los del CVaR representa una manera efectiva de establecer un control del riesgo mercado más exhaustivo.

Agradecimientos

El presente trabajo ha sido apoyado por el proyecto de investigación clave SIP-20150596 de la Escuela Superior de Economía, Instituto Politécnico Nacional. De la misma manera, los autores agradecemos al dictaminador sus valiosas observaciones y recomendaciones.

Bibliografía

- [1] Acerbi, C. y D. Tasche, *On the Coherence of Expected Shortfall*, Journal of Banking & Finance, **26**(7) (2002), 1487-1503.
- [2] Acerbi, C., *Spectral measures of risk: A coherent representation of subjective risk aversion*, Journal of Banking & Finance, **26**(7) (2002), 1505-1518.
- [3] Artzner, P., F. Delbaen, J. M. Eber, y D. Heath, *Coherent Measures of Risk*, Mathematical Finance, **9**(3) (1999), 203-228.

- [4] Cho, W. N., *Robust portfolio optimization using conditional value at risk*, Final Report, Department of Computing, Imperial College London, 2008.
- [5] Climent-Hernández, J. A., F. Venegas-Martínez y F. Ortiz-Arango, *Portafolio óptimo y productos estructurados en mercados alpha-estables: un enfoque de minimización de riesgo*, Revista Nicolaita de Estudios Económicos, **10**(2) (2015), 81-106.
- [6] CNBV, Comisión Nacional Bancaria y de Valores (2014). "Disposiciones De Carácter General Aplicables A Las Instituciones De Crédito", disponible en: <http://www.cnbv.gob.mx/SECTORES-SUPERVISADOS/Paginas/default.aspx>
- [7] Cornuejols, G. y Tütüncü, R., *Optimization Methods in Finance*, Cambridge University Press, Cambridge, UK, 2006.
- [8] Hubbard, Douglas (2006). *How to Measure Anything: Finding the Value of Intangibles in Business*, John Wiley & Sons.
- [9] Krokmal, P., J. Palmquist, y S. Uryasev, *Portfolio optimization with conditional value-at-risk objective and constraints*, Journal of Risk, **4**(2) (2002), 43–48.
- [10] Jorion, P. (2007). *Value at Risk: The New Benchmark for Managing Financial Risk*. Third Edition. McGraw-Hill.
- [11] Markowitz, Harry M., *Portfolio Selection*, Journal of Finance. **7**(1) (1952), 77–91.
- [12] Quaranta, A. G. y A. Zaffaroni, *Robust optimization of conditional value at risk and portfolio selection*, Journal of Banking and Finance, **32**(10) (2008), 2046–2056.
- [13] Rockafellar R. T. y S. Uryasev, *Optimization of conditional value-at-risk*, Journal of Risk, **2**(3) (2000), 21–41.
- [14] Rockafellar R. T. y S. Uryasev, *Conditional Value-at-Risk for General Loss Distributions*, Journal of Banking and Finance, **26**(7) (2002), 1443-1471.
- [15] Salahi, M., Farshid Mehrdoust, y Farzaneh Piri, *CVaR Robust Mean-CVaR Portfolio Optimization*, ISRN Applied Mathematics, **2013**, Article ID 570950, 9 pages, 2013. doi:10.1155/2013/570950.

- [16] Szegö, G., Measures of risk, *Journal of Banking and Finance*, **26**(7) (2002), 1253-1272.
- [17] Venegas-Martínez, F., Riesgos financieros y económicos, productos derivados y decisiones económicas bajo incertidumbre. Segunda edición, Cengage, México (2008).
- [18] Wang, J. Mean-Variance-VaR Based Portfolio Optimization, Valdosta State University, 2000.
- [19] Zhu, S., y M. Fukushima, *Worst-Case Conditional Value-at-Risk with Application to Robust Portfolio Management*, *Operations Research*, **57**(5) (2009), 1155-1168.

Escuela Superior de Economía, Instituto Politécnico Nacional
Plan de Agua Prieta no. 66, Col. Plutarco Elías Calles, Delegación Miguel Hidalgo
Ciudad de México, C.P. 11340

amortiz@ipn.mx

luisal.zac@gmail.com

Universidad Autónoma Metropolitana–Unidad Xochimilco, Departamento de
Producción Económica

Calzada del Hueso 1100, Col. Villa Quietud, Delegación Coyoacán,
Ciudad de México, C.P. 04960

arnava@correo.xoc.uam.mx

Topología

Capítulo 8

Funciones de Whitney admisibles

José Gerardo Ahuatzí Reyes, David Herrera Carrasco,
Fernando Macías Romero
FCFM, BUAP

Resumen

Un continuo es un espacio métrico no vacío, compacto y conexo. Sea X un continuo. Denotaremos por 2^X al hiperespacio de los cerrados de X no vacíos y por $C(X)$ al hiperespacio de subcontinuos de X . Sea $\mathcal{H} = 2^X$ o $\mathcal{H} = C(X)$. Una función de Whitney para \mathcal{H} es una función que asigna a cada $A \in \mathcal{H}$ un valor no negativo, asigna a los conjuntos unitarios el valor nulo y es estrictamente creciente según la contención y el orden de los números reales. Una función de Whitney μ es admisible si existe una deformación de \mathcal{H} que mueve a cada punto $A \in \mathcal{H}$ desde un conjunto unitario de tal forma que la trayectoria para A determinada por esta deformación es estrictamente creciente, donde no se anula, según μ . En este capítulo se presentan algunos resultados acerca de las funciones de Whitney admisibles que han sido desarrollados previamente por varios autores. Los primeros resultados expuestos presentan algunas condiciones necesarias para la existencia de funciones de Whitney admisibles. Posteriormente, se muestra que cualquier función de Whitney para el hiperespacio de subcontinuos de cualquier dendrita dada es admisible. Por último, se muestra que, dado un continuo localmente conexo X que no posee arcos libres y dada una función de Whitney admisible μ para $C(X)$, los conjuntos de la forma $\mu^{-1}(t)$, con $t \in (0, \mu(X))$, son cubos de Hilbert; más aún, se prueba, bajo las mismas hipótesis, que la función $\mu|_{\mu^{-1}(0, \mu(X))} : \mu^{-1}(0, \mu(X)) \rightarrow (0, \mu(X))$ es una fibración trivial cuyas fibras son cubos de Hilbert.

1 Introducción

Dado un espacio métrico Y , consideraremos el siguiente conjunto

$$2^Y = \{A : A \text{ es subconjunto de } Y \text{ compacto y no vacío}\},$$

provisto con la métrica de Hausdorff. Un **hiperespacio** de Y es cualquier subconjunto de 2^Y . De particular interés resultan para este capítulo los hiperespacios 2^Y ,

$C(Y)$ y $F_1(Y)$, en donde

$$C(Y) = \{A \in 2^Y : A \text{ es conexo}\} \text{ y}$$

$$F_1(Y) = \{\{p\} : p \in Y\}.$$

Un **continuo** es un espacio métrico distinto del vacío, compacto y conexo. Una **dendrita** es un continuo localmente conexo que no contiene curvas cerradas simples. Sea X un continuo. Un **arco libre** de X es un arco A tal que $A - \{p, q\}$ es abierto en X , donde p y q son los puntos extremos de A . Sea \mathcal{H} un hiperespacio de X . Una función continua $\mu : \mathcal{H} \rightarrow [0, \infty)$ es una **función de Whitney** para \mathcal{H} si cumple las siguientes condiciones:

- (1) Para cualesquiera $A, B \in \mathcal{H}$ tales que $A \subsetneq B$, se satisface que $\mu(A) < \mu(B)$.
- (2) $\mu(A) = 0$ si y solo si $A \in \mathcal{H} \cap F_1(X)$.

H. Whitney [21] mostró en 1932 que, para cualquier continuo X , existe una función de Whitney para 2^X y, en consecuencia, para cualquier hiperespacio de X ; sin embargo, tal resultado se desarrolló con propósitos alejados del estudio de los hiperespacios. Unos años más tarde, en 1942, J. Kelley [14] utilizó por primera vez este tipo de funciones para investigar la estructura de los hiperespacios (véase [18, pág. 18]). En los años subsecuentes, las funciones de Whitney han mostrado ser una herramienta muy eficaz en la teoría de hiperespacios de continuos, así como en la teoría de continuos misma.

En 1982, J. T. Goodykoontz, Jr., y S. B. Nadler, Jr., [8] hallaron algunas condiciones sobre una función de Whitney μ , para el hiperespacio 2^X o para $C(X)$ de un continuo localmente conexo arbitrario X que no posee arcos libres, bajo las cuales los conjuntos de la forma $\mu^{-1}(t_0)$ son cubos de Hilbert, siempre que $t_0 \in (0, \mu(X))$. A una función que satisficiera tales condiciones le denominaron admisible. La siguiente definición precisa este concepto.

Definición 1.1. Sean X un continuo y $\mathcal{H} = C(X)$ o $\mathcal{H} = 2^X$. Una función de Whitney μ para \mathcal{H} es **admisible** si existe una función continua $h : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ que satisface las siguientes condiciones:

- i) Para cada $A \in \mathcal{H}$, se cumple que $h(A, 1) = A$ y $h(A, 0) \in F_1(X)$.
- ii) Si $A \in \mathcal{H}$ y $t \in [0, 1]$ son tales que $\mu(h(A, t)) > 0$, entonces, para cada $s \in [0, t)$, se cumple la desigualdad $\mu(h(A, s)) < \mu(h(A, t))$.

A h se le denomina **deformación μ -admisible**.

Tres años después de la publicación de [8], H. Kato [13] mostró que, bajo las mismas condiciones (es decir, si X es un continuo localmente conexo que no posee arcos libres y μ es una función de Whitney admisible para $C(X)$), la función $\mu|_{\mu^{-1}(0, \mu(X))} : \mu^{-1}(0, \mu(X)) \rightarrow (0, \mu(X))$ es una fibración trivial cuyas fibras son cubos de Hilbert (véase la definición).

Este capítulo tiene el objetivo de desarrollar las demostraciones de los dos resultados anteriores, así como de algunos otros resultados que fueron originalmente demostrados en [8]. Entre estos últimos destacan algunas condiciones necesarias para la existencia de funciones de Whitney admisibles, así como el hecho que todas las funciones de Whitney para el hiperespacio de subcontinuos de una dendrita arbitraria son admisibles. Para la demostración de este último resultado seguiremos la prueba hallada en [11], por estar más acorde al contexto de este capítulo.

2 Algunas observaciones previas

Los primeros dos resultados que demostramos en este capítulo son dos observaciones sencillas que hablan acerca del comportamiento de las funciones de Whitney admisibles y de la naturaleza topológica de la propiedad de poseer una función de este tipo. El primero muestra que las deformaciones admisibles, aunque no necesariamente dejan fijos los puntos de $F_1(X)$, sí mueven estos puntos únicamente dentro del conjunto $F_1(X)$.

Observación 2.1. Sea X un continuo y μ una función de Whitney para \mathcal{H} . Supóngase que μ es admisible. Sea $h : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ una deformación μ -admisible. Si $h(A, t') \in F_1(X)$ y $t, t' \in [0, 1]$ son tales que $t \leq t'$, entonces $h(A, t) \in F_1(X)$. En particular, $h(F_1(X) \times [0, 1]) \subset F_1(X)$.

Demostración. Supongamos que existen $A \in \mathcal{H}$ y $t, t' \in [0, 1)$ tales que $t < t'$, $h(A, t') \in F_1(X)$ y $h(A, t) \notin F_1(X)$. Luego, $\mu(h(A, t')) = 0$ y $s > 0$, donde $s = \mu(h(A, t))$. Como $\mu(h(\{A\} \times [t, t']))$ es un subconjunto conexo de \mathbb{R} , es un intervalo. Así, $[0, s] \subset \mu(h(\{A\} \times [t, t']))$. Sean $s'' \in (0, s)$ y $t'' \in [t, t']$ tales que $s'' = \mu(h(A, t''))$. Luego, $\mu(h(A, t)) > 0$, $t < t''$ y $\mu(h(A, t'')) < \mu(h(A, t))$, lo cual contradice el hecho que h es μ -admisible. \square

Sean Y_1 y Y_2 espacios métricos. Sea $\mathcal{H}(Y_1) = 2^{Y_1}$ o $\mathcal{H}(Y_1) = C(Y_1)$. Dada una función continua $f : Y_1 \rightarrow Y_2$, denotaremos por $f^* : \mathcal{H}(Y_1) \rightarrow \mathcal{H}(Y_2)$ la **función inducida** por f a $\mathcal{H}(Y_1)$, es decir, la función dada, para cada $A \in \mathcal{H}(Y_1)$, por

$$f^*(A) = f(A)$$

(a cada cerrado o subcontinuo de Y_1 , según sea el caso, se le asigna su imagen bajo f). Como la compacidad y la conexidad se preservan bajo funciones continuas,

f^* está siempre bien definida. Asimismo, es inmediato que $f^*(F_1(Y_1)) \subset F_1(Y_2)$. Además, se sabe que f^* es una función continua y que si f es un homeomorfismo, entonces f^* es un homeomorfismo y $(f^*)^{-1} = (f^{-1})^*$ (véase [3, Teorema 1.101 y corolario 1.102]; las demostraciones de tales resultados aplican también a espacios métricos). Esta construcción será empleada de manera repetida a lo largo del capítulo.

El siguiente resultado muestra, entre otras cosas, que la propiedad de poseer una función de Whitney admisible es una propiedad topológica.

Teorema 2.2. *Sean X y Y continuos homeomorfos y $f : Y \rightarrow X$ un homeomorfismo. Sea $\mathcal{H}(X) = C(X)$ o $\mathcal{H}(X) = 2^X$. Sea $\mu : \mathcal{H}(X) \rightarrow [0, \infty)$ una función continua y $\omega = \mu \circ f^* : \mathcal{H}(Y) \rightarrow [0, \infty)$. Si μ es una función de Whitney, entonces ω es una función de Whitney; si, además, μ es admisible, entonces ω es admisible.*

Demostración. Supongamos que μ es una función de Whitney. Como f^* es continua, ω es una composición de funciones continuas y, por tanto, ω es continua. Note además que, si B_1 y B_2 son elementos de $\mathcal{H}(Y)$ tales que $B_1 \subsetneq B_2$, entonces $f^*(B_1)$ y $f^*(B_2)$ son elementos de $\mathcal{H}(X)$ que cumplen $f^*(B_1) \subsetneq f^*(B_2)$ y, por ende, $\omega(B_1) = \mu(f^*(B_1)) < \mu(f^*(B_2)) = \omega(B_2)$. Además, dado $B \in \mathcal{H}(X)$, se cumple que $B \in F_1(Y)$ si y solo si $f^*(B) \in F_1(X)$ y esto último ocurre si y solo si $\mu(f^*(B)) = 0$, es decir, si y solo si $\omega(B) = 0$. Así, ω es una función de Whitney para $\mathcal{H}(Y)$.

Supongamos ahora que, además, μ es admisible. Sea $h : \mathcal{H}(X) \times [0, 1] \rightarrow \mathcal{H}(X)$ una deformación μ -admisible. Sea $g : \mathcal{H}(Y) \times [0, 1] \rightarrow \mathcal{H}(Y)$ la función dada por

$$g(B, t) = (f^*)^{-1} \circ h(f^*(B), t)$$

Vamos a mostrar que g es una deformación ω -admisible. Note primero que g es una composición de funciones continuas y, por ende, es continua. Dado cualquier $B \in \mathcal{H}(Y)$, se cumple que $h(f^*(B), 0) \in F_1(X)$ y así $g(B, 0) \in F_1(Y)$ (porque f^{-1} preserva cardinalidades); asimismo, $h(f^*(B), 1) = f^*(B)$, por lo cual $g(B, 1) = (f^*)^{-1}(f^*(B)) = B$. Además, dado cualquier $t \in [0, 1]$,

$$\omega(g(B, t)) = \mu \circ f^* \circ (f^*)^{-1} \circ h(f^*(B), t) = \mu \circ h(f^*(B), t)$$

Así, si B y t son tales que $\omega(g(B, t)) > 0$, entonces $\mu(h(f^*(B), t)) > 0$ y, en consecuencia, cada $s \in [0, t]$ satisface

$$\omega(g(B, s)) = \mu(h(f^*(B), s)) < \mu(h(f^*(B), t)) = \omega(g(B, t))$$

Por lo tanto, g es ω -admisible y ω es admisible. □

3 Condiciones necesarias para la existencia de funciones de Whitney admisibles

Aunque todo continuo admite funciones de Whitney, aún continuos relativamente simples, como son el continuo $\sin(1/x)$ y la circunferencia unitaria, no admiten funciones de Whitney admisibles (véase el Teorema 3.3 y el Corolario 3.9). En esta sección vamos a mostrar algunas condiciones necesarias sobre un continuo dado para la existencia de funciones de Whitney admisibles para alguno de sus hiperespacios.

Lema 3.1 ([3, Teorema 1.103]). *Sea X un espacio métrico. Entonces, la función $\phi : X \rightarrow F_1(X)$, que a cada $x \in X$ le asigna $f(x) = \{x\}$, es un homeomorfismo.*

Teorema 3.2 ([12, Corolario 14.10]). *Los hiperespacios 2^X y $C(X)$ de cualquier continuo X son continuos arco conexos.*

Teorema 3.3. *Sean X un continuo y $\mathcal{H} = 2^X$ o $\mathcal{H} = C(X)$. Si existe una función de Whitney admisible para \mathcal{H} , entonces X es arco conexo.*

Demostración. Supongamos que μ es una función de Whitney admisible para \mathcal{H} . Sea $h : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ una deformación μ -admisible. Dado $t \in [0, 1]$, sea $h_t : \mathcal{H} \rightarrow \mathcal{H}$ la función dada, para cada $A \in \mathcal{H}$, por $h_t(A) = h(A, t)$; es inmediato que h_t es continua. Como cada $A \in \mathcal{H}$ satisface $h_0(A) \in F_1(X)$, se tiene que $h_0(\mathcal{H}) \subset F_1(X)$. Como \mathcal{H} es un continuo arco conexo (Teorema 3.2) y la imagen continua de un continuo arco conexo es un continuo arco conexo, tenemos que $h_0(\mathcal{H})$ es un continuo arco conexo.

Sea $x \in X$. De manera análoga a h_0 , h^x es continua y $h^x([0, 1])$ es arco conexo, donde $h^x : [0, 1] \rightarrow \mathcal{H}$ es la función dada, para cada $t \in [0, 1]$, por $h^x(t) = h(\{x\}, t)$. Como $h^x(1) = \{x\}$, tenemos que $\mu(h^x(1)) = 0$; más aún, por la observación 2.1, $h^x([0, 1]) \subset F_1(X)$. Por otro lado, si $h^x(1) \notin h_0(\mathcal{H})$, entonces $h^x(1) \neq h_0(\{x\}) = h^x(0)$ y, en consecuencia, existe un arco α_x contenido en $h^x([0, 1])$ de $h^x(1)$ a $h^x(0)$; así, α_x es un arco contenido en $F_1(X)$ que va de $\{x\}$ al elemento $h^x(0)$ de $h_0(\mathcal{H})$. Por el contrario, si $h^x(1) \in h_0(\mathcal{H})$, sea $\alpha_x = \{x\}$. Como la unión de espacios arco conexos que se intersectan es arco conexa, $h_0(\mathcal{H}) \cup \alpha_x$ es un subconjunto de $F_1(X)$ que es arco conexo, en cualquiera de los dos casos. Del mismo modo, dado $y \in X$, $h_0(\mathcal{H}) \cup \alpha_x \cup \alpha_y$ es un subconjunto de $F_1(X)$ que es arco conexo, donde α_y se define de manera análoga a α_x . Por tanto, existe un arco contenido en $F_1(X)$ que va de $\{x\}$ a $\{y\}$. Esto muestra que $F_1(X)$ es arco conexo. Como X es homeomorfo a $F_1(X)$ (Teorema 3.1), se concluye que X es arco conexo. \square

Lema 3.4 ([3, Lema 1.109]). *Sean $a, b \in \mathbb{R}$ tales que $a < b$. Entonces, la función $\iota : [a, b] \rightarrow C([a, b])$ dada por*

$$\iota(t) = [a, t]$$

es continua.

Lema 3.5 ([18, Lema 1.43]). *Sea X un continuo. Si Λ es un subconjunto conexo de 2^X tal que $\Lambda \cap C(X) \neq \emptyset$, entonces $\bigcup \Lambda$ es un subconjunto conexo de X .*

Lema 3.6 ([18, Lema 1.48]). *Sea X un continuo. Entonces, la función $\bigcup : 2^{2^X} \rightarrow 2^X$, que a cada $A \in 2^{2^X}$ le asigna $\bigcup A$, es continua.*

Teorema 3.7. *Supongamos que existe una función de Whitney admisible para $\mathcal{H} = 2^X$ o $C(X)$. Entonces, X es contráctil si y solo si \mathcal{H} es contráctil.*

Demostración. Supongamos que \mathcal{H} es contráctil. Luego, existen $A_0 \in \mathcal{H}$ y una función continua $F' : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ tal que $F'(A, 0) = A$ y $F'(A, 1) = A_0$, para cada $A \in \mathcal{H}$. Como todo espacio contráctil es conexo por trayectorias ([1, teorema 3.1.30]), existe una función continua $\alpha : [0, 1] \rightarrow \mathcal{H}$ tal que $\alpha(0) = A_0$ y $\alpha(1) = X$. Sea $F : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ la función

$$F(A, t) = \begin{cases} F'(A, 2t), & 0 \leq t \leq \frac{1}{2} \\ \alpha(2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Como cualquier $A \in \mathcal{H}$ satisface que $F'(A, 2(\frac{1}{2})) = F'(A, 1) = A_0$ y $\alpha(2(\frac{1}{2}) - 1) = \alpha(0) = A_0$, la función α está bien definida y es continua.

Sea $G : \mathcal{H} \times [0, 1] \rightarrow 2^X$ la función

$$G(A, t) = \bigcup \{F(A, t') : t' \in [0, t]\}.$$

Afirmación. G está bien definida, es continua y $G(\mathcal{H} \times [0, 1]) \subset \mathcal{H}$.

Sea $\gamma : \mathcal{H} \times [0, 1] \rightarrow C(\mathcal{H}) \times C([0, 1])$ la función

$$\gamma(A, t) = (\{A\}, [0, t]).$$

Note que $\gamma = (\phi, \iota)$, donde $\phi : \mathcal{H} \rightarrow C(\mathcal{H})$ y $\iota : [0, 1] \rightarrow C([0, 1])$ son las funciones $\phi(A) = \{A\}$ y $\iota(t) = [0, t]$. Como estas dos últimas funciones son continuas (por los lemas 3.1 y 3.4), γ es continua. Sea $\theta : C(\mathcal{H}) \times C([0, 1]) \rightarrow C(\mathcal{H} \times [0, 1])$ la función $\theta(A, B) = A \times B$. Por el Lema 4.3, θ es continua. Además, por el Lema 3.6, la función unión $\bigcup : 2^{2^X} \rightarrow 2^X$ es continua. Como cada $(A, t) \in \mathcal{H} \times I$ satisface $F^* \circ \theta \circ \gamma(A, t) = F^*(\{A\} \times [0, t]) = \{F(A, t') : t' \in [0, t]\}$, se tiene que $G = \bigcup \circ F^* \circ \theta \circ \gamma$, y así G está bien definida y es continua.

Por otro lado, si $\mathcal{H} = 2^X$, es inmediato que $G(\mathcal{H} \times [0, 1]) \subset \mathcal{H}$. Supongamos que $\mathcal{H} = C(X)$. Luego, cualquier $(A, t) \in \mathcal{H} \times [0, 1]$ satisface que $F(A, 0) \in C(X)$,

$\{A\} \times [0, t]$ es conexo y, por consiguiente, que $F^*(\{A\} \times [0, t])$ es un subconjunto conexo de 2^X que intersecta a $C(X)$; aplicando el Lema 3.5, se obtiene que $G(A, t)$ es conexo, es decir, $G(A, t) \in C(X)$. Por tanto, también en este caso, $G(\mathcal{H} \times [0, 1]) \subset \mathcal{H}$. Esto prueba la Afirmación.

Sea $h : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ una deformación μ -admisibles. Sea $f : F_1(X) \times [0, 1] \rightarrow 2^X$ la función dada por

$$f(\{x\}, t) = \begin{cases} h(\{x\}, 1 - 2t), & 0 \leq t \leq \frac{1}{2}, \\ h(G(\{x\}, 2t - 1), 0), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

Como cada $x \in X$ satisface $h(\{x\}, 1 - 2(\frac{1}{2})) = h(\{x\}, 0)$ y $h(G(\{x\}, 2(\frac{1}{2}) - 1), 0) = h(G(\{x\}, 0), 0) = h(F(\{x\}, 0)) = h(\{x\}, 0)$, la función está bien definida y es continua. Note que $f(\{x\}, 0) = h(\{x\}, 1) = \{x\}$. Asimismo, como $F(\{x\}, 1) = X$, se cumple que $G(\{x\}, 1), 0 = X$ y así $f(\{x\}, 1) = h(X, 0)$. Por último, por la observación 2.1, cada $(\{x\}, t) \in F_1(X) \times [0, 1]$ satisface que $h(\{x\}, 1 - 2t) \in F_1(X)$, si $0 \leq t \leq \frac{1}{2}$, y, directamente porque h es admisible, $h(G(\{x\}, 2t - 1), 0) \in F_1(X)$, si $\frac{1}{2} \leq t \leq 1$. En consecuencia, $f(\{x\}, t) \in F_1(X)$, en cualquier caso. De este modo $f(F_1(X) \times [0, 1]) \subset F_1(X)$. Por tanto, f es una contracción de $F_1(X)$.

Supongamos ahora que X es contráctil. Sean $x_0 \in X$ y $f' : X \times [0, 1] \rightarrow X$ una función continua tal que cualquier $x \in X$ cumple $f'(x, 0) = x$ y $f'(x, 1) = x_0$. Sea $G' : 2^X \times I \rightarrow 2^X$ la función

$$G'(A, t) = \{f'(a, t) : a \in A\}$$

Como $G' = (f')^* \circ \theta(Id_{C(X)}, \phi')$, donde θ es la función descrita en la parte previa de esta demostración y $\phi' : [0, 1] \rightarrow C([0, 1])$ es la función $\phi'(t) = \{t\}$ (la cual es continua por el Lema 3.1), se tiene que G' es continua. Además, cada $A \in 2^X$ satisface que $G'(A, 0) = \{a : a \in A\} = A$ y $G'(A, 1) = \{x_0\}$. Más aún, si $A \in C(X)$, entonces $\theta(A, \phi'(t)) = A \times \{t\}$ es conexo y así, como $G'(A, t)$ es la imagen continua de $\theta(A, \phi'(t))$ bajo f , es conexo; por consiguiente, $G'(A, t) \in C(X)$. Lo anterior muestra que $G'(C(X)) \subset C(X)$. Por lo tanto, G' es una contracción de 2^X y la restricción correspondiente de G' a $C(X)$ es una contracción de $C(X)$. De este modo, \mathcal{H} es contráctil. □

Dado un continuo X , un hiperespacio de crecimiento es un subconjunto \mathcal{C} de 2^X tal que cualesquiera $A \in \mathcal{C}$ y $B \in 2^X$ cumplen la siguiente implicación: si $A \subset B$ y cada componente de B intersecta a A , entonces $B \in \mathcal{C}$.

Para una prueba detallada del teorema 3.8, véase [4, Teorema 4.14].

Teorema 3.8 ([14], comentario posterior al Teorema 4.5). *Si X es un continuo localmente conexo, entonces cada hiperespacio de crecimiento de X es un retracto absoluto.*

Corolario 3.9. Sean X un continuo y $\mathcal{H} = 2^X$ o $\mathcal{H} = C(X)$. Si existe una función de Whitney admisible para \mathcal{H} y si X es un continuo localmente conexo, entonces X es contráctil.

Demostración. Veamos primero que \mathcal{H} es un hiperespacio de crecimiento. Es claro que 2^X lo es. Supongamos que $A \in C(X)$ y $B \in 2^X$ son tales que $A \subset B$ y cada componente de B intersecta a A . Sean B_1 y B_2 componentes de B . Luego $B_1 \cap A \neq \emptyset$ y $B_2 \cap A \neq \emptyset$. Así, $B_1 \cup B_2 \cup A$ es conexo. Sea C la componente de B que contiene a $B_1 \cup B_2 \cup A$. Como cada par de componentes de B distintas son ajenas, se sigue que $B_1 = C = B_2$. Por tanto, B solo tiene una componente, es decir, $B \in C(X)$. De este modo, $C(X)$ es un hiperespacio de crecimiento.

De este modo, por el Teorema 3.8, \mathcal{H} es un retracto absoluto. En particular, \mathcal{H} es contráctil ([10, Teorema 7.1]). Así, por el Teorema 3.7, X es contráctil. \square

4 El hiperespacio de subcontinuos de una dendrita admite funciones de Whitney admisibles

De acuerdo al Corolario 3.9, incluso dentro de la clase de los continuos localmente conexos existen continuos cuyos hiperespacios no admiten funciones de Whitney admisibles, a saber, los continuos localmente conexos que no son contraíbles (por ejemplo, la circunferencia unitaria y el arete hawaiano). Sin embargo existen otros espacios en la misma clase para los cuales toda función de Whitney (para el hiperespacio de cerrados o el de subcontinuos) es admisible. En esta sección se muestra que cada **dendrita**, es decir, cada continuo localmente conexo que no contiene curvas cerradas simples, posee tal propiedad (para el hiperespacio de subcontinuos).

