

mixba'al

www.doi.org/10.24275/uami/dcbi/mix/v12n1

e-ISSN: 2007-7874



Revista de divulgación e investigación del Departamento de Matemáticas de la UAM-I, dedicada a promover las ideas y la creación de todos aquellos interesados en el quehacer matemático.



UNIVERSIDAD AUTÓNOMA METROPOLITANA

Directorio

José Antonio de los Reyes Heredia
Rector General

Rodrigo Díaz Cruz
Rector Unidad Iztapalapa

Jesús Alberto Ochoa Tapia
Director de la División de Ciencias Básicas e Ingeniería.
UAM-Iztapalapa

Patricia Saavedra Barrera
Jefa del Departamento de Matemáticas.
UAM-Iztapalapa

Mixba'al, Revista Metropolitana de Matemáticas, es una publicación anual del Departamento de Matemáticas de la UAM-Iztapalapa

Coordinador Editorial

Mario Pineda Ruelas
mpr@xanum.uam.mx

Comité Editorial

Elsa Baez Juárez
ebaeb@cua.uam.mx

Shirley Bromberg Silverstein
stbsster@gmail.com

Judith Campos Cordero
judith@ciencias.unam.mx

Martín Celli Siboni
celli@xanum.uam.mx

Pedro L. del Ángel Rodríguez
luis@ciemat.mx

Begoña Fernández
bff@ciencias.unam.mx

Silvia Gavito Ticozzi
sgt@correo.azc.uam.mx

Gustavo Izquierdo Buenrostro
iubg@xanum.uam.mx

L. Héctor Juárez Valencia
hctj@xanum.uam.mx

Jorge A. León Vázquez
jleon@ctrl.cinvestav.mx

Roberto Quezada Batalla
roqb@xanum.uam.mx

Edith Corina Sáenz Valadez
ecsv@ciencias.unam.mx

Martha L. Shaid Sandoval Miranda
marlisha@gmail.com

Ekaterina Todorova
todorova@ciemat.mx

Luis Miguel Villegas Silva
villegas63@gmail.com

Diseño logo Michael Rivera Arce
Dibujo portada Sofia Peláez

MIXBA'AL. Vol. 12, No. 1, septiembre de 2021 a septiembre de 2022, es una publicación anual de la Universidad Autónoma Metropolitana a través de la Unidad Iztapalapa, División de Ciencias Básicas e Ingeniería, Departamento de Matemáticas. Prolongación Canal de Miramontes 3855, Col. Ex Hacienda San Juan de Dios, Delegación Tlalpan, C.P. 14387, México, Ciudad de México y Av. San Rafael Atlixco, No. 186, Edificio AT, tercer piso, Col. Vicentina, Delegación Iztapalapa, C.P. 09340, México, Ciudad de México. Tel. 5804 4658. Página electrónica de la revista: <http://mat.izt.uam.mx/mat/index.php/revistamixba-al>. Correos electrónicos: mixbaal2009@gmail.com, mixb@xanum.uam.mx. Coordinador Editorial Mario Pineda Ruelas. Certificado de Reserva de Derechos al Uso Exclusivo de Título No. 04-2010-072017382600-203, ISSN: 2007-7874, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última actualización de este número Mario Pineda Ruelas, Departamento de Matemáticas, División de Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana-Iztapalapa. Av. San Rafael Atlixco No. 186, Edificio AT, tercer piso, Colonia Vicentina, Delegación Iztapalapa, C.P. 09340, México, Ciudad de México. Fecha de última modificación 31 de agosto de 2021. Tamaño del archivo 12.4 MB. Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor responsable de la publicación. Queda estrictamente prohibida la reproducción total o parcial de los contenidos e imágenes de la publicación sin previa autorización de la Universidad Autónoma Metropolitana. Contacto: Departamento de Matemáticas, Universidad Autónoma Metropolitana, Unidad Iztapalapa. Tel: (01) 55 5804 4654.

CONTENIDO

- 7 Dra. Rosa Obdulia González Robles (1952-2021). A la memoria de Luly.
M.J. Arroyo Paniagua
- 13 In memoriam: Francisco Hugo Martínez Ortíz.
Comité Editorial
- 15 Tres teoremas sobre cardinales medibles.
Franklin Galindo
- 33 Operadores simétricos multivaluados y el problema de momentos de Hamburger truncado.
Josué Rios-Cangas
- 47 El método de descomposición de Adomian para el cálculo de integrales gaussianas: ecuación de calor y de Black-Sholes.
Oswaldo González-Gaxiola
- 55 Caminando entre medidas invariantes y topologías: el Teorema de Haar Inverso.
Saúl Pilatowsky Cameo
- 71 El Teorema de metrización de Urysohn.
Fidel Casarrubias Segura
- 81 La fórmula de descomposición para la valoración de opciones de compra bajo el modelo de Heston.
Raúl Merino

A LOS AUTORES

Mixba'al, Revista Metropolitana de Matemáticas, es una revista de publicación anual de divulgación e investigación en matemáticas en el sentido más amplio, concebida con el propósito de apoyar la comunicación entre la comunidad matemática de habla hispana. Los artículos sometidos, pueden ser trabajos de investigación o trabajos que presenten de manera original algún tema de las matemáticas, por ejemplo, demostraciones nuevas de resultados conocidos; artículos panorámicos sobre un área de investigación; la presentación distinta de algún tema vinculado a la docencia; aplicaciones o aspectos lúdicos de las mismas, entre otros.

Los trabajos sometidos deben estar escritos en español impecable, aunque eventualmente podrán aceptarse trabajos en inglés. Éstos deben ser enviados a cualquiera de los editores en formato pdf y en un plazo no mayor a dos semanas, el comité editorial decidirá si el trabajo es acorde a la línea editorial de la revista; es caso que así sea, se enviará a arbitraje.

Cuando un trabajo reciba arbitraje favorable, se le comunicará al autor y en un plazo no mayor a tres semanas, deberá reenviarlo al editor responsable, atendiendo las sugerencias del árbitro y en el formato oficial que la revista le proporcionará. El trabajo debe estar organizado de la siguiente forma: título, resumen no mayor a 100 palabras, clasificación de la AMS 2020, palabras clave, introducción y en seguida el desarrollo del mismo. El formato oficial de la revista es una estupenda guía.



DRA. ROSA OBDULIA GONZÁLEZ ROBLES
(1952-2021)
A LA MEMORIA DE LULY

M. J. ARROYO PANIAGUA



Luly ingresó al Departamento de Matemáticas de la UAM Iztapalapa desde su fundación en el año de 1974 y en los cuarenta y siete años que trabajó en él siempre lo hizo con gran compromiso institucional. A esta institución le aportó un perfil y experiencia laboral únicos que no es frecuente encontrar, pues era una estudiosa de la estadística aplicada, una excelente profesora, poseía una gran capacidad para dirigir y coordinar a otros. Sabía trabajar en equipo y muy importante fue su aprecio por el desarrollo del trabajo interdisciplinario, en donde ella ponía toda su experiencia al tiempo que buscaba entender las necesidades de otros especialistas de muy diversos campos, como la sociología, la antropología, la fisiología, la neurociencia y la biotecnología.

Sin duda, su tenacidad, disciplina, franqueza, amor al trabajo, su liderazgo, formaron parte del legado familiar, que ella supo florecer y acrecentar durante su vida. Luly nació en un rancho en el ejido de Real del Castillo, mejor conocido como “Ojos Negros” que se encuentra a una distancia de 32 kilómetros de la Ciudad de Ensenada en el estado de Baja California. Actualmente es parte de la ruta del queso y del vino. Sus padres junto con sus dos primeros hijos llegaron a Ojos Negros oriundos del Estado de Jalisco, ya que el gobierno de Lázaro Cárdenas deseaba poblar esa zona y otorgaba tierras para su cultivo.

En ese rancho se formó una familia de ocho hermanos, dos hombres y seis mujeres, ella fue la número siete y la menor de las mujeres. En Ojos Negros solo existía una primaria, y para asistir a ella, la familia caminaba varios kilómetros. Para sus padres, de ideas liberales y trabajadores, era fundamental darles a todos sus hijos educación y ya que los mayores crecían, con mucho esfuerzo, decidieron que la madre se trasladara con los hijos a una casita en las afueras de Ensenada, el padre les visitaba con la frecuencia que el trabajo del campo le permitía. Así, Luly entraría a la Primaria “Segundo Ayuntamiento” en el turno de la mañana con sus hermanos menores y los mayores estudiarían en la Secundaria “Estatad No. 9”, que iniciaba labores a las dos de la tarde en el mismo edificio. Posteriormente, ingresó a la “Preparatoria Ensenada” que ahora pertenece al Colegio de Bachilleres. El deseo de sus padres se logró pues los ocho hijos estudiaron sus carreras. Las tres hermanas mayores se dedicaron al magisterio y fueron las primeras en entrar a trabajar para apoyar el estudio de los hermanos menores. Luly para contribuir a este esfuerzo familiar pidió una beca al “Club de Mujeres Profesionistas y de Negocios” de Ensenada. Cuando empezó a trabajar dando clases, les escribió para que le fuera interrumpida y los recursos se dieran a otras personas que lo necesitaran.

Con ese bagaje Luly llegó a la Ciudad de México a estudiar la Licenciatura en Matemáticas en la Facultad de Ciencias de la UNAM para después ingresar a la

Maestría en Estadística Aplicada del Colegio de Posgraduados. Por su carácter fuerte y decisión, concretó su maestría a pesar del ambiente académico que en esa institución imperaba, en la que había profesores que no veían bien que las mujeres se preparan y a las que sin el menor reparo menospreciaban, esa actitud hostil hizo que Luly buscara asesoría con el Dr. Alberto Castillo, a quien siempre mencionaba como su mentor y gran amigo. Varios años después, su pasión por la docencia la motivó a enriquecer su experiencia y formación y concluyó los estudios de la Especialidad en Docencia con Base en Competencias, la Maestría en Educación y el Doctorado en Evaluación Educativa en la Universidad Anáhuac.

Su ingreso a la UAM como Ayudante fue motivado por quienes fueran sus maestros en la Facultad de Ciencias, los doctores Alberto Ruiz Moncayo y Arturo Fregoso Urbina. Con el ímpetu de iniciar una nueva institución de educación superior basada en la estructura departamental, y el pronto inicio de clases, formó parte del grupo responsable de impartir las clases de matemáticas al alumnado que entraría a las licenciaturas de la División de Ciencias Sociales y Humanidades. Desde ese momento, aun con su juventud, comprendió la importancia que tendrían los cursos de matemáticas en la formación de los estudiantes en otras disciplinas. Inició desde ese entonces, una colaboración académica estrecha con el Dr. Adolfo Mir Araujo que se convertiría con el tiempo en una gran amistad, eso mismo sucedió con el Dr. Javier Vivaldo Lima.

A Luly le gustaba dar clase, continuamente diseñaba ejemplos y prácticas nuevas para sus alumnos, siempre llegaba puntual a los salones y durante años la vimos caminar por los pasillos jalando una maleta en la que llevaba su computadora y un cañón. Al diseñar sus clases, buscaba ejemplos adecuados y vigentes dependiendo del enfoque del curso, proponía clases interactivas, en las que los estudios de caso motivaran a sus estudiantes, promoviendo su reflexión y participación.

Con argumentos e insistencia ante las autoridades, hizo posible que al crecer el uso del cómputo en la Unidad, se contara con la paquetería de software estadístico para que las clases y las prácticas de los estudiantes se hicieran procesando bases de datos que permitieran mejorar el planteamiento de las hipótesis, su análisis y las conclusiones de los trabajos; tenía muy claro que nuestros egresados debían contar con las capacidades necesarias para poder insertarse con éxito en el mercado laboral.

Era una profesora exigente, a la vez de motivadora y muy franca, que reconocía el esfuerzo e interés de los estudiantes, correspondiéndoles imponiéndose más horas de dedicación, buscando siempre estrategias para mejorar el proceso de enseñanza y aprendizaje.

Durante toda su trayectoria docente, Luly impartió muchísimos cursos de estadística para las licenciaturas y la Maestría en Economía que ofrece la División de Ciencias Sociales y Humanidades, los cursos de diseño experimental y de bioestadística para las licenciaturas y el Posgrado de Biotecnología bajo la responsabilidad de la División de Ciencias Biológicas y de la Salud; obviamente también hizo lo propio para las licenciaturas de la División de Ciencias Básicas e Ingeniería. Los colegas del Departamento de Biotecnología le reconocen una participación fundamental en la formulación del plan y de los programas de estudios de su Posgrado, y en pro de la formación de sus egresados, al enseñarles a saber hacer diseños experimentales, cómo entenderlos y aplicarlos en su especialidad. Solo interrumpió la impartición de cursos para este posgrado durante el disfrute de sus periodos sabáticos.

Ella participó en muchísimas de las comisiones encargadas de crear, adecuar o modificar los programas de estudio de diseño de experimentos y de estadística para las licenciaturas de las tres divisiones de la Unidad. Durante muchos periodos, también se desempeñó como Coordinadora de los cursos que el Departamento de Matemáticas imparte para las licenciaturas de la División de Ciencias Sociales y Humanidades. En esa labor, siempre tejió puentes de comunicación entre los responsables de las diferentes dependencias. Su experiencia y conocimiento muchas veces fue convocada

para conciliar intereses académicos y administrativos que surgían de esta relación, aun cuando ya no fungiera como Coordinadora.

Por su labor docente, el Consejo Divisional de la División de Ciencias Básicas e Ingeniería, le otorgó el Premio a la Docencia en el año 2015. Al saber de su pronta e inesperada partida, muchos de los que fueron sus alumnos, reconocieron su entrega como profesora al escribir en las redes sociales su reconocimiento y aprecio en forma espontánea.

No solamente impartió cursos, ella se metió de lleno al muestreo y a las encuestas; en la Unidad Iztapalapa fue impartido varias veces el “Diplomado sobre encuestas y su realización” en el que Luly tuvo una gran participación. También impartió diferentes cursos de formación en técnicas de análisis cuantitativo para personal del Instituto Nacional de Evaluación Educativa y de la Unidad Cuajimalpa de la UAM, entre otras instituciones.

Luly estableció con muchos colegas una interacción académica formal en múltiples rincones de la Unidad y fuera de ella. Por su capacidad y su curiosidad por el saber, a partir de la puesta en marcha del Posgrado en Biotecnología en 1991, entabló con su profesorado y muchos de sus egresados, además de su amistad, colaboraciones que derivaron en la publicación de artículos de investigación en revistas indizadas. Así como un manual y un libro de texto para el diseño experimental. Este último lo escribió en coautoría con la Dra. Flor de María Cuervo López y el Dr. Alberto Castillo Morales. Además de los colegas de biotecnología, Luly colaboró en muy diversas investigaciones con el Dr. Javier Velázquez Moctezuma y su grupo de especialistas. Asimismo, trabajó con colegas del Departamento de Antropología de la Unidad Iztapalapa, de la Universidad de Aguascalientes, de la Universidad Autónoma de Baja California Sur, de la Universidad de Salamanca y de la Universidad Complutense de Madrid. En los últimos años, también lo hizo con el Dr. José Segovia Vila del CINVESTAV.

Los trabajos de investigación en los que colaboró se encuentran publicados en revistas tales como: *Livestock Research Development*, *Journal of Neuroscience Research*, *The World Journal of Biological Psychiatry*, *The Open Sleep Journal*, *Bull Environ Contam Toxicol*, *Brain and Development* (special section: *Developmental Neuropathology*), *Neuroscience & Medicine*, *Apoptosis*, *Experimental Cell Research*, *Sleep Medicine*, *Journal of Psychology & Psychotherapy*, *Environmental Technology*, *Gene Therapy*, *Cytotherapy* y *OncoTargets and Therapy*.

En 1996, colabora en uno de los primeros estudios sobre egresados en el mercado laboral que encabezó la Dra. Giovanna Valenti de la Unidad Xochimilco de la UAM. A partir de este primer proyecto, junto con el Dr. Castillo, inicia las bases para crear una metodología estadística que permite realizar diversos estudios en las instituciones de educación superior para conocer las condiciones de inicio de las personas que ingresan a estas. Esta metodología la aplican junto con el Dr. Mir en diversos proyectos encomendados por las universidades tecnológicas.

Posteriormente, para el Consejo Regional del área Metropolitana (CRAM) de la ANUIES, Luly gestiona nuevos proyectos enfocados a conocer las competencias en inglés como una segunda lengua; los conocimientos y habilidades en matemáticas; la comprensión auditiva, de lectura y la expresión escrita del alumnado. Todos incluyen la exploración de las variables socio-familiares y de la educación preuniversitaria. Estos proyectos ofrecen información que apoyan a las instituciones para generar estrategias que hagan más eficiente el tránsito curricular de los estudiantes en la educación superior. Para realizarse, cada uno de estos proyectos requirieron de un trabajo continuo durante dos o tres años y de la coordinación e integración de equipos de académicos de diversas especialidades y de varias instituciones.

El último proyecto que encabezó fue el dedicado al estudio de las Tecnologías de Información y de Comunicación que inició por encargo del CRAM de la ANUIES en

el año 2018 el cual estaba en una etapa muy avanzada al momento de su partida, el que ya no concluirá como ella lo hubiera deseado.

Por todo este trabajo, se le reconoce como una especialista en el análisis sociológico de los datos y los sociólogos afirman que su trabajo sentó directrices teóricas y metodológicas en el campo de la investigación educativa en nuestro país.

Su producción como coautora y/o editora en libros fue vasta, todos ellos producto de estudios y trabajo de varios años, entre ellos encontramos: “Los egresados de la UAM en el mercado de trabajo. Investigación evaluativa sobre la calidad de la oferta de servicios educativos”, “Metodología estadística para la realización de estudios de egresados en una institución de educación superior”, “Factores que Influyen en la decisión de estudiar en las Universidades Tecnológicas”, “Los Egresados de las Universidades Tecnológicas. Formación profesional y situación laboral”, “Competencia Lingüística en Inglés de Estudiantes de Primer Ingreso a Instituciones de Educación Superior del área Metropolitana de la Ciudad de México”, “Conocimientos y habilidades en matemáticas de los estudiantes de primer ingreso a las instituciones de educación superior del área metropolitana de la Ciudad de México”, “Habilidades lingüísticas de los estudiantes de primer ingreso a las Instituciones de Educación Superior. Área Metropolitana de la Ciudad de México”, “Ejercicios para el taller de Diseño experimental”, libros editados por la UAM, la UAMI, la ANUIES y por la Coordinación de Universidades Tecnológicas de la Subsecretaría de Educación Superior de la SEP. Entre sus coautores menciono a los doctores Adolfo Mir Araujo, Javier Vivaldo Lima, Giovanna Valenti Nigrini, Irma Munguía Zatarain, Manuel Jorge González Montesinos, Manuel Meda Vidal, Carmen López Laiseca y Javier Alfaro Pastor, Shirley Bromberg Silvertein, María José Arroyo Paniagua, y siempre trabajando codo a codo con ella, el Dr. Alberto Castillo.

Todos los que tuvimos de una u otra forma la posibilidad de trabajar con ella en algún cometido, la recordamos como una profesional rigurosa, buscadora de aplicaciones prácticas y metodológicas. En cada uno de los proyectos que emprendía, ponía toda su energía, compromiso y conocimientos, ya fuera un estudio de egresados, o sobre evaluación educativa, una investigación en fisiología, neurociencias o biotecnología, y obviamente estudiaba y aprendía con sus colegas. Para ella era necesario comprender los diversos problemas que abordan las diferentes líneas de investigación a las que enriquecía y fundamentaba con su aporte estadístico.

Fuera del trabajo, a Luly le gustaba hacer muy diversas actividades, amaba pintar cuadros de flores, con colores vivos, contrastantes; era una lectora apasionada, devoraba los libros y uno podía pasar horas en una conversación hablando de ellos con ella; si uno quería ir al cine, la persona adecuada para dar recomendaciones de la película que merecía ser vista era Luly, pues era una cinéfila empedernida al igual que amaba al teatro. Pocos sabrán que hace muchos años se presentó en el Teatro del Fuego Nuevo de nuestra querida Unidad Iztapalapa, la obra “Don Juan y las Matemáticas”. Todos los actores eran profesores y estudiantes del Departamento de Matemáticas, entre los cuales se encontraba Luly plena de ánimo y alegría; esta actividad fue parte del programa de la Primera Semana de las Matemáticas que organiza anualmente el Departamento. Fue una viajera incansable, se sorprendía y apreciaba las diferentes culturas, gustaba de la buena comida, ya fuera francesa, italiana, china y por supuesto la mexicana, la compartía en un buen restaurante o la cocinaba, pues tenía muy buen sazón y era algo que disfrutaba hacer con algunas amigas y amigos a los que siempre abrió las puertas de su casa con gran generosidad.

En la Ciudad de México hizo una extensión de la familia con muchas personas, los valores familiares inculcados fueron compartidos en innumerables charlas, siempre reconoció lo que su familia le entregó. Todos sus amigos fuimos testigos del inmenso cariño que la unió siempre a su hijo Rodrigo, quien con su querida Fortuna, le dieron

la gran felicidad de tener nietos y a los que adoraba. Igualmente, pudimos ver lo felices que fueron ella y su compañero, nuestro amigo José.

Somos muchos las compañeras y compañeros del trabajo que encontramos en ella a una colega profesional, generosa y solidaria con la que compartimos nuestro trabajo diario y gran parte de nuestras vidas en la UAM. Todos, su familia, sus amigos y colegas fuimos sorprendidos por su pronta partida. Al reunirnos para recordar su vida y celebrar el haberla conocido y gozado de su cariño y amistad, constatamos lo que siempre sentimos respecto a ella, que se entregó a los que quería, dándose toda, sin reservas.

Dirección de la autora:

Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina,
Alcaldía Iztapalapa, C.P. 09340, CDMX, México.
e-mail: mja@xanum.uam.mx



IN MEMORIAM: FRANCISCO HUGO MARTÍNEZ ORTÍZ (1947–2021)

El profesor Francisco Hugo Martínez Ortíz o simplemente Hugo, como le llamábamos sus colegas del Departamento de Matemáticas de la UAM-Iztapalapa, nació en Teposcolula, Oaxaca, el 14 de julio de 1947. Obtuvo el título de Matemático en la Facultad de Ciencias de la UNAM el 14 de noviembre de 1970. Armando Santamaría Borja, exprofesor del departamento y compañero de Hugo desde la preparatoria recuerda: *“Conocí a Hugo en marzo de 1963, cuando ambos ingresamos al turno vespertino del plantel 7 de la Escuela Nacional Preparatoria de la UNAM, dos años después ingresamos a la Facultad de Ciencias y debo decir que Hugo fué el único alumno de su generación que completó sus estudios en los cuatro años que establece el plan de estudios. Por esta razón se le otorgó una beca y una plaza en el Instituto de Matemáticas Aplicadas y Sistemas (IIMAS) y así empezó a dar clases en la UNAM. Era muy disciplinado, me decía que todos los días se levantaba a las cinco de la mañana para estudiar dos horas y después se iba a la universidad. Era un estudiante brillante y excelente ser humano. En 1974, el Dr. Alberto Ruiz Moncayo, primer Jefe del Departamento de Matemáticas, lo invitó a formar parte del personal académico de la recién fundada Universidad Autónoma Metropolitana, Hugo aceptó y fué miembro de la planta de profesores fundadores de la Unidad Iztapalapa de la UAM. Era un profesor que impartía una excelente docencia, lo cual era reconocido y muy bien valorado por sus alumnos. Fué invitado por el Departamento de Irrigación de la Universidad Autónoma Chapingo como profesor de asignatura y fué distinguido en varias ocasiones por los propios alumnos como su mejor profesor de la generación. Destacó por sus notables cualidades didácticas, su excelente calidad humana y, sin duda alguna, por su talento matemático, al grado tal que con frecuencia llegó a decirse que él era un Doctor, sin el grado correspondiente.”*

Por su parte, Gabriel López Garza agrega: *“Hugo no sólo fué mi vecino de cubículo y amigo, sino que escribimos un libro en colaboración. De él puedo decir que no conozco a ningún otro profesor que diera sus clases con mejor ánimo y entusiasmo (y dio clases por más de cuarenta años). También puedo decir con certeza que su saber en todos los cursos que dictó siempre fué profundo y sabio, muy por encima del conocimiento que dan los grados académicos. Su saber era el que sólo se logra con pasión por las matemáticas y con el ejercicio del pensamiento por largos años de entrega por el sólo placer de saber y entender más cada día. Como humano era un ser bondadoso y amable, de quien nunca percibí sino palabras y opiniones positivas hacia los demás y de los demás hacia él. Lamento mucho su pérdida y extrañaré su entusiasmo y energía para enfrentar las labores cotidianas”.*

Antes de ingresar como profesor de la UAM, Hugo aprobó todos los cursos del programa de Maestría en Matemáticas de la UNAM y fué investigador del Centro de Investigación en Matemáticas Aplicadas y Sistemas (CIMAS), actualmente IIMAS, y del Instituto de Geofísica, ambos de la UNAM. Ya como miembro del departamento de matemáticas, fué coordinador de los cursos de apoyo a la División de Ciencias Sociales y Humanidades durante ocho años. Sus intereses matemáticos se concentraron en Ecuaciones Diferenciales Ordinarias y Parciales. Publicó varios trabajos de docencia y divulgación (monografías, notas de curso, artículos de docencia y un libro), entre ellos: Aplicación de las series de Fourier a la solución de problemas con valores en la frontera y/o iniciales (Instituto de Geofísica–UNAM), Teoría de Sturm-Liouville (IIMAS–UNAM), Sobre el Teorema Espectral (IIMAS–UNAM), Notas sobre el “Método de

Ritz” (IIMAS–UNAM), Ecuaciones Diferenciales Parciales (libro en coautoría con Gabriel López Garza, UAM–2014), Método de la función de Green para ecuaciones diferenciales ordinarias lineales de segundo orden (artículo de divulgación en Mixba’al, Revista Metropolitana de Matemáticas, No 1, Vol. VI, 2015, 53-73).

Las apreciaciones de Armando y Gabriel sobre la calidad humana de Hugo y su compromiso con su trabajo como profesor, son compartidas por todos sus colegas del Departamento de Matemáticas, algunos de los cuales fueron sus alumnos. Por su calidad didáctica, sus cursos fueron seleccionados y video grabados como parte del material didáctico de un proyecto divisional de docencia que esperamos no quede en el olvido. Extrañaremos su buen humor, su modestia y su pasión por las matemáticas, en particular su compromiso con la enseñanza, que son un ejemplo a seguir por las nuevas generaciones.

Comité Editorial



TRES TEOREMAS SOBRE CARDINALES MEDIBLES

FRANKLIN GALINDO

RESUMEN. El estudio de los «cardinales grandes» es uno de los principales temas de investigación de la teoría de conjuntos y de la teoría de modelos que ha contribuido con el desarrollo de dichas disciplinas. Existe una gran variedad de tales cardinales, por ejemplo cardinales inaccesibles, débilmente compactos, Ramsey, medibles, supercompactos, etc. Tres valiosos teoremas clásicos sobre cardinales medibles son los siguientes: (i) *compacidad débil*, (ii) *Si κ es un cardinal medible, entonces κ es un cardinal inaccesible y existen κ cardinales inaccesibles menores que κ* , y (iii) *Si existe un cardinal medible, entonces el axioma de constructibilidad ($V=L$) es falso*. El objetivo de este artículo es presentar una demostración de cada uno de estos tres teoremas en el contexto de la teoría de modelos usando ideas del texto [3]. Tales demostraciones tienen en común el uso del método de construcción de modelos llamado ultraproductos, de lógicas infinitarias o fragmentos de la lógica de segundo orden y del axioma de elección. Cardinales grandes y/o ultraproductos son importantes en teoría de conjuntos, teoría de modelos, análisis matemático, teoría de la medida, probabilidades, topología, análisis funcional, física, teoría de números, finanzas, etc.

1. INTRODUCCIÓN

El estudio de los «cardinales grandes» es uno de los principales temas de investigación de la teoría de conjuntos y de la teoría de modelos que ha contribuido con el desarrollo de dichas disciplinas. Existe una gran variedad de tales cardinales, por ejemplo cardinales inaccesibles, débilmente compactos, Ramsey, medibles, supercompactos, etc. Tres valiosos teoremas clásicos sobre cardinales medibles son los siguientes (ver [3,6,16]): (i) *compacidad débil*, (ii) *Si κ es un cardinal medible, entonces κ es un cardinal inaccesible y existen κ cardinales inaccesibles menores que κ* , y (iii) *Si existe un cardinal medible, entonces el axioma de constructibilidad ($V=L$) es falso*. El objetivo de este artículo es presentar una demostración de cada uno de estos tres teoremas en el contexto de la teoría de modelos usando ideas del texto «Model Theory» de Chang y Keisler [3]. Tales demostraciones tienen en común el uso del método de construcción de modelos llamado ultraproductos, de lógicas infinitarias o fragmentos de la lógica de segundo orden y del axioma de elección. Vale la pena resaltar que los cardinales grandes y/o los ultraproductos son importantes en teoría de conjuntos, teoría de modelos, análisis matemático, teoría de la medida, probabilidades, topología, análisis funcional, física, teoría de números, finanzas, etc. (ver [2,3,6,13,15,16,19])

Existen otras demostraciones interesantes de los teoremas que se demostrarán en este artículo que también usan ultraproductos, pero que son distintas a las que aquí se presentarán. Algunas de ellas se pueden encontrar en los siguientes textos o notas [6,13,15,16].

Se ofrecerán en este artículo dos referencias de problemas abiertos sobre cardinales medibles, cardinales grandes y ultraproductos.

El orden expositivo del contenido de este artículo es el siguiente: En la siguiente sección (2) se presentarán algunos conceptos preliminares y el teorema fundamental de ultraproductos (teorema de Lós). En la sección 3 se presentará el primer teorema sobre cardinales medibles (i). En la sección 4 se presentará el segundo teorema sobre cardinales medibles (ii). Y en la última sección (5) se presentará el tercer teorema sobre

cardinales medibles (iii) y al final de la sección se ofrecerán dos referencias de problemas abiertos sobre sobre cardinales medibles, cardinales grandes y ultraproductos.

2. CONCEPTOS PRELIMINARES Y EL TEOREMA FUNDAMENTAL DE ULTRAPRODUCTOS

En esta sección se presentará el método de construcción de modelos llamado ultraproductos en el contexto de la teoría de modelos siguiendo principalmente el texto [3], cap.4]. Esto significa que se usará informalmente la teoría axiomática de conjuntos de Zermelo-Fraenkel con el axioma de elección (ZFC) tal como es presentada y desarrollada en los textos [5,8,12,13,15,17], entre otros. Específicamente se presentará esta sección siguiendo el artículo [10] (el cual está en la web), en dicho artículo se describe a los ultraproductos, a una demostración del teorema fundamental de los ultraproductos, y a una prueba del teorema de compacidad usando ultraproductos. El método de ultraproductos se originó con Skolem en 1930 y luego fue desarrollado por Loś en 1955 (ver [3,13,18]).

La presentación se realizará en el siguiente orden expositivo: En las subsecciones 2.1 y 2.2 se ofrecerán algunos conceptos y resultados necesarios para definir los ultraproductos, por ejemplo «filtro», «ultrafiltro», «lema de Zorn», «teorema del ultrafiltro», etc. Y se mencionarán sin definir otros conceptos y resultados de la teoría de modelos que son necesarios para desarrollar este artículo pero que son básicos y se presupone que el lector tiene este conocimiento, por ejemplo la sintaxis y la semántica de la lógica de primer orden y algunas relaciones entre estructuras. Y en la subsección 2.3 se definirán los ultraproductos y se formulará el teorema fundamental de los ultraproductos siguiendo ideas principalmente de [3,13,16,21].

2.1. Filtros, ultrafiltros, lema de Zorn, teorema del ultrafiltro.

Definición 1. ■ Un *filtro* sobre un conjunto no vacío S es una colección \mathcal{F} de subconjuntos de S tal que:

- (i) $S \in \mathcal{F}$ y $\emptyset \notin \mathcal{F}$.
- (ii) Si $X \in \mathcal{F}$ y $Y \in \mathcal{F}$, entonces $X \cap Y \in \mathcal{F}$.
- (iii) Si $X \in \mathcal{F}$ y $X \subseteq Y$, entonces $Y \in \mathcal{F}$.

- Sea \mathcal{F} un filtro sobre un conjunto S . \mathcal{F} es un *ultrafiltro* si para cualquier $X \subseteq S$ se cumple que:

$$X \in \mathcal{F} \leftrightarrow S \setminus X \notin \mathcal{F}.$$

- Sea \mathcal{F} un ultrafiltro sobre un conjunto S . \mathcal{F} es *no principal* si y sólo si $\forall i \in S (\{i\} \notin \mathcal{F})$.

Ejemplos de filtros (ver [3,13]): (1) *filtro trivial*: $\mathcal{F} = \{S\}$. (2) Para cada $B \subseteq S$, $B \neq \emptyset$, el filtro $\mathcal{F}_B = \{Z \subseteq S : B \subseteq Z\}$ se llama *filtro principal* generado por B . Para $B = \{a\} \subseteq S$, \mathcal{F}_B se escribe \mathcal{F}_a , $\mathcal{F}_a = \{Z \subseteq S : a \in Z\}$. Nótese que \mathcal{F}_a es un ultrafiltro principal. (3) Sea S un conjunto infinito, el filtro $\mathcal{F} = \{X \in P(S) : |S \setminus X| < \aleph_0\}$ se llama *filtro de Fréchet*. Nótese que el filtro de Fréchet no es principal. Dado cualquier conjunto infinito S siempre se puede construir un filtro no principal sobre S , que extiende al filtro de Fréchet sobre S usando la propiedad de intersección finita, tal como lo afirma el teorema que viene a continuación (Teorema 2). Pregunta: Ya se demostró anteriormente que existen ultrafiltros principales, mediante un ejemplo, \mathcal{F}_a . Pero, ¿Existen ultrafiltros no principales?. La existencia de tales entidades matemáticas sólo se puede garantizar usando el lema de Zorn, no hay otra manera (ver [13], p. 75)]. A continuación se presentará (después del Teorema 2) el teorema del ultrafiltro, el cual permitirá contar con tales entidades (ultrafiltros no principales) las cuales son fundamentales para la investigación matemática, por ejemplo para la construcción del cuerpo ordenado y no arquimediano de los hiper-reales del análisis no estándar de Robinson [2], y jugarán un importante papel en la definición de los cardinales medibles.

El teorema del ultrafiltro se prueba a partir del lema de Zorn, lema que se formulará también en este trabajo.

Una familia G de conjuntos tiene la *propiedad de intersección finita* si para cualquier conjunto finito $H = \{X_1, \dots, X_n\} \subseteq G$ se cumple que $H_1 \cap \dots \cap H_n \neq \emptyset$.

- TEOREMA 2. 1. Si Δ es una familia de filtros sobre S , entonces $\bigcap \Delta$ es un filtro sobre S .
2. Si Ω es una \subseteq -cadena de filtros sobre S (es decir, $\forall X, Y \in \Omega (X \subseteq Y \text{ o } Y \subseteq X)$), entonces $\bigcup \Omega$ es un filtro sobre S .
3. Si $G \subseteq P(S)$ tiene la propiedad de intersección finita, entonces existe un filtro \mathcal{F} tal $G \subseteq \mathcal{F}$ ($\mathcal{F} = \{X \subseteq S : \exists Z_1, \dots, Z_n \in G (Z_1 \cap \dots \cap Z_n \subseteq X)\}$).

Una prueba de este resultado puede encontrarse (entre otros) en [13, p.74] y [3, p. 212].

TEOREMA 3 (Teorema del ultrafiltro(Tarski)). *Todo filtro se puede extender a un ultrafiltro.*

Como ya se dijo antes la demostración del teorema del ultrafiltro requiere del lema de Zorn, una versión de dicha prueba puede encontrarse (entre otros) en [13, p. 75] y [3, p. 214]. A continuación se enuncia el lema de Zorn después de presentar unas definiciones previas:

Definición 4. Sea A un conjunto y R una relación binaria en A (es decir, $R \subseteq A \times A$)

1. R es *reflexiva* si y sólo si $\forall x \in A (xRx)$
2. R es *simétrica* si y sólo si $\forall x, y \in A (xRy \rightarrow yRx)$
3. R es *transitiva* si y sólo si $\forall x, y, z \in A (xRy \wedge yRz \rightarrow xRz)$
4. R es *antisimétrica* si y sólo si $\forall x, y \in A (xRy \wedge yRx \rightarrow x = y)$.
5. R es una relación de *equivalencia* si R es una relación reflexiva, simétrica y transitiva.

Definición 5. 1. Un *orden parcial* es un par (P, \leq) donde P es un conjunto no vacío y \leq es una relación en P que es reflexiva, antisimétrica y transitiva.

2. Dado un orden parcial (P, \leq) se dice que $p < q \leftrightarrow p \leq q \wedge p \neq q$.
3. Sean (P, R) un orden parcial y $D \subseteq P$. $x \in P$ es un elemento *mínimo* (*máximo*) de D si $x \in D \wedge$ no existe ningún $y \in D$ tal que $y \neq x \wedge yRx$ (xRy). x es una *cota inferior* (*superior*) de D si $\forall y \in D (xRy \vee y = x)$ ($yRx \vee y = x$). x es un *ínfimo* (*supremo*) de D si x es cota inferior (*superior*) de $D \wedge$ para todo $y \in P$, si y es una cota inferior (*superior*) de D , entonces $yRx \vee y = x$ ($xRy \vee y = x$). x es un *menor* (*mayor*) elemento de D si $x \in D \wedge \forall y \in D (xRy \vee y = x)$ ($yRx \vee y = x$).
4. Sea (P, R) un orden parcial. (P, R) es un *orden total* (o lineal) si la relación R satisface la propiedad de tricotomía: $\forall x, y \in P (xRy \vee yRx \vee x = y)$.
5. Sea (P, R) un orden parcial. (P, R) es un *buen orden* si para todo $X \subseteq P$ se cumple que: Si $X \neq \emptyset$, entonces X tiene un menor elemento. (Nótese que todo conjunto bien ordenado es un conjunto totalmente ordenado).

(Un orden parcial u orden total o buen orden (P, \leq) a veces se denotará por P)

LEMA 6 (Lema de Zorn). *Sea (K, S) un conjunto parcialmente ordenado tal que cada $X \subseteq K$ totalmente ordenado tiene una cota superior en K . Entonces K tiene un elemento máximo.*

Es conocido que el lema de Zorn es equivalente al axioma de elección. Donde el axioma de elección es la siguiente sentencia: *Todo conjunto tiene una función selectora.* (Dado un conjunto Z se dice que la función f es una función de elección - o una función selectora - para Z si el $\text{dom}(f) = Z - \{\emptyset\}$ y para todo $W \in \text{dom}(f)$, se tiene que $f(W) \in W$). Una prueba de la equivalencia puede encontrarse (entre otros) en [5,

p. 83-85] y [8, p. 151-153]. También es conocido que el lema de Zorn es equivalente al principio del buen orden (*Todo conjunto se puede bien ordenar*). Una prueba de tal equivalencia puede encontrarse (entre otros) en [5, p. 82-85] y [8, p. 151-153, 196-197].

2.2. Conceptos básicos de la teoría de modelos: sintaxis y semántica de la lógica de primer orden, relaciones entre estructuras. Los conceptos básicos de la teoría de modelos, como por ejemplo la sintaxis y la semántica de la lógica de primer orden y las relaciones básicas entre estructuras, que se usarán en este artículo se suponen conocidas por parte del lector, no se describirán, son las mismas nociones que están formuladas y explicadas en los textos contemporáneos de lógica matemática o de teoría de modelos, por ejemplo en [3,4,7,9,19,21]. Un resumen de tales conceptos puede conseguirse en [10], entre otros. Tales nociones son: «lenguajes de primer orden», «estructuras», «formalización de un lenguaje (en primer orden)», «fórmula», «sentencia», «satisfacción y verdad en una estructura», «modelo», «contradicción», «consecuencia lógica», relaciones entre estructuras: «isomorfismo», «subestructura», «inmersión», «submodelo elemental», «inmersión elemental», «equivalencia elemental», etc.

2.3. El teorema fundamental de ultraproductos. ¿Cómo se contruyen las estructuras (modelos) llamadas ultraproductos? La construcción se realiza usando (entre otros) a los ultrafiltros, en este artículo se presenta tal construcción siguiendo principalmente los textos [3, cap. 4], [21, p. 129-135] y [13, p. 158-161].

Supongamos que I es un subconjunto no vacío, D es un filtro sobre I y que para cada $i \in I$, A_i es un subconjunto no vacío. Entonces se considera el producto cartesiano de los conjuntos A_i , es decir:

$$C = \prod_{i \in I} A_i = \{f : f : I \longrightarrow \cup_{i \in I} A_i \wedge \forall i \in I (f(i) \in A_i)\}.$$

Ahora se define una relación de equivalencia \sim en $\prod_{i \in I} A_i$ de la siguiente manera:

$$f \sim g \iff \{i \in I : f(i) = g(i)\} \in D.$$

Se cumple que la relación \sim es reflexiva, simétrica y transitiva, por lo tanto \sim es una relación de equivalencia. Entonces se considera el conjunto cociente de $\prod_{i \in I} A_i$ determinado por \sim (llamado *producto reducido de los A_i modulo D*):

$$\prod_{i \in I} A_i / \sim = \{[f] : f \in \prod_{i \in I} A_i\}$$

Se denotará al conjunto cociente $\prod_{i \in I} A_i / \sim$ por $\prod_D A_i$ y la clase de equivalencia $[f]$ se denotará por f_D ($\forall f \in \prod_{i \in I} A_i$). Cuando D es un ultrafiltro $\prod_D A_i$ es llamado un *ultraproducto*. Y en el caso de que todos los conjuntos A_i sean iguales, digamos, $A_i = A$ ($\forall i \in I$), $\prod_D A$ es llamada la *ultrapotencia de A modulo D* .

Ahora se definirá el *producto reducido de modelos* para un lenguaje \mathcal{L} fijo:

(\mathcal{L} es un conjunto, de cualquier cardinalidad, de símbolos de relación (R_0, R_1, \dots), símbolos de función (F_0, F_1, \dots) y de símbolos constantes (c_0, c_1, \dots). Donde el subconjunto de \mathcal{L} de símbolos de relación puede ser vacío, el subconjunto de \mathcal{L} de símbolos de función puede ser vacío, y el subconjunto de \mathcal{L} de símbolos constantes puede ser vacío. Cada símbolo de relación y cada símbolo de función tiene asociado un número natural ≥ 1 , su número de argumentos, de modo que tenemos símbolos de relación o de función unarios, binarios, ..., n-arios, etc.)

Sea I un conjunto no vacío, D un filtro sobre I y para cada $i \in I$ sea \mathfrak{A}_i una estructura para un lenguaje \mathcal{L} . Supongamos que para cada símbolo relacional P de \mathcal{L} la interpretación de P en \mathfrak{A}_i es R_i , los símbolos de función F son interpretados en \mathfrak{A}_i por G_i y los símbolos constantes c son interpretados en \mathfrak{A}_i por a_i .

El *producto reducido* $\prod_D \mathfrak{A}_i$ es una estructura para \mathcal{L} definida de la siguiente manera:

(i) El universo de $\prod_D \mathfrak{A}_i$ es $\prod_D A_i$.