Un **dendroide** es un continuo arco conexo y hereditariamente unicoherente. Sea X un dendroide. Dados elementos distintos p y q de X , existe un único arco en X cuyos puntos extremos son p y q (véase [16, Teorema 1]); a dicho arco le denotaremos por $[p, q]$. Si $p = q$, $[p, q]$ denotará el conjunto $\{p\}$. Se dice que un continuo X es **suave en un punto** $p \in X$ si para cualquier sucesión $\{a_n\}$ tal que $\lim a_n = a$, para algún $a \in X$, se cumple que $\lim [p, a_n] = [p, a]$.

Lema 4.1 ([6, Corolario 4]). *Cualquier dendrita es un dendroide y es suave en cada uno de sus puntos.*

Como cualquier dendrita X es un continuo únicamente arco conexo, cualquier subconjunto U de X arco conexo y cualesquiera $x, y \in U$ cumplen que $[x, y] \subset U$. Este hecho se utilizará en repetidas ocasiones en la prueba del siguiente lema.

Lema 4.2. *Sea X una dendrita. Entonces, la función $\alpha : X \times X \rightarrow C(X)$ dada, para cualesquiera $x, y \in X$, por $\alpha(x, y) = [x, y]$ es continua.*

Demostración. Sea $\{(x_n, y_n)\}_{n \in \mathbb{N}}$ una sucesión en $X \times X$. Vamos a probar que $\lim [x_n, y_n] = [x, y]$. Tenemos los dos casos siguientes:

- (i) $x = y$. Sea $\varepsilon > 0$. Luego, existe un abierto U de X conexo y tal que $x \in U \subset B(x, \varepsilon)$. Sea $N \in \mathbb{N}$ tal que, si $n \geq N$, entonces $[x_n, y_n] \subset U$. Como U es arco conexo (Teorema 5.9), se tiene que $[x_n, y_n] \subset U \subset B(x, \varepsilon)$. Así, si $n \geq N$, entonces $H_d(\{x\}, [x_n, y_n]) < \varepsilon$. Esto muestra que $\lim [x_n, y_n] = [x, y]$.
- (ii) $x \neq y$. Sean U y V abiertos de X ajenos, conexos y tales que $x \in U$, $y \in V$, $U \cap \{p, y\} = \emptyset$ y $V \cap \{p, x\} = \emptyset$. Sea $N \in \mathbb{N}$ tal que si $n \geq N$, entonces $x_n \in U$ y $y_n \in V$. Sea $n \geq N$. Si $x_n \in [x, y]$, sea $x'_n = x_n$. Supongamos que $x_n \notin [x, y]$. Como U es arco conexo (Teorema 5.9), $[x_n, x] \subset U$. Sea $x'_n \in [x_n, x]$ tal que $[x_n, x'_n] \cap [x, y] = \{x'_n\}$. Así, en cualquiera de los dos casos anteriores, $[x_n, x'_n] \subset U$, $[x_n, x'_n] \cap [x, y] = \{x'_n\}$ y $[x_n, x'_n] \cup [x'_n, y] = [x_n, y]$ (esto último porque la unión es un arco). Análogamente, existe $y'_n \in [x, y]$ tal que $[y'_n, y_n] \subset V$, $[x, y] \cap [y'_n, y_n] = \{y'_n\}$ y $[x, y'_n] \cup [y'_n, y_n] = [x, y_n]$. Note que $[x_n, x'_n] \cap [y'_n, y_n] \subset U \cap V = \emptyset$, por lo cual $[x_n, x'_n] \cup [x'_n, y'_n] \cup [y'_n, y_n]$ es un arco, el cual, por la propiedad de unicidad de arcos de X , es $[x_n, y_n]$. Por otro lado, note que, por la arco conexidad de U y de V , $[x_n, x'_n] \subset U$ y $[y'_n, y_n] \subset V$. Así, si $p \notin [x'_n, y'_n]$, entonces $p \in [x_n, x'_n]$ o bien $p \in [y'_n, y_n]$ y, por ende, $p \in U$ o $p \in V$. Como esto último no es posible, $p \in [x'_n, y'_n]$ y, por consiguiente, $p \in [x_n, y_n]$. Como n representa a cualquier número natural mayor o igual que N , podemos concluir que $\lim [x_n, y_n] = \lim ([x_n, p] \cup [p, y_n]) = (\lim [x_n, p]) \cup (\lim [p, y_n]) = [x, p] \cup [p, y] = [x, y]$ (la penúltima igualdad se cumple por [18, 4.15]). □

Lema 4.3. Sean Y_1 y Y_2 continuos. Sea $\theta : C(Y_1) \times C(Y_2) \rightarrow C(Y_1 \times Y_2)$ la función

$$\theta(A, B) = A \times B.$$

Entonces, θ es continua.

Demostración. Como la conexidad y la compacidad se conservan bajo productos, θ está bien definida. Sean d_1 y d_2 las métricas de Y_1 y de Y_2 , respectivamente. Sean $\rho : Y_1 \times Y_2 \rightarrow [0, \infty)$ y $R : (C(Y_1) \times C(Y_2)) \times (C(Y_1) \times C(Y_2))$ las funciones

$$\begin{aligned} \rho((x_1, x_2), (y_1, y_2)) &= \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}, \\ R((A_1, A_2), (B_1, B_2)) &= \max\{H_{d_1}(A_1, B_1), H_{d_2}(A_2, B_2)\}. \end{aligned}$$

Se sabe que ρ y R son métricas compatibles con las topologías producto de $X \times X$ y $C(X) \times C(X)$, respectivamente.

Supongamos que los elementos (A, B) y (E, F) de $C(Y_1) \times C(Y_2)$ y $\varepsilon > 0$ son tales que $R((A, B), (E, F)) < \varepsilon$. Luego, $A \subset N_{d_1}(E, \varepsilon)$ y $B \subset N_{d_2}(F, \varepsilon)$. Sea $(x, y) \in A \times B$. Sean $z \in E$ y $w \in F$ tales que $d_1(x, z) < \varepsilon$ y $d_2(y, w) < \varepsilon$. Luego, $(z, w) \in E \times F$ y $\rho((x, y), (z, w)) < \varepsilon$. Así, $A \times B \subset N_\rho(E \times F, \varepsilon)$. Del mismo modo se prueba que $E \times F \subset N_\rho(A \times B, \varepsilon)$. Por tanto, $H_\rho(A \times B, E \times F) < \varepsilon$.

El párrafo anterior muestra, en particular, que si $\{(A_n, B_n)\}_{n \in \mathbb{N}}$ es una sucesión en $C(Y_1) \times C(Y_2)$ tal que $\lim(A_n, B_n) = (A_0, B_0)$, para algún $(A_0, B_0) \in C(Y_1) \times C(Y_2)$, entonces $\lim(A_n \times B_n) = A_0 \times B_0$. Por lo tanto, θ es continua. \square

Observación 4.4. Dados un continuo X , una función de Whitney admisible μ para $C(X)$ y cualquier $t_0 \in (0, \mu(X))$, se cumple que $\mu^{-1}(t_0)$ separa a $C(X)$ en dos subconjuntos abiertos (y disjuntos), a saber, $\mu^{-1}([0, t_0])$ y $\mu^{-1}(t_0, \mu(X)]$. Así, si \mathcal{K} es un subconjunto conexo de $C(X)$ y existen $A, B \in \mathcal{K}$ tales que $\mu(A) \leq t_0$ y $\mu(B) \geq t_0$, entonces existe $D \in \mathcal{K}$ tal que $D \in \mu^{-1}(t_0)$. Este hecho se usará de forma muy frecuente en las demostraciones de los siguientes resultados.

La prueba del siguiente teorema, el cual es el resultado principal de esta sección, es original de Illanes y Leonel [11]. Aunque el resultado puede ser extendido a los dendroides suaves [8, Teorema 2.17], los argumentos de [8] requieren de conceptos alejados de los objetivos de este capítulo.

Teorema 4.5. *Sea X una dendrita. Entonces, toda función de Whitney para $C(X)$ es admisible.*

Demostración. Sea $\mu : C(X) \rightarrow [0, \infty)$ una función de Whitney. Fijemos $p_0 \in X$. Dados $A \in C(X)$ y $a, b \in A - \{p_0\}$, sean $a' \in [p_0, a]$ y $b' \in [p_0, b]$ tales que $[p_0, a'] \cap A = \{a'\}$ y $[p_0, b'] \cap A = \{b'\}$. Si $a' \neq b'$, entonces existe un arco α contenido en A con puntos extremos a' y b' . Note que $\alpha = [a', b']$ y $[p_0, a'] \cap [a', b'] = [p_0, a'] \cap A = \{a'\}$. Así, $[p_0, a'] \cup [a', b']$ es un arco con puntos extremos p_0 y b' y, por consiguiente, $[p_0, a'] \cup [a', b'] = [p_0, b']$. Pero $[p_0, b'] \cap A = \{b'\}$, lo cual contradice que $[a', b'] \subset A \cap ([p_0, a'] \cup [a', b'])$. Así, $a' = b'$. De este modo, podemos considerar la función $f : C(X) \rightarrow X$, que a cada $A \in C(X)$ asigna $f(A)$, el único punto de A tal que $[p_0, f(A)] \cap A = \{f(A)\}$.

Afirmación 1. f es continua.

Supongamos que $\{A_k\}_{k \in \mathbb{N}}$ es una sucesión tal que $\lim A_k = A$, para algún $A \in C(X)$. Si $p_0 \in A$, entonces existe una sucesión $\{p_k\}_{k \in \mathbb{N}}$ tal que $\lim p_k = p$ y, para cada $k \in \mathbb{N}$, $p_k \in A_k$. Luego, $\lim [p_0, p_k] = \{p_0\}$. Como $f(A_k) \in [p_0, p_k]$, tenemos que $\lim f(A_k) = p_0$. Supongamos que $p_0 \notin A$. Sea $a_k \in A_k$, para cada $k \in \mathbb{N}$, tal que $\lim a_k = a$, para alguna $a \in A$. Luego, $f(A_k) \in [p_0, a_k]$. Sea $\{f(A_{n_k})\}_{n \in \mathbb{N}}$ una

subsucesión convergente de $\{f(A_k)\}_{k \in \mathbb{N}}$. Sea $a' \in X$ tal que $\lim f(A_{n_k}) = a'$. Note que $a' \in A$. Sea $b_k \in A_k$ tal que $\lim b_k = f(A)$. Como $f(A_k) \in [p_0, b_k]$, tenemos que $[p_0, a'] = \lim [p_0, f(A_{n_k})] \subset \lim [p_0, b_k] = [p_0, f(A)]$. Así, $f(A) = a'$. Por tanto, toda subsucesión convergente de $\{f(A_k)\}$ converge a $f(A)$, lo cual implica que $\{f(A_k)\}$ converge a $f(A)$. Esto muestra que f es continua.

Sea $g : C(X) \rightarrow [0, 1]$ la función dada, para cada $A \in C(X)$ por

$$g(A) = \sup\{\mu([f(A), a]) : a \in A\}.$$

Afirmación 2. g es continua y, para cada $A \in C(X)$, existe $a_0 \in A$ tal que $g(A) = \mu([f(A), a_0])$.

Para mostrar que g es continua, vamos a expresarla como la composición de funciones continuas. Sea $\gamma : C(X) \rightarrow C(X) \times C(X)$ la función

$$\gamma(A) = (\{f(A)\}, A)$$

Note que $\gamma = (\phi \circ f, Id_{C(X)})$, donde $\phi : X \rightarrow C(X)$ es la función dada por $\phi(x) = \{x\}$, la cual es continua, por el Lema 3.1. De este modo, γ es la composición de funciones continuas y, por tanto, es continua. Sean $\theta : C(X) \times C(X) \rightarrow C(X \times X)$ y $\alpha : X \times X \rightarrow C(X)$ las funciones

$$\begin{aligned} \theta(A, B) &= A \times B, \\ \alpha(x, y) &= [x, y]. \end{aligned}$$

Por los lemas 4.3 y 4.2, respectivamente, las funciones θ y α son continuas. Sea $g_0 = \mu \circ \alpha : X \times X \rightarrow [0, \mu(X)]$. Luego, la función inducida por g_0 a $C(X \times X)$, $g_0^* : C(X \times X) \rightarrow C([0, \mu(X)])$, es continua y

$$g_0^*(D) = \{\mu([x, y]) : (x, y) \in D\},$$

para cada $D \in C(X \times X)$. Así, para cada $A \in C(X)$,

$$\begin{aligned} \sup \circ g_0^* \circ \theta \circ \gamma(A) &= \sup\{\mu([x, y]) : (x, y) \in \theta(\gamma(A))\} \\ &= \sup\{\mu([x, y]) : (x, y) \in (\{f(A)\}, A)\} \\ &= \sup\{\mu([f(A), a]) : a \in A\}; \\ &= g(A). \end{aligned}$$

Por tanto, $g = \sup \circ g_0^* \circ \theta \circ \gamma$. Esto prueba que g es continua. Además, como cada $A \in C(X)$ cumple que $g_0^* \circ \theta \circ \gamma(A)$ es cerrado en $[0, \mu(X)]$ (de hecho, un subcontinuo),

se tiene que $g(A) \in g_0^* \circ \theta \circ \gamma(A)$. Así, existe $a_0 \in A$ tal que $g(A) = \mu([f(A), a_0])$. Esto concluye la demostración de la Afirmación 2.

Sea $h : C(X) \times [0, 1] \rightarrow C(X)$ la función definida por

$$h(A, t) = \{a \in A : \mu([f(A), a]) \leq tg(A)\}$$

Vamos a mostrar algunas propiedades de h .

Afirmación 3. h está bien definida.

Sea $(A, t) \in C(X) \times [0, 1]$. Dado $a \in h(A, t)$, cualquier $b \in [f(A), a]$ cumple que $[f(A), b] \subset [f(A), a]$ y, por ende, que $\mu([f(A), b]) \leq \mu([f(A), a]) \leq tg(A)$ y $b \in h(A, t)$; así, $[f(A), a] \subset h(A, t)$. Así, $h(A, t) = \bigcup\{[f(A), a] : a \in h(A, t)\}$ y, como cada $[f(A), a]$ es conexo y contiene a $f(A)$, el conjunto $h(A, t)$ es conexo. Por otro lado, sean $\{a_n\}_{n \in \mathbb{N}}$ una sucesión en $h(A, t)$ (el cual es subconjunto de A) y $a \in A$ tales que $\lim a_n = a$. Como μ es continua y X es suave en cada uno de sus puntos, entonces $\mu([f(A), a]) = \lim \mu([f(A), a_n]) \leq tg(A)$. Así, $a \in h(A, t)$. Esto muestra que $h(A, t)$ es cerrado en A y, por consiguiente, en X . Por tanto, $h(A, t) \in C(X)$.

Afirmación 4. h es continua.

Sean $\{(A'_n, t'_n)\}_{n \in \mathbb{N}}$ y (A, t) una sucesión y un punto en $C(X) \times [0, 1]$, respectivamente, tales que $\lim (A'_n, t'_n) = (A, t)$. Sea $\{(A_k, t_k)\}_{k \in \mathbb{N}}$ una subsucesión de $\{(A'_n, t'_n)\}_{n \in \mathbb{N}}$ tal que $\{h(A_k, t_k)\}_{k \in \mathbb{N}}$ converge a algún $B \in C(X)$. Vamos a probar que $B = h(A, t)$. Sea $b \in B$. Luego, existe una sucesión $\{b_k\}_{k \in \mathbb{N}}$ en X tal que $\lim b_k = b$ y, para cada $k \in \mathbb{N}$, $b_k \in h(A_k, t_k)$. Por la continuidad de f , se tiene que $\lim f(A_k) = f(A)$ y, por el Lema 4.2, $\lim [f(A_k), b_k] = [f(A), b]$. De este modo, por la continuidad de μ y de g , se cumple que

$$\mu([f(A), b]) = \mu(\lim [f(A_k), b_k]) = \lim \mu([f(A_k), b_k]) \leq \lim t_k g(A_k) = tg(A).$$

Así, $b \in h(A, t)$. Por tanto, $B \subset h(A, t)$. Recíprocamente, supongamos que $a \in h(A, t)$. Vamos a probar que $a \in B$. Como $h(A, t) \subset A$ y $\lim A_k = A$, existe una sucesión $\{x_k\}_{k \in \mathbb{N}}$ en X tal que $\lim x_k = a$ y, para cada $k \in \mathbb{N}$, $x_k \in A_k$. Dado $k \in \mathbb{N}$, vamos a definir el punto y_k de la siguiente manera. Si $\mu([f(A_k), x_k]) \leq t_k g(A_k)$, entonces sea $y_k = x_k$. Por el contrario, si $t_k g(A_k) < \mu([f(A_k), x_k])$, entonces, $\mu(\alpha(f(A_k), f(A_k))) = \mu([f(A_k), f(A_k)]) = 0 \leq t_k g(A_k) < \mu(\alpha(f(A_k), x_k))$; así, el continuo $\alpha(\{f(A_k)\} \times [f(A_k), x_k])$ cumple las hipótesis de la observación 4.4 con $t_0 = t_k g(A_k)$ y, por consiguiente, existe $y_k \in [f(A_k), x_k]$ tal que $\mu([f(A_k), y_k]) = t_k g(A_k)$. Note que, en cualquier caso, $y_k \in h(A_k, t_k)$. En consecuencia, si $y_k = x_k$

para una cantidad infinita de $k \in \mathbb{N}$, entonces $a \in \liminf h(A_k, t_k) = B$, porque $\lim x_k = a$. Así, podemos suponer que existe $N \in \mathbb{N}$ tal que, si $k \geq N$, entonces $y_k \neq x_k$. Así, para cada $k \geq N$, $\mu([f(A_k), y_k]) = t_k g(A_k)$. En particular, por la continuidad de g ,

$$\lim \mu([f(A_k), y_k]) = \lim t_k g(A_k) = tg(A)$$

Por otro lado, como X es compacto, existen $y \in X$ y una subsucesión $\{y_{j_k}\}_{k \in \mathbb{N}}$ de $\{y_k\}_{k \in \mathbb{N}}$ tales que $\lim y_{j_k} = y$. Podemos suponer que $j_k \geq N$, para cada $k \in \mathbb{N}$. Note que $y \in \liminf h(A_k, t_k) = B$. Además, como $\lim [f(A_{j_k}), x_{j_k}] = [f(A), a]$ y cada y_{j_k} es elemento de $[f(A_{j_k}), x_{j_k}]$, tenemos que $y \in [f(A), a]$. De este modo, $[f(A), y] \subset [f(A), a]$. Además, por la continuidad de μ y de g ,

$$\mu([f(A), y]) = \mu(\lim [f(A_{j_k}), y_{j_k}]) = \lim \mu([f(A_{j_k}), y_{j_k}])$$

Luego, $\mu([f(A), y]) = tg(A)$. Como $a \in h(A, t)$, $tg(A) \geq \mu([f(A), a])$. Por tanto, $[f(A), y] \subset [f(A), a]$ y $\mu([f(A), y]) \geq \mu([f(A), a])$, lo cual implica que $[f(A), y] = [f(A), a]$. Así, $a = y \in B$. Esto muestra que $h(A, t) \subset B$. Por lo tanto, $B = h(A, t)$. Esto muestra que h es continua.

Afirmación 5. Para cada $A \in C(X)$, $h(A, 1) = A$ y $h(A, 0) = \{f(A)\}$.

Por la elección de g , cada $a \in A$ satisface $\mu([f(A), a]) \leq g(A)$ y, por ende, se tiene que $h(A, 1) = \{a \in A : \mu([f(A), a]) \leq g(A)\} = A$. Además, si $a \in A$ es tal que $\mu([f(A), a]) = 0$, entonces $[f(A), a] \in F_1(X)$ y, por ende, $[f(A), a] = \{f(A)\}$ y $a = f(A)$. De este modo, $h(A, 0) = \{a \in A : \mu([f(A), a]) = 0\} = \{f(A)\}$.

Afirmación 6. Si $A \in C(X)$ y $t \in [0, 1]$ son tales que $h(A, t) > 0$, entonces $\mu(h(A, s)) < \mu(h(A, t))$, para cada $s \in [0, t)$.

Supongamos que $A \in C(X)$ y $t \in [0, 1]$ son tales que $h(A, t) > 0$. Sea $s \in [0, t)$. Como $sg(A) \leq tg(A)$, se tiene que $\{a \in A : \mu([f(A), a]) \leq sg(A)\} \subset \{a \in A : \mu([f(A), a]) \leq tg(A)\}$, es decir, $h(A, s) \subset h(A, t)$. Si $h(A, s) \subsetneq h(A, t)$, entonces $\mu(h(A, s)) < \mu(h(A, t))$. Supongamos que $h(A, s) = h(A, t)$. Por la Afirmación 2, existe $a_0 \in A$ tal que $\mu([f(A), a_0]) = g(A)$. Como $tg(A) \leq g(A)$, se tiene que $\mu(\alpha(f(A), f(A))) = \mu([f(A), f(A)]) = 0 \leq tg(A) \leq \mu([f(A), a_0]) = \mu(\alpha(f(A), a_0))$; así, el continuo $\alpha(\{f(A)\} \times [f(A), a_0])$ cumple las hipótesis de la observación 4.4 con $t_0 = tg(A)$ y, por consiguiente, existe $a_1 \in [f(A), a_0]$ tal que $\mu([f(A), a_1]) = tg(A)$. Note que $a_1 \in h(A, t) = h(A, s)$. Así, $\mu([f(A), a_1]) \leq sg(A)$. Por tanto, $tg(A) \leq sg(A)$. Como $s < t$, lo anterior implica que $g(A) = 0$ y, en consecuencia, $h(A, t) = \{a \in A : \mu([f(A), a]) \leq 0\} = h(A, 0)$. Luego, por la Afirmación 5,

$h(A, t) = \{f(A)\}$ y, por ende, $\mu(h(A, t)) = 0$. Esto contradice la hipótesis inicial de que $\mu(h(A, t)) > 0$. Por lo tanto, $\mu(h(A, s)) < \mu(h(A, t))$.

Por las afirmaciones 3, 4, 5 y 6, se tiene que h es μ -admisibles. Esto concluye la demostración de que μ es admisible. \square

5 Niveles de Whitney que son cubos de Hilbert

Determinar un modelo geométrico para cierto **nivel de Whitney** (es decir, un conjunto de la forma $\mu^{-1}(t)$, donde μ es una función de Whitney y t es un número no negativo), generalmente no es una tarea sencilla (salvo para el valor 0 y el valor asignado por μ al espacio total). En esta sección vamos a mostrar que para cada elemento de cierta clase de continuos localmente conexos, casi todos los niveles de Whitney de su hiperespacio de continuos son homeomorfos al cubo de Hilbert.

Para probar el teorema 5.2 utilizaremos el siguiente lema, el cual es un sencillo resultado que se deja demostrar al lector.

Lema 5.1. *Sean X un espacio métrico, x un elemento de X y $\{x_n\}_{n \in \mathbb{N}}$ una sucesión en X . Entonces, $\lim x_n = x$ si y solo si, para cualquier subsucesión $\{x_{k_n}\}_{n \in \mathbb{N}}$ de $\{x_n\}_{n \in \mathbb{N}}$, existe una subsucesión $\{x_{j_n}\}_{n \in \mathbb{N}}$ de $\{x_{k_n}\}_{n \in \mathbb{N}}$ tal que $\lim_{x_{j_n}} = x$.*

Teorema 5.2. *Si μ es una función de Whitney admisible para $\mathcal{H} = C(X)$ o 2^X , entonces, para cualquier $t_0 \in (0, \mu(X))$, $\mu^{-1}(t_0)$ es un retracto de $\mu^{-1}([t_0, \mu(X)])$.*

Demostración. Sea $t_0 \in (0, \mu(X))$. Sea $h : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ un deformación μ -admisibles. Sea $\mathcal{B} = \mu^{-1}([t_0, \mu(X)])$. Sea $A \in \mathcal{B}$. Como $h(A, 1) = A$ y $h(A, 0) \in F_1(X)$, se cumple que $\mu(h(A, 1)) \geq t_0$ y $\mu(h(A, 0)) = 0$. Note que $h(\{A\} \times [0, 1])$ es conexo, porque h es continua. Así, por la observación 4.4, existe $t_A \in [0, 1]$ tal que $\mu(h(A, t_A)) = t_0$. Además, para cualquier número real $s \in [0, 1]$ distinto de t_A , se cumple que $0 \leq s < t_A$ o $t_A < s \leq 1$ y, por ende, $\mu(h(A, s)) < \mu(h(A, t_A))$ o $\mu(h(A, t_A)) < \mu(h(A, s))$. Por tanto, t_A es el único elemento de $[0, 1]$ tal que $\mu(h(A, t_A)) = t_0$. Sea $\eta : \mathcal{B} \rightarrow [0, 1]$ la función dada, para cada $A \in \mathcal{B}$, por $\eta(A) = t_A$.

Vamos a mostrar que η es continua. Para tal efecto, sea $\{A_n\}_{n \in \mathbb{N}}$ una sucesión en \mathcal{B} tal que $\lim A_n = A$, para algún $A \in \mathcal{B}$. Sea $\{A_{k_n}\}_{n \in \mathbb{N}}$ una subsucesión de $\{A_n\}_{n \in \mathbb{N}}$. Como $[0, 1]$ es compacto, existe una subsucesión $\{A_{j_n}\}_{n \in \mathbb{N}}$ de $\{A_{k_n}\}_{n \in \mathbb{N}}$ tal que $\lim_{n \rightarrow \infty} \eta(A_{j_n}) = s_A$, para algún $s_A \in [0, 1]$. Así, $\lim(A_{j_n}, \eta(A_{j_n})) = (A, s_A)$. Como $\mu \circ h$ es continua, esto implica que $t_0 = \lim \mu(h(A_{j_n}, \eta(A_{j_n}))) = \mu(h(A, s_A))$. Por la propiedad de unicidad de t_A , tenemos que $s_A = t_A$. Así, $\lim \eta(A_{j_n}) = \eta(A)$. Aplicando el lema 5.1, obtenemos que η es continua.

Sea $r : \mathcal{B} \rightarrow \mu^{-1}(t_0)$ la función tal que, para cada $A \in \mathcal{B}$, $r(A) = h(A, \eta(A))$. Como r es una composición de funciones continuas, se cumple que r es continua. Note que, para cada $A \in \mu^{-1}(t_0)$, $\mu(h(A, 1)) = \mu(A) = t_0$, por lo cual $\eta(A) = 1$ y $r(A) = h(A, 1) = A$. De este modo, r es una retracción de \mathcal{B} sobre $\mu^{-1}(t_0)$. \square

Teorema 5.3 ([10, Proposición 7.7]). *Sean X un espacio topológico y $Y \subset X$. Si Y es un retracto de X y X es un retracto absoluto, entonces Y es un retracto absoluto.*

Teorema 5.4. *Sean X un continuo localmente conexo y μ una función de Whitney para \mathcal{H} , donde $\mathcal{H} = 2^X$ o $\mathcal{H} = C(X)$. Si μ es admisible, entonces, para cada $t_0 \in (0, \mu(X))$, $\mu^{-1}(t_0)$ es un retracto absoluto.*

Demostración. Supongamos que μ es admisible. Sea $t_0 \in (0, \mu(X))$ y sea $\mathcal{B} = \mu^{-1}([t_0, \mu(X)])$. Note que, dados $A \in \mathcal{B}$ y $C \in \mathcal{B}$ tales que $A \subset C$, se cumple que $t_0 \leq \mu(A)$ y $\mu(A) \leq \mu(C)$, lo cual implica que $C \in \mathcal{B}$. Así, \mathcal{B} es un hiperespacio de crecimiento. Aplicando el Teorema 3.8, obtenemos que \mathcal{B} es un retracto absoluto. Como $\mu^{-1}(t_0)$ es un retracto de \mathcal{B} (teorema 5.2), se tiene, por el teorema 5.3, que $\mu^{-1}(t_0)$ es un retracto absoluto. \square

Sean X un continuo, μ una función de Whitney para $C(X)$, $t \in [0, \mu(X)]$, $K \in C(X)$. En lo que sigue, utilizaremos la siguiente notación:

$$\begin{aligned} C(X, K) &= \{A \in C(X) : K \cap A \neq \emptyset\}, \\ C^K(X) &= \{A \in C(X) : K \subset A\}, \\ C^K(X, t) &= \mu^{-1}(t) \cap C^K(X), \end{aligned}$$

Si $K = \{p\}$, para algún $p \in X$, entonces se obviarán los paréntesis en la notación anterior, reemplazando K por p .

Teorema 5.5 (de Lynch, [12, Teorema 66.4]). *Sean X un continuo, μ una función de Whitney para $C(X)$ y $t_0 \in [0, \mu(X)]$. Si $E \in C(X)$ es tal que $\mu(E) \leq t_0$, entonces $C^E(X, t_0)$ es un retracto absoluto.*

Sean X un espacio métrico compacto y A un subconjunto cerrado de X . Decimos que A es un **Z-conjunto** en X si Id_X es el límite uniforme de funciones continuas cuyas imágenes no intersectan a A . Decimos que una función continua f entre espacios métricos compactos X_1 y X_2 es una **Z-función** si $f(X_1)$ es un Z-conjunto en X_2 .

Teorema 5.6 (de Toruńczyk, [12, Teorema 9.3]). *Sea X un retracto absoluto. Si la función identidad sobre X es un límite uniforme de Z-funciones, entonces X es un cubo de Hilbert.*

Dado un continuo X , un **arco ordenado** es un subcontinuo \mathcal{A} de 2^X tal que cualesquiera $A, B \in \mathcal{A}$ satisfacen que $A \subset B$ o $B \subset A$.

Lema 5.7 ([8, Proposición 3.1]). *Sea μ una función de Whitney para $C(X)$. Sea $\mathcal{S} \subset C(X)$ y sea $\sigma : \mathcal{S} \rightarrow C(C(X))$ una función continua tal que, para cada $B \in \mathcal{S}$, $\sigma(B)$ es un arco ordenado en $C(X)$. Sea $t_0 \in [0, \mu(X)]$ y supóngase que, para cada $B \in \mathcal{S}$, $\sigma(B) \cap \mu^{-1}(t_0) \neq \emptyset$. Entonces, para cada $B \in \mathcal{S}$, $\sigma(B) \cap \mu^{-1}(t_0)$ consiste exactamente de un punto, denotado por $\sigma_{t_0}(B)$, y la función $\sigma_{t_0} : \mathcal{S} \rightarrow \mu^{-1}(t_0)$ es continua.*

Un **árbol** es una gráfica finita que no contiene curvas cerradas simples.

Lema 5.8 ([3, Lema 1.42]). *Sea X un espacio topológico arco conexo. Si F es un subconjunto finito de X , entonces existe un árbol T contenido en X , tal que $F \subset T$.*

Teorema 5.9 ([17, Teorema 8.26]). *Todo abierto conexo de un continuo localmente conexo es arco conexo.*

Teorema 5.10 ([17, Teorema 8.9]). *Si X es un continuo localmente conexo, entonces, para cualquier $\varepsilon > 0$, existe una colección finita X_1, X_2, \dots, X_n de continuos localmente conexos tal que $X = \bigcup_{i=1}^n X_i$ y, para cada $i \in \{1, \dots, n\}$, $\text{diám}(X_i) < \varepsilon$.*

Teorema 5.11 ([15, §49, II, Teorema 1]). *Sea X un espacio métrico. Si X_1, X_2, \dots, X_n son cerrados de X localmente conexos, entonces $\bigcup_{i=1}^n X_i$ es localmente conexo.*

Lema 5.12. *Sean X un continuo localmente conexo y K un cerrado de X con interior no vacío. Si K no contiene arcos libres de X , entonces existe una función continua $g : C(X) \rightarrow C(X) - C^K(X)$ tal que, si $A, B \in C(X)$ y $A \subset B$, entonces $H_d(A, g(A)) < \varepsilon$ y $g(A) \subset g(B)$.*

Demostración. Fijemos $p \in \text{int}_X(A)$ y sea $\varepsilon > 0$, tal que $B_d(p, \varepsilon) \subsetneq \text{int}_X(A)$. Por el Teorema 5.10, existen subcontinuos localmente conexos X_1, X_2, \dots, X_m de X , tales que $X = \bigcup_{i=1}^m X_i$ y $\text{diám}(X_i) < \frac{\varepsilon}{4}$, para cada $i \in \{1, 2, \dots, m\}$. Sean $\Lambda_0 = \{i \in \{1, \dots, m\} : p \in X_i\}$ y $S = \bigcup_{i \in \Lambda_0} X_i$. Como $p \in \bigcap_{i \in \Lambda_0} X_i$, se tiene que S es un continuo y, para cada $x \in S$, $d(p, x) < \frac{\varepsilon}{4}$; es decir, $S \subset B_d(p, \frac{\varepsilon}{4}) \subsetneq X$.

Sean $\Lambda_1 = \{i \in \{1, \dots, m\} : p \notin X_i \text{ y } X_i \cap S \neq \emptyset\}$ y $\Lambda_2 = \{i \in \{1, \dots, m\} : X_i \cap S = \emptyset\}$. Obsérvese que S y $\bigcup_{i \in \Lambda_2} X_i$ son subconjuntos cerrados y ajenos de X y $\Lambda_0 \cup \Lambda_1 \cup \Lambda_2 = \{1, \dots, m\}$.

Note que $X - S \subset \bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i$ y, como $\bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i$ es cerrado en X , $\text{cl}_X(X - S) \subset \bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i$. Pero, como $S \cap \bigcup_{i \in \Lambda_2} X_i = \emptyset$, se cumple que $\text{cl}_X(X - S) \subset \bigcup_{i \in \Lambda_1} X_i$. Así, $\text{Fr}_X(S) \subset \bigcup_{i \in \Lambda_1} X_i$. En particular, como $\text{Fr}_X(S) \neq \emptyset$ (porque $S \neq \emptyset$, $S \subsetneq X$ y X es conexo), se tiene $\Lambda_1 \neq \emptyset$.

Para cada $i \in \Lambda_1$, fijemos $p_i \in X_i \cap S$. Note que, por el Teorema 5.11, se tiene que S es localmente conexo y así, por el Teorema 5.9, S es arco conexo. Luego, por el Lema 5.8, existe un árbol T contenido en S , tal que $\{p\} \cup \{p_i : i \in F\} \subset T$.

Sea $Y = T \cup (\bigcup_{i \in F} X_i)$. Como cada $i \in F$ cumple que $p_i \in T \cap X_i$, se tiene que Y es un continuo. Además, por el Teorema 5.11, Y es localmente conexo. Obsérvese que, para cada $i \in F$ y $x \in X_i$ se cumple $d(x, p) \leq d(x, p_i) + d(p, p_i) < \frac{\varepsilon}{2}$; es decir, $\bigcup_{i \in F} X_i \subset B_d(p, \frac{\varepsilon}{2})$. Luego, $Y \subset B_d(p, \frac{\varepsilon}{2}) \subset \text{int}_X(K)$.

Consideremos la función $\phi : Y \rightarrow C(Y)$, dada por $\phi(y) = \{y\}$. Por el Teorema 3.1, ϕ es continua. Como Y es un subconjunto cerrado de $S \cup Y$ y $C(Y)$ es un retracto absoluto (véase la demostración del Corolario 3.9), por [12, Teorema 9.1] podemos extender ϕ a una función continua $r : S \cup Y \rightarrow C(Y)$.

Como $\text{Fr}_X(S) \subset Y$, $r(y) = \{y\}$, para cada $y \in \text{Fr}_X(S)$. Luego, la función $f : X \rightarrow C(X)$ dada por

$$f(x) = \begin{cases} r(x), & x \in S \cup Y, \\ \{x\}, & x \in X - (S \cup Y), \end{cases}$$

está bien definida y es continua.

Note que $p \in X - \bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i = S - \bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i$. Así, S es una vecindad de p en X . Si $S \subset Y$, entonces $S - \bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i \subset Y - \bigcup_{i \in \Lambda_1 \cup \Lambda_2} X_i \subset T$, por lo cual T es una vecindad de p en X ; como $T \subset Y \subset A$, esto contradice la hipótesis de que A no contiene arcos libres de X . Por lo tanto, $S \not\subset Y$.

Sea $q \in S - Y$. Luego, si $x \in S \cup Y$, $f(x) = r(x) \subset Y$ y así $q \notin f(x)$; si $x \in X - (S \cup Y)$, entonces $x \notin S$ y $q \notin f(x) = \{x\}$. De este modo, $q \notin f(X)$.

Sea $g : 2^X \rightarrow 2^X$ la función

$$g(B) = \bigcup \{f(b) : b \in B\}$$

Note que $g = \bigcup \circ f^*$, por lo cual, g está bien definida y es continua. Sea $B \in 2^X$. Como $g(B) \subset f(X)$, tenemos que $q \notin g(B)$. Así, $S \not\subset g(B)$. Además, si $B \in C(X)$, entonces $f^*(B)$ es conexo y, como $f(x) \in C(X)$, tenemos que $f^*(B) \cap C(X) \neq \emptyset$. Así, por el Teorema 3.5 $g(B) \in C(X)$. Por tanto, $g(C(X)) \subset C(X)$. Además, .

Tomemos $A \in C(X)$. Supongamos que $x \in A$. Si $x \in S \cup Y$, entonces $x \in f(x) = r(x) \subset Y$ y, como $S \cup Y \subset B_d(p, \frac{1}{2}\varepsilon)$, si y es cualquier punto de $r(x)$, entonces $d(x, y) < d(x, p) + d(p, y) < \varepsilon$ y $x \in N_d(f(x), \varepsilon) \subset N_d(g(A), \varepsilon)$. Por el contrario, si $x \in X - (S \cup Y)$, entonces $x \in f(x) = \{x\}$ y así $x \in N_d(f(x), \varepsilon) \subset N_d(g(A), \varepsilon)$. Por tanto, $A \subset N_d(g(A), \varepsilon)$. Recíprocamente, supongamos que $x \in g(A)$. Luego, $x \in f(a)$, para alguna $a \in A$. Si $a \in S \cup Y$, entonces $x \in f(a) \subset Y$. De nuevo porque $S \cup Y \subset B_d(p, \frac{1}{2}\varepsilon)$, se tiene que $d(x, a) \leq d(x, p) + d(p, a) < \varepsilon$. Si $a \in S \cup Y$, entonces $f(a) = \{a\}$ y así $x \in g(A)$. Por tanto, $g(A) \subset N_d(A, \varepsilon)$. Considerando las contenciones mostradas, se concluye que $H_d(A, g(A)) < \varepsilon$. □

Lema 5.13 (véase [3, Lemas 1.121 (3) y 1.122]). *Dado un continuo con métrica convexa d , la función $\Phi : 2^X \times [0, \infty) \rightarrow 2^X$ dada por*

$$\Phi(A, t) = C_d(A, t)$$

es continua y $\Phi(C(X) \times [0, \infty)) \subset C(X)$.

Lema 5.14. *Sean X un continuo localmente conexo y K un cerrado de X con interior no vacío. Si μ es una función de Whitney admisible para $C(X)$ y K no contiene arcos libres de X , entonces, para cualquier $t_0 \in (0, \mu(X))$, $C^K(X, t_0)$ es un Z -conjunto de $\mu^{-1}(t_0)$.*

Demostración. Supongamos que μ es una función de Whitney admisible para $C(X)$ y que K no contiene arcos libres de X . Sea $t_0 \in (0, \mu(X))$. Sea $\varepsilon > 0$. Para demostrar el lema, basta exhibir una función continua $\sigma_{t_0} : \mu^{-1}(t_0) \rightarrow \mu^{-1}(t_0) - C^K(X, t_0)$ tal que, para cada $A \in \mu^{-1}(t_0)$, $H_d(\sigma_{t_0}(A), A) < 3\varepsilon$. Para tal efecto, sea $h : C(X) \times [0, 1] \rightarrow C(X)$ una deformación μ -admisible (la cual existe por el Teorema 4.5).

Afirmamos que existe $s_0 \in (0, 1)$ tal que $\lambda > t_0$, donde $\lambda = \inf\{\mu(C_d(h(B, s), \varepsilon)) : B \in \mu^{-1}(t_0), s \in [s_0, 1]\}$. Supongamos que esto no es así. Sea $\zeta = \inf\{\mu(C_d(B, \varepsilon)) : B \in \mu^{-1}(t_0)\}$. Note que, como $\mu^{-1}(t_0)$ es compacto y, para cada $B \in \mu^{-1}(t_0)$, $B \subsetneq C_d(B, \varepsilon)$, se tiene $\zeta > t_0$. Sea $n \in \mathbb{N}$ tal que $1/n < \zeta - t_0$. Como μ es uniformemente continua, existe $\rho > 0$ tal que si $H_d(A, B) < \rho$ entonces $|\mu(A) - \mu(B)| < 1/(2n)$. Sea $\Phi_h : C(X) \times [0, 1] \rightarrow C(X)$ la función dada por $\Phi_h(B, s) = C_d(h(B, s), \varepsilon)$. Como Φ_h es la composición de funciones continuas (h y la función Φ del Lema 5.13) y su dominio es compacto, Φ_h es uniformemente continua. Así, existe $\eta > 0$ tal que si $1 - \eta < s' < 1$, entonces $H_d(\Phi_h(B, s'), C_d(B, \varepsilon)) < \rho$ y así $|\mu(\Phi_h(B, s')) - \mu(C_d(B, \varepsilon))| < 1/(2n)$. Por hipótesis, existen $1 - 1/n < s_n < 1$ y $B_n \in \mu^{-1}(t_0)$ tales que $\mu(\Phi_h(B_n, s_n)) < t_0 + 1/(2n)$. Luego, $|\mu(\Phi_h(B_n, s_n)) - \mu(C_d(B_n, \varepsilon))| < 1/(2n)$. Así, $\mu(C_d(B_n, \varepsilon)) < 1/(2n) + \mu(\Phi_h(B_n, s_n)) < 1/(2n) + t_0$, si $\mu(\Phi_h(B_n, s_n)) < t_0$, o bien $|\mu(C_d(B_n, \varepsilon)) - t_0| < 1/n$, si $\mu(\Phi_h(B_n, s_n)) \geq t_0$. En ambos casos $\zeta \leq \mu(C_d(B_n, \varepsilon)) < 1/n + t_0$, lo cual no es posible, por la elección de n . Esto prueba nuestra afirmación.

Como h es uniformemente continua, existe $\nu > 0$ tal que si $B_1, B_2 \in C(X)$ y $t_1, t_2 \in [0, 1]$ son tales que $d(B_1, B_2) + |t_1 - t_2| < \nu$, entonces $H_d(h(B_1, t_1), h(B_2, t_2)) < \varepsilon$. Fijemos $s \in (\max\{1 - \nu, s_0\}, 1)$. Como $\mu^{-1}(t_0)$ es compacto, tenemos que $\xi = \sup\{\mu(h(B, s)) : B \in \mu^{-1}(t_0)\} < t_0$. Como μ es uniformemente continua, existe $\delta \in (0, \varepsilon)$ tal que si $A_1, A_2 \in C(X)$ y $H_d(A_1, A_2) < \delta$, entonces $|\mu(A_1) - \mu(A_2)| < \min\{\frac{1}{2}(\lambda - t_0), \frac{1}{2}(t_0 - \xi)\}$. Por el Lema 5.12, existe una función continua $g : C(X) \rightarrow C(X) - C^K(X)$ tal que, si $A, B \in C(X)$ y $A \subset B$, entonces $H_d(A, g(A)) < \delta$ y $g(A) \subset g(B)$. Así, $-\frac{1}{2}(\lambda - t_0) < \mu(g(A)) - \mu(A)$, para cada $A \in C(X)$. Por tanto, para cada $B \in \mu^{-1}(t_0)$ y $s \in [s_0, 1)$, se cumple que

$t_0 < \frac{1}{2}\lambda + \frac{1}{2}t_0 = \lambda - \frac{1}{2}(\lambda - t_0) < \mu(C_d(h(B, s), \varepsilon)) - \frac{1}{2}(\lambda - t_0) < \mu(g(C_d(h(B, s), \varepsilon)))$. Análogamente, $\mu(g(A)) - \mu(A) < \frac{1}{2}(t_0 - \xi)$, para cada $A \in C(X)$. Luego, para cada $B \in \mu^{-1}(t_0)$, $t_0 > \frac{1}{2}t_0 + \frac{1}{2}\xi = \frac{1}{2}(t_0 - \xi) + \xi > \mu(g(h(B, s))) - \mu(h(B, s)) + \xi > \mu(h(g(B, s)))$

Así, para cada $B \in \mu^{-1}(t_0)$,

$$\mu(h(g(B, s))) < t_0 < \mu(g(C_d(h(B, s), \varepsilon))) \tag{1}$$

$$H_d(h(B, s), B) < \varepsilon \tag{2}$$

Consideremos la función $\sigma : \mu^{-1}(t_0) \rightarrow C(C(X))$ que a cada $B \in \mu^{-1}(t_0)$ le asigna el arco ordenado $\sigma(B) = \{g(C_d(h(B, s), t)) : t \geq 0\}$. Note que

$$\sigma(B) = (g \circ \Phi)^* \circ \theta(\phi(h(B, s)), [0, \infty)),$$

donde $\theta : C(X) \times C(X) \rightarrow C(X \times X)$ y $\phi : C(X) \rightarrow C(C(X))$ son las funciones continuas dadas por

$$\theta(A_1, A_2) = A_1 \times A_2,$$

$$\phi(A_1) = \{A_1\}$$

(véase los lemas 4.3 y 3.1). De este modo, σ es una función continua. Luego, por (1), σ cumple las hipótesis del Teorema 5.7. Así, la función $\sigma_{t_0} : \mu^{-1}(t_0) \rightarrow \mu^{-1}(t_0)$ que a cada $B \in \mu^{-1}(t_0)$ asigna $\sigma_{t_0}(B)$, el único punto de $\sigma(B) \cap \mu^{-1}(t_0)$, es una función continua. Sea $t_B \geq 0$ tal que $\sigma_{t_0}(B) = g(C_d(h(B, s), t_B))$. Como $g(C(X)) \cap C^K(X) = \emptyset$, se satisface que $\sigma_{t_0}(B) \notin C^K(X)$ y, por (1), se tiene que $0 < t_B < \varepsilon$. Por lo tanto,

$$\begin{aligned} H_d(\sigma_{t_0}(B), B) &\leq H_d(\sigma_{t_0}(B), C_d(h(B, s), t_B)) + H_d(C_d(h(B, s), t_B), h(B, s)) \\ &\quad + H_d(h(B, s), B) \\ &< \delta + t_B + \varepsilon < 3\varepsilon, \end{aligned}$$

y $\sigma_{t_0}(\mu^{-1}(t_0)) \subset \mu^{-1}(t_0) - C^K(X, t_0)$. Esto prueba que $C^K(X, t_0)$ es un Z-conjunto de $\mu^{-1}(t_0)$. □

Teorema 5.15. *Sean X un continuo localmente conexo y μ una función de Whitney para $C(X)$. Si μ es admisible y X no contiene arcos libres, entonces, para cualquier $t_0 \in (0, \mu(X))$, $\mu^{-1}(t_0)$ es un cubo de Hilbert.*

Demostración. Supongamos que μ es admisible y que X no contiene arcos libres. Sea $t_0 \in (0, \mu(X))$. Sea $\varepsilon > 0$. Por el Teorema 5.4, $\mu^{-1}(t_0)$ es un retracto absoluto. Luego, por el Teorema de Toruńczyk, para demostrar el teorema basta exhibir una

Z-función $\sigma_{t_0} : \mu^{-1}(t_0) \rightarrow \mu^{-1}(t_0)$ tal que, para cualquier $A \in \mu^{-1}(t_0)$, se cumple que $H_d(A, \sigma_{t_0}(A)) < 2\varepsilon$.

De acuerdo con la demostración del Lema 5.14, existen una deformación μ -admisibles $h : C(X) \times [0, 1] \rightarrow C(X)$, $s \in (0, 1)$ y $\lambda > t_0$ tales que, para cada $B \in \mu^{-1}(t_0)$,

$$\lambda \leq \mu(C_d(h(B, s), \varepsilon)), \quad (3)$$

$$H_d(B, h(B, s)) < \varepsilon. \quad (4)$$

Sea $\xi = \sup\{\mu(h(B, s)) : B \in \mu^{-1}(t_0)\}$. Como $s < 1$ y h es admisible, para cada $B \in \mu^{-1}(t_0)$ se tiene

$$\mu(h(B, s)) < t_0. \quad (5)$$

Por la compacidad de $\mu^{-1}(t_0)$, $\xi < t_0$. Además, de (5) y (3) obtenemos que, para cada $B \in \mu^{-1}(t_0)$,

$$\mu(h(B, s)) < t_0 < \mu(C_d(h(B, s), \varepsilon)). \quad (6)$$

Consideremos la función $\sigma : \mu^{-1}(t_0) \rightarrow C(C(X))$ que a cada $B \in \mu^{-1}(t_0)$ le asigna el arco ordenado $\sigma(B) = \{C_d(h(B, s), t) : t \geq 0\}$. De manera análoga a la función σ construida en la demostración del Lema 5.14, esta función es continua. De este modo, y por (6), σ cumple las hipótesis del Lema 5.7, y así la función $\sigma_{t_0} : \mu^{-1}(t_0) \rightarrow \mu^{-1}(t_0)$ que a cada $B \in \mu^{-1}(t_0)$ le asigna $\sigma_{t_0}(B)$, el único punto de $\sigma(B) \cap \mu^{-1}(t_0)$, es continua.

Dado $B \in \mu^{-1}(t_0)$, sea $t_B \geq 0$ tal que $\sigma_{t_0}(B) = C_d(h(B, s), t_B)$. Como $\mu^{-1}(t_0)$ y $\mu^{-1}([0, \xi])$ son subconjuntos compactos y ajenos de $C(X)$, se tiene que $\gamma > 0$, donde

$$\gamma = \inf \{H_d(B_1, B_2) : B_1 \in \mu^{-1}(t_0), B_2 \in \mu^{-1}([0, \xi])\}.$$

Así, y como $h(B, s) \in \mu^{-1}([0, \xi])$ y $\sigma_{t_0}(B) \in \mu^{-1}(t_0)$,

$$H_d(h(B, s), C_d(h(B, s), t_B)) \geq \gamma. \quad (7)$$

Note que, si $t_B < \gamma$, entonces $H_d(h(B, s), C_d(h(B, s), t_B)) < \gamma$, lo cual contradice (7). Así, cada $B \in \mu^{-1}(t_0)$ satisface $t_B \geq \gamma$.

Vamos a mostrar que σ_{t_0} es una Z-función. Como X es compacto, existen $p_1, p_2, \dots, p_k \in X$ tales que $X \subset \bigcup_{i=1}^k B_d(p_i, \gamma/2)$. Para cada $i \in \{1, 2, \dots, k\}$, sea $K_i = C_d(\{p_i\}, \gamma/2)$. Sea $B_0 \in \omega^{-1}(t_0)$. Sea $x_0 \in h(B_0, s)$. Luego, existe $j \in \{1, 2, \dots, k\}$ tal que $x_0 \in B_d(p_j, \gamma/2)$, es decir, tal que $d(x_0, p_j) < \gamma/2$. Así, para cualquier $y \in K_j$,

$$d(y, x_0) \leq d(y, p_j) + d(x_0, p_j) \leq \gamma/2 + d(x_0, p_j) < \gamma/2 + \gamma/2 = \gamma \leq t_{B_0},$$

En consecuencia, $K_j \subset B_d(x_0, t_{B_0}) \subset C_d(h(B_0, s), t_{B_0}) = \sigma_{t_0}(B_0)$. Esto muestra que $\sigma_{t_0}(B_0) \in C^{K_j}(X, t_0)$. Por tanto,

$$\sigma_{t_0}(\mu^{-1}(t_0)) \subset \bigcup_{i=1}^k C^{K_i}(X, t_0).$$

Por otro lado, como X no contiene arcos libres, cada K_i es un cerrado de X con interior no vacío que no contiene arcos libres de X . Así, por el Lema 5.14, $C^{K_i}(X, t_0)$ es un Z -conjunto de $\mu^{-1}(t_0)$. Como la unión de Z -conjuntos es a su vez un Z -conjunto, la unión $\bigcup_{i=1}^k C^{K_i}(X, t_0)$ es un Z -conjunto de $\mu^{-1}(t_0)$. Asimismo, como los subconjuntos cerrados de Z -conjuntos son Z -conjuntos, $\sigma_{t_0}(\mu^{-1}(t_0))$ es un Z -conjunto de $\mu^{-1}(t_0)$. Así, σ_{t_0} es una Z -función. Además, cada $B \in \mu^{-1}(t_0)$ satisface $H_d(\sigma_{t_0}(B), h(B, s)) \leq t_B$ y, por (6), $t_B < \varepsilon$; así, y aplicando (4), tenemos que

$$H_d(\sigma_{X, t_0}(B), B) \leq H_d(\sigma_{t_0}(B), h(B, s)) + H_d(h(B, s), B) < t_B + \varepsilon < 2\varepsilon.$$

Esto concluye la demostración de que $\mu^{-1}(t_0)$ es un cubo de Hilbert. □

6 Funciones de Whitney y fibraciones

Para concluir este capítulo, vamos a mostrar que cualquier función de Whitney para el hiperespacio de subcontinuos de un continuo localmente conexo que no posee arcos libres presenta un comportamiento muy similar a la primera proyección del producto topológico de un intervalo abierto con un cubo de Hilbert, en un sentido que precisaremos más adelante (previo al Teorema 6.3).

Sean Y un espacio métrico y $\varepsilon > 0$. Una función continua $f : Y \rightarrow Y$ **mueve puntos no más de ε** si, para cada $y \in Y$, $d(y, f(y)) < \varepsilon$. Una homotopía $h : Y \times [0, 1] \rightarrow Y$ mueve puntos no más de ε si, para cualesquiera $y \in Y$ y $t \in [0, 1]$, $d(y, h(y, t)) < \varepsilon$.

Dados dos espacios métricos E y B y una función continua $p : E \rightarrow B$, decimos que p es **fuertemente regular** si la imagen inversa de cada compacto es compacta y para cualesquiera $b_0 \in B$ y $\varepsilon > 0$ existe una vecindad $U(b_0)$ en B tal que, para cada $b \in U(b_0)$, existen funciones continuas $g_{b, b_0} : p^{-1}(b) \rightarrow p^{-1}(b_0)$ y $g_{b_0, b} : p^{-1}(b_0) \rightarrow p^{-1}(b)$ tales que $g_{b_0, b}$ y g_{b, b_0} mueven puntos no más que ε y $g_{b, b_0} \circ g_{b_0, b}$ y $g_{b_0, b} \circ g_{b, b_0}$ son homotópicas a las funciones identidad en $p^{-1}(b_0)$ y $p^{-1}(b)$, respectivamente, por medio de homotopías que mueven puntos no más que ε .

Teorema 6.1 ([18, 14,44]). *Si X es cualquier continuo y μ es cualquier función de Whitney para $C(X)$, entonces la función $\mu^{-1} : [0, \mu(X)] \rightarrow C(C(X))$ dada por*

$$\mu^{-1}(t) = \{A \in C(X) : \mu(A) = t\}$$

es continua.

Teorema 6.2. *Sea X un continuo localmente conexo y sea $\mathcal{H} = 2^X$ o $\mathcal{H} = C(X)$. Si μ es una función de Whitney admisible para \mathcal{H} , entonces $\mu|_{\mu^{-1}((0, \mu(X)))} : \mu^{-1}((0, \mu(X))) \rightarrow (0, \mu(X))$ es una función fuertemente regular cuyas fibras son retracts absolutos.*

Demostración. Supongamos que μ es una función de Whitney admisible para \mathcal{H} . Sea $h : \mathcal{H} \times [0, 1] \rightarrow \mathcal{H}$ una deformación μ -admisible. Note que, para cada $t \in (0, \mu(X))$, se cumple que $\mu^{-1}(t)$ es compacto y que, por el Teorema 5.4, $\mu^{-1}(t)$ es un retracto absoluto. Para probar la parte restante de la condición de regularidad fuerte vamos a utilizar algunas funciones auxiliares y demostrar algunas propiedades relacionadas con estas.

Sean $t \in (0, \mu(X))$ y $\mathcal{B} = \mu^{-1}([t, \mu(x)])$. Si $t < \mu(X)$, sea $r_t : \mathcal{B} \rightarrow \mu^{-1}(t)$ la retracción construida en la demostración del Teorema 5.2; es decir, para cada $A \in \mathcal{B}$,

$$r_t(A) = h(A, \theta_t(A)),$$

donde $\theta_t(A)$ es el único elemento de $[0, 1]$ tal que $\omega(h(A, \theta_t(A))) = t$. Si $t = \mu(X)$, entonces $\mu^{-1}(t) = \mathcal{B} = \{X\}$ y, en tal caso, podemos definir $r_t : \mathcal{B} \rightarrow \mu^{-1}(t)$ haciendo $\theta_t(X) = 1$ y $r_t(X) = X = h(X, \theta_t(X))$.

Como X es localmente conexo, podemos considerar a d como una métrica conexa. Así, C_d es continua. Como X es compacto, existe una cota para la métrica d , digamos M . Luego, para cada $A \in \mu^{-1}([0, t])$, $C_d(A, 2M) = X$ y, en consecuencia, $\mu(C_d(A, 0)) \leq t \leq \mu(C_d(A, 2M))$. Como C_d es continua, el conjunto

$$\{C_d(A, z) : z \in [0, 2M]\}$$

es conexo. Así, por la observación 4.4, existe $z \in [0, 2M]$ tal que $\mu(C_d(A, z)) = t$. De este modo, podemos considerar la función $\Psi_t : \mu^{-1}([0, t]) \rightarrow [0, \infty)$ dada, para cada $A \in \mu^{-1}([0, t])$, por

$$\Psi_t(A) = \inf\{z \in [0, 2M] : \mu(C_d(A, z)) = t\}.$$

Sea $R_t : \mathcal{B} \rightarrow \omega^{-1}(t)$ la función dada, para cada $A \in \mathcal{B}$, por

$$R_t(A) = C_d(A, \Psi_t(A)).$$

Afirmación 1. Ψ_t y R_t son continuas.