(ii) Sea P un símbolo de relación n -ario de \mathcal{L} . La interpretación de P en $\prod_D \mathfrak{A}_i$ es la relación n -ária S definida de la siguiente manera:

$$S(f_D^1, \dots, f_D^n) \longleftrightarrow \{i \in I : R_i(f^1(i), \dots, f^n(i))\} \in D.$$

(iii) Sea F un símbolo de función n -ario de \mathcal{L} . Entonces F es interpretado en $\prod_D \mathfrak{A}_i$ por la función n -ária H definida de la siguiente manera:

$$H(f_D^1, \dots, f_D^n) = \langle G_i(f^1(i), \dots, f^n(i)) : i \in I \rangle_D$$

(iv) Sea c una constante de \mathcal{L} . Entonces la interpretación de c en $\prod_D \mathfrak{A}_i$ es un $b \in \prod_D \mathfrak{A}_i$ que se define como sigue:

$$b = \langle a_i : i \in I \rangle_D.$$

Se cumple que las definiciones realizadas anteriormente de $S(f_D^1, \dots, f_D^n)$ y $H(f_D^1, \dots, f_D^n)$ dependen solamente de las clases de equivalencia y no de sus representantes f^1, \dots, f^n . Es decir, $f^1 \sim g^1, \dots, f^n \sim g^n$, entonces :

$$\{i \in I : R_i(f^1(i), \dots, f^n(i))\} \in D \leftrightarrow \{i \in I : R_i(g^1(i), \dots, g^n(i))\} \in D$$

y

$$\langle G_i(f^1(i), \dots, f^n(i)) : i \in I \rangle_D = \langle G_i(g^1(i), \dots, g^n(i)) : i \in I \rangle_D.$$

Ahora de formulará el teorema fundamental de ultraproductos (Loś, 1955, [Lo]) y dos corolarios del mismo:

TEOREMA 7 (Teorema fundamental de ultraproductos (Loś)). *Sea \mathfrak{B} el ultraproducto $\prod_D \mathfrak{A}_i$, donde I es el conjunto de índices de los \mathfrak{A}_i . Entonces :*

(i) *Para cada término $t(x_1, \dots, x_n)$ de \mathcal{L} y elementos $f_D^1, \dots, f_D^n \in B = \prod_D A_i$ se tiene que:*

$$t_{\mathfrak{B}}[f_D^1, \dots, f_D^n] = \langle t_{\mathfrak{A}_i}[f^1(i), \dots, f^n(i)] : i \in I \rangle_D$$

(ii) *Dada una fórmula $\varphi(x_1, \dots, x_n)$ de \mathcal{L} y $f_D^1, \dots, f_D^n \in B$ se tiene que:*

$$\mathfrak{B} \models \varphi[f_D^1, \dots, f_D^n] \leftrightarrow \{i \in I : \mathfrak{A}_i \models \varphi[f^1(i), \dots, f^n(i)]\} \in D$$

(iii) *Para cada sentencia φ de \mathcal{L} se tiene que:*

$$\mathfrak{B} \models \varphi \leftrightarrow \{i \in I : \mathfrak{A}_i \models \varphi\} \in D.$$

(Inuitivamente (iii) afirma que φ es verdadera en el ultraproducto $\mathfrak{B} = \prod_D \mathfrak{A}_i$ si y sólo si φ es verdadera en “casi todos” los factores \mathfrak{A}_i de \mathfrak{B})

Una prueba de este teorema puede encontrarse en [3,10,13,21], entre otros.

COROLARIO 8. *Sea \mathfrak{A} una estructura para un lenguaje \mathcal{L} y sea $\prod_D \mathfrak{A}$ una ultrapotencia de \mathfrak{A} . Entonces $\mathfrak{A} \equiv \prod_D \mathfrak{A}$.*

Una prueba de este corolario puede encontrarse en [21], entre otros.

Para enunciar el segundo corolario se requiere de una definición previa: Sea I un conjunto no vacío, D un filtro sobre I y \mathfrak{A} una estructura para un lenguaje \mathcal{L} . La *inmersión natural* de \mathfrak{A} dentro de $\prod_D \mathfrak{A}$ es la función $d : \mathfrak{A} \rightarrow \prod_D \mathfrak{A}$ definida como sigue: $\forall a \in A (d(a) = \langle a : i \in I \rangle_D)$. El rango de d es denotado por $d(A)$ y la restricción de $\prod_D \mathfrak{A}$ a $d(A)$ es denotada por $d(\mathfrak{A})$.

COROLARIO 9. *Sea \mathfrak{A} una estructura para un lenguaje \mathcal{L} y D un ultrafiltro. Entonces la inmersión natural de \mathfrak{A} dentro de la ultrapotencia $\prod_D \mathfrak{A}$ es una inmersión elemental.*

Una prueba de este corolario puede encontrarse en [3,10,13,21], entre otros.

3. PRIMER TEOREMA SOBRE CARDINALES MEDIBLES: COMPACIDAD DÉBIL

Se iniciará esta sección presentando la definición de «cofinalidad», «cardinal regular», «cardinal inaccesible», «cardinal medible» y «lógicas infinitarias», las cuales son necesarias para desarrollar la misma.

Sea α un ordinal límite. Decimos que $\beta < \alpha$ es *cofinal* con α si existe una función creciente $f : \beta \rightarrow \alpha$ tal que para todo $\xi < \alpha$, existe un $\delta < \beta$ tal $f(\delta) \geq \xi$ (es decir, la imagen de f es no acotada en α). Dado α , la *cofinalidad* de α , $\text{cof}(\alpha)$, es el menor ordinal cofinal con α . Con respecto a la cofinalidad se cumple lo siguiente: $\text{cof}(\alpha)$ es el menor cardinal β tal que existe una partición de α en β pedazos cada uno de los cuales tiene cardinalidad estrictamente menor que α . Un cardinal infinito κ es *regular* si es igual a su cofinalidad. Decimos que es *singular* en caso contrario ($\text{cof}(\alpha) < \alpha$). Un cardinal κ es un cardinal *límite fuerte* si para todo cardinal $\theta < \kappa$ se tiene que $2^\theta < \kappa$. Un cardinal $\kappa > \aleph_0$ es *inaccesible* si es regular y límite fuerte (Nótese que si se quita la condición de que $\kappa > \aleph_0$ se tiene que \aleph_0 es un cardinal inaccesible).

Sea α un cardinal infinito. Un filtro D sobre I se llama α -*completo* si y sólo si: $X \subseteq D$ y $|X| < \alpha$ implica $\bigcap X \in D$. Un cardinal $\alpha > \aleph_0$ se dice que es *medible* si y sólo si existe un ultrafiltro no principal y α -completo sobre α (Nótese que si se quita la condición de que $\kappa > \aleph_0$ se tiene que \aleph_0 es un cardinal medible).

Se definen ahora los lenguajes infinitarios $\mathcal{L}_{\kappa\kappa}$, donde κ es un cardinal mayor o igual que \aleph_0 , siguiendo a [1,3,5,9]:

Considérese un lenguaje de primer orden \mathcal{L} (conjunto de símbolos relacionales, funcionales y constantes) y también considérese las reglas que se usan para formalizar el mismo, es decir, para construir las fórmulas del lenguaje \mathcal{L} . A la lista de símbolos lógicos agréguese los siguientes nuevos símbolos: \bigwedge (conjuntor infinito) y \bigvee (disyuntor infinito), y sustitúyase la lista numerable de variables (que se usa normalmente) por una lista de variables de cardinal κ ($\text{VAR}_{\mathcal{L}_{\kappa\kappa}}$). Para construir el conjunto de fórmulas de $\mathcal{L}_{\kappa\kappa}$ se utilizan las mismas reglas que se usan para construir las fórmulas de lenguaje \mathcal{L} (ver alguna definición usual), más las siguientes dos nuevas reglas:

($\odot 1$) Si Φ es un conjunto de fórmulas de $\mathcal{L}_{\kappa\kappa}$ tal que $|\Phi| < \kappa$, entonces $\bigwedge \Phi$ y $\bigvee \Phi$ también son formulas de $\mathcal{L}_{\kappa\kappa}$.

($\odot 2$) Si φ es una fórmula de $\mathcal{L}_{\kappa\kappa}$ y V es un conjunto de variables tal que $|V| < \kappa$, entonces $(\forall V)\Phi$ y $(\exists V)\Phi$ también son fórmulas de $\mathcal{L}_{\kappa\kappa}$.

Semántica de $\mathcal{L}_{\kappa\kappa}$: Una estructura para evaluar las formulas del lenguaje $\mathcal{L}_{\kappa\kappa}$ es igual que una estructura para el lenguaje \mathcal{L} , pues tiene los mismos símbolos no lógicos. Sea ϕ una fórmula de $\mathcal{L}_{\kappa\kappa}$, \mathfrak{C} una estructura para \mathcal{L} y $s : \text{VAR}_{\mathcal{L}_{\kappa\kappa}} \rightarrow C$. La definición de $\mathfrak{C} \models \phi[s]$ (s satisface a ϕ en \mathfrak{C}) es igual que la definición usual que se hace para \mathcal{L} , más las siguientes reglas correspondientes a los nuevas fórmulas construídas mediante ($\odot 1$) y ($\odot 2$):

($\odot 1.1$) $\mathfrak{C} \models \bigwedge \Phi[s] \iff \mathfrak{C} \models \varphi[s]$, para toda $\varphi \in \Phi$.

($\odot 1.2$) $\mathfrak{C} \models \bigvee \Phi[s] \iff \mathfrak{C} \models \varphi[s]$, para alguna $\varphi \in \Phi$.

($\odot 2.1$) $\mathfrak{C} \models ((\forall V)\chi)[s] \iff \mathfrak{C} \models \chi[s']$ para toda $s' : \text{VAR}_{\mathcal{L}_{\kappa\kappa}} \rightarrow C$ que difiere de s a lo sumo en los valores de las variables de V .

($\odot 2.2$) $\mathfrak{C} \models ((\exists V)\chi)[s] \iff \mathfrak{C} \models \chi[s']$ para alguna $s' : \text{VAR}_{\mathcal{L}_{\kappa\kappa}} \rightarrow C$ que difiere de s a lo sumo en los valores de las variables de V .

Nótese que si $\kappa = \aleph_0$, entonces $\mathcal{L}_{\aleph_0\aleph_0}$ es la lógica de primer orden usual.

Es importante destacar, porque se utilizará más adelante, que con dicho lenguaje infinitario, en particular con $\kappa > \aleph_0$ se puede caracterizar el concepto de «relación bien fundamentada» (una relación binaria R es bien fundamentada si no existen cadenas infinitas descendientes con respecto a R , por ejemplo « \in » en \mathbf{V}) y el concepto de «relación bien ordenada». En efecto: La siguiente sentencia afirma que la relación binaria que nombra el símbolo relacional binario $P(x, y)$ es bien fundamentada:

$$\text{RBF: } (\forall x_0 x_1 x_2 \dots) \neg \bigwedge \{P(x_{n+1}, x_n) : n \in \aleph_0\},$$

y cuando se adiciona (con la conjunción finita \wedge) a la sentencia anterior los axiomas de orden total escritos con el lenguaje \mathcal{L} ,

$$\text{RBO: } [(\forall x_0 x_1 x_2 \dots) \neg \bigwedge \{P(x_{n+1}, x_n) : n \in \aleph_0\}] \wedge \text{Axiomas de orden total},$$

se tiene que la relación determinada por $P(x, y)$ es un buen orden en la estructura respectiva donde $P(x, y)$ sea interpretado y se satisfaga la sentencia RBO.

El teorema fundamental de ultraproductos tiene versiones más fuertes. En efecto, se cumple que los ultraproductos también preservan las fórmulas infinitarias (si se le agrega una hipótesis adicional al ultrafiltro, δ -completo, δ un cardinal mayor o igual que \aleph_0), ese es el significado del siguiente resultado, una prueba del mismo puede encontrarse en [3], p. 231-232]:

LEMA 10 (Los ultraproductos también preservan las fórmulas infinitarias). *Sea \mathfrak{B} el ultraproducto $\prod_D \mathfrak{A}_i$, donde I es el conjunto de índices de los \mathfrak{A}_i , y D es un ultrafiltro δ -completo (δ un cardinal mayor o igual que \aleph_0). Entonces:*

(i) *Dada una fórmula $\varphi(x_1, \dots, x_n, x_{n+1}, \dots)$ de $\mathcal{L}_{\delta\delta}$ y $f_D^1, \dots, f_D^n, f_D^{n+1}, \dots \in B$ se tiene que:*

$$\mathfrak{B} \models \varphi[f_D^1, \dots, f_D^n, f_D^{n+1}, \dots] \leftrightarrow$$

$$\{i \in I : \mathfrak{A}_i \models \varphi[f^1(i), \dots, f^n(i), f^{n+1}(i), \dots]\} \in D$$

(ii) *Para cada sentencia φ de $\mathcal{L}_{\delta\delta}$ se tiene que:*

$$\mathfrak{B} \models \varphi \leftrightarrow \{i \in I : \mathfrak{A}_i \models \varphi\} \in D$$

El siguiente teorema es una versión del teorema de compacidad (de la lógica de primer orden) para conjuntos de sentencias de lenguajes infinitarios cuyo cardinal sea un cardinal medible. La demostración supone ultraproductos y lenguajes infinitarios (preservados por ultraproductos). Con dicho teorema se abre la puerta para la definición de un nuevo tipo de cardinal grande: «cardinal débilmente compacto». Se cumple que «cardinal medible» implica «cardinal débilmente compacto» (y algo más fuerte), pero «cardinal débilmente compacto» no implica «cardinal medible» (ver [3], p. 243] y [5], p. 132]). Es decir, la hipótesis «existe un cardinal medible» es más fuerte que la hipótesis «existe un cardinal débilmente compacto». Los cardinales débilmente compactos se pueden caracterizar también con propiedades de combinatoria infinita tipo Ramsey (ver [5], p. 118]), después de la demostración del teorema se comentará brevemente este asunto.

TEOREMA 11 (Teorema de compacidad débil). *Sea η un cardinal medible, y Γ un conjunto de sentencias de $\mathcal{L}_{\eta\eta}$ tal que $|\Gamma| = \eta$ y cualquier subconjunto $\Gamma_0 \subseteq \Gamma$ tal que $|\Gamma_0| < \eta$ tiene un modelo. Entonces Γ tiene un modelo.*

Demostración. Como η es medible se tiene que existe un ultrafiltro H no principal y η -completo sobre η . Como H no contiene conjuntos unitarios y es η -completo, entonces H no contiene conjuntos de cardinal menor que η (para ver la intuición de este hecho se puede pensar que $\eta = \aleph_0$, aunque obviamente si η es medible es mayor que \aleph_0 por definición, es una sólo una idea para la intuición, y luego considerar que H es el filtro de Fréchet sobre \aleph_0 , de este modo el hecho descrito se aprecia claramente. Una demostración mas general y rigurosa es la siguiente: Sea $A \in H$ tal que $|A| < \eta$. Para todo $x \in A$ se cumple que $\{x\} \notin H$ porque H es ultrafiltro no principal y entonces por definición no contiene conjuntos unitarios. Entonces para todo $x \in A(A \setminus \{x\} \in H)$ ya que H es ultrafiltro. En consecuencia, como H es η -completo se cumple que:

$\bigcap_{x \in A} (A \setminus \{x\}) \in H$. Por lo tanto: $\emptyset = (\bigcap_{x \in A} A \setminus \{x\}) \cap A \in H$, pues H es un filtro. Contradicción). Por lo tanto para cada $\delta < \eta$,

$$\{\theta : \delta < \theta < \eta\} \in H,$$

ya que el complemento de $\{\theta : \delta < \theta < \eta\}$ no pertenece a H porque tiene cardinal menor que η , y entonces como H es un ultrafiltro $\{\theta : \delta < \theta < \eta\} \in H$.

Sea la siguiente enumeración de Γ : $\Gamma = \{\phi_\beta : \beta < \eta\}$. Por hipótesis, para cada $\beta < \eta$, existe un modelo \mathfrak{C}_β del conjunto $\{\phi_\delta : \delta < \beta\}$. Sea

el ultraproducto $\mathfrak{D} = \prod_H \mathfrak{C}_\beta$. Entonces para cada $\phi_\delta \in \Gamma$, se tiene que,

$$\{\theta < \eta : \mathfrak{C}_\theta \models \phi_\delta\} \supseteq \{\theta : \delta < \theta < \eta\} \in H.$$

En consecuencia, como los ultraproductos preservan las formulas infinitarias (por el Lema 10), se cumple que el ultraproducto \mathfrak{D} es un modelo de Γ . Lo que se quería demostrar. \square

Definición 12. Un cardinal $\kappa > \aleph_0$ es *débilmente compacto* si para todo conjunto Γ de sentencias de $\mathcal{L}_{\kappa\kappa}$ tal que $|\Gamma| = \kappa$ ocurre lo siguiente: Si cada subconjunto de Γ de cardinalidad menor que κ tiene un modelo, entonces Γ tiene un modelo.

Es decir, κ es débilmente compacto si cumple con el teorema anteriormente demostrado, en otras palabras, un cardinal medible es débilmente compacto.

Nótese que como la lógica de primer orden $\mathcal{L}_{\aleph_0\aleph_0}$ satisface el teorema de compacidad esto implica que \aleph_0 es débilmente compacto, si se elimina la restricción de que el cardinal débilmente compacto debe ser mayor que \aleph_0 , es decir, la definición de cardinal débilmente compacto es una generalización de una propiedad de \aleph_0 para cardinales no numerables, algo análogo ocurre con la noción de cardinal inaccesible y la de cardinal medible.

Como se dijo anteriormente todo cardinal medible es débilmente compacto, pero lo inverso no ocurre, débilmente compacto no implica medible, una demostración de ello puede encontrarse en [3], p. 243].

La relación combinatoria $\kappa \rightarrow \kappa_2^2$ significa que para toda partición en dos clases del conjunto de subconjuntos de dos elementos de κ existe un subconjunto $H \subseteq \kappa$ cuyos subconjuntos de dos elementos están todos en el misma clase y H tiene cardinal κ . Esta definición se puede reexpresar de la siguiente manera teniendo presente que $[A]^2 = \{\{x, y\} : x \in A \wedge y \in A\}$ y que $F''[A] = \{F(x) : x \in A\}$. Entonces $\kappa \rightarrow \kappa_2^2$ significa que para toda función $F : [\kappa]^2 \rightarrow \{0, 1\}$ existe un subconjunto $H \subseteq \kappa$ tal que $|H| = \kappa$ y existe un $i \in \{0, 1\}$ tal que $F''[H]^2 = \{i\}$.

Se cumple el siguiente teorema (ver [5], p. 118]):

Si κ es inaccesible, entonces κ es débilmente compacto si y sólo si $\kappa \rightarrow \kappa_2^2$.

Para culminar con los cardinales débilmente compactos vale la pena agregar lo siguiente: *cardinal débilmente compacto* implica estrictamente *cardinal inaccesible*, es decir, la hipótesis «Existe un cardinal débilmente compacto» es más fuerte estrictamente que la hipótesis «Existe un cardinal inaccesible» (ver [5], p. 132)].

4. SEGUNDO TEOREMA SOBRE CARDINALES MEDIBLES: «CARDINAL MEDIBLE» ES MÁS FUERTE ERICTAMENTE QUE «CARDINAL INACCESIBLE».

El segundo ejemplo de aplicación del método de ultraproductos en la teoría de conjuntos con cardinales medibles es la demostración del teorema: *Si α es un cardinal medible, entonces α es un cardinal inaccesible y además α es el α -ésimo cardinal inaccesible, es decir, existen α cardinales inaccesibles menores que α .* Esto significa (entre otros) que la hipótesis conjuntista «Existen cardinales medibles» es estrictamente más fuerte que la hipótesis «existen cardinales inaccesibles» (ver [3,5]). Es conocido que la existencia de cardinales inaccesibles no se puede demostrar de los axiomas estándar de la teoría de conjuntos, por el Segundo teorema de incompletitud de Gödel(1931)

(ver [3,5,11,13,17]). De modo que es claro que tampoco se puede demostrar de los axiomas estándar de la teoría de conjuntos la existencia de los cardinales medibles ni la existencia de ningún otro cardinal más fuerte que inaccesible.

Vale la pena resaltar que en la prueba del teorema de esta sección se usarán lógicas infinitarias (como en el teorema anterior) y lógicas de segundo orden de un tipo específico, Σ_1^1 fórmulas, las cuales son preservadas por los ultraproductos. Algunos resultados y definiciones que se presuponen en la prueba se enuncian a continuación:

El siguiente resultado se refiere a la expansión de productos reducidos, una prueba del mismo puede encontrarse en [3], p. 216-217]:

LEMA 13 (Teorema de expansión de productos reducidos). *Sea un lenguaje \mathcal{L}' que expande a un lenguaje \mathcal{L} . Sea I un conjunto no vacío y para cada $i \in I$ sea \mathfrak{A}_i una estructura para \mathcal{L} , y \mathfrak{B}_i una expansión de \mathfrak{A}_i para \mathcal{L}' (es decir, \mathfrak{B}_i restringida a \mathcal{L} es \mathfrak{A}_i , en otras palabras, los universos de \mathfrak{B}_i y \mathfrak{A}_i son iguales, y las interpretaciones de los símbolos de \mathcal{L} en ambas también son iguales. \mathfrak{B}_i difiere de \mathfrak{A}_i sólo en la interpretación de los nuevos símbolos de $\mathcal{L}' \setminus \mathcal{L}$). Sea D un filtro sobre I . Entonces el producto reducido $\prod_D \mathfrak{B}_i$ es una expansión del producto reducido $\prod_D \mathfrak{A}_i$. (Es decir, $\prod_D \mathfrak{B}_i$ restringida a \mathcal{L} es $\prod_D \mathfrak{A}_i$)*

Las Σ_1^1 fórmulas y los ultraproductos:

Un tipo específico de fórmulas de segundo orden son preservadas por los ultraproductos, las Σ_1^1 fórmulas, tal resultado se expresará a continuación mediante una definición y un lema, una demostración de dicho lema puede encontrarse en [3], p. 222]:

Sea \mathcal{L} un lenguaje de primer orden, y sea una expansión de \mathcal{L} , $\mathcal{L}' = \mathcal{L} \cup \{Q_1, \dots, Q_n, G_1, \dots, G_m\}$, donde los Q_j y los G_i son símbolos de relación y de función, respectivamente, que no ocurren en \mathcal{L} . Una Σ_1^1 fórmula sobre \mathcal{L} es una fórmula ϱ de la siguiente forma:

$$(\exists Q_1 \dots Q_n G_1 \dots G_m)\phi,$$

donde ϕ es una fórmula (en primer orden) del lenguaje expandido \mathcal{L}' . De modo que una Σ_1^1 fórmula es una fórmula de segundo orden donde todos los cuantificadores sobre relaciones y funciones ocurren al inicio de dicha fórmula, y además de eso tales cuantificadores sólo son existenciales. La definición de satisfacibilidad de una Σ_1^1 fórmula se define como sigue: Si ϕ es una sentencia, entonces ϱ ocurre en una estructura \mathfrak{C} para \mathcal{L} si y sólo si existe una expansión $\mathfrak{C}' = (\mathfrak{C}, P_1, \dots, P_n, F_1, \dots, F_m)$ de \mathfrak{C} para \mathcal{L}' tal que ϕ es verdad en \mathfrak{C}' . Si ϕ tiene una variable libre z , entonces $\mathfrak{C} \models \varrho[b]$ si y sólo si existe una expansión \mathfrak{C}' de \mathfrak{C} para \mathcal{L}' tal que $\mathfrak{C}' \models \phi[b]$.

LEMA 14 (Los ultraproductos también preservan las Σ_1^1 fórmulas). *Sea \mathfrak{B} el ultraproducto $\prod_D \mathfrak{A}_i$, donde I es el conjunto de índices de los \mathfrak{A}_i , $f_D^1, \dots, f_D^k \in B$, y $\varrho(z_1, \dots, z_k)$ es una Σ_1^1 fórmula. Entonces:*

Si,

$$\{i \in I : \mathfrak{A}_i \models \varrho[f^1(i), \dots, f^k(i)]\} \in D,$$

Entonces,

$$\mathfrak{B} \models \varrho[f_D^1, \dots, f_D^k].$$

El siguiente resultado sobre cardinales medibles se usa en el teorema que se demostrará, una prueba del mismo puede encontrarse en [3], p. 233]:

LEMA 15. *Sea η un cardinal medible, y sea H un ultrafiltro sobre η , no principal y η -completo. Se forma la ultrapotencia $\mathfrak{C} = \prod_H \langle \eta, < \rangle$. Entonces:*

(i) \mathfrak{C} es una estructura bien ordenada de tipo de orden mayor que η .

(ii) Para cualquier $\delta < \eta$, $d(\delta)$ es el δ -ésimo elemento de \mathfrak{C} .

La demostración del siguiente teorema usa ideas de la prueba que se encuentra en el texto [3], p. 233-236]. Otra prueba de este teorema que también usa ultraproductos pero que es distinta a la que se realizará aca se puede encontrar en [15], p. 313].

TEOREMA 16. *Sea η un cardinal medible. Entonces η es un cardinal inaccesible y además η es el η -ésimo cardinal inaccesible, es decir, existen η cardinales inaccesibles menores que η .*

Demostración. Sea H un ultrafiltro no principal y η -completo sobre η . Se considera el modelo $\mathfrak{C} = \langle \eta, <, \rho \rangle_{\rho \in \eta}$ y se forma la ultrapotencia $\mathfrak{D} = \prod_H \mathfrak{C}$. Por el Lema anterior (15) \mathfrak{D} es una estructura bien ordenada de tipo de orden mayor que η . Sea $\gamma > \eta$ el tipo de orden de \mathfrak{D} y para cada $\delta < \gamma$, sea $\hat{\delta}$ el δ -ésimo elemento de \mathfrak{D} . Como por el Corolario 9 el rango de la inmersión natural $d : \langle \eta, <, \rho \rangle_{\rho \in \eta} \rightarrow \prod_H \mathfrak{D}$, $\text{rango}(d) = \{d(\delta) : \delta \in \eta\}$, es un segmento inicial de \mathfrak{D} , entonces se cumple que $d(\delta) = \hat{\delta}$, para todo $\delta < \eta$. Por esta razón cada constante c_δ , $\delta < \eta$, es interpretada por δ en la estructura \mathfrak{C} y por $d(\delta) = \hat{\delta}$ en la ultrapotencia \mathfrak{D} .

(I) Ahora se probará que el cardinal η es un cardinal inaccesible:

(I.1) Demostración de que η es un cardinal regular:

(Por reducción al absurdo) Supóngase que η no es un cardinal regular, es decir, η es un cardinal singular. Entonces existe un ordinal $\alpha < \eta$ cofinal con η , es decir, existe una función creciente $U : \alpha \rightarrow \eta$ cuya imagen es no acotada en η . Entonces se define la siguiente función $U' : \eta \rightarrow \eta$ de la siguiente manera: $U'(\beta) = 0$, para todo $\eta > \beta \leq \alpha$. Y $U'(\beta) = U(\beta)$, para todo $\beta < \alpha$. Con esta función U' se forma el modelo (\mathfrak{C}, U') , y la ultrapotencia,

$$\prod_H (\mathfrak{C}, U') = (\mathfrak{D}, W).$$

Para cada $\beta < \eta$, se tiene que:

$$W(\hat{\beta}) = W(d(\beta)) \overset{\spadesuit}{\equiv} d(U'(\beta)) = U'(\hat{\beta}) < \hat{\eta}.$$

\spadesuit : Como la inmersión natural d es una inmersión elemental, se está aplicando la cláusula (2) de la definición de inmersión elemental.

Por lo tanto, en la ultrapotencia (\mathfrak{D}, W) es verdadera la siguiente sentencia:

$$\exists x \forall y (y < c_\alpha \rightarrow U'(y) < x), (\bullet)$$

para $x = \hat{\eta}$.

Sin embargo, como el rango de U' es cofinal en η , se cumple que la siguiente sentencia es verdad en la estructura (\mathfrak{C}, U') :

$$\forall x \exists y (y < c_\alpha \wedge x < U'(y)), (\circ)$$

y por lo tanto (por la inmersión natural d) tal sentencia es también verdadera en la ultrapotencia (\mathfrak{D}, W) . Pero las sentencias (\bullet) y (\circ) se contradicen mutuamente. Por lo tanto, η no puede ser un cardinal singular, en consecuencia η es un cardinal regular.

(I.2) Demostración de que η es un cardinal límite fuerte:

Hay que probar que para todo cardinal $\beta < \eta$ se tiene que $2^\beta < \eta$. Se hará la prueba por reducción al absurdo. Supóngase que existe un cardinal κ tal que,

$$\kappa < \eta \leq 2^\kappa.$$

En consecuencia existe un función inyectiva $W : \eta \rightarrow P(\kappa)$. Sea $T \subseteq \eta \times \kappa$ una relación binaria, «la representante de W », que se define como sigue:

$$T(\theta, \sigma) \iff \sigma \in W(\theta).$$

Entonces se forma la estructura (\mathfrak{C}, T) y se considera la ultrapotencia:

$$\prod_H (\mathfrak{C}, T) = (\mathfrak{D}, Q).$$

Sea W' una función de dominio γ (γ es el tipo de orden de \mathfrak{D} fijado anteriormente) definida de la siguiente manera:

$$W'(\delta) = \{\theta < \gamma : Q(\hat{\delta}, \hat{\theta})\}.$$

Se cumple que la función W' es una función inyectiva de γ en $P(\kappa)$, porque las dos sentencias siguientes ocurren en la estructura (\mathfrak{C}, T) y por lo tanto (por la inmersión natural d , $(\mathfrak{C}, T) \preceq (\mathfrak{D}, Q)$) en la ultrapotencia (\mathfrak{D}, Q) :

$$\begin{aligned} & \forall x \forall y (T(x, y) \rightarrow y < c_\kappa), \\ & \forall x \forall y [(x \neq y) \rightarrow \exists z \neg (T(x, z) \leftrightarrow T(y, z))]. \end{aligned}$$

También se cumple que $W(\delta) = W'(\delta)$, para toda $\delta < \eta$, ya que d es una inmersión elemental de (\mathfrak{C}, T) en (\mathfrak{D}, Q) . De lo anterior se puede inferir que

el conjunto $Z = W'(\eta)$ no está en el rango de W , a pesar de que $Z \in P(\kappa)$. De modo que la siguiente sentencia,

$$\exists x \forall y [T(x, y) \leftrightarrow \bigvee \{y \equiv c_\pi : \pi \in Z\}],$$

es falsa en la estructura (\mathfrak{C}, T) , pero es verdadera en la ultrapotencia (\mathfrak{D}, Q) considerando a $x = \hat{\eta}$. Esto contradice el que d sea una inmersión elemental de (\mathfrak{C}, T) en (\mathfrak{D}, Q) . Por lo tanto η es un cardinal límite fuerte.

(II) Demostración de que η es el η -ésimo cardinal inaccesible:

Esta parte de la demostración se realiza considerando el lema que afirma que las Σ_1^1 fórmulas son preservadas por los ultraproductos (Lema 14).

Sea Θ la clase de todos los cardinales inaccesibles, Δ la clase de todos los ordinales los cuales no son cardinales regulares, y Ω la clase de todos los ordinales los cuales no son cardinales límite fuerte. En consecuencia $\alpha \in \Theta$ si y sólo si $\alpha \notin \Delta \cup \Omega$. Se probará que $\Theta \cap \eta$ es cofinal con η . En consecuencia, dado que η es regular, se cumple que $|\Theta \cap \eta| = \eta$, por lo tanto η es el η -ésimo cardinal inaccesible, lo que se quiere demostrar, es decir, (II) ocurre.

Como $\eta \subseteq \Theta \cup \Delta \cup \Omega$, es suficiente con demostrar que para cada $\beta < \eta$,

(1) existe un θ tal que $\beta \leq \theta < \eta$ y $\theta \notin \Delta \cup \Omega$.

Supóngase que para algún $\beta < \eta$, (1) no ocurre. Entonces para todo $\theta < \eta$ se cumple que:

(2) $\theta < \beta$ o $\theta \in \Delta$ o $\theta \in \Omega$.

Existe una Σ_1^1 fórmula $\phi_\Delta(v)$ tal que para cualquier modelo $\langle \eta', < \rangle$ donde η' es un ordinal, y cualquier $\tau \in \eta'$,

$$(3) \tau \in \Delta \text{ si y sólo si } \langle \eta', < \rangle \models \phi_\Delta[\tau].$$

$\phi_\Delta(v)$ se obtiene formalizando la siguiente sentencia:

« Existe un $z < v$ y existe una función $J : z \rightarrow v$ tal que el rango de J es cofinal con v ».

También existe una Σ_1^1 fórmula $\varphi_\Omega(v)$ tal que para cualquier modelo $\langle \eta', < \rangle$ donde η' es un ordinal, y cualquier $\tau \in \eta'$,

$$(4) \tau \in \Omega \text{ si y sólo si } \langle \eta', < \rangle \models \varphi_\Omega[\tau].$$

$\phi_\Delta(v)$ se obtiene formalizando la siguiente sentencia:

« Existe un $z < v$ y existe una relación $T \subseteq v \times z$ tal que T representa a una función inyectiva de v en $P(z)$ ».

Anteriormente (en esta demostración) se ha explicado como definir una relación T que representa a una función inyectiva $W : v \rightarrow P(z)$ y también se ha explicado como decir que W es inyectiva usando a T .

A partir de (2), (3) y (4) se tiene que la fórmula,

$$(5) \quad v < c_\beta \vee \phi_\Delta(v) \vee \varphi_\Omega(v),$$

es verdadera en $\langle \eta, < \rangle$ para todo $v \in \eta$.

Moviendo los cuantificadores de segundo orden para el inicio se puede apreciar que (5) es equivalente a una Σ_1^1 fórmula. Por lo tanto, por el Lema 14 (Las fórmulas Σ_1^1 son preservadas por ultraproductos), para cualquier $g_H \in B$ la fórmula (5) es satisfecha por g_H en $\langle D, < \rangle$, donde $\langle D, < \rangle$ es la ultrapotencia correspondiente a la estructura $\langle \eta, < \rangle$, construida con el ultrafiltro H no principal y η -completo sobre η . Dado que $\langle D, < \rangle$ es isomorfa a $\langle \gamma, < \rangle$, la fórmula (5) es verdad en $\langle \gamma, < \rangle$, para toda $v \in \gamma$. Poniendo $v = \eta$ se tiene que:

$$\eta < \beta \quad \text{o} \quad \phi_\Delta(\eta) \quad \text{o} \quad \varphi_\Omega(\eta).$$

Usando (3) y (4), con $\gamma = \eta'$, se tiene que:

$\eta < \beta$ o $\eta \in \Delta$ o $\eta \in \Omega$. Pero esto contradice la hipótesis de que $\beta < \eta$ y que η es un cardinal inaccesible. Con esto termina la demostración del Teorema. \square

5. TERCER TEOREMA SOBRE CARDINALES MEDIBLES: EL TEOREMA DE SCOTT.

El tercer teorema sobre cardinales medibles que se demostrará (usando ultraproductos) es el teorema de Scott [22]: *Si existen cardinales medibles, entonces el axioma de constructibilidad ($V=L$) es falso.*

Es conocido que el axioma de constructibilidad implica la «hipótesis generalizada del continuo» ($\forall \alpha \in \mathbf{Ord} (2^{\aleph_\alpha} = \aleph_{\alpha+1})$) y al axioma de elección (ver [11,13,17]).

También es conocido que Gödel demostró que se cumple $\langle \mathbf{L}, \in \rangle \models \mathbf{V} = \mathbf{L}$ (ver [11,13,17]), entonces el teorema de Scott implica que en \mathbf{L} no existen cardinales medibles.

La prueba del teorema de Scott se realiza utilizando ideas principalmente de Chang y Keisler en [3, p. 238-239]. Como en la demostración de los dos teoremas anteriores sobre cardinales medibles también se usan en esta prueba lenguajes infinitarios (los cuales son preservados por los ultraproductos). Otra demostración de este teorema que también usa ultraproductos pero que es distinta a la que se realizará aca se puede encontrar en [15, p. 311].

A continuación se enuncian algunos resultados y definiciones previas:

\mathbf{V} es la clase de los conjuntos bien fundamentados («La jerarquía acumulativa de conjuntos de Zermelo») que se define usando inducción transfinita en la clase de los ordinales:

$$\begin{aligned} V_0 &= \emptyset \\ V_{\alpha+1} &= P(V_\alpha) \\ V_\lambda &= \bigcup_{\beta < \lambda} V_\beta, \quad \lambda \text{ límite.} \\ \mathbf{V} &= \bigcup_{\alpha \in \mathbf{Ord}} V_\alpha. \end{aligned}$$

Si $x \in \mathbf{V}$ entonces el *rango* de x , $\rho(x)$, es el menor ordinal α tal que $x \in V_{\alpha+1}$.

Para cada $\alpha \in \mathbf{Ord}$ ocurre que V_α es un conjunto transitivo (X es *transitivo* si $\forall z(z \in X \rightarrow z \subseteq X)$). También se cumple que: (i) Para cada $\alpha, \beta \in \mathbf{Ord}$: Si $\alpha \leq \beta$, entonces $V_\alpha \subseteq V_\beta$. (ii) Para cada ordinal α : $V_\alpha \cap \mathbf{Ord} = \alpha$. Y (iii) si η es un cardinal medible, entonces $\langle V_\eta, \in \rangle$ es un modelo de *ZFC*. (ver [3,13,15]).

\mathbf{L} es la clase de los conjuntos constructibles de Gödel que se define (informalmente) usando inducción transfinita en la clase de los ordinales:

Antes de dar la definición se introduce el concepto de definibilidad en una estructura (ver [3,5]): Consideremos una estructura $\mathfrak{A} = \langle A, < R_\beta^{\mathfrak{A}} >_{\beta \in \gamma}, < f_\mu^{\mathfrak{A}} >_{\mu \in \delta}, < c_\xi^{\mathfrak{A}} >_{\xi \in \eta}$ para un lenguaje \mathcal{L} . Se dice que un subconjunto $B \subseteq A$ es *definible* en \mathfrak{A} si existe una

fórmula $\varphi(x)$ del lenguaje \mathcal{L} tal que $B = \{z \in A : \mathfrak{A} \models \varphi[z]\}$. Se dice que B es definible en \mathfrak{A} con parámetros si existe fórmula $\varphi(x, x_1, \dots, x_n)$ del lenguaje \mathcal{L} y existen $a_1, \dots, a_n \in A$ tal que: $B = \{z \in A : \mathfrak{A} \models \varphi[z, a_1, \dots, a_n]\}$.

$$L_0 = \emptyset$$

$$L_{\alpha+1} = \{X \subseteq L_\alpha : X \text{ es definible en la estructura } \langle L_\alpha, \in, \langle b : b \in L_\alpha \rangle \rangle\}$$

$$L_\lambda = \bigcup_{\beta \in \lambda} L_\beta, \quad \lambda \text{ límite.}$$

$$\mathbf{L} = \bigcup_{\alpha \in \text{Ord}} L_\alpha$$

Es claro que en el paso sucesor la expresión « X es definible en la estructura $\langle L_\alpha, \in, \langle b : b \in L_\alpha \rangle \rangle$ » supone que se tiene un lenguaje de primer orden con identidad, cuyos símbolos no lógicos son: Una constante \underline{b} para cada $b \in L_\alpha$ y un símbolo relacional binario $\underline{\in}$ para la relación de pertenencia \in . Sin embargo también se puede hablar de «definible con parámetros» y eliminar las nuevas constantes agregadas al lenguaje inicial. Formalizaciones en ZF de la definición intuitiva de \mathbf{L} pueden encontrarse en [13,17].

Se cumple que para $\alpha \in \mathbf{Ord}$: L_α es transitivo y $L_\alpha \subseteq V_\alpha$. También ocurre que para cada $n \in \aleph_0$ ($L_n = V_n$). Y también se cumple que: (i) $L_\omega = V_\omega$. Y (ii) para cada ordinal α : $L_\alpha \cap \mathbf{Ord} = \alpha$.

Sea α un cardinal infinito. $H(\alpha)$ es el conjunto de todos los conjuntos hereditariamente de cardinal menor que α , es decir, $H(\alpha) = \{x : |TC(x)| < \alpha\}$, donde dado un conjunto D , $TC(D)$ es la *clausura transitiva* de D , es decir, $TC(D)$ es el menor conjunto transitivo (con respecto a la relación inclusión) que contiene a D . Se cumple que $z \in H(\alpha)$ si y sólo si existe un conjunto transitivo w tal que $z \subseteq w$ y $|w| < \alpha$. Entre las propiedades $H(\alpha)$ se encuentran las siguientes (ver [3], p. 237] y [17], p. 130-133]):

- LEMA 17. 1. $H(\alpha)$ es un subconjunto transitivo de V_α .
 2. Si $\alpha \leq \beta$, entonces $H(\alpha) \subseteq H(\beta)$.
 3. $\alpha \subseteq H(\alpha)$.
 4. Si $\alpha > \aleph_0$ es un cardinal regular, entonces $\langle H(\alpha), \in \rangle$ es un modelo de $ZF - P$.
 5. $H(\alpha) \cap \mathbf{Ord} = \alpha$.

Como se definió anteriormente en este artículo una relación binaria E es *bien fundamentada* si no existen secuencias infinitas decrecientes con respecto a E , es decir, si no existen secuencias $\{x_n : n \in \aleph_0\}$ tal que: $\dots, x_4 E x_3, x_3 E x_2, x_2 E x_1, x_1 E x_0$.

Sea X un conjunto y E una relación binaria bien fundamentada sobre X . La estructura $\langle X, E \rangle$ se llama estructura o modelo *bien fundamentado*.

Sea $\langle X, E \rangle$ una estructura bien fundamentada. Se dice que $a \in X$ es un *ordinal* de $\langle X, E \rangle$ si y sólo si ocurre lo siguiente:

$$\langle X, E \rangle \models \forall x \forall y [(x \in a \wedge y \in a \rightarrow x \in y \vee y \in x \vee x \equiv y) \wedge$$

$$(x \in a \wedge y \in x \rightarrow y \in a)].$$

Hecho \star : Dada la definición anterior se cumple que u es un ordinal de $\langle H(\alpha), \in \rangle$ si y sólo si $u \in \alpha$. (ver [3], p. 238].

El siguiente resultado afirma que las relaciones bien fundamentadas sobre un conjunto dado X bien ordenan los ordinales de dicho conjunto X si $\langle X, E \rangle \models ZF - P$, una prueba del mismo puede encontrarse en [3], p. 238]:

LEMA 18. *Sea $\langle X, E \rangle$ un modelo bien fundamentado de $ZF - P$. Entonces el conjunto de los ordinales de $\langle X, E \rangle$ está bien ordenado por E .*

Sea $\langle X, E \rangle$ un modelo bien fundamentado de $ZF - P$. El *tipo de orden de* $\langle X, E \rangle$ es el tipo de orden del conjunto de ordinales de $\langle X, E \rangle$ según E .

Formulación del axioma de constructibilidad (la versión de dicho axioma que se utilizará en esta demostración):

Axioma de constructibilidad: Para cualquier cardinal regular $\mu > \aleph_0$ (todo modelo bien fundamentado $\langle A, S \rangle$ de $ZF - P$ de tipo de orden μ es isomorfo a $\langle H(\mu), \in \rangle$).