Note que, si $C_d(A, \Psi_t(A)) \neq X$, entonces, para cada $z > \Psi_t(A, z)$, $C_d(A, \Psi_t(A)) \subsetneq$

$C_d(A, z)$ y, en consecuencia, $\mu(C_d(A, \Psi_t(A))) < \mu(C_d(A, z))$; así, en tal caso, se tiene que $\mu(C_d(A, \Psi_t(A))) = t$. Por otro lado, si $C_d(A, \Psi_t(A)) = X$, entonces, para cada $z > \Psi_t(A)$, $C_d(A, z) = X$, por lo cual $\mu(C_d(A, \Psi_t(A))) = \mu(X) = t$.

Veamos que Ψ_t es continua. Para tal efecto, sea $\{A_n\}$ una sucesión en \mathcal{B} tal que $\lim A_n = A$, para alguna $A \in \mathcal{B}$. Supongamos que $\lim \Psi_t(A_n)$ no existe o es diferente de $\Psi_t(A)$. Luego, por el Lema 5.1, existe una subsucesión $\{A_{k_n}\}_{n \in \mathbb{N}}$ tal que cada una de sus subsucesiones no tiene límite o su límite es distinto de $\Psi_t(A)$. Como $[0, 2M]$ es compacto, existe una subsucesión $\{\Psi_t(A_{l_n})\}_{n \in \mathbb{N}}$ de $\{\Psi_t(A_{k_n})\}_{n \in \mathbb{N}}$ tal que $\lim \Psi_t(A_{l_n}) = z_0$, para algún $z_0 \in [0, 2M]$. Luego, por la continuidad de μ y de C_d (Lema 5.13),

$$t = \lim \mu(C_d(A_{l_n}, \Psi_t(A_{l_n}))) = \mu(C_d(A, z_0))$$

y, por ende $\Psi_t(A) \leq z_0$. Se sigue de lo anterior y de la elección de $\{A_{k_n}\}_{n \in \mathbb{N}}$ que $\Psi_t(A) < z_0$. Note que, si $C_d(A, \Psi_t(A)) \neq X$, entonces $C_d(A, \Psi_t(A)) \subsetneq C_d(A, z_0)$ y, por ende,

$$t = \mu(C_d(A, \Psi_t(A))) < \mu(C_d(A, z_0)) = t,$$

lo cual es una contradicción. Así $C_d(A, \Psi_t(A)) = X$ y, en consecuencia, $t = \mu(X)$. Como $\lim \mu(C_d(A_{l_n}, \Psi_t(A))) = \mu(X)$, tenemos que $\lim C_d(A_{l_n}, \Psi_t(A)) = X$ (el lector puede verificar que, en general, si $B \in \mathcal{H}$ y $\{B_n\}_{n \in \mathbb{N}}$ es una sucesión en \mathcal{H} tal que todos sus elementos son subconjuntos de B y $\lim \mu(B_n) = \mu(B)$, entonces $\lim B_n = B$). Sea $\varepsilon > 0$. Luego, existe $N \in \mathbb{N}$ tal que, si $n \geq N$, entonces $H_d(C_d(A_{l_n}, \Psi_t(A)), X) < \varepsilon$ y, por consiguiente,

$$X \subset C_d(C_d(A_{l_n}, \Psi_t(A)), \varepsilon) \subset C_d(A_{l_n}, \Psi_t(A) + \varepsilon);$$

de esto último se sigue que $C_d(A_{l_n}, \Psi_t(A) + \varepsilon) = X$ y $\mu(C_d(A_{l_n}, \Psi_t(A) + \varepsilon)) = t$. Esto implica que, para cada $n \geq N$, $\Psi_t(A_{l_n}) \leq \Psi_t(A) + \varepsilon$. Así, $z_0 \leq \Psi_t(A) + \varepsilon$. Como ε fue elegido arbitrariamente, se tiene que $z_0 \leq \Psi_t(A)$, lo cual contradice la desigualdad $\Psi_t(A) < z_0$, obtenida con anterioridad. De este modo, $\lim \Psi_t(A_n) = \Psi_t(A)$. Por lo tanto, Ψ_t es continua.

Como R_t es una composición de funciones continuas (entre ellas, C_d y Ψ_t), se tiene claramente que R_t es continua.

Ahora vamos a construir, con ayuda de las funciones r_t y R_t , las funciones que servirán para mostrar la parte restante de las condiciones de la regularidad fuerte. Para cada $t_1, t_2 \in (0, \mu(X))$, con $t_1 < t_2$, sea $g_{t_1, t_2} : \mu^{-1}(t_1) \rightarrow \mu^{-1}(t_2)$ la función dada, para cada $A \in \mu^{-1}(t_1)$, por

$$g_{t_1, t_2}(A) = R_{t_2}(A);$$

asimismo, sea $g_{t_2, t_1} : \mu^{-1}(t_2) \rightarrow \mu^{-1}(t_1)$ la función dada, para cada $E \in \mu^{-1}(t_2)$, por

$$g_{t_2, t_1}(E) = r_{t_1}(E).$$

Lo que sigue es hallar abiertos adecuados, alrededor de cada punto dado de $\mu((0, \mu(X)))$, en el cual las familias de funciones g_{t_1, t_2} y g_{t_2, t_1} satisfagan la condición de regularidad fuerte que estamos buscando. Para lograr esto, sean $t_0 \in (0, \omega(X))$ y $\varepsilon > 0$. Sean $\tau_1, \tau_2 > 0$ tales que $\tau_1 < t_0 < \tau_2$. Como h es uniformemente continua, existe $\delta \in (0, 1)$ tal que si $H_d(A_1, A_2) + |s_1, s_2| < \delta$, entonces $H_d(h(A_2, s_1), h(A_2, s_2)) < \frac{\varepsilon}{2}$.

Afirmación 2. Existe ρ tal que, para cualesquiera $t_1, t_2 \in (t_0 - \rho, t_0 + \rho) \cap (0, \mu(X))$ con $t_1 < t_2$, g_{t_1, t_2} mueve puntos no más de $\frac{\varepsilon}{2}$ y, para cada $A \in \mu^{-1}(t_1)$, $\Psi_{t_2}(A) < \frac{\varepsilon}{2}$.

Note que, para cada $A \in \mu^{-1}((0, \mu(X)))$, $A \subsetneq C_d(A, \frac{\varepsilon}{2})$, así que $\mu(A) < \mu(C_d(A, \frac{\varepsilon}{2}))$. Dado $t \in (0, \mu(X))$, sea

$$\eta(t) = \inf\{\mu(C_d(A, \frac{\varepsilon}{2})) : A \in \mu^{-1}(t)\}.$$

Como $\mu^{-1}(t)$ es compacto y $\mu \circ C_d$ es continua, tenemos que $\eta(t) > t$, para cada $t \in (0, \mu(X))$, para cada $t \in (0, \mu(X))$. Consideremos la función $\eta : [\tau_1, \tau_2] \rightarrow [0, \mu(X)]$ dada por la asociación anterior. Como η es una composición de funciones continuas, η es continua (de hecho, $\eta = \inf \circ (\mu \circ C_d)^* \circ \theta(\mu^{-1}, \{\frac{\varepsilon}{2}\})$, donde $\theta : C(C(X)) \times C([0, \frac{\varepsilon}{2}]) \rightarrow C(C(X) \times [0, \frac{\varepsilon}{2}])$ es la función dada por $\theta(A, B) = A \times B$, la cual es continua por el Lema 4.3; el amable lector puede comprobar que la función que a cada $A \in 2^{\mathbb{R}}$ le asigna $\inf(A)$ es continua, además de que, según el Teorema 6.1, la función que a cada $t \in [0, \mu(X)]$ le asocia $\mu^{-1}(t)$ también es continua). Así, y como $[\tau_1, \tau_2]$ es compacto, existe $\rho_1 > 0$ tal que

$$\rho_1 < \inf(\{\eta(t) - t : t \in [\tau_1, \tau_2]\} \cup \{t_0 - \tau_1, \tau_2 - t_0\}).$$

De este modo, para cualesquiera $t_1, t_2 \in (t_0 - \frac{\rho_1}{2}, t_0 + \frac{\rho_1}{2})$ con $t_1 < t_2$, se tiene que $t_2 - t_1 < \rho_1 < \eta(t_1) - t_1$ y, en consecuencia, $t_2 < \eta(t_1)$. Luego, para cada $A \in \omega^{-1}(t_1)$, $t_2 < \mu(C_d(A, \frac{\varepsilon}{2}))$, lo cual implica que $\Psi_{t_2}(A) < \frac{\varepsilon}{2}$. Así, para cada $A \in \omega^{-1}(t_1)$, $C_d(A, \Psi_{t_2}(A)) \subsetneq N_d(A, \frac{\varepsilon}{2})$ y, por consiguiente, $H_d(A, R_{t_2}(A)) < \frac{\varepsilon}{2}$. Por lo tanto, g_{t_1, t_2} mueve puntos no más de $\frac{\varepsilon}{2}$.

Afirmación 3. Existe $\rho > 0$ tal que, para cualesquiera $t_1, t_2 \in (t_0 - \rho, t_0 + \rho) \cap (0, \mu(X))$ con $t_1 < t_2$, g_{t_2, t_1} mueve puntos no más de $\frac{\varepsilon}{2}$ y, para cada $A \in \mu^{-1}(t_2)$, $1 - \theta_{t_1}(A) < \delta$.

Note que $1 - \delta \in (0, 1)$ y, por consiguiente, $\omega(h(A, 1 - \delta)) < \omega(h(A, 1)) = \omega(A)$, para cada $A \in \mu^{-1}((0, \mu(X)))$. Dado $t \in (0, \mu(X))$, sea

$$\xi(t) = \sup\{\mu(h(A, 1 - \delta)) : A \in \mu^{-1}(t)\}.$$

Como $\mu^{-1}(t)$ es compacto y $\mu \circ h$ es continua, tenemos que $\xi(t) < t$, para cada $t \in (0, \mu(X))$. Consideremos la función $\xi : [\tau_1, \tau_2] \rightarrow [0, \mu(X)]$ dada por la asociación anterior. Como ξ es una composición de funciones continuas, ξ es continua (de forma análoga a η , podemos expresar $\xi = \sup \circ (\mu \circ h)^* \circ \theta(\mu^{-1}, \{1 - \delta\})$; además, al igual que la función ínf, la función $\sup : 2^{\mathbb{R}} \rightarrow \mathbb{R}$ es continua). Así, y como $[\tau_1, \tau_2]$ es compacto, existe $\rho_2 > 0$ tal que

$$\rho_2 < \inf(\{t - \xi(t) : t \in [\tau_1, \tau_2]\} \cup \{t_0 - \tau_1, \tau_2 - t_0\}).$$

Así, para cualesquiera $t_1, t_2 \in (t_0 - \frac{\rho_2}{2}, t_0 + \frac{\rho_2}{2})$ con $t_1 < t_2$, tenemos que $t_2 - t_1 < \rho_2 < t_2 - \xi(t_2)$ y, por ende, $\xi(t_2) < t_1$. Por tanto, para cada $A \in \mu^{-1}(t_2)$, $\mu(h(A, 1 - \delta)) < t_1 = \omega(h(A, \theta_{t_1}(A)))$, lo cual implica que $1 - \delta < \theta_{t_1}(A)$, es decir, $1 - \theta_{t_1}(A) < \delta$. Así, para cada $A \in \mu^{-1}(t_2)$, $H_d(A, g_{t_2, t_1}(A)) = H_d(h(A, 1), h(A, \theta_{t_1}(A))) < \frac{\varepsilon}{2}$; esto es, g_{t_2, t_1} mueve puntos no más de $\frac{\varepsilon}{2}$.

Afirmación 4. Existe $\rho > 0$ tal que, para cualesquiera $t_1, t_2 \in (t_0 - \rho, t_0 + \rho) \cap (0, \mu(X))$ con $t_1 < t_2$, $g_{t_2, t_1} \circ g_{t_1, t_2}$ y $g_{t_1, t_2} \circ g_{t_2, t_1}$ son homotópicas a la funciones identidad en $\mu^{-1}(t_1)$ y $\mu^{-1}(t_2)$, respectivamente, por medio de homotopías que mueven puntos no más de ε .

Para cualquier $t_1, t_2 \in (0, \mu(X))$, consideremos la función $H_{t_1, t_2} : \mu^{-1}(t_1) \times [0, 1] \rightarrow \mu^{-1}(t_1)$ dada, para cada $A \in \mu^{-1}(t_1)$ y cada $s \in [0, 1]$, por

$$H_{t_1, t_2}(A, s) = r_{t_1}(C_d(A, s\Psi_{t_2}(A))).$$

Asimismo, sea $H_{t_2, t_1} : \mu^{-1}(t_2) \times [0, 1] \rightarrow \mu^{-1}(t_2)$ la función dada, para cada $A \in \mu^{-1}(t_2)$ y cada $s \in [0, 1]$, por

$$H_{t_2, t_1}(A, s) = R_{t_2}(h(A, 1 - s(1 - \theta_{t_1}(A)))).$$

Es claro que, para cada $A \in \mu^{-1}(t_1)$,

$$H_{t_1, t_2}(A, 0) = r_{t_1}(C_d(A, 0)) = r_{t_1}(A) = A,$$

$$H_{t_1, t_2}(A, 1) = r_{t_1}(C_d(A, \Psi_{t_2}(A))) = r_{t_1} \circ R_{t_2}(A) = g_{t_2, t_1} \circ g_{t_1, t_2}(A),$$

y, para cada $A \in \mu^{-1}(t_2)$,

$$H_{t_2, t_1}(A, 0) = R_{t_2}(h(A, 1)) = R_{t_2}(A) = A,$$

$$H_{t_2, t_1}(A, 1) = R_{t_2}(h(A, \theta_{t_1}(A))) = R_{t_2} \circ r_{t_1}(A) = g_{t_1, t_2} \circ g_{t_2, t_1}(A).$$

Así, H_{t_1,t_2} es una homotopía entre $Id_{\mu^{-1}(t_1)}$ y $g_{t_2,t_1} \circ g_{t_1,t_2}$, y H_{t_2,t_1} es una homotopía entre $Id_{\mu^{-1}(t_2)}$ y $g_{t_1,t_2} \circ g_{t_2,t_1}$.

Por otro lado, por las afirmaciones 2 y 3, existe ρ tal que, para cualesquiera $t_1, t_2 \in (t_0 - \rho, t_0 + \rho)$ con $t_1 < t_2$, las funciones g_{t_1,t_2} y g_{t_2,t_1} mueven puntos no más de $\frac{\varepsilon}{2}$ y si $A_1 \in \mu^{-1}(t_1)$ y $A_2 \in \mu^{-1}(t_2)$, entonces $\Psi_{t_2}(A_1) < \frac{\varepsilon}{2}$ y $1 - \theta_{t_1}(A_2) < \delta$.

Sean $t_1, t_2 \in (t_0 - \rho, t_0 + \rho)$, con $t_1 < t_2$. Dados $A \in \mu^{-1}(t_2)$ y $s \in [0, 1]$, note que $s(1 - \theta_{t_1}(A)) < \delta$. Sea $B = h(A, 1 - s(1 - \theta_{t_1}(A)))$. Luego, $H_d(h(A, 1), B) < \frac{\varepsilon}{2}$. Note que $t_1 \leq \mu(B) \leq t_2$ y, por consiguiente, $\mu(B) \in (t_0 - \rho, t_0 + \rho)$, por lo cual

$$H_d(B, H_{t_2,t_1}(A, s)) = H_d(B, g_{\mu(B),t_2}(B)) < \frac{\varepsilon}{2}.$$

Por lo tanto, $H_d(A, H_{t_2,t_1}(A, s)) \leq H_d(A, B) + H_d(B, H_{t_2,t_1}(A, s)) < \varepsilon$. Análogamente, dados $E \in \mu^{-1}(t_1)$ y $s \in [0, 1]$, se cumple que $s\Psi_{t_2}(E) < \frac{\varepsilon}{2}$ y, en consecuencia, $H_d(E, F) < \frac{\varepsilon}{2}$, donde $F = C_d(E, s\Psi_{t_2}(E))$. Además, $t_1 \leq \mu(F) \leq t_2$, por lo cual $\mu(F) \in (t_0 - \rho, t_0 + \rho)$. Así,

$$H_d(F, H_{t_1,t_2}(E, s)) = H_d(F, g_{\mu(F),t_1}(F)) < \frac{\varepsilon}{2}.$$

Por lo tanto, $H_d(E, H_{t_1,t_2}(E, s)) \leq H_d(E, F) + H_d(F, H_{t_1,t_2}(E, s)) < \varepsilon$. Esto muestra que H_{t_1,t_2} y H_{t_2,t_1} son homotopías que mueven puntos no más de ε .

Por las Afirmaciones 2, 3 y 4, existe $\rho > 0$ tal que, para cualesquiera $t_1, t_2 \in (t_0 - \rho, t_0 + \rho) \cap (0, \mu(X)]$ con $t_1 < t_2$, las funciones g_{t_1,t_2} y g_{t_2,t_1} mueven puntos no más de ε , y las funciones $g_{t_2,t_1} \circ g_{t_1,t_2}$ y $g_{t_1,t_2} \circ g_{t_2,t_1}$ son homotópicas a las funciones identidad en $\mu^{-1}(t_1)$ y $\mu^{-1}(t_2)$, respectivamente, por medio de homotopías que mueven puntos no más de ε . En particular, esto es válido si t es cualquier punto de $(t_0 - \rho, t_0 + \rho) \cap (0, \mu(X)]$ y hacemos $t_1 = t_0$ y $t_2 = t$, si $t_0 \geq t$, o $t_1 = t$ y $t_2 = t_0$, si $t < t_0$. Así, el abierto de $(0, \mu(X)]$ descrito por la expresión $(t_0 - \rho, t_0 + \rho) \cap (0, \mu(X)]$ satisface las condiciones de regularidad fuerte para el punto t_0 y el valor ε dado. Por lo tanto, $\mu|_{\mu^{-1}((0, \mu(X))]} : \mu^{-1}((0, \mu(X))) \rightarrow (0, \mu(X)]$ es una función fuertemente regular. □

Sean E, B y F espacios métricos y sea $p : E \rightarrow B$ una función continua y sobreyectiva. Se dice que p es una **fibración localmente trivial** con fibra F si, para cada punto $b \in B$, existe una vecindad abierta U de b y un homeomorfismo $h : p^{-1}(U) \rightarrow U \times F$ tal que $p_U = q \circ h$, donde $p_U = p|_{p^{-1}(U)} : p^{-1}(U) \rightarrow U$ y $q : U \times F \rightarrow U$ es la proyección correspondiente. Se dice que p es una **fibración trivial** con fibra F si existe un homeomorfismo $h : E \rightarrow B \times F$ tal que $p = q \circ h$, donde $q : B \times F \rightarrow B$ es la proyección correspondiente. Es claro que una fibración trivial es una fibración localmente trivial. El enunciado recíproco no siempre es

cierto, pero el siguiente resultado muestra una condición suficiente para asegurar su validez.

Teorema 6.3 ([19, Corolario 11.6]). *Sea $p : E \rightarrow B$ una fibración localmente trivial, con B un espacio métrico separable y localmente compacto. Si B es contráctil, entonces p es una fibración trivial.*

El siguiente resultado es un caso especial de [5, Teorema 1 (1)] que resulta de este último y de considerar el hecho de que cada función fuertemente regular entre espacios métricos compactos, cuyas fibras son retracts absolutos por vecindades y cuyo codominio es de dimensión finita, resulta ser una fibración [7, Teorema 1] (fibración de Hurewicz; para esta noción, véase [2, 4.3.11]).

Teorema 6.4 ([5, Teorema 1 (1)]). *Sean E y B dos espacios métricos compactos. Si $p : E \rightarrow B$ es una función fuertemente regular cuyas fibras son Q -variedades, $\dim B < \infty$ y B es localmente conexo por trayectorias, entonces p es una fibración localmente trivial.*

Teorema 6.5. *Sea X un continuo localmente conexo que no posee arcos libres. Si μ es una función de Whitney admisible para $C(X)$, entonces $\mu|_{\mu^{-1}((0, \mu(X)))} : \mu^{-1}((0, \mu(X))) \rightarrow (0, \mu(X))$ es una fibración trivial con fibras Q , donde Q es el cubo de Hilbert.*

Demostración. Note que, por el Teorema 6.2, $\mu|_{\mu^{-1}((0, \mu(X)))} : \mu^{-1}((0, \mu(X))) \rightarrow (0, \mu(X))$ es una función fuertemente regular. Se sigue de forma inmediata que $\mu|_{\mu^{-1}((0, \mu(X)))} : \mu^{-1}((0, \mu(X))) \rightarrow (0, \mu(X))$ es una función fuertemente regular. Además, por el Teorema 5.15, las fibras de $\mu|_{\mu^{-1}((0, \mu(X)))}$ son cubos de Hilbert y, por consiguiente, Q -variedades. Asimismo, $(0, \mu(X))$ es arco conexo y $\dim((0, \mu(X))) = 1$. De este modo, aplicando el Teorema 6.4, obtenemos que $\mu|_{\mu^{-1}((0, \mu(X)))}$ es una fibración localmente trivial. Como $(0, \mu(X))$ es un espacio métrico separable, localmente compacto y contráctil, se tiene, por el Teorema 6.3, que $\mu|_{\mu^{-1}((0, \mu(X)))}$ es una fibración trivial. \square

Dados dos espacios métricos X y Y , decimos que X está ε -**homotópicamente dominado** por Y si existen funciones $f : X \rightarrow Y$ y $g : Y \rightarrow X$ tales que existe una ε -homotopía entre Id_X y $g \circ f$, es decir, tales que existe una función continua $h : X \times [0, 1] \rightarrow X$ que cumple, para cada $x \in X$, $h(x, 0) = x$, $h(x, 1) = g \circ f(x)$ y el diámetro del conjunto $\{h(x, t) : t \in [0, 1]\}$ es menor que ε .

Teorema 6.6 ([9, Teorema 7.2]). *Sea X un espacio métrico separable. Entonces, X es un retracto absoluto por vecindades si y solo si existe una métrica de X tal que, para cada $\varepsilon > 0$, existe un retracto absoluto por vecindades Z tal que X es ε -homotópicamente dominado por Z .*

Teorema 6.7. *Sea X un continuo localmente conexo y sea μ una función de Whitney admisible para $\mathcal{H} = 2^X$ o $\mathcal{H} = C(X)$. Si $\mu|_{\mu^{-1}([0, \mu(X)])} : \mu^{-1}([0, \mu(X)]) \rightarrow [0, \mu(X)]$ es una función fuertemente regular, entonces X es un retracto absoluto.*

Demostración. Supongamos que $\mu|_{\mu^{-1}([0, \mu(X)])}$ es una función fuertemente regular. Sea $\varepsilon > 0$. Note que, en particular, existe $r_\varepsilon > 0$ tal que, si $t \in [0, r_\varepsilon)$, entonces existen funciones continuas $f : \mu^{-1}(0) \rightarrow \mu^{-1}(t)$ y $g : \mu^{-1}(t) \rightarrow \mu^{-1}(0)$ tales que $f \circ g$ es homotópica a $Id_{\mu^{-1}(0)}$ por medio de una homotopía que mueve puntos no más de $\frac{\varepsilon}{3}$. Además, note que una homotopía que mueve puntos no más de $\frac{\varepsilon}{3}$ es una ε -homotopía. Así, $\mu^{-1}(0)$ está ε -dominada por $\mu^{-1}(\frac{1}{2}r_\varepsilon)$, el cual, por el Teorema 5.4, es un retracto absoluto. De este modo, aplicando el Teorema 6.6, obtenemos que $\mu^{-1}(0)$ es un retracto absoluto por vecindades. Como $\mu^{-1}(0) = F_1(X)$ y $F_1(X)$ es homeomorfo a X , concluimos que X es un retracto absoluto por vecindades. Más aún, como X es contráctil (según el Teorema 5.4), tenemos que X es un retracto absoluto (véase [20, Corolario 1.4.5]). \square

Lema 6.8. *Sea X un espacio topológico y sea K su cono topológico. Si X es un espacio métrico compacto, entonces K es un continuo. Si, además, X es localmente conexo, entonces K es localmente conexo.*

Demostración. Para la primera parte del enunciado, véase [17, 3.15]. Supongamos que, además, X es localmente conexo. Note que K es la imagen de X bajo una función continua, a saber, la proyección natural $q : X \rightarrow K$ del cociente $K = (X \times [0, 1]) / X \times \{1\}$. Como X es compacto y K es metrizable, q es una función cerrada. Aplicando [17, Proposición 8.16], se obtiene que K es localmente conexo. \square

Teorema 6.9 ([20, Teorema 1.4.6]). *Sea X un espacio métrico separable. Entonces, X es un retracto absoluto por vecindades si y solo si el cono topológico sobre X es un retracto absoluto.*

Corolario 6.10 (al Teorema 6.7). *Si K es el cono topológico sobre un espacio métrico compacto y localmente conexo el cual no es un retracto absoluto por vecindades, entonces cada función de Whitney admisible para $C(K)$ o para 2^K no es una función fuertemente regular.*

Demostración. Sea X el espacio del cual K es el cono topológico. Note que, por el Lema 6.8, K es metrizable, compacto, conexo y localmente conexo. Así, K es un continuo localmente conexo. Sea $\mathcal{H} = C(K)$ o $\mathcal{H} = 2^K$. Sea $\mu : \mathcal{H} \rightarrow [0, \mu(K)]$ una función de Whitney admisible. Supongamos que μ es fuertemente regular. Se sigue inmediatamente que $\mu|_{\mu^{-1}([0, \mu(K)])} : \mu^{-1}([0, \mu(K)]) \rightarrow [0, \mu(K)]$ es fuertemente regular. Aplicando el Teorema 6.7, obtenemos que K es un retracto absoluto. Sin

embargo, como X no es un retracto absoluto por vecindades, K no es un retracto absoluto (Teorema 6.9). Esta contradicción muestra que μ no es fuertemente regular. \square

Agradecimientos

Agradecemos al amable árbitro, cuyos valiosos comentarios y sugerencias permitieron mejorar este capítulo.

Bibliografía

- [1] M. R. Adhikari. *Basic Algebraic Topology and its Applications*. Springer India, 2016.
- [2] M. Aguilar, S. Gitler, y C. Prieto. *Topología Algebraica. Un enfoque homotópico*. McGraw Hill/Interamericana Editores S.A. de C.V., 1998.
- [3] J. G. Ahuatzi Reyes. Los continuos enrejados tienen n -ésimo hiperespacio único. Tesis de licenciatura, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2014.
- [4] J. G. Ahuatzi Reyes, D. Herrera Carrasco, y F. Macías Romero. Continuos localmente conexos sin hiperespacio único $C_n(X)$. *Matemáticas y sus aplicaciones 6*, págs. 215–240. Dirección de Fomento Editorial de la BUAP, Textos Científicos, 2015.
- [5] T. A. Chapman and S. Ferry. Hurewicz fiber maps with ANR fibers. *Topology*, 16:131–143, 1977.
- [6] J. J. Charatonik and C. E. Eberhart. On smooth dendroids. *Fund. Math.*, 47:297–322, 1970.
- [7] S. Ferry. Strongly regular mappings with compact ANR fibers are Hurewicz fiberings. *Pacific Journal of Mathematics*, 75(2):373–382, October 1978.
- [8] J. T. Goodykoontz and S. B. Nadler, Jr. Whitney levels in hyperspaces of certain Peano continua. *Trans. Amer. Math. Soc.*, 274:671–694, 1982.
- [9] O. Hanner. Some theorems on absolute neighborhood retracts. *Arkiv för Matematik*, 1:389–408, 1951.
- [10] S. T. Hu. *Theory of Retracts*. Wayne State University Press, 1965.

- [11] A. Illanes and R. Leonel. Whitney equivalent continua. *Topology Proc.*, 39:293–315, 2012.
- [12] A. Illanes and S. B. Nadler, Jr. *Hyperspaces: Fundamentals and Recent Advances*, volume 126 of *Monographs and Textbooks in Pure and Applied Math.* Marcel Dekker, Inc., New York, Basel, 1999.
- [13] H. Kato. Concerning hyperspaces of certain peano continua and strong regularity of whitney maps. *Pacific Journal of Mathematics*, 119(1):159–167, 1985.
- [14] J. L Kelley. Hyperspaces of a continuum. *Trans. Amer. Math. Soc.*, 52:23–36, 1942.
- [15] K. Kuratowski. *Topology*, volume II. Academic Press, New York-London and PWN, Warszawa, 1968. New edition, revised and augmented. Translated from the French by A. Kirkor.
- [16] T. B. Muenzenberger, R. E. Smithson, and L. E. Ward, Jr. Characterizations of arboroids and dendritic spaces. *Pacific Journal of Mathematics*, 102(1):107–121, 1982.
- [17] S. B. Nadler, Jr. *Continuum Theory: An Introduction*, volume 158 of *Monographs and Textbooks in Pure and Applied Mathematics*. New York: Marcel Dekker Inc., 1992.
- [18] S. B. Nadler, Jr. *Hyperspaces of Sets: A Text with Research Questions*, volumen 33 de *Aportaciones Matemáticas: Textos*. Sociedad Matemática Mexicana, 2006.
- [19] N. Steenrod. *Topology of fibre bundles*. Princeton University Press, 1951.
- [20] J. van Mill. *The Infinite-Dimensional Topology of Function Spaces*, volume 64 of *North-Holland Mathematical Library*. North-Holland, 2001.
- [21] H. Whitney. Regular Families of Curves, I. *Proc. Nac. Acad. Sci.*, 18(3):275–278, 1932.
- [22] H. Whitney. Regular families of curves. *Annals of Mathematics*, 34(2):244–270, 1933.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570

215571041@alumnos.fcfm.buap.mx

dherrera@fcfm.buap.mx

fmacias@fcfm.buap.mx

Capítulo 9

Tercer producto simétrico único

Vianey Córdova Salazar, David Herrera Carrasco,
Fernando Macías Romero
FCFM, BUAP

Resumen

Un *continuo* X es un espacio métrico, no vacío, que es compacto y conexo. Dado $n \in \mathbb{N}$, consideremos el hiperespacio $F_n(X)$, conocido como el *n -ésimo producto simétrico de X* , con la métrica de Hausdorff. Un continuo X tiene *n -ésimo producto simétrico único* si para cualquier continuo Y tal que $F_n(X)$ es homeomorfo a $F_n(Y)$, entonces X es homeomorfo a Y . En este capítulo exponemos modelos para los hiperespacios, en especial para el segundo y tercer producto simétrico, además mencionamos resultados recientes sobre el tercer producto simétrico único.

1 Introducción

Un *continuo* X es un espacio métrico, no vacío, que es compacto y conexo. Dado un continuo X , los *hiperespacios* son familias de subconjuntos de X con alguna característica particular, considerados con la métrica de Hausdorff, [21, Theorem 1.2], [9]. Denotamos por 2^X y $F_n(X)$ a los espacios $\{A \subset X : A \text{ es un conjunto cerrado en } X \text{ y no vacío}\}$ y $\{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\}$, respectivamente. El hiperespacio $F_n(X)$ es conocido como el *n -ésimo producto simétrico de X* . El hiperespacio $F_1(X)$ es una copia isométrica de X encajada en cada n -ésimo producto simétrico. EL concepto del n -ésimo producto simétrico fue introducido por K. Borsuk y S. Ulam en [4]. El tema del n -ésimo producto simétrico fue retomado por Alejandro Illanes en el año 2002 en el artículo *Dendrites with unique hyperspace* $F_2(X)$, [19]; después en el año 2006, Enrique Castañeda y Alejandro Illanes, mostraron que las gráficas finitas tienen producto simétrico único, [7]. En el 2009 G. Acosta, R. Hernández-Gutiérrez, V. Martínez-de-la-Vega, publicaron en la revista Glasnik Math., el artículo *Dendrites and symmetric products*, [1]; y en el mismo año D. Herrera-Carrasco, M. de J. López, F. Macías-Romero mostraron que las dendritas cuyo conjunto de puntos extremos es cerrado tienen producto simétrico único [16]. En el año 2012, D. Herrera-Carrasco, F. Macías-Romero, F. Vázquez-Juárez,

generalizaron el resultado obtenido en [16, Teorema 3.7], ellos mostraron que para $n \in \mathbb{N} - \{2, 3\}$, la clase de los continuos localmente conexos tal que $\mathcal{E}_n(X)$ es denso en $F_n(X)$ tienen n -ésimo producto simétrico único, más aún, mostraron que la clase de los continuos casi enrejado localmente conexos tienen n -ésimo producto simétrico único, para $\mathbb{N} - \{2, 3\}$, vea [18, Teorema 4.3 y Corolario 4.4]. En el año 2013, se publican dos artículos *Continua with unique symmetric product* y *Rigidity of symmetric products*, cuyos autores son J. G. Anaya, E. Castañeda-Alvarado, A. Illanes y R. Hernández-Gutiérrez, V. Martínez-de-la-Vega, respectivamente, vea [2] y [15]. En el año 2015, se muestra que los continuos enrejados tienen segundo y tercer producto simétrico único [13], cabe mencionar que en el tema de n -ésimo producto simétrico en los casos $n = 2$ y $n = 3$ se usan técnicas distintas a los casos cuando $n \in \mathbb{N} - \{2, 3\}$.

2 Continuos e hiperespacios

A continuación enunciamos algunos conceptos y resultados que son necesarios para el desarrollo de este capítulo. En todo este trabajo, si X es un continuo y A un subconjunto de X , los símbolos $\text{cl}_X(A)$, $\text{Bd}_X(A)$ e $\text{int}_X(A)$ denotan la cerradura de A , la frontera de A y el interior de A en X , respectivamente. También, $A \subset Y \subset X$, los símbolos $\text{cl}_Y(A)$, $\text{Bd}_Y(A)$ e $\text{int}_Y(A)$ denotan la cerradura de A , la frontera de A y el interior de A en el subespacio Y de X , respectivamente. La cardinalidad de un conjunto A se representa por $|A|$. Si d es la métrica de X , $\varepsilon > 0$ y $t \in X$, el conjunto $\{x \in X : d(t, x) < \varepsilon\}$ lo denotamos por el conjunto $B_d(t, \varepsilon)$ o $B(t, \varepsilon)$ cuando no exista una posible confusión. Un espacio topológico es *no degenerado* si tiene más de un punto. La notación y definiciones que en este capítulo refieren, son tomadas como en [21].

Definición 2.1. Sean X un espacio topológico y $x \in X$. El espacio topológico X es *localmente conexo en el punto x* si para cada conjunto abierto U en X , tal que $x \in U$ existe un conjunto abierto y conexo V en X , tal que $x \in V \subset U$. Si X es localmente conexo en cada uno de sus puntos decimos que X es *localmente conexo*.

En esta sección estudiamos los conceptos de hiperespacio, métrica de Hausdorff y Topología de Vietoris.

Definición 2.2. Sean X un continuo con métrica d y $A, B \subset X$, no vacíos, denotamos por $d(A, B)$ a la *distancia de A a B* definida como $d(A, B) = \inf\{d(a, b) : a \in A \text{ y } b \in B\}$; y a la distancia de un punto p a un conjunto C por $d(p, C) = d(\{p\}, C)$.

Definición 2.3. Si $A \in 2^X$, la *nube* en X con centro en A y de radio $\varepsilon > 0$ es $N(\varepsilon, A) = \{x \in X : \text{existe } a \in A \text{ tal que } d(a, x) < \varepsilon\}$.

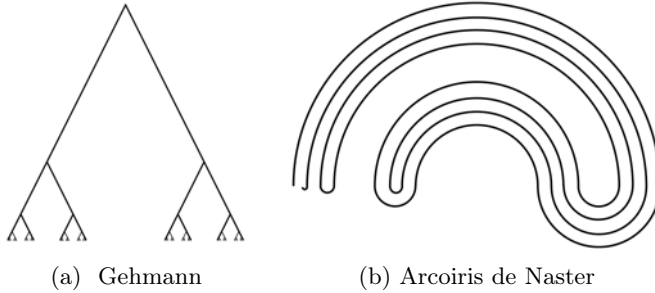


Figura 1

Si $A, B \in 2^X$, definimos $H(A, B) = \inf\{\varepsilon > 0 : A \subset N(\varepsilon, B) \text{ y } B \subset N(\varepsilon, A)\}$. En [24, Teorema 0.2], se muestra que H es una métrica para 2^X , dicha métrica es llamada la *métrica de Hausdorff*.

Definición 2.4. Sean X un continuo, $n \in \mathbb{N}$ y U_1, U_2, \dots, U_n subconjuntos de X , no vacíos. El *vietórico* de U_1, U_2, \dots, U_n , denotado por $\langle U_1, U_2, \dots, U_n \rangle$, es el conjunto $\{A \in 2^X : A \subset \bigcup_{i=1}^n U_i \text{ y } A \cap U_i \neq \emptyset, \text{ para cada } i \in \{1, 2, \dots, n\}\}$.

La familia de subconjuntos de 2^X dada por $\mathcal{B} = \{\langle U_1, U_2, \dots, U_n \rangle : n \in \mathbb{N} \text{ y los conjuntos } U_1, U_2, \dots, U_n \text{ son abiertos de } X\}$ es una familia de abiertos en 2^X que genera la misma topología que la dada por la métrica de Hausdorff, [21, Teorema 3.1]; y la topología generada por \mathcal{B} se le llama *Topología de Vietoris*, [24, Definición 0.12].

Dado un continuo X , los *hiperespacios* son familias de subconjuntos de X , con alguna característica particular, considerados con la métrica de Hausdorff. Para nuestro estudio los siguientes hiperespacios de X serán de gran importancia.

$$2^X = \{A \subset X : A \text{ es un conjunto cerrado en } X \text{ y no vacío}\},$$

$$F_n(X) = \{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\}.$$

Para una mejor lectura de este capítulo, usaremos la notación $\langle U_1, \dots, U_m \rangle_n$ para hacer referencia a los vietóricos en el n -ésimo producto simétrico. Es decir, Si X es un continuo, $U_1, U_2, \dots, U_m \subset X$ y $n, m \in \mathbb{N}$, definimos:

$$\langle U_1, \dots, U_m \rangle_n = \{A \in F_n(X) : A \subset \bigcup_{i=1}^m U_i \text{ y } A \cap U_i \neq \emptyset, \text{ para cada } i \in \{1, \dots, m\}\}.$$

Observemos que los conjuntos de la forma $\langle U_1, U_2, \dots, U_m \rangle_n$, donde U_1, U_2, \dots, U_m son abiertos en X y m es un número natural, forman una base de la topología del

n -ésimo producto simétrico, [21, Teorema 1.2]; es decir, una base para la topología inducida por la métrica de Hausdorff.

Teorema 2.5. [22, Corolario 1.8.8 y 1.8.9] Sean X un continuo y $n \in \mathbb{N}$. Los hiperespacios 2^X y $F_n(X)$ son continuos.

Teorema 2.6. [12, Lema 1.25] Sean X un continuo y $n, m \in \mathbb{N}$. Si A_1, \dots, A_n y B_1, \dots, B_m son subconjuntos de X tales que, para cada $j \in \{1, \dots, n\}$, existe $k_j \in \{1, \dots, m\}$ tal que $A_j \subset B_{k_j}$ y para cada $k \in \{1, \dots, m\}$, existe $j_k \in \{1, \dots, n\}$ tal que $A_{j_k} \subset B_k$, entonces $\langle A_1, \dots, A_n \rangle_n \subset \langle B_1, B_2, \dots, B_m \rangle_n$.

Teorema 2.7. [23, Lema 1] Sean X un continuo y $n, m \in \mathbb{N}$ tales que $m \leq n$. Si C_1, \dots, C_m son subconjuntos conexos de X , entonces $\langle C_1, \dots, C_m \rangle_n$ es un subconjunto conexo de $F_n(X)$.

Teorema 2.8. Sean X un continuo localmente conexo, $n \in \mathbb{N}$ y $A \in F_n(X)$. Si $A = \{x_1, \dots, x_m\}$ con $m \leq n$ y \mathcal{U} es una vecindad de A en $F_n(X)$, entonces existen subconjuntos U_1, \dots, U_m abiertos en X , conexos y disjuntos por pares tales que para cada $i \in \{1, \dots, m\}$, se cumple que $x_i \in U_i$ y $A \in \langle U_1, \dots, U_m \rangle_n \subset \text{int}_{F_n(X)}(\mathcal{U})$.

Demostración. Supongamos que $A = \{x_1, \dots, x_m\}$ y que \mathcal{U} es una vecindad de A en $F_n(X)$. Como X es un espacio Hausdorff, existen subconjuntos C_1, \dots, C_m abiertos en X y ajenos por pares tales que $x_i \in C_i$, para cada $i \in \{1, \dots, m\}$. Por otra parte, como $A \in \text{int}_{F_n(X)}(\mathcal{U})$, existen subconjuntos V_1, \dots, V_l abiertos en X tales que $A \in \langle V_1, \dots, V_l \rangle_n \subset \text{int}_{F_n(X)}(\mathcal{U})$. Para cada $i \in \{1, \dots, m\}$, sea $U_i = C_i \cap \left[\bigcap \{V_j : j \in \{1, \dots, l\} \text{ y } x_i \in V_j\} \right]$. Observemos que los subconjuntos U_1, \dots, U_m son abiertos en X y ajenos por pares.

Veamos que $A \in \langle U_1, \dots, U_m \rangle_n \subset \langle V_1, \dots, V_l \rangle_n$. Es claro que $A \in \langle U_1, \dots, U_m \rangle_n$. Para demostrar que $\langle U_1, \dots, U_m \rangle_n \subset \langle V_1, \dots, V_l \rangle_n$, veamos primero que para cada $j \in \{1, \dots, l\}$, existe $k_j \in \{1, \dots, m\}$ tal que $U_{k_j} \subset V_j$. Sea $j \in \{1, \dots, l\}$. Como $A \in \langle V_1, \dots, V_l \rangle_n$, inferimos que $A \cap V_j \neq \emptyset$. Sea $p_j \in A \cap V_j$. Dado que $p_j \in A$, existe $k_j \in \{1, \dots, m\}$ tal que $p_j \in U_{k_j}$. Observemos que $p_j \in C_{k_j}$. Por definición de los conjuntos U_1, \dots, U_m , tenemos que $U_{k_j} = C_{k_j} \cap \left[\bigcap \{V_j : j \in \{1, \dots, l\} \text{ y } p_j \in V_j\} \right]$. Así, $U_{k_j} \subset C_{k_j} \cap V_j$. Por tanto, $U_{k_j} \subset V_j$.

Ahora veamos que para cada $i \in \{1, \dots, m\}$, existe $k_i \in \{1, \dots, l\}$ tal que $U_i \subset V_{k_i}$. La veracidad de este inciso es inmediata de las definiciones de los conjuntos U_1, \dots, U_m . Por el Teorema 2.6, concluimos que $\langle U_1, \dots, U_m \rangle_n \subset \langle V_1, \dots, V_l \rangle_n$. Por último, observemos que dado que X es un continuo localmente conexo, podemos suponer que los subconjuntos U_1, \dots, U_m son conexos. \square

Teorema 2.9. Sean X un continuo y $n \in \mathbb{N}$. Entonces X es localmente conexo si y sólo si $F_n(X)$ es localmente conexo.

Demostración. Supongamos que X es localmente conexo. Sean $A \in F_n(X)$ y \mathcal{V} un abierto en $F_n(X)$ tal que $A \in \mathcal{V}$. Supongamos que $A = \{x_1, \dots, x_m\}$ con $m \leq n$. Por el Teorema 2.8, existen V_1, \dots, V_m abiertos en X , conexos y disjuntos dos a dos, tales que para cada $i \in \{1, \dots, m\}$, tenemos que $x_i \in V_i$ y $A \in \langle V_1, \dots, V_m \rangle_n \subset \mathcal{V}$. Notemos que $\langle V_1, \dots, V_m \rangle_n$ es abierto en $F_n(X)$, por el Teorema 2.7, tenemos que $\langle V_1, \dots, V_m \rangle_n$ es conexo en $F_n(X)$. Así, $F_n(X)$ es localmente conexo. Recíprocamente, supongamos que $F_n(X)$ es localmente conexo. Sean $p \in X$ y U un abierto en X con $p \in U$. Luego, $\{p\} \in \langle U \rangle_n$. Además, $\langle U \rangle_n$ es abierto en $F_n(X)$. Como $F_n(X)$ es localmente conexo, existe un abierto y conexo \mathcal{V} en $F_n(X)$ tal que $\{p\} \in \mathcal{V} \subset \langle U \rangle_n$. Por [12, Lema 1.49], tenemos que $(\bigcup_{V \in \mathcal{V}} V) \cap U$ es abierto y conexo en X . Notemos que $p \in (\bigcup_{V \in \mathcal{V}} V) \cap U \subset U$. Así, X es localmente conexo. \square

Definición 2.10. Sean A , un subconjunto no vacío de un espacio topológico X y β un número cardinal. Se dice que A es de *orden menor o igual* a β en X , denotado por $\text{ord}(A, X) \leq \beta$, si para cualquier conjunto abierto U de X con $A \subset U$ existe un conjunto abierto V de X , tal que $A \subset V \subset U$ y $|\text{Bd}(V)| \leq \beta$. Si $A = \{p\}$ en lugar de escribir $\text{ord}(\{p\}, X) \leq \beta$ solo se escribirá como $\text{ord}(p, X) \leq \beta$. Se dice que A es de *orden* β en X , denotado por $\text{ord}(A, X) = \beta$, si $\text{ord}(A, X) \leq \beta$ y para cualquier número cardinal $\alpha < \beta$, tenemos que $\text{ord}(A, X) \not\leq \alpha$.

Teorema 2.11. *Sea X un espacio topológico. Si Y y A son subconjuntos de X , entonces se cumple lo siguiente:*

(i) $\text{Bd}_Y(A \cap Y) \subset \text{Bd}_X(A)$.

(ii) Si $\text{cl}_X(A) \subset \text{int}_X(Y)$, entonces $\text{Bd}_Y(A) = \text{Bd}_X(A)$.

Demostración. Para demostrar (i), notemos que $\text{Bd}_Y(A \cap Y) = \text{cl}_Y(A) \cap \text{cl}_Y(Y - A) \subset \text{cl}_X(A) \cap \text{cl}_X(Y - A) \subset \text{cl}_X(A) \cap \text{cl}_X(X - A) = \text{Bd}_X(A)$.

Ahora probemos (ii). Sabemos que $\text{Bd}_X(A) = \text{cl}_X(A) \cap \text{cl}_X(X - A) = \text{cl}_X(A) \cap (\text{cl}_X(X - Y) \cup \text{cl}_X(Y - A)) = (\text{cl}_X(A) \cap (X - \text{int}_X(Y))) \cup (\text{cl}_X(A) \cap \text{cl}_X(Y - A)) = \text{cl}_X(A) \cap \text{cl}_X(Y - A) = Y \cap \text{cl}_X(A) \cap \text{cl}_X(Y - A) = \text{cl}_Y(A) \cap \text{cl}_Y(Y - A) = \text{Bd}_Y(A)$. \square

Una primera propiedad del concepto de orden es la siguiente.

Teorema 2.12. *Sea X un continuo. Si A es un subespacio de X y $p \in A$, entonces $\text{ord}(p, A) \leq \text{ord}(p, X)$. Más aún, si A es una vecindad de p en X , entonces $\text{ord}(p, A) = \text{ord}(p, X)$.*

Demostración. Sea U un subconjunto abierto de A con $p \in U$. Así, existe un subconjunto abierto V de X con $U = A \cap V$. Como $p \in V$, existe un subconjunto abierto W de X tal que $p \in W \subset V$ y $|\text{Bd}_X(W)| \leq \text{ord}(p, X)$. Por el Teorema 2.11,

tenemos que $\text{Bd}_A(A \cap W) \subset \text{Bd}_X(W)$. Así, $|\text{Bd}_A(A \cap W)| \leq \text{ord}(p, X)$. Como $A \cap W$ es abierto en A y $p \in A \cap W \subset A \cap V = U$, concluimos que $\text{ord}(p, A) \leq \text{ord}(p, X)$.

Supongamos que A es una vecindad de p en X . Por el párrafo anterior, basta probar que $\text{ord}(p, A) \leq \text{ord}(p, X)$. Sea U un subconjunto abierto de X con $p \in U$. Como $\text{int}_X(A) \cap U$ es un subconjunto abierto de A con $p \in \text{int}_X(A) \cap U$, existe un subconjunto abierto W de A tal que $p \in W \subset \text{int}_X(A) \cap U$ y $|\text{Bd}_A(W)| \leq \text{ord}(p, A)$. Como $W \subset \text{int}_X(A) \subset A$, tenemos que W es abierto en $\text{int}_X(A)$ y así en X . Por el Teorema 2.11, se cumple que $\text{Bd}_X(W) = \text{Bd}_A(W)$. Como $p \in W \subset U$ y $|\text{Bd}_X(W)| = |\text{Bd}_A(W)| \leq \text{ord}(p, A)$, tenemos que $\text{ord}(p, X) \leq \text{ord}(p, A)$. Esto concluye la prueba del teorema. \square

Definición 2.13. Dado $n \in \mathbb{N}$, al producto topológico de n intervalos $[0, 1]$ se denota con I^n , donde $I^n = \prod_{k=1}^n I_k$ y $I_k = [0, 1]$, para cada $k \in \{1, 2, \dots, n\}$. Una *n-celda* es un espacio topológico homeomorfo a I^n .

Definición 2.14. Una *gráfica finita* es un continuo que puede escribirse como la unión de una cantidad finita de arcos, tales que cualesquiera dos de ellos son ajenos o bien se intersectan en uno o en sus dos puntos extremos, únicamente.

Definición 2.15. Sea X un continuo. Un punto p en X es un punto *ordinario* de X si $\text{ord}(p, X) = 2$. El punto p es un punto de *ramificación* de X si $\text{ord}(p, X) > 2$. Un punto p es un punto *extremo* de X si $\text{ord}(p, X) = 1$. La colección de puntos ordinarios de X , se denota por $O(X)$; la colección de puntos de ramificación de X , se denota por $R(X)$ y la colección de puntos extremos de X , se denota por $E(X)$. De esta forma un continuo X puede expresarse de la siguiente manera $X = E(X) \cup O(X) \cup R(X)$.

Definición 2.16. Sea X un continuo, un *arco libre* en X es un arco J con puntos extremos x y y tal que $J - \{x, y\}$ es un conjunto abierto en X . Un *arco libre maximal* en X es un arco libre en X que es maximal con respecto a la inclusión. Un *ciclo* en X es una curva cerrada simple J contenida en X para la cual existe $z \in X$ tal que $J \cap R(X) = \{z\}$ y $J - \{z\}$ es abierto en X .

Notemos que si X no es una curva cerrada simple y J es un ciclo en X , entonces $J \cap R(X) = \{x\}$ si y solo si $J - \{x\}$ es un conjunto abierto en X .

$$\mathcal{A}_R(X) = \{J \subset X : J \text{ es un ciclo en } X\},$$

$$\mathcal{A}_E(X) = \{J \subset X : J \text{ es un arco libre maximal en } X \text{ y } |J \cap R(X)| = 1\} \text{ y}$$

$$\mathcal{A}_S(X) = \{J \subset X : J \text{ es un arco libre maximal en } X\} \cup \mathcal{A}_R(X),$$

notemos que si $J, K \in \mathcal{A}_S(X)$ y $J \neq K$, entonces $\text{int}_X(J) \cap K = \emptyset$. Sea $x, y \in X$ decimos que x es *adyacente* a y en X o que x y y son *adyacentes* en X si

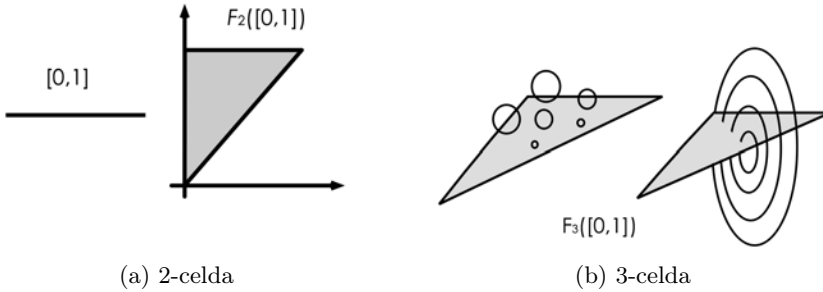


Figura 2: Modelos del segundo y tercer producto simétrico para el intervalo unitario

existe $J \in \mathcal{A}_S(X)$ el cual no es un ciclo, x y y son los puntos extremos de J . Un *arco externo* es un elemento de $\mathcal{A}_E(X)$ y un *arco interno* es un elemento de $\mathcal{A}_S(X) - (\mathcal{A}_E(X) \cup \mathcal{A}_R(X))$.

3 Modelos

En la teoría de los hiperespacios es muy útil tener ideas geométricas de cómo se ven éstos. Un modelo para un determinado hiperespacio $\mathcal{K}(X)$ es un espacio topológicamente equivalente, donde los elementos son puntos en lugar de subconjuntos. Los modelos de los hiperespacios, desde el punto de vista geométrico, son una herramienta para sugerir propiedades en ellos. El lector puede encontrar más información de este tema en [20]. Sin ahondar tanto en el tema de modelos de productos simétricos, en este capítulo mostraremos ejemplos de los modelos para el segundo y tercer producto simétrico para continuos como el intervalo unitario, la circunferencia y el triodo simple.

3.1 Intervalo unitario

Consideremos la función $\varphi : F_2(X) \rightarrow \mathbb{R}^2$ definida por $\varphi(\{a, b\}) = (\max\{a, b\}, \min\{a, b\})$, notemos que φ es un encaje cuya imagen es el triángulo convexo $\{(a, b) \in \mathbb{R}^2 : 0 \leq a \leq b \leq 1\}$, así, $F_2([0, 1])$ es una 2-celda, vea (a) de la Figura 2.

Para mostrar el modelo del tercer producto simétrico del intervalo unitario, $[0, 1]$, consideremos la función $\varphi : F_3(X) \rightarrow \mathbb{R}^2$ dada por $\varphi(A) = (\max(A), \min(A))$. Notemos que φ es una función continua y su imagen es el triángulo T de la Figura 2. Dado $(a, b) \in T$, notemos que $\varphi^{-1}((a, b)) = \{\{a, b, c\} : a \leq c \leq b\}$. En el caso $a < b$, como c corre sobre el intervalo y $\{a, a, b\} = \{a, b, b\}$, el conjunto $\{\{a, b, c\} : a \leq c \leq b\}$ es una curva cerrada simple. En el caso $a = b$, tenemos que $\varphi^{-1}((a, b)) = \{\{a\}\}$. Notemos que tenemos que poner una curva cerrada simple en cada punto $(a, b) \in T$

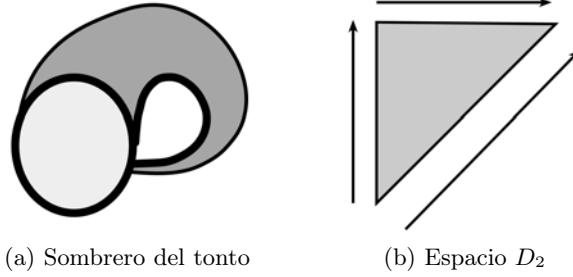


Figura 3

tal que $a < b$. Esto puede realizarse tomando un solido de revolución obtenido mediante la rotación de T alrededor de su diagonal. Así, el modelo para $F_3([0, 1])$ es una 3-celda, vea (b) de la Figura 2.

Para los casos $n \geq 4$, Borsuk and S. Ulam, en [4, Teorema 7], prueban que $F_n([0, 1])$ no es una n -celda. Un estudio detallado de los hiperespacios $F_n([0, 1])$ fue hecho por R. N. Anderson, M. M. Marjanović y R. M. Shori en [3]. En particular, en el Teorema 2.1 de [3] se probó que $F_4([0, 1])$ es homeomorfo al *cono*(D_2) \times $[0, 1]$, donde D_2 es el *sombrero del tonto*, vea (a) de la Figura 3.

Recordemos que el espacio D_2 es el espacio que se obtiene de identificar las aristas de un triángulo convexo de acuerdo a las flechas mostradas en (b) de la Figura 3.

Como primer paso tenemos que identificar dos de la flechas para obtener un cono; después, identificamos el vertice del cono a un punto en su base para obtener el espacio (a) en la Figura 3; finalmente, identificamos las dos curvas cerradas simples marcadas en negrita en la parte (a) de la Figura 3.

Es sencillo ver que D_2 puede construirse en \mathbb{R}^3 . Así, $F_4([0, 1])$ es encajado en \mathbb{R}^5 . Lo cual responde a la pregunta hecha en [20, Question 13.1] para el caso $n = 4$; dicha pregunta sigue abierta para $n \geq 5$.

Pregunta 3.1. *¿ $F_n([0, 1])$ es encajable en \mathbb{R}^{n+1} para todo $n \geq 5$?*

3.2 Circunferencia unitaria

El producto simétrico $F_2(S^1)$ es la banda de Moebius. Para esto tomemos $\mathcal{NA} = \{\{p, q\} \in F_2(S^1) : p \neq -q\}$ y sean las funciones $A : \mathcal{NA} \rightarrow C(S^1)$, $m : \mathcal{NA} \rightarrow S^1$, $L : \mathcal{NA} \rightarrow \mathbb{R}$ y $\varphi : \mathcal{NA} \rightarrow \mathbb{R}^2$ las funciones dadas por

$$A(\{p, q\}) = \text{el arco mas corto que une a } p \text{ y } q \text{ en } S^1,$$

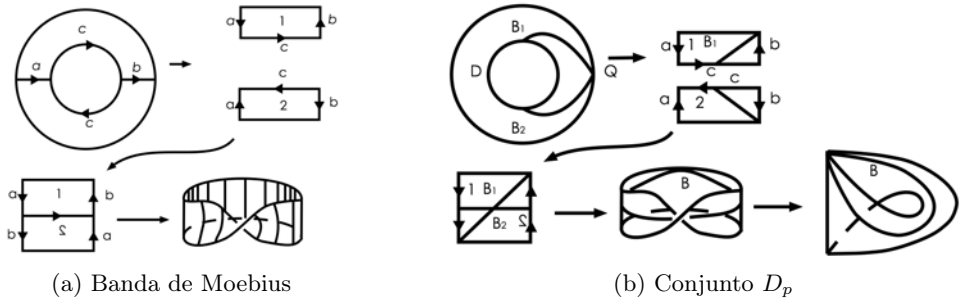


Figura 4: Esta figura muestra la construcción de la banda de Moebius y de cómo se representa en M el conjunto D_p .

$$\begin{aligned}
 m(\{p, q\}) &= \text{el punto medio de } A(\{p, q\}), \\
 L(\{p, q\}) &= \text{la longitud de } A(\{p, q\}) \text{ y} \\
 \varphi(\{p, q\}) &= (1 - (\frac{1}{2\pi}L(\{p, q\})))m(\{p, q\}).
 \end{aligned}$$

Notemos que φ es un homeomorfismo entre \mathcal{NA} y el anillo $R = \{z \in \mathbb{R}^2 : \frac{1}{2} < |z| \leq 1\}$. Si queremos extender φ al conjunto $\mathcal{A} = \{\{z, -z\} \in F_2(S^1) : z \in S^1\}$, por la continuidad y dependiendo de como nos aproximemos a $\{z, -z\}$ por elementos $\{p, q\} \in \mathcal{NA}$, debemos definir $\varphi(\{z, -z\})$ con dos valores, $\varphi(\{z, -z\}) = \pm w$, donde $2w$ es el punto obtenido de rotar z por $\frac{\pi}{2}$. Para resolver esta ambigüedad, necesitamos identificar los puntos w y $-w$. Notemos que los puntos w son los puntos de la circunferencia $B = \{u \in \mathbb{R}^2 : ||u|| = \frac{1}{2}\}$. El espacio cociente obtenido de $R_0 = R \cup B$ por la identificación es la *banda de Moebius* M , véase (a) de la Figura 4. Se inicia con el anillo y se corta por las dos flechas a y b , luego hacemos las identificaciones mostradas en (a) de la Figura 4.