Observación: Se puede apreciar que la versión del axioma de constructibilidad que se acaba de formular es diferente a la que se usa normalmente en teoría de conjuntos: $\mathbf{V} = \mathbf{L}$, expresión que denota a la proposición $\forall x \exists \alpha \in \mathbf{Ord}(x \in L_\alpha)$. La razón de la elección es que la versión del axioma de constructibilidad que utilizaremos conviene más que la versión $\mathbf{V} = \mathbf{L}$ a los fines de poder aplicar las técnicas que usaremos en la demostración del teorema de Scott. Como se dijo anteriormente existen pruebas distintas del teorema de Scott de la que haremos en este artículo las cuales usan la versión $\mathbf{V} = \mathbf{L}$ del axioma de constructibilidad, por ejemplo en [6,13,15,16], sin embargo, se puede demostrar en ZFC que la versión del axioma de constructibilidad que utilizaremos es equivalente a la versión $\mathbf{V} = \mathbf{L}$. No daremos una demostración rigurosa de tal equivalencia pues dispersaría al lector de la idea principal del artículo ya que dicha demostración requiere de la introducción de conceptos y resultados adicionales (que no están en el artículo), no obstante, mencionaremos algunas ideas fundamentales que se pueden aplicar en la demostración de dicha equivalencia y ofreceremos un esbozo de una prueba que es original del autor de este artículo: Primero que nada es conveniente decir, siguiendo a [3], que dicha versión afirma intuitivamente que $H(\mu)$ es muy estrecho ¿y qué significa esto? se puede explicar muy bien con la siguiente proposición que también es una versión equivalente del axioma de constructibilidad, la cual llamaremos «axioma de constructibilidad (versión 3)»: *Para cualquier cardinal regular $\mu > \aleph_0$ (no existe un subconjunto propio $M \subset H(\mu)$ tal que M es transitivo, $\alpha \subseteq M$ y $\langle M, \in \rangle$ es un modelo de $ZF - P$).* Con respecto a $H(\mu)$ es bueno tener presente la cláusula (4) del lema anterior 17, cuya prueba se puede encontrar en los textos mencionados cuando se enunció el lema. Ahora bien, una prueba de las equivalencias mencionadas se puede hacer usando los siguientes resultados (y las ideas contenidas en el mismos): (a) Para cualquier cardinal regular $\kappa > \aleph_0$: $L_\kappa \models ZF - P$. Una prueba de dicho resultado (a) puede encontrarse en [20], p.84]. Y (b) Si $\mathbf{V} = \mathbf{L}$, entonces para todo κ , $H(\kappa) = L_\kappa$. Una prueba de tal resultado (b) se puede encontrar en [20], p. 84-85]. Entre las ideas y resultados contenidos en las demostraciones de (a) y (b) están: (1) La relación funcional de los constructibles $\alpha \mapsto L_\alpha$ es absoluta para modelos transitivos de $ZF - P$, (2) absolutividad en general, (3) teorema de reflexión, (4) teorema de Löwenheim-Skolem, (5) teorema de colapso transitivo de Mostowski y (6) lema de condensación. Finalmente, para probar que las tres versiones del axioma de constructibilidad son equivalentes, usando las ideas y los resultados anteriores (entre otros), se puede proceder así: Primero se prueba que: (I) axioma de constructibilidad \Rightarrow axioma de constructibilidad (versión 3). Luego se demuestra que: (II) axioma de constructibilidad (versión 3) $\Rightarrow \mathbf{V} = \mathbf{L}$. Y por último se demuestra que: (III) $\mathbf{V} = \mathbf{L} \Rightarrow$ axioma de constructibilidad. Esbozo de la prueba de (I): Sea μ un cardinal regular tal que $\mu > \aleph_0$. Y supongamos que existe un $M \subset H(\mu)$ tal que M es transitivo, $\alpha \subseteq M$ y $\langle M, \in \rangle$ es un modelo de $ZF - P$. Entonces, por la hipótesis, $\langle M, \in \rangle$ es isomorfo a $\langle H(\mu), \in \rangle$. Sea π la función biyectiva de M en $H(\mu)$ que establece el isomorfismo. En consecuencia π es la función identidad, es decir, $\pi(x) = x$ ($\forall x \in M$), y por lo tanto, $M = H(\mu)$, pues dos clases transitivas distintas no pueden ser \in -isomorfas (ver teorema 26 del texto [15], p. 74]). Esbozo de la prueba de (II): Sea $x \in \mathbf{V}$. Consideramos la clausura transitiva del conjunto unitario $\{x\}$, es decir, $TC(\{x\})$. Sea α el cardinal de $TC(\{x\})$ y α^+ es cardinal sucesor de α . Ocurre que

$x \in H(\alpha^+)$. Consideramos $\langle H(\alpha^+), \in \rangle$. Por el Lema 17 tenemos que $\langle H(\alpha^+), \in \rangle$ es un modelo de $ZF - P$. Por otro lado tenemos que, por la absolutividad de $\alpha \mapsto L_\alpha$, $L_{\alpha^+} \subseteq H(\alpha^+)$. También ocurre que $\langle L_{\alpha^+}, \in \rangle$ es un modelo de $ZF - P$ (resultado (a) anteriormente referido). Entonces $x \in L_{\alpha^+}$, pues si esto no ocurriera se contradice la hipótesis. Por lo tanto, $x \in \mathbf{L}$. Esbozo de la prueba de (III): Sea μ un cardinal regular tal que $\mu > \aleph_0$. Y sea $\langle A, S \rangle$ un modelo bien fundamentado de $ZF - P$ de tipo de orden μ . Entonces por el teorema del colapso transitivo de Mostowski se tiene que existe una clase transitiva M tal que $\langle A, S \rangle$ y $\langle M, \in \rangle$ son isomorfas. Como $\alpha \mapsto L_\alpha$ es absoluta, se tiene que $L_\mu \subseteq M$. Sea $x \in M$. Entonces, como por hipótesis $\mathbf{V} = \mathbf{L}$, $x \in \mathbf{L}$. Y como $\rho(x) < \mu$ y $L_\mu = \{z \in \mathbf{L} : \rho(z) < \mu\}$, se tiene que $x \in L_\mu$. De modo que $M \subseteq L_\mu$. En consecuencia, $M = L_\mu$. Por otro lado, por el resultado (b) referido anteriormente, ocurre que $M = L_\mu = H(\mu)$. Luego, $\langle A, S \rangle$ y $\langle H(\mu), \in \rangle$ son isomorfas. Fin del esbozo de la prueba (III), y con esto también termina el esbozo de la prueba de las equivalencias de las tres versiones mencionadas del Axioma de constructibilidad.

TEOREMA 19 (Teorema de Scott). *Si existe un cardinal medible, entonces el axioma de constructibilidad es falso. (En consecuencia el axioma de constructibilidad implica que no existen cardinales medibles).*

Demostración. Sea η el primer cardinal medible. Y sea $\theta = |2^{2^{2^\eta}}|^+$. Entonces $V_{\eta+3}$ es un conjunto transitivo de cardinalidad menor que θ , pues:

$$|V_{\eta+1}| = |P(V_\eta)| = 2^{|V_\eta|} (\eta \text{ es inaccesible, } |V_\eta| = \eta) = 2^\eta, \quad (\text{ver } [\underline{15}], \text{ p. 72])}$$

$$|V_{\eta+2}| = |P(V_\eta + 1)| = 2^{|V_{\eta+1}|} = 2^{2^\eta},$$

$$|V_{\eta+3}| = |P(V_{\eta+2})| = 2^{|V_{\eta+2}|} = 2^{2^{2^\eta}} < \theta.$$

Y también θ es un cardinal regular porque todo cardinal sucesor es regular (ver [5], p. 92]).

En consecuencia $V_{\eta+3} \in H(\theta)$ y $\langle H(\theta), \in \rangle$ es un modelo de $ZF - P$, porque θ es un cardinal regular no numerable (Lema 17, cláusula 4).

Sea H un ultrafiltro no principal y η -completo sobre η . Con H y $\langle H(\theta), \in \rangle$ se forma la ultrapotencia,

$$\langle \mathfrak{D}, R \rangle = \prod_H \langle H(\theta), \in \rangle.$$

Entonces como el concepto «relación bien fundamentada» se puede caracterizar con una fórmula del lenguaje infinitario $\mathcal{L}_{\aleph_1 \aleph_1}$,

$$\text{RBF: } (\forall x_0 x_1 x_2 \dots) \neg \bigwedge \{P(x_{n+1}, x_n) : n \in \aleph_0\},$$

y los ultraproductos preservan las formulas infinitarias (Lema 10), entonces $\langle \mathfrak{D}, R \rangle$ es un modelo bien fundamentado de $ZF - P$ (los axiomas de $ZF - P$ se cumplen en $\langle \mathfrak{D}, R \rangle$ porque existe la inmersión (elemental) natural d de $\langle H(\theta), \in \rangle$ en $\langle \mathfrak{D}, R \rangle$). En consecuencia, por el Lema 18, el conjunto de los ordinales de $\langle \mathfrak{D}, R \rangle$ está bien ordenado por R .

Proposición: El conjunto de los ordinales de $\langle \mathfrak{D}, R \rangle$ tiene tipo de orden θ .

Prueba de la proposición: La inmersión natural d es un isomorfismo de $\langle H(\theta), \in \rangle$ en $\langle \mathfrak{D}, R \rangle \upharpoonright d(H(\theta))$. En consecuencia el conjunto de los ordinales de $\langle \mathfrak{D}, R \rangle$ tiene tipo de orden al menos θ (tomar en cuenta el **Hecho** \star). Sea z cualquier ordinal de $\langle \mathfrak{D}, R \rangle$. Entonces se concluye que $z = g_H$, para alguna función $g : \eta \rightarrow \theta$.

(Nota: Usando la definición de «ordinal de una estructura bien fundamentada» y el Lema 18 de buen orden se puede probar que todo ordinal z de la ultrapotencia $\langle \mathfrak{D}, R \rangle = \prod_H \langle H(\theta), \in \rangle$ tiene la forma g_H , para alguna función $g : \eta \rightarrow \theta$.)

Como $\text{cof}(\theta) > \eta$, entonces g no es cofinal en θ es decir, el rango de g está acotado en θ . Por lo tanto existe un $\delta < \theta$ tal que $g : \eta \rightarrow \delta$, es decir, $g \in {}^\eta\delta$. De lo anterior se sigue que si wRz , entonces $w = j_H$ para alguna $j \in {}^\eta\delta$. En consecuencia el conjunto $\{w : wRz\} \leq \delta^\eta$, y como $\delta \leq 2^{2^{2^\eta}}$ (por definición de cardinal sucesor (ver [5], p. 76)), entonces $\delta^\eta < \theta$: En efecto, como $\delta \leq 2^{2^{2^\eta}}$, entonces por propiedades de potencias de cardinales (ver [13], p. 29) se tiene que $(\delta)^\eta \leq (2^{2^{2^\eta}})^\eta = 2^{(2^{2^\eta}) \cdot \eta} = 2^{(2^{2^\eta})} < \theta$. Esto muestra que cualquier ordinal de (\mathfrak{D}, R) tiene menos de θ predecesores. En consecuencia el conjunto de los ordinales de (\mathfrak{D}, R) tiene a lo sumo tipo de orden θ . Por lo tanto, el tipo de orden del conjunto de los ordinales de (\mathfrak{D}, R) es θ . Ha culminado la prueba de la proposición.

Sea $\phi(z)$ una fórmula de la teoría de conjuntos que afirma « z es el primer cardinal medible». Cuando se escribe en detalle dicha fórmula se puede apreciar que sus cuantificadores pueden ser restringidos a $P(P(P(z)))$ (ver [3], p. 239). En consecuencia, en el modelo $\langle H(\theta), \in \rangle$ un elemento a satisface $\phi(z)$ si y solamente si a es realmente el primer cardinal medible, $a = \eta$. De modo que en la ultrapotencia (\mathfrak{D}, R) el único elemento que satisface $\phi(z)$ es $d(\eta)$. El ordinal $d(\eta)$ es mayor que el η -ésimo ordinal de (\mathfrak{D}, R) ¿por qué? por el Lema 15 y considerando la estructura $\langle \eta, < \rangle$ se tiene que la ultrapotencia que se forma con la misma usando el ultrafiltro H no principal y η -completo sobre η , $(\mathfrak{C}, K) = \prod_H \langle \eta, \in \rangle$, es una estructura bien ordenada de tipo de orden mayor que η y para cualquier $\delta < \eta$, $d(\delta)$ es el δ -ésimo elemento de $\prod_H \langle \eta, \in \rangle$. Se cumple que $\prod_H \langle \eta, \in \rangle$ está inmersa en la ultrapotencia $(\mathfrak{D}, R) = \prod_H \langle H(\theta), \in \rangle$ ¿cuál es una inmersión? La siguiente: Para cada $f \in {}^\eta\eta$ se considera $f_H \in (\mathfrak{C}, K) = \prod_H \langle \eta, \in \rangle$ y $f_H \in (\mathfrak{D}, R) = \prod_H \langle H(\theta), \in \rangle$, dicha correspondencia es una inmersión. Como tal inmersión preserva el orden y la ultrapotencia (\mathfrak{C}, K) esta bien ordenada y su tipo de orden es mayor que η , entonces (\mathfrak{C}, K) es un segmento inicial de los ordinales de la ultrapotencia bien fundamentada (\mathfrak{D}, R) (todo ordinal z de las ultrapotencias $(\mathfrak{D}, R) = \prod_H \langle H(\theta), \in \rangle$ o $(\mathfrak{C}, K) = \prod_H \langle \eta, \in \rangle$ tiene la forma g_H , para alguna función $g : \eta \rightarrow \delta$). Por lo tanto existe una función $u \in {}^\eta\eta$ tal que la clase de equivalencia u_H es el η -ésimo ordinal de (\mathfrak{C}, R) y de (\mathfrak{D}, R) , es decir, u_H tiene η predecesores. Considerando a $d(\eta)$ en la ultrapotencia (\mathfrak{D}, R) se tiene que ella es la clase de equivalencia de la función constante en η : $\langle \eta : \delta \in \eta \rangle$. Si se llama a tal función constante c se puede reescribir la misma así: $c : \eta \rightarrow H(\theta)$ tal que para cada $\delta \in \eta$ ($c(\delta) = \eta$). Por otro lado, u_H en (\mathfrak{D}, R) es la clase de equivalencia de la función $u \in {}^\eta\eta$. Es decir, para cada $\delta \in \eta$ ($u(\delta) \in c(\delta) = \eta$). Por lo tanto, $\eta = \{\delta : u(\delta) \in c(\delta) = \eta\} \in H$. Entonces, por la definición de ultrapotencia, se cumple que: $u_H R d(\eta)$. En consecuencia las estructuras $\langle H(\theta), \in \rangle$ y (\mathfrak{D}, R) no son isomorfas, pues cualquier isomorfismo asigna η al η -ésimo ordinal de (\mathfrak{D}, R) , $\hat{\eta}$, y $\hat{\eta}$ no satisface $\phi(z)$. Por lo tanto, el axioma de constructibilidad es falso. Con esto termina la demostración del Teorema. \square

Problemas abiertos: Vale la pena destacar que algunos problemas abiertos de teoría de conjuntos y/o teoría de modelos relacionados con cardinales medibles, cardinales grandes y ultraproductos pueden encontrarse en los textos [16] y [3], p. 597-602 (Apéndice B)].

REFERENCIAS

- [1] Bell, J., *Infinitary Logic*, Enciclopedia de Filosofía de la universidad de Stanford, 2016, <https://plato.stanford.edu/entries/logic-infinitary/>
- [2] Corbillón, M., *Análisis real no estándar*, Tesis de licenciatura en Matemáticas, Tutor: Dr. Josep Maria Font Llovet, Facultat de Matemàtiques, Universitat de Barcelona, 2015.
- [3] Chang, C., Keisler, H., *Model Theory*, Dover Publications, 2012.
- [4] Di Prisco, C., *Introducción a la Lógica Matemática*, Emalca Amazonia, 2009.
- [5] Di Prisco, C., *Teoría de Conjuntos*, Universidad Central de Venezuela: Consejo de Desarrollo Científico y Humanístico, 2009.
- [6] Di Prisco, C., *Inmersiones elementales y grandes cardinales*, Notas no publicadas, 1982.
- [7] Enderton, H., *Una Introducción Matemática a la Lógica*, Universidad Nacional Autónoma de México, 2004.

- [8] Enderton, H., *Elements of Set Theory*, Academic Press, 1977.
- [9] Ebbinghaus, H., Flum, J., Thomas, W., *Mathematical Logic*, Springer, 1996.
- [10] Galindo, F., *Una presentación de la demostración directa del teorema de compacidad de la lógica de primer orden que usa el método de ultraproductos*, Una Investigación, Vol. VIII, N° 15 (2016).
- [11] Gödel, K., *Obras completas*, Alianza, 1981.
- [12] Hrbacek, K., Jech, T., *Introduction to set theory*, Marcel Dekker, 1999.
- [13] Jech, T., *Set Theory*, Springer, 2000.
- [14] Jech, T., *The Axiom of Choice*, Dover Publications, 2008.
- [15] Jech, T., *Set Theory*, Academic Press, 1978.
- [16] Kanamori, A., *The Higher Infinite. Large Cardinal in Set Theory from their Beginnings*, Springer, 1997.
- [17] Kunen, K., *Set Theory. An Introduction to Independence Proofs*, College Publications, 2011.
- [18] Loś, J., *Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres*, en «Mathematical Interpretation of Formal Systems» (T. Skolem et al., eds.), North-Holland Publishing Co., 1955, 98-113.
- [19] Manzano, M., *Teoría de Modelos*, Alianza, 1989.
- [20] Marks, A., *Set Theory*, Notas, Mathematics UCLA, 2020. (Estan en la web).
- [21] Mendelson, E., *Introduction to Mathematical Logic*, Chapman and Hall/CRL, 2009.
- [22] Scott, D., *Measurable cardinals and constructive sets*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astron. Phys. 9, 521-524, 1961.

Dirección del autor:

Universidad Central de Venezuela,
Facultad de Humanidades y Educación,
Escuela de Filosofía,
Departamento de Lógica y Filosofía de la Ciencia.
Venezuela.
Colaborador Visitante del Departamento de Matemáticas
del Instituto Venezolano de Investigaciones Científicas.
Venezuela.
e-mail: franklin.galindo@ucv.ve



OPERADORES SIMÉTRICOS MULTIVALUADOS Y EL PROBLEMA DE MOMENTOS DE HAMBURGER TRUNCADO

JOSUÉ I. RIOS-CANGAS

RESUMEN. La teoría de operadores lineales multivaluados generaliza la teoría clásica de operadores lineales y surge por la necesidad de trabajar con el adjunto de un operador lineal no densamente definido en el espacio. Una clase particular de los operadores multivaluados son los simétricos multivaluados y mediante estos operadores se da una caracterización de todas las medidas que resuelven el problema de momentos de Hamburger truncado.

1. INTRODUCCIÓN

Los operadores lineales multivaluados (o también llamados relaciones lineales) en un espacio de Hilbert \mathcal{H} son conjuntos lineales en $\mathcal{H} \oplus \mathcal{H}$, donde $\mathcal{H} \oplus \mathcal{H}$ representa la suma ortogonal de \mathcal{H} consigo mismo (ver [3]). De manera particular una relación lineal es la gráfica de un operador lineal y por esta razón las relaciones lineales se consideran como una extensión de los operadores lineales. Es de interés mencionar que von Neumann no solo fue el pionero en la teoría de extensiones de operadores, sino también en la teoría de relaciones lineales. Ciertamente la noción de relaciones lineales surge en [20] por la necesidad de trabajar con el adjunto de un operador no densamente definido, es decir, cuya cerradura de su dominio no coincide con \mathcal{H} . La teoría de relaciones lineales fue desarrollada posteriormente en [3, 6, 9] (citas más recientes se pueden encontrar en [4, 16]).

La importancia de estudiar las relaciones simétricas viene del estudio de los operadores simétricos, debido a que están íntimamente conectados con los sistemas conservativos. Estos sistemas se generalizan a sistemas en donde la energía no necesariamente se conserva en el tiempo (por ejemplo, los sistemas disipativos para los cuales la energía no aumenta). Cabe señalar que uno de los precursores de las relaciones simétricas fue sin duda R. Arens [3] y en [7, 12] se estudian aspectos importantes de la teoría de relaciones simétricas. De igual importancia en el estudio de la teoría de relaciones simétricas aparece el estudio de la teoría de extensiones simétricas, la cual abarca la teoría clásica de von Neumann sobre extensiones simétricas de operadores simétricos [19]. La teoría de extensiones simétricas para relaciones lineales tuvo origen en un trabajo de Dijksma y Snoo [9] (c.f. [1, 10]).

El objetivo principal de este artículo es mostrar la teoría de operadores multivaluados, particularmente la teoría de operadores multivaluados simétricos. En la Sección 2 se plantea la teoría de operadores lineales en espacios de Hilbert donde se repasan de manera sencilla algunos conceptos y resultados sobre: operadores cerrados, operadores acotados, el adjunto de un operador, operadores simétricos y autoadjuntos. Además se introducen ejemplos prácticos de interés al lector (Ejemplos 1.3). La sección continúa con la gráfica de un operador lineal para familiarizarse con la notación de operadores multivaluados.

La Sección 3.2 presenta la teoría de operadores multivaluados e inicia con el espectro de estos operadores y sus propiedades. Un operador multivaluado cerrado se puede descomponer canónicamente en su parte operador y su parte multivaluada (ver [9]). Esta descomposición permite dar similitudes entre los espectros de un

2010 *Mathematics Subject Classification.* 47A06; 47B25; 44A60.

Palabras clave. Operadores multivaluados, operadores simétricos no densamente definidos, el problema clásico de momentos.

operador multivaluado y de su parte operador (Lema [7](#) y Teorema [8](#)). De hecho bajo ciertas condiciones, el Teorema [9](#) muestra que es espectro de un operador multivaluado coincide con el espectro de su parte operador en cierto espacio de Hilbert. La sección prosigue con la teoría de operadores multivaluados simétricos, donde se muestran distintas caracterizaciones de estos operadores (Proposiciones [10](#), [11](#), [12](#) y Teorema [13](#)).

Es de interés señalar que el problema de momentos fue fundamental para el desarrollo del análisis en el período que comprende de 1894, cuando Stieltjes escribió sus famosas memorias [18](#), hasta la década de 1950 cuando Krein completó una serie de notas relacionadas con este tema (c.f. [17](#)). Básicamente el problema de momentos de Hamburger indica que si para una sucesión $\{s_j\}_{j \geq 0}$ de números reales, es posible encontrar una medida μ sobre \mathbb{R} , tal que

$$(1) \quad s_j = \int_{\mathbb{R}} \lambda^j d\mu(\lambda),$$

además de saber si la existencia de μ es única. En [17](#), [2](#) se muestran condiciones necesarias y suficientes para que el problema de momentos [1](#) sea determinado (cuando existe una única medida) o indeterminado (cuando existen infinitas medidas). Si la sucesión de momentos es finita, es decir, si $\{s_j\}_{j=0}^N$ se dice que el problema de momentos de Hamburger [1](#) es truncado y este problema tiene solución cuando N es un número par (c.f. [8](#), Sec. 10] y [15](#), Sec. 9.1]). Mediante la teoría de operadores multivaluados simétricos, la última sección de este trabajo se centra en caracterizar todas las medidas μ que resuelven el problema de momentos de Hamburger truncado.

2. PRELIMINARES

2.1. Operadores lineales. Sea $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert separable sobre el campo \mathbb{C} , con producto interno anti-lineal en su primer argumento y norma $\|f\| = \langle f, f \rangle^{1/2}$, con $f \in \mathcal{H}$.

Una aplicación $T: \text{dom } T \subset \mathcal{H} \rightarrow \mathcal{H}$ es un *operador lineal* (o simplemente operador) si para cada $f, g \in \text{dom } T$ y $\alpha, \beta \in \mathbb{C}$, se cumple que $T(\alpha f + \beta g) = \alpha T f + \beta T g$, donde

$$\begin{aligned} \text{dom } T &:= \{f \in \mathcal{H} : T f \in \mathcal{H}\} & \text{ran } T &:= \{T f : f \in \text{dom } T\} \\ \ker T &:= \{f \in \text{dom } T : T f = 0\} \end{aligned}$$

representan el *dominio*, *rango* y *núcleo* de T , respectivamente, y estos resultan ser conjuntos lineales en \mathcal{H} .

Para dos operadores lineales T, S y $\alpha, \beta \in \mathbb{C}$, el operador $\alpha T + \beta S$ tiene $\text{dom}(\alpha T + \beta S) = \text{dom } T \cap \text{dom } S$ y actúa como $(\alpha T + \beta S)f = \alpha T f + \beta S f$. El operador TS tiene $\text{dom}(TS) = \{f \in \text{dom } S : S f \in \text{dom } T\}$ y $TS f = T(S f)$. Además, $T|_{\mathcal{C}}$ es el operador T con dominio restringido a un conjunto lineal $\mathcal{C} \subset \text{dom } T$. En este sentido, $T \subset S$ si $S|_{\text{dom } T} = T$ y en este caso se dice que S es extensión de T . Por otra parte, T es *invertible* si tiene núcleo trivial y en cuyo caso la inversa se denota por T^{-1} .

Lo siguiente denota dos tipos de cerradura:

1. Un operador T es *cerrado* si y solo si para toda sucesión $\{x_n\}_{n \in \mathbb{N}} \subset \text{dom } T$ tal que $(x_n, T x_n) \rightarrow (x, y)$ se tiene $x \in \text{dom } T$ y $y = T x$.
2. Un operador T es *cerrable* si y solo si para toda sucesión $\{x_n\}_{n \in \mathbb{N}} \subset \text{dom } T$ tal que $(x_n, T x_n) \rightarrow (0, y)$ se tiene $y = 0$.

Cuando T es cerrado se escribe $T = \overline{T}$ y cuando T es cerrable se tiene que \overline{T} es un operador lineal. Es claro que el núcleo de un operador cerrado es cerrado.

Un operador T se llama *acotado* si existe $c > 0$ tal que $\|T f\| \leq c \|f\|$, para todo $f \in \text{dom } T$. Es fácil ver que un operador es acotado si y solo si es continuo. Además, la *norma* de un operador T viene dada por

$$\|T\| := \sup_{\substack{f \in \text{dom } T, \\ f \neq 0}} \frac{\|T f\|}{\|f\|},$$

la cual cumple las siguientes equivalencias:

$$\|T\| = \sup_{\substack{f \in \text{dom } T, \\ \|f\| < 1}} \|Tf\| = \sup_{\substack{f \in \text{dom } T, \\ \|f\| \leq 1}} \|Tf\| = \sup_{\substack{f \in \text{dom } T, \\ \|f\| = 1}} \|Tf\| .$$

Observación 1. Si un operador T cumple dos de las siguientes tres propiedades, entonces T cumple las tres propiedades (c.f. [5, Sec. 3.2]):

1. T es cerrado.
2. T es acotado.
3. $\text{dom } T$ es cerrado.

Denote por $\mathcal{B}(\mathcal{H})$ como el conjunto de todos los operadores acotados con dominio todo \mathcal{H} y por consiguiente cerrados. El conjunto $\mathcal{B}(\mathcal{H})$ es un espacio lineal normado y es de importancia mencionar que en este espacio se definen los siguientes tres tipos de convergencia.

Para $T, T_n \in \mathcal{B}(\mathcal{H})$, con $n \in \mathbb{N}$, se dice que:

1. $\{T_n\}$ converge uniformemente a T si $\|T - T_n\| \rightarrow 0$.
2. $\{T_n\}$ converge fuertemente a T si $\|Tf - T_n f\| \rightarrow 0$, para cada f en \mathcal{H} .
3. $\{T_n\}$ converge débilmente a T si $\langle g, Tf - T_n f \rangle \rightarrow 0$, para cada $f, g \in \mathcal{H}$.

La convergencia uniforme, fuerte y débil, se denotan por

$$u\text{-}\lim_{n \rightarrow \infty} T_n = T; \quad s\text{-}\lim_{n \rightarrow \infty} T_n = T; \quad w\text{-}\lim_{n \rightarrow \infty} T_n = T,$$

respectivamente. Además, la convergencia uniforme implica la convergencia fuerte y la convergencia fuerte implica la convergencia débil.

Ahora bien, un operador es densamente definido si tiene dominio denso en \mathcal{H} . Para un operador densamente definido T se denota su *adjunta* como T^* , con dominio

$$(2) \quad \text{dom } T^* = \{h \in \mathcal{H} : \exists k \in \mathcal{H} \text{ tal que } \langle Tf, h \rangle = \langle f, k \rangle, \forall f \in \text{dom } T\},$$

de manera que $T^*h = k$. El adjunto T^* es un operador lineal cerrado en \mathcal{H} (c.f. [14, Sec. 1.2]) y puede tener dominio no denso en \mathcal{H} (ver Ejemplo 3).

Permita introducir una parte de la teoría de operadores simétricos (ver [14, Secs. 3.1 y 3.2] para más referencia).

Un operador T es *simétrico* o *hermitiano* si su forma sesquilineal es real, es decir, si $\langle f, Tf \rangle \in \mathbb{R}$, para todo $f \in \text{dom } T$. Esto es equivalente a decir que $\langle f, Tg \rangle = \langle Tf, g \rangle$, para todo $f, g \in \text{dom } T$.

Los operadores simétricos son caracterizados por tener espectro real. Además, un operador densamente definido T es simétrico si $T \subset T^*$ y autoadjunto si $T = T^*$.

Ejemplo 1. El operador identidad $If = f$ y el operador cero $Of = 0$, para todo $f \in \mathcal{H}$, son ejemplos sencillos de operadores en $\mathcal{B}(\mathcal{H})$, con norma $\|I\| = 1$, $\|O\| = 0$ y adjunto $I^* = I$, $O^* = O$, es decir, son autoadjuntos.

Ejemplo 2. Sea $\mathcal{H} = L_2(\mathbb{R})$ el espacio de las funciones cuadrado integrable en \mathbb{R} , con producto interno $\langle f, g \rangle = \int \overline{f(t)}g(t)dt$. Considere el operador de multiplicación $Jf(t) = tf(t)$ cuyo dominio viene dado por

$$\text{dom } J = \left\{ f \in L_2(\mathbb{R}) : \int (1 + t^2) |f(t)|^2 dt < \infty \right\},$$

el cual es denso en $L_2(\mathbb{R})$.

El operador J es autoadjunto y no acotado. En efecto,

$$\langle Jf, f \rangle = \int \overline{tf(t)}f(t)dt = \int \overline{f(t)}tf(t)dt = \langle f, Jf \rangle, \quad \forall f \in \text{dom } J$$

de donde se sigue que $J \subset J^*$. Luego, si $g \in \text{dom } J^*$, entonces $\langle Jf, g \rangle = \langle f, J^*g \rangle$, para todo $f \in \text{dom } J$, que verifica $tg(t) = J^*g(t) \in L_2(\mathbb{R})$. De este modo, $g \in \text{dom } J$ y por

lo tanto $J^* = J$. Continuando, defina para $n \in \mathbb{N}$

$$f_n(t) = \begin{cases} 1, & \text{si } n \leq t < n+1 \\ 0, & \text{en otro caso} \end{cases}$$

Así, $\|f_n\| = 1$ y uno calcula que $\|Jf_n\| > n \rightarrow \infty$. Por lo tanto, J es no acotado.

Ejemplo 3. Considere un conjunto lineal \mathcal{D} en \mathcal{H} y $0 \neq g \in \mathcal{H}$. Sea \mathcal{F} un funcional lineal no continuo sobre \mathcal{D} y defina el operador lineal T con $\text{dom } T = \mathcal{D}$, tal que $Tf = \mathcal{F}(f)g$.

El operador T no es acotado ni cerrable. Ciertamente, como \mathcal{F} no es continuo existe $\{f_n\}_{n \in \mathbb{N}} \subset \mathcal{D}$, con $\|f_n\| = 1$, tal que $0 \neq |\mathcal{F}(f_n)| \rightarrow \infty$. Así, $\|Tf_n\| = |\mathcal{F}(f_n)| \|g\| \rightarrow \infty$, es decir, T no es acotado. Además, haciendo $h_n = \mathcal{F}(f_n)^{-1} f_n$ se tiene que $h_n \rightarrow 0$ y $Th_n = \mathcal{F}(h_n)g = g \neq 0$, es decir, T no es cerrable.

Si \mathcal{D} es denso en \mathcal{H} , entonces T^ no es densamente definido y $T^* = O_{\{g\}^\perp}$.* Desde luego, como el producto interno es continuo y \mathcal{F} es discontinuo, la aplicación $f \rightarrow \langle h, Tf \rangle = \mathcal{F}(f) \langle h, g \rangle$ es continua si y solo si $h \perp g$. Por lo tanto, $\text{dom } T^* = \{g\}^\perp$ y $T^*h = 0$, para todo $h \in \text{dom } T^*$.

2.2. La gráfica de un operador. Este tema se aborda de manera análoga a [5, Cap. 3]. Sea $\mathcal{H} \oplus \mathcal{H}$ la suma ortogonal de dos copias de \mathcal{H} (c.f. [5, Sec. 2.3]), es decir,

$$(3) \quad \mathcal{H} \oplus \mathcal{H} := \left\{ \begin{pmatrix} f \\ g \end{pmatrix} : f, g \in \mathcal{H} \right\},$$

que es un espacio de Hilbert con producto interno

$$\left\langle \begin{pmatrix} f \\ g \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle = \langle f, h \rangle + \langle g, k \rangle, \quad \begin{pmatrix} f \\ g \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \in \mathcal{H} \oplus \mathcal{H}.$$

La convergencia en $\mathcal{H} \oplus \mathcal{H}$ implica la convergencia en cada una de sus entradas.

La gráfica de un operador lineal T viene dada por

$$\mathcal{G}(T) := \left\{ \begin{pmatrix} f \\ Tf \end{pmatrix} \in \mathcal{H} \oplus \mathcal{H} : f \in \text{dom } T \right\},$$

la cual resulta ser un conjunto lineal en $\mathcal{H} \oplus \mathcal{H}$. Sin embargo no todo conjunto lineal en $\mathcal{H} \oplus \mathcal{H}$ resulta ser la gráfica de un operador lineal.

PROPOSICIÓN 1. *Un conjunto lineal $\mathcal{G} \subset \mathcal{H} \oplus \mathcal{H}$ es la gráfica de un operador lineal si y solo si*

$$(4) \quad \left\{ \begin{pmatrix} f \\ g \end{pmatrix} \in \mathcal{G} : f = 0 \right\} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}.$$

Demostración. Suponga que (4) se cumple y considere las aplicaciones lineales $\pi, \rho: \mathcal{G} \rightarrow \mathcal{H}$ tales que

$$\pi \begin{pmatrix} f \\ g \end{pmatrix} = f \quad ; \quad \rho \begin{pmatrix} f \\ g \end{pmatrix} = g, \quad \begin{pmatrix} f \\ g \end{pmatrix} \in \mathcal{G}.$$

De este modo, $\ker \pi = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ y se verifica que $T = \rho\pi^{-1}$ es un operador lineal con $\text{dom } T = \pi\mathcal{G}$. Por lo tanto $\mathcal{G}(T) = \mathcal{G}$. La prueba inversa es directa. \square

Para efectos prácticos, se introducen las aplicaciones $\mathbb{U}, \mathbb{W}: \mathcal{H} \oplus \mathcal{H} \rightarrow \mathcal{H} \oplus \mathcal{H}$, que actúan como

$$\mathbb{U} \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} g \\ f \end{pmatrix} \quad ; \quad \mathbb{W} \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} g \\ -f \end{pmatrix}, \quad \begin{pmatrix} f \\ g \end{pmatrix} \in \mathcal{H} \oplus \mathcal{H}$$

las cuales son operadores lineales que satisfacen $\mathbb{U}^2 = I = -\mathbb{W}^2$ y $\mathbb{U}\mathbb{W} = -\mathbb{W}\mathbb{U}$.

Para dos operadores T, S , lo siguiente se cumple de manera sencilla:

- (a) $T \subset S$ si y solo si $\mathcal{G}(T) \subset \mathcal{G}(S)$.
- (b) T es cerrado si y solo si $\mathcal{G}(T)$ es cerrada.
- (c) Si T es invertible entonces $\mathcal{G}(T^{-1}) = \mathbb{U}\mathcal{G}(T)$.

Para un conjunto lineal $\mathcal{G} \subset \mathcal{H} \oplus \mathcal{H}$, denote el conjunto lineal

$$-\mathcal{G} := \left\{ \begin{pmatrix} f \\ -g \end{pmatrix} \in \mathcal{H} \oplus \mathcal{H} : \begin{pmatrix} f \\ g \end{pmatrix} \in \mathcal{G} \right\}.$$

PROPOSICIÓN 2. Si $\mathcal{G} \subset \mathcal{H} \oplus \mathcal{H}$ es un conjunto lineal, entonces:

(a) $(\mathbb{W}\mathcal{G})^\perp = \mathbb{W}(\mathcal{G}^\perp)$.

(b) $\overline{\mathbb{W}\mathcal{G}} = \mathbb{W}\overline{\mathcal{G}}$.

(c) $(\mathbb{U}\mathcal{G})^\perp = \mathbb{U}(\mathcal{G}^\perp)$.

(d) $\overline{\mathbb{U}\mathcal{G}} = \mathbb{U}\overline{\mathcal{G}}$.

(e) $\mathbb{U}^2\mathcal{G} = \mathbb{W}^2\mathcal{G} = \mathcal{G}$.

(f) $\mathbb{U}\mathbb{W}\mathcal{G} = \mathbb{W}\mathbb{U}\mathcal{G} = -\mathcal{G}$.

Demostración. Note que $\begin{pmatrix} f \\ g \end{pmatrix} \in (\mathbb{W}\mathcal{G})^\perp$ si y solo si $\forall \begin{pmatrix} h \\ k \end{pmatrix} \in \mathbb{W}\mathcal{G}, \begin{pmatrix} k \\ -h \end{pmatrix} \in \mathcal{G}$ y

$$\left\langle \begin{pmatrix} g \\ -f \end{pmatrix}, \begin{pmatrix} k \\ -h \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} f \\ g \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle = 0,$$

si y solo si $\begin{pmatrix} g \\ -f \end{pmatrix} \in (\mathcal{G})^\perp$, o bien $\begin{pmatrix} f \\ g \end{pmatrix} \in \mathbb{W}(\mathcal{G})^\perp$. Esto prueba el punto (a). Ahora,

$$\overline{\mathbb{W}\mathcal{G}} = \left((\mathbb{W}\mathcal{G})^\perp \right)^\perp = \mathbb{W} \left((\mathcal{G}^\perp)^\perp \right) = \mathbb{W}\overline{\mathcal{G}},$$

es decir, (b). Los puntos (c), (d) se siguen de manera análoga y (e), (f) se siguen directamente de la definición, notando que \mathcal{G} es conjunto lineal. \square

Los operadores \mathbb{U}, \mathbb{W} y sus propiedades serán de gran utilidad en la secuela.

3. OPERADORES LINEALES MULTIVALUADOS

Esta sección presenta una manera de extender la teoría espectral de operadores lineales en espacios de Hilbert.

3.1. Espectro de un operador multivaluado. Se llama *operador lineal multivaluado* o *relación lineal* a un conjunto lineal T en $\mathcal{H} \oplus \mathcal{H}$ (ver (3)), con

$$(5) \quad \begin{aligned} \text{dom } T &:= \left\{ f \in \mathcal{H} : \begin{pmatrix} f \\ g \end{pmatrix} \in T \right\} & \text{ran } T &:= \left\{ g \in \mathcal{H} : \begin{pmatrix} f \\ g \end{pmatrix} \in T \right\} \\ \text{ker } T &:= \left\{ f \in \mathcal{H} : \begin{pmatrix} f \\ 0 \end{pmatrix} \in T \right\} & \text{mul } T &:= \left\{ g \in \mathcal{H} : \begin{pmatrix} 0 \\ g \end{pmatrix} \in T \right\} \end{aligned}$$

los cuales son conjuntos lineales en \mathcal{H} . La notación (5) es habitual en el caso de operadores lineales a excepción de $\text{mul } T$, que representa el *multivaluado* de T .

Una relación T es *cerrada* si $T = \overline{T}$, la cual es subespacio (conjunto lineal cerrado) en $\mathcal{H} \oplus \mathcal{H}$. Uno verifica fácilmente que si T es cerrada, entonces $\text{ker } T$ y $\text{mul } T$ son subespacios de \mathcal{H} .

Observación 2. La gráfica de un operador lineal es un caso particular de una relación. Una consecuencia de la Proposición 1 es que una relación T es la gráfica de un operador si y solo si $\text{mul } T = 0$. En el contexto de operadores multivaluados se identifica a cada operador con su gráfica.

Para dos relaciones T, S y $\zeta \in \mathbb{C}$, se consideran las relaciones:

$$\begin{aligned} T + S &:= \left\{ \begin{pmatrix} f \\ g+h \end{pmatrix} : \begin{pmatrix} f \\ g \end{pmatrix} \in T, \begin{pmatrix} f \\ h \end{pmatrix} \in S \right\} & \zeta T &:= \left\{ \begin{pmatrix} f \\ \zeta g \end{pmatrix} : \begin{pmatrix} f \\ g \end{pmatrix} \in T \right\} \\ ST &:= \left\{ \begin{pmatrix} f \\ k \end{pmatrix} : \begin{pmatrix} f \\ g \end{pmatrix} \in T, \begin{pmatrix} g \\ k \end{pmatrix} \in S \right\} & T^{-1} &:= \left\{ \begin{pmatrix} g \\ f \end{pmatrix} : \begin{pmatrix} f \\ g \end{pmatrix} \in T \right\} \end{aligned}$$

que representan la suma de relaciones, multiplicación por escalar, composición de relaciones e inversa de una relación, respectivamente. Note que $T^{-1} = \mathbb{U}T$ y esto extiende la noción de la inversa de un operador lineal.

Dos relaciones T y S son linealmente independientes si $T \cap S = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$. Además se asume que los símbolos $\dot{+}$, \oplus , y \ominus representan la notación estándar, es decir,

$$\begin{aligned} T \dot{+} S &= \left\{ \begin{pmatrix} f+h \\ g+k \end{pmatrix} : \begin{pmatrix} f \\ g \end{pmatrix} \in T, \begin{pmatrix} h \\ k \end{pmatrix} \in S \text{ y } T \cap S = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \right\} . \\ T \oplus S &= T \dot{+} S, \text{ con } T \subset S^\perp . \\ T \ominus S &= T \cap S^\perp . \end{aligned}$$

Hablando de manera estricta, el símbolo \oplus en este contexto difiere de la expresión $\mathcal{H} \oplus \mathcal{H}$. Esto no debería causar confusión al usar el mismo símbolo.

Se denota la *adjunta* de una relación T como la relación lineal

$$T^* := \left\{ \begin{pmatrix} h \\ k \end{pmatrix} \in \mathcal{H} \oplus \mathcal{H} : \langle h, g \rangle = \langle k, f \rangle, \quad \forall \begin{pmatrix} f \\ g \end{pmatrix} \in T \right\} .$$

Este concepto extiende al de [2], sin pedir restricción sobre la densidad del dominio.

TEOREMA 3. *Para relaciones T, S y $0 \neq \alpha \in \mathbb{C}$, lo siguiente se cumple:*

- (a) $T^* = (\mathbb{W}T)^\perp$
- (b) $T^{**} = \overline{T}$
- (c) $(T^*)^{-1} = (T^{-1})^*$
- (d) $(\alpha T)^* = \overline{\alpha} T^*$
- (e) $S \subset T \Rightarrow T^* \subset S^*$
- (f) $\ker T^* = (\text{ran } T)^\perp$

Demostración. Note que $\begin{pmatrix} h \\ k \end{pmatrix} \in T^*$ si y solo si para todo $\begin{pmatrix} f \\ g \end{pmatrix} \in T$, $\langle k, f \rangle = \langle h, g \rangle$, es decir, $\left\langle \begin{pmatrix} h \\ k \end{pmatrix}, \begin{pmatrix} -g \\ f \end{pmatrix} \right\rangle = 0$, esto si y solo si $\begin{pmatrix} h \\ k \end{pmatrix} \in (\mathbb{W}T)^\perp$ y por lo tanto [a]. Luego, se sigue de la Proposición 2 que

$$T^{**} = \left(\mathbb{W}(\mathbb{W}T)^\perp \right)^\perp = (T^\perp)^\perp = \overline{T}$$

que produce [b]. Además, $(T^*)^{-1} = \mathbb{U}(\mathbb{W}T)^\perp = (\mathbb{W}U)^\perp = (T^{-1})^*$, es decir, [c]. Es simple calcular por contención que $(\alpha T)^* \subset \overline{\alpha} T^*$ y como $\alpha \neq 0$, se tiene que $\overline{\alpha} T^* = \overline{\alpha} (\alpha^{-1} \alpha T)^* \subset \overline{\alpha} \overline{\alpha^{-1}} (\alpha T)^*$ e implica [d]. Si $S \subset T$ entonces $\mathbb{W}S \subset \mathbb{W}T$. Así, $(\mathbb{W}T)^\perp \subset (\mathbb{W}S)^\perp$ y se tiene [e]. Por último, $\begin{pmatrix} f \\ 0 \end{pmatrix} \in T^*$ si y solo si $\langle f, k \rangle = \langle 0, h \rangle = 0$, para todo $\begin{pmatrix} h \\ k \end{pmatrix} \in T$, esto si y solo si $f \in (\text{ran } T)^\perp$, de donde se llega a [f]. \square

La igualdad [a] implica que T^* es una relación cerrada. Además se sigue de las igualdades [c] y [f] que

$$(6) \quad \overline{\text{dom } T} = \overline{\text{ran } T^{-1}} = \left(\ker (T^{-1})^* \right)^\perp = (\text{mul } T^*)^\perp ,$$

es decir, si un operador no está densamente definido en \mathcal{H} entonces su adjunta tiene multivaluado no trivial. De esto proviene la condición de densidad en [2].

Observación 3. Dos conjuntos lineales cerrados G, L que son linealmente independientes satisfacen

$$G \cap L = (G^\perp \dot{+} L^\perp)^\perp .$$

Ciertamente, si $f \in G \cap L$, entonces $\langle f, h \rangle = 0$ para todo $h \in G^\perp \dot{+} L^\perp$, lo cual produce $G \cap L \subset (G^\perp \dot{+} L^\perp)^\perp$. Inversamente, note que $G^\perp, L^\perp \subset G^\perp \dot{+} L^\perp$ y como los conjuntos son cerrados, $(G^\perp \dot{+} L^\perp)^\perp \subset G, L$, es decir, $(G^\perp \dot{+} L^\perp)^\perp \subset G \cap L$.