A continuación, veamos cómo se representa en M , el conjuntos $D_p = \{\{p, z\} : z \in S^1\}$. De acuerdo con la definición de φ , la imagen de este conjunto consiste de dos arcos B_1 y B_2 del anillo R_0 . Si seguimos las transformaciones que hemos realizado para obtener la banda de Moebius M , se puede observar que el conjunto D_p es homeomorfo a la curva cerrada simple B que toca a la frontera de M en exactamente un punto. Esta curva cerrada simple se representa al final de la imagen (b) en la Figura 4. Es importante señalar que en esta representación de la banda de Moebius la curva B se puede dibujar totalmente sin líneas punteadas.

Como hemos visto, algunos modelos de hiperespacios son fáciles de construir; hay otros ejemplos más complicados para los que es necesario una aproximación específica. K. Borsuk en 1949 publica un artículo, vea [5], en el cual afirma que

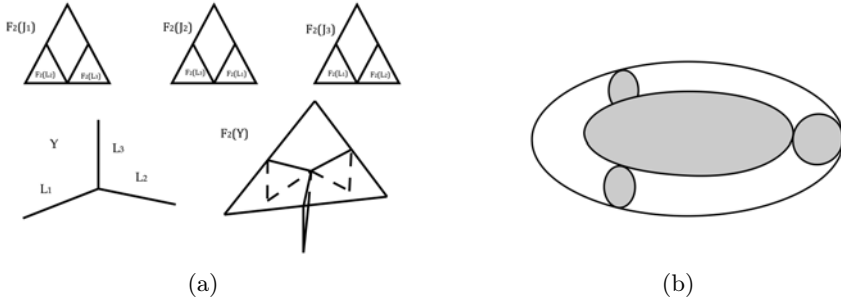


Figura 5

$F_3(S^1)$ es homeomorfo a $S^1 \times S^2$, donde S^n es la esfera unitaria en \mathbb{R}^{n+1} , lo cual es un error. Tres años después, R. Bott, en [6], corrigió este hecho probando que $F_3(S^1)$ es homeomorfo a S^3 . Para los casos $n \geq 4$ no se han construido modelos para $F_n(S^1)$, sin embargo, en [25] y [28] se han estudiado varias propiedades topológicas de estos espacios.

4 Triodo simple

Consideremos al triodo simple Y como en la imagen (a) de la Figura 5. Sea $J_1 = L_1 \cup L_2$, $J_2 = L_2 \cup L_3$ y $J_3 = L_3 \cup L_1$. Notemos que $F_2(T) = F_2(L_1) \cup F_2(L_2) \cup F_2(L_3)$. Como cada J_i es un arco, nosotros sabemos que $F_2(J_i)$ puede verse como un triángulo convexo. Así, para el modelo de $F_2(Y)$ necesitamos tomar tres triángulos $F_2(J_1)$, $F_2(J_2)$ y $F_2(J_3)$ e identificar los puntos que representan elementos de $F_2(Y)$ que aparecen en más de un triángulo. Por ejemplo, $F_2(J_1) \cup F_2(J_2) = F_2(L_3)$ es el subtriángulo en ambos triángulos $F_2(J_1)$ y $F_2(J_2)$. En la imagen (a) de la Figura 5, se mostraron los triángulos $F_2(J_1)$, $F_2(J_2)$ y $F_2(J_3)$ con las partes que necesitan para ser identificadas. El espacio resultante es el triángulo convexo con tres alas unidas a él. Enrique Castañeda en [27], encontró un modelo para $F_3(Y)$. Demostró que $F_3(Y)$ es el cono sobre un toro con cuatro discos conectados a él, uno como un «ecuador» y los otros tres como los «meridianos», vea la imagen (b) de la Figura 5.

5 Hiperespacio único

En el año 2015, Luis A. Guerrero Méndez, David Herrera Carrasco y Fernando Macías Romero mostraron, entre otras cosas, que los continuos enrejados tienen segundo y tercer producto simétrico único, vea [13]. Recientemente sabemos que los continuos casi enrejados localmente conexos tienen segundo producto simétrico

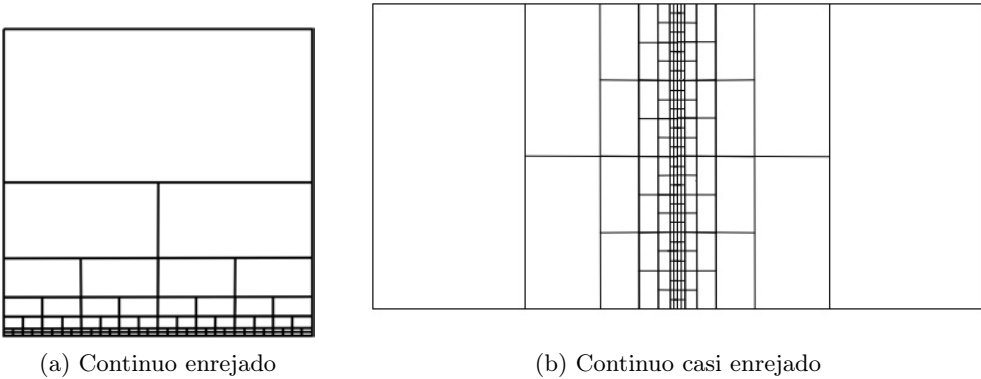


Figura 6: Ejemplos de continuos enrejado y casi enrejado

único para esto el lector puede revisar [17], en el caso $n = 3$, los autores de este capítulo, tienen una conjetura positiva a la pregunta de que si los continuos casi enrejados localmente conexos tienen tercer producto simétrico único. Cabe mencionar que en la historia de la teoría de hiperespacio único, los casos $n = 2$ y $n = 3$ usan una técnica distinta a la de los casos $n \geq 4$.

Definición 5.1. Un continuo X tiene *n-ésimo producto simétrico único* si para cualquier continuo Y tal que $F_n(X)$ es homeomorfo a $F_n(Y)$, entonces X es homeomorfo a Y .

¿Bajo qué condiciones un continuo X tiene *n-ésimo producto simétrico único*?

Dado un continuo X , sea

$$\mathcal{G}(X) = \{x \in X : x \text{ tiene una vecindad } G \text{ en } X \text{ tal que } G \text{ es una gráfica finita} \}$$

$$\text{y sea } \mathcal{P}(X) = X - \mathcal{G}(X).$$

Definición 5.2. Un continuo X es *casi enrejado* si el conjunto $\mathcal{G}(X)$ es denso en X .

Definición 5.3. Un continuo casi enrejado X es *enrejado* si X tiene una base de vecindades \mathcal{B} tal que para cada elemento $U \in \mathcal{B}$, se tiene que $U - \mathcal{P}(X)$ es conexo.

En las Figuras (a) y (b), de la Figura 6, mostramos ejemplos de continuos enrejados y casi enrejados Dado un continuo X y $n \in \mathbb{N}$, sea

$$P_n(X) = \{A \in F_n(X) : A \cap \mathcal{P}(X) \neq \emptyset\},$$

$$R_n(X) = \{A \in F_n(X) : A \cap R(X) \neq \emptyset\},$$

$$\Lambda_n(X) = F_n(X) - (R_n(X) \cup P_n(X)),$$

$$\mathcal{E}_n(X) = \{A \in F_n(X) : A \text{ tiene una vecindad en } F_n(X) \text{ la cual es una } n\text{-celda}\}.$$

A continuación exponemos algunas propiedades que el conjunto $\mathcal{E}_n(X)$ posee, además de resultados necesarios para una mejor lectura de esta sección.

Teorema 5.4. [1, Teorema 2.1] Sean X un continuo y $n \in \mathbb{N}$. Dado $i \in \{1, \dots, n\}$, sea J_i un arco en X con puntos extremos a_i y b_i . Si los conjuntos J_1, \dots, J_n son disjuntos por pares, entonces $\langle J_1, \dots, J_n \rangle_n$ es una n -celda en $F_n(X)$ cuyo interior como variedad es el conjunto $\langle J_1 - \{a_1, b_1\}, \dots, J_n - \{a_n, b_n\} \rangle_n$.

Teorema 5.5. Sean X y Y continuos y $n \in \mathbb{N}$. Si $h : F_n(X) \rightarrow F_n(Y)$ es un homeomorfismo, entonces $h(\mathcal{E}_n(X)) = \mathcal{E}_n(Y)$.

Demostración. Sea $B \in h(\mathcal{E}_n(X))$. Existe $A \in \mathcal{E}_n(X)$ tal que $B = h(A)$. Como $A \in \mathcal{E}_n(X)$, existe una vecindad V de A en $F_n(X)$ tal que V es homeomorfo a I^n . Como $h : F_n(X) \rightarrow F_n(Y)$ es un homeomorfismo, tenemos que $h(V)$ es una vecindad de $h(A)$ en $F_n(Y)$ y V es homeomorfo a $h(V)$. Así, $h(V)$ es homeomorfo a I^n . Luego, B tiene una vecindad en $F_n(Y)$ que es una n -celda. Por lo tanto, $h(\mathcal{E}_n(X)) \subset \mathcal{E}_n(Y)$.

Por otro lado, como $h^{-1} : F_n(Y) \rightarrow F_n(X)$ es un homeomorfismo, de manera similar a como se demostró que $h(\mathcal{E}_n(X)) \subset \mathcal{E}_n(Y)$, tenemos que $h^{-1}(\mathcal{E}_n(Y)) \subset \mathcal{E}_n(X)$. Luego, $\mathcal{E}_n(Y) = h(h^{-1}(\mathcal{E}_n(Y))) \subset h(\mathcal{E}_n(X))$. Por lo tanto, $h(\mathcal{E}_n(X)) = \mathcal{E}_n(Y)$. \square

Teorema 5.6. [14, Lema 2] Si X es un continuo enrejado, entonces X es un continuo localmente conexo.

Teorema 5.7. [14, Lema 8] Sean X un continuo localmente conexo, $\{J_m\}_{m=1}^{\infty}$ una sucesión en $\mathcal{A}_S(X)$ de elementos distintos por pares y $x_m \in J_m$ para cada $m \in \mathbb{N}$. Si $\{x_m\}_{m=1}^{\infty}$ converge a x , para algún $x \in X$, entonces $\{J_m\}_{m=1}^{\infty}$ converge a $\{x\}$.

Teorema 5.8. [14, Lema 10] Si X es un continuo localmente conexo y J un arco libre, entonces existe $K \in \mathcal{A}_s(X)$ tal que $J \subset K$.

Teorema 5.9. Si X es un continuo localmente conexo, entonces $\bigcup \{\text{int}_X(J) : J \in \mathcal{A}_S(X)\} = X - (\mathcal{P}(X) \cup R(X))$.

Demostración. Supongamos $x \in \bigcup \{\text{int}_X(J) : J \in \mathcal{A}_S(X)\}$, así existe $J \in \mathcal{A}_S(X)$ tal que $x \in \text{int}_X(J)$. Luego, $x \in \mathcal{G}(X)$, por otra parte observemos que $\text{ord}(x, J) = 1$ o $\text{ord}(x, J) = 2$. Por el Teorema 2.12, tenemos que $\text{ord}(x, X) = 1$ o $\text{ord}(x, X) = 2$.

Luego, $x \notin R(X)$. Por lo tanto $x \in X - (\mathcal{P}(X) \cup R(X))$. Por otro lado, supongamos que $x \in X - (\mathcal{P}(X) \cup R(X))$, así, existe una gráfica finita T en X tal que $x \in \text{int}_X(T)$. Como $x \notin R(X)$, existe un arco libre I de X tal que $x \in \text{int}_X(I)$. Por el Teorema 5.8, existe $K \in \mathcal{A}_S(X)$ tal que $I \subset K$. Así, $x \in \text{int}_X(K)$ y así $x \in \bigcup \{\text{int}_X(J) : J \in \mathcal{A}_S(X)\}$. \square

Teorema 5.10. [18, Teorema 3.1] *Si X es un continuo localmente conexo, entonces las siguientes condiciones son equivalentes.*

- i) *El continuo X es casi enrejado,*
- ii) *el conjunto $\mathcal{E}_n(X)$ es denso en $F_n(X)$, para toda $n \in \mathbb{N}$,*
- iii) *para todo subconjunto abierto no vacío de X contiene un arco libre de X .*

Teorema 5.11. [18, Teorema 3.8] *Sea X es un continuo localmente conexo y $n \in \mathbb{N}$. Si $\mathcal{E}_n(X)$ es denso en $F_n(X)$, entonces*

- i) $\mathcal{E}_n(X) \subset \Lambda_n(X)$,
- ii) $\mathcal{E}_n(X) = \Lambda_n(X) - F_{n-1}(X)$, para todo $n \in \mathbb{N} - \{2, 3\}$.

Teorema 5.12. *Si X un continuo localmente conexo tal que X no es ni un arco ni una curva cerrada simple y $n \in \mathbb{N}$, entonces las componentes de $\Lambda_n(X)$ son subconjuntos no vacíos de la forma $\langle \text{int}_X(I_1), \dots, \text{int}_X(I_m) \rangle_n$, donde $m \leq n$, los conjuntos $\text{int}_X(I_1), \dots, \text{int}_X(I_m)$ son disjuntos por pares y $I_j \in \mathcal{A}_S(X)$, para cada $j \in \{1, \dots, m\}$.*

Demostración. Sean $I_1, \dots, I_n \in \mathcal{A}_S(X)$. Notemos que $\text{int}_X(I_1), \dots, \text{int}_X(I_m)$ son subconjuntos de X abiertos y conexos tales que para $i, j \in \{1, \dots, n\}$, $\text{int}_X(I_i) \cap \text{int}_X(I_j) = \emptyset$ con $i \neq j$. Por el Teorema 2.7, tenemos que $\langle \text{int}_X(I_1), \dots, \text{int}_X(I_m) \rangle_n$ es un conjunto conexo de $F_n(X)$, además, $\langle \text{int}_X(I_1), \dots, \text{int}_X(I_m) \rangle_n$ es un subconjunto abierto en $F_n(X)$. Supongamos que $\{I_1, \dots, I_m\} \neq \{J_1, \dots, J_m\}$ así, $\langle \text{int}_X(I_1), \dots, \text{int}_X(I_m) \rangle_n \cap \langle \text{int}_X(I_1), \dots, \text{int}_X(I_m) \rangle_n = \emptyset$. Por el Teorema 5.9, tenemos que $X - (\mathcal{P}(X) \cup R(X)) = \bigcup \{\text{int}_X(I) : I \in \mathcal{A}_S(X)\}$, así, la unión de todos los elementos de la forma $\langle \text{int}_X(I_1), \dots, \text{int}_X(I_m) \rangle_n$ es igual a $\Lambda_n(X)$. Esto completa la prueba del teorema. \square

El siguiente teorema nos dice cuando el conjunto $\mathcal{E}_n(X)$ es un conjunto abierto en $F_n(X)$, para cada $n \in \mathbb{N}$. Dicho teorema el lector lo puede encontrar en [17, Lemma 2.1].

Teorema 5.13. *Si X es un continuo localmente conexo y $n \in \mathbb{N}$, entonces $\mathcal{E}_n(X)$ es un subconjunto abierto en $F_n(X)$.*

Sea X un continuo y W un conjunto abierto en X . Para cualquier subconjunto U de X definimos $c(U, W, X)$ como el número de componentes de $U \cap W$, si este número es finito y $c(U, W, X) = \infty$, en otro caso. Para cada $p \in c_X(W)$, definimos

$$v(p, W, X) = \text{mín}(\{m \in \mathbb{N} : \text{existe una base } \mathfrak{B} \text{ de vecindades de } p \text{ en } X \text{ tal que } c(U, W, X) = m, \text{ para cada } U \in \mathfrak{B}\} \cup \{\infty\}).$$

Dado $A \in F_n(X)$, de acuerdo a [13, Remark 3.5], denotamos

$$v(A) = v(A, \mathcal{E}_n(X), F_n(X)).$$

Teorema 5.14. [13, Teorema 3.4] *Si X es un continuo localmente conexo, casi enrejado y $n \in \{2, 3\}$, entonces $\mathcal{E}_n(X) = \Lambda_n(X)$.*

Demostración. Por los Teoremas 5.10 y 5.11, tenemos que $\mathcal{E}_n(X) \subset \Lambda_n(X)$. Sea $A \in \Lambda_n(X)$, así, $A \subset (E(X) \cup O(X)) - \mathcal{P}(X)$. Por el Teorema [14, Teorema 4], existe una gráfica finita G contenida en X tal que $A \subset \text{int}(G)$. Luego, $A \in \Lambda_n(G)$. Por [7, Lema 5.1], tenemos que $A \in \mathcal{E}_n(G)$. Además, sabemos que la intersección de un conjunto abierto y una n -celda contiene una n -celda, implicamos que $A \in \mathcal{E}_n(X)$. Por lo tanto, $\Lambda_n(X) \subset \mathcal{E}_n(X)$. \square

Teorema 5.15. [13, Teorema 3.8] *Sean X y Y continuos localmente conexos casi enrejados y $n \in \{2, 3\}$. Si $h : F_n(X) \rightarrow F_n(Y)$ es un homeomorfismo, entonces*

- (a) $h(R_n(X) \cup P_n(X)) = R_n(Y) \cup P_n(Y)$;
- (b) si $A \in F_n(X)$, entonces $v(A) = v(h(A))$;
- (c) $h(R_n(X) - P_n(X)) = R_n(Y) - P_n(Y)$;
- (d) $h(F_1(R(X) \cap \mathcal{G}(X))) = F_1(R(Y) \cap \mathcal{G}(Y))$;
- (e) $h(P_n(X)) = P_n(Y)$;
- (f) si X es un continuo enrejado, entonces $h(F_1(\mathcal{P}(X))) \subset F_1(\mathcal{P}(Y))$.

La siguiente propiedad es una técnica ampliamente utilizada en las pruebas de muchos resultados en el estudio de la unicidad de hiperespacios.

Definición 5.16. Sea \mathcal{C} una clase de continuos y $n \in \mathbb{N}$. La clase \mathcal{C} es F_n -cerrada si para cada $X \in \mathcal{C}$ y para cada continuo Y tal que $F_n(X)$ es homeomorfo a $F_n(Y)$, entonces $Y \in \mathcal{C}$.

Las siguientes clases de continuos son F_n -cerradas.

- (a) La clase de los continuos localmente conexo es F_n -cerrada, para cada $n \in \mathbb{N}$, [8, Teorema 6.3].
- (b) La clase de los continuos enrejados es F_n -cerrada, para cada $n \in \{2, 3\}$, [13, Teorema 3.10].

Recientemente David Herrera Carrasco, María de Jesús López Toriz y Fernando Macías Romero, mostraron el siguiente teorema.

Teorema 5.17. [17, Teorema 3.1] *La clase de los continuos localmente conexos casi enrejados es F_n -cerrada, para cada $n \in \mathbb{N}$.*

Teorema 5.18. [17, Teorema 4.1] *Sea X un continuo localmente conexo casi enrejado. Entonces $x \in \mathcal{P}(X)$ si y solo si existe una sucesión de elementos distintos por pares contenida en $\mathcal{A}_S(X)$ la cual converge a $\{x\}$.*

Teorema 5.19. [17, Teorema 4.2] *Sean X y Y continuos casi enrejados y localmente conexos. Si $h : F_2(X) \rightarrow F_2(Y)$ es un homeomorfismo y $J \in \mathcal{A}_S(X)$, entonces $h(\langle J \rangle_2) = \langle L_1, L_2 \rangle_2$, para algún $L_1, L_2 \in \mathcal{A}_S(Y)$.*

Teorema 5.20. [17, Teorema 4.3] *Sean X un continuo casi enrejado y localmente conexo, $J, K \in \mathcal{A}_S(X)$ y \mathfrak{K} una colección de componentes de $\mathcal{E}_2(X) - \langle \text{int}_X(J), \text{int}_X(K) \rangle_2$. Sea $A = \{z \in K : \text{existe } x \in \text{int}_X(J) \text{ tal que } \{x, z\} \in \text{cl}_{F_2(X)}(\bigcup \mathfrak{K})\}$ y supongamos que $A \neq \emptyset$. Entonces $A \subset \text{Bd}_X(K)$ y $\langle J, A \rangle_2 \subset \langle J, K \rangle_2 \cap \text{cl}_{F_2(X)}(\bigcup \mathfrak{K}) \subset \langle J, \text{Bd}_X(K) \rangle_2 \cup \langle \text{Bd}_X(J), K \rangle_2$.*

Los Teoremas 5.17, 5.18, 5.19 y 5.20 son de gran importancia en el desarrollo de la teoría del segundo producto simétrico único, además son de gran utilidad en el desarrollo del tercer producto simétrico único. Para el caso del tercer producto simétrico único se tienen los siguientes resultados.

Lema 5.21. *Sean X y Y continuos casi enrejados y localmente conexos. Si $h : F_2(X) \rightarrow F_2(Y)$ es un homeomorfismo y $J \in \mathcal{A}_S(X)$, entonces existen $L_1, L_2, L_3 \in \mathcal{A}_S(Y)$ tales que $h(\langle J \rangle_2) = \text{cl}_{F_3(X)}(\langle \text{int}_X(L_1), \text{int}_X(L_2), \text{int}_X(L_3) \rangle_3)$.*

Demostración. Sea $J \in \mathcal{A}_S(X)$, por los Teoremas 5.12 y 5.14, tenemos que $h(\langle \text{int}_X(J) \rangle_3) = \langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3$, para algún $K, L, M \in \mathcal{A}_S(Y)$. Así, $h(\langle J \rangle_3) = h(\text{cl}_{F_3(X)}(\langle \text{int}_X(J) \rangle_3)) = \text{cl}_{F_3(Y)}(h(\langle \text{int}_X(J) \rangle_3))$ y $\text{cl}_{F_3(Y)}(h(\langle \text{int}_X(J) \rangle_3)) = \text{cl}_{F_3(Y)}(\langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3)$. \square

Teorema 5.22. [11, Teorema 3.1] *Sean Y un continuo casi enrejado y localmente conexo, $K, L, M \in \mathcal{A}_S(Y)$, \mathfrak{K} una colección de componentes de $\mathcal{E}_3(Y) - \langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3$. Sean $A = \{z \in M : \text{existe } k \in \text{int}_Y(K), l \in \text{int}_Y(L)$*

tal que $\{k, l, z\} \in \text{cl}_{F_3(Y)}(\bigcup \mathfrak{K})$, $B = \{z \in \text{Bd}_Y(M) : \text{existe } k \in \text{Bd}_Y(K), l \in \text{int}_Y(L) \text{ tal que } \{k, l, z\} \in \text{cl}_{F_3(Y)}(\bigcup \mathfrak{K})\}$, y supongamos que $A \neq \emptyset$ y $B \neq \emptyset$. Entonces se cumple lo siguiente.

(a) $A \subset \text{Bd}_Y(M)$.

(b) Para cada $z \in A$, se cumple que $\text{cl}_{F_3(Y)}(\langle \text{int}_Y(K), \text{int}_Y(L), \{z\} \rangle_3) \subset \text{cl}_{F_3(Y)}(\langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3) \cap \text{cl}_{F_3(Y)}(\bigcup \mathfrak{K})$.

(c) Para cada $b_1 \in \text{Bd}_Y(K)$ y $b_2 \in B$, tenemos que $\text{cl}_{F_3(Y)}(\langle \{b_1\}, \text{int}_Y(L), \{b_2\} \rangle_3) \subset \text{cl}_{F_3(Y)}(\langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3) \cap \text{cl}_{F_3(Y)}(\bigcup \mathfrak{K})$,

Teorema 5.23. [11, Teorema 3.2] Sean X y Y continuos casi enrejados localmente conexos. Si $h : F_3(X) \rightarrow F_3(Y)$ es un homeomorfismo y $J \in \mathcal{A}_S(X)$ tal que $h(\langle \text{int}_X(J) \rangle_3) = \langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3$, para algún $K, L, M \in \mathcal{A}_S(Y)$, entonces

(a) $h(\langle J \rangle_3 \cap \langle R(X) \cap \mathcal{G}(X) \rangle_3) = \text{cl}_{F_3(Y)}(\langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3) \cap \langle R(Y) \cap \mathcal{G}(Y) \rangle_3$.

(b) $h(\langle J \rangle_3 \cap \langle \mathcal{P}(X) \rangle_3) = \text{cl}_{F_3(Y)}(\langle \text{int}_Y(K), \text{int}_Y(L), \text{int}_Y(M) \rangle_3) \cap \langle \mathcal{P}(Y) \rangle_3$.

(c) Si $p, q \in E(J)$ tal que $p \in \mathcal{P}(X)$ y $q \in R(X) \cap \mathcal{G}(X)$, entonces $h(\{p, q\}) = \{a, b, c\}$ tal que $\{a, b, c\} \subset \mathcal{P}(Y) \cup (R(Y) \cap \mathcal{G}(Y))$, $\{a, b, c\} \cap \mathcal{P}(Y) \neq \emptyset$ and $\{a, b, c\} \cap (R(Y) \cap \mathcal{G}(Y)) \neq \emptyset$.

Teorema 5.24. [11, Teorema 3.3] Sean X y Y continuos casi enrejados localmente conexos. Si $h : F_3(X) \rightarrow F_3(Y)$ es un homeomorfismo y $J \in \mathcal{A}_S(X)$, entonces existe $K \in \mathcal{A}_S(Y)$ tal que $h(\langle J \rangle_3) = \langle K \rangle_3$.

Bibliografía

- [1] G. Acosta, R. Hernández-Gutiérrez, V. Martínez-de-la-Vega, *Dendrites and symmetric products*, Glas. Math. Ser. III 44 (1) (2009) 195–210.
- [2] J. G. Anaya, E. Castañeda-Alvarado, A. Illanes, *Continua with unique symmetric product*. Comment. Math. Univ. Carolin. 54 (2013) 397–406.
- [3] R. N. Andersen, M. M. Marjanović and R. M. Schori, *Symmetric products and higher dimensional products*, Topol. Proc. 18 (1993), 717.
- [4] K. Borsuk, S. Ulam, *On symmetric product of topological spaces*, Bull. Amer. Math. Soc. 37 (1931) 875–882.

- [5] K. Borsuk, *On the third symmetric potency of the circumference*, Fund. Math. 36 (1949), 235–244.
- [6] R. Bott, *On the third symmetric potency of S^1* , Fund. Math. 39 (1952), 364–368.
- [7] E. Castañeda, A. Illanes, *Finite graphs have unique symmetric products*, Topol. Appl. 153 (2006) 1434–1450.
- [8] J. J. Charatonik, A. Illanes, *Local connectedness in hyperspaces*, Rocky Mt. J. Math. 36 (2006) 811–856.
- [9] V. Córdova-Salazar, *Elementos básicos de hiperespacios de continuos*, Tesis de Licenciatura en Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2011.
- [10] V. Córdova-Salazar, *Gráficas finitas con n -ésimo hiperespacio único*, Tesis de Maestría en Ciencias Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2014.
- [11] V. Córdova-Salazar, D. Herrera-Carrasco, F. Macías-Romero, *Almost meshed locally connected continua have unique third symmetric product*, enviado a Topol. Appl.
- [12] L. A. Guerrero Méndez, *Dendritas y Productos Simétricos*, Tesis de Maestría en Ciencias Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2012.
- [13] L. A. Guerrero-Méndez, D. Herrera-Carrasco, M. de J. López, F. Macías-Romero, *Meshed continua have unique second and third symmetric products*, Topol. Appl. 191 (2015) 16–27.
- [14] R. Hernández-Gutiérrez, A. Illanes, V. Martínez-de-la-Vega, *Uniqueness of hyperspaces for Peano continua*, Rocky Mt. J. Math. 43 (5) (2013) 1583–1624.
- [15] R. Hernández-Gutiérrez, V. Martínez-de-la-Vega, *Rigidity of symmetric products*, Topol. Appl. 160 (2013) 1577–1587.
- [16] D. Herrera-Carrasco, M. de J. López, F. Macías-Romero, *Dendrites with unique symmetric products*, Topol. Proc. 34 (2009) 175–190.
- [17] D. Herrera-Carrasco, M. de J. López, F. Macías-Romero, *Almost meshed locally connected continua have unique second symmetric product*, Topol. Appl. 209 (2016) 1–13.

- [18] D. Herrera-Carrasco, F. Macías-Romero, F. Vázquez-Juárez, *Peano continua with unique symmetric products*, J. Math. Res. 4 (4) (2012) 1–9.
- [19] A. Illanes, *Dendrites with unique hyperspace $F_2(X)$* , JP J. Geom. Topol. 2 (1) (2002) 75–96.
- [20] A. Illanes, *Models of Hyperspaces*, Topol. Proc. 41 (2013) 39–64.
- [21] A. Illanes, S. B. Nadler Jr., *Hyperspaces Fundamentals and Recent Advances*, Monographs and Textbooks in Pure and Applied Math., Vol. 216, Marcel Dekker, Inc., New York, 1999.
- [22] S. Macías, *Topics on continua*, Pure and Applied Mathematics Series, Vol. 275, Chapman, Hall/CRC, Taylor, Francis Group, Boca Raton, 2005.
- [23] J. M. Martínez-Montejano, *Non-confluence of the natural map of products onto symmetric products*, Lectures Notes in Pure and Applied Mathematics, 230, 229–236.
- [24] S. B. Nadler Jr., *Hyperspaces of sets. A text whit research questions*, Aportaciones Matemáticas, Serie Textos N. 33, Sociedad Matemática Mexicana, ISBN: 968-36-3594-6, 2006.
- [25] C. Tuffley, *Finite subset spaces of S^1* , Algebr. Geom. Topol. 2 (2002), 1119–1145.
- [26] F. Vázquez Juárez, *Dendritas cuyo conjunto de puntos extremos es cerrado*, Tesis de Maestría en Ciencias Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2010.
- [27] P. I. Vidal Escobar, *Los hiperespacios de continuos desde el punto de vista de sus modelos geométricos*, Tesis de Licenciatura, Universidad Autónoma del Estado de México, 2009.
- [28] W. Wu, *Note sur les produits essentiels symétriques des espaces topologique, I*, Comptes Rendus des Séances de l' Académie des Sciences 16 (1947), 1139–1141.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Ciudad Universitaria, C.P. 72570, Puebla, México.

vcordova@fcfm.buap.mx

dherrera@fcfm.buap.mx

fmacias@fcfm.buap.mx

Capítulo 10

Caracterización de los hiperespacios de un continuo localmente conexo

Lázaro Flores De Jesús, David Herrera Carrasco,
Fernando Macías Romero
FCFM, BUAP

Resumen

Un continuo es un espacio métrico no degenerado, compacto y conexo. Dado un continuo X , un hiperespacio es una familia de subconjuntos de X con ciertas características. Un teorema importante dentro de la teoría de los continuos localmente conexos es el teorema de Curtis y Schori, el cual afirma que si X es un continuo localmente conexo, entonces el hiperespacio 2^X es homeomorfo al cubo de Hilbert y también, si X no contiene arcos libres, entonces el hiperespacio $C(X)$ es homeomorfo al cubo de Hilbert. En este trabajo presentaremos una prueba del teorema de Curtis y Schori, así como el acervo necesario para su comprensión.

1 Introducción

Un **continuo** es un espacio métrico X no degenerado, compacto y conexo. Dado un continuo X , una **vecindad** de un punto $x \in X$ es un subconjunto V de X tal que existe un abierto U de X con $x \in U \subset V$. Un continuo X es **localmente conexo** en un punto $x \in X$ si para cada vecindad U de x existe un conjunto abierto y conexo V tal que $x \in V \subset U$. Si X es localmente conexo en cada uno de sus puntos, X es localmente conexo.

Dado un continuo X , un hiperespacio de X es una familia de subconjuntos de X que cumplen ciertas características. Entre los hiperespacios más conocidos se encuentran los siguientes:

$$\begin{aligned} 2^X &= \{A \subset X : A \text{ es no vacío y cerrado}\}, \\ C(X) &= \{A \in 2^X : A \text{ es compacto}\}, \\ F_n(X) &= \{A \in 2^X : A \text{ tiene a lo más } n \text{ puntos}\}. \end{aligned}$$

Los espacios anteriores resultan ser nuevamente continuos considerando en ellos la métrica de Hausdorff, [11, Teorema 2.2].

Son ejemplos de continuos localmente conexos los siguientes:

- i) continuos que se pueden escribir como una unión finita de arcos tales que cualesquiera dos de ellos se intersectan en uno o ambos de sus puntos extremos (**gráficas finitas**),
- ii) continuos que son homemorfos a $[0, 1]^n$ (**n -celdas**),
- iii) continuos que son homemorfos a $[0, 1]^\infty = I^\infty$ (**cubo de Hilbert**),
- iv) el punto peludo o también conocido como F_ω ([18, Ejemplo 2.6]).

Por otro lado, el continuo que es la cerradura de $\{(x, \sin(\frac{1}{x})) : x \in (0, 1]\}$ no es un continuo localmente conexo. Respecto a los continuos localmente conexos y sus hiperespacios tenemos el siguiente resultado.

Teorema 1.1. [15, Teorema 1.92] *Sea X un continuo localmente, entonces los hiperespacios 2^X y $C(X)$ son continuos localmente conexos.*

El presente trabajo esta dividido en cuatro secciones, en la segunda sección se presentan las definiciones y resultados que nos permitirán comprender las secciones posteriores. En la tercera sección se hará un tratamiento más a fondo sobre los continuos localmente conexos, presentaremos algunas de sus principales características. Finalmente en la cuarta sección se presentará la prueba del teorema de Curtis y Schori.

En cada resultado que se presente se incluirá su demostración o se dará una referencia en la cual el lector puede encontrar la demostración.

2 Conceptos básicos y resultados auxiliares

Las definiciones y resultados que se presentan en esta sección son necesarios para poder exponer este capítulo. De hecho los conceptos que aquí se definen los estamos considerando como se definen en [11] y [14].

Definición 2.1. Sean X y Y espacios topológicos, $A \subset Y$ y $f : A \rightarrow X$ una función continua. Una función $F : Y \rightarrow X$ es una **extensión continua** de f a Y si F es continua y $F|_A = f$. Un espacio normal X es un **extensor absoluto** si, para cada espacio normal Y , cada subconjunto cerrado A de Y y cada función continua $f : A \rightarrow X$, la función f tiene una extensión continua a Y .

Definición 2.2. Un subconjunto cerrado A de un espacio topológico Y es un **retracto** de Y si la función identidad Id_A en A tiene una extensión continua a Y . Un

espacio normal X es un **retracto absoluto** si, para cada espacio normal Y y cada subconjunto cerrado B de Y homeomorfo a X , se satisface que B es un retracto de Y .

Teorema 2.3 (Borsuk). *Un espacio métrico compacto K es un retracto absoluto si y solo si K es un extensor absoluto.*

Demostración. Sea K un espacio métrico compacto y supongamos que K es un retracto absoluto. Por el teorema de metrización de Urysohn, [12, pág. 241], podemos asumir que existen $K' \subset I^\infty$ y $\varphi : K' \rightarrow K$ un homeomorfismo. Como K es un retracto absoluto existe una retracción $r : I^\infty \rightarrow K'$.

Sea B un subconjunto cerrado de un espacio métrico M y sea $f : B \rightarrow K$ una función continua. Como la función f es continua, tenemos que $\varphi^{-1} \circ f : B \rightarrow K'$ es continua.

Notemos que $\varphi^{-1} \circ f = (f_i)_{i=1}^\infty$ donde $f_i : B \rightarrow [0, 1]_i$ para cada i .

Por el teorema de extensión de Tietze [12, pág. 127], cada f_i puede ser extendida a una función continua $g_i : M \rightarrow [0, 1]_i$. Consideremos $g = (g_i)_{i=1}^\infty : M \rightarrow I^\infty$, entonces $\varphi \circ r \circ g : M \rightarrow K$ es una extensión continua de f , es decir, $\varphi \circ r \circ g|_B = f$ pues si $x \in B$ entonces $(\varphi \circ r \circ g)(x) = (\varphi \circ r)(g(x)) = (\varphi \circ r)(\varphi^{-1} \circ f(x)) = \varphi((\varphi^{-1} \circ f)(x)) = f(x)$.

Por lo tanto, K es un extensor absoluto.

Ahora supongamos que K es un extensor absoluto. Supongamos que K es homeomorfo a un subespacio cerrado K' de un espacio métrico Y . Entonces existe $\varphi : K' \rightarrow K$ homeomorfismo. Por lo tanto, $\varphi \circ Id_{K'} : K' \rightarrow K$ es continua y como K es un extensor absoluto, $\varphi \circ Id_{K'}$ se puede extender a una función continua $f : Y \rightarrow K$ y $f|_{K'} = \varphi \circ Id_{K'}$. Así, $\varphi^{-1} \circ f : Y \rightarrow K' \subset Y$ es la función buscada, ya que si $k \in K'$ se tiene que $(\varphi^{-1} \circ f)(k) = \varphi^{-1}(f(k)) = \varphi^{-1}((\varphi \circ Id_{K'})(k)) = Id_{K'}(k) = k$. \square

Corolario 2.4. *Sea K un espacio métrico compacto encajado en I^∞ . Si K es un retracto de I^∞ , entonces K es un retracto absoluto.*

Demostración. Por el teorema 2.3 tenemos que cualquier retracto de I^∞ es un extensor absoluto. Por lo tanto, por el teorema 2.3, cualquier retracto de I^∞ es un retracto absoluto. \square

Teorema 2.5 (Wojdyslawski). [19, Teorema II, Teorema II_m] *Si X es un continuo localmente conexo, entonces 2^X y $C(X)$ son retracts absolutos.*

3 Continuos localmente conexos y métricas convexas

Dentro de esta sección se presentaran algunos resultados relacionados con los continuos localmente conexos, y tambien se presentara la relación que hay entre estos

y las métricas convexas.

Una caracterización de los espacios localmente conexo en términos de las componentes de los conjuntos abiertos, nos la da el siguiente resultado.

Teorema 3.1. [14, Ejercicio 5.22] *Un espacio métrico X es un espacio localmente conexo si y solo si para cada abierto U y cada componente C de U , se tiene que C es abierto.*

En particular, según este último teorema, las componentes de los espacios localmente conexo son conjuntos abiertos.

Existen dos formas naturales de conexidad local: Sea X un espacio topológico y $p \in X$; X es *localmente conexo* en p si p posee una base de vecindades formada por vecindades abiertas conexas; X es *conexo en pequeño (cik)* en p si p posee una base de vecindades formada por vecindades conexas (esto es, conjuntos conexas que contienen a p en sus interiores en X). Es cierto que si X es localmente conexo en p , entonces X es *cik* en p . Sin embargo, el inverso es falso aun para continuos. Sin embargo, si un espacio topológico es *cik* en *todo* punto, entonces es localmente conexo en todos sus puntos. Lo anterior queda formalizado en la definición 3.2 y en el teorema 3.3.

Definición 3.2. Un espacio métrico X es **conexo en pequeño en un punto** $x \in X$ si para cada vecindad N de x existe una vecindad conexa G de x tal que $x \in G^\circ \subset G \subset N$. Si X es conexo en pequeño en cada uno de sus puntos, se dice que X es **conexo en pequeño**.

Teorema 3.3. [14, Ejercicio 5.22] *Un espacio métrico X es localmente conexo si y solo si X es conexo en pequeño.*

La propiedad de ser un espacio localmente conexo es una propiedad topológica, es decir, se conserva bajo homeomorfismos, como lo muestra el siguiente resultado.

Teorema 3.4. *Sean X y Y espacios topológicos. Si X es un espacio localmente conexo y $f : X \rightarrow Y$ es una función continua, suprayectiva y cerrada, entonces Y es un espacio localmente conexo.*

Demostración. Sean X y Y espacios topológicos, con X un espacio localmente conexo y $f : X \rightarrow Y$ una función suprayectiva, continua y cerrada. Veamos que Y es localmente conexo.

Sea $y \in Y$ y U un abierto en Y tal que $y \in U$. Como f es suprayectiva, existe $x \in X$ tal que $f(x) = y$, también como f es continua se tiene que $f^{-1}(U)$ es un abierto en X y además $x \in f^{-1}(U)$. Como X es localmente conexo, existe V abierto

en X y conexo tal que $x \in V \subset f^{-1}(U)$, entonces $y \in f(V) \subset f(f^{-1}(U))$. Como f es suprayectiva se cumple que $f(f^{-1}(U)) = U$.

Por otro lado, $f(V)$ es un abierto en Y , pues f es cerrada y por lo tanto abierta, y $f(V)$ es conexo ya que f es continua. Así, para cada $y \in Y$ y cada abierto U que lo contiene, existe un conjunto abierto y conexo V en Y tal que $y \in V \subset U$. Por lo tanto, Y es localmente conexo en cada uno de sus puntos y así es localmente conexo. \square

Los continuos localmente conexos poseen una propiedad muy importante, conocida como la propiedad \mathcal{S} , la cual presentamos a continuación, ya que nos sera de gran utilidad mas adelante.

Definición 3.5. Sea X un espacio métrico. Un subconjunto B no vacío de X tiene la **propiedad \mathcal{S}** si para cualquier $\varepsilon > 0$ existen $\{A_1, \dots, A_n\}$ subconjuntos conexos de B de diámetro menor que ε tales que $B = \bigcup_{i=1}^n A_i$.

Como una observación inmediata, la propiedad \mathcal{S} no es una propiedad topológica. Por ejemplo, el intervalo abierto $(0, 1)$ tiene la propiedad \mathcal{S} y el espacio \mathbb{R} no tiene la propiedad \mathcal{S} .

Para acercarnos a nuestro objetivo lo primero que podemos observar es que los espacios métricos que tienen la propiedad \mathcal{S} son espacios localmente conexo, como se muestra a continuación.

Teorema 3.6. *Un espacio métrico X que tiene la propiedad \mathcal{S} es un espacio localmente conexo.*

Demostración. Basta demostrar que X es conexo en pequeño (vea el teorema 3.3). Sea $p \in X$ y N una vecindad de p . Así, existe $\varepsilon > 0$ tal que $B_{\frac{\varepsilon}{2}}(p) \subset N$. Como X tiene la propiedad \mathcal{S} , existen A_1, \dots, A_n subconjuntos conexos de X tales que $X = \bigcup_{i=1}^n A_i$ y para cada $i \in \{1, \dots, n\}$, se tiene que $\text{diám}(A_i) < \frac{\varepsilon}{2}$.
Sea

$$G = \bigcup \{A_i : p \in \overline{A_i}\}.$$

Veamos que G es conexo. Supongamos, por el contrario, que G no es conexo. Así, existe una separación (S, T) tal que S y T son no vacíos, abiertos en X , ajenos y $G = S \cup T$. Como $p \in X$, existe $k \in \{1, \dots, n\}$ tal que $p \in A_k \subset \overline{A_k}$. Luego, $A_k \subset G$. Supongamos, sin perder generalidad, que $A_k \subset S$. Como $T \neq \emptyset$, existe $j \in \{1, \dots, n\}$ tal que $A_j \subset T$. Notemos que $p \in \overline{A_j}$. Luego, existe una sucesión $\{x_m\}_{m=1}^{\infty}$ en A_j tal que $x_m \rightarrow p$. De aquí, $p \in \overline{T}$. Pero como $p \in S$, se cumple que $\overline{T} \cap S \neq \emptyset$ lo que niega nuestra hipótesis, por lo tanto G es conexo.

Ahora veamos que $p \in G^\circ$. Para esto, supongamos lo contrario, que $p \in X - G^\circ = \overline{X - G}$. Así, existe una sucesión $\{y_m\}_{m=1}^\infty$ en $X - G$ tal que $y_m \rightarrow p$. Notemos que, para cada $m \in \mathbb{N}$, el punto y_m cumple que

$$y_m \notin G. \quad (1)$$

Por otro lado, como la sucesión $\{y_m\}_{m=1}^\infty$ está en $X = \bigcup_{i=1}^n A_i$, existe una subsucesión $\{y_{m_k}\}_{k=1}^\infty$ de A_{k_0} , para algún $k_0 \in \{1, \dots, n\}$. Como $y_{m_k} \rightarrow p$, el punto $p \in \overline{A_{k_0}}$ y, por lo tanto, $A_{k_0} \subset G$. Luego, $\{y_{m_k}\}_{k=1}^\infty$ esta contenida en $A_{k_0} \subset G$, esto contradice a (1). Por lo tanto, $p \in G^\circ$.

Finalmente, veamos que $G \subset N$. Para esto, sea $g \in G$. Notemos que $g \in A_i$ para algún A_i con $p \in \overline{A_i}$. Además $d(g, p) \leq \text{diám}(A_i) < \frac{\varepsilon}{2}$. Con esto, obtenemos $G \subset B_{\frac{\varepsilon}{2}}(p)$.

En resumen, $p \in G^\circ \subset G \subset B_{\frac{\varepsilon}{2}}(p) \subset N$ y G es una vecindad conexa de N que contiene a p . Como p es arbitrario, tenemos que X es conexo en pequeño y por lo tanto es localmente conexo. \square

En relación con el resultado anterior tenemos que para espacios métricos compactos la propiedad \mathcal{S} es equivalente a ser un espacio localmente conexo.

Teorema 3.7. *Un espacio métrico compacto no vacío X es un espacio localmente conexo si y solo si tiene la propiedad \mathcal{S} .*

Demostración. Supongamos que X es un espacio localmente conexo. Sea $\varepsilon > 0$ y $x \in X$, por definición, existe V_x subconjunto abierto en X tal que V_x es conexo y $x \in V_x \subset B_{\frac{\varepsilon}{3}}(x)$. La colección $\mathcal{L} = \{V_x : x \in X\}$ es una cubierta abierta para X . Por la compacidad de X existe una colección finita V_{x_1}, \dots, V_{x_n} de \mathcal{L} tales que $X = \bigcup_{i=1}^n V_{x_i}$ con $\text{diám}(V_{x_i}) < \varepsilon$ para toda $i \in \{1, \dots, n\}$, por lo tanto se cumple la definición 3.5.

La reciproca se obtiene aplicando el teorema 3.6. \square

Teorema 3.8. *Sea X un espacio métrico. Si Y es un subconjunto de X que tiene la propiedad \mathcal{S} y Z es un subconjunto de X tal que $Y \subset Z \subset \overline{Y}_X$, entonces Z tiene la propiedad \mathcal{S} y de aquí, Z es un espacio localmente conexo.*

Demostración. Sea X un espacio métrico, Y un subconjunto de X que tiene la propiedad \mathcal{S} y Z un subconjunto de X tal que $Y \subset Z \subset \overline{Y}_X$. Sea $\varepsilon > 0$. Como Y tiene la propiedad \mathcal{S} , existen A_1, \dots, A_n subconjuntos conexos de Y tales que $Y = \bigcup_{i=1}^n A_i$ y para toda $i \in \{1, \dots, n\}$, tenemos que $\text{diám}(A_i) < \varepsilon$.

Ahora para cada $i \in \{1, \dots, n\}$, sea $B_i = \overline{(A_i)}_Z$. Por [8, Teorema 1.6, pág. 109], cada B_i es conexo.

Por otro lado, para todo $i \in \{1, \dots, n\}$, tenemos que

$$B_i = \overline{(A_i)}_Z = \overline{(A_i)}_X \cap Z \subset Z.$$

Luego $\bigcup_{i=1}^n B_i \subset Z$. Notemos que

$$\overline{(Y)}_Z = \overline{(\bigcup_{i=1}^n A_i)}_Z = \bigcup_{i=1}^n \overline{(A_i)}_Z = \bigcup_{i=1}^n B_i,$$

y como por hipótesis $Z \subset \overline{Y}_X$, tenemos que

$$Z \subset \overline{Y}_X \cap Z = \overline{Y}_Z = \bigcup_{i=1}^n B_i.$$

Así, $Z = \bigcup_{i=1}^n B_i$.

Ahora para todo $i \in \{1, \dots, n\}$, tenemos que

$$B_i = \overline{(A_i)}_Z = \overline{(A_i)}_X \cap Z \subset \overline{(A_i)}_X.$$

Luego,

$$\text{diám}(B_i) \leq \text{diám}(\overline{(A_i)}_X) = (A_i)_X < \varepsilon.$$

Así, $Z = \bigcup_{i=1}^n B_i$, y para todo $i \in \{1, \dots, n\}$, tenemos que B_i es conexo y $\text{diám}(B_i) < \varepsilon$. Por lo tanto Z tiene la propiedad S , y por el teorema 3.6, Z es un espacio localmente conexo. \square

Teorema 3.9. [14, Teorema 8.9] *Si X es un espacio métrico que tiene la propiedad S , entonces para cualquier $\varepsilon > 0$, el espacio X es unión finita de subconjuntos conexos los cuales tienen la propiedad S y diámetro menor que ε .*

Observación 3.10. La condición de subconjuntos abiertos en el teorema 3.9 puede ser cambiada por subconjuntos cerrados. Por [8, Teorema 1.6, pág. 109] y por el teorema 3.9, se cumple, para cada $i \in \{1, \dots, n\}$, que

$$\overline{S\left(A_i, \frac{\varepsilon}{3}\right)}$$

es un subconjunto conexo cerrado. Por el teorema 3.8, estos subconjuntos tienen la propiedad S ; como

$$\text{diám}\left(S\left(A_i, \frac{\varepsilon}{3}\right)\right) = \text{diám}\left(\overline{S\left(A_i, \frac{\varepsilon}{3}\right)}\right),$$

obtenemos que

$$\text{diám}\left(\overline{S\left(A_i, \frac{\varepsilon}{3}\right)}\right) < \varepsilon.$$

Por lo tanto, si X es un espacio métrico con la propiedad \mathcal{S} y $\varepsilon > 0$ entonces X se puede ver como la unión finita de subconjuntos cerrados (ó abiertos) los cuales tienen la propiedad \mathcal{S} y de diámetro menor que ε .

Teorema 3.11. *Si X es un continuo localmente conexo, entonces para cualquier $\varepsilon > 0$, el continuo X es unión finita de subcontinuos localmente conexos de diámetro menor que ε .*

Demostración. Supongamos que X es un continuo localmente conexo. Por el teorema 3.7, el continuo X tiene la propiedad \mathcal{S} . Por la observación 3.10, tenemos que X es unión de subconjuntos cerrados conexos con diámetro menor que ε que tienen la propiedad \mathcal{S} . Notemos que estos subconjuntos son compactos. Finalmente, aplicando el teorema 3.6 a cada uno de estos subconjuntos, obtenemos el resultado deseado. \square

Un aspecto importante que hay que resaltar en los continuos localmente conexos es el que se presenta a continuación.

Teorema 3.12. *[14, Teorema 8.23] Todo continuo localmente conexo X es arco-conexo.*

Un aspecto relevante de los continuos localmente conexos es que se les puede dotar de una métrica que nos permite hablar de los «puntos medios» de los segmentos entre puntos del espacio. Esta noción se desarrolla a continuación, y también damos algunos resultados interesantes.

Definición 3.13. Sea X un espacio topológico. Una **métrica convexa** para el espacio X es una métrica, d , que induce la topología en X y para la cual los puntos medios siempre existen, es decir, para cualesquiera $x, y \in X$, existe $m \in X$ tal que

$$d(x, m) = \frac{1}{2}d(x, y) = d(m, y).$$

Un aspecto importante de los continuos localmente conexos es el siguiente:

Teorema 3.14. *[13, Teorema 4] Todo continuo localmente conexo admite una métrica convexa.*

Definición 3.15. Sean (X, d) un espacio métrico, $r > 0$ y $A \in 2^X$ la **d-bola cerrada generalizada en X centrada en A de radio r** , la cual denotamos por $C_d(r, A)$, es definida como sigue:

$$C_d(r, A) = \{x \in X : d(x, A) \leq r\}.$$

Lema 3.16. [14, Proposición 10.5] Sea X un continuo con una métrica convexa d . Si $r > 0$ es fijo, entonces para cualesquiera $A, B \in 2^X$,

$$H[C_d(r, A), C_d(r, B)] \leq H(A, B)$$

Lema 3.17. [14, Proposición 10.6] Sea X un continuo con una métrica convexa d . Entonces, para cualquier $A \in C(X)$ y $r > 0$, tenemos que $C_d(r, A) \in C(X)$.

Definición 3.18. Sean X y Y espacios métricos, X con métrica acotada, y $f, g : Y \rightarrow X$ funciones continuas, denotaremos por

$$d_\infty(f, g) = \sup\{d(f(y), g(y)) : y \in Y\}.$$

Sean X y Y espacios métricos y $\{f_n\}_{n \in \mathbb{N}}$ una sucesión de funciones continuas de X en Y . Diremos que $f : X \rightarrow Y$ es el **límite uniforme** de $\{f_n\}_{n \in \mathbb{N}}$ si $\lim_{n \rightarrow \infty} d_\infty(f_n, f) = 0$.

Definición 3.19. Sean X un espacio métrico compacto y A un subconjunto cerrado de X . Diremos que A es un **Z-conjunto** en X si Id_X es el límite uniforme de funciones continuas cuyas imágenes no intersectan a A . También diremos que una función continua f entre espacios métricos compactos X_1 y X_2 es una **Z-función** si $f(X_1)$ es un Z-conjunto en X_2 .

El siguiente resultado nos da la seguridad que que podemos realizar algunas operaciones entre Z-conjuntos obteniendo nuevamente Z-conjuntos.

Teorema 3.20. Sea X un espacio métrico compacto. Entonces, se cumplen las siguientes afirmaciones:

- i) Un subconjunto cerrado de un Z-conjunto en X , es a su vez un Z-conjunto en X .
- ii) La unión finita de Z-conjuntos en X es un Z-conjunto en X .

Demostración. (1). Supongamos que A es un Z-conjunto en X y B es un subconjunto cerrado de A . Sea $\varepsilon > 0$. Luego, existe una función continua $f_\varepsilon : X \rightarrow X - A$ tal que $d_\infty(f_\varepsilon, Id_X) < \varepsilon$. Observemos que $X - A \subset X - B$. De esta manera, podemos considerar que $f_\varepsilon : X \rightarrow X - B$. Además, B es un subconjunto cerrado de X . Por lo tanto, B es un Z-conjunto en X .

(2). Bastará probar la afirmación para dos elementos. Sean A_1 y A_2 dos Z-conjuntos en X y $\varepsilon > 0$. Tomemos una función continua $f_1 : X \rightarrow X - A_1$, tal que $d_\infty(f_1, Id_X) < \frac{\varepsilon}{2}$. Observemos que, dado $a \in A$, se cumple que $d(f_1(X), A_1) \leq d(f_1(a), a) \leq d_\infty(f_1, Id_X)$. Así, $d(f_1(X), A_1) < \frac{\varepsilon}{2}$. Como X es compacto, $f_1(X)$

es compacto. Además, $f_1(X) \cap A_1 = \emptyset$. Luego, $d(f_1(X), A_1) > 0$. Así, existe una función continua $f_2 : X \rightarrow X - A_2$, tal que $d_\infty(f_2, Id_X) < \frac{1}{2}d(f_1(X), A_1)$.

Observemos que para cualquier $x \in X$ se cumple que $d(f_1(x), f_2(f_1(x))) < \frac{1}{2}d(f_1(X), A_1) < d(f_1(X), A_1)$ y, por tanto, $f_2(f_1(x)) \notin A_1$. Se sigue de esto último que $f_2(f_1(X)) \subset X - A_1$. Como también $f_2(X) \subset X - A_2$, podemos considerar la función continua $g = f_2 \circ f_1 : X \rightarrow X - (A_1 \cup A_2)$.

Observemos también que

$$\begin{aligned} d(g(x), x) &\leq d(g(x), f_1(x)) + d(f_1(x), x) = d(f_2(f_1(x)), f_1(x)) + d(f_1(x), x) \\ &< \frac{1}{2}d(f_1(X), A_1) + \frac{\varepsilon}{2} < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \frac{3\varepsilon}{4} < \varepsilon, \end{aligned}$$

para cualquier $x \in X$. Luego, $d_\infty(g, Id_X) < \varepsilon$. Por tanto, $A_1 \cup A_2$ es un Z -conjunto en X . \square

Finalizamos esta sección con un teorema importante, que nos permitirá dar la prueba del teorema de Curtis y Schori

Teorema 3.21 (Toruńczyk). [17, Teorema 1] *Sea Y un retracto absoluto. Si la función identidad en Y es un límite uniforme de Z -funciones, entonces Y es el cubo de Hilbert.*

4 El teorema de Curtis y Schori

Dado un hiperespacio de un continuo X , se pueden considerar subfamilias de subconjuntos que cumplan otra cierta característica, en la siguiente definición especificamos a que nos referimos con esto.

Definición 4.1. Dado X un continuo y un subconjunto A de X , consideremos los conjuntos $2_A^X = \{B \in 2^X : A \subset B\}$ y $C_A(X) = \{B \in C(X) : A \subset B\}$. Los subconjuntos 2_A^X y $C_A(X)$ son llamados los **hiperespacios de contención** para A en 2^X y $C(X)$, respectivamente.