COROLARIO 4. *Si S, T son relaciones cerradas linealmente independientes entonces*

$$(S \dot{+} T)^* = S^* \cap T^* .$$

Demostración. Del punto (a) del Teorema 3 y de la Proposición 2, uno calcula de manera sencilla que

$$\begin{aligned} (S \dot{+} T)^* &= \mathbb{W}(S \dot{+} T)^\perp = (\mathbb{W}S \dot{+} \mathbb{W}T)^\perp \\ &= (S^{*\perp} \dot{+} T^{*\perp})^\perp = S^* \cap T^*, \end{aligned}$$

la última igualdad se sigue de la Observación 3. \square

Una relación es *acotada* si resulta ser un operador lineal acotado. Cabe aclarar que en el contexto de relaciones lineales existen diferentes maneras de definir relaciones acotadas (ver por ejemplo 7).

Se denota el conjunto *cuasi-regular* de una relación T como

$$\hat{\rho}(T) := \{\zeta \in \mathbb{C} : (T - \zeta I)^{-1} \text{ es acotada}\}.$$

PROPOSICIÓN 5. *El conjunto cuasi-regular de una relación es abierto.*

Demostración. Para $\zeta \in \hat{\rho}(T)$ se tiene la existencia de $c_\zeta > 0$ tal que

$$(7) \quad \|k\| \leq c_\zeta \|h\|, \quad \forall \begin{pmatrix} h \\ k \end{pmatrix} \in (T - \zeta I)^{-1}.$$

Considere la bola abierta $\mathbb{B}_{c_\zeta^{-1}}(\zeta)$ y $\eta \in \mathbb{B}_{c_\zeta^{-1}}(\zeta)$. Uno calcula de manera sencilla que para $\begin{pmatrix} f \\ g \end{pmatrix} \in (T - \eta I)^{-1}$ se tiene $\begin{pmatrix} f + (\eta - \zeta)g \\ g \end{pmatrix} \in (T - \zeta I)^{-1}$. Así de 7,

$$\frac{1}{c_\zeta} \|g\| \leq \|f + (\eta - \zeta)g\| \leq \|f\| + |\eta - \zeta| \|g\|.$$

De esta manera, $\|g\| \leq c' \|f\|$ con $c'^{-1} = c_\zeta^{-1} - |\eta - \zeta| > 0$ y por consiguiente $(T - \eta I)^{-1}$ es acotada. Por lo tanto $\mathbb{B}_{c_\zeta^{-1}}(\zeta) \subset \hat{\rho}(T)$. \square

Para $\zeta \in \hat{\rho}(T)$ se tiene que $\text{ran}(T - \zeta I)$ es cerrado si y solo si T es cerrada. Además, $\dim(\text{ran}(T - \zeta I)^\perp)$ es contante en cada componente conexa de $\hat{\rho}(T)$ [13, Sec. 2].

El conjunto *regular* de una relación T se denota como

$$\rho(T) := \{\zeta \in \mathbb{C} : (T - \zeta I)^{-1} \in \mathcal{B}(\mathcal{H})\}.$$

Claramente el conjunto regular está contenido en el cuasi-regular y también es abierto. Además, si una relación no es cerrada entonces su conjunto regular es vacío.

Considere los siguientes conjuntos:

$$\begin{aligned} \sigma(T) &:= \mathbb{C} \setminus \rho(T) && (\text{espectro}) \\ \hat{\sigma}(T) &:= \mathbb{C} \setminus \hat{\rho}(T) && (\text{núcleo espectral}) \\ \sigma_p(T) &:= \{\zeta \in \mathbb{C} : \ker(T - \zeta I) \neq \{0\}\} && (\text{espectro puntual}) \\ \sigma_p^\infty(T) &:= \{\zeta \in \sigma_p(T) : \dim \ker(T - \zeta I) = \infty\} && (\text{espectro puntual no discreto}) \\ \sigma_c(T) &:= \left\{ \zeta \in \mathbb{C} : \text{ran}(T - \zeta I) \neq \overline{\text{ran}(T - \zeta I)} \right\} && (\text{espectro continuo}) \end{aligned}$$

TEOREMA 6. *Si T es una relación cerrada entonces $\sigma_p(T) \cup \sigma_c(T) = \hat{\sigma}(T)$.*

Demostración. Si $\zeta \in \hat{\rho}(T)$ entonces $\text{ran}(T - \zeta I)$ es cerrado y además

$$(8) \quad \ker(T - \zeta I) = \text{mul}(T - \zeta I)^{-1} = \{0\},$$

esto implica que $\zeta \notin \sigma_p(T) \cup \sigma_c(T)$. Por otra parte, si $\zeta \notin \sigma_p(T) \cup \sigma_c(T)$ el mismo cálculo de 8 muestra que $(T - \zeta I)^{-1}$ es un operador cerrado, ya que T es cerrada, con dominio cerrado y por lo tanto acotado (ver Observación 1), es decir, $\zeta \in \hat{\rho}(T)$. \square

Observación 4. Como en el caso de operadores lineales en espacios de Hilbert de dimensión finita, las relaciones lineales son cerradas y sus espectros puntual no discreto y continuo son vacíos. Además, el núcleo espectral coincide con el espectro.

Es conveniente descomponer a una relación cerrada T como $T = T_{\odot} \oplus T_{\infty}$, donde

$$(9) \quad T_{\infty} := \left\{ \begin{pmatrix} 0 \\ g \end{pmatrix} \in T \right\} \quad \text{y} \quad T_{\odot} := T \ominus T_{\infty}$$

son cerradas llamadas la *parte multivaluada* y la *parte operador* de T , respectivamente. Note que $\text{ran } T_{\odot} \perp \text{ran } T_{\infty} = \text{mul } T$ y $\text{dom } T = \text{dom } T_{\odot}$.

LEMA 7. Si T es una relación cerrada, entonces $\hat{\rho}(T) \subset \hat{\rho}(T_{\odot})$.

Demostración. Uno verifica de manera sencilla que $(T_{\odot} - \zeta I)^{-1} \subset (T - \zeta I)^{-1}$, para $\zeta \in \mathbb{C}$. Así, si $\zeta \in \hat{\rho}(T)$ entonces $(T - \zeta I)^{-1}$ es acotada al igual que $(T_{\odot} - \zeta I)^{-1}$, es decir, $\zeta \in \hat{\rho}(T_{\odot})$. \square

Es simple calcular que para una relación cerrada T con $\text{dom } T \subset (\text{mul } T)^{\perp}$,

$$(10) \quad T - \zeta I = (T_{\odot} - \zeta I) \oplus T_{\infty}, \quad \zeta \in \mathbb{C}.$$

TEOREMA 8. Si T es una relación cerrada tal que $\text{dom } T \subset (\text{mul } T)^{\perp}$, entonces $\hat{\rho}(T) = \hat{\rho}(T_{\odot})$.

Demostración. Es suficiente mostrar debido al Lema 7 que $\hat{\rho}(T_{\odot}) \subset \hat{\rho}(T)$. Esto se sigue directamente si $(T - \zeta I)^{-1}$ es acotada, para todo $\zeta \in \hat{\rho}(T_{\odot})$. Note de (10) que

$$(T - \zeta I)^{-1} = (T_{\odot} - \zeta I)^{-1} \oplus (T_{\infty})^{-1}.$$

De este modo, para $\begin{pmatrix} h \\ k \end{pmatrix} \in (T - \zeta I)^{-1}$, existen $\begin{pmatrix} r \\ k \end{pmatrix} \in (T_{\odot} - \zeta I)^{-1}$ y $\begin{pmatrix} s \\ 0 \end{pmatrix} \in (T_{\infty})^{-1}$ tales que $h = r + s$. Como $\zeta \in \hat{\rho}(T_{\odot})$, existe $C > 0$ tal que $\|k\| \leq C \|r\|$. Así,

$$\begin{aligned} \|k\| &\leq C (\|r\| + \|s\|) \\ &= C \|r + s\| = C \|h\|. \end{aligned}$$

Por lo tanto $(T - \zeta I)^{-1}$ es acotada. \square

Observación 5. La propiedad del Teorema 8 no siempre se cumple para el conjunto regular de una relación. Ciertamente, para A autoadjunta se tiene $\text{dom } A \subset (\text{mul } A)^{\perp}$ (ver Proposición 11). Si $\text{mul } A \neq \{0\}$ entonces $\text{dom}(A_{\odot} - \zeta I)^{-1} \neq \mathcal{H}$, para todo $\zeta \in \mathbb{C}$ y por ende $\rho(A_{\odot}) = \emptyset$. Sin embargo $\mathbb{C} \setminus \mathbb{R} \subset \rho(A)$ (consecuencia del Teorema 13).

El siguiente resultado se presenta con efecto de interés al lector, el cual puede encontrarse en [13, Teo. 2.10].

TEOREMA 9. Si T es una relación cerrada con $\text{dom } T \subset (\text{mul } T)^{\perp}$, entonces:

$$\begin{aligned} \hat{\sigma}(T) &= \hat{\sigma}(T_T) & \sigma_c(T) &= \sigma_c(T_T) \\ \sigma(T) &= \sigma(T_T) & \sigma_p(T) &= \sigma_p(T_T) \end{aligned}$$

donde $T_T = T_{\odot}$ en el espacio de Hilbert $(\text{mul } T)^{\perp} \oplus (\text{mul } T)^{\perp}$.

La afirmación anterior significa que las propiedades espectrales de una relación T con $\text{dom } T \subset (\text{mul } T)^{\perp}$, son las mismas que las de su parte operador T_{\odot} visto en el espacio $(\text{mul } T)^{\perp} \oplus (\text{mul } T)^{\perp}$.

3.2. Operadores multivaluados simétricos. Esta parte extiende el concepto de operador simétrico. Particularmente el concepto de operador autoadjunto.

Una relación A es llama *simétrica* si $A \subset A^*$ y *autoadjunta* si $A = A^*$. Los siguientes resultados exhiben caracterizaciones de las relaciones simétricas.

PROPOSICIÓN 10. Para una relación lineal A , los siguientes son equivalentes:

- (a) A es simétrica.
- (b) Para cada $\begin{pmatrix} f \\ g \end{pmatrix}$ y $\begin{pmatrix} h \\ k \end{pmatrix}$ en A , se cumple que $\langle f, k \rangle = \langle g, h \rangle$.
- (c) Para cada $\begin{pmatrix} f \\ g \end{pmatrix}$ en A , se tiene que $\langle f, g \rangle \in \mathbb{R}$.

Demostración. (a) \Rightarrow (b) Si $\begin{pmatrix} f \\ g \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \in A$ entonces $\begin{pmatrix} h \\ k \end{pmatrix} \in A^*$, es decir, $\langle h, g \rangle = \langle k, f \rangle$.
 (b) \Rightarrow (c) Para $\begin{pmatrix} f \\ g \end{pmatrix} \in A$, se tiene $\langle f, g \rangle = \langle g, f \rangle$. Por lo tanto $\langle f, g \rangle \in \mathbb{R}$. (c) \Rightarrow (a)
 Si $\begin{pmatrix} f \\ g \end{pmatrix} \in A$ entonces para todo $\begin{pmatrix} h \\ k \end{pmatrix} \in A$ y $\alpha \in \mathbb{C}$, se tiene que $\begin{pmatrix} f \\ g \end{pmatrix} + \alpha \begin{pmatrix} h \\ k \end{pmatrix} \in A$ y $\langle f + \alpha h, g + \alpha k \rangle \in \mathbb{R}$. Como $\langle f, g \rangle, \langle h, k \rangle \in \mathbb{R}$, se sigue que

$$\alpha \langle f, k \rangle + \bar{\alpha} \langle h, g \rangle = \bar{\alpha} \langle k, f \rangle + \alpha \langle g, h \rangle .$$

Así, para $\alpha \in \{1, i\}$, uno llega a $\text{Im} \langle f, k \rangle = \text{Im} \langle g, h \rangle$ y $\text{Re} \langle f, k \rangle = \text{Re} \langle g, h \rangle$, es decir, $\langle f, k \rangle = \langle g, h \rangle$. Por lo tanto $\begin{pmatrix} f \\ g \end{pmatrix} \in A^*$. \square

Observación 6. Existe otra caracterización de relaciones simétricas con respecto a su conjunto cuasi-regular. A saber, A es simétrica si y solo si $\mathbb{C} \setminus \mathbb{R} \subset \hat{\rho}(A)$ y

$$\|(A - \zeta I)^{-1}\| \leq 1/|\text{Im} \zeta| ,$$

para todo $\zeta \in \mathbb{C} \setminus \mathbb{R}$ [13, Obs. 3.2].

PROPOSICIÓN 11. Si A es una relación simétrica entonces $\overline{\text{dom} A} \subset (\text{mul} A)^\perp$.

Demostración. Si $A \subset A^*$ entonces $\text{mul} A \subset \text{mul} A^*$ y tomando complemento ortogonal se llega a $(\text{mul} A^*)^\perp \subset (\text{mul} A)^\perp$. Por lo tanto de (6) se sigue la afirmación. \square

Como consecuencia del resultado anterior, una relación simétrica cerrada cumple las propiedades espectrales del Teorema 9.

PROPOSICIÓN 12. Si A es una relación simétrica cerrada con dominio todo el espacio, entonces A es un operador autoadjunto y acotado.

Demostración. Como $\text{dom} A = \mathcal{H}$ se sigue de (6) y la Observación 2 que A^* es un operador y, además, $A = A^*$. Por último, A es un operador cerrado con dominio cerrado y por tanto acotado. \square

La sección concluye con la siguiente caracterización de relaciones autoadjuntas.

TEOREMA 13. Una relación es autoadjunta si y solo si es simétrica cerrada con espectro real.

Demostración. Si $A = A^*$ entonces es simétrica y cerrada. Ahora, para $\zeta \in \hat{\rho}(A) \setminus \mathbb{R}$ de la Observación 6 se tiene que $\bar{\zeta} \in \hat{\rho}(A) \setminus \mathbb{R}$ y $(A - \bar{\zeta} I)^{-1}$ es un operador. De este modo como A es cerrada,

$$\mathcal{H} = (\text{mul}(A - \bar{\zeta} I)^{-1})^\perp = (\ker(A - \bar{\zeta} I))^\perp = \text{ran}(A - \zeta I) ,$$

es decir, $\zeta \in \rho(A)$. Como $\hat{\rho}(A)$ es abierto y $\dim(\text{ran}(A - \zeta I))^\perp$ permanece constante en componentes conexas de $\hat{\rho}(A)$, si $\zeta \in \hat{\rho}(A) \cap \mathbb{R}$, entonces $\zeta \in \rho(A)$. Por lo tanto, $\hat{\rho}(A) = \rho(A)$, que implica $\sigma(A) \subset \mathbb{R}$.

Inversamente, si $A \subset A^*$ es cerrada con espectro real, entonces $\pm i \in \rho(A)$ que implica $\text{ran}(A - iI) = \mathcal{H}$ y $\ker(A^* - iI) = \{0\}$. Por consiguiente, si $\begin{pmatrix} f \\ g \end{pmatrix} \in A^*$ entonces existe $\begin{pmatrix} h \\ g - if \end{pmatrix} \in (A - iI)$, es decir,

$$(11) \quad \begin{pmatrix} h \\ g - i(f - h) \end{pmatrix} \in A \subset A^* .$$

Por linealidad, $\begin{pmatrix} f - h \\ i(f - h) \end{pmatrix} \in A^*$ que equivale a $f - h \in \ker(A^* - iI)$. Por lo tanto $f = h$ y de (11) se obtiene que $\begin{pmatrix} f \\ g \end{pmatrix} \in A$. \square

El teorema anterior es equivalente a decir que una relación A es autoadjunta si y solo si es simétrica cerrada, con $\hat{\rho}(A) = \rho(A)$.

4. EL PROBLEMA DE MOMENTOS DE HAMBURGER TRUNCADO

Para $n \in \mathbb{N}$, considere $\{s_j\}_{j=0}^{2n} \subset \mathbb{R}$, con $s_0 = 1$, tal que

$$(12) \quad \sum_{j,k=0}^n s_{j+k} \bar{\alpha}_j \alpha_k > 0,$$

para toda sucesión no trivial $\{\alpha_j\}_{j=0}^n \subset \mathbb{C}$.

Problema 1. Resolver si existe una medida μ sobre \mathbb{R} , tal que

$$s_j = \int_{\mathbb{R}} \lambda^j d\mu(\lambda), \quad j = 0, 1, \dots, 2n.$$

El problema anterior se le conoce como el problema de momentos de Hamburger truncado para $2n$ -sucesiones positivas definidas y este problema tiene solución debido a (12) (c.f [15], Sec. 9.1).

Lo siguiente es dar concretamente la medida μ que resuelve el Problema 1 y determinar su unicidad.

Sea $\mathbb{C}_n[t]$ el espacio de los polinomios complejos de grado $\leq n$, equipado con producto interno

$$\langle f, g \rangle = \sum_{j,k=0}^n s_{j+k} \bar{\alpha}_j \beta_k, \quad f = \sum_{j=0}^n \alpha_j t^j, \quad g = \sum_{j=0}^n \beta_j t^j \in \mathbb{C}_n[t]$$

y base ortonormal $\{e_j\}_{j=0}^n$, calculada mediante proceso de Gram-Schmidt, donde e_j es un polinomio de grado j , con coeficiente principal positivo.

Considere sobre $\mathbb{C}_n[t]$ al operador de multiplicación $Jf(t) = tf(t)$, el cual claramente tiene dominio

$$(13) \quad \text{dom } J = \text{span} \{e_j\}_{j=0}^{n-1} = \mathbb{C}_n[t] \ominus \{e_n\}.$$

PROPOSICIÓN 14. *El operador de multiplicación es simétrico. Además, existen $a_j \in \mathbb{R}$, $b_j > 0$, con $j = 0, 1, \dots, n-1$, tales que*

$$Je_j = b_j e_{j+1} + a_j e_j + b_{j-1} e_{j-1}. \quad (b_{-1} = 0)$$

Demostración. Es sencillo verificar que $\langle Jf, f \rangle = \langle f, Jf \rangle$, para todo $f \in \text{dom } J$, lo que implica que J es simétrico. Además para $j = 0, 1, \dots, n-1$, se tiene que $Je_j = \sum_{k=0}^n \langle Je_j, e_k \rangle e_k$. Como J es simétrico, $\langle Je_j, e_k \rangle = \langle e_j, Je_k \rangle = 0$ para $1 < |j-k|$, ya que Je_j es de grado $j+1$ y Je_k es de grado $k+1$. De este modo

$$Je_j = \langle Je_j, e_{j+1} \rangle e_{j+1} + \langle Je_j, e_j \rangle e_j + \langle Je_j, e_{j-1} \rangle e_{j-1}.$$

Observe que Je_j y e_{j+1} son polinomios de grado $j+1$ con coeficiente principal positivo y entonces $\langle Je_j, e_{j+1} \rangle > 0$. Por lo tanto, haciendo $b_j = \langle Je_j, e_{j+1} \rangle$, $a_j = \langle Je_j, e_j \rangle$ se llega a (13). \square

El resultado anterior asegura que la representación matricial de J con respecto a la base $\{e_j\}_{j=0}^n$, está dada por

$$J = \begin{pmatrix} a_0 & b_0 & 0 & \dots & 0 & * \\ b_0 & a_1 & b_1 & \dots & 0 & * \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & * \\ 0 & 0 & 0 & \dots & b_{n-1} & * \end{pmatrix}.$$

Además, (6) señala que J^* es un operador multivaluado con $\text{mul } J^* = \text{span} \{e_n\}$.

Sea $J_0 \supset J$ el operador con $\text{dom } J_0 = \mathbb{C}_n[t]$ y representación matricial

$$J_0 = \begin{pmatrix} a_0 & b_0 & 0 & \dots & 0 & 0 \\ b_0 & a_1 & b_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & b_{n-1} & 0 \end{pmatrix},$$

que representa una extensión autoadjunta de J . Estos operadores serán identificados con su gráfica para trabajar con operadores multivaluados en $\mathbb{C}_n[t] \oplus \mathbb{C}_n[t]$.

Considere el operador multivaluado

$$Z := \text{span} \left\{ \begin{pmatrix} 0 \\ e_n \end{pmatrix} \right\} \subset \mathbb{C}_n[t] \oplus \mathbb{C}_n[t].$$

TEOREMA 15. *El operador de multiplicación satisface $J = J_0 \cap Z^*$.*

Demostración. Si $f \in \text{dom } J$, entonces $\langle f, e_n \rangle = \langle Jf, 0 \rangle = 0$. Así, $\begin{pmatrix} f \\ Jf \end{pmatrix} \in Z^*$ y $J \subset J_0 \cap Z^*$. Para la otra contención, si $\begin{pmatrix} f \\ g \end{pmatrix} \in J_0 \cap Z^*$, entonces $f \in \text{dom } J_0$ y $\langle f, e_n \rangle = \langle g, 0 \rangle = 0$, es decir $f \in \text{dom } J$, lo que lleva a $J_0 \cap Z^* \subset J$. \square

Uno verifica de manera sencilla que J_0 y Z son linealmente independientes.

COROLARIO 16. *El adjunto de J viene dado por $J^* = J_0 \dot{+} Z$.*

Demostración. Uno tiene del Corolario [4](#) que $J_0 \dot{+} Z = (J_0 \dot{+} Z)^{**} = (J_0 \cap Z^*)^*$. Por lo tanto, del Teorema [15](#) se tiene la afirmación. \square

El siguiente resultado es consecuencia directa del Teorema [15](#) y de [\[11\]](#) Teo. 2.4]

COROLARIO 17. *Todas las extensiones autoadjuntas del operador de multiplicación J son perturbaciones unidimensionales del operador J_0 y están en correspondencia uno-a-uno con $\tau \in \mathbb{R} \cup \{\infty\}$, las cuales vienen dadas por*

$$(14) \quad J_\tau = \left\{ \left(J_0 f(t) + \tau \begin{pmatrix} f(t) \\ \langle f, e_n \rangle e_n \end{pmatrix} \right) : f \in \mathbb{C}_n[t] \right\}, \quad \tau \neq \infty$$

que son operadores lineales, a diferencia de

$$J_\infty = J \dot{+} Z$$

que resulta ser un operador multivaluado.

Las extensiones [\(14\)](#) tienen representación matricial

$$(15) \quad J_\tau = \begin{pmatrix} a_0 & b_0 & 0 & \dots & 0 & 0 \\ b_0 & a_1 & b_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & b_{n-1} & \tau \end{pmatrix}$$

mientras que la representación matricial de la parte operador de J_∞ es

$$J_{\infty \odot} = \begin{pmatrix} a_0 & b_0 & 0 & \dots & 0 & 0 & * \\ b_0 & a_1 & b_1 & \dots & 0 & 0 & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-2} & b_{n-2} & * \\ 0 & 0 & 0 & \dots & b_{n-2} & a_{n-1} & * \\ 0 & 0 & 0 & \dots & 0 & 0 & * \end{pmatrix}.$$

En virtud del Teorema [9](#) y de la Proposición [11](#), el espectro de J_∞ coincide con el espectro de

$$(16) \quad J_{\infty\odot} = \begin{pmatrix} a_0 & b_0 & 0 & \dots & 0 & 0 \\ b_0 & a_1 & b_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-2} & b_{n-2} \\ 0 & 0 & 0 & \dots & b_{n-2} & a_{n-1} \end{pmatrix},$$

en el espacio $\mathbb{C}_{n-1}[t]$.

Lo siguiente es demostrar la solución al Problema [1](#). Debido al teorema espectral, cada extensión autoadjunta [\(15\)](#) de J tiene representación

$$J_\tau = \sum_{\lambda \in \sigma(J_\tau)} \lambda P_\lambda, \quad (\tau \in \mathbb{R})$$

donde P_λ es la proyección ortogonal sobre el auto-vector unitario de $\lambda \in \sigma(J_\tau)$. Defina la medida espectral (ver [\[14\]](#), Caps. 4 y 5)

$$E_\tau(M) := \sum_{\lambda \in \sigma(J_\tau) \cap M} P_\lambda, \quad M \in \mathcal{B}(\mathbb{R}),$$

donde $\mathcal{B}(\mathbb{R})$ representa la σ -álgebra de Borel en \mathbb{R} . Así,

$$(17) \quad J_\tau = \int_{\mathbb{R}} \lambda dE_\tau(\lambda).$$

Ahora bien, denote

$$\mu_\tau(\lambda) := \langle 1, E_\tau(\lambda)1 \rangle,$$

la cual representa una medida de probabilidad, ya que $\mu_\tau(\mathbb{R}) = \langle 1, 1 \rangle = s_0 = 1$.

De esta manera se cumple que para $j, k = 0, 1, \dots, n$,

$$(18) \quad \begin{aligned} s_{j+k} &= \langle t^j, t^k \rangle = \langle J^j 1, J^k 1 \rangle \\ &= \langle J_\tau^j 1, J_\tau^k 1 \rangle = \langle 1, J_\tau^{j+k} 1 \rangle \\ &= \int_{\mathbb{R}} \lambda^{j+k} d\langle 1, E_\tau(\lambda)1 \rangle = \int_{\mathbb{R}} \lambda^{j+k} d\mu_\tau(\lambda). \end{aligned}$$

Con esto se ha demostrado lo siguiente:

TEOREMA 18. *El Problema [1](#) tiene infinitas soluciones parametrizadas por $\tau \in \mathbb{R}$ de la siguiente manera:*

$$s_j = \int_{\mathbb{R}} \lambda^j d\mu_\tau(\lambda), \quad j = 0, 1, \dots, 2n.$$

De manera análoga a [\(17\)](#) se tiene la existencia de una medida espectral E_∞ para [\(16\)](#) en $\mathbb{C}_{n-1}[t]$, tal que $J_{\infty\odot} = \int_{\mathbb{R}} \lambda dE_\infty(\lambda)$. La sección finaliza con el siguiente resultado que se demuestra siguiendo las mismas líneas de [\(18\)](#), con la medida $\mu_\infty(\lambda) := \langle 1, E_\infty(\lambda)1 \rangle$.

COROLARIO 19. *El Problema [1](#) satisface*

$$s_j = \int_{\mathbb{R}} \lambda^j d\mu_\infty(\lambda), \quad j = 0, 1, \dots, 2(n-1).$$

La condición [\(13\)](#) fue la que condujo a trabajar el problema de momentos de Hamburger truncado con la teoría de operadores multivaluados. Gracias a esta teoría se han podido caracterizar todas las medidas que resuelven el Problema [1](#).

AGRADECIMIENTOS. El autor expresa su gratitud al árbitro anónimo cuyos comentarios pertinentes llevaron a una mejor presentación de este trabajo.

REFERENCIAS

- [1] Keshav Raj Acharya, *Self-adjoint extension and spectral theory of a linear relation in a Hilbert space*, ISRN Math. Anal. (2014), Art. ID 471640, 5. MR 3178945
- [2] N. I. Akhiezer, *The classical moment problem and some related questions in analysis*, Translated by N. Kemmer, Hafner Publishing Co., New York, 1965. MR 0184042 (32 #1518)
- [3] Richard Arens, *Operational calculus of linear relations*, Pacific J. Math. **11** (1961), 9–23. MR 0123188 (23 #A517)
- [4] Jussi Behrndt, Seppo Hassi, and Henk de Snoo, *Boundary value problems, Weyl functions, and differential operators*, Monographs in Mathematics, vol. 108, Birkhäuser/Springer, Cham, [2020] ©2020. MR 3971207
- [5] M. Sh. Birman and M. Z. Solomjak, *Spectral theory of selfadjoint operators in Hilbert space*, Mathematics and its Applications (Soviet Series), D. Reidel Publishing Co., Dordrecht, 1987, Translated from the 1980 Russian original by S. Khrushchëv and V. Peller. MR 1192782 (93g:47001)
- [6] Earl A. Coddington, *Extension theory of formally normal and symmetric subspaces*, American Mathematical Society, Providence, R.I., 1973, Memoirs of the American Mathematical Society, No. 134. MR 0477855
- [7] Ronald Cross, *Multivalued linear operators*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 213, Marcel Dekker, Inc., New York, 1998. MR 1631548
- [8] V. A. Derkach and M. M. Malamud, *The extension theory of Hermitian operators and the moment problem*, J. Math. Sci. **73** (1995), no. 2, 141–242, Analysis. 3. MR 1318517
- [9] A. Dijksma and H. S. V. de Snoo, *Self-adjoint extensions of symmetric subspaces*, Pacific J. Math. **54** (1974), 71–100. MR 0361889 (50 #14331)
- [10] Mercedes Fernandez Miranda and Jean-Philippe Labrousse, *The Cayley transform of linear relations*, Proc. Amer. Math. Soc. **133** (2005), no. 2, 493–499. MR 2093073
- [11] Seppo Hassi and Henk de Snoo, *One-dimensional graph perturbations of selfadjoint relations*, Ann. Acad. Sci. Fenn. Math. **22** (1997), no. 1, 123–164. MR 1430397 (97m:47025)
- [12] H. Langer and B. Textorius, *On generalized resolvents and Q -functions of symmetric linear relations (subspaces) in Hilbert space*, Pacific J. Math. **72** (1977), no. 1, 135–165. MR 0463964 (57 #3902)
- [13] Josué I. Rios-Cangas and Luis O. Silva, *Dissipative extension theory for linear relations*, Expo. Math. **38** (2020), no. 1, 60–90. MR 4082306
- [14] Konrad Schmüdgen, *Unbounded self-adjoint operators on Hilbert space*, Graduate Texts in Mathematics, vol. 265, Springer, Dordrecht, 2012. MR 2953553
- [15] Konrad Schmüdgen, *The moment problem*, Graduate Texts in Mathematics, vol. 277, Springer, Cham, 2017. MR 3729411
- [16] Yuming Shi, Guixin Xu, and Guojing Ren, *Boundedness and closedness of linear relations*, Linear Multilinear Algebra **66** (2018), no. 2, 309–333. MR 3750592
- [17] Barry Simon, *The classical moment problem as a self-adjoint finite difference operator*, Adv. Math. **137** (1998), no. 1, 82–203. MR 1627806 (2001e:47020)
- [18] T.-J. Stieltjes, *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. **8** (1894), no. 4, J1–J122. MR 1508159
- [19] J. von Neumann, *Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren*, Math. Ann. **102** (1930), no. 1, 49–131. MR 1512569
- [20] John von Neumann, *Functional Operators. II. The Geometry of Orthogonal Spaces*, Annals of Mathematics Studies, no. 22, Princeton University Press, Princeton, N. J., 1950. MR 0034514

Dirección del autor:

Universidad Autónoma Metropolitana,
 Unidad Iztapalapa,
 División de Ciencias Básicas e Ingeniería,
 Departamento de Matemáticas.
 Av. San Rafael Atlixco 186, Col. Vicentina,
 Alcaldía Iztapalapa, C.P. 09340, CDMX, México.
 e-mail: jottsmok@xanum.uam.mx



EL MÉTODO DE DESCOMPOSICIÓN DE ADOMIAN PARA EL CÁLCULO DE INTEGRALES GAUSSIANAS: ECUACIÓN DE CALOR Y DE BLACK-SCHOLES

OSWALDO GONZÁLEZ-GAXIOLA

RESUMEN. Derivada por los economistas Fischer Black, Myron Scholes y Robert Merton en 1973, la ecuación de Black-Scholes es una ecuación diferencial parcial cuya solución es utilizada para determinar cuánto vale en un momento dado una opción de compra o venta en principio tipo europea. En el presente trabajo proponemos una adaptación del método de descomposición de Adomian, que permite la evaluación de ciertas integrales gaussianas en términos de una serie convergente. Además haremos la aplicación del método expuesto para encontrar soluciones a la ecuación de Black-Scholes bajo cierta restricción impuesta por la condición inicial.

1. INTRODUCCIÓN

La mayor parte de los fenómenos que surgen en el mundo real son descritos por ecuaciones diferenciales no lineales (tanto ordinarias como parciales) y en algunos casos por ecuaciones integrales. Sin embargo, los métodos más comunes desarrollados hasta ahora en matemáticas se utilizan para resolver ecuaciones diferenciales lineales. En la década de los años 80, George Adomian introdujo un nuevo método para resolver ecuaciones diferenciales no lineales, tanto parciales como ordinarias e incluso estocásticas [1, 2, 3]. Este método ha sido denominado como el método de descomposición de Adomian (ADM). Se ha demostrado que el método de descomposición resuelve de manera eficiente, fácil y precisa una gran clase de ecuaciones diferenciales lineales y no lineales ordinarias, parciales, integrales, de orden entero o fraccional, deterministas o estocásticas [4, 5, 6, 7, 8]. El método se adapta muy bien tanto a problemas físicos como de otras áreas del conocimiento, ya que no requiere de linealizaciones, perturbaciones y otras suposiciones restrictivas que pueden cambiar la naturaleza del problema que se está modelando y resolviendo. ADM tiene la ventaja de que converge a la solución exacta en una gran mayoría de casos importantes en las aplicaciones y se puede manejar de manera relativamente fácil pues el método genera a través de su algoritmo una solución en forma de una serie cuyos términos son determinados por una relación recursiva. Algunos trabajos fundamentales sobre diversos aspectos de las modificaciones de ADM así como el estudio de la convergencia se pueden ver en [9, 10, 11, 12, 13].

Por otra parte, en 1973, Robert C. Merton, junto a Fisher Black y Myron Scholes desarrolló el modelo de Black-Scholes. Merton ayudó a introducir el cálculo estocástico en la economía financiera, lo que permitió que el comportamiento de los precios fuese descrito con el lenguaje preciso de la probabilidad. Al modelo obtenido Robert Merton denominó la ecuación de Black-Scholes [14]. En 1997 el Premio Nobel de Economía fue para Robert C. Merton y Myron S. Scholes, por contribuir a las ciencias económicas, con un nuevo método para determinar el valor de los derivados, antes Fisher Black falleció y por lo tanto no pudo recibir este premio.

En el presente trabajo haremos una adaptación de la versión estándar de ADM para calcular integrales gaussianas y así poder hallar la solución de la ecuación de calor y

2010 *Mathematics Subject Classification.* 35K05, 35Q91 91G15.

Palabras clave. Método de descomposición de Adomian, ecuación de Black-Scholes, integrales gaussianas.

de ello se derivará una alternativa de solución para la ecuación de Black-Scholes en su versión lineal.

2. UNA BREVE DESCRIPCIÓN DEL MÉTODO DE DESCOMPOSICIÓN DE ADOMIAN

El método que a continuación expondremos es un método poderoso para resolver ecuaciones diferenciales no lineales. El método se basa en descomponer la solución de una ecuación diferencial lineal o no lineal en una serie de funciones, cada término de esta serie se obtiene de manera polinomial y recursiva partiendo de una condición inicial que requiere ser una función analítica, la técnica es muy simple en una formulación abstracta pero la dificultad surge al calcular cada uno de los polinomios especialmente en los modelos descritos por ecuaciones diferenciales no lineales.

Para establecer la idea básica de una manera sencilla, consideramos la ecuación diferencial ya sea ordinaria o parcial, en general no lineal, en la forma [15]:

$$(1) \quad Fu(x, t) = g(x, t)$$

con la condición inicial

$$(2) \quad u(x, 0) = f(x),$$

donde F representa un operador diferencial (en general, no lineal) que envuelve tanto términos lineales como no lineales y entonces la ecuación (1) puede escribirse como

$$(3) \quad L_t u(x, t) + Ru(x, t) + Nu(x, t) = g(x, t)$$

donde $L_t = \frac{\partial}{\partial t}$, R es un operador lineal que involucra derivadas parciales con respecto a x y N es un operador no lineal; g es un término no homogéneo independiente de u . Despejando $L_t u(x, t)$,

$$(4) \quad L_t u(x, t) = g(x, t) - Ru(x, t) - Nu(x, t)$$

Como L es invertible, operando en (4) con el inverso $L_t^{-1}(\cdot) = \int_0^t (\cdot) dz$ tenemos que,

$$(5) \quad L_t^{-1} L_t u(x, t) = L_t^{-1} g(x, t) - L_t^{-1} Ru(x, t) - L_t^{-1} Nu(x, t)$$

luego, una expresión equivalente a (5) es

$$(6) \quad u(x, t) = f(x) + L_t^{-1} g(x, t) - L_t^{-1} Ru(x, t) - L_t^{-1} Nu(x, t)$$

donde $f(x)$ es la constante de integración (con respecto a t) que satisface $L_t f = 0$. Para problemas con valor inicial en $t = t_0$, tendríamos convenientemente definido L^{-1} para $L = \frac{\partial^n}{\partial t^n}$ como la integral n -veces iterada definida de t_0 a t .

Por el método (ADM) asumimos la solución de (1), (2) en forma de serie para la función desconocida $u(x, t)$ dada por,

$$(7) \quad u(x, t) = \sum_{n=0}^{\infty} u_n(x, t)$$

El término no lineal $Nu(x, t)$ por medio de ADM se descompone como

$$(8) \quad Nu(x, t) = \sum_{n=0}^{\infty} A_n(u_0, u_1, \dots, u_n)$$

donde la sucesión $\{A_n\}_{n=0}^{\infty}$, es la llamada sucesión de polinomios de Adomian y son dados por la fórmula

$$(9) \quad A_n = \frac{1}{n!} \frac{d^n}{d\alpha^n} [N(\sum_{k=0}^n \alpha^k u_k)]|_{\alpha=0}.$$

En [16] podemos ver que el cálculo explícito de cada uno de los A_n es realmente sencillo y se tiene:

$$\begin{aligned} A_0(u_0) &= N(u_0) \\ A_1(u_0, u_1) &= N'(u_0)u_1 \\ A_2(u_0, u_1, u_2) &= N'(u_0)u_2 + \frac{u_1^2}{2!} N''(u_0) \end{aligned}$$

$$\begin{aligned}
 A_3(u_0, \dots, u_3) &= N'(u_0)u_3 + N''(u_0)u_1u_2 + \frac{u_1^3}{3!}N'''(u_0) \\
 A_4(u_0, \dots, u_4) &= u_4N'(u_0) + (\frac{1}{2!}u_2^2 + u_1u_3)N''(u_0) + \frac{u_1^2u_2}{2!}N'''(u_0) + \frac{u_1^4}{4!}N^{(iv)}(u_0) \\
 &\vdots
 \end{aligned}$$

Ahora, sustituyendo (7), (8) y (9) en la ecuación (6) obtenemos,

$$(10) \quad \sum_{n=0}^{\infty} u_n(x, t) = f(x) + L_t^{-1}g(x, t) - L_t^{-1}R \sum_{n=0}^{\infty} u_n(x, t) - L_t^{-1} \sum_{n=0}^{\infty} A_n(u_0, u_1, \dots, u_n),$$

de donde obtenemos el algoritmo recursivo

$$(11) \quad \begin{cases} u_0(x, t) = f(x) + L_t^{-1}g(x, t), \\ u_{n+1}(x, t) = -L_t^{-1}Ru_n(x, t) - L_t^{-1}A_n(u_0, u_1, \dots, u_n), \quad n = 0, 1, 2, \dots \end{cases}$$

Con el algoritmo recursivo establecido en la ecuación (11), la aproximación a n -términos de la solución de (1)-(2) la obtenemos como

$$(12) \quad S_k(x, t) = \sum_{n=0}^k u_n(x, t) \quad \text{con} \quad u(x, t) = \lim_{k \rightarrow \infty} S_k(x, t)$$

Como sabemos, cuantos más términos se agreguen a la solución aproximada, más precisa será.

La descomposición de la solución en serie, generalmente converge muy rápido. Las condiciones para las cuales el método converge no son objetivo del presente estudio y han sido estudiadas principalmente en los trabajos [10], [11], [12] y [13] así como en varias de sus referencias.

3. LA ECUACIÓN DE CALOR Y ADM

Ahora veremos que ADM resulta ser un procedimiento efectivo para la evaluación de ciertas integrales que presentan cierto grado de dificultad como son las integrales gaussianas. Para abordar este problema, comenzamos por considerar la ecuación de calor para una varilla infinita con una condición inicial arbitraria:

$$(13) \quad \begin{cases} \frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}, \quad x \in \mathbb{R}, \quad t > 0, \quad k > 0. \\ u(x, 0) = f(x), \\ |u(x, t)| \leq M \quad \text{para toda } x \in \mathbb{R} \text{ y } t > 0. \end{cases}$$

La representación explícita de la solución al problema de valor inicial (13) es dada por

$$(14) \quad u(x, t) = \frac{1}{\sqrt{4k\pi t}} \int_{-\infty}^{\infty} f(\xi) e^{-\frac{(x-\xi)^2}{4k t}} d\xi, \quad -\infty < x < \infty, \quad t > 0.$$

En el integrando de (14) la función $G(x, t) = \frac{1}{\sqrt{4k\pi t}} e^{-\frac{x^2}{4k t}} d\xi$ es conocida como la solución fundamental de la ecuación de calor. Para mayor referencia al respecto ver [17].

Puede suceder que la integral en el lado derecho de (14) no sea expresable en términos de funciones elementales y menos aún pueda ser calculada de manera exacta por medio de los métodos elementales del cálculo integral. En el presente trabajo buscamos evaluar este tipo de integrales que involucran la función de error usando el método de descomposición de Adomian. El presente trabajo no es el primero en abordar el problema de evaluar integrales a través de un método de descomposición, para mayor información podemos ver [18], [19] y sus referencias.

Consideremos el problema de valor inicial dado por la ecuación de calor (13). Primero, definimos los operadores lineales L y R en el contexto de ADM como

$$(15) \quad L_t = \frac{\partial}{\partial t} \quad \text{y} \quad R = \frac{\partial^2}{\partial x^2}.$$

Con esto podemos escribir la ecuación diferencial parcial de (13) como

$$(16) \quad L_t u = kRu.$$

Ahora vamos a considerar el operador inverso de L_t , esto es, $L_t^{-1}(u) = \int_0^t \frac{\partial u}{\partial \xi} d\xi$ para aplicarlo a ambos lados de (16) para obtener, usando el teorema fundamental del cálculo y la condición inicial $f(x) = u(x, 0)$:

$$(17) \quad u(x, t) = f(x) + kL_t^{-1}Ru.$$

Haciendo uso de (7), descomponemos la solución en serie de componentes u_n , esto es,

$$(18) \quad u(x, t) = \sum_{n=0}^{\infty} u_n(x, t).$$

Sustituyendo (18) en (17), obtenemos

$$(19) \quad \sum_{n=0}^{\infty} u_n(x, t) = f(x) + k \int_0^t \sum_{n=0}^{\infty} \frac{\partial^2 u_n}{\partial x^2}(x, \xi) d\xi.$$

Finalmente, considerando (19) e igualando sumandos del mismo orden obtenemos el algoritmo para encontrar las componentes de la serie (18):

$$(20) \quad \begin{cases} u_0(x, t) = f(x), \\ u_{n+1}(x, t) = k \int_0^t \frac{\partial^2 u_n}{\partial x^2}(x, \xi) d\xi, \quad n = 0, 1, 2, \dots \end{cases}$$

Consecuentemente, tenemos

$$\begin{aligned} u_0(x, t) &= f(x), \\ u_1(x, t) &= kf^{(2)}(x)t, \\ u_2(x, t) &= k^2 f^{(4)}(x) \frac{t^2}{2!}, \\ &\vdots \\ u_n(x, t) &= k^n f^{(2n)}(x) \frac{t^n}{n!} \\ &\vdots \end{aligned}$$

Del esquema anterior la solución de la ecuación de calor es dada por

$$(21) \quad u(x, t) = \sum_{n=0}^{\infty} k^n f^{(2n)}(x) \frac{t^n}{n!},$$

donde $f^{(j)}(x)$ significa la j -ésima derivada de f respecto a x .

Como resumen de todo lo anterior, podemos establecer el siguiente teorema.

TEOREMA 1. *Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función $C^\infty(\mathbb{R})$. Entonces para $-\infty < x < \infty$ se tiene*

$$(22) \quad \int_{-\infty}^{\infty} f(\xi) e^{-\frac{(\xi-x)^2}{4kt}} d\xi = \sqrt{4k\pi t} \sum_{n=0}^{\infty} f^{(2n)}(x) \frac{(kt)^n}{n!}, \quad t > 0, \quad k > 0.$$

Demostración. La demostración se sigue directamente de utilizando (21) y (14). \square

COROLARIO 2. *Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función $C^\infty(\mathbb{R})$. Entonces para $b > 0$ se tiene*

$$(23) \quad \int_{-\infty}^{\infty} f(x) e^{-bx^2} dx = \sqrt{\frac{\pi}{b}} \sum_{n=0}^{\infty} \frac{f^{(2n)}(0)}{4^n b^n n!}.$$

La fórmula (23) nos proporciona una herramienta útil para calcular integrales gaussianas en las cuales la función f cumplan con la condición de diferenciabilidad.

4. APLICACIÓN A LA ECUACIÓN DE BLACK-SCHOLES

La ecuación de Black-Scholes es una ecuación diferencial parcial de segundo orden de tipo parabólico utilizada en matemáticas financieras modernas para determinar el precio de activos financieros, y esta es dada por [14, 20]:

$$(24) \quad V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \delta)SV_S - rV = 0,$$

la ecuación (24) es una ecuación diferencial parcial parabólica definida en

$$D_V = \{(S, t) : S > 0, 0 < t \leq T\}.$$

En la ecuación de Black-Scholes (24) tenemos que $V = V(S, t)$ es el precio de una opción o también conocida como función de pago, $S = S(t)$ representa el precio de un activo subyacente al tiempo t , r es la tasa de interés del mercado de deuda libre de riesgo, δ representa los dividendos que se reparten de forma continua y σ es la volatilidad de la acción, medida como la desviación estándar de los logaritmos de la cotización de la acción y que aquí supondremos constante.

El modelo dado por (24) tiene varias suposiciones técnicas y propias del lenguaje financiero, las principales son que r , es conocida y constante en el tiempo, no hay costos de transacción en la compra o venta del activo por la opción, no hay impuestos u otras tasaciones y no existen oportunidades de arbitraje sin riesgo.

Siguiendo la idea dada en [21], consideremos los siguientes cambios de variables; para las variables independientes

$$S = e^x, \quad t = T - \frac{2\tau}{\sigma^2},$$

y para la variable dependiente

$$v(x, \tau) = V(S, t) = V(e^x, T - \frac{2\tau}{\sigma^2}).$$

Ahora, calculando tenemos:

$$\begin{aligned} \frac{\partial V}{\partial t} &= \frac{\partial v}{\partial \tau} \frac{\partial \tau}{\partial t} = -\frac{\sigma^2}{2} \frac{\partial v}{\partial \tau}, \\ \frac{\partial V}{\partial S} &= \frac{\partial v}{\partial x} \frac{\partial x}{\partial S} = \frac{1}{S} \frac{\partial v}{\partial x}, \\ \frac{\partial^2 V}{\partial S^2} &= \frac{\partial}{\partial S} \left(\frac{\partial V}{\partial S} \right) = \frac{1}{S^2} \left[\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x} \right]. \end{aligned}$$

Sustituyendo en (24) obtenemos

$$(25) \quad v_\tau = v_{xx} + \left(\frac{r - \delta}{\sigma^2/2} - 1 \right) v_x - \frac{r}{\sigma^2/2} v.$$

Ahora definimos los coeficientes como

$$\kappa = \frac{r - \delta}{\sigma^2/2}, \quad l = \frac{\delta}{\sigma^2/2},$$

de donde podemos reescribir la ecuación (25) como

$$(26) \quad v_\tau = v_{xx} + (\kappa - 1)v_x - (\kappa + l)v.$$

La ecuación (26) tiene coeficientes constantes.

Además proponemos el siguiente cambio de variable dependiente

$$(27) \quad v(x, \tau) = e^{-\gamma x - (\beta^2 + l)\tau} u(x, \tau),$$

donde

$$\gamma = \frac{1}{2}(\kappa - 1), \quad \text{y} \quad \beta = \frac{1}{2}(\kappa + 1) \quad \text{y por lo tanto} \quad \beta^2 = \gamma^2 + \kappa.$$

Derivando obtenemos

$$v_\tau = e^{-\gamma x - (\beta^2 + l)\tau} (u_\tau - (\beta^2 + l)u),$$

$$v_x = e^{-\gamma x - (\beta^2 + l)\tau} (u_x - \gamma u), \quad v_{xx} = e^{-\gamma x - (\beta^2 + l)\tau} (u_{xx} + \gamma^2 u - 2\gamma u_x).$$

Sustituyendo en la ecuación (26)

$$(28) \quad u_\tau = u_{xx} + (\kappa - 1 - 2\gamma)u_x + \gamma(2\gamma - \kappa + 1)u = u_{xx}.$$

Por lo tanto, la ecuación que satisface la variable dependiente $u(x, \tau)$ es la forma adimensional de la ecuación de calor:

$$(29) \quad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

Como consecuencia de lo anterior, la ecuación de Black-Scholes (24) es equivalente a la ecuación de calor definida en

$$D_u = \{(x, \tau) : -\infty < x < \infty, 0 \leq \tau \leq \frac{\sigma^2}{2}T\}.$$

En general, la ecuación de calor está definida para todo $\tau \geq 0$ pero considerando que una opción expira al tiempo de madurez $t = T$ el dominio de definición es el conjunto D_u .

Debido al Teorema 1 si se conoce la condición inicial suficientemente suave $u(x, 0) = g(x)$, entonces la solución de la ecuación de Black-Scholes es dada por

$$(30) \quad u(x, \tau) = \sqrt{4\pi\tau} \sum_{n=0}^{\infty} g^{(2n)}(x) \frac{\tau^n}{n!}, \quad \tau \geq 0.$$

Finalmente, no debemos olvidar que la relación entre las variables originales (S, t) en D_V y las transformadas (x, τ) en D_u esta dada por:

$$x = \ln(S), \quad \tau = \frac{\sigma^2}{2}(T - t).$$

Cabe aclarar que la solución que hemos obtenido a partir del método de descomposición expuesto tiene como limitante el requisito de una condición inicial con alto grado de diferenciabilidad. Además, los mercados del mundo real tienen comportamientos no lineales [22, 23], un modelo con volatilidad no constante que resulta modificada por el mercado financiero tiene diferentes tratamientos y muchos de ellos matemáticamente muy sofisticados. Al verse modificada la volatilidad el modelo afecta el costo de las transacciones y se presenta iliquidez financiera, entre otras posibles afectaciones. En [24] presentamos una versión no lineal del modelo de Black-Scholes derivada de los costos de transacción así como también de la iliquidez del mercado, y la solución de dicho modelo a través del mismo método aquí empleado, en él se considera la volatilidad como una función que depende del tiempo, del precio del activo subyacente y de la prima de la opción.

5. CONCLUSIONES

En el presente trabajo hemos expuesto de manera breve el método de descomposición estándar de Adomian (ADM), también basados en el método y en la ecuación de difusión del calor hemos obtenido una manera práctica de evaluar integrales gaussianas a través de la serie proporcionada por ADM sin tener que recurrir a discretizaciones, perturbaciones u otras suposiciones que en ocasiones distorsionan la naturaleza del problema. Además, hemos dado una breve introducción al modelo matemático establecido por la ecuación de Black-Scholes en su versión lineal, la cual puede transformarse en una ecuación de difusión y dependiendo de la condición inicial puede ser resuelta por medio del método estudiado. Finalmente, podemos concluir que el método que hemos mostrado es una herramienta poderosa y no sofisticada que puede utilizarse además de para el cálculo de integrales gaussianas y resolver algunas ecuaciones diferenciales que surgen en las aplicaciones de las matemáticas.

REFERENCIAS

- [1] Adomian, G., *Solving Frontier Problems of Physics: The Decomposition Method*, Boston, MA: Kluwer Academic Publishers, 1994.
- [2] Adomian, G. *A review of the decomposition method in applied mathematics*. J. Math. Anal. Appl., 135, No. 2, 501-544, 1988.
- [3] Adomian, G., *Nonlinear stochastic operator equations*, New York, Academic Press, 1986.
- [4] Gejji, V. D. & Jafari, H. *Adomian decomposition: a tool for solving a system of fractional differential equations*. J. Math. Anal. Appl., 301, 508-518, 2005.
- [5] Khuri, S. A. *A Laplace decomposition algorithm applied to class of nonlinear differential equations*. J. Math. Appl., 4, 141-155, 2001.
- [6] Khuri, S. A. *A new approach to Bratu's problem*. Appl. Math. Comput., 147, 131-136, 2004.
- [7] González-Gaxiola, O. & Biswas, A. *Akhmediev breathers, Peregrine solitons and Kuznetsov-Ma solitons in optical fibers and PCF by Laplace-Adomian decomposition method*. Optik, 172, 930-939, 2018.
- [8] Adomian, G. & Serrano, S. E. *Stochastic contaminant transport equation in porous media*. Appl. Math. Lett., 11, 53-55, 1998.
- [9] Venkatarangan, S. N. & Rajalakshmi, K. *A modification of Adomian's solution for nonlinear oscillatory systems*. Comput. Math. Appl., 29, 67-73, 1995.
- [10] Cherruault, Y. *Convergence of Adomian's method*. Kybernetes, 18, No. 2, 31-38, 1989.
- [11] Cherruault, Y. & G. Adomian, G. *Decomposition methods: a new proof of convergence*. Math. Comput. Modelling, 18, No. 12, 103-106, 1993.
- [12] Abbaoui, K. & Cherruault Y. *Convergence of Adomian's method applied to differential equations*. Comput. Math. Appl., 28, No. 5, 103-109, 1994.
- [13] Abbaoui, K. & Cherruault Y. *New ideas for proving convergence of decomposition methods*. Comput. Math. Appl., 29, No. 7, 103-108, 1995.
- [14] Black, F. & Scholes, M. *The pricing of option and corporate liabilities*. The Journal of Political Economy, 81, No. 3, 637-654, 1973.
- [15] Wazwaz, A. M., *Partial Differential Equations: Methods and Applications*, Lisse, Netherlands: Balkema Publishers, 2002.
- [16] Wazwaz, A. M. *A new algorithm for calculating adomian polynomials for nonlinear operators*. Appl. Math. and Computation, 111, No. 1, 53-69, 2000.
- [17] Strauss, W. A., *Partial Differential Equations: An Introduction*, John Wiley and Sons, Inc., 1992.
- [18] Adomian, G. & Rach, R. *Evaluation of integrals by decomposition*. J. of Comput. and Applied Math., 23, 99-101, 1986.
- [19] Sandoval-Hernández, M. A., Vazquez-Leal, H., Filobello-Nino, U. & Hernández-Martínez, L. *New handy and accurate approximation for the Gaussian integrals with applications to science and engineering*. Open Math., 17, 1774-1793, 2019.
- [20] Cox, J. C., Ross, S. A. & Rubinstein, M. *Option pricing: A simplified approach*. J. of Financial Economics, 7, 229-263, 1979.
- [21] Gal, C. G., Gal, S. G. & Goldstein, J. A. *Burgers and Black-Merton-Scholes equations with real time variable and complex spatial variable*. Applicable Analysis, 92, No. 8, 1766-1786, 2013.
- [22] Frey, R. & Stremme, A. *Market volatility and feedback effects from dynamic hedging*. Math. Financ., 4, 351-374, 1997.
- [23] Liu, H. & Yong, J. *Option pricing with an illiquid underlying asset market*. J. Econ. Dyn. Control, 29, 2125-2156, 2005.
- [24] González-Gaxiola, O., Ruiz de Chávez J. & Santiago, J. A. *A nonlinear option pricing model through the Adomian decomposition method*. Int. J. Appl. Comput. Math., 2, 453-467, 2016.

Dirección del autor:

Universidad Autónoma Metropolitana,
 Unidad Cuajimalpa,
 División de Ciencias Naturales e Ingeniería,
 Departamento de Matemáticas Aplicadas y Sistemas.
 Av. Vasco de Quiroga 4871, Santa Fe, Cuajimalpa
 C.P. 05300, Ciudad de México.
 e-mail: ogonzalez@cua.uam.mx



CAMINANDO ENTRE MEDIDAS INVARIANTES Y TOPOLOGÍAS: EL TEOREMA DE HAAR INVERSO

SAÚL PILATOWSKY CAMEO

RESUMEN. En este artículo presentamos el teorema de Haar inverso, originalmente demostrado por Weil en 1940. El teorema de Haar garantiza que en grupos equipados con cierta topología, siempre existe una medida que no cambia al trasladarla por medio de la operación del grupo, es decir, una medida invariante. Esta medida –llamada medida de Haar– guarda una relación estrecha con la topología del grupo. El teorema de Haar inverso garantiza que, en cierto sentido, todas las medidas invariantes son medidas de Haar. Dado un grupo –sin topología– y una medida invariante, se puede construir una topología de manera que la medida invariante es, esencialmente, la medida de Haar.

1. INTRODUCCIÓN

Detrás del teorema de Haar está la idea de *compatibilidad de estructuras*. Cuando se estudian espacios que poseen diferentes estructuras a la vez –algebraicas, topológicas y de medida–, se busca imponer condiciones sobre ellos que las hagan compatibles unas con las otras. Por ejemplo, si se tiene un grupo equipado con una topología, es natural pedir que las operaciones de grupo sean continuas, o si tenemos una medida μ en un grupo G , buscar que la medida de cualquier conjunto medible $A \subseteq G$ no cambie al operar sobre todos sus elementos, es decir, que

$$\mu(A) = \mu(xA),$$

para cada $x \in G$, donde $xA = \{xa \mid a \in A\}$. A esta última propiedad se le llama *invarianza*.

La medida de Haar es una medida definida sobre un grupo equipado con una topología tal que las tres estructuras –la algebraica, la topológica y la de medida– son compatibles entre sí. Esta medida es invariante y, además, se puede determinar completamente a partir de la medida de los conjuntos compactos y de los conjuntos abiertos de la topología. El teorema de Haar garantiza la existencia de estas medidas cuando previamente se tiene una topología localmente compacta y un grupo que son compatibles entre sí. En la primera sección de este artículo presentamos brevemente este resultado, aunque sin demostración. El enfoque principal de este trabajo es el teorema de Haar inverso, que presentamos en la segunda sección. Este teorema nos permite construir topologías localmente compactas que son compatibles con el álgebra de grupo cuando previamente se tiene una medida invariante, transformando la medida invariante en una medida de Haar.

A lo largo de este trabajo, utilizaremos la palabra *compatible* de manera repetida para significar diferentes relaciones. Hablaremos de medidas de Radon como aquellas medidas que son compatibles con una topología, grupos topológicos como aquellos que tienen una topología compatible con las operaciones de grupo, y medidas invariantes como aquellas medidas que son compatibles con la estructura algebraica. Iremos especificando los detalles conforme requiramos los conceptos, así que el lector no debe preocuparse si no conoce algunos de ellos.

2010 *Mathematics Subject Classification*. 28C10, 43A05, 22D99.

Palabras clave. medida de Haar, medida invariante, topología de Weil.

2. UNA PROBADITA DEL TEOREMA DE HAAR

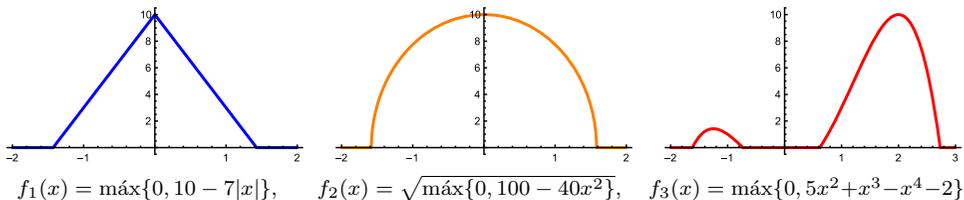
2.1. Medidas de Radon (topología + medida). Estudiaremos exclusivamente espacios topológicos que son de Hausdorff (puntos distintos tienen vecindades ajenas) y localmente compactos (cada punto tiene una base local de vecindades compactas). Los espacios euclidianos \mathbb{R}^n , así como cualquier subconjunto cerrado o abierto de estos, son espacios de este tipo. Otro ejemplo se obtiene al tomar cualquier conjunto y equiparlo con la topología discreta.

Fijemos un espacio topológico de Hausdorff y localmente compacto X . El primer objetivo es especificar a qué nos referimos al pedir que la estructura topológica de X sea compatible con las estructuras de medida que definiremos allí. Denotamos por $\mathcal{B}(X)$ a la σ -álgebra más pequeña que contiene a todos los conjuntos abiertos de X , y la llamamos la σ -álgebra de Borel. Es fácil notar que $\mathcal{B}(X)$ contiene a todos los conjuntos cerrados –pues son los complementos de los abiertos–, y por lo tanto $\mathcal{B}(X)$ contiene a los compactos de X , que siempre son cerrados en espacios de Hausdorff. Habiendo definido la σ -álgebra, nos adentramos a construir sobre ella medidas que también sean compatibles con la topología.

Sobre X , nos interesa integrar aquellas funciones continuas $f: X \rightarrow [0, \infty)$ que son cero fuera de un conjunto compacto. A las funciones en

$$\mathcal{C}_0(X) = \left\{ f: X \rightarrow [0, \infty) \mid \begin{array}{l} f \text{ es continua y existe un compacto } \\ S \subseteq X \text{ tal que } f(x) = 0 \text{ si } x \notin S \end{array} \right\}$$

se les llama funciones continuas de soporte compacto. Por ejemplo, en $X = \mathbb{R}$ con la topología usual, las funciones $f_1, f_2, f_3: \mathbb{R} \rightarrow [0, \infty)$ dadas por



son continuas, y el conjunto compacto $S = [-2, 3]$ satisface que $f_1(x) = f_2(x) = f_3(x) = 0$ si $x \notin S$, por lo que $f_1, f_2, f_3 \in \mathcal{C}_0(\mathbb{R})$. Sin embargo, la función $f_4(x) = 42$ no es un elemento de $\mathcal{C}_0(\mathbb{R})$, pues, aunque sí es continua, el único conjunto $S \subseteq \mathbb{R}$ tal que $f_4(x) = 0$ si $x \notin S$ es $S = \mathbb{R}$, que no es compacto.

Es sencillo verificar que si $f, g \in \mathcal{C}_0(X)$ y $c \in [0, \infty)$, entonces $f + g \in \mathcal{C}_0(X)$ y $cf \in \mathcal{C}_0(X)$. Una *integral* en $\mathcal{C}_0(X)$ es un funcional $I: \mathcal{C}_0(X) \rightarrow [0, \infty)$ tal que $I(f + g) = I(f) + I(g)$ (es *aditivo*) y $\forall c \in [0, \infty): I(cf) = cI(f)$ (es *homogéneo*).

De nuevo, veamos un ejemplo en $X = \mathbb{R}$ con la topología euclidiana. La integral de Lebesgue $I(f) = \int_{-\infty}^{\infty} f(x) dx$ para $f \in \mathcal{C}_0(\mathbb{R})$ es una integral en este sentido. En efecto, la aditividad, homogeneidad y el hecho de que $I(f) \geq 0$ deben ser obvios. Lo que tal vez no resulta evidente es por qué $I(f) < \infty$. Si $f \in \mathcal{C}_0(\mathbb{R})$, entonces existe $S \subseteq \mathbb{R}$ compacto tal que $f(x) = 0$ si $x \notin S$. Como S es compacto en \mathbb{R} , debe ser acotado, es decir, existe $N > 0$ tal que $S \subseteq [-N, N]$. Además, siendo f continua, existe $M = \max\{f(x) \mid x \in [-N, N]\}$. Así,

$$I(f) = \int_{-\infty}^{\infty} f(x) dx = \int_{-N}^N f(x) dx \leq \int_{-N}^N M dx = 2NM < \infty.$$

Resulta que, en general, las integrales $I: \mathcal{C}_0(X) \rightarrow [0, \infty)$ siempre son las integrales de Lebesgue inducidas por una medida en $\mathcal{B}(X)$. Más aún, esta medida tiene ciertas propiedades que la hacen compatible con la topología.

TEOREMA 1 (de representación de Riesz-Markov-Kakutani (RMK)).

Para toda integral $I: \mathcal{C}_0(X) \rightarrow [0, \infty)$ existe una única medida $\mu: \mathcal{B}(X) \rightarrow [0, \infty]$ tal

que

$$\int_X f(x) d\mu(x) = I(f)$$

para cada $f \in \mathcal{C}_0(X)$, y que satisface las siguientes condiciones:

- (I) Todo compacto $C \subseteq X$ tiene medida finita $\mu(C) < \infty$. (Finitud local)
 (II) Para todo $E \in \mathcal{B}(X)$ abierto o con $\mu(E) < \infty$,

$$\mu(E) = \sup\{\mu(K) \mid K \subseteq E \text{ compacto}\}. \quad (\text{Regularidad interior})$$

- (III) Para todo $E \in \mathcal{B}(X)$,

$$\mu(E) = \inf\{\mu(U) \mid U \supseteq E \text{ abierto}\}. \quad (\text{Regularidad exterior})$$

Demostración. Ver [13], pp. 42-47]. □

A las medidas que satisfacen las propiedades (I), (II) y (III) se les conoce como *medidas de Radon*. Podemos determinar el valor de una medida de Radon a través de los conjuntos importantes topológicamente, es decir, aquellos abiertos y aquellos compactos. La medida de cualquier conjunto $E \in \mathcal{B}(X)$ es el ínfimo de las medidas de todos los conjuntos abiertos que lo contienen y, si E es abierto o tiene medida finita, esta es el supremo de las medidas de todos los conjuntos compactos que contiene. Esto compatibiliza la estructura topológica con la de medida.

La medida de Lebesgue usual de \mathbb{R} es justamente la medida correspondiente a la integral $I(f) = \int_{-\infty}^{\infty} f(x) dx$ a través del teorema 1, y por lo tanto es una medida de Radon. Otro ejemplo en cualquier espacio topológico localmente compacto y de Hausdorff X se obtiene fijando $z \in X$ y definiendo la integral $I_z: \mathcal{B}(X) \rightarrow [0, \infty)$ por $I_z(f) = f(z)$ para cada $f \in \mathcal{C}_0(X)$. Por el teorema 1, existe una única medida de Radon $\delta_z: \mathcal{B}(X) \rightarrow [0, \infty]$ tal que $\int_X f(x) d\delta_z(x) = f(z)$. Esta medida se llama la medida de Dirac centrada en z .

A continuación, veremos cómo se compatibiliza una estructura topológica con una algebraica.

2.2. Grupos topológicos (topología + álgebra). Nos interesan espacios topológicos que, además de topología, tengan una estructura de grupo. Pero debemos asegurarnos que el producto $(x, y) \mapsto xy$ y la inversión $x \mapsto x^{-1}$ respeten la topología, es decir, sean funciones continuas. Diremos que un grupo G equipado con alguna topología de Hausdorff τ es un *grupo topológico* si las funciones $x \mapsto x^{-1}$ (de G a G) y $(x, y) \mapsto xy$ (de $G \times G$ a G) son continuas, donde $G \times G$ es dotado con la topología producto usual.

Los números reales \mathbb{R} con la topología euclidiana forman un grupo topológico, pues las funciones $x \mapsto -x$ y $(x, y) \mapsto x + y$ son continuas. Similarmente ocurre para el grupo de reales positivos $\mathbb{R}_+ = (0, \infty)$, pero con el producto $x \mapsto 1/x$ y $(x, y) \mapsto xy$. En general, cualquier grupo con la topología discreta –como el grupo de enteros \mathbb{Z} con la topología heredada de \mathbb{R} – es un grupo topológico, pues toda función cuyo dominio es discreto es continua. El grupo general lineal, de matrices invertibles de $n \times n$ con coeficientes en \mathbb{C} , es un grupo topológico si se le dota con la topología euclidiana de $(\mathbb{C}^n)^n$.

Es importante notar que la topología debe ser de Hausdorff para que un grupo sea un grupo topológico. Por ejemplo, si tomamos el grupo de números reales \mathbb{R} con la topología indiscreta $\tau_{\text{ind}} = \{\emptyset, \mathbb{R}\}$, que no es de Hausdorff, entonces $(\mathbb{R}, \tau_{\text{ind}})$ no es un grupo topológico, a pesar de que las funciones $x \mapsto -x$ y $(x, y) \mapsto x + y$ sí son continuas en este caso. Cabe mencionar, sin embargo, que algunos autores no imponen este requerimiento.

En los grupos topológicos, las operaciones de grupo preservan los conjuntos abiertos:

TEOREMA 2. Sean (G, τ) un grupo topológico y $A \in \tau$. Entonces

$$A^{-1} = \{a^{-1} \mid a \in A\}, \quad xA = \{xa \mid a \in A\} \quad y \quad Ax = \{ax \mid a \in A\}$$

son elementos de τ para todo $x \in G$.

Demostración. Verifiquemos que xA es un conjunto abierto. Los demás casos son similares. Definimos la función $f: G \rightarrow G$ por $f(y) = x^{-1}y$. Tenemos que

$$f^{-1}(A) = \{y \in G \mid f(y) \in A\} = \{y \in G \mid x^{-1}y \in A\} = \{y \in G \mid y \in xA\} = xA.$$

Es sencillo verificar que $y \mapsto (z, y)$ es una función continua para todo $z \in G$ y, como G es un grupo topológico, $(x, y) \mapsto xy$ también lo es. Por lo tanto, la composición $f: y \mapsto (x^{-1}, y) \mapsto x^{-1}y$ es una función continua. Por definición de continuidad, $xA = f^{-1}(A)$ es un conjunto abierto. \square

Del teorema anterior se sigue que si A es un elemento de $\mathcal{B}(G)$, entonces xA , Ax y A^{-1} también son elementos de $\mathcal{B}(G)$. Resulta natural preguntarse si existen medidas de Radon que le asignen la misma medida a estos conjuntos, sin importar quién es x . Es decir, buscamos medidas en G que sean compatibles tanto con el álgebra como con la topología.

2.3. La medida de Haar (topología + álgebra \rightarrow medida). Sea G un grupo topológico localmente compacto y $\mu: \mathcal{B}(G) \rightarrow [0, \infty]$ una medida. Decimos que μ es una *medida invariante izquierda (derecha)* si para todo $A \in \mathcal{B}(G)$ y $x \in G$, $\mu(xA) = \mu(A)$ ($\mu(Ax) = \mu(A)$). Una medida $\mu: \mathcal{B}(G) \rightarrow [0, \infty]$ es una *medida de Haar izquierda (derecha)* si μ es una medida de Radon no nula e invariante izquierda (derecha).

Dado que las medidas de Radon de G están completamente determinadas por las integrales de las funciones $\mathcal{C}_0(G)$ (ver teorema [1](#)), no es difícil probar (ver [\[10\]](#) 15.8, p. 193) que si $\mu: \mathcal{B}(G) \rightarrow [0, \infty]$ es una medida de Radon, entonces μ es una medida de Haar izquierda si y solamente si para cada $f \in \mathcal{C}_0(G)$ y $y \in G$,

$$\int_G f(yx) d\mu(x) = \int_G f(x) d\mu(x),$$

y similarmente para el caso derecho.

La medida de Lebesgue usual en \mathbb{R} es el primer ejemplo de una medida Haar izquierda (y derecha), pues, por el teorema de cambio de variable, tomando $u = y + x$, se tiene $du = dx$, y luego

$$\int_{-\infty}^{\infty} f(y+x) dx = \int_{-\infty}^{\infty} f(x+y) dx = \int_{-\infty}^{\infty} f(u) du = \int_{-\infty}^{\infty} f(x) dx$$

para cada $f \in \mathcal{C}_0(\mathbb{R})$ y $y \in \mathbb{R}$.

Otro ejemplo se obtiene al tomar el grupo de reales positivos $\mathbb{R}_+ = (0, \infty)$ cuya operación es el producto. Utilizando el teorema de cambio de variable con $u = yx$ y $du = y dx$ obtenemos

$$\int_0^{\infty} f(yx) dx = \int_0^{\infty} f(u) \frac{1}{y} du = \frac{1}{y} \int_0^{\infty} f(x) dx \neq \int_0^{\infty} f(x) dx,$$

si $y \neq 1$. Esto nos dice que la integral de Lebesgue usual $I(f) = \int_0^{\infty} f(x) dx$ no corresponde a una medida de Haar izquierda en este grupo. Sin embargo, no todo está perdido, hay muchas otras medidas que podríamos definir. Consideremos la integral $H: \mathcal{B}(\mathbb{R}_+) \rightarrow [0, \infty)$ dada por $H(f) = \int_0^{\infty} \frac{1}{x} f(x) dx$. Por el teorema [1](#), existe una única medida de Radon $\mu: \mathcal{B}(\mathbb{R}_+) \rightarrow [0, \infty)$ tal que $\int_0^{\infty} \frac{1}{x} f(x) dx = H(f) = \int_{\mathbb{R}_+} f(x) d\mu(x)$ para cada $f \in \mathcal{C}_0(\mathbb{R}_+)$. La medida μ sí es una medida de Haar izquierda en \mathbb{R}_+ pues, para cada $f \in \mathcal{C}_0(\mathbb{R}_+)$ y $y \in \mathbb{R}_+$,

$$\int_{\mathbb{R}_+} f(yx) d\mu(x) = \int_0^{\infty} \frac{1}{x} f(yx) dx = \int_0^{\infty} \frac{y}{u} f(u) \frac{1}{y} du = \int_0^{\infty} \frac{1}{u} f(u) du = \int_{\mathbb{R}_+} f(x) d\mu(x).$$

Un hecho con tintes milagrosos es que todos los grupos localmente compactos siempre poseen una medida de Haar, ¡y es única! –Bueno, casi única–. Este es el famoso teorema de Haar:

TEOREMA 3 (teorema de Haar). *Sea G un grupo topológico localmente compacto. Existe una medida de Haar izquierda (derecha) $\mu: \mathcal{B}(G) \rightarrow [0, \infty]$ que es única salvo por múltiplos reales positivos. Es decir, si tenemos otra medida de Haar izquierda (derecha) $\mu': \mathcal{B}(G) \rightarrow [0, \infty]$, existe $c > 0$ tal que $\mu' = c\mu$.*

Este teorema, originalmente postulado en un caso particular por Haar en 1933 [8], ha sido motivo de miles de artículos y centenas de libros de texto. Cualquiera de sus posibles demostraciones es extensa, por lo que no lo probaremos aquí. Se pueden encontrar tratamientos muy accesibles en los libros [10, 7, 4, 5] basadas en la primera demostración dada por Weil [15] y en la demostración posterior de Cartan [3]. Una prueba menos conocida es la de Bredon [2] que se encuentra tratada con detalle en la referencia [12].

3. EL TEOREMA DE HAAR INVERSO

El teorema de Haar permite construir una medida invariante a partir de una topología localmente compacta. El teorema de Haar inverso, originalmente demostrado por Weil en 1940 [15, ap. I] nos permite ir de regreso, es decir, construir una topología localmente compacta a partir de una medida invariante. En lo que resta del artículo, nos enfocaremos en demostrar este resultado. Seguiremos, en esencia, el camino original de Weil que se encuentra detallado en el libro de Halmos [9, cap. XII], aunque haremos algunas modificaciones sustanciales para evitar ciertos conceptos que han caído en desuso.

$$\begin{array}{ccc} \text{topología} + \text{álgebra} & \xrightarrow{\text{teorema de Haar}} & \text{medida} \\ \text{medida} + \text{álgebra} & \xrightarrow{\text{teorema de Haar inverso}} & \text{topología} \end{array}$$

3.1. Grupos medibles (medida + álgebra). Para poder definir una medida de Haar (una medida de Radon invariante) lo primero que fue necesario hacer fue armonizar la estructura topológica con la estructura algebraica. Esto lo hicimos pidiendo que las operaciones de grupo fueran continuas. Ahora queremos hablar de medidas invariantes sin una topología, por lo que debemos definir una σ -álgebra de tal manera que las operaciones de grupo respeten, ya no lo abierto, sino que lo medible. Recordemos que dados dos conjuntos X y Y con σ -álgebras \mathcal{A}_X y \mathcal{A}_Y , respectivamente, una función $f: X \rightarrow Y$, es *medible* de \mathcal{A}_X a \mathcal{A}_Y si para cada $A \in \mathcal{A}_Y$, se tiene $f^{-1}(A) \in \mathcal{A}_X$.

Diremos que un grupo G equipado con una σ -álgebra \mathcal{A} de subconjuntos de G es un *grupo medible* si la función $x \mapsto x^{-1}$ es medible de \mathcal{A} a \mathcal{A} y la función $(x, y) \mapsto xy$ es medible de $\mathcal{A} \otimes \mathcal{A}$ a \mathcal{A} , donde $\mathcal{A} \otimes \mathcal{A}$ es la σ -álgebra producto, que es la σ -álgebra más pequeña que contiene al conjunto de rectángulos $\{A \times B \mid A, B \in \mathcal{A}\}$.

Como ejemplo, consideremos el grupo de números reales \mathbb{R} con la suma y la σ -álgebra de Borel, $\mathcal{A} = \mathcal{B}(\mathbb{R})$. Veamos que $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es un grupo medible. Como \mathbb{R} es un grupo topológico, las funciones $x \mapsto -x$ y $(x, y) \mapsto x + y$ son continuas. Toda función continua es medible respecto a la σ -álgebra de Borel. Entonces $x \mapsto -x$ es medible de $\mathcal{B}(\mathbb{R})$ a $\mathcal{B}(\mathbb{R})$, y $(x, y) \mapsto x + y$ es medible de $\mathcal{B}(\mathbb{R} \times \mathbb{R})$ a $\mathcal{B}(\mathbb{R})$. Pero nosotros necesitamos que $(x, y) \mapsto x + y$ sea medible de $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ —no $\mathcal{B}(\mathbb{R} \times \mathbb{R})$ —a $\mathcal{B}(\mathbb{R})$. Por suerte, en los números reales, ocurre que $\mathcal{B}(\mathbb{R} \times \mathbb{R}) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. De hecho, es sencillo generalizar este argumento para cualquier grupo topológico segundo numerable G , pues en este caso siempre ocurre que $\mathcal{B}(G \times G) = \mathcal{B}(G) \otimes \mathcal{B}(G)$ (c.f. [6, 4.1.7, p. 119]).

Si bien, en general, una σ -álgebra es una estructura muy diferente a una topología, resulta que muchos resultados sobre grupos topológicos se pueden calcar para grupos medibles. Por ejemplo, si en el teorema [2] sustituimos τ por \mathcal{A} , se obtiene el siguiente resultado.

TEOREMA 4. Sean (G, \mathcal{A}) un grupo medible y $A \in \mathcal{A}$. Entonces

$$A^{-1} = \{a^{-1} \mid a \in A\}, \quad xA = \{xa \mid a \in A\} \quad y \quad Ax = \{ax \mid a \in A\}$$

son elementos de \mathcal{A} para todo $x \in G$.

Demostración. Calcar la demostración del teorema [2](#) pero sustituyendo

<i>conjunto abierto</i>	por	<i>conjunto medible,</i>	
<i>función continua</i>	por	<i>función medible,</i>	
<i>y grupo topológico</i>	por	<i>grupo medible.</i>	□

Sabiendo que las operaciones de grupo preservan la mensurabilidad, podemos volver a hablar de medidas invariantes, sin necesitar una topología. Una medida $\mu: \mathcal{A} \rightarrow [0, \infty]$ sobre un grupo medible (G, \mathcal{A}) es *invariante izquierda (derecha)* si $\mu(xA) = \mu(A)$ ($\mu(Ax) = \mu(A)$) para todo $A \in \mathcal{A}$ y $x \in G$.

Por ejemplo, en el grupo medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, la medida de Lebesgue es una medida invariante izquierda y derecha. Otro ejemplo sencillo es tomar un grupo finito $G = \{x_1, \dots, x_n\}$, con la σ -álgebra discreta $\mathcal{A} = \{A \mid A \subseteq G\}$. Es fácil verificar que G es un grupo medible utilizando que $\mathcal{A} \otimes \mathcal{A}$ es la σ -álgebra discreta de $G \times G$. En este caso, la medida de conteo $\mu: G \rightarrow \{0, 1, \dots, n\}$, donde $\mu(A)$ es igual al número de elementos de A , es una medida invariante izquierda, pues si $A \subseteq G$, entonces xA tiene el mismo número de elementos que A para todo $x \in G$.

A partir de ahora solo trabajaremos con invarianza izquierda, pero cabe mencionar que todos los siguientes resultados se pueden reescribir para medidas invariantes derechas, cambiando el orden de todos los productos. Antes de empezar a *topologizar* los grupos medibles con medidas invariantes izquierdas, presentamos un par de resultados y notaciones importantes. Dado X un conjunto y \mathcal{A} una σ -álgebra, cuando digamos que $f: X \rightarrow [-\infty, \infty]$ es medible en \mathcal{A} , tomaremos a los conjuntos de Borel de los reales extendidos $[-\infty, \infty]$ como la σ -álgebra para el contradominio, que es la σ -álgebra generada por los conjuntos de la forma $[a, b)$ con $-\infty \leq a < b \leq \infty$.

El siguiente teorema nos permite calcular una especie de *convolución* de conjuntos con una medida invariante izquierda y σ -finita.

TEOREMA 5 (del promedio). Sean (G, \mathcal{A}) un grupo medible y $\mu: \mathcal{A} \rightarrow [0, \infty]$ una medida invariante izquierda y σ -finita. Para cualesquiera $A, B \in \mathcal{A}$, la función $x \mapsto \mu(x^{-1}A \cap B)$ es medible y

$$\int_G \mu(x^{-1}A \cap B) d\mu(x) = \mu(A)\mu(B^{-1}).$$

Demostración. Ver [9](#), Sec. 59, teo. F, p. 261] o [12](#), 3.2.11]. □

Con el teorema del promedio podemos mostrar que si bien las inversiones o traslaciones derechas pueden no preservar la medida de un conjunto bajo una medida invariante izquierda y σ -finita, estas operaciones sí preservan su positividad.

COROLARIO 6. Sean (G, \mathcal{A}) un grupo medible y $\mu: \mathcal{A} \rightarrow [0, \infty]$ una medida invariante izquierda y σ -finita. Si $A \in \mathcal{A}$ satisface $\mu(A) > 0$, entonces $\mu(A^{-1}) > 0$ y $\mu(Aa) > 0$ para todo $a \in G$.

Demostración. El teorema del promedio nos dice que

$$\int_G \mu(x^{-1}A \cap A^{-1}) d\mu(x) = \mu(A)^2 > 0.$$

Por lo tanto, existe $x \in G$ tal que $B = x^{-1}A \cap A^{-1}$ tiene medida positiva, y entonces debe ocurrir que $\mu(A^{-1}) > 0$. Si repetimos el mismo argumento para el conjunto $a^{-1}B$ con medida $\mu(a^{-1}B) = \mu(B) > 0$, debe existir $y \in G$ tal que $C = y^{-1}a^{-1}B \cap (a^{-1}B)^{-1}$ tiene medida positiva. Dado que $C \subseteq B^{-1}a \subseteq Aa$, concluimos que $\mu(Aa) > 0$. □

El hecho de que la familia de conjuntos de medida positiva se mantenga invariante ante cualquiera de las operaciones de grupo resultará muy importante. Lo que haremos en la siguiente sección es postular como abiertos precisamente a ciertos conjuntos de medida positiva y con ello construir una topología para el grupo medible.

3.2. La topología de Weil (medida + álgebra → topología). Recordemos que es posible construir topologías a partir de bases locales de vecindades abiertas. Pidiendo ciertas condiciones sobre una familia de subconjuntos de un grupo, podemos garantizar que existe una topología donde este es un grupo topológico que tiene a esa familia como una base local de vecindades abiertas.

TEOREMA 7. *Sea G un grupo y consideremos una familia no vacía \mathcal{V} de subconjuntos de G tal que:*

- (a) *Para todo $V \in \mathcal{V}$, $e \in V$ (e es el elemento neutro de G)*
- (b) *Dados $U, V \in \mathcal{V}$, existe $W \in \mathcal{V}$ tal que $W \subseteq V \cap U$.*
- (c) *Para cada $U \in \mathcal{V}$ y $x \in U$, existe $V \in \mathcal{V}$ con $xV \subseteq U$.*
- (d) *Para cada $U \in \mathcal{V}$ existe $V \in \mathcal{V}$ tal que $VV^{-1} \subseteq U$.*
- (e) *Dados $U \in \mathcal{V}$ y $x \in G$, existe $V \in \mathcal{V}$ tal que $V \subseteq xUx^{-1}$.*
- (f) *Si $x \in G$ y $x \neq e$, entonces existe $V \in \mathcal{V}$, tal que $x \notin V$.*

Entonces la familia

$$\mathcal{V}(x) = \{xV \mid V \in \mathcal{V}\}$$

es una base local de vecindades abiertas para cada $x \in G$ en una única topología con la cual G es un grupo topológico.

Demostración. Ver [12] 1.1.12] o [1] 1.3.12, p. 22]. □

Para construir una topología en un grupo medible que lo convierta en un grupo topológico, podemos construir una familia \mathcal{V} como la del teorema 7. Utilicemos el siguiente ejemplo para motivar el procedimiento que desarrollaremos más adelante.

Consideremos el grupo medible \mathbb{R} con la suma y la σ -álgebra de Borel. Sobre él tenemos la medida invariante izquierda $\mu: \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ dada por la medida usual de Lebesgue. Ya sabemos que la topología usual hace que \mathbb{R} sea un grupo topológico. Veamos si podemos construir una base local de vecindades abiertas para esta topología utilizando solo la medida μ . Lo primero que se debería venir a la mente cuando uno dice *base local de vecindades abiertas para \mathbb{R}* es la familia de intervalos (bolas) abiertos. Consideremos

$$\mathcal{V} = \{B(\varepsilon, 0) \mid \varepsilon > 0\}, \quad \text{donde } B(\varepsilon, x) = \{y \in \mathbb{R} \mid |y - x| < \varepsilon\} = (x - \varepsilon, x + \varepsilon).$$

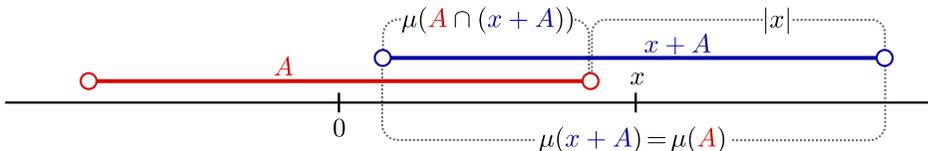
Es sencillo verificar que la familia \mathcal{V} satisface las condiciones del teorema 7 y entonces la familia

$$\mathcal{V}(x) = \{x + V \mid V \in \mathcal{V}\} = \{B(\varepsilon, x) \mid \varepsilon > 0\}$$

es una base local de vecindades abiertas para cada $x \in \mathbb{R}$ en una única topología con la cual \mathbb{R} es un grupo topológico, que es la topología usual de \mathbb{R} .

Observemos que para construir las bolas $B(\varepsilon, x)$ utilizamos el valor absoluto $|\cdot|$, que es una norma en \mathbb{R} . ¿Habría forma de calcular esta norma utilizando la medida de Lebesgue μ ? Es decir, dado $x \in \mathbb{R}$, ¿podemos calcular $|x|$ utilizando solo μ ?

Si tomamos $A = (-M, M) \in \mathcal{B}(\mathbb{R})$ donde $M > 0$, obtenemos $x + A = (x - M, x + M)$. Si $|x| \geq 2M$, entonces $A \cap (x + A) = \emptyset$, y por lo tanto $\mu(A \cap (x + A)) = 0$. Pero si $|x| < 2M$, el siguiente esquema debería convencer al lector de que $\mu(A \cap (x + A)) + |x| = \mu(A)$.



Por lo tanto,

$$|x| = \mu(A) - \mu(A \cap (x + A)) \quad (\text{si } |x| < 2M).$$

Regresemos al caso general. Fijemos un grupo medible (G, \mathcal{A}) y una medida invariante izquierda $\mu: \mathcal{A} \rightarrow [0, \infty]$ cualesquiera. Inspirados por el ejemplo anterior, definimos

$$\|x\|_\mu^A = \mu(A) - \mu(A \cap xA)$$

para cada $A \in \mathcal{A}$ con $0 < \mu(A) < \infty$ y $x \in G$.

LEMA 8. *Para cualesquiera $x, y \in G$ y $A \in \mathcal{A}$ con $0 < \mu(A) < \infty$ se satisfacen las siguientes condiciones:*

- (I) $\|e\|_\mu^A = 0$.
- (II) $\|x\|_\mu^A = \|x^{-1}\|_\mu^A$. (Simetría)
- (III) $\|xy\|_\mu^A \leq \|x\|_\mu^A + \|y\|_\mu^A$. (Desigualdad del triángulo)

Demostración. Verifiquemos cada inciso por separado.

- (I) Directamente de la definición

$$\|e\|_\mu^A = \mu(A) - \mu(A \cap eA) = \mu(A) - \mu(A \cap A) = \mu(A) - \mu(A) = 0.$$

- (II) Como μ es invariante izquierda,

$$\mu(A \cap xA) = \mu(x^{-1}(A \cap xA)) = \mu(x^{-1}A \cap x^{-1}xA) = \mu(x^{-1}A \cap A),$$

y entonces

$$\|x\|_\mu^A = \mu(A) - \mu(A \cap xA) = \mu(A) - \mu(x^{-1}A \cap A) = \|x^{-1}\|_\mu^A.$$

- (III) Primero notemos que $A \supseteq A \cap (x^{-1}A \cup yA) = (A \cap x^{-1}A) \cup (A \cap yA)$, y entonces

$$\begin{aligned} \mu(A) &\geq \mu((A \cap x^{-1}A) \cup (A \cap yA)) \\ &= \mu(A \cap x^{-1}A) + \mu(A \cap yA) - \mu(A \cap x^{-1}A \cap yA). \end{aligned}$$

Ahora, de nuevo por la invarianza izquierda de μ se tiene

$$\mu(A \cap xyA) = \mu(x^{-1}A \cap yA),$$

y como $A \cap x^{-1}A \cap yA \subseteq x^{-1}A \cap yA$, obtenemos

$$\begin{aligned} \|xy\|_\mu^A &= \mu(A) - \mu(x^{-1}A \cap yA) \leq \mu(A) - \mu(A \cap x^{-1}A \cap yA) \\ &\leq \mu(A) - \mu(A \cap x^{-1}A) + \mu(A) - \mu(A \cap yA) = \|x^{-1}\|_\mu^A + \|y\|_\mu^A. \end{aligned}$$

Por la simetría, $\|x^{-1}\|_\mu^A = \|x\|_\mu^A$, dándonos la desigualdad deseada. \square

Podríamos estar tentados a llamar a las funciones $\|\cdot\|_\mu^A$ *normas de grupo*, en similitud a las normas en espacios vectoriales. Sin embargo, las normas en espacios vectoriales cumplen la propiedad de que $\|v\| = 0$ si y solamente si v es el vector cero. Aquí, puede pasar que $\|x\|_\mu^A = 0$ aunque $x \neq e$. Por ejemplo, en un grupo de medida finita $0 < \mu(G) < \infty$ (más adelante construimos un ejemplo explícito de este tipo), se tiene $\|x\|_\mu^G = \mu(G) - \mu(G) = 0$ para todo $x \in G$. Por ello, a las funciones $\|\cdot\|_\mu^A$ que cumplen las propiedades del lema [8](#) se les llama *seudonormas de grupo* [11](#).

Con las pseudonormas $\|\cdot\|_\mu^A$ definiremos una familia de bolas, de la manera usual. Para cada $\varepsilon > 0$ y $A \in \mathcal{A}$ con $\varepsilon < \mu(A) < \infty$, escribimos

$$B_\mu(\varepsilon, A) = \{x \in G \mid \|x\|_\mu^A < \varepsilon\}.$$

Pedimos que $\varepsilon < \mu(A)$ porque es el caso interesante. Cuando $\varepsilon > \mu(A)$, se tiene $\|x\|_\mu^A \leq \mu(A) < \varepsilon$ sin importar quién es $x \in G$. Denotamos por \mathcal{V}_μ a la familia de todas las bolas $B_\mu(\varepsilon, A)$, es decir,

$$\mathcal{V}_\mu = \{B_\mu(\varepsilon, A) \mid A \in \mathcal{A}, \varepsilon > 0 \text{ y } \varepsilon < \mu(A) < \infty\}.$$

Notemos para todo $B_\mu(\varepsilon, A) \in \mathcal{V}_\mu$, se tiene $B_\mu(\varepsilon, A) \in \mathcal{A}$, pues la función

$$f(x) = \mu(xA \cap A) = \mu(x^{-1}A \cap A)$$

es una función medible (ver el teorema 5) y $B_\mu(\varepsilon, A) = f^{-1}([0, \mu(A) - \varepsilon])$. También notemos que por la simetría de $\|\cdot\|_\mu^A$, se tiene $B_\mu(\varepsilon, A) = B_\mu(\varepsilon, A)^{-1}$. Los conjuntos $B_\mu(\varepsilon, A)$ en \mathcal{V}_μ serán vecindades para e en la topología que buscamos.

TEOREMA 9. La familia \mathcal{V}_μ satisface los incisos (a), (c) y (d) del teorema 7.

Demostración. Estos incisos se deducen a partir de las propiedades del lema 8. Sea $B_\mu(\varepsilon, A) \in \mathcal{V}_\mu$.

(a) Como $\|e\|_\mu^A = 0 < \varepsilon$, tenemos $e \in B_\mu(\varepsilon, A)$.

(c) Para $x \in B_\mu(\varepsilon, A)$, definimos $\delta = \varepsilon - \|x\|_\mu^A > 0$. Por la definición de \mathcal{V}_μ , se tiene $\mu(A) > \varepsilon$, y entonces $\mu(A) > \delta$. Por lo tanto, podemos considerar la bola $B_\mu(\delta, A) \in \mathcal{V}_\mu$. Si $y \in B_\mu(\delta, A)$,

$$\|xy\|_\mu^A \leq \|x\|_\mu^A + \|y\|_\mu^A < \|x\|_\mu^A + \delta = \varepsilon,$$

y así $xy \in B_\mu(\varepsilon, A)$. Por lo tanto, $x B_\mu(\delta, A) \subseteq B_\mu(\varepsilon, A)$.

(d) Se tiene $B_\mu(\varepsilon/2, A) B_\mu(\varepsilon/2, A)^{-1} \subseteq B_\mu(\varepsilon, A)$, pues si

$$x \in B_\mu(\varepsilon/2, A) B_\mu(\varepsilon/2, A)^{-1},$$

entonces $x = ab^{-1}$ para ciertos $a, b \in B_\mu(\varepsilon/2, A)$. Luego,

$$\|x\|_\mu^A \leq \|a\|_\mu^A + \|b\|_\mu^A < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

y por lo tanto $x \in B_\mu(\varepsilon, A)$. \square

Para demostrar los demás incisos, necesitaremos hacer un poco más de trabajo, e imponer algunas condiciones adicionales sobre la medida invariante. Regresemos al ejemplo de los números reales.

Consideremos la medida invariante izquierda de Lebesgue $\mu: \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$. Tomando $M > 0$ y $A = (-M, M) \in \mathcal{B}(\mathbb{R})$, ya vimos que $\|x\|_\mu^A = \min\{|x|, 2M\}$. Por lo tanto, si $0 < \varepsilon < \mu(A) = 2M$,

$$B_\mu(\varepsilon, A) = \{x \in \mathbb{R} \mid \|x\|_\mu^A < \varepsilon\} = \{x \in \mathbb{R} \mid |x| < \varepsilon\} = (-\varepsilon, \varepsilon).$$

Escribamos $C - C = \{x - y \mid x, y \in C\}$ para cualquier $C \subseteq \mathbb{R}$. Entonces

$$A - A = \{x - y \mid -M < x < M, -M < y < M\} = (-2M, 2M) \supseteq (-\varepsilon, \varepsilon) = B_\mu(\varepsilon, A).$$

Por otro lado, si definimos $E = (-\varepsilon/4, \varepsilon/4) \in \mathcal{B}(\mathbb{R})$, tenemos que

$$E - E = \{x - y \mid x, y \in E\} = (-\varepsilon/2, \varepsilon/2) \subseteq (-\varepsilon, \varepsilon) = B_\mu(\varepsilon, A).$$

Es decir, hemos visto que

$$E - E \subseteq B_\mu(\varepsilon, A) \subseteq A - A$$

donde $E \in \mathcal{B}(\mathbb{R})$ satisface $0 < \mu(E) < \infty$. Resulta que esto siempre ocurre, independientemente del grupo y de quién es A , siempre y cuando la medida μ sea σ -finita.

En un grupo G cualquiera, dados $A, B \subseteq G$, escribimos

$$AB = \{ab \mid a \in A, b \in B\} \quad \text{y} \quad A^{-1} = \{a^{-1} \mid a \in A\},$$

y así, en particular, el conjunto

$$AA^{-1} = \{a_1 a_2^{-1} \mid a_1, a_2 \in A\}$$

generaliza la notación $A - A$ que teníamos para \mathbb{R} . El resultado que encontramos en el ejemplo anterior se generaliza de la siguiente manera.

LEMA 10. Sean (G, \mathcal{A}) un grupo medible y $\mu: \mathcal{A} \rightarrow [0, \infty]$ una medida invariante izquierda y σ -finita. Si $B_\mu(\varepsilon, A) \in \mathcal{V}_\mu$, entonces:

(i) $B_\mu(\varepsilon, A) \subseteq AA^{-1}$.

(ii) Dado $F \in \mathcal{A}$ con $\mu(F) > 0$, existe $E \subseteq F$ medible con $0 < \mu(E) < \infty$ y tal que

$$EE^{-1} \subseteq B_\mu(\varepsilon, A).$$

Al inciso (ii) se le conoce como el lema de fragmentación de Weil.

Demostración.

- (i) Sea $x \in B_\mu(\varepsilon, A)$. Por definición, $0 < \mu(A) - \varepsilon < \mu(A \cap xA)$ y entonces $A \cap xA \neq \emptyset$. Luego, existe $a \in A \cap xA$, y por lo tanto $a = xa'$ para algún $a' \in A$. Así, $x = aa'^{-1} \in AA^{-1}$.
- (ii) Ver [9] Sec. 62, teo. A, p. 270] o [12] 3.3.12]. \square

El lema [10] nos permite demostrar dos más de las propiedades que necesitamos para poder aplicar el teorema [7].

TEOREMA 11. *Sea (G, \mathcal{A}) un grupo medible y $\mu: \mathcal{A} \rightarrow [0, \infty]$ una medida invariante izquierda y σ -finita. La familia \mathcal{V}_μ satisface los incisos (b) y (e) del teorema [7].*

Demostración. El lema [10] nos dice que todo conjunto en \mathcal{V}_μ contiene a (y está contenido en) un conjunto de la familia

$$\mathcal{N} = \{AA^{-1} \mid A \in \mathcal{A}, \quad 0 < \mu(A) < \infty\}.$$

Por lo tanto, basta que veamos que la familia \mathcal{N} satisface los incisos (b) y (e).

- (b) Sean $U = AA^{-1} \in \mathcal{N}$ y $V = BB^{-1} \in \mathcal{N}$. Buscamos $W = CC^{-1} \in \mathcal{N}$ con $W \subseteq U \cap V$. Como $\mu(A) > 0$, por el corolario [6], $\mu(A^{-1}) > 0$. Luego $\mu(A^{-1})\mu(B) > 0$ y, por el teorema del promedio (teorema [5]),

$$\int_G \mu(x^{-1}A^{-1} \cap B^{-1}) d\mu(x) > 0.$$

Entonces debe existir algún $x \in G$ tal que $\mu(x^{-1}A^{-1} \cap B^{-1}) > 0$. Tomando $D = Ax \cap B$, tenemos $D^{-1} = x^{-1}A^{-1} \cap B^{-1}$, y entonces $\mu(D) > 0$. Como μ es σ -finita, existe $C \subseteq D$ medible con $0 < \mu(C) < \infty$. Así, $CC^{-1} \in \mathcal{N}$ y además, dado que $D \subseteq B$, tenemos $DD^{-1} \subseteq BB^{-1}$ y, dado que $Dx^{-1} \subseteq A$, tenemos

$$DD^{-1} = Dx^{-1}xD^{-1} = Dx^{-1}(Dx^{-1})^{-1} \subseteq AA^{-1}.$$

Por lo tanto,

$$W = CC^{-1} \subseteq DD^{-1} \subseteq AA^{-1} \cap BB^{-1} = U \cap V.$$

- (e) Si $U = AA^{-1} \in \mathcal{N}$ y $x \in G$ entonces tomando $V = (xA)(xA)^{-1} \in \mathcal{N}$, tenemos

$$V = xAA^{-1}x^{-1} = xUx^{-1}. \quad \square$$

Resta verificar que la familia \mathcal{V}_μ satisface el inciso (f) del teorema [7]. Este inciso es el que nos garantiza que la topología generada por el teorema [7] sea de Hausdorff, como requerimos en la definición de grupo topológico. Desafortunadamente, no cualquier medida invariante y σ -finita hace que \mathcal{V}_μ satisfaga esta propiedad.

Un ejemplo sencillo es tomar el grupo medible $(\mathbb{R}, \mathcal{A}_{\text{ind}})$, con la σ -álgebra indiscreta $\mathcal{A}_{\text{ind}} = \{\emptyset, \mathbb{R}\}$ y la medida invariante izquierda $\mu_{\text{ind}}: \mathcal{A}_{\text{ind}} \rightarrow \{0, 1\}$ dada por $\mu_{\text{ind}}(\emptyset) = 0$ y $\mu_{\text{ind}}(\mathbb{R}) = 1$. Para todo $x \in \mathbb{R}$ se tiene

$$\|x\|_{\mu_{\text{ind}}}^{\mathbb{R}} = \mu_{\text{ind}}(\mathbb{R}) - \mu_{\text{ind}}(\mathbb{R} \cap (x + \mathbb{R})) = \mu_{\text{ind}}(\mathbb{R}) - \mu_{\text{ind}}(\mathbb{R}) = 1 - 1 = 0.$$

Por lo tanto, si $0 < \varepsilon < \mu_{\text{ind}}(\mathbb{R}) = 1$, se tiene $B_{\mu_{\text{ind}}}(\varepsilon, \mathbb{R}) = \mathbb{R}$, y luego $\mathcal{V}_{\mu_{\text{ind}}} = \{\mathbb{R}\}$. Si $x \neq 0$, entonces $x \in V$ para todo $V \in \mathcal{V}_{\mu_{\text{ind}}} = \{\mathbb{R}\}$, contrario al inciso (f) del teorema [7]. Sin embargo, evidentemente, la única topología en la cual la familia $\{\mathbb{R}\}$ es una base local de vecindades abiertas es la topología indiscreta, que precisamente no es de Hausdorff.

La solución a este obstáculo es directamente imponer que se satisfaga el inciso (f) del teorema [7]. Diremos que una medida invariante izquierda $\mu: \mathcal{A} \rightarrow [0, \infty]$ en un grupo medible (G, \mathcal{A}) es *separadora* si \mathcal{V}_μ satisface el inciso (f) del teorema [7]. Es sencillo verificar que esto es equivalente a pedir que para cada $x \in G$ con $x \neq e$, exista $A \in \mathcal{A}$ con $\mu(A \cap xA) < \mu(A) < \infty$. Esta propiedad en realidad no es muy restrictiva; imponerla solo nos permite excluir ciertas medidas triviales, como μ_{ind} . Por ejemplo, la medida de Lebesgue $\mu: \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ es separadora, pues para cada

$x \in \mathbb{R}$ con $x \neq 0$, tomando $A = (-|x|/2, |x|/2)$ obtenemos $A \cap (x + A) = \emptyset$ y entonces $\mu(A \cap (x + A)) = 0 < |x| = \mu(A)$.

Si una medida es invariante izquierda, σ -finita, separadora y no nula diremos que es una *medida de Weil*. Hemos demostrado que si (G, \mathcal{A}) es un grupo medible y $\mu: \mathcal{A} \rightarrow [0, \infty]$ una medida de Weil, entonces se satisfacen todos los incisos del teorema 7. Este teorema requiere, adicionalmente, que la familia \mathcal{V}_μ sea no vacía. Por ello pedimos que la medida de Weil μ sea no nula. Esta condición, junto con la σ -finitud, garantiza la existencia de algún $A \in \mathcal{A}$ con $0 < \mu(A) < \infty$, y en consecuencia $B_\mu(\varepsilon, A) \in \mathcal{V}_\mu$ para todo $\varepsilon \in (0, \mu(A))$. Así, podemos aplicar el teorema 7, dándonos una única topología τ_μ con base local de vecindades abiertas

$$\{xV \mid V \in \mathcal{V}_\mu\} \quad (\mathcal{V}_\mu = \{B_\mu(\varepsilon, A) \mid A \in \mathcal{A}, 0 < \varepsilon < \mu(A) < \infty\}).$$

Esta topología se llama la *topología de Weil* inducida por μ .

Recordemos que las topologías en las cuales existía una única medida de Haar (salvo múltiplos reales positivos) eran aquellas localmente compactas. La topología de Weil es *casi* localmente compacta, es decir, equipado con ella, el grupo medible es un subgrupo denso de un grupo topológico localmente compacto. Para demostrar este hecho, introducimos el siguiente concepto.

Dado un grupo topológico G , decimos que $V \subseteq G$ es *totalmente acotado* si para cualquier vecindad U de e , existen $x_1, \dots, x_n \in G$ tales que

$$V \subseteq \bigcup_{i=1}^n x_i U.$$

Esta definición generaliza aquella de los espacios métricos, donde se dice que un conjunto es totalmente acotado si se puede cubrir con una cantidad finita de bolas de cualquier radio.

Resulta que un grupo topológico es un subgrupo denso de un grupo localmente compacto si existe una vecindad V de e que sea totalmente acotada. Este hecho fue originalmente demostrado por Weil en [14, cap. 3]. Se puede consultar la demostración en la referencia [1, p. 398]. Veremos que la topología inducida por una medida de Weil satisface siempre esta propiedad.

LEMA 12. Sean (G, \mathcal{A}) un grupo medible y $\mu: \mathcal{A} \rightarrow [0, \infty]$ una medida de Weil. Existe una vecindad $V \in \mathcal{V}_\mu$ que es totalmente acotada en la topología de Weil inducida por μ . Más aún, si $E \in \mathcal{A}$ es totalmente acotado en esta topología, entonces $\mu(E) < \infty$.

Demostración. Lo primero que haremos es construir $W \in \mathcal{V}_\mu$ de medida finita. Como μ es σ -finita y no nula, existe $A \in \mathcal{A}$ tal que $0 < \mu(A) < \infty$. Por el corolario 6, $\mu(A^{-1}) > 0$, y por la σ -finitud podemos suponer, sin pérdida de generalidad, que $\mu(A^{-1}) < \infty$. Fijemos un $\varepsilon \in (0, \mu(A))$ y sea $W = B_\mu(\varepsilon, A) \in \mathcal{V}_\mu$. Del lema 10 deducimos que $W \subseteq AA^{-1}$ y, por definición, para cualquier $x \in W$ se satisface que $\mu(A) - \mu(x^{-1}A \cap A) = \|x\|_\mu^A < \varepsilon$. Despejando, obtenemos que

$$\frac{\mu(x^{-1}A \cap A)}{\mu(A) - \varepsilon} > 1$$

para cualquier $x \in W$. Del teorema del promedio (teorema 5) se sigue que

$$\begin{aligned} \mu(W) &= \int_W 1 \, d\mu(x) \leq \int_W \frac{\mu(x^{-1}A \cap A)}{\mu(A) - \varepsilon} \, d\mu(x) \leq \int_G \frac{\mu(x^{-1}A \cap A)}{\mu(A) - \varepsilon} \, d\mu(x) \\ &= \frac{\mu(A)\mu(A^{-1})}{\mu(A) - \varepsilon} < \infty. \end{aligned}$$

Así, hemos obtenido una vecindad $W \in \mathcal{V}_\mu$ de medida finita. Notemos que esto demuestra la segunda parte del enunciado, pues si $E \in \mathcal{A}$ es totalmente acotado, se puede cubrir con una cantidad finita de conjuntos de la forma xW ($x \in G$), que tienen medida finita.

Veamos que en efecto existe una vecindad $V \in \mathcal{V}_\mu$ totalmente acotada. Como en la demostración del teorema 9, inciso (d), podemos construir $V \in \mathcal{V}_\mu$ tal que $VV^{-1} \subseteq W$. Demostraremos que V es totalmente acotada. Por contradicción, supongamos que existe una vecindad U de e tal que $V \not\subseteq \bigcup_{i=1}^n x_i U$ para cualesquiera $x_1, \dots, x_n \in G$. Como \mathcal{V}_μ es una base local para e , podemos suponer que $U \in \mathcal{V}_\mu$.

Por medio de la propiedad de U construyamos una sucesión de puntos en V . Tomemos $x_1 = e \in V$. Como $V \not\subseteq x_1 U$, existe $x_2 \in V$ con $x_2 \notin x_1 U$. De nuevo, como $V \not\subseteq x_1 U \cup x_2 U$ existe $x_3 \in V$ con $x_3 \notin x_1 U \cup x_2 U$. Siguiendo de esta manera, para cada $n \in \mathbb{N}$, dado que $V \not\subseteq \bigcup_{i=1}^n x_i U$, podemos tomar $x_{n+1} \in V$ con $x_{n+1} \notin \bigcup_{i=1}^n x_i U$.

Por el lema 10 (II), existe $F \in \mathcal{A}$ con $0 < \mu(F) < \infty$, $F \subseteq V^{-1}$ y $FF^{-1} \subseteq U$. Si $n > m$, entonces $x_n \notin x_m U \supseteq x_m FF^{-1}$, es decir, no existe $z \in F$ tal que $x_n z \in x_m F$, por lo que $x_n F \cap x_m F = \emptyset$. Con lo anterior, hemos construido una familia de conjuntos ajenos con medida finita y positiva $\{x_n F \mid n \in \mathbb{N}\}$. Además, se tiene

$$x_n F \subseteq x_n V^{-1} \subseteq VV^{-1} \subseteq W$$

para todo $n \in \mathbb{N}$, por lo que

$$\mu(W) \geq \mu\left(\bigcup_{n \in \mathbb{N}} x_n F\right) = \sum_{n=1}^{\infty} \mu(x_n F) = \sum_{n=1}^{\infty} \mu(F) = \infty,$$

lo cual contradice que $\mu(W) < \infty$. Esta contradicción viene de suponer que la vecindad V no es totalmente acotada, por lo que debe serlo. \square

Hemos llegado, al fin, al esperado teorema de Haar inverso. A partir de una medida de Weil pudimos construir una topología que es densamente encajable en un grupo localmente compacto. Por el teorema de Haar, este grupo localmente compacto debe tener una medida de Haar izquierda. Lo que resulta fantástico es que esta medida de Haar está completamente determinada por la medida de Weil con la que empezamos. Esto nos dice que no solo ocurre que las topologías localmente compactas producen medidas de Radon invariantes, sino que además las medidas invariantes en cualquier grupo medible son, en esencia, medidas de Radon en una topología localmente compacta.

TEOREMA 13 (de Haar inverso de Weil). *Consideremos un grupo medible (G, \mathcal{A}) y una medida de Weil $\mu: \mathcal{A} \rightarrow [0, \infty]$. Con la topología de Weil inducida por μ , G es un subgrupo denso de un grupo topológico localmente compacto \widehat{G} . Más aún, si $f \in \mathcal{C}_0(\widehat{G})$, la restricción $f|_G$ es medible en \mathcal{A} y*

$$\int_G f(x) d\mu(x) = \int_{\widehat{G}} f(x) d\widehat{\mu}(x),$$

donde $\widehat{\mu}: \widehat{G} \rightarrow [0, \infty]$ es la medida de Haar izquierda de \widehat{G} .

[Notemos que por el teorema de representación de RMK (teorema 11), la medida de Haar $\widehat{\mu}$ está totalmente determinada por las integrales de las funciones $f \in \mathcal{C}_0(\widehat{G})$. Lo que nos dice este teorema es que μ determina completamente a $\widehat{\mu}$.]

Demostración. El lema 12 nos garantiza que existe un grupo topológico localmente compacto \widehat{G} del cual G es un subgrupo denso. La siguiente afirmación nos permite conectar los conjuntos totalmente acotados de G con los conjuntos compactos de \widehat{G} .

Afirmación: Si $E \subseteq G$ está contenido en un compacto de \widehat{G} , entonces E es totalmente acotado en G .

En efecto, pensemos que $E \subseteq \widehat{C} \subseteq \widehat{G}$ con \widehat{C} compacto, y sea $V \subseteq G$ una vecindad abierta de e . Denotemos por $\widehat{F} \subseteq \widehat{G}$ a la cerradura de E en \widehat{G} . Entonces \widehat{F} es compacto, pues es cerrado dentro del compacto \widehat{C} . Sea $\widehat{V} \subseteq \widehat{G}$ una vecindad abierta de e tal que $\widehat{V} \cap G = V$. Si $y \in \widehat{F}$, como $y\widehat{V}^{-1}$ es una vecindad abierta de y , existe $x \in E \cap y\widehat{V}^{-1}$, y entonces $y \in x\widehat{V}$. Por lo tanto, la familia de conjuntos $x\widehat{V}$ con $x \in E$ forman una cubierta abierta del compacto \widehat{F} , y deben existir $x_1, x_2, \dots, x_n \in E$ tales que

$\widehat{F} \subseteq \bigcup_{i=1}^n x_i \widehat{V}$. Luego

$$E \subseteq \widehat{F} \cap G \subseteq \bigcup_{i=1}^n x_i \widehat{V} \cap G = \bigcup_{i=1}^n x_i V,$$

es decir, E es totalmente acotado, demostrando la afirmación.

Para demostrar el resto del teorema, requeriremos del siguiente hecho sobre las funciones en $\mathcal{C}_0(\widehat{G})$. Toda $f \in \mathcal{C}_0(\widehat{G})$ es uniformemente continua por la izquierda (derecha), es decir, para todo $\varepsilon > 0$ existe una vecindad V de e tal que para cada $x, y \in G$, si $x \in yV$ (Vy), entonces $|f(x) - f(y)| < \varepsilon$. [10, 15.4, p. 185].

Tomemos $f \in \mathcal{C}_0(\widehat{G})$ y veamos que $f|_G: G \rightarrow [0, \infty)$ es medible en \mathcal{A} , es decir, que para $A \in \mathcal{B}([0, \infty))$, se tiene $(f|_G)^{-1}(A) \in \mathcal{A}$. Dado que $\mathcal{B}([0, \infty))$ es la σ -álgebra generada por los intervalos de la forma $[\alpha, \infty)$ con $\alpha \geq 0$, basta demostrar que $E \in \mathcal{A}$, donde $E = (f|_G)^{-1}([\alpha, \infty)) = \{x \in G \mid f(x) > \alpha\}$. En efecto, para cada $n \in \mathbb{N}$ definamos $E_n = \{x \in G \mid f(x) \geq \alpha + 1/n\}$. Como f es uniformemente continua por la izquierda, existe una vecindad $W \subseteq G$ de e que, sin pérdida de generalidad, está en $\mathcal{V}_\mu \subseteq \mathcal{A}$, tal que si $x \in yW$, entonces $|f(x) - f(y)| < 1/n$. Luego $E_n W \subseteq E$, pues si $y \in E_n$ y $x \in yW$, tenemos

$$f(x) = f(y) + f(x) - f(y) \geq \alpha + \frac{1}{n} + f(x) - f(y) \geq \alpha + \frac{1}{n} - |f(x) - f(y)| > \alpha.$$

Ahora, recordemos las funciones en $\mathcal{C}_0(\widehat{G})$ valen cero fuera de un conjunto compacto de \widehat{G} . Como $f \in \mathcal{C}_0(\widehat{G})$ y $f(x) \geq \alpha + 1/n > 0$ para todo $x \in E_n$, tenemos que E_n debe estar contenido en un conjunto compacto de \widehat{G} y entonces la afirmación garantiza que E_n es totalmente acotado en G . Sean $x_1, x_2, \dots, x_m \in G$ tales que

$$E_n \subseteq E'_n \subseteq E_n W \subseteq E,$$

donde $E'_n = \bigcup_{i=1}^m x_i W \in \mathcal{A}$. Tenemos que

$$E = \bigcup_{n \in \mathbb{N}} E_n \subseteq \bigcup_{n \in \mathbb{N}} E'_n \subseteq E.$$

y entonces $E = \bigcup_{n \in \mathbb{N}} E'_n \in \mathcal{A}$. Así, $f|_G$ es medible en \mathcal{A} .

De nuevo, como $f \in \mathcal{C}_0(\widehat{G})$, el conjunto $S = \{x \in G \mid f(x) \neq 0\}$ debe de estar contenido en un conjunto compacto de \widehat{G} , y entonces, por la afirmación, S es totalmente acotado en G . Por el lema [12], $\mu(S) < \infty$, y entonces

$$0 \leq \int_G f(x) d\mu(x) = \int_S f(x) d\mu(x) \leq \int_S M d\mu(x) = \mu(S)M < \infty,$$

con $M = \max\{f(x) \mid x \in S\}$. En consecuencia, podemos definir $I: \mathcal{C}_0(\widehat{G}) \rightarrow [0, \infty)$ por $I(f) = \int_G f(x) d\mu(x)$. Es sencillo verificar que I es una integral (i.e. es aditiva y homogénea). Por el teorema de representación de RMK (teorema [1]), existe una única medida de Radon $\widehat{\mu}: \mathcal{B}(\widehat{G}) \rightarrow [0, \infty]$ tal que

$$\int_G f(x) d\mu(x) = \int_{\widehat{G}} f(x) d\widehat{\mu}(x)$$

para cada $f \in \mathcal{C}_0(\widehat{G})$.

Resta verificar que $\widehat{\mu}$ es una medida de Haar izquierda en \widehat{G} . Recordemos que por el teorema de representación de RMK (teorema [1]), los valores de $\widehat{\mu}$ están completamente determinados por las integrales de las funciones en $\mathcal{C}_0(\widehat{G})$, y basta que veamos que

$$\int_{\widehat{G}} f(\widehat{y}\widehat{x}) d\widehat{\mu}(\widehat{x}) = \int_{\widehat{G}} f(\widehat{x}) d\widehat{\mu}(\widehat{x})$$

para todo $\widehat{y} \in \widehat{G}$ y $f \in \mathcal{C}_0(\widehat{G})$.

Fijemos $f \in \mathcal{C}_0(\widehat{G})$. Si $\widehat{y} = y \in G$, definiendo $L_y: G \rightarrow G$ por $L_y(x) = yx$ y considerando la medida imagen $\mu L_y^{-1}: \mathcal{A} \rightarrow [0, \infty]$ (c.f. [4, 2.6.8, p. 76]), se obtiene

$\mu L_y^{-1}(A) = \mu(y^{-1}A) = \mu(A)$ para cada $A \in \mathcal{A}$ (por la invarianza izquierda de μ), y entonces

$$\begin{aligned} \int_{\widehat{G}} f(y\widehat{x}) d\widehat{\mu}(\widehat{x}) &= \int_G f(yx) d\mu(x) = \int_G (f \circ L_y)(x) d\mu(x) = \int_G f(x) d(\mu L_y^{-1})(x) \\ &= \int_G f(x) d\mu(x) = \int_{\widehat{G}} f(\widehat{x}) d\widehat{\mu}(\widehat{x}). \end{aligned}$$

Para $\widehat{y} \notin G$, debemos hacer un poco más de trabajo. Tomemos $\varepsilon > 0$ y veamos que

$$\left| \int_{\widehat{G}} f(\widehat{y}\widehat{x}) d\widehat{\mu}(\widehat{x}) - \int_{\widehat{G}} f(\widehat{x}) d\widehat{\mu}(\widehat{x}) \right| \leq \varepsilon.$$

Como \widehat{G} es localmente compacto, existe una vecindad compacta $\widehat{C} \subseteq \widehat{G}$ de e , y como $f \in \mathcal{C}_0(\widehat{G})$, existe un compacto $\widehat{S} \subseteq \widehat{G}$ tal que $f(\widehat{x}) = 0$ si $\widehat{x} \notin \widehat{S}$. El conjunto

$$\widehat{N} = \widehat{y}^{-1}\widehat{C}\widehat{S}$$

es compacto (el producto de compactos es compacto [10, 4.4, p. 17]). Como $\widehat{\mu}$ es una medida de Radon, es localmente finita [ver teorema 1(1)], y entonces $\widehat{\mu}(\widehat{N}) < \infty$. Sea $\delta > 0$ tal que $\delta \widehat{\mu}(\widehat{N}) \leq \varepsilon$. Como f es uniformemente continua por la derecha, existe una vecindad $\widehat{V} \subseteq \widehat{G}$ de e tal que

$$|f(\widehat{v}) - f(\widehat{w})| < \delta \quad \text{si} \quad \widehat{w} \in \widehat{V}\widehat{v}.$$

Sin pérdida de generalidad, podemos suponer que $\widehat{V} \subseteq \widehat{C}^{-1}$. Como G es denso, existe $y \in G \cap \widehat{V}\widehat{y}$ y, entonces, para cada $\widehat{x} \in G$ tenemos $y\widehat{x} \in \widehat{V}\widehat{y}\widehat{x}$, de donde obtenemos que $|f(\widehat{y}\widehat{x}) - f(y\widehat{x})| < \delta$.

Además, notemos que si $\widehat{x} \notin \widehat{N}$, entonces $\widehat{y}\widehat{x} \notin \widehat{C}\widehat{S} \supseteq \widehat{S}$, por lo que $f(\widehat{y}\widehat{x}) = 0$, y también

$$y\widehat{x} \notin y\widehat{y}^{-1}\widehat{C}\widehat{S} \supseteq y\widehat{y}^{-1}\widehat{V}^{-1}\widehat{S} \supseteq \widehat{S}, \quad (\text{pues } e \in y\widehat{y}^{-1}\widehat{V}^{-1}, \text{ ya que } y \in \widehat{V}\widehat{y})$$

por lo que $f(y\widehat{x}) = 0$. Por lo tanto, $|f(\widehat{y}\widehat{x}) - f(y\widehat{x})| = 0$ para todo $\widehat{x} \notin \widehat{N}$ y entonces

$$\begin{aligned} \left| \int_{\widehat{G}} f(\widehat{y}\widehat{x}) d\widehat{\mu}(\widehat{x}) - \int_{\widehat{G}} f(\widehat{x}) d\widehat{\mu}(\widehat{x}) \right| &= \left| \int_{\widehat{G}} f(\widehat{y}\widehat{x}) d\widehat{\mu}(\widehat{x}) - \int_{\widehat{G}} f(y\widehat{x}) d\widehat{\mu}(\widehat{x}) \right| \\ &\leq \int_{\widehat{G}} |f(\widehat{y}\widehat{x}) - f(y\widehat{x})| d\widehat{\mu}(\widehat{x}) \\ &= \int_{\widehat{N}} |f(\widehat{y}\widehat{x}) - f(y\widehat{x})| d\widehat{\mu}(\widehat{x}) \\ &\leq \int_{\widehat{N}} \delta d\widehat{\mu}(\widehat{x}) = \delta \widehat{\mu}(\widehat{N}) \leq \varepsilon. \quad \square \end{aligned}$$

AGRADECIMIENTOS

Agradezco a la Dra. Natalia Jonard Pérez por sus detallados comentarios, a la Dra. Judith Campos Cordero por haberme invitado a realizar este artículo y al árbitro anónimo por sus valiosas correcciones.

REFERENCIAS

- [1] Arkhangel'skii, A. V. & Tkachenko, M., Topological Groups and Related Structures. Atlantis Press, Francia, 2008, ISBN 9789491216350.
- [2] Bredon, G. E., *A new treatment of the Haar integral*. The Michigan Math. Jour., 10, no. 4, 365–373, 1963, DOI:10.1307/mmj/1028998972, <https://projecteuclid.org/euclid.mmj/1028998972>.
- [3] Cartan, H., *Sur la mesure de Haar*. C. R. Acad. Sci. Paris, 211, 759–762, 1940.
- [4] Cohn, D., Measure Theory. Birkhäuser, Estados Unidos, 2ª ed., 2013, ISBN 9781461469551.
- [5] Diestel, J. & Spalsbury, A., The Joys of Haar Measure. American Mathematical Society, Estados Unidos, 2014, ISBN 9781470414115.
- [6] Dudley, R. M., Real Analysis and Probability. Cambridge University Press, Cambridge New York, 2002, ISBN 0521007542.
- [7] Folland, G. B., Real Analysis: Modern Techniques and Their Applications. PAM, Wiley, 2ª ed., 1999, ISBN 9780471317166.

- [8] Haar, A., *Der Massbegriff in der Theorie der Kontinuierlichen Gruppen*. The Ann. of Math., 34, no. 1, 147, 1933, DOI:10.2307/1968346, <https://www.jstor.org/stable/1968346>.
- [9] Halmos, P., *Measure Theory*. Springer Science, Estados Unidos, 1950, ISBN 9780387900889.
- [10] Hewitt, E. & Ross, K. A., *Abstract Harmonic Analysis*, vol. 1. Springer, Estados Unidos, 1979, ISBN 9780387941905.
- [11] Pettis, B. J., *On continuity and openness of homomorphisms in topological groups*. The Ann. of Math, 52, no. 2, 293, 1950, DOI:10.2307/1969471, <https://www.jstor.org/stable/1969471>
- [12] Pilatowsky Cameo, S., *El teorema de Haar y su inverso*. Tesis de Licenciatura, Universidad Nacional Autónoma de México, Ciudad Universitaria, 2021. Disponible en TESIUNAM: https://tesiunam.dgb.unam.mx/F/?func=service&doc_number=000811324&service_type=MEDIA
- [13] Rudin, W., *Real and Complex Analysis*. McGraw-Hill, Estados Unidos, 1987, ISBN 9780070542341.
- [14] Weil, A., *Sur les Espaces à Structure Uniforme et sur la Topologie Générale*. Hermann, Francia, 1937.
- [15] Weil, A., *L'Intégration dans les Groupes Topologiques et ses Applications*. Hermann, Francia, 1940, ISBN 9782705611453.

Dirección del autor:

Facultad de Ciencias,

Universidad Nacional Autónoma de México,

Ciudad Universitaria, 04510 CDMX.

ORCID: <https://orcid.org/0000-0001-9843-6491>

e-mail: saulpilatowsky@ciencias.unam.mx



EL TEOREMA DE METRIZACIÓN DE URYSOHN

FIDEL CASARRUBIAS SEGURA

RESUMEN. Demostramos que en cualquier espacio topológico normal T_1 que tiene una base a lo más numerable existe una pseudométrica que define a su topología. Con esto damos la demostración del clásico teorema de metrización de Urysohn que trata sobre la metrizabilidad de espacios topológicos regulares T_0 que tienen una base a lo más numerable.

1. INTRODUCCIÓN

El concepto de distancia entre dos puntos es algo muy intuitivo. Frechet dió una formulación matemática de la distancia entre dos puntos, llamada métrica, abstrayéndola de la recta real, del plano y del espacio tridimensional. Los espacios métricos, como llamamos hoy día a las parejas formadas por un conjunto no vacío y una métrica, aparecen en muchas áreas de las matemáticas como herramientas fundamentales.

La distancia entre puntos se puede utilizar para definir la distancia de un punto a un conjunto y con ello definir una topología; para hacerlo basta acordar que todos los puntos cuya distancia al conjunto A es igual a cero, son cercanos a A y definir la cerradura de A como el conjunto de tales puntos. Por esta razón podemos pensar a la noción de espacio topológico como una axiomatización de la idea de cercanía de un punto a un conjunto. Los espacios topológicos cuya topología proviene de una métrica son llamados *metrizables*. Estos espacios tienen muy buenas propiedades, y en cierta forma, podemos pensar que su estructura topológica es sencilla.

¿Cuándo un espacio topológico es metrizable? Tratar de responder a esta pregunta conlleva la idea de buscar condiciones (necesarias y) suficientes que nos permitan construir una métrica y reconstruir a la topología con esa métrica. Hacer esto es crear un teorema de metrización.

Probablemente el teorema de metrización más conocido es el teorema de metrización de Urysohn. Este teorema fue originalmente demostrado para espacios topológicos normales T_1 segundo-numerables en el famoso artículo *Zum Metrizizationproblem* ([4]), y poco tiempo después fue extendido a la clase de los espacios T_3 segundo-numerables por Tychonoff.

La prueba de Urysohn en realidad muestra que todo espacio normal, T_1 , segundo-numerable es un espacio pseudometrizable; es decir, un espacio en el cual es posible definir una pseudométrica de manera tal que la topología original del espacio es igual a la topología inducida por la pseudométrica.

Nuestro propósito principal en esta nota es dar una demostración simple y detalla de este clásico teorema de (pseudo)metrización. Exhibimos explícitamente la técnica empleada por Urysohn para construir una pseudométrica que induzca la topología de cualquier espacio normal segundo-numerable.

La prueba de Urysohn hace uso de su muy famoso lema. En esta nota, también proporcionamos una prueba del lema de Urysohn; y la hacemos en dos partes. Primero aislamos en un resultado auxiliar la técnica para crear funciones continuas a partir de cubiertas abiertas particulares, que están indexadas por subconjuntos densos de la recta real. Después demostramos el lema de Urysohn haciendo uso de este resultado

2010 *Mathematics Subject Classification.* 54A05, 54D10, 54D15, 54E35.

Palabras clave. Espacio topológico, espacio regular, espacio normal, metrizabilidad.

técnico. Esta manera de presentar la demostración del lema de Urysohn la hace más asequible para estudiantes de cursos básicos de topología.

Presentamos también la demostración original (en lenguaje moderno) de Tychonoff acerca de que todo espacio regular segundo-numerable es un espacio normal, resultado con el cual se extiende el teorema de metrización de Urysohn de la clase de los espacios T_4 segundo-numerables a la clase de los espacios regulares T_0 que son segundo-numerables.

2. TERMINOLOGÍA Y NOTACIÓN

Una topología τ en un conjunto no vacío X es una colección de subconjuntos de X que es cerrada bajo uniones de cualquier cantidad de conjuntos y cerrada bajo intersecciones de cantidades finitas de conjuntos (y contiene a los conjuntos \emptyset y X). A la pareja ordenada (X, τ) se le llama *espacio topológico*.

Una *base* para una topología τ es una subcolección $\mathcal{B} \subseteq \tau$ tal que para todo $U \in \tau$ y todo elemento $x \in U$, existe $B \in \mathcal{B}$ tal que $x \in B \subseteq U$. Un espacio topológico (X, τ) es *segundo-numerable* si τ tiene una base a lo más numerable.

En esta nota distinguimos a los espacios normales de los espacios T_4 . Para ser más precisos, diremos que un espacio topológico (X, τ) es *normal* si para cualquier par de subconjuntos cerrados y ajenos F y G existen abiertos ajenos U y V tales que $F \subseteq U$ y $G \subseteq V$. Y diremos que (X, τ) es un espacio T_4 si es normal y además satisface el axioma de separación T_1 .

También distinguiremos a los espacios regulares de los espacios T_3 . Un espacio topológico (X, τ) es *regular* si para cada subconjunto cerrado F de X y cada punto $x \in X \setminus F$, existen abiertos ajenos A, B tales que $x \in A$ y $F \subseteq B$. Y X es un espacio T_3 si es un espacio regular T_0 .

Recuerde que un espacio X es T_1 (respectivamente, T_0) si para cualquier par de elementos diferentes x y y en X es posible hallar un abierto U tal que $U \cap \{x, y\} = \{x\}$ (respectivamente, tal que $|U \cap \{x, y\}| = 1$).

En algunas partes de esta nota hacemos uso del resultado que muestra que el límite uniforme de una sucesión de funciones continuas real valuadas y definidas en un espacio topológico X , es una función continua de valores reales. La demostración de este hecho básico puede ser consultada en [1, 6.4.3].

3. EL LEMA DE URYSOHN

El lema de Urysohn [5] es uno de los resultados más relevantes y fundamentales, y más conocido, de la topología general. En él se demuestra que en los espacios normales hay una gran cantidad de funciones continuas de valores reales con buenas propiedades de separación. La «maquinaria» que construye esas funciones continuas es aislada en el siguiente lema técnico.

LEMA 1 ([2]). *Suponga que (X, τ) es un espacio topológico y que D es un subconjunto denso de la recta real \mathbb{R} que tiene asociada una cubierta abierta $\mathcal{U} = \{U_\alpha\}_{\alpha \in D}$ de X con las siguientes propiedades:*

1. Si $\alpha, \beta \in D$ y $\alpha < \beta$ entonces $\text{cl}(U_\alpha) \subseteq U_\beta$.
2. $\bigcap_{\alpha \in D} U_\alpha = \emptyset$.

Entonces la función $f : X \rightarrow \mathbb{R}$ definida por:

$$f(x) = \inf\{\alpha \in D : x \in U_\alpha\}$$

es continua.

¹Si X es regular y T_0 , y x y y son elementos distintos de X , entonces por el axioma T_0 existe un abierto U tal que $|U \cap \{x, y\}| = 1$. Si $x \in U$ y $y \notin U$ entonces $x \notin \text{cl}(\{y\})$. Como X es regular, existen abiertos ajenos V y W tales que $x \in V$ y $y \in \text{cl}(\{y\}) \subseteq W$. De manera análoga, si $y \in U$ y $x \notin U$ entonces existen abiertos ajenos V y W tales que $y \in V$ y $x \in \text{cl}(\{x\}) \subseteq W$. Así, X es un espacio Hausdorff.

Demostración. Debido a que \mathcal{U} es una cubierta de X , cada conjunto $A_x = \{\alpha \in D : x \in U_\alpha\}$ es no vacío, y también es acotado inferiormente en \mathbb{R} . En efecto, si A_x no es acotado inferiormente entonces para cualquier elemento $\beta \in D$ es posible hallar un elemento $\alpha \in A_x$ de modo que $\alpha < \beta$. Aplicando (1) tenemos que $\text{cl}(U_\alpha) \subseteq U_\beta$. Luego como $\alpha \in A_x$, se tiene que $x \in U_\alpha \subseteq \text{cl}(U_\alpha) \subseteq U_\beta$. Entonces $\beta \in A_x$. Por lo tanto, $A_x = D$ y por ello $x \in \bigcap_{\alpha \in D} U_\alpha$; lo que contradice la condición 2.

El que los conjuntos A_x sean no vacíos y acotados inferiormente garantiza que el número real $f(x) = \inf\{\alpha \in D : x \in U_\alpha\}$ está bien definido.

Verifiquemos ahora que f es continua. Suponga que $x_0 \in X$ es un elemento cualquiera y que $\epsilon > 0$ es un número positivo arbitrario. Como $f(x_0) = \inf A_{x_0} < f(x_0) + \epsilon$ existe $\beta_0 \in A_{x_0}$ tal que $f(x_0) \leq \beta_0 < f(x_0) + \epsilon$. Por ser D denso en \mathbb{R} podemos elegir $\beta_1, \beta_2 \in D$ tales que $f(x_0) - \epsilon < \beta_1 < \beta_2 < f(x_0)$. La definición de $f(x_0)$ implica que $x_0 \notin U_{\beta_2}$. Por (1), $\beta_1 < \beta_2$ implica que $\text{cl}(U_{\beta_1}) \subseteq U_{\beta_2}$. Luego $x_0 \notin \text{cl}(U_{\beta_1})$. Por lo tanto $U_{\beta_0} \setminus \text{cl}(U_{\beta_1})$ es un abierto de X que contiene a x_0 . Resulta además que $f[U_{\beta_0} \setminus \text{cl}(U_{\beta_1})] \subseteq (f(x_0) - \epsilon, f(x_0) + \epsilon)$. Efectivamente, si $x \in U_{\beta_0} \setminus \text{cl}(U_{\beta_1})$, entonces $\beta_0 \in A_x = \{\beta \in D : x \in U_\beta\}$ y $A_x \subseteq (\beta_1, \infty)$. Luego, $\beta_1 \leq f(x) \leq \beta_0$ porque $f(x) = \inf A_x$, β_1 es cota inferior de A_x y β_0 es elemento de A_x . Consecuentemente, $f[U_{\beta_0} \setminus \text{cl}(U_{\beta_1})] \subseteq [\beta_1, \beta_0] \subseteq (f(x_0) - \epsilon, f(x_0) + \epsilon)$. Por lo tanto f es continua en x_0 . \square

El lema de Urysohn trabaja en espacios normales, en su demostración usamos la siguiente propiedad de estos espacios topológicos, la cual es muy sencilla de demostrar:

Si F es un subconjunto cerrado y U es un subconjunto abierto de un espacio normal X tal que $F \subseteq U$, entonces existe un abierto V de X tal que $F \subseteq V \subseteq \text{cl}(V) \subseteq U$.

En efecto, por la normalidad de X existen abiertos ajenos V y W que contienen, respectivamente, a los subconjuntos cerrados ajenos F y $X \setminus U$. Debido a que $X \setminus W$ es cerrado, podemos concluir que $F \subseteq V \subseteq \text{cl}(V) \subseteq X \setminus W \subseteq U$, como se desea.

La razón por la cuál se elige la anterior propiedad de espacios normales es la siguiente. Para aplicar el lema [1](#) en un espacio normal debemos construir una cubierta abierta de dicho espacio normal. La construcción se hace seleccionando en pasos sucesivos a conjuntos abiertos. La condición 1 del lema [1](#) nos dice que en cada paso podemos aplicar la propiedad de espacios normales: $F \subseteq V \subseteq \text{cl}(V) \subseteq U$, y elegir al abierto V que nos proporciona esa propiedad, para que éste forme parte de la cubierta. En cada uno de esos pasos, la aplicación de dicha propiedad de espacios normales, y por consecuencia, la elección del abierto V , está supeditada a la forma en que está acomodado un determinado número racional en el intervalo $[0, 1]$ en relación a una cantidad finita de otros números racionales. Lo curioso es que para descubrir ese acomodo, esencialmente aplicaremos el método que nos enseñaron en educación básica para saber, por ejemplo, dónde debemos colocar al racional $\frac{5}{7}$ en el $[0, 1]$ cuando el intervalo $[0, 1]$ está «dividido en octavos». De esta manera, como alguna vez escuche decir a un buen amigo, «la prueba del lema de Urysohn es esencialmente saber cómo están “acomodados” los números racionales en $[0, 1]$ ».

LEMA 2 (de Urysohn). *Sea X un espacio normal. Supongamos que F y G son subconjuntos cerrados X que son ajenos. Entonces existe una función continua $f : X \rightarrow [0, 1]$ tal que $f[F] \subseteq \{0\}$ y $f[G] \subseteq \{1\}$.*

Demostración. La idea clave para la demostración es construir una cubierta abierta para X que satisfaga las hipótesis del lema [1](#). Para ello utilizaremos un subconjunto denso especial de la recta real y a los subconjuntos cerrados ajenos F y G del espacio normal X . Una vez hecho esto el lema [1](#) nos construye la función continua que necesitamos.

Definición de un subconjunto denso D de \mathbb{R} adecuado. Consideremos una enumeración inyectiva [2](#) $D_0 = \{q_n : n = 0, 1, 2, \dots\}$ del conjunto $\mathbb{Q} \cap [0, 1]$ tal que $q_0 = 0$

²Considere una biyección f de $\mathbb{N} \cup \{0\}$ en $\mathbb{Q} \cap [0, 1]$ tal que $f(0) = 0$ y $f(1) = 1$.

y $q_1 = 1$. Definamos $D = (-\infty, 0) \cup D_0 \cup (1, \infty)$. Es fácil demostrar que D es un subconjunto denso de \mathbb{R} .

Ahora construiremos a una cubierta abierta \mathcal{U} de X a partir de los subconjuntos cerrados F y G , indexada por el conjunto D , y que cumpla las condiciones del lema [1](#).

Construcción de la cubierta $\mathcal{U} = \{U_\alpha : \alpha \in D\}$. Primero definamos

$$U_\alpha = \begin{cases} \emptyset & \alpha < 0, \\ X & \alpha > 1. \end{cases}$$

Con los elementos de \mathcal{U} que tenemos definidos hasta este momento podemos ya verificar que la colección \mathcal{U} será una cubierta de X y que se cumplirá la condición 2 del lema [1](#). Observe que estos elementos cumplen también la condición 1.

De esta manera, nuestra única preocupación ahora es construir a los restantes elementos de la cubierta \mathcal{U} con el único requisito de que ellos cumplan la condición 1 del lema [1](#). Es decir, debemos construir una colección $\mathcal{U}_0 = \{U_{q_n} : n = 0, 1, 2, \dots\}$ de subconjuntos abiertos de X que tengan la siguiente propiedad:

(a) si $r < s$ con $r, s \in D_0$ entonces $\text{cl}(U_r) \subseteq U_s$.

Como X es un espacio normal, y F es un cerrado tal que $F \subseteq X \setminus G$, existe un subconjunto abierto V de X tal que $F \subseteq V \subseteq \text{cl}(V) \subseteq X \setminus G$. Defina $U_{q_0} = V$ y $U_{q_1} = X \setminus G$.

Supongamos ahora que $n \geq 2$ y que hemos construido los subconjuntos abiertos $U_{q_0}, U_{q_1}, U_{q_2}, \dots, U_{q_n}$ de tal forma que ellos satisfacen la condición (a). La idea para construir al subconjunto abierto $U_{q_{n+1}}$ radica en saber «cómo está acomodado el racional q_{n+1} en $[0, 1]$ » en relación a los números racionales q_0, q_1, \dots, q_n . Para saberlo, simplemente aplicamos nuestros conocimientos de educación básica y definimos

$$r = \text{máx}\{q_k : k \in \{0, 1, \dots, n\} \ \& \ q_k < q_{n+1}\}$$

y

$$s = \text{mín}\{q_l : l \in \{0, 1, \dots, n\} \ \& \ q_{n+1} < q_l\}.$$

Es claro que $r, s \in \{q_0, q_1, \dots, q_n\}$ y que $r < q_{n+1} < s$. Por nuestra hipótesis de recursión, tenemos que $\text{cl}(U_r) \subseteq U_s$. Como X es un espacio normal, existe W abierto en X tal que $\text{cl}(U_r) \subseteq W \subseteq \text{cl}(W) \subseteq U_s$. Defina $U_{q_{n+1}} = W$.

No es difícil verificar ahora que los conjuntos

$$\{U_{q_0}, U_{q_1}, U_{q_2}, \dots, U_{q_n}, U_{q_{n+1}}\}$$

satisfacen las condiciones requeridas. Esto completa la construcción recursiva de la familia $\mathcal{U}_0 = \{U_r : r \in D\} \subseteq \tau_X$ que satisface la condición (a).

Definamos $\mathcal{U} = \{U_\alpha : \alpha \in \mathbb{R} \setminus [0, 1]\} \cup \mathcal{U}_0$. Es fácil verificar que la colección \mathcal{U} satisface las hipótesis del lema [1](#). La función $f : X \rightarrow \mathbb{R}$ que se define en el lema [1](#) a partir de la cubierta \mathcal{U} que hemos construido es continua. Además $f[X] \subseteq [0, 1]$, $f(F) \subseteq \{0\}$ y $f(G) \subseteq \{1\}$. Efectivamente, si existiera un elemento $x \in X$ para el cual $f(x) = \inf\{\alpha \in D : x \in U_\alpha\} < 0$, entonces existe $\beta < 0$ tal que $\beta \in \{\alpha \in D : x \in U_\alpha\}$; pero entonces $x \in U_\beta = \emptyset$, lo cual es evidentemente una contradicción. Por otro lado, como para toda $\alpha \in (1, \infty)$, $U_\alpha = X$, se tiene que $(1, \infty) \subseteq \{\alpha \in D : x \in U_\alpha\}$. Luego, $f(x) = \inf\{\alpha \in D : x \in U_\alpha\} \leq 1 = \inf(1, \infty)$. En consecuencia, $f[X] \subseteq [0, 1]$. Por otra parte, si $x \in F$ entonces $x \in U_{q_0}$. Luego, $0 = q_0 \in \{\alpha \in D : x \in U_\alpha\}$. Por ello, $f(x) = \inf\{\alpha \in D : x \in U_\alpha\} \leq 0$. Es decir, $f(x) = 0$. Finalmente, si $x \in G$ entonces $x \notin U_{q_1}$, y por ello, $x \notin U_\alpha$ para todo $\alpha \leq 1$. En consecuencia, $\{\alpha \in D : x \in U_\alpha\} = (1, \infty)$. Y entonces $f(x) = 1$. \square

Uno de los corolarios inmediatos del lema de Urysohn es la siguiente caracterización de la normalidad en términos de funciones continuas.

COROLARIO 3. *Las siguientes condiciones son equivalentes para cualquier espacio topológico (X, τ) .*

1. (X, τ) es un espacio normal;
2. para cualesquiera subconjuntos cerrados ajenos F_1 y F_2 de X existe una función $f : X \rightarrow [0, 1]$ tal que $f[F_1] \subseteq \{0\}$ y $f[F_2] \subseteq \{1\}$.

4. EL TEOREMA DE PSEUDOMETRIZACIÓN DE URYSOHN PARA ESPACIOS NORMALES

Si X es un conjunto no vacío, entonces una función ρ de valores reales no negativos definida en $X \times X$ es llamada *pseudométrica* si cumple las siguientes tres condiciones: (1) $\rho(x, x) = 0$ para cada $x \in X$, (2) $\rho(x, y) = \rho(y, x)$ para todo $x, y \in X$; (3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ para cualesquiera $x, y, z \in X$.

Si ρ es una pseudométrica sobre X , entonces definimos a las bolas abiertas como los conjuntos $B(x, r) = \{y \in X : \rho(x, y) < r\}$, donde $x \in X$ y $r > 0$. La colección $\tau_\rho = \{\emptyset\} \cup \{A : (\forall x \in A) (\exists r > 0 : B(x, r) \subseteq A)\}$ es la topología sobre X inducida por la pseudométrica ρ .

La diferencia entre una métrica y una pseudométrica es simplemente la condición (1), la cual se sustituye en el caso de una métrica por la condición: (1') $\rho(x, y) = 0$ si y sólo si $x = y$ ³. Desde el punto de vista topológico esta diferencia es codificada por el axioma de separación T_0 .

PROPOSICIÓN 4 (Kolmogorov). *Las siguientes condiciones son equivalentes para cualquier pseudométrica ρ sobre X :*

- (1) ρ es una métrica;
- (2) (X, τ_ρ) es un espacio T_0 .

Demostración. (1) \Rightarrow (2). Si $x, y \in X$ son elementos diferentes, entonces $U = B(x, r) \in \tau_\rho$ donde $r = d(x, y)$. Es claro que $|U \cap \{x, y\}| = 1$. Esto muestra que (X, τ_ρ) es un espacio topológico T_0 .

(2) \Rightarrow (1). Suponga que $x, y \in X$ son tales que $x \neq y$. Como (X, τ_ρ) es T_0 , existe un abierto U tal que $|U \cap \{x, y\}| = 1$. Podemos suponer sin perder generalidad que $x \in U$. Entonces existe $r > 0$ tal que $B(x, r) \subseteq U$. Como $y \notin U$, tenemos que $y \notin B(x, r)$. Consecuentemente, $\rho(x, y) \geq r > 0$. Esto demuestra que la pseudométrica ρ es una métrica. \square

Un espacio topológico (X, τ) es *pseudometrizable* si existe una pseudométrica ρ de modo que $\tau_\rho = \tau$. En el siguiente teorema se resume el método de Urysohn para pseudometrizar a cualquier espacio normal segundo-numerable.

TEOREMA 5. *Todo espacio normal, T_1 y segundo numerable es pseudometrizable.*

Demostración. Supongamos que (X, τ) es un espacio normal que es segundo numerable. Fijemos una base numerable \mathcal{B} para τ . Definamos como \mathcal{D} al siguiente conjunto:

$$\mathcal{D} = \{(U, V) \in \mathcal{B} \times \mathcal{B} \mid \text{cl}(U) \subseteq V\}.$$

Debido a que \mathcal{B} es numerable, la colección \mathcal{D} también lo es. Podemos entonces suponer que $\mathcal{D} = \{D_n : n \in \mathbb{N}\}$. Como cada elemento $D_n \in \mathcal{D}$ es una pareja ordenada $D_n = (U, V)$ escribiremos $D_n = (U_n, V_n)$ para enfatizar que los elementos U_n y V_n forman la pareja ordenada D_n ; de esta manera

$$\mathcal{D} = \{D_n = (U_n, V_n) \in \mathcal{B} \times \mathcal{B} \mid \text{cl}(U_n) \subseteq V_n; n \in \mathbb{N}\}.$$

Aplicando el lema de Urysohn, podemos fijar para todo número natural n una función continua $f_n : X \rightarrow [0, 1]$ tal que $f_n[\text{cl}(U_n)] \subseteq \{0\}$ y $f_n[X \setminus V_n] \subseteq \{1\}$.

Afirmación 1. La relación $\rho(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} |f_n(x) - f_n(y)|$ define una pseudométrica $\rho : X \times X \rightarrow \mathbb{R}$ en X .

³Es claro que toda métrica es una pseudométrica, pero la función constante cero es una pseudométrica que no es una métrica en cualquier conjunto con al menos dos elementos.

Demostración de la Afirmación 1. Para cualesquiera $x, y \in X$, la sucesión

$$\left(\sum_{i=1}^n \frac{1}{2^i} |f_i(x) - f_i(y)| \right)_{n \in \mathbb{N}}$$

es acotada superiormente en \mathbb{R} porque $\sum_{i=1}^n \frac{1}{2^i} |f_i(x) - f_i(y)| \leq \sum_{i=1}^n \frac{1}{2^{i-1}} < 2$ para toda $n \in \mathbb{N}$. Así la serie $\sum_{n=1}^{\infty} \frac{1}{2^n} |f_n(x) - f_n(y)|$ es convergente.

Por otro lado, es fácil probar que $\rho(x, y) \geq 0$, $\rho(x, x) = 0$, $\rho(x, y) = \rho(y, x)$ y además $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ para cualesquiera $x, y, z \in X$; esto es, ρ es una pseudométrica en X . \square

La idea ahora es demostrar que la pseudométrica ρ genera a la topología de X . Para hacerlo, primero demostraremos que si fijamos la segunda variable en la función ρ entonces la función resultante es continua.

Afirmación 2. Para cada $z \in X$, la función $f_z : X \rightarrow \mathbb{R}$ definida por $f_z(x) = \rho(x, z)$ es una función continua de (X, τ) en $(\mathbb{R}, \tau_{\mathbb{R}})$.

Demostración de la Afirmación 2. Fijemos un elemento cualquiera $z \in X$. Para cada $k \in \mathbb{N}$ y cada $x \in X$, definimos $h_{z,k}(x) = \frac{1}{2^k} |f_k(x) - f_k(z)|$. No es difícil darse cuenta que cada función $h_{z,k} : X \rightarrow \mathbb{R}$ es una función continua en (X, τ) . Definamos ahora $g_{z,n} = \sum_{k=1}^n h_{z,k}$ ($n \in \mathbb{N}$). Como $g_{z,n}$ es suma de funciones continuas, cada $g_{z,n}$ es una función continua.

La sucesión de funciones continuas $(g_{z,n})_{n \in \mathbb{N}}$ converge uniformemente en X a la función f_z : supongamos que $\epsilon > 0$. Fijemos un $N \in \mathbb{N}$ de modo que $\frac{1}{2^{N-1}} < \frac{\epsilon}{2}$. Aplicando lo establecido en los párrafos anteriores, es fácil verificar que $|g_{z,n}(x) - g_{z,m}(x)| < \frac{\epsilon}{2}$ para todo $m, n \in \mathbb{N}$ con $n, m \geq N$ y para cualquier $x \in X$. Resulta que $|g_{z,n}(x) - f_z(x)| < \epsilon$ para todo $n \geq N$ y todo elemento $x \in X$. Efectivamente, supongamos que $x \in X$ y que $n \geq N$ son elementos cualesquiera. Como $f_z(x) = \lim_{n \rightarrow \infty} g_{z,n}(x)$, existe $M \in \mathbb{N}$ tal que $|g_{z,m}(x) - f_z(x)| < \frac{\epsilon}{2}$ para todo $m \geq M$. Fijemos $m \in \mathbb{N}$ de modo que $m \geq N + M$. Entonces $n, m \geq N$ y $m > M$. Luego $|g_{z,n}(x) - f_z(x)| \leq |g_{z,n}(x) - g_{z,m}(x)| + |g_{z,m}(x) - f_z(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Por lo tanto, $(g_{z,m})_{m \in \mathbb{N}}$ converge uniformemente en X a la función f_z . \square

Como f_z es el límite uniforme de una sucesión de funciones continuas real valuadas, f_z es continua para todo $z \in X$. Obsérvese que $f_z(x) = \rho(x, z)$ para cada $x \in X$.

Para finalizar la demostración, verificaremos que la topología τ_ρ inducida por ρ en X coincide con la topología τ . Para hacerlo es suficiente demostrar la siguiente afirmación.

Afirmación 3.

- Para cada $z \in X$ y $r > 0$, existe un $B \in \mathcal{B}$ tal que $z \in B \subseteq B(z, r) = \{y \in X : \rho(z, y) < r\}$.
- Para cada $B \in \mathcal{B}$ y cada $x \in B$, existe una $\epsilon > 0$ tal que $B(x, \epsilon) \subseteq B$.

Demostración de la afirmación 3. (a) Supongamos que $z \in X$. Por la afirmación 2, la función $f_z : X \rightarrow \mathbb{R}$ definida por $f_z(x) = \rho(x, z)$ es continua en X cuando en X se considera a la topología τ . Consecuentemente, por la continuidad de f_z en z , si $r > 0$ existe un elemento $B \in \mathcal{B}$ tal que $z \in B$ y $f_z(B) \subseteq (f_z(z) - r, f_z(z) + r) = (-r, r)$. Entonces, para cada $x \in B$ se tiene que $\rho(x, z) = f_z(x) \in (-r, r)$; por lo tanto, $z \in B \subseteq B(z, r)$.

(b) Supongamos, por el contrario, que existen $B \in \mathcal{B}$ y $x \in B$ tales que $B(x, \epsilon) \setminus B \neq \emptyset$ para cada $\epsilon > 0$. En particular, para toda $n \in \mathbb{N}$, podemos fijar $x_n \in B(x, \frac{1}{2^n}) \setminus B$. Como $x \in B$, B es abierto básico y X es normal T_1 , existe $U \in \mathcal{B}$ tal que $x \in U \subseteq \text{cl}(U) \subseteq B$. Luego la pareja ordenada (U, B) pertenece a \mathcal{D} . Por ello, existe $N \in \mathbb{N}$ tal que $(U, B) = D_N = (U_N, V_N)$. Como para cada $n \in \mathbb{N}$ sucede que $x_n \in X \setminus B$, tenemos que $f_N(x) = 0$ y $f_N(x_n) = 1$ para toda $n \in \mathbb{N}$. En particular, $f_N(x) = 0$ y $f_N(x_N) = 1$. En consecuencia, $\rho(x, x_N) = \sum_{n=1}^{\infty} \frac{1}{2^n} |f_n(x) - f_n(x_N)| \geq \frac{1}{2^N} |f_N(x) - f_N(x_N)| = \frac{1}{2^N}$;

lo cual contradice la elección de x_N en $B(x, \frac{1}{2^N}) \setminus B$. Esta contradicción muestra que (b) es cierta. \square

Verifiquemos que $\tau = \tau_\rho$. Si $U \in \tau$ es no vacío, para toda $x \in U$ podemos fijar un elemento B de la base \mathcal{B} de modo que $x \in B \subseteq U$. Por el inciso (b) de la afirmación 3, existe $\epsilon_x > 0$ tal que $x \in B(x, \epsilon) \subseteq U$. Esto muestra que $U \in \tau_\rho$. Entonces $\tau \subseteq \tau_\rho$. Por otra parte, si $U \in \tau_\rho$ es no vacío, entonces para toda $x \in U$ existe $r_x > 0$ tal que $B(x, r_x) \subseteq U$. Aplicando el inciso (a) de la afirmación 3, para toda $x \in U$ podemos fijar un elemento B_x de la base \mathcal{B} de modo que $x \in B_x \subseteq B(x, r_x)$. Entonces $U = \bigcup_{x \in U} B_x \in \tau$. Por lo tanto, $\tau_\rho \subseteq \tau$. \square

Como todo espacio pseudometrizable T_0 es metrizable, a partir del teorema 5 podemos concluir el teorema de metrización de Urysohn para espacios T_4 .

TEOREMA 6 (Urysohn 4). *Todo espacio normal, T_1 y segundo numerable es metrizable.*

5. EL TEOREMA DE METRIZACIÓN DE URYSOHN PARA ESPACIOS T_3

En la parte final de 4 Urysohn planteó la pregunta de si era posible modificar la hipótesis de normalidad en el teorema 6 a una hipótesis más débil. En 3, Tychonoff resolvió positivamente este planteamiento demostrando el siguiente resultado.

TEOREMA 7 (Tychonoff). *Todo espacio segundo numerable regular es normal.*

Demostración. Supongamos que (X, τ) es un espacio regular segundo numerable. Supongamos también que \mathcal{B} es una base numerable de (X, τ) y que F y G son cualesquiera subconjuntos de X , cerrados, ajenos y no vacíos.

Para cada $x \in F$, por la regularidad de (X, τ) , existen abiertos ajenos A_x, B_x tales que $x \in A_x$ y $G \subseteq B_x$. Debido a que \mathcal{B} es una base para τ , para cada $x \in F$, existe $U_x \in \mathcal{B}$ tal que $x \in U_x \subseteq A_x$. Obsérvese que $\text{cl}(U_x) \cap G = \emptyset$ para cada $x \in F$.

Definamos $\mathcal{F} = \{U_x : x \in F\}$. Como $\mathcal{F} \subseteq \mathcal{B}$, tenemos que \mathcal{F} es numerable. Luego podemos numerar a sus elementos: $\mathcal{F} = \{U_n : n \in \mathbb{N}\}$. Entonces $F \subseteq \bigcup_{n=1}^{\infty} U_n$, y además, $\text{cl}(U_n) \cap G = \emptyset$ para cada $n \in \mathbb{N}$.

De manera totalmente análoga a lo antes hecho, podemos construir una colección $\mathcal{G} = \{V_n : n \in \mathbb{N}\} \subseteq \mathcal{B}$ tal que $G \subseteq \bigcup_{n=1}^{\infty} V_n$, y además, $\text{cl}(V_n) \cap F = \emptyset$ para cada $n \in \mathbb{N}$.

Definamos ahora $C_1 = U_1$, $D_1 = V_1 \setminus \text{cl}(U_1)$, y para cada $n \geq 2$, $C_n = U_n \setminus \bigcup_{i=1}^{n-1} \text{cl}(D_i)$ y $D_n = V_n \setminus \bigcup_{i=1}^n \text{cl}(C_i)$.

Claramente cada conjunto C_n , y cada conjunto D_n , es abierto en X . Además, $C_n \subseteq U_n$ y $D_n \subseteq V_n$ para cada $n \in \mathbb{N}$. Definamos $U = \bigcup_{n=1}^{\infty} C_n$ y $V = \bigcup_{n=1}^{\infty} D_n$. Resulta que U y V son abiertos. Para terminar demostraremos que U y V son conjuntos ajenos y que $F \subseteq U$ y $G \subseteq V$. Efectivamente, supongamos por el contrario que existe $x \in U \cap V$. Entonces existen $n, m \in \mathbb{N}$ tales que $x \in C_n \cap D_m$. Si $n \leq m$ entonces $x \in C_n \cap (V_m \setminus \bigcup_{i=1}^m \text{cl}(C_i)) \subseteq C_n \cap (V_m \setminus \text{cl}(C_n)) \subseteq (C_n \setminus C_n) = \emptyset$, lo cual es imposible. Luego, debe ocurrir que $n > m$. Pero entonces $x \in C_n \cap D_m = (U_n \setminus \bigcup_{i=1}^{n-1} \text{cl}(D_i)) \cap D_m \subseteq (U_n \setminus D_m) \cap D_m = \emptyset$, lo cual es absurdo. Lo anterior muestra que necesariamente $U \cap V = \emptyset$. Por otro lado, si $z \in F$ entonces existe $n \in \mathbb{N}$ tal que $z \in U_n$. Si $n = 1$ entonces $z \in U_1 = C_1 \subseteq U$. Supongamos que $n \geq 2$. Como $z \in F$, $z \notin \text{cl}(V_i)$ para cada $i \in \mathbb{N}$. Debido a que $D_i \subseteq V_i$ para toda $i \in \mathbb{N}$, podemos concluir que $z \notin \bigcup_{i=1}^{n-1} \text{cl}(D_i)$. Luego, $z \in U_n \setminus \bigcup_{i=1}^{n-1} \text{cl}(D_i) = C_n \subseteq U$. Por otra parte, si $z \in G$ es cualquier elemento, entonces existe $m \in \mathbb{N}$ tal que $z \in V_m$. Como $z \in G$, $z \notin \text{cl}(U_i)$ para toda $i \in \mathbb{N}$. Como $C_i \subseteq U_i$ para toda $i \in \mathbb{N}$, tenemos que $z \in V_m \setminus \bigcup_{i=1}^m \text{cl}(C_i) = D_m \subseteq V$. \square

Algunos autores enuncian el teorema 7 para espacios T_3 , es decir, para espacios regulares que además son espacios T_0 4. Se puede notar en la prueba que dimos que

⁴El que un espacio sea regular no implica que éste sea T_0 . Como ejemplo de ello considere al conjunto \mathbb{R} con la topología indiscreta $\tau = \{\emptyset, \mathbb{R}\}$.

no es necesario el uso del axioma T_0 . También es importante mencionar que de esta prueba (la cual es original de Tychonoff) se extrajo la siguiente caracterización de los espacios normales.

PROPOSICIÓN 8. *Las siguientes condiciones son equivalentes para cualquier espacio topológico (X, τ) :*

1. (X, τ) es un espacio normal,
2. para cualquier subconjunto cerrado F y cualquier abierto U con $F \subseteq U$, existe una sucesión $(W_n)_{n \in \mathbb{N}}$ de subconjuntos abiertos de X tal que $F \subseteq \bigcup_{n=1}^{\infty} W_n$ y $\text{cl}(W_n) \subseteq U$ para toda $n \in \mathbb{N}$.

Demostración. (1) \Rightarrow (2). Supongamos que F es cerrado y que U es un abierto tal que $F \subseteq U$. Como X es un espacio normal, podemos encontrar un abierto V de X tal que $F \subseteq V \subseteq \text{cl}(V) \subseteq U$. Definamos ahora $W_n = V$ para toda $n \in \mathbb{N}$. Es claro que la sucesión $(W_n)_{n \in \mathbb{N}}$ tiene las propiedades deseadas.

(2) \Rightarrow (1). Consideremos subconjuntos cerrados ajenos F y G de X . Definamos $U = X \setminus G$. Como $F \subseteq U$ y U es abierto, por hipótesis, existe una sucesión de subconjuntos abiertos $(U_n)_{n \in \mathbb{N}}$ de X tal que $F \subseteq \bigcup_{n=1}^{\infty} U_n$ y $\text{cl}(U_n) \subseteq X \setminus G$ para cada $n \in \mathbb{N}$. De manera análoga, existe una sucesión $(V_n)_{n \in \mathbb{N}}$ de abiertos tales que $G \subseteq \bigcup_{n=1}^{\infty} V_n$ y $\text{cl}(V_n) \subseteq X \setminus F$ para cada $n \in \mathbb{N}$.

Definamos ahora $C_1 = U_1$, $D_1 = V_1 \setminus \text{cl}(U_1)$, y para cada $n \geq 2$,

$$C_n = U_n \setminus \bigcup_{i=1}^{n-1} \text{cl}(D_i) \quad \text{y} \quad D_n = V_n \setminus \bigcup_{i=1}^n \text{cl}(C_i).$$

Claramente cada conjunto C_n , y cada conjunto D_n , es abierto en X . Además, $C_n \subseteq U_n$ y $D_n \subseteq V_n$ para cada $n \in \mathbb{N}$. Definamos ahora $U = \bigcup_{n=1}^{\infty} C_n$ y $V = \bigcup_{n=1}^{\infty} D_n$. Resulta que U y V son abiertos, son conjuntos ajenos y además $F \subseteq U$ y $G \subseteq V$ (vea la prueba del teorema [7](#)). Por lo tanto, X es un espacio normal. \square

Para los fines de esta nota la consecuencia más importante del teorema [7](#) es el siguiente teorema de pseudometrización para espacios regulares T_0 que son segundo numerables.

TEOREMA 9 (de metrización de Urysohn). *Todo espacio segundo numerable, regular y T_0 es metrizable separable.*

Demostración. Supongamos que (X, τ) es un espacio regular, T_0 y segundo numerable. Aplicando el teorema [7](#) podemos concluir que (X, τ) es un espacio normal T_1 . Aplicando ahora el teorema [5](#), tenemos que X es pseudometrizable.

Debido a que todo espacio pseudometrizable T_0 es metrizable, el espacio (X, τ) es metrizable. Éste es separable porque todo espacio segundo-numerable lo es. Esto termina la demostración del teorema. \square

REFERENCIAS

- [1] Casarrubias-Segura, F & Tamariz-Mascarúa, A. *Elementos de topología general*. Aportaciones matemáticas. Textos 37. Nivel medio. Instituto de matemáticas. UNAM. 2019.
- [2] Dávila-Albarrán, J. L. *Lema de Urysohn y sus aplicaciones*. Tesis de Máster. Posgrado de Matemáticas, Máster de Matemática Avanzada. Curso 2012-2013. Universidad de Murcia. España.
- [3] Tychonoff, A. N. *Über einen Metrisationssatz von P. Urysohn*, *Mathematische Annalen*, 95, 994–996. 1925.
- [4] Urysohn, P. *Zum Metrisationproblem*. *Mathematische Annalen*, 94, 309–315, 1925.
(Traducción al español: *Lecturas básicas en Topología General*, Comunicaciones 28, Aportaciones Matemáticas. Sociedad Matemática Mexicana. 2000. Páginas 295–301).
- [5] Urysohn, P. *Über die Mächtigkeit der zusammenhängenden Räume*, *Mathematische Annalen*, 94, 262–295, 1925.

Dirección del autor:

Universidad Nacional Autónoma de México,
Facultad de Ciencias,
Departamento de Matemáticas.
Av. Universidad 3000, Circuito Exterior S/N
Alcaldía Coyoacán, C.P. 04510.
Ciudad Universitaria, Ciudad de México. México.
e-mail: fcasarrubiass@ciencias.unam.mx.



LA FÓRMULA DE DESCOMPOSICIÓN PARA LA VALORACIÓN DE OPCIONES DE COMPRA BAJO EL MODELO DE HESTON.

RAÚL MERINO

RESUMEN. En este artículo se presenta una breve introducción a los derivados financieros. Por una parte, recorremos parte de la historia y características de los derivados financieros. Por otra, se introduce el modelo de Black-Scholes-Merton, pilar básico de la valoración de opciones, y el modelo de Heston, uno de los modelos de volatilidad estocástica más populares. Además, se explica la teoría de la fórmula de descomposición que nos permite aproximar el precio de una opción en el modelo de Heston mediante el uso de la fórmula de Itô.

1. INTRODUCCIÓN

Es fácil pensar que los derivados financieros, y en general, cualquier producto financiero, es una invención de la economía moderna. Aunque más allá de la visión cinéfila de 'Wall Street' o los errores derivados de las crisis financieras; los derivados surgieron tan pronto como las personas fuimos capaces de hacer promesas creíbles. Los primeros derivados financieros sirvieron para garantizar el suministro de productos, facilitar el comercio y asegurar a los agricultores contra la pérdida de cosechas. La primera evidencia de un contrato derivado se encuentra en la Ley 46 del código de Hammurabi, situado entre 1782 y 1750 a.C.

Para entender bien qué es un derivado financiero, hemos de dividirlos en dos categorías diferentes. Por un lado, tenemos la 'fuente primaria', a la que llamaremos 'subyacente', como pueden ser las acciones, bonos, materias primas, divisas, etc. Por otro lado, su contrato 'derivado', es un acuerdo en obtener un derecho en la entrega de un subyacente, o una cantidad monetaria dependiente de su evolución, en un momento futuro determinado.

Los contratos derivados más usuales son los futuros y las opciones. Un futuro es un acuerdo legal para comprar (o vender) un activo a un precio predeterminado en un momento específico en el futuro. Mientras que una opción otorga el derecho, pero no la obligación, de comprar (o vender) un activo a un precio predeterminado en un momento específico en el futuro. El precio acordado en el contrato se le llama 'Strike' y 'Moneyness' a la ratio entre el precio del activo respecto el Strike.

Una de las primeras historias relacionadas con la especulación de derivados se debe a Tales de Mileto. A fin de demostrar que los filósofos podrían hacerse ricos si ese fuese su objetivo, Tales utilizó sus conocimientos para predecir cómo sería la cosecha de aceitunas al verano siguiente. Al ver que sería abundante, decidió arrendar todas las prensas de oliva por adelantado. Como nadie sabía si la cosecha sería buena o mala, Tales compró los derechos a un precio relativamente bajo. Cuando la cosecha resultó ser abundante, la demanda de las prensas fue alta y Tales realquiló las prensas a precios mucho más elevados, obteniendo un beneficio considerable.

Pese a la antigüedad en el uso de contratos derivados, su modelización no comenzó hasta 1900. Louis Bachelier introdujo el primer modelo en su tesis 'Théorie de la spéculation', [5], siendo la piedra angular en la teoría de la valoración de derivados.

En su tesis, Bachelier se dio cuenta que había un equilibrio entre compradores y vendedores.

2010 *Mathematics Subject Classification.* 60H05, 91G20, 91G30.

Palabras clave. Cálculo de opciones, modelos de volatilidad estocástica.

Parece que en el mercado, el conjunto de los especuladores, no puede creer ni en una subida ni en una caída del mercado, ya que, por cada precio cotizado, hay tantos compradores como vendedores.

En particular, se dio cuenta de la necesidad de utilizar martingalas para describir los movimientos de los precios.

La esperanza matemática de los especuladores es nula.

Aunque el método para describir la evolución de los activos es similar a las metodologías actuales, Bachelier desarrolló su modelo bajo hipótesis de equilibrio, mientras que hoy en día se utilizan hipótesis de no arbitraje.

A pesar de las ideas novedosas utilizadas en la tesis de Bachelier, esta permaneció desconocida durante décadas. En 1953, Kendall, [26], sin ser consciente de la tesis de Bachelier, analizó 22 series de precios con el propósito de encontrar un modelo. En ese momento, se dio cuenta de que, entre los intervalos, había cambios aleatorios, descartando un efecto sistemático. Además, fue el primero en observar la dependencia temporal de la varianza empírica. Unos años más tarde, en 1959, Osborne, [34], encontró que los retornos logarítmicos del precio siguen un movimiento browniano.

A mediados de 1950, el estadista Jimmy Savage recuperó el trabajo de Bachelier y lo envió a diferentes amigos. Una de esas cartas llegó a Paul Samuelson, quien estaba trabajando en problemas relacionados con la valoración de opciones y warrants. Paul Samuelson se inspiró en el trabajo de Bachelier y relacionó el precio de las opciones con el uso de martingalas en [37].

Unos años más tarde, en 1973, abrió la primera bolsa de opciones listadas, el CBOE, Chicago Board Options Exchange. Ese mismo año, se publicó el modelo de Black-Scholes-Merton usando hipótesis de no-arbitraje, ver [9] y [30]. Este modelo es la base de gran parte de los desarrollos acontecidos en años posteriores. El modelo de Black-Scholes-Merton, o modelo BSM, proporciona una fórmula analítica que describe con parsimonia el precio de las opciones. La fórmula depende de la Moneyness, el tipo de interés y la volatilidad constante. La volatilidad constante es una de los principales inconvenientes del modelo. Pese a ello, se considera el modelo ‘base’ de mercado. Tanto es así, que el precio de las opciones cotizadas viene dado mediante la volatilidad de la fórmula de BSM. Estas volatilidades se les llama volatilidades implícitas, i.e. es la volatilidad constante que introducida en la fórmula de BSM obtiene el precio de mercado de la opción.

En los años siguientes, aparecieron nuevas versiones del modelo BSM intentando incluir ciertos patrones que se observaban en el mercado. El modelo CEV, Constant Elasticity Variance, fue presentado en un artículo sin publicar en [13], ver también [8]. La principal mejora de este modelo es explicar la relación inversa entre el nivel del activo y la varianza de sus retornos. El siguiente año, Merton, [31], publicó el primer modelo de difusión con saltos. Este modelo extiende el modelo de BSM para procesos estocásticos con trayectorias no continuas, explicando posibles vibraciones ‘anormales’ debido a la llegada de nueva información con un efecto no marginal en el precio del activo. Los dos modelos introducen una asimetría en la estructura de volatilidades implícitas, también conocida como ‘skew’. La asimetría significa que la volatilidad implícita decrece a medida que crece la Moneyness.

El 19 de octubre de 1987, también conocido como Lunes Negro, fue una de las mayores crisis financieras conocidas. Todos los mercados internacionales experimentaron grandes pérdidas, por ejemplo, el Dow Jones perdió en un día un 22.6% y se observó un aumento considerable de la volatilidad. Poco después del colapso, los traders observaron que las opciones cuyo Strike estaba alejado del valor actual del activo cotizaban inusualmente más caras comparadas con Strikes similares al valor del activo. Este fenómeno se llama ‘la sonrisa de volatilidad’.

Una extensión natural del modelo de BSM es considerar que la volatilidad sigue un proceso estocástico. Este tipo de modelos se llaman modelos de volatilidad estocástica y surgieron con la finalidad de explicar mejor las superficies que veían los traders.

Hubieron varios intentos para lograr obtener una solución, uno de los primeros fue [24], seguido de [25], [40] y [38] que intentaron resolver el problema mediante métodos numéricos. Usando un enfoque diferente, [23] obtuvo una aproximación del precio como una expansión de Taylor en el caso en que el activo y la varianza no estuvieran correlacionados. En [39] se propuso un modelo donde la volatilidad depende de un proceso de Ornstein-Uhlenbeck aritmético no correlacionado con el activo. Mediante integración numérica es posible encontrar el precio de las opciones. En 1993 se publicó el modelo de Heston, [22], que se ha convertido en uno de los modelos más populares debido a que proporciona una solución semi-analítica y a sus buenas propiedades estadísticas. Heston propuso un modelo donde el proceso de varianza se modela mediante un proceso CIR, [14], y con una correlación arbitraria entre el activo y la varianza. Para un enfoque más detallado, ver [18] y las referencias que contiene.

Además de observar la ‘sonrisa de volatilidad’, las opciones que están cerca de su vencimiento se negocian con volatilidades implícitas más altas. Los traders, conscientes de la posibilidad de un gran movimiento de mercado, o salto en el precio de los activos, solicitan precios más altos. Los modelos de volatilidad estocástica no son capaces de reproducir toda la superficie de volatilidad cuando se ven volatilidades tan altas en los vencimientos cortos. Para mejorarlos, se introdujeron los modelos estocásticos de volatilidad con saltos. El primer modelo se le atribuye a [6], quien incorporó al modelo de Heston un proceso de saltos en el activo de manera similar a [31]. Este modelo se ajusta mejor a la superficie de mercado. Existen diferentes alternativas utilizando amplitudes de salto diferentes, ver [41] o [27]. También es posible extender este modelo agregando saltos al proceso de varianza, por ejemplo, un modelo introducido por [16]. Sin embargo, según varios estudios empíricos, estos modelos tienden a sobreajustarse a los precios de mercado.

A pesar de la extensa literatura en los modelos de volatilidad estocástica, su uso es complejo. Por un lado, estos modelos necesitan calibrarse, es decir, se deben encontrar qué parámetros minimizan el error entre los precios del modelo y las opciones de mercado. Por otro lado, el cálculo es computacionalmente más intenso. [15], [17] y [36] propusieron un modelo diferente, el modelo de volatilidad local. Definieron una volatilidad instantánea única que es una función determinista del tiempo y el precio del activo consistente con los precios de las opciones de mercado.

Los modelos de volatilidad local son autoconsistentes, libres de arbitraje y se pueden calibrar con precisión para toda la superficie de volatilidad. Sin embargo, como señala [21], el comportamiento dinámico de la sonrisa y las asimetrías producidas por este modelo pueden ser contrarias al movimiento observado en el mercado, obteniendo peores coberturas que utilizando el modelo de BSM. En [21], se introduce el modelo SABR, este modelo se puede clasificar como un modelo de volatilidad local estocástico. Este modelo consiste en modelar el precio del activo con el modelo de CEV con un proceso de volatilidad estocástica exponencial. El principal éxito de este modelo es obtener un modelo capaz de ajustar la ‘sonrisa de volatilidad’ con pocos parámetros y con una aproximación sencilla de calcular de la volatilidad implícita.

La última tendencia en modelización consiste en considerar que la volatilidad sigue un proceso Volterra, en particular, se sustituye el browniano por un browniano fraccionario. Inicialmente, los pioneros de estos modelos, ver [10] y [11], supusieron que la volatilidad tenía memoria, es decir, era un proceso persistente con un parámetro de Hurst entre $H \in (1/2, 1)$. En cambio, hoy en día, se están considerando los modelos antipersistentes, llamados modelos de volatilidad ‘rough’. En estos modelos se observa una consistencia entre la serie temporal de la volatilidad realizada y la volatilidad fraccionaria aproximada. Esto significa que debería ser un modelo más consistente para valores opciones de mercado, ver , por ejemplo, [3], [7] y [19].

2. PRELIMINARES

2.1. Derivados financieros. A menudo, las personas necesitan firmar acuerdos donde se comprometen a intercambiar diferentes activos o cantidades de dinero en momentos de tiempo diferentes. Por ejemplo, cuando alguien necesita comprar una casa, obtiene un préstamo hipotecario en el que recibe una gran suma de dinero a cambio de pagos mensuales futuros. Estos acuerdos son contratos financieros y a cada parte se le denomina contraparte.

Cuando un contrato financiero depende de la evolución de un activo, se le denomina contrato derivado. Esto se debe a que el precio del contrato se ‘deriva’ de la evolución del activo, también llamado subyacente. Nos referimos al precio del subyacente en el momento t como S_t . Por ejemplo, un contrato derivado nos permitiría recibir una cantidad de efectivo hoy a cambio de 50 kg de arroz en el futuro. Este producto puede resultar interesante para agricultores, pudiendo adelantar parte de los ingresos de realizar la cosecha. Los derivados más simples son los futuros.

Definición 1. Un contrato de futuros es un acuerdo entre dos partes, donde una de ellas se compromete a comprar un activo a un precio específico a la otra, en un momento específico en el futuro. El momento en que se entrega el activo se denomina vencimiento y se indica con T . El precio especificado se conoce como K , también llamado precio de entrega. El valor del contrato a vencimiento es $S_T - K$ y su valor en el momento inicial es cero.

Un contrato de futuro es un juego de suma cero, una contraparte gana dinero y la otra lo pierde. Para valorar este contrato, no es necesario un modelo o el uso de probabilidades. Mediante hipótesis de no arbitraje se puede demostrar que un contrato de futuros es justo cuando el precio de entrega es el precio *forward*.

Definición 2. El precio *forward* de un activo con valor actual S_t y vencimiento T es

$$F(t, T) = S_t e^{r(T-t)}$$

donde r es la tasa libre de riesgo.

No todos los contratos derivados pueden cotizarse sin el uso de un modelo. Todos aquellos donde una contraparte tiene la posibilidad de ejercitar o no el contrato, se debe especificar un modelo. Este tipo de productos se llaman opciones.

Definición 3. Una opción de compra europea es un acuerdo legal que otorga al tenedor el derecho, pero no la obligación, de comprar una unidad de un activo subyacente por un precio de ejercicio predeterminado K en la fecha de vencimiento T . Si S_T es el precio del activo subyacente en la fecha vencimiento T , entonces el valor, o pago, de una opción de compra es

$$(S_T - K)_+ = \begin{cases} S_T - K & \text{si } S_T > K, \\ 0 & \text{si } S_T \leq K. \end{cases}$$

Si en el vencimiento, el precio del activo es mayor al precio acordado, el tenedor ejercerá la opción y obtendrá una ganancia. De lo contrario, como el tenedor obtendría una pérdida, este decidiría no ejercer la opción.

Definición 4. Una opción de venta europea es un acuerdo legal que otorga al tenedor el derecho, pero no la obligación, de vender una unidad de un activo subyacente por un precio de ejercicio predeterminado K en la fecha de vencimiento T . Si S_T es el precio del activo subyacente en la fecha vencimiento T , entonces el valor, o pago, de una opción de venta es

$$(K - S_T)_+ = \begin{cases} K - S_T & \text{si } S_T < K, \\ 0 & \text{si } S_T \geq K. \end{cases}$$

Al contrario que los futuros, para adquirir una opción se debe pagar una prima. Este tipo de opciones se llaman *plain vanilla* y son las más sencillas. En inglés se utiliza el adjetivo *plain vanilla* para referirse a la versión más sencilla o común de algo, en analogía al sabor del helado de vainilla que se hizo ampliamente disponible y económico con el desarrollo de la vainilla artificial. Este tipo de opciones pueden ser *européas* o *americanas*. Las opciones *européas* sólo se pueden ejecutar en el vencimiento, mientras que las *americanas* se pueden ejecutar en cualquier momento anterior al vencimiento. Existen gran variedad de tipos de opciones con funciones de pago diferentes, a estas opciones se les conoce como opciones *exóticas* y se suelen contratar de manera privada entre dos contrapartes. Este tipo de contratación se llama OTC o *Over-The-Counter*.

Las opciones se clasifican dependiendo de cómo sea el precio de activo respecto de ejercicio:

Definición 5. Una opción está

- *In the money* o *ITM* si se ejecutará hoy, se obtendría un valor positivo. En el caso de una opción de compra si $S_t > K$.
- *At the money* o *ATM* si se ejecutará hoy, el valor del activo S_t tiene el mismo nivel que el precio de ejercicio K . Es decir, si $S_t = K$.
- *Out the money* o *OTM* si se ejecutará hoy, el valor de la opción sería cero. En el caso de una opción de compra si $S_t < K$.

2.2. Breve introducción matemática. Para ser capaces de modelar la evolución de activos, es necesario introducir algunos elementos de cálculo estocástico. La teoría es extensa y, en este caso, introduciremos brevemente qué es un movimiento browniano, un proceso de Itô, la fórmula de Itô y la fórmula de Feynman-Kac.

Definición 6 (Movimiento browniano/Proceso de Wiener). Un movimiento browniano, o proceso de Wiener, es un proceso estocástico $(X_t)_{t \geq 0}$ tal que:

1. Las trayectorias $t \mapsto X_t$ son continuas con probabilidad 1.
2. Para una secuencia finita de tiempo $t_0 < t_1 < \dots < t_n$, los incrementos

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$$

son independientes. Es decir, el comportamiento de un incremento no afecta al comportamiento de otros incrementos.

3. Tiene incrementos estacionarios. Es equivalente a decir que la distribución de los incrementos sólo depende de la longitud del incremento y no sobre el tiempo. Se puede escribir que para cualquier $t_1 < t_2$, $X_{t_2} - X_{t_1} \sim X_{t_2-t_1} - X_0$.

La definición de movimiento browniano induce su propia distribución.

TEOREMA 7. Si $(X_t)_{t \geq 0}$ es un movimiento browniano, entonces $X_t - X_0$ es una variable aleatoria cuyos valores siguen una distribución normal con media rt y varianza $\sigma^2 t$, donde r y σ son constantes positivas.

Demostración. Se puede ver la demostración en [12], Teorema 2.2.1. □

Definición 8. Un movimiento browniano estándar es un movimiento browniano tal que $X_0 = 0$ c.s., $r = 0$ y $\sigma^2 = 1$. Nos referimos al movimiento browniano estándar como $(W_t)_{t \geq 0}$.

Así como sabemos integrar respecto el tiempo, también estamos interesados en integrar respecto al movimiento browniano, es decir, ser capaces de resolver integrales estocásticas. Desafortunadamente, la definición de este tipo de integrales falla porque, aunque las trayectorias del movimiento browniano son continuas, no son de variación acotada ni diferenciables. Para poder definir este tipo de integrales, necesitamos que el integrando sea de cuadrado integrable y que sea adaptado a la información que genera el movimiento browniano. Este tipo de procesos se llaman adaptados.

Cuando un proceso se escribe como la suma de una integral respecto el tiempo y otra integral respecto el movimiento browniano se llama procesos de Itô y son la base del cálculo estocástico.

Definición 9 (Proceso de Itô). Un proceso de Itô $(X_t)_{0 \leq t \leq T}$ es un proceso del tipo:

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s,$$

donde para todo $t \leq T$, tenemos que

1. X_0 es \mathcal{F}_0 -medible.
2. (K_t) y (H_t) son procesos \mathbb{F} -adaptados.
3. $\int_0^T |K_s| ds < \infty$ \mathbb{P} c.s.
4. (Condición de cuadrado integrable) $\int_0^T |H_s|^2 ds < \infty$ \mathbb{P} c.s.

Este tipo de procesos tienen la particularidad que tiene variación cuadrática diferente de cero. Es decir, si en análisis se considera que $(dt)^2 = 0$, en este caso $(dW_t)^2 = dt$. Para poder trabajar con estos procesos, se necesita trabajar con una fórmula equivalente a la regla de la cadena, pero que funcione con los procesos de Itô. Esta fórmula es la fórmula de Itô y es uno de los teoremas fundamentales del cálculo estocástico.

TEOREMA 10 (Fórmula de Itô). *Dado un proceso de Itô $(X_t)_{0 \leq t \leq T}$,*

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s,$$

y $f(t, x) \in \mathcal{C}^{1,2}$, entonces

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \partial_s f(s, X_s) ds + \int_0^t \partial_x f(s, X_s) dX_s \\ &+ \frac{1}{2} \int_0^t \partial_x^2 f(s, X_s) d[X, X]_s \end{aligned}$$

donde la variación cuadrática se define como

$$[X, X]_t := \int_0^t H_s^2 ds.$$

Demostración. Se puede ver la demostración en [35], Capítulo II, Teorema 32. \square

La fórmula de Feynman-Kac nos permite expresar el precio de una opción como una solución de una ecuación diferencial en derivadas parciales. Es una consecuencia de la conexión entre las ecuaciones diferenciales estocásticas y ciertas ecuaciones diferenciales parciales parabólicas.

TEOREMA 11 (La Fórmula de Feynman-Kac). *Sea $f(t, x)$ una función acotada, sea $\varphi(x)$ una función dos veces diferenciable, sea $V(t, x)$ una función acotada inferiormente para toda $x \in \mathbb{R}$. Dadas unas funciones $\mu(t, x)$ y $\sigma^2(t, x)$ Lipschitz continua y $u(t, x) \in \mathcal{C}^{1,2}$. Consideremos la ecuación diferencial en derivadas parciales*

$$\partial_t u(t, x) + \mu(t, x) \partial_x u(t, x) + \frac{1}{2} \sigma^2(t, x) \partial_x^2 u(t, x) - V(t, x) u(t, x) + f(t, x) = 0,$$

definida para todo $x \in \mathbb{R}$ y $t \in [0, T]$, sujeta a la condición terminal

$$u(x, T) = \varphi(x).$$

La solución de esta ecuación diferencial en derivadas parciales es la fórmula de Feynman-Kac y viene dada por la esperanza condicionada

$$u(t, x) = \mathbb{E} \left[\int_t^T e^{-\int_t^r V(\tau, X_\tau) d\tau} f(r, X_r) dr + e^{-\int_t^T V(\tau, X_\tau) d\tau} \varphi(X_T) | X_t = x \right]$$

sobre la medida de probabilidad subyacente tal que X_t es un proceso de Itô definido por la ecuación

$$dX_t = \mu(t, X)dt + \sigma(t, X)dW_t,$$

donde W_t es un proceso de Wiener y la condición inicial de X_t es $X_t = x$.

Demostración. Se puede ver la demostración en [33], Teorema 8.2.1. □

3. EL MODELO DE BLACK-SCHOLES-MERTON

Consideremos que el vencimiento de una opción T está fijado y $t \in [0, T]$. En un modelo financiero sencillo, existen dos tipos de activos: una cuenta bancaria B_t que representa a un activo libre de riesgo y un activo con riesgo S_t . El precio de B_t viene dado por

$$(1) \quad B_t = e^{rt}$$

donde $r \geq 0$ es la tasa libre de riesgo. Además, es la solución de una Ecuación Diferencial Ordinaria:

$$(2) \quad dB_t = rB_t dt.$$

El precio de un activo con riesgo S_t viene dado por

$$(3) \quad dS_t = \mu S_t dt + \sigma S_t dW_t$$

donde μ es la tasa de crecimiento del activo, $\sigma > 0$ es la volatilidad y W_t es un movimiento browniano. Bajo argumentos de no arbitraje se puede ver que $\mu = r$.

Usando la fórmula de Itô, la ecuación diferencial estocástica (3) tiene la solución

$$(4) \quad S_t = S_0 e^{(r - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

donde $S_0 > 0$ es el valor actual del activo.

Este modelo tiene una solución analítica en el caso de opciones europeas.

TEOREMA 12. *El precio de una opción de compra europea es*

$$C_{BS}(t, T, S_t, K, r, \sigma) = S_t \Phi(d_+) - K e^{-r\tau} \Phi(d_-)$$

donde S_t es el precio actual del subyacente, σ es la volatilidad constante, K es el precio de ejercicio, $\tau = T - t$ es el tiempo hasta vencimiento, r es el tipo de interés libre de riesgo y Φ es la distribución acumulativa de una normal estándar. Los símbolos d_+ y d_- se refieren a las funciones siguientes

$$d_{\pm} = \frac{\ln(S_t/K) + (r \pm \frac{\sigma^2}{2})(T - t)}{\sigma \sqrt{T - t}}$$

Demostración. Queremos calcular cuál es el precio de una opción,

$$C_{BS}(t, T, S_t, K, r, \sigma) = e^{-r(T-t)} \mathbb{E} [(S_T - K)_+].$$

El precio del activo subyacente en la fecha vencimiento, S_T , viene dado por

$$S_T = S_t e^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(W_T - W_t)}.$$

Utilizando las propiedades del movimiento browniano, se puede escribir $W_T - W_t = \sqrt{T - t}Z$ donde Z es una distribución de probabilidad normal, es decir, $Z \sim \mathcal{N}(0, 1)$. Podemos reescribir el precio de una opción cómo

$$\begin{aligned} C_{BS}(t, T, S_t, K, r, \sigma) &= e^{-r(T-t)} \mathbb{E} \left[\left(S_t e^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma \sqrt{T-t}Z} - K \right)_+ \right] \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} \left(S_t e^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma \sqrt{T-t}z} - K \right)_+ n(z) dz \end{aligned}$$

donde n es la función de densidad de la normal. La integral es diferente de cero cuando $z \geq -d_-$, por lo tanto, se puede cambiar el dominio de la integral, obteniendo

$$\begin{aligned}
 C_{BS}(t, T, S, K, r, \sigma) &= e^{-r(T-t)} \int_{-d_-}^{\infty} \mathbb{E} \left[\left(S_t e^{(r-\frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}z} - K \right)_+ n(z) dz \right] \\
 &= S_t e^{-\frac{1}{2}\sigma^2(T-t)} \int_{-d_-}^{\infty} e^{\sigma\sqrt{T-t}z} n(z) dz - K e^{-r(T-t)} \int_{-d_-}^{\infty} n(z) dz \\
 &= \frac{S_t}{\sqrt{2\pi}} \int_{-d_-}^{\infty} e^{-\frac{1}{2}(z-\sigma(T-t))^2} dz - K e^{-r(T-t)} \Phi(d_-) \\
 &= \frac{S_t}{\sqrt{2\pi}} \int_{-d_- - \sigma\sqrt{T-t}}^{\infty} e^{-\frac{1}{2}u^2} dz - K e^{-r(T-t)} \Phi(d_-) \\
 &= S_t \Phi(d_+) - K e^{-r\tau} \Phi(d_-).
 \end{aligned}$$

□

Definición 13. Dada una opción europea cotizada en el mercado con precio C , strike K y vencimiento T , la volatilidad implícita $IV = \sigma(K, T)$ es la volatilidad que utilizamos en la fórmula de BSM para obtener el precio de mercado de la opción. Es decir,

$$C_{\text{Mercado}} = C_{BS}(0, T, S_0, K, r, \sigma(K, T))$$

donde S_0 es el precio del activo en el momento de negociación.

La solución para la volatilidad implícita es única debido a que la fórmula de BSM es monótonamente creciente en σ con límites superiores o inferiores dependiendo de si es una opción de compra o de venta. Por el teorema de la función inversa, para cada precio hay una volatilidad implícita.

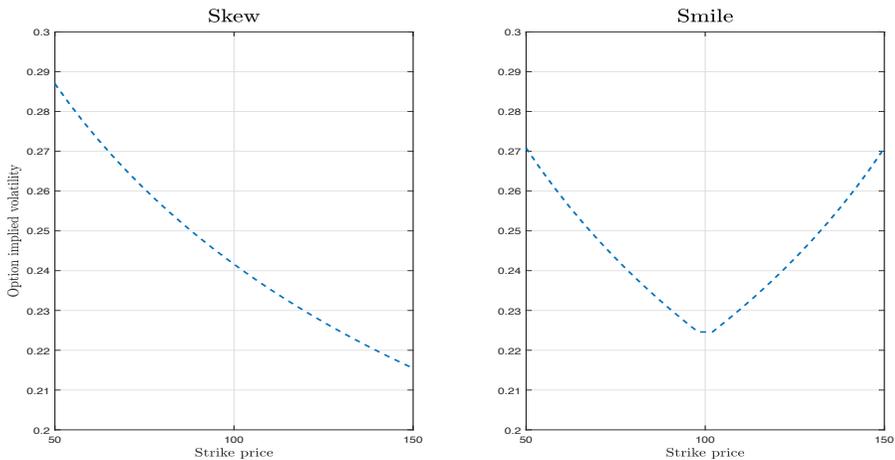


FIGURA 1. Ejemplo de Skew y Smile.

4. EL MODELO DE HESTON

El modelo de Heston, [22], es un modelo de volatilidad estocástica donde la varianza instantánea sigue un proceso de reversión a la media. La dinámica del modelo para el subyacente tiene en cuenta la asimetría y el exceso de curtosis observada en los retornos de los activos financieros además de ajustar los precios de mercado.

El proceso que modela el precio, S , sigue el siguiente sistema de ecuaciones diferenciales estocásticas

$$\begin{aligned}
 dS_t &= rS_t dt + \sigma_t S_t \left(\rho dW_t + \sqrt{1 - \rho^2} d\tilde{W}_t \right), \\
 d\sigma_t^2 &= \kappa (\theta - \sigma_t^2) dt + \nu \sqrt{\sigma_t^2} dW_t.
 \end{aligned}$$

El proceso σ^2 modela la varianza del precio, $\theta > 0$ es el nivel medio de la varianza a largo plazo, $\kappa > 0$ es la velocidad a la que la varianza vuelve a la media, $\nu > 0$ es la volatilidad de la varianza y r es la tasa de interés libre de riesgo. Las condiciones iniciales para el proceso de volatilidad σ y el proceso de precio S son $\sigma_0 > 0$ y $s_0 > 0$, respectivamente. Supondremos que se cumple la condición de Feller $2\kappa\theta \geq \nu^2$ que evita que la varianza sea cero.

Heston expresó el precio de una opción de compra de manera similar a la fórmula de BSM, en su caso

$$C_{Heston} = SP_1 - Ke^{-r(T-t)}P_2$$

donde P_1 es la derivada respecto el valor del subyacente y P_2 es la probabilidad condicionada de que el activo sea más grande que el precio de ejercicio K en la fecha de vencimiento. Las probabilidades P_1 y P_2 dependen de las funciones características φ_1 y φ_2 definidas mediante la inversa de la transformada de Fourier

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-iu \ln K} \varphi_j(S_0, \sigma_0^2, t, T, u)}{iu} du \right] \quad j = 1, 2.$$

Heston asume que las funciones características φ_1 y φ_2 tiene la forma funcional

$$\varphi_j(x, v, t, T, u) = \exp [C(\tau, \gamma) + D(\tau, \gamma)v + i\gamma x]$$

donde

$$\begin{aligned} C(\tau, \gamma) &= r\gamma i\tau + \frac{\kappa\theta}{\nu^2} \left\{ (b_j - \rho\nu\gamma i + d)\tau - 2 \ln \left[\frac{1 - ge^{d\tau}}{1 - g} \right] \right\}, \\ D(\tau, \gamma) &= \frac{b_j - \rho\nu\gamma i + d}{\nu^2} \left[\frac{1 - e^{d\tau}}{1 - ge^{d\tau}} \right] \end{aligned}$$

con

$$\begin{aligned} g &= \frac{b_j - \rho\nu\gamma i + d}{b_j - \rho\nu\gamma i - d}, & d &= \sqrt{(\rho\nu\gamma i - b_j)^2 - \nu^2(2u_j\gamma i - \gamma^2)} \\ u_1 &= 0,5, & u_2 &= -0,5, & a &= \kappa\theta, & b_1 &= \kappa + \lambda - \rho\nu, & b_2 &= \kappa + \lambda \end{aligned}$$

Existen diversas mejoras a la fórmula propuesta por Heston, como puede ser *the little Heston trap* o la fórmula de Lewis. Se puede ver un resumen en [32].

5. FÓRMULA DE DESCOMPOSICIÓN PARA EL MODELO DE HESTON.

Existen diversos métodos para encontrar el valor de una opción para el modelo de Heston: desde fórmulas semi-analíticas hasta diferentes esquemas Monte Carlo. Además de estos métodos, se puede construir el precio de una opción a partir de una adaptación de la fórmula de BSM y la suma de términos extras que explican la componente estocástica de la volatilidad. Este método inspirado en [23] fue utilizado por primera vez en [1] donde se utilizan técnicas de Malliavin, más tarde en [2], [4] y [20] se usó utilizando el cálculo de Itô.

Consideramos \mathcal{F}_t a la información disponible en el momento t , definimos la esperanza condicionada respecto la filtración natural $\{\mathcal{F}_t, t \geq 0\}$ de S como $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot | \mathcal{F}_t]$. Esta esperanza condicionada es el valor esperado en relación a la información disponible en el momento t . La esperanza condicionada de $\mathbb{E}_t[X]$ es la proyección de X al espacios de los procesos adaptados, es decir, a los procesos definidos en L^2 .

En el caso de un modelo de volatilidad no correlacionado, la fórmula para una opción de compra europea viene dado por

$$(5) \quad V_t = E_t[BS(t, S_t, \bar{\sigma}_t)],$$

donde $\bar{\sigma}^2(t)$ es la varianza media futura definida por

$$(6) \quad \bar{\sigma}_t^2 := \frac{1}{T-t} \int_t^T \sigma_s^2 ds.$$

La igualdad (5) se llama la fórmula de Hull-White. Existen diferentes métodos para extender la fórmula para modelos correlacionados. La idea principal de [2] se basa en expandir el precio de una opción respecto a la proyección adaptada de la varianza media futura, es decir,

$$(7) \quad v_t^2 := \mathbb{E}_t(\bar{\sigma}_t^2) = \frac{1}{T-t} \int_t^T \mathbb{E}_t[\sigma_s^2] ds.$$

Definimos

$$(8) \quad M_t = \int_0^T \mathbb{E}_t[\sigma_s^2] ds.$$

Se puede reescribir la proyección adaptada de la varianza media futura como

$$v_t^2 := \mathbb{E}_t(\bar{\sigma}_t^2) = \frac{1}{T-t} \left[M_t - \int_0^t \sigma_s^2 ds \right].$$

Entonces, se obtiene que

$$\begin{aligned} dv_t^2 &= \frac{dt}{(T-t)^2} \left[M_t - \int_0^t \sigma_s^2 ds \right] + \frac{1}{T-t} [dM_t - \sigma_t^2 dt] \\ &= \frac{1}{T-t} [dM_t + (v_t^2 - \sigma_t^2) dt]. \end{aligned}$$

Sin pérdida de generalidad, podemos reescribir el modelo de precios usando el logaritmo del precio $X_t = \log(S_t)$ obteniendo

$$(9) \quad dX_t = \left(r - \frac{1}{2}\sigma_t^2 \right) dt + \sigma_t \left(\rho dW_t + \sqrt{1-\rho^2} d\tilde{W}_t \right).$$

Observación 1. A fin de generalizar la fórmula de descomposición, se asume que el vencimiento, T , el strike, z , y el tipo de interés, r , están fijados. Consideramos la función de pagos $A(t, T, x, z, r, y) = A(t, x, y)$. La función $A(t, x, y)$ pertenece al espacio $\mathcal{C}^{1,2,2}((0, T) \times (0, \infty) \times (0, \infty))$ donde A es una vez diferenciable respecto a t y dos veces diferenciable respecto a x e y . Se asume que las derivadas son continuas.

A continuación, demostraremos cómo podemos descomponer el precio de una opción a partir de su valor terminal.

TEOREMA 14. Sea X_t el proceso asociado al logaritmo del precio según (9), $\{B_t, t \in [0, T]\}$ una semimartingala continua respecto a la filtración \mathcal{F}_t^W y sea el funcional $A(t, x, y)$ según la observación [1]. Para cada $t \in [0, T]$, tenemos la fórmula

siguiente

$$\begin{aligned}
\mathbb{E}_t \left[e^{-r(T-t)} A(T, X_T, v_T^2) B_T \right] &= A(t, X_t, v_t^2) B_t \\
&+ \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \partial_y A(u, X_u, v_u^2) B_u \frac{1}{T-u} (v_u^2 - \sigma_u^2) du \right] \\
&+ \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} A(u, X_u, v_u^2) dB_u \right] \\
&+ \frac{1}{2} \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} (\partial_x^2 - \partial_x) A(u, X_u, v_u^2) B_u (\sigma_u^2 - v_u^2) du \right] \\
&+ \frac{1}{2} \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \partial_y^2 A(u, X_u, v_u^2) B_u \frac{1}{(T-u)^2} d[M, M]_u \right] \\
&+ \rho \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \partial_{x,y}^2 A(u, X_u, v_u^2) B_u \frac{\sigma_u}{T-u} d[W, M]_u \right] \\
&+ \rho \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \partial_x A(u, X_u, v_u^2) \sigma_u d[W, B]_u \right] \\
&+ \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \partial_y A(u, X_u, v_u^2) \frac{1}{T-u} d[M, B]_u \right].
\end{aligned}$$

Demostración. Si aplicamos la fórmula de Itô al proceso $e^{-rt} A(t, S_t, v_t^2) B_t$, se obtiene

$$\begin{aligned}
e^{-rT} A(T, X_T, v_T^2) B_T &= e^{-rt} A_{BS}(t, X_t, v_t^2) B_t \\
&- r \int_t^T e^{-ru} A(u, X_u, v_u^2) B_u du \\
&+ \int_t^T e^{-ru} \partial_u A(u, X_u, v_u^2) B_u du \\
&+ \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u dX_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) B_u dv_u^2 \\
&+ \int_t^T e^{-ru} A(u, X_u, v_u^2) dB_u \\
&+ \frac{1}{2} \int_t^T e^{-ru} \partial_x^2 A(u, X_u, v_u^2) B_u d[X, X]_u \\
&+ \frac{1}{2} \int_t^T e^{-ru} \partial_y^2 A(u, X_u, v_u^2) B_u d[v^2, v^2]_u \\
&+ \int_t^T e^{-ru} \partial_{x,y}^2 A(u, X_u, v_u^2) B_u d[X, v^2]_u \\
&+ \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) d[X, B]_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) d[v^2, B]_u.
\end{aligned}$$

Desarrollando, obtenemos

$$\begin{aligned}
e^{-rT} A(T, X_T, v_T^2) B_T &= e^{-rt} A(t, X_t, v_t^2) B_t \\
&- r \int_t^T e^{-ru} A(u, X_u, v_u^2) B_u du \\
&+ \int_t^T e^{-ru} \partial_u A(u, X_u, v_u^2) B_u du \\
&+ \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u \left(r - \frac{1}{2} \sigma_u^2 \right) du \\
&+ \rho \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u \sigma_u dW_u \\
&+ \sqrt{1 - \rho^2} \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u \sigma_u d\widetilde{W}_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) \frac{B_u}{T - u} dM_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) B_u \frac{v_u^2 - \sigma_u^2}{T - u} du \\
&+ \int_t^T e^{-ru} A(u, X_u, v_u^2) dB_u \\
&+ \frac{1}{2} \int_t^T e^{-ru} \partial_x^2 A(u, X_u, v_u^2) B_u \sigma_u^2 du \\
&+ \frac{1}{2} \int_t^T e^{-ru} \partial_y^2 A(u, X_u, v_u^2) \frac{B_u}{(T - u)^2} d[M, M]_u \\
&+ \rho \int_t^T e^{-ru} \partial_{x,y}^2 A(u, X_u, v_u^2) \frac{B_u \sigma_u}{T - u} d[W, M]_u \\
&+ \sqrt{1 - \rho^2} \int_t^T e^{-ru} \partial_{x,y}^2 A(u, X_u, v_u^2) \frac{B_u \sigma_u}{T - u} d[\widetilde{W}, M]_u \\
&+ \rho \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) \sigma_u d[W, B]_u \\
&+ \sqrt{1 - \rho^2} \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) \sigma_u d[\widetilde{W}, B]_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) \frac{1}{T - u} d[M, B]_u.
\end{aligned}$$

Añadiendo y quitando el termino

$$\frac{1}{2} \int_t^T e^{-ru} \partial_x^2 A(u, X_u, v_u^2) B_u v_u^2 du$$

tenemos que

$$\begin{aligned}
e^{-rT} A(T, X_T, v_T^2) B_T &= e^{-rt} A(t, X_t, v_t^2) B_t \\
&- \boxed{r \int_t^T e^{-ru} A(u, X_u, v_u^2) B_u du} \\
&+ \boxed{\int_t^T e^{-ru} \partial_u A(u, X_u, v_u^2) B_u du} \\
&+ \boxed{\int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u \left(r - \frac{1}{2} \sigma_u^2 \right) du} \\
&+ \rho \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u \sigma_u dW_u \\
&+ \sqrt{1 - \rho^2} \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) B_u \sigma_u d\widetilde{W}_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) \frac{B_u}{T - u} dM_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) B_u \frac{v_u^2 - \sigma_u^2}{T - u} du \\
&+ \int_t^T e^{-ru} A(u, X_u, v_u^2) dB_u \\
&+ \frac{1}{2} \int_t^T e^{-ru} \partial_x^2 A(u, X_u, v_u^2) B_u \sigma_u^2 du \\
&+ \frac{1}{2} \int_t^T e^{-ru} \partial_y^2 A(u, X_u, v_u^2) \frac{B_u}{(T - u)^2} d[M, M]_u \\
&+ \rho \int_t^T e^{-ru} \partial_{x,y}^2 A(u, X_u, v_u^2) B_u \frac{\sigma_u}{T - u} d[W, M]_u \\
&+ \rho \int_t^T e^{-ru} \partial_x A(u, X_u, v_u^2) \sigma_u d[W, B]_u \\
&+ \int_t^T e^{-ru} \partial_y A(u, X_u, v_u^2) \frac{1}{T - u} d[M, B]_u \\
&+ \boxed{\frac{1}{2} \int_t^T e^{-ru} \partial_x^2 A(u, X_u, v_u^2) B_u v_u^2 du} \\
&- \frac{1}{2} \int_t^T e^{-ru} \partial_x^2 A(u, X_u, v_u^2) B_u v_u^2 du.
\end{aligned}$$

Si agrupamos los términos que están en recuadros, tenemos la fórmula de Feynman-Kac asociada al proceso (9), estos términos se anulan. Multiplicando por e^{-rt} y usando esperanzas condicionadas la demostración finaliza. \square

A fin de abreviar la formulación, definimos los operadores siguientes respecto al logaritmo del precio:

- $\Lambda := \partial_x$,
- $\Gamma := (\partial_x^2 - \partial_x)$,
- $\Gamma^2 = \Gamma \circ \Gamma$.

El siguiente corolario nos servirá de base para descomponer en el precio de una opción, así como también obtener una aproximación.

COROLARIO 15. Sea la función A y el proceso B como en el Teorema [14](#). Supongamos que la función A cumple

$$(10) \quad \partial_y \bar{A}(t, x, y) = \frac{(T-t)}{2} (\partial_x^2 - \partial_x) A(t, x, y).$$

Consideremos que $A_t := A(t, X_t, v_t^2) \forall t \in [0, T]$. Para todo $t \in [0, T]$, tenemos que

$$\begin{aligned} e^{-r(T-t)} E_t [A_T B_T] &= A_t B_t \\ &+ \frac{\rho}{2} E_t \left[\int_t^T e^{-r(u-t)} \Lambda \Gamma A_u B_u \sigma_u d[W, M]_u \right] \\ &+ \frac{1}{8} E_t \left[\int_t^T e^{-r(u-t)} \Gamma^2 A_u B_u d[M, M]_u \right] \\ &+ \rho E_t \left[\int_t^T e^{-r(u-t)} \Lambda A_u \sigma_u d[W, B]_u \right] \\ &+ \frac{1}{2} E_t \left[\int_t^T e^{-r(u-t)} \Gamma A_u d[M, B]_u \right] \\ &+ E_t \left[\int_t^T e^{-r(u-t)} A_u dB_u \right]. \end{aligned}$$

Demostración. Substituyendo [\(10\)](#) en el Teorema [14](#) y usando las definiciones de Λ y Γ se completa la prueba. \square

Ahora ya tenemos todos los elementos para poder aplicar la fórmula de descomposición a una opción de compra (o de venta). Realicemos el cambio de variable

$$C_{\overline{BS}}(t, x, y) := C_{BS}(t, e^x, \sqrt{y}).$$

Consideremos que $C_{SV}(t, x, y)$ es el precio de una opción de compra respecto un modelo de volatilidad estocástica. Como hemos hecho anteriormente, podemos hacer un cambio de variables

$$C_{\overline{SV}}(t, x, y) := C_{SV}(t, e^x, \sqrt{y}).$$

En el vencimiento, todas las fórmulas valen lo mismo, el valor del subyacente es conocido, es decir:

$$V_T = C_{SV}(T, S_T, \sigma_T) = C_{\overline{SV}}(T, X_T, v_T^2) = C_{\overline{BS}}(T, X_T, v_T^2).$$

Si aplicamos el Corolario [15](#) con $A = C_{\overline{BS}}$ y $B \equiv 1$, se puede descomponer el precio de una opción de compra en un modelo de volatilidad estocástica, es decir $C_{SV}(t, S_t, \sigma)$, con la fórmula:

$$\begin{aligned} V_t &= C_{BS}(t, S_t, v_t) \\ &+ \frac{1}{8} \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \Gamma^2 C_{\overline{BS}}(u, X_u, v_u^2) d[M, M]_u \right] \\ &+ \frac{\rho}{2} \mathbb{E}_t \left[\int_t^T e^{-r(u-t)} \Lambda \Gamma C_{\overline{BS}}(u, X_u, v_u^2) \sigma_u d[W, M]_u \right] \\ &= C_{BS}(t, S_t, v_t) + (I) + (II). \end{aligned}$$

Si miramos la demostración con cierta calma, se puede ver que no hemos utilizado casi ninguna hipótesis sobre la volatilidad, exceptuando que sea positiva y que la solución de la ecuación diferencial estocástica exista y sea única. Por lo tanto, esta descomposición funciona para cualquier modelo de volatilidad estocástica. Desgraciadamente es una solución complicada de calcular numéricamente.

Si aplicamos el Corolario 10 a los términos (I) y (II), se puede conseguir una aproximación del modelo de volatilidad que es equivalente a ‘congelar’ la derivada en el momento inicial. En este caso, obtendríamos

$$\begin{aligned}
 V_t &= C_{BS}(t, S_t, v_t) \\
 &+ \frac{1}{8} \Gamma^2 C_{BS}(t, X_t, v_t^2) \mathbb{E}_t \left[\int_t^T d[M, M]_u \right] \\
 &+ \frac{\rho}{2} \Lambda \Gamma C_{BS}(t, X_t, v_t^2) \mathbb{E}_t \left[\int_t^T \sigma_u d[W, M]_u \right] \\
 &+ \epsilon_t.
 \end{aligned}$$

Al aplicar el corolario, aparecen nuevos términos que son descartados y considerados como el error ϵ_t de la aproximación. Este error se puede acotar en función del modelo de volatilidad estocástica seleccionado. En 4, se obtiene que el error en el modelo de Heston se puede acotar por

$$|\epsilon_t| \leq \left(\nu^2 (|\rho| + \nu)^2 \right) \left(\frac{1}{r} \wedge (T - t) \right) \Pi(\kappa, \theta),$$

donde $\Pi(\kappa, \theta)$ es una constante positiva que depende de κ y θ .

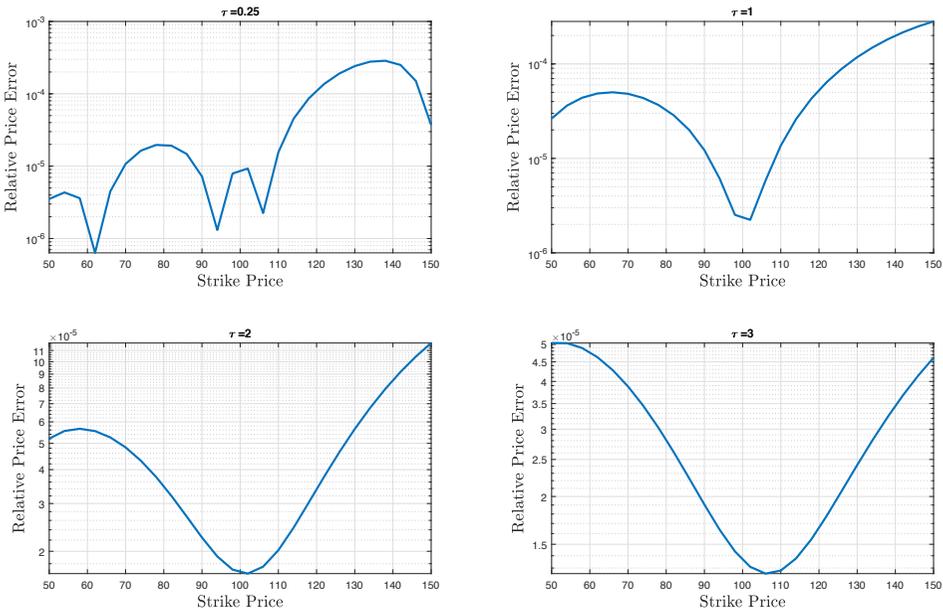


FIGURA 2. Error de la fórmula de descomposición para diferentes vencimientos.

Es posible mejorar la aproximación, sólo es necesario aplicar el Corolario 15 a cada nuevo término. Se pueden seleccionar qué términos aproximar para conseguir el orden deseado. En 20 se proponen nuevas fórmulas de aproximación, así como una comparativa con otros métodos existentes en la literatura.

Esta técnica nos permite expresar el precio de la opción en un modelo de volatilidad estocástica como la fórmula de BSM, respecto la proyección de la volatilidad futura, y una corrección que depende de sus derivadas. Además, es bastante genérica y se puede utilizar para otros modelos. Por ejemplo, para el modelo de Bates como en 28 o para un modelo de volatilidad local como el CEV, ver 29.

6. CONCLUSIONES

Sin lugar a duda, el siglo XX ha sido el siglo de oro de los derivados financieros. En este artículo, se explica una breve historia de los derivados financieros con las referencias clásicas de la literatura así como las nuevas tendencias. Además de introducir las definiciones básicas de las opciones de compra, se explica el modelo de Black-Scholes-Merton que es uno de los pilares básicos de la valoración de activos, siendo la base de modelos más sofisticados. Se introduce también el modelo de Heston, uno de los modelos de volatilidad estocástica más populares con su solución original. Adicionalmente, se presenta la teoría de la fórmula de descomposición, la cual nos permite aproximar el precio de una opción mediante la fórmula Black-Scholes-Merton, respecto la proyección de la volatilidad futura, y unos términos adicionales que explican la componente de la volatilidad estocástica. Se observa la calidad numérica del ajuste en varios casos.

AGRADECIMIENTOS. El autor expresa su gratitud a Jorge A. León y al árbitro anónimo por sus comentarios y sugerencias a fin de mejorar la calidad del artículo.

REFERENCIAS

- [1] Alòs E., *A generalization of the Hull and White formula with applications to option pricing approximation*, Finance and Stochastics, (2006).
- [2] Alòs E., *A decomposition formula for option prices in the Heston model and applications to option pricing approximation*, Finance and Stochastics, (2012).
- [3] Alòs E., León J. A., Vives J., *On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility*, Finance and Stochastics, (2007).
- [4] Alòs E., de Santiago R., Vives J., *Calibration of stochastic volatility models via second-order approximation: The Heston case*, International Journal of Theoretical and Applied Finance, (2015).
- [5] Bachelier L., *Théorie de la spéculation*, PhD thesis, Paris: Gauthier-Villars, 1900.
- [6] Bates D. S., *Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options*, Review of Financial Studies, (1996).
- [7] Bayer C., Friz P., Gatheral J., *Pricing under rough volatility*, Quantitative Finance, (2016).
- [8] Beckers S., *The constant elasticity of variance model and its implications for option pricing*, The Journal of Finance, (1980).
- [9] Black F. S., Scholes M. S., *The pricing of options and corporate liabilities*, Journal of Political Economy, (1973).
- [10] Comte F., Renault E., *Long memory in continuous-time stochastic volatility models*, Mathematical Finance, An International Journal of Mathematics, Statistics and Financial Economics, (1998).
- [11] Comte F., Coutin L., Renault E., *Affine fractional stochastic volatility models*, Annals of Finance, (2012).
- [12] Corcuera J. M., *Quantitative finance*, (2018).
- [13] Cox J. C., *Notes on option pricing I: Constant elasticity of variance diffusions*, The Journal of Portfolio Management, (1975).
- [14] Cox J. C., Ingersoll J. E., Ross S. A., *A theory of the term structure of interest rates*, Econometrica. Journal of the Econometric Society, (1985).
- [15] Derman E., Kani I., *Riding on a smile*, Risk, (1994).
- [16] Duffie D., Pan J., Singleton K., *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica, Journal of the Econometric Society, (2000).
- [17] Dupire B., *Pricing with a smile*, Risk, (1994).
- [18] Fouque J-P., Papanicolaou G., Sircar K. R., *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, Cambridge, U.K., (2000).
- [19] Gatheral J., Jaisson T., Rosenbaum M., *Volatility is rough*, Quantitative Finance, (2018).
- [20] Gulisashvili A., Lagunas M., Merino R., Vives J., *Higher order approximation of call option prices under stochastic volatility models*, Journal of Computational Finance, (2020).
- [21] Hagan P. S., Kumar D., Lesniewski A., Woodward D. E., *Managing smile risk*, Wilmott Magazine, (2002).
- [22] Heston S. L., *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies, (1993).
- [23] Hull J. C., White A. D., *The pricing of options on assets with stochastic volatilities*, Journal of Finance, (1987).
- [24] Johnson H. E., *Option pricing when the variance is changing*, Graduate School of Management, (1979).

- [25] Johnson H. E., Shanno D., *Option pricing when the variance is changing*, Graduate School of Administration, (1985).
- [26] Kendall M.G., *The analysis of economic time-series part I: Prices*, Journal of the Royal Statistical Society, Series A (General), (1953).
- [27] Kou S. G., *A jump-diffusion model for option pricing*, Management Science, (2002).
- [28] Mrázek M., Pospíšil J., Sobotka T., Vives J., *Decomposition formula for jump diffusion models*, International Journal of Theoretical and Applied Finance, (2018).
- [29] Merino R., Vives J., *Option price decomposition in spot-dependent volatility models and some applications*, International Journal of Stochastic Analysis, (2017).
- [30] Merton R. C., *Theory of rational option pricing*, Economics and Management Science, (1973).
- [31] Merton R. C., *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Economics, (1976).
- [32] Mrázek M., Pospíšil J., *Calibration and simulation of Heston model*, Open Mathematics, (2017).
- [33] Øksendal B., *Stochastic differential equations: An introduction with applications*, Universitext, Springer-Verlag, Berlin, (2003).
- [34] Osborne MFM., *Brownian motion in the stock market*, Operations Research, (1959).
- [35] Protter P., *Stochastic integration and differential equations*, Springer-Verlag Berlin Heidelberg, (2004).
- [36] Rubinstein M., *Implied binomial trees*, Journal of Finance, (1994).
- [37] Samuelson P., *Rational theory of warrant pricing*, Industrial Management Review, (1965).
- [38] Scott L. O., *Option pricing when the variance changes randomly: Theory, estimation, and an application*, The Journal of Financial and Quantitative Analysis, (1987).
- [39] Stein J., Stein E., *Stock price distributions with stochastic volatility: An analytic approach*, Review of Financial Studies, (1991).
- [40] Wiggins J. B., *Option values under stochastic volatility: Theory and empirical estimates*, Journal of Financial Economics, (1987).
- [41] Yan G., Hanson F. B., *Option pricing for a stochastic-volatility jump-diffusion model with log-uniform jump-amplitude*, Proceedings of American Control Conference, IEEE, (2006).

Dirección del autor:

VidaCaixa S.A.,
Unidad de Gestión de Riesgo de Mercado,
C/Juan Gris, 2-8, 08014 Barcelona, Spain.
e-mail: raul.merino85@gmail.com

mixba'al

CONTENIDO

- 7 Dra. Rosa Obdulia González Robles (1952-2021). A la memoria de Luly.
M.J. Arroyo Paniagua
- 13 In memoriam: Francisco Hugo Martínez Ortíz.
Comité Editorial
- 15 Tres teoremas sobre cardinales medibles.
Franklin Galindo
- 33 Operadores simétricos multivaluados y el problema de momentos de Hamburger truncado.
Josué Rios-Cangas
- 47 El método de descomposición de Adomian para el cálculo de integrales gaussianas: ecuación de calor y de Black-Sholes.
Oswaldo González-Gaxiola
- 55 Caminando entre medidas invariantes y topologías: el Teorema de Haar Inverso.
Saúl Pilatowsky Cameo
- 71 El Teorema de metrización de Urysohn.
Fidel Casarrubias Segura
- 81 La fórmula de descomposición para la valoración de opciones de compra bajo el modelo de Heston.
Raúl Merino