Teorema 4.2. [11, Teorema 11.2] *Sea X un continuo localmente conexo no degenerado. Si K es un subconjunto cerrado de X tal que $K^\circ \neq \emptyset$, entonces 2_K^X es un Z -conjunto en 2^X ; también, si K no contiene arcos libres en X , $C_K(X)$ es un Z -conjunto en $C(X)$.*

Demostración. En vista de la definición de un Z -conjunto, comenzamos notando que por [10, Teorema 1.2.19], 2_K^X es cerrado en 2^X ; también, por [10, Teorema 1.2.19], tenemos que $C_K(X)$ es cerrado en $C(X)$.

Ahora, sea d una métrica para X . Para determinar que tan cerca está una función a la función identidad en 2^X o $C(X)$, usaremos la métrica de Hausdorff H .

Primero probaremos el teorema para 2^X_K , para esto tenemos dos casos.

Caso 1. K contiene un arco libre en X .

Sea $\varepsilon > 0$. Entonces, dado que K contiene un arco libre en X , K contiene un arco libre J en X tal que

$$\text{diám}_d(J) < \varepsilon.$$

Denotemos por p y q los puntos extremos de J , sea $\varphi : 2^J \rightarrow 2^J$ la función garantizada en [11, Lema 11.1]. Definimos $f_\varepsilon : 2^X \rightarrow 2^X$ como sigue:

$$f_\varepsilon(B) = \begin{cases} B, & \text{si } B \cap J = \emptyset, \\ (B - J^\circ) \cup \varphi(B \cap J), & \text{si } B \cap J \neq \emptyset. \end{cases}$$

Veamos que f_ε es continua en 2^X . Sea $B \in 2^X$. Entonces tenemos los dos casos siguientes:

Caso (a) $B \cap J = \emptyset$.

Sea $\{B_n\}_{n=1}^\infty$ una sucesión en 2^X que converge a B . Como $B_n \rightarrow B$ existe $N \in \mathbb{N}$ tal que para todo $n \geq N$ se cumple que $B_n \cap J = \emptyset$, así tenemos que $f_\varepsilon(B_n) = B_n$ y por lo tanto $f_\varepsilon(B_n)$ converge a $B = f_\varepsilon(B)$.

Caso (b) $B \cap J \neq \emptyset$.

Sea $\{B_n\}_{n=1}^\infty$ una sucesión en 2^X que converge a B . Tenemos que demostrar que $f_\varepsilon(B_n) \rightarrow f_\varepsilon(B)$.

Notemos que si $a^+, a^-, a^0 \in J^\circ$, entonces $a_n^+ \rightarrow a^+$, donde $a_n^+ = \inf(B_n \cap [0, 1])$, $a^+ = \inf(B \cap [0, 1])$, $a_n^- \rightarrow a^-$, donde $a_n^- = \sup(B_n \cap [-1, 0])$, $a^- = \sup(B \cap [-1, 0])$ y $a_n^0 \rightarrow a^0$, donde $a_n^0 = \{|b| : b \in B_n \cap J\}$, $a^0 = \{|b| : b \in B \cap J\}$.

Por definición tenemos que

$$\varphi(B \cap J) = \gamma(B \cap J), \text{ si } (B \cap J) \cap \left[-\frac{1}{2}, \frac{1}{2}\right] = \emptyset, \text{ o}$$

$$\varphi(B \cap J) = [(B \cap J) - (-1, 1)] \cup \{-1, 1\}, \text{ si } 0 \in B \cap J, \text{ o}$$

$$\varphi(B \cap J) = [\gamma(B \cap J) - (2a^0 - 1, 1 - 2a^0)] \cup \{2a^0 - 1, 1 - 2a^0\}, \text{ si } 0 < a^0 \leq \frac{1}{2},$$

donde

$$\gamma(B \cap J) = (B \cap J) \cup \{2a^+ - 1\}, \text{ si } (B \cap J) \subset (0, 1], \text{ o}$$

$$\gamma(B \cap J) = (B \cap J) \cup \{2a^- + 1\}, \text{ si } (B \cap J) \subset [-1, 0), \text{ o}$$

$$\gamma(B \cap J) = (B \cap J) \cup \{2a^+ - 1, 2a^- + 1\}, \text{ si } (B \cap J) \cap [-1, 0) \neq \emptyset \text{ y } (B \cap J) \cap (0, 1] \neq \emptyset.$$

De la misma manera se define $\varphi(B_n \cap J)$.

Sea $y \in f_\varepsilon(B)$ y denotemos por $E(J)$ al conjunto de puntos extremos de J .

(i) Supongamos que $y \in \varphi(B \cap J)$; primero supongamos que $y \in \gamma(B \cap J)$.

Supongamos que $y \in B \cap J^\circ$. Como J es un arco libre, existe una sucesión $\{y_n\}_{n=1}^\infty$

tal que para toda n se cumple que $y_n \in B_n \cap J^\circ$ y $y_n \rightarrow y$.

Supongamos que $y \notin \{2a^+ - 1, 2a^- + 1\}$, como $a_n^+ \rightarrow a^+$ y $a_n^- \rightarrow a^-$ existe N tal que si $n \geq N$, entonces $y_n \in B_n \cap J^\circ$, $y_n \notin \{2a_n^+ - 1, 2a_n^- + 1\}$ y tal que $y_n \rightarrow y$, para $n \geq N$, $y_n \in \gamma(B_n \cap J)$, así $y_n \in \varphi(B_n \cap J)$. De donde $y_n \in f_\varepsilon(B_n)$.

Ahora supongamos que $y \in B \cap J^\circ$ y $y = 2a^+ - 1$, entonces $y_n = 2a_n^+ - 1 \rightarrow 2a^+ - 1$. Si $y_n = 2a^- + 1$, entonces $y_n = 2a_n^- + 1 \rightarrow 2a^- + 1$. Por lo tanto $y_n \rightarrow y$ y $y_n \in \gamma(B_n \cap J)$, de donde $y_n \in \varphi(B_n \cap J)$. Así $y_n \in f_\varepsilon(B_n)$.

Ahora tomemos $y \in E(J)$. Entonces existe $\{y_n\}_{n=1}^\infty$ tal que $y_n \rightarrow y$ con $y_n \in B_n$. Tenemos los siguientes casos $y_n \in B_n \cap J^\circ$ o $y_n \in B_n - J^\circ$. De cualquier manera $y_n \in f_\varepsilon(B_n)$.

(ii) Finalmente supongamos que $y \in B - J^\circ$ y $y \notin E(J)$. Entonces existe $\{y_n\}_{n=1}^\infty$ tal que $y_n \rightarrow y$ y $y_n \in B_n - J^\circ$, entonces $y_n \in f_\varepsilon(B_n)$.

Por lo anterior, hemos demostrado que $f_\varepsilon(B) = (B - J^\circ) \cup \varphi(B \cap J) \subset \liminf(f_\varepsilon(B_n))$.

Ahora demostraremos que $\limsup((B - J^\circ) \cup \varphi(B \cap J)) \subset (B - J^\circ) \cup \varphi(B \cap J)$. Sea $x \in \limsup((B_n - J^\circ) \cup \varphi(B_n \cap J))$ y supongamos que $x \in B_n \cap J$ y $x \in J^\circ$, entonces existe una sucesión de números naturales $n_1 < n_2 < \dots$ y puntos $x_{n_k} \in B_{n_k} \cap J$ para cada $k \in \mathbb{N}$, tales que $x_{n_k} \rightarrow x$. Ya tenemos que $x \in J$, falta ver que $x \in B$. Supongamos que $x \notin B$, entonces tenemos que $d(x, B) > 0$. Sea $\varepsilon_1 = \frac{d(x, B)}{2}$.

Como $B_n \rightarrow B$ existe $N \in \mathbb{N}$ tal que para cada $n \geq N$ se cumple que $H(B_n, B) < \varepsilon_1$, así para cada $n \geq N$ tenemos que $B_n \in N(\varepsilon_1, B)$ y $B \in N(\varepsilon_1, B_n)$. Como $n_1 < n_2 < \dots$ es una sucesión de números naturales estrictamente creciente tenemos que existe n_r tal que $n_r > N$ y por lo tanto $x_{n_r} \in B_{n_r}$ y además $x_{n_r} \in N(\varepsilon_1, B)$, así existe $b \in B$ tal que $d(x_{n_r}, b) < \varepsilon_1$ y $d(x_{n_r}, x) < \varepsilon_1$. Por lo tanto tenemos que $d(x, b) \leq d(x, x_{n_r}) + d(x_{n_r}, b) < 2\varepsilon_1 = d(x, B)$, lo cual es una contradicción. Por lo tanto $x \in B$, y así $x \in B \cap J$, por lo cual $x \in \varphi(B \cap J)$.

Por otro lado si $x \in B_n \cap J$ y $x \in E(J)$, tenemos al igual que en el párrafo anterior, que $x \in B$ y por lo tanto $x \in B \cap J$, así $x \in \varphi(B \cap J)$.

Finalmente cuando $x \in B_n - J^\circ$ y $x \notin E(J)$, existe una sucesión $\{y_n\}_{n=1}^\infty$ tal que $y_n \rightarrow x$ con $y_n \in B_n - J^\circ$, por lo tanto $x \in B - J^\circ$ y así $x \in f_\varepsilon(B)$ Por lo anterior concluimos que $\limsup((B - J^\circ) \cup \varphi(B \cap J)) \subset (B - J^\circ) \cup \varphi(B \cap J)$.

Por lo tanto $\lim_{n \rightarrow \infty} f_\varepsilon(B_n) = f_\varepsilon(B)$.

Finalmente, f_ε es ε -cerca a la función identidad en 2^X (con H) dado que $\text{diám}_d(J) < \varepsilon$. Por lo tanto, hemos probado que 2_K^X es un Z -conjunto en 2^X en el caso cuando K contiene un arco libre en X .

Caso 2. K no contiene arcos libres en X .

Probaremos que para cada $\varepsilon > 0$, existe una función continua g_ε , de 2^X en $2^X - 2_K^X$ tal que g_ε es ε -cerca a la función identidad en 2^X (con H).

Sea $\varepsilon > 0$. Recordemos de las hipótesis del teorema que $K^\circ \neq \emptyset$; sea $p \in K^\circ$. Por el

teorema 3.11, $X = \bigcup_{i=1}^n X_i$, donde $n < \infty$, cada X_i es un continuo localmente conexo y $diám_d(X_i) < \frac{\varepsilon}{4}$ para cada i .

Definimos el siguiente conjunto, llamado la estrella de p con respecto a X_1, \dots, X_n , como sigue:

$$St(p) = \bigcup \{X_i : p \in X_i\}.$$

Sin pérdida de generalidad (recuerde que $p \in K^\circ$), supongamos que ε es lo suficientemente pequeño tal que $St(p) \subset K$ y $St(p) \neq X$. Sea

$$\mathcal{C} = \{X_j : p \notin X_j \text{ y } X_j \cap St(p) \neq \emptyset\}.$$

Dado que $St(p) \neq X$ y $X = \bigcup_{i=1}^n X_i$ es conexo, se tiene que $\mathcal{C} \neq \emptyset$. Para cada $X_j \in \mathcal{C}$, sea $p_j \in X_j \cap St(p)$; note que los puntos p_j realmente existen (dado que $\mathcal{C} \neq \emptyset$). Por [11, Lema 10.7], $St(p)$ es un continuo localmente conexo. Por lo tanto, por el teorema 3.12, existe un arco A_j en $St(p)$ de p a cada uno de los puntos p_j elegidos anteriormente. Sea $A = \bigcup A_j$ y sea

$$Y = A \cup (\bigcup \mathcal{C}).$$

Se sigue nuevamente de [11, Lema 10.7] que Y es un continuo localmente conexo. Por lo tanto, por el teorema 2.5, $C(Y)$ es un AR .

Notemos que

$$\overline{[X - St(p)]} \cap St(p) \subset Y. \tag{2}$$

En efecto, sea $z \in \overline{[X - St(p)]} \cap St(p)$. Sea $\{z_k\}_{k=1}^\infty$ una sucesión en $X - St(p)$ tal que $\{z_k\}_{k=1}^\infty$ converge a z . Dado que $X = \bigcup_{i=1}^n X_i$ y $n < \infty$, existe m tal que $z_k \in X_m$ para una cantidad infinita k . Esto implica que X_m tiene las siguientes tres propiedades: (i) $z \in X_m$; (ii) $p \notin X_m$ (dado que $z_k \notin St(p)$ para cualquier k); (iii) $X_m \cap St(p) \neq \emptyset$ (por (i) dado que $z \in St(p)$). Por (ii) y (iii), $X_m \in \mathcal{C}$. Por lo tanto, por (i), $z \in Y$. Esto prueba 2.

Definimos $\alpha : Y \rightarrow C(Y)$ como sigue:

$$\alpha(y) = \{y\} \text{ para todo } y \in Y.$$

Entonces, dado que $C(Y)$ es un AR , por el teorema 2.3, α puede ser extendida a una función continua $\beta : St(p) \cup Y \rightarrow C(Y)$. Además tenemos que $i : \overline{X - [St(p) \cup Y]} \rightarrow C(X)$ dada por $i(x) = \{x\}$ para cada $x \in \overline{X - [St(p) \cup Y]}$ es una función continua. Notemos que

$$(St(p) \cup Y) \cap \overline{X - [St(p) \cup Y]} = Fr(St(p) \cup Y) \subset Fr(St(p)) \cup Fr(Y).$$

Por 2 tenemos que $Fr(St(p)) \subset Y$ y, como Y es cerrado, $Fr(Y) \subset Y$. Por lo tanto, $(St(p) \cup Y) \cap \overline{X - [St(p) \cup Y]} \subset Y$.

Así, si $x \in (St(p) \cup Y) \cap \overline{X - [St(p) \cup Y]}$, se tiene que $x \in Y$ y por lo tanto $\beta(x) = \{x\} = i(x)$.

Por [8, Teorema 9.4, pág. 83] existe una función continua γ , dada por

$$\gamma(x) = \begin{cases} \beta(x), & \text{si } x \in St(p) \cup Y \\ \{x\}, & \text{si } x \in \overline{X - [St(p) \cup Y]}. \end{cases}$$

la cual es una extensión de β e i .

Ahora, utilizando la función γ definimos la función g_ε dada por: para cada $B \in 2^X$, sea

$$g_\varepsilon(B) = \bigcup \{\gamma(b) : b \in B\}.$$

Probaremos que g_ε tiene las siguientes tres propiedades:

- (a) g_ε es una función de 2^X en 2^X y g_ε es continua;
- (b) $K \not\subset g_\varepsilon(B)$ para todo $B \in 2^X$;
- (c) g_ε es ε -cercana a la función identidad en 2^X (con H).

Prueba de (a): Sea $B \in 2^X$. Dado que γ es una función continua de X en $C(X)$, $\gamma(B)$ es un subconjunto compacto no vacío de $C(X)$; por lo tanto, $\gamma(B) \in 2^{2^X}$. Entonces, dado que $g_\varepsilon(B) = \bigcup \gamma(B)$, observamos que por [3, Teorema 3.26], $g_\varepsilon(B) \in 2^X$. Esto prueba que g_ε envía a 2^X en 2^X . El hecho de que g_ε es continua se sigue de la continuidad de γ y de [3, Teorema 3.26] como sigue. Sea u la función unión descrita en [3, Teorema 3.26]. Sea $\gamma^* : 2^X \rightarrow 2^{2^X}$ definida por

$$\gamma^*(B) = \gamma(B) \text{ para todo } B \in 2^X.$$

Observamos que $g_\varepsilon = u \circ \gamma^*$; también, γ^* es continua (por la continuidad de γ y por [11, Teorema 1.3]) y u es continua. Por lo tanto, g_ε es continua. Esto prueba (a).

Prueba de (b): La razón de que (b) es verdadero es que $St(p) \not\subset Y$. Aquí los detalles. Primero probemos que

$$St(p) \not\subset Y. \tag{3}$$

Para usarse en la prueba de (3), sea $U = X - \bigcup \{X_i : p \notin X_i\}$. Note que

$$U \subset St(p) - \bigcup \mathcal{C}.$$

Por lo tanto, para probar (3), es suficiente demostrar que $U - A \neq \emptyset$ (dado que $Y = A \cup (\bigcup \mathcal{C})$). Recuerde que A fue definido como la unión de los arcos en $St(p)$ y que $St(p) \subset K$; también, recuerde nuestra hipótesis de que K no contiene arcos libres en X . Por lo tanto, A es la unión finita de arcos cada uno de los cuales tiene interior vacío en X . Por lo tanto, por el teorema de Baire [12, pág. 414], $A^\circ = \emptyset$. Entonces, dado que U es claramente no vacío y abierto en X , tenemos que $U \not\subset A$; i. e., $U - A \neq \emptyset$. Por lo tanto hemos probado (3).

Ahora, completemos la prueba de (b). Por 3, existe un punto $q \in St(p) - Y$. Recuerde la fórmula para γ y el hecho de que β es una función que va de $St(p) \cup Y$ en $C(Y)$. Entonces podemos observar que $q \notin \gamma(x)$ para cualquier $x \in X$. Por lo tanto, por la fórmula para g_ε , $q \notin g_\varepsilon(B)$ para cualquier $B \in 2^X$. Por lo tanto, dado que $q \in St(p) \subset K$, $K \not\subset g_\varepsilon(B)$ para cualquier $B \in 2^X$. Esto prueba (b).

Prueba de (c): Observemos que el $diám_d[St(p) \cup Y] < \varepsilon$. Por lo tanto, por la fórmula para γ y el hecho de que β es una función de $St(p) \cup Y$ en $C(Y)$, observamos que

$$diám[\{x\} \cup \gamma(x)] < \varepsilon \text{ para todo } x \in X.$$

Entonces, para cualquier $B \in 2^X$, se sigue que $B \subset N_d(\varepsilon, g_\varepsilon(B))$ y $g_\varepsilon(B) \subset N_d(\varepsilon, B)$. Por lo tanto, $H(g_\varepsilon(B), B) < \varepsilon$ para todo $B \in 2^X$, [11, Ejercicio 2.9]. Esto prueba (c).

Por (a), (b) y (c), 2_K^X es un Z -conjunto en 2^X .

La prueba del teorema para $C_K(X)$ es una adaptación de lo hecho para 2^X . Considere la función $g_\varepsilon|_{C(X)}$, donde g_ε es como se definió anteriormente. Probaremos que $g_\varepsilon|_{C(X)}$ es una función de $C(X)$ en $C(X)$. Sea $B \in C(X)$. Entonces, dado que $\gamma : X \rightarrow C(X)$ es continua, $\gamma(B)$ es un subcontinuo de $C(X)$; i. e., $\gamma(B) \in C[C(X)]$. Entonces, dado que $g_\varepsilon = \bigcup \gamma(B)$, por [11, Ejercicio 11.5] tenemos que $g_\varepsilon(B) \in C(X)$. Por lo tanto, en vista de los incisos (a), (b) y (c) anteriores, se sigue que $C_K(X)$ es un Z -conjunto en $C(X)$. \square

Existen tres partes del teorema de Curtis y Schori; las dos primeras partes son de importancia primaria, mientras que la tercera parte es un caso especial del teorema de Edwards, el cual enunciamos en el teorema 4.3.

Con respecto a la terminología en la tercera parte del teorema de Curtis y Schori, un *factor del cubo de Hilbert* es un espacio, Y , tal que $Y \times I^\infty \approx I^\infty$.

Para la demostración de la parte (3) del teorema de Curtis y Schori se necesita el siguiente teorema.

Teorema 4.3 (Edwards). [9] *Todo retracto absoluto es un factor del cubo de Hilbert.*

Teorema 4.4 (Curtis y Schori). *Sea X un continuo localmente conexo no degenerado. Entonces*

- (1) 2^X es el cubo de Hilbert,
 (2) $C(X)$ es el cubo de Hilbert cuando no existen arcos libres en X , y
 (3) $C(X) \times I^\infty$ es homeomorfo a I^∞ .

Demostración. (1) La prueba se basa en el teorema de Toruńczyk en 3.21. Recurriendo a la primera hipótesis de 3.21, notamos primeramente que por el teorema 2.5, 2^X y $C(X)$ son retracts absolutos.

Verificaremos la segunda hipótesis de el teorema 3.21 para 2^X y después para $C(X)$. Para esto, asumamos por el teorema 3.14 que d es una métrica convexa para X .

Sea $\varepsilon > 0$. De acuerdo al teorema 3.21, debemos probar que existe un Z -función de 2^X en 2^X que es ε -cercana a la función identidad en 2^X . Definimos $\Phi_\varepsilon : 2^X \rightarrow 2^X$ como sigue (vea la definición 3.15):

$$\Phi_\varepsilon(A) = C_d(\varepsilon, A) \text{ para todo } A \in 2^X.$$

Por el teorema 3.16, Φ_ε es continua. Observemos que $A \subset C_d(\varepsilon, A)$, por lo tanto Φ_ε es ε -cercana a la función identidad en 2^X (con H). Finalmente, demostremos que Φ_ε es una Z -función. Dado que X es compacto, existe una cantidad finita de puntos, p_1, \dots, p_n de X tales que

$$X = \bigcup_{i=1}^n C_d\left(\frac{\varepsilon}{2}, \{p_i\}\right).$$

Sea $K_i = C_d\left(\frac{\varepsilon}{2}, \{p_i\}\right)$ para cada $i \in \{1, \dots, n\}$. Por la primera parte del teorema 4.2, $2_{K_i}^X$ es un Z -conjunto en 2^X para cada i . Por lo tanto, por el teorema 3.20, $\bigcup_{i=1}^n 2_{K_i}^X$ es un Z -conjunto en 2^X . Para cada $A \in 2^X$, se tiene que existe j tal que $\Phi_\varepsilon(A) \in 2_{K_j}^X$; en otras palabras,

$$\Phi_\varepsilon(2^X) \subset \bigcup_{i=1}^n 2_{K_i}^X.$$

Entonces, por el teorema 3.20 tenemos que un subconjunto cerrado de un Z -conjunto es un Z -conjunto, así $\Phi_\varepsilon(2^X)$ es un Z -conjunto en 2^X . Por lo tanto, hemos probado que Φ_ε es una Z -función.

Por lo tanto, habiendo verificado la hipótesis de 3.21, tenemos que 2^X es el cubo de Hilbert. Esto prueba la parte (1) del teorema.

(2) Ahora, demostremos la parte (2) del teorema. Supongamos que no existen arcos libres en X . La prueba de que $C(X)$ satisface la segunda hipótesis del teorema

3.21 es una simple adaptación de lo que ya hicimos para 2^X . A saber, sea Φ_ε como se definió anteriormente, y sea $\varphi_\varepsilon = \Phi_\varepsilon|_{C(X)}$. Por el teorema 3.17, φ_ε es una función de $C(X)$ en $C(X)$. De las propiedades de Φ_ε , obtenemos que φ_ε es continua y que φ_ε es ε -cercana a la función identidad en $C(X)$. Para ver que φ_ε es una Z -función, sea $W_i = C_d(\frac{\varepsilon}{2}, \{p_i\})$ para cada $i \in \{1, \dots, n\}$. Entonces por la segunda parte del teorema 4.2, $C_{K_i}(X)$ es un Z -conjunto en $C(X)$ para cada i . Por lo tanto, por el teorema 3.20, $\bigcup_{i=1}^n C_{K_i}(X)$ es un Z -conjunto en $C(X)$. Para cada $B \in C(X)$, como $\frac{\varepsilon}{2} < \varepsilon$, se tiene que existe j tal que $\varphi_\varepsilon(B) \in C_{K_j}(X)$; en otras palabras,

$$\varphi_\varepsilon(C(X)) \subset \bigcup_{i=1}^n C_{K_i}(X).$$

Entonces, por el teorema 3.20 tenemos que un subconjunto cerrado de un Z -conjunto es un Z -conjunto, así $\varphi_\varepsilon(C(X))$ es un Z -conjunto en $C(X)$. Por lo tanto, hemos probado que φ_ε es una Z -función.

Por lo tanto, habiendo verificado la hipótesis de 3.21, nuevamente tenemos que $C(X)$ es el cubo de Hilbert. Esto prueba la parte (2) del teorema.

(3) Para la parte (3) del teorema observemos que por el teorema 2.5, $C(X)$ es un retracto absoluto, también por el corolario 2.4, I^∞ es un retracto absoluto. Por [1, Teorema 7.1, pág. 92], $C(X) \times I^\infty$ es un retracto absoluto. Por lo tanto, por el teorema 4.3, tenemos que $C(X) \times I^\infty$ es homeomorfo al cubo de Hilbert. \square

Agradecimientos

Los autores agradecen a los árbitros por haber dedicado su valioso tiempo para la revisión del trabajo, ya que con sus sugerencias y comentarios enriquecieron sustancialmente el contenido de este.

Bibliografía

- [1] K. Borsuk, *Theory of Retracts*, Monografie Matematyczne, Vol. 44, Polish Scientific Publishers, Warszawa, Poland, 1967.
- [2] F. Casarrubias Segura, A. Tamariz Mascarúa, *Elementos de Topología General*, Aportaciones Matemáticas, Serie Textos 37, Sociedad Matemática Mexicana, México D.F., 2012.
- [3] V. Córdova Salazar, *Elementos básicos de Hiperespacios de Continuos*, Tesis de Licenciatura en Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2011.

- [4] D. W. Curtis y R. M. Schori, 2^X and $C(X)$ are homeomorphic to the Hilbert cube, Bull. Amer. Math. Soc. 80, 1974, 927-931.
- [5] D. W. Curtis y R. M. Schori, *Hyperspaces of Peano continua are Hilbert cubes*, Fund. Math. 101, 1978, 19-38.
- [6] R. Duda, *On convex metric spaces III*, Fund. Math. 51, 1962, 23-33.
- [7] R. Duda, *On convex metric spaces V*, Fund. Math. 68, 1970, 87-106.
- [8] J. Dugundji, *Topology*, Allyn and Bacon, Boston, London, Sydney, Toronto, 1966.
- [9] R. D. Edwards, *Characterizing infinite dimensional manifolds topologically*, Lecture Notes in Mathematics, Vol. 770, Séminaire Bourbaki (Ed. by A. Dold and B. Eckmann), Springer-Verlag, Berlin, 1980, 278-302
- [10] L. Flores De Jesús, *Continuos localmente conexos*, Tesis de Licenciatura en Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2015.
- [11] A. Illanes y S. B. Nadler, Jr., *Hyperspaces, Fundamentals and Recent Advances*. Monographs and Textbooks in Pure and Applied Mathematics, Vol. 216, Marcel Dekker, Inc. New York and Basel, 1999.
- [12] K. Kuratowski, *Topology*, Vol. I, Acad. Press, New York, N.Y., 1966.
- [13] E. E. Moise, *Grille decomposition and convexification theorems for compact metric locally connected continua*, Bull. Amer. Math. Soc. 55 (1949), 1111-1121.
- [14] S. B. Nadler Jr., *Continuum Theory, An Introduction*. Monographs and Textbooks in Pure and Applied Mathematics, Vol. 158, Marcel Dekker, Inc., New York, 1992.
- [15] S. B. Nadler Jr., *Hyperspaces of sets*. Monographs and Textbooks in Pure and Applied Mathematics, Vol. 49, Marcel Dekker, Inc. New York and Basel, 1978.
- [16] R. M. Schori y J. E. West, 2^I is homeomorphic to the Hilbert cube, Bull. Amer. Math. Soc. 78, 1972, 402-406.
- [17] H. Toruńczyk, *On CE-images of the Hilbert cube and characterization of Q -manifolds*, Fund. Math. 106, 1980, 31-40.

- [18] F. Vázquez Juárez, *Propiedades de las dendritas cuyo conjunto de puntos extremos es cerrado*, Tesis de Maestría en Ciencias Matemáticas, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, 2010.
- [19] M. Wojdyslawski, *Rétractes absolus et hyperespaces des continus*, Fund. Math. 32, 1939, 184-192.

Facultad de Ciencias Físico Matemáticas, BUAP
Avenida San Claudio y 18 Sur, Colonia San Manuel,
Puebla, Pue. C.P. 72570

216470050@alumnos.fcfm.buap.mx

dherrera@fcfm.buap.mx

fmacias@fcfm.buap.mx

Índice de autores

- Ahuatzi Reyes, José Gerardo, 199
- Córdova Salazar, Vianey, 231
Carrasco Pacheco, José Luis, 123
Cejudo Castilla, César, 31
Contreras Carreto, Agustín, 155
- de Gante Coronel, Elizabeth, 155
- Flores De Jesús, Lázaro, 249
- García Ramírez, Ángel Raúl, 53
Gutiérrez García, Ismael, 83
- Hernández Morales, José Margarito, 123
Hernández Cervantes, Álvaro, 109
Herrera Carrasco, David, 199, 231, 249
- Jiménez Pozo, Miguel Antonio, 109
- López Andrade, Carlos Alberto, 53
- Macías Prado, María del Rocío, 155
Macías Romero, Fernando, 199, 231, 249
- Ortiz Ramírez, Ambrosio, 177
- Ramírez Contreras, Juan Manuel, 31
Rodríguez Nava, Abigail, 177
- Sánchez Zacateco, Luis Alberto, 177
- Torresblanca Badillo, Anselmo, 83
- Villa Hernández, David, 31

Matemáticas y sus aplicaciones 7
su composición, diseño y cuidado, estuvo a cargo
de Fernando Macías Romero
se terminó de imprimir el 7 de diciembre de 2016, en los talleres de
El Errante editor, S. A. de C. V., sito en privada Emiliano Zapata 5947,
Col. San Baltazar Campeche, Puebla, Pue., México.

El tiraje consta de 500 ejemplares